In host evolution of *Exophiala dermatitidis* in cystic fibrosis lung micro-environment

Tania Kurbessoian[1], Daniel Murante[2], Alex Crocker[2], Deborah A. Hogan[2], Jason E. Stajich[1]

[1]*Department of Microbiology and Plant Pathology and Institute of Integrative Genome Biology, University of California, 92521, Riverside, CA, USA*

[2]*Department of Microbiology and Immunology, Geisel School of Medicine, Dartmouth, 03755, Hanover, NH, USA*

* correspondence to: Jason E. Stajich, jason.stajich@ucr.edu

Keywords: Exophiala dermatitidis, black yeast, cystic fibrosis, population genomics

Abstract

Individuals with cystic fibrosis (CF) are susceptible to chronic lung infections that lead to inflammation and irreversible lung damage. While most respiratory infections that occur in CF are caused by bacteria, some are dominated by fungi such as the slow-growing black yeast *Exophiala dermatitidis*. Here, we analyze isolates of *E. dermatitidis* cultured from two samples, collected from a single subject two years apart. One isolate genome was sequenced using long-read Nanopore technology as an in-population reference to use in comparative single nucleotide polymorphism (SNP) and insertion-deletion (INDEL) variant analyses of twenty-three isolates. We then used population genomics and phylo-genomics to compare the isolates to each other as well as the type strain *E. dermatitidis NIH/UT8656*. Within the CF lung population, three *E. dermatitidis* clades were detected, each with varying mutation rates. Overall, the isolates were highly similar suggesting that they were recently diverged. All isolates were MAT 1-1, which was consistent with their high relatedness and the absence of evidence for mating or recombination between isolates. Phylogenetic analysis grouped sets of isolates into clades that contained isolates from both early and late time points indicating there are multiple persistent lineages. Functional assessment of variants unique to each clade identified alleles in genes that encode transporters, cytochrome P450 oxidoreductases, iron acquisition and DNA repair processes. Consistent with the genomic heterogeneity, isolates showed some stable phenotype heterogeneity in melanin production, subtle differences in antifungal minimum inhibitory concentrations, growth on different substrates. The persistent population heterogeneity identified in lung-derived isolates is an important factor to consider in the study of chronic fungal infections, and the analysis of changes in fungal pathogens over time may provide important insights into the physiology of black yeasts and other slow-growing fungi *in vivo*.

Introduction

Cystic fibrosis (CF) is an autosomal recessive disorder caused by mutations in the cystic fibrosis transmembrane regulator (CFTR) gene that impair the balance of salts and water across epithelia. In the lungs, these ion transport defects cause viscous mucus which contributes to respiratory infections that cause most of the morbidity and mortality in CF populations (Riordan *et al.* 1989; Davis 2006; Ferec and Cutting 2012). Microbial colonization of mucosal plugs results in recurring infection and inflammation that cause irreversible lung damage and declining lung function (Turcios 2020). Bacteria, particularly *Staphylococcus aureus, Pseudomonas aeruginosa,*

Burkholderia cepacia and *Stenotrophomonas maltophilia* are pathogens that frequently dominate CF respiratory infections (Burns *et al.* 1998; Mariani-Kurkdjian and Bingen 2003; Horré *et al.* 2004; Steinkamp *et al.* 2005; Tunney *et al.* 2008; Pihet *et al.* 2009; Zhao *et al.* 2012) and are sometimes isolated with various species of fungi. Clinically significant fungi in CF lung infections include *Exophiala dermatitidis*, *Scedosporium apiospermum*, *Aspergillus fumigatus*, *Candida albicans* and *Clavispora (Candida) lusitaniae* (Kusenbach *et al.* 1992; Cimon *et al.* 2000; Defontaine *et al.* 2002; Horré *et al.* 2004, 2009; Parize *et al.* 2014; Chen *et al.* 2018; Demers *et al.* 2018; Jong *et al.* 2020). The consequences of fungal infection on CF outcomes is not well understood but is influenced by the genotypes of both the host and microbes (Burns *et al.* 1998; Nagano *et al.* 2007; Pihet *et al.* 2009; Packeu *et al.* 2012).

Exophiala dermatitidis, previously named *Hormiscium dermatitidis* and *Wangiella dermatitidis*, are taxonomically classified in Phylum Ascomycota, Order Chaetothyriales, Family Herpotrichiellaceae. To date, about 40 species in the *Exophiala* genus have been identified, seventeen of which are known to cause disease in mammals. Among these, *E. dermatitidis* is the most clinically prevalent with reported mortality rates of 25–80% in systemic and invasive cases, even though fatal systemic cases are relatively rare (Kirchhoff *et al.* 2019). Clinical presentations of this fungus include phaeohyphomycosis, keratitis, chromoblastomycosis and even several neural diseases and meningitis (Revankar *et al.* 2002; Matos *et al.* 2002; Uijthof *et al.* 2009; Revankar and Sutton 2010; Seyedmousavi *et al.* 2014; Song *et al.* 2017; Kirchhoff *et al.* 2019; Lavrin *et al.* 2020). The first instance of *E. dermatitidis* to be isolated from a sputum culture procured from a cystic fibrosis patient was in 1990 (Haase *et al.* 1990; Kusenbach *et al.* 1992). Many studies have since isolated *E. dermatitidis* from CF sputum cultures (Rath *et al.* 1997; Diemert *et al.* 2001; Horré *et al.* 2004; Griffard *et al.* 2010; Packeu *et al.* 2012), some of which have led to the development of later state mycosis disease (Sudfeld *et al.* 2010; Kondori *et al.* 2011; Song *et al.* 2017; Grenouillet *et al.* 2018), treatment (Packeu *et al.* 2012; Mukai *et al.* 2014) or even death (Klasinc *et al.* 2019).

Exophiala species are black yeasts which are classified by three defining features. They produce melanin through 1-8 dihydroxynapthalene (DHN) biosynthesis pathway, exhibit morphological plasticity or meristematic growth (yeast cells, hyphae or even pseudohyphae), and have membrane associated carotenoids and an intracellular mycosporine-like amino acids (Hoog *et al.* 2000; Nosanchuk and Casadevall 2003, 2006; Saunte *et al.* 2012; Smith and Casadevall 2019). All of these properties likely contribute to the extreme resistance to environmental stresses including desiccation, UV or solar exposure, and ionizing radiation. These resistance traits may also contribute to success in growth in mammalian hosts and their ability to cause disease in susceptible hosts. The *E. dermatitidis* strain NIH/UT8656 genome was sequenced and assembled into 11 complete and contiguous chromosomes (Robertson *et al.* 2012; Chen *et al.* 2014; Schultzhaus *et al.* 2020; Malo *et al.* 2021) which has enabled comparative genomics and identification of the genes which may underlie its resilience (Robertson *et al.* 2012; Schultzhaus *et al.* 2020; Malo *et al.* 2021) and success in human host colonization (Kondori *et al.* 2011; Kirchhoff *et al.* 2019).

Reservoirs of *E. dermatitidis* have long been associated with hot and humid tropical origins (Sudhadham *et al.* 2008). *E. dermatitidis* are isolated from many man-made substrates found in humidifiers (Nishimura and Miyaji 1982), saunas (Matos *et al.* 2002), and dishwashers (Zupančič *et al.* 2016; Babič *et al.* 2018). Babic et al. (Babič *et al.* 2018) concluded that

Exophiala tends to be found in locations with oligotrophic conditions or where rubber seals and humidity act as an enrichment or trapping mechanism which supports *Exophiala* persistence. *E. dermatitidis* and related species are also found in broad environmental niches including wasp nest (Conti-Díaz *et al.* 1977), healthy bats (Reiss and Mok 1979), lesions of toads (Frank *et al.* 1970), rotten wood (Dixon *et al.* 1980) which suggests human patients acquire infections from environmental exposure (Sudhadham *et al.* 2011).

Successful microbial invasions require iron, a critical growth-limiting factor, which must be sequestered from the surrounding environment through the release of iron chelating proteins called siderophores (Neilands 1995; Mossialos and Amoutzias 2009). For example, *Pseudomonas aeruginosa* produces the fluorescent siderophores, pyoverdine and pyochelin, which are used to sequester iron from the lung environment and as cofactors for respiratory proteins needed for surface motility and biofilm maturation (Haase *et al.* 1990; Banin *et al.* 2005; Matilla *et al.* 2007). Another example, the fungus *Aspergillus fumigatus* produces two hydroxamate-type siderophores: triacetylfusarinine (SidD) and ferricrocin (SidC), while also producing fusarinine and hydroxyferricrocin type siderophores (Schrettl *et al.* 2007; Haas 2014). In a study by Zajc et al., (Zajc *et al.* 2019), indicated 10 predicted siderophores in four *Exophiala dermatitidis* genomes, two of which include triacetylfusarinine and ferricrocin. Ferricrocin is an internal siderophore used to store iron and is essential for sexual development and contributes to oxidative stress resistance (Schrettl *et al.* 2007; Tyrrell and Callaghan 2016). Triacetylfusarinine is used to facilitate hyphal growth under iron-depleted conditions (Schrettl *et al.* 2007; Tyrrell and Callaghan 2016).

The *E. dermatitis* locus (HMPREF1120_07636) depicts the non-ribosomal peptide synthase SidC necessary for siderophore synthesis (Zajc *et al.* 2019; Malo *et al.* 2021). Polymicrobial infections persist in the CF lung environment through production and scavenging of extracellular siderophores which aid microbes in competition for resources (Tyrrell and Callaghan 2016; Sass *et al.* 2019; Yan *et al.* 2022). Microbes can use, obtain, and sequester iron in siderophores from hemoglobin found in red blood cells and lactoferrin contained in mucosal secretions. The competition and use of iron is an important dynamic in polymicrobial infections including fungal-bacteria competition within plant and animal hosts, and may sometimes assist in promoting the growth of their host (Crowley *et al.* 1991; Johnson 2008; Aznar *et al.* 2014; Kim 2018; Mochochoko *et al.* 2021; Pohl Carolina H. and Noverr Mairi C.).

In light of several studies that chronic CF-related fungal infections diversify over time (Kim *et al.* 2015; Demers *et al.* 2018; Jones *et al.* 2020; Ross *et al.* 2021), here we report both phenotypic and genomic diversity among *E. dermatitidis* isolates from a single individual. Population genomic analyses identified multiple lineages that persisted over two years. This data set allows us to test the hypothesis that, as seen in previous studies, the CF lung environment supports stably diverged populations of a clonally derived yeast.

Methods

Sputum-derived isolate cultures

Frozen sputum was obtained from a specimen bank in which samples were obtained in accordance with protocols approved by the Dartmouth-Hitchcock Institutional Review Board. Aliquots of sputum were plated onto Sabouraud Dextrose Agar (SAB) medium as described

previously (Grahl *et al.* 2018). Individual isolates were obtained and banked; isolate identifiers are listed in **Supplemental Table 1.**

DNA extraction and sequencing

Exophiala isolates were grown in Yeast Peptone Dextrose media (YPD) for approximately 24 hours in 5 ml roller-drum cultures at 37 □C. Cells were spun down for 5 minutes at 5000 RCF and washed thrice with deionized water. Genomic DNA was extracted from cell pellets using the MasterPure yeast DNA purification kit (Epicentre). Melanin was removed from genomic DNA using the *OneStep*™ PCR Inhibitor Removal kit (Zymo Research). Genomic DNA was measured by Nanodrop and diluted to ~20ng/µl. DNA extractions were sent to Novogene, (Novogene Corporation Inc., Cambridge, United Kingdom) for 2x150bp sequencing on an Illumina NovoSeq 6000. DNA from isolate DCF04 was also extracted and sequenced on Oxford Nanopore (ONT) platform with library preparation and sequencing following manufacturer's directions (Oxford Nanopore, Oxford United Kingdom). Flow cell versions FAK67997 and FAK73296 were used along with base-calling using guppy (*v. 3.4.4+a296cb*) (Wick *et al.* 2019).

Genome assembly and annotation

Genome assemblies were constructed for the twenty-three *E. dermatitidis* isolates from Illumina sequencing. One isolate, Ex4, was also re-sequenced using Oxford Nanopore technology. All genomes were *de novo* assembled with AAFTF pipeline (*v.0.2.3*) (Palmer and Stajich 2022) which performs read QC and filtering with BBTools bbduk (*v.38.86)* (Bushnell 2014) followed by SPAdes (*v.3.15.2*) (Bankevich *et al.* 2012) assembly using default parameters, followed by screening to remove short contigs < 200 bp and contamination using NCBI's VecScreen. The BUSCO ascomycota_odb9 database (Manni *et al.* 2021) was used to determine how complete the assembly was for all 23 isolates of *E. dermatitidis.* A hybrid assembly of isolate DCF04/Ex4 was generated using MaSUrCa (*v.3.3.4*) (Zimin *et al.* 2013) as the assembler using both Nanopore and Illumina sequencing reads. General default parameters were used except: CA_PARAMETERS=cgwErrorRate=0.15, NUM_THREADS=16, and JF_SIZE=200000000. The updated genome was then scaffolded to strain NIH/UT8656 accession GCF_000230625 using Ragtag (*v.1.0.2*) (Alonge *et al.* 2019) which uses minimap2 (*v. 2.17-r941)* (Li 2018) to further link scaffolds based on shared co-linearity of these isolates' genomes.

We predicted genes in this near complete genome assembly with Funannotate (*v1.8.1*) (Palmer and Stajich 2020). A masked genome was created by first generating a library of sequence repeats with the RepeatModeler pipeline (Smit and Hubley 2008). These species-specific predicted repeats were combined with fungal repeats in the RepBase (Bao *et al.* 2015) to identify and mask repetitive regions in the genome assembly with RepeatMasker (*v.4-1-1*) (SMIT A. F. A 2004). To predict genes, *ab initio* gene predictors SNAP (*v.2013_11_29*) (Korf 2004) and AUGUSTUS (*v.3.3.3*) (Stanke *et al.* 2006) were trained using the Funannotate 'train' command based on the full-length transcripts constructed by Genome-Guided run of Trinity (*v.2.11.0*) (Grabherr *et al.* 2011) using RNA-Seq from published E. *dermatitis* SRA accession SRS282040. The assembled transcripts were aligned with PASA (*v.2.4.1*) (Haas *et al.* 2008) to produce full-length spliced alignments and predicted open reading frames for training the *ab initio* predictors and as informant data for gene predictions. Additional gene models were predicted by GeneMark.HMM-ES (*v.4.62_lic*) (Brůna *et al.* 2020), and GlimmerHMM (*v.3.0.4*) (Majoros *et al.* 2004) that utilize a self-training procedure to optimize *ab initio* predictions. Additional exon evidence to provide hints to gene predictors was generated by DIAMOND BLASTX alignment

of SwissprotDB proteins and polished by Exonerate (*v.2.4.0*) (Slater and Birney 2005). Finally, EvidenceModeler (*v.1.1.1*) (Haas *et al.* 2008) generated consensus gene models in Funannotate were constructed using default evidence weights. Non-protein-coding tRNA genes were predicted by tRNAscan-SE (*v.2.0.9*) (Lowe and Chan 2016).

The annotated genome was processed with antiSMASH (*v.5.1.1*) (Blin *et al.* 2021) to predict secondary metabolite biosynthesis gene clusters. These annotations were also incorporated into the functional annotation by Funannotate. Putative protein functions were assigned to genes based on sequence similarity to InterProScan5 (*v.5.51-85.0*) (Jones *et al.* 2014), Pfam (*v.35.0*) (Finn *et al.* 2014), Eggnog (*v.2.1.6-d35afda*) (Huerta-Cepas *et al.* 2019), dbCAN2 (*v.9.0*) (Zhang *et al.* 2018) and MEROPS (*v.12.0*) (Rawlings *et al.* 2018) databases relying on NCBI BLAST (*v.2.9.0+)* (Sofi *et al.* 2022) and HMMer (*v.3.3.2*) (Potter *et al.* 2018). Gene Ontology terms were assigned to protein products based on the inferred homology based on these sequence similarity analyses. The final annotation produced by Funannotate was deposited in NCBI as a genome assembly with gene model annotation.

Copy number variation (CNV) was examined by plotting window-based read coverage of the short-read alignments of each isolate. The depth of coverage was calculated using mosdepth (Pedersen and Quinlan 2018), and visualized with R using the ggplot2 package (Wickham 2016).

The Mating Type (MAT) locus was identified through searching for homologous MAT genes (HMPREF08862) and (HMPREF05727) in this study's 23 *E dermatitidis* isolate genomes with cblaster (Gilchrist *et al.* 2021). A homothallic black yeast, *Capronia coronata* CBS 617.96 (AMWN00000000.1) (Teixeira *et al.* 2017), which has both MAT genes, was also incorporated into the analyses and visualization. The identified homologous regions were examined for their conserved synteny of the MAT locus using clinker (Gilchrist and Chooi 2021) and a custom Biopython script (Cock *et al.* 2009) (Kurbessoian 2022) to extract the annotated region of the genome which contained the locus.

Identification of telomeric repeat sequences was performed using FindTelomeres.py script (https://github.com/JanaSperschneider/FindTelomeres). Briefly, this searches for chromosomal assembly with a regular expression pattern for telomeric sequences at the 5' and 3' end of each scaffold. Telomere repeat sequences were also predicted using A Telomere Identification toolkit (tidk) (*v.0.1.5*) "explore" option (https://github.com/tolkit/telomeric-identifier).

Identification of sequence variation

Sequence variation among isolates was assessed using the best practices of the Genome Analysis ToolKit GATK (*v. 4.0.4.0*) (McKenna *et al.* 2010; Franke and Crowgey 2020) to identify SNPs and Insertion/Deletions (INDEL). Illumina paired-end reads were aligned to isolate DCF04 assembly with BWA (*v.0.7.17*) (Li and Durbin 2010) and processed with Samtools (*v.1.8*) (Li *et al.* 2009) and Picard Toolkit (Institute; Toolkit) AddOrReplaceReadGroups and MarkDuplicates (*v.2.18.3*). The alignments were further improved by realigning reads near inferred INDELs using GATK tools RealignerTargetCreator, IndelRealigner, and PrintReads. Genotypes were inferred with the GATK Haplotype and GenotypeGVCF methods to produce a single VCF file of the identified variants. Low quality SNPs were further filtered using GATK VariantFiltration and finally SelectVariants was used with the parameters: mapping quality (score < 40), quality by depth (<2 reads), Strand Odds Ratio (SQR > 4.0), Fisher Strand Bias (> 200), and Read Position

Rank Sum Test (< -20) to retain only high-quality polymorphisms. Finally, an additional stringent series of three filtering steps implemented in bcftools (*v. 1.12*)(Li et al. 2019) was used on the VCF file to remove calls that were below the 1000 quality score threshold, where any individual isolate had a "no call", and where the standard deviation in read depth (DP) was above or below a standard deviation value of 1 for an individual SNP. SnpEff (*v.4.3r*) (Cingolani *et al.* 2012) was used to score the impact of the identified variants using the Funannotate annotated DCF04 genome GFF3 file.

Variant calling was performed on two sets of individuals, one limited to the twenty-three CF patient population isolates with reads aligned to the DCF04 isolate and one using the *E. dermatitidis* NIH/UT8656 type strain. Pairwise isolate comparisons of SNP and INDEL were counted to generate isolate correlation heatmaps for both variant types using a UPGMA clustering. A custom script make_diagonal.sh uses plink (*v.2.00a3LM_AVX2*) (Chang *et al.* 2015) to count all pairwise differences between individuals in the VCF files stratified by SNPs or INDELs. A custom Perl script transformed pairwise counts into a matrix of isolated differences observed for both SNP and INDEL variants. Counts were summarized as heatmaps with a R script. To summarize the matrix plots, a distance plot using regression statistics was applied on both SNP counts and INDEL counts. A regression plot and statistics for the slope of the progression line, Pearson's R, R-squared and the *p*-value were computed with a R script. All scripts developed for this manuscript are available at the Github (Kurbessoian 2022) project linked in this paper.

A calculation of the population mutation rate was performed on each isolate based on the number of SNPs shared among a pair of isolates. The formula to calculate the mutation rate per year for each isolate is as follows: (SNP Pairwise Value) / (Adjusted Genome Length) / Pair / Year. The time between isolated collections was 22 months. The value used for the "Adjusted Genome Length" was collected from running the assembler AAFTF pipeline (*v.0.2.3*) (Palmer and Stajich 2022). A one-way ANOVA was run on the grouped calculated mutation rates for each isolate to determine significance.

Genome comparison with dot-plot was constructed with D-GENIES (Cabanettes and Klopp 2018) using minimap2 and default parameters through the website for the tool. Longer isoform proteins were extracted for each strain genome annotation in order to call a more accurate gene count. Using Orthofinder (Emms and Kelly 2019) and DIAMOND ultra-sensitive parameters (Buchfink *et al.* 2015), assessment of overlap in the predicted protein-coding gene sequences from the genomes of *E. dermatitidis DCF04* and *E. dermatitidis NIH/UT8656* protein genomes was generated.

Phylogenetics relationships of the isolates

SNPs from polymorphic sites were extracted from the VCF files as multi-fasta files using BCFTools (Li *et al.* 2019) and a custom script make_strain_tree.sh. A Maximum Likelihood phylogenetic tree was constructed from the multi-fasta file using IQTree (*v. 2.0.4*) (Minh *et al.* 2020) and the model parameters [-m GTR+ASC]. The chosen nucleotide substation model was GTR+F+ASC selected based on Bayesian information criteria (BIC). Statistical support for the tree nodes was evaluated from 1000 bootstrap replicates using UFBoot ultra-fast bootstrapping approximation (Hoang *et al.* 2018). The tree was visualized using iTOL (Letunic and Bork 2016).

Phenotype assays

Exophiala isolates were streaked from -80˚C onto yeast extract peptone dextrose (YPD) plates (2% glucose, 2% yeast extract, 1% peptone) and allowed to grow for 48 hours at 37 □C. Overnight cultures were started from YPD patches inoculated into 5 ml of liquid YPD and grown for approximately 24 hours in 5 ml rolling barrel cultures at 30 □C. For MIC assays, cultures were spun down for 5 minutes at 5000 RCF and washed thrice in deionized water. Cells were counted on a hemocytometer and added to a final concentration of 1000 colony forming units per well in a 96-well flat-bottom dish, then grown at 37˚C for 72 hours before measuring final MIC.

Results

A molecular and culture-based analysis of a series of sputum samples identified an individual with CF with a chronic lung infection caused by *E. dermatitidis* (Grahl *et al.* 2018). *E. dermatitidis* isolates were recovered from banked sputum samples, collected two years apart. *Staphylococcus aureus* and *Candida albicans* were also identified in clinical cultures from the patient in the intervening years between the two timepoints (**Figure 1**). The subject's antimicrobial use history included Aztreonam, Azithromycin, Tobramycin, Ciprofloxacin, and Doxycycline and the patient's lung function, measured by percent predicted forced expiratory volume (%FEV1), ranged between 80 and 49% during this time period. While *E. dermatitidis* was not detected in the first clinical microbiological analysis, perhaps due to its extremely slow growth out of clinical samples (Grahl *et al.* 2018) or suppression by bacteria, it was detected in the second clinical analysis. Twenty-three isolates (eleven from the early time point and twelve from the late timepoint) were selected for further population genomic study.

Sequencing and assembly of E. dermatitidis isolates

To gain information on the genetic variation and potential population structure for the recovered *E. dermatitidis* isolates, we sequenced and assembled the genomes of the twenty-three isolates (**Supplemental Table 2**). The depth of coverage ranged from 11-47x coverage across all 23 Illumina sequenced samples. The BUSCO (Benchmarking Universal Single Copy Orthologs) is another tool used to determine genome assembly completeness; within our dataset it ranged from 99-99.3 % complete, 98.3-99.3 single copies present, 0-1 ranged in duplicates, with a range of 0-0.3 in missing BUSCOs. Contig counts ranged from 44-1422, while the average genome assembly size is 26,706,323 Mbp, L50 about 1,078,277 and N50 of 9.

To further examine the fine-scale variation within the CF isolates we sought to generate a within population high-quality reference genome. The isolate DCF04 was sequenced using Oxford Nanopore (ONT) long reads to a genome depth of coverage at 11x and a hybrid genome assembly constructed from both ONT and Illumina reads (**Table 1; Supplemental Table 2).** The 26.6 Mb assembly contained only 3.19% of identified repetitive elements. The candidate telomeric repeat units "TTTAGGG/CCCTAA" were identified as repeat arrays at both ends of five scaffolds, but also found as single pairs in the remaining 4 scaffolds, as would be expected for 9 complete chromosomes (**Supplemental Table 3**). Sixty-four tRNA models were predicted from the genome. While the total number of genes predicted was 10030, 9599 of which are protein coding. 15 secondary metabolite clusters, 37 biosynthetic enzymes, and 49 small COGs were predicted with antiSMASH.

Table 1. Genome assembly summary statistics for reference isolate *E. dermatitidis* DCF04.

Genome Assembly Statistics	***E. dermatitidis* DCF04**
Scaffold Count	44
Total length	26,633,774 bp
Minimum length	501
Maximum length	4,314,646
Mean contig length	59,1861.64
Scaffold L50	4
Scaffold N50	3,672,342
Scaffold L90	7
Scaffold N90	2,864,055
Contig Count	165
Contig L50	21
Contig N50	451,966
BUSCO Complete %	99%
BUSCO Duplicate %	0%
BUSCO Single %	99%
GC%	51.40

Comparing genome assembly and annotation of DCF04 to NIH/UT8656

Our DCF04 isolate is genetically close to the public strain *Exophiala dermatitidis NIH/UT8656* sequenced with Sanger sequencing technology (BioProject: PRJNA225511, Assembly: GCF_000230625.1). To assess the differences between the two genome assemblies we compiled summary statistics (**Table 1** and **Supplemental Table 2**) and interrogated the predicted gene content of both genomes. The DCF04 assembly had 165 contigs linked into 44 scaffolds, while NIH/UT8656 comprised 238 contigs linked into 10 scaffolds. Note that DCF04 scaffolds were derived by a comparative assembly against the NIH/UT8656 assembly to achieve best assembly after checking for rearrangements. The total length of the genome assembly is nearly the same for both DCF04 at 26.6Mb and NIH/UT8656 at 26.4Mb. The summary statistics for scaffold L50 and N50 are also nearly identical for DCF04 at 4 and 3.7Mb and NIH/UT8656 were 4 and 3.6Mb. A dot-plot comparing the two genome assemblies revealed minimal rearrangements or discontinuity suggesting high similarity of the two isolated genomes (**Supplemental Figure 1**).

The genome content was further compared using OrthoFinder (Emms and Kelly 2019) (**Supplemental Table 4).** OrthoFinder identified 8,256 orthologous groups or 17,640 orthologous protein-coding genes between the two genomes. Both strains had a number of unassigned genes that were given orthogroups, 705 for DCF04 and 475 for NIH/UT8656, along with 58 genes that were assigned, 34 DCF04 genes and 24 NIH/UT8656 genes. Of the 34 isolate-specific assigned genes in DCF04, 21 had identifiable fungal homologs with NIH/UT8656 including ABC multidrug transporters, AAT family amino acid transporter, 5-oxoprolinase and hypothetical protein/P-loop containing nucleoside triphosphate hydrolase protein. Two of the 13 assigned protein-coding genes resulted in a ribonuclease HI protein, while the remaining assigned 11 resulted in hypothetical protein matches. 4 out of the 5 orthogroups specific to NIH/UT8656 (22 protein-coding genes) were identical to the 4 seen in DCF04, while 1 orthogroup matched to a hypothetical protein/DUF300-domain containing protein. When analyzing the unassigned genes, these functionally had no known paralog on NCBI. We believe these results reflect differences in gene prediction pipelines as much as it could be due to gene content differences (Weisman *et al.* 2022).

All Ex CF isolates are MAT1-1 mating type

The small-scale genome synteny evaluation tool clinker was used to visualize slices of the genome adjacent to the identified MAT loci. All the lung isolates including DCF04/Ex4 *E. dermatitidis* isolate encoded a MAT1-1 locus (**Figure 2, Supplemental Figure 2, Supplemental Table 5**). For all other *E. dermatitidis* isolates and *Capronia* species analyzed, *SLA2*, a SRC-like adapter protein, and *APN2*, apurinic-apyrimidinic endonuclease 2, flanked the MAT loci in all isolates, and a hypothetical protein between SLA2 and MAT1-1-4 was also present across isolates. The MAT locus found in strain NIH/UT8656 have the MAT1-2 mating type, while our clinical isolates have the MAT 1-1 mating type. Consistent with MAT loci described in *E. dermatitidis* (Metin *et al.* 2019) DCF04 had both MAT1-1-4 and MAT1-1-1 genes (**Figure 2**). The homothallic outgroup *Capronia coronata CBS617.96* genome (Teixeira *et al.* 2017) contains both MAT1-1 and MAT1-2 genes.

Chromosome copy number variation across E. dermatitidis isolates

Copy number variation of full or partial chromosomes was evaluated by calculating depth of coverage using 10kb sliding windows (**Figure 3).** The read depths of windows across *E. dermatitidis* isolate genomes from this study were compared across all 9 chromosomes. Visual scanning of the plots identified an anomaly of at least 1.5x higher coverage on chromosome 5 in Ex3 (**Figure 3A**). A similar but much smaller region of chromosome 5 appears to have 2-2.25x coverage and may be duplicated as an aneuploid in Ex13. Additional partial 1.25-1.5x coverage for part of the left arm of chromosome 2 in isolate Ex18 is also observed.

The plots of isolates Ex15, Ex18 and Ex20 indicate 0.5x less normalized coverage in chromosomes 1 and 3 (**Figure 3B and Figure 3C**), a possible sign of segmental chromosomal aneuploidy event. The continuation of the coverage pattern between chromosomes 1 & 3 seem to complement each other for these isolates. Haploid organisms, like *Exophiala dermatitidis*, could benefit from genomic plasticity through expansion or contraction to enable adaptation in a new environment (Selmecki *et al.* 2010; Legrand *et al.* 2019). Noting that Ex3 was isolated in the early time point, this genome copy may contribute to an adaptive mechanism to support colonization and persistence in the human host. Other isolates sharing potential aneuploidies (Ex15, Ex18, Ex20) are from the late time point and also have the highest mutation rate among

this population as observed from the phylogenetic tree and mutation rate calculation (**Figure 4 and Table 4**).

SNP genotyping and SNP-based phylogenetic analysis of 23 E. dermatitidis isolates

We used the DCF04/Ex4 isolate as a reference for variant identification within the 23 CF isolate collection. The DCF04/Ex4 VCF file generated 441 variants between twenty-two isolates and the DCF04 reference isolate. A phylogenetic tree was generated using the filtered SNP data results and rooted with the earliest diverging group containing Ex1, Ex2, Ex12 and Ex14 (**Figure 4**). When repeated for out-population reference strain NIH/UT8656 the analysis resulted in about ~11,000 variants between the NIH/8656 strain and the CF isolates (Kurbessoian 2022). Another phylogenetic tree was generated using a secondary dataset created with the NIH strain to be used as a root (**Supplemental Figure 4**).

Three clades (Clade I, Clade II, and Clade III) were identified based on the tree. Clade I is composed of two early collected and two late collected isolates of *E. dermatitidis* Ex1, Ex2, and Ex12, and Ex14. Clade II is composed of three subgroups, one of which contains only early isolates and a second group with both an early isolate and late isolates. The third group in Clade II contains two early collected isolates Ex6 and Ex10. Clade III has two main groups composed of both early and late isolates. The subgroup of Clade III had a long branch length suggesting more divergence. Interestingly, the CNV plot (**Figure 3**) showed these three isolates, Ex15, Ex20 and Ex18 contained similar CNV differences in chromosome 1 and 3 when compared to the other isolates.

Non-synonymous and synonymous SNP and INDEL pairwise differences

A dissimilarity matrix was constructed comparing the overlaps of SNPs and INDELs collated for all pairs of isolates. As expected, isolates that are closest to each other in the SNP-based phylogenetic tree had fewer differences in their SNP composition (**Figure 5A**). Our analysis was further supported by the observation that the number of SNP differences correlated with the number of INDELs detected in pairwise analyses (**Figure 5B**). Within Clade I and Clade II, isolates had few SNP and INDEL differences, indicating the evolutionary distance between them was small. The SNP and INDEL counts within Clade III were much higher indicating a higher divergence within this group of isolates as compared to other groups.

Table 2. Genes with SNP variants found stratified by clade.

Clade	Isolates	Nonsynonymous mutations of interest in clade	Number of NonSynonymous/Synonymous SNPs unique to clade
I	**Ex1, Ex2, Ex12, Ex14**	MADS box Transcription factor, G4 quadruplex binding protein, MFS transporter SP family sugar:H+ symporter, Bud site selection Bud4, DNA polymerase alpha subunit A, GTPase-activiting protein	59/7
II	a) **Ex3, Ex4, Ex8**	a. All hypothetical.	56/8

	b) **Ex22, Ex17, Ex16, Ex7, Ex19, Ex23** c) **Ex6, Ex10**	b. Ex7 - Ex22 GO terms indicate Signal Transduction Mechanism and Transcription Factor, Ex7-Ex17 Zinc finger binding protein TFIIIA, Ex16 queuine tRNA-ribosyltransferase, MFS transporter SP family sugar:H+ symporter, NAD-dependent histone deacetylase SIR2 and cytochrome P450, family 7, subfamily B (oxysterol 7-alpha-hydroxylase) c. None.	
III	a) **Ex9, Ex13, Ex11** b) **Ex18, Ex15, Ex20, Ex21, Ex5**	a. Ex13 & Ex9, Ex11 Ap-3 complex subunit delta, DNA repair protein RAD50, and Regulator of nonsense transcripts 1-like protein. b. Ex18 & Ex15, Ex20 an MFS transporter DHA2 family methylenomycin A resistance protein, sulfite reductase (ferredoxin), glycerol ethanol-ferric requiring protein, polyketide synthase, MFS transporter SP family solute carrier family 2 and a DNA repair protein RAD20	125/20

Non-synonymous SNP Differences seen between clades

The human-readable snpEff table which best described the variant analysis along with functional gene annotation can be found in **Supplemental Table 6**. Clade I is composed of four isolates: two (Ex1, Ex2) and two isolates from the late time point (Ex12, Ex14). Two genes with mutations are of note, a MADS-box transcription factor (HMPREF1120_06786) and a G4 quadruplex nucleic acid binding protein (HMPREF1120_02174). The MADS-box transcription factor is part of the MADS-box proteins with a highly conserved 56 amino acid DNA-binding domain, some containing a weakly conserved K-box domain that is involved in the dimerization of transcription factors (Shore and Sharrocks 1995). In fungi, isolates with knocked-out MADS-box genes have reduced virulence (Damveld *et al.* 2005; Qu *et al.* 2014; Xiong *et al.* 2016). Mutations in *E. dermatitidis* MADS-box regions may increase their pervasiveness and tolerance of the lung environment. Previous studies have found that G4 quadruplex nucleic acid binding proteins. The helical complex is formed through guanine rich nucleic acid sequences and is found at the telomeric regions of chromosomes.

Clade II contains 11 isolates, six from the early time point and five from the later time point, that fall into three different subgroups. Ex3, Ex4/DCF04, and Ex8 fell into one group, Ex22, Ex17, Ex16, Ex7, Ex19 and Ex23 fell into the second group, and Ex6 and Ex10 formed a third group. Group one and group three both contain only early isolates, while the second group contains the majority of the later isolates.

Clade III, which contains eight isolates, also divided into two groups: one composed of Ex18, Ex15, Ex20, Ex21, and the other group of Ex9, Ex13 and Ex11. Five are later isolated while the

other three are from the early isolations. The most significant group is the second consisting of Ex9, Ex13 and Ex11. Four genes were found to have distinct mutational differences when comparing Ex13 (a late isolate) to the early isolates Ex9 and Ex11. These four genes are the Ap-3 complex subunit delta (HMPREF1120_00143), DNA repair protein RAD50 (HMPREF1120_05599), a regulator of nonsense transcripts 1-like protein (HMPREF1120_06837) which is similar to helicase-RNA complex, and a gene with no identified function (HMPREF1120_02556). The second group had three isolates (Ex18, Ex15 and Ex20) with a higher mutation rate than *E. dermatitidis* in Clade III. There are about 13 instances of hypothetical proteins, while the other 12 instances are predicted genes. Only six of the twelve are genes of note: an MFS transporter DHA2 family methylenomycin A resistance protein (HMPREF1120_00012), sulfite reductase (ferredoxin) (HMPREF1120_00943), glycerol ethanol-ferric requiring protein (HMPREF1120_01007), polyketide synthase (HMPREF1120_03173), MFS transporter SP family solute carrier family 2 (HMPREF1120_06771) and a DNA repair protein RAD50 (HMPREF1120_05599). Research on polyketide synthases in micro-colonial fungi and *E. dermatitidis* have been found to impact phenotypes and adjust the melanin synthesis pathway, resistance or susceptibility to antifungals and extreme environment adaptability (Paolo *et al.* 2006).

Analysis of SNP differences between clades

Finally, a single non-synonymous mutation in a gene orthologous to *S. cerevisiae MRS4* (HMPREF1120_06597), a mitochondrial iron transporter. One allele of *MRS4* allele encoded a protein that was identical in sequence to the allele in the reference NIH strain UT8656, present in the isolates in Clade I (Ex1, Ex2, Ex12, Ex14), and a subclade of clade II (Ex3, Ex4/DCF04, Ex8). The remainder of the isolates had a second allele with a non-synonymous mutation in the 40^{th} amino acid position, converting a glutamic acid residue to glycine. The functional consequences and differences of these Mrs4 alleles will be described in a separate manuscript.

Testing for enrichment of evolutionary patterns within clades

A pairwise comparison of the synonymous and nonsynonymous SNP differences was performed on fifteen pairs of isolates identified as early and late members of the same lineage. **Table 3** summarizes the significant results with a focus on candidate genes that may relate to lung pathogenicity. A comprehensive list of gene differences among all pairwise comparisons is in **Supplemental Table 7** and includes hypothetical proteins with no identified function. This analysis tested for differences between early and late isolates found in the same clade to focus on changes that may have occurred within the host.

Table 3. Representative genes with non-synonymous SNP variants of interest in pairwise comparisons of *E. dermatitidis* CF lung strains.

Early vs. Late	Function

Ex5 & Ex18	MFS transporter, DHA2 family, methylenomycin A resistance protein (HMPREF1120_00012) DNA repair protein RAD50 (HMPREF1120_05599) cytochrome P450 oxidoreductase (HMPREF1120_01361) sulfite reductase (ferredoxin) (HMPREF1120_00943) sulfite oxidase (HMPREF1120_01306) MFS transporter, DHA1 family, multidrug resistance protein polyketide synthase (HMPREF1120_6570) MFS transporter, SP family, sugar:H+ symporter (HMPREF1120_04157) tyrosinase (HMPREF1120_04514)
Ex9 & Ex13	G2/mitotic-specific cyclin 3/4 (HMPREF1120_00797) cytochrome P450 oxidoreductase (HMPREF1120_01361) DEAD box RNA helicase HelA (HMPREF1120_02010) MFS transporter, DHA1 family, multidrug resistance protein (HMPREF1120_01715) MFS transporter, SP family, sugar:H+ symporter (HMPREF1120_04157) DNA repair protein RAD50 (HMPREF1120_05599) ISU1 - iron-binding protein (HMPREF1120_06751) MFS transporter, SIT family, siderophore-iron:H+ symporter (HMPREF1120_07838)

The analysis considered all pairwise combinations between the early Ex1 and Ex2 and the late Ex12 and Ex14 isolates from Clade I. These isolates show very little genetic differentiation (about 6-9 variant SNPs and 6-10 variant INDELs). The contrast of early Ex1 vs Ex2 isolates to the late isolates (Ex12,Ex14) found variants in three genes: AFG2- an ATPase (HMPREF1120_06104), a DNA polymerase (HMPREF1120_07994), GYP1- a GTPase activating protein (HMPREF1120_08601), and MFS transporter (HMPREF1120_04157) homologs. These variants are an interesting observation and warrants extra analysis in the future.

Pairwise comparisons of the Clade II isolates contrasted the early Ex7 with each of the four late isolates Ex16, Ex17, Ex19, and Ex23. There was more variation among these isolates (about 11-169 variant SNPs and 3-16 variant INDELs) than observed in Clade I isolates. Isolate Ex23 also appeared to have substantially more differences from all others indicating it may be more distantly related or where its corresponding ancestral early strain was not sampled. A variant of note, NAD-dependent histone deacetylase SIR2, is involved with chromosomal remodeling specifically with phenotype transcription modification (Freire-Benéitez *et al*. 2016).

Comparison of Clade III group 1 members Ex5 (early) to Ex15,18,20,21 identified the highest number of variants as compared to the other clades (from 62-138 variant SNPs and 7-10 variant INDELs). Clade III group 2 members Ex9 (early), Ex11 (early) and Ex13 had a similar number of variants as Clades I & II (6-15 variant SNPs and 0-17 variant INDELs). The five pairwise comparisons for group 1 revealed a non-synonymous mutation in the RAD50-DNA repair protein (HMPREF1120_05599). Group 1 members in the phylogenetic tree also appear to have a long branch length indicating a potentially higher diversification rate than the other clades. It may be that the RAD50 changes could have an impact on DNA repair and contribute to the higher mutation rate observed. In addition, non-synonymous mutations were identified in a cytochrome P450 oxidoreductase (HMPREF1120_01361), MFS transporter sugar to H+ symporter (HMPREF1120_04157) and GTPase activating proteins (HMPREF1120_08601). SNP variants with predicted functional impact on interactions with the host and CF lungs were

found in the sulfite oxidase (HMPREF1120_01306), sulfite reductase (HMPREF1120_00943), polyketide synthase (HMPREF1120_03173), and tyrosinase (HMPREF1120_04514) genes.

Interestingly, the group 2 (Clade III) isolates Ex9, Ex11 and Ex13 had more variants in iron-binding (HMPREF1120_06751) and siderophore related (HMPREF1120_07838) genes than group 1. We took a candidate gene approach and tested if the siderophore NPRS SidC (HMPREF1120_07636) had accumulated any specific mutations in these lineages, but we failed to identify any non-synonymous variants in this gene across the CF lung isolates. Though, the presence of SNP variants in iron-binding and siderophore transporters indicates there are other means *E. dermatitidis* is obtaining the elusive iron molecules from its environment. Further analysis of these variants will point to supporting phenotypic differences allowing for effective colonization of CF lung environments.

Mutation rate calculation to test for different rates of evolution between clades

We calculated a mutation rate for each of the three clades (**Table 4**) by taking an average of all the pairwise comparisons with a clade. The Clade I average of the four pairwise calculations for all combinations of early and late isolates was 6.78E-08 variants/pair/2 years. The average rate for the five comparisons of early to late isolates in Clade II was 2.20E-07 variants/pair/2 years. The average rate for the six comparisons of early to late isolates in Clade III the first group values were averaged and determined to be 7.39E-07 variants/pair/2 years, while the second group values were averaged and determined to be 1.13E-07 variants/pair/2 years. The Clade III group 1 isolates appear to be a faster evolving group and may have acquired variants allowing improved adaptation to the lung environment. One-way ANOVA test revealed a statistically significant difference between all Clades (F-value = 7.0711, p-value = 0.00104) and a Tukey post-hoc test across the clades indicated Clade III (Group 1) has a significantly different mutation rate. Calculations underlying the mutation rate values are detailed in **Supplemental Table 8**.

Table 4. Average mutation rates for each clade/group.

Groups or Clades	Pairs	Mutation Rates
Clade I	Ex1+Ex12, Ex1+Ex14, Ex2+Ex12, Ex2+Ex12	6.76E-08 variants/pair/months
Clade II	Ex7+Ex16, Ex7+Ex17, Ex7+Ex19, Ex7+Ex22, Ex7+Ex23	2.20E-07 variants/pair/months
Clade III (Group I)	Ex5+Ex15, Ex5+Ex18, Ex5+Ex20, Ex5+Ex21	7.39E-07 variants/pair/months
Clade III (Group II)	Ex9+Ex11, Ex9+Ex13, Ex11+Ex13	1.13E-07 variants/pair/months

Heterogeneous phenotypes observed in the E. dermatitidis isolates

The twenty-three *E. dermatitidis* isolates (Ex1-11 from the early time point and isolates Ex12-23 from the late time point) were heterogeneous for traits in a number of ways including pigmentation, antifungal sensitivity, and auxotrophy. We noted differences in melanin pigmentation across isolates (**Figure 6A**) that did not correlate with either time point or

phylogenetic clade. Within Clade 1, Ex1, Ex2, and Ex12 had light melanization, but isolate Ex14 was one of the strongest melanin producers (**Figure 6B**). In a search for genetic determinants responsible for this phenotypic range, we identified three candidate intergenic mutations. Of particular interest is a mutation in the 5' UTR of a putative iron-dependent biphenyl-2,3-dioxygenase potentially involved in the degradation of phenolic compounds. Melanin is synthesized through the oxidation of various aromatic compounds, and thus disruption or enhancement of biphenyl oxidation may be a mechanism by which melanin production was enhanced in Ex14.

The 23 isolates also varied in sensitivity to itraconazole, a recommended treatment for *E. dermatitidis* infections (Fothergill *et al.* 2009; Mukai *et al.* 2014), over a ~10-fold range (minimum inhibitory concentrations (MIC) from 0.0625-0.5 µg/ml) (**Figure 7**). Heterogeneity in amino acid auxotrophy, scored as no growth on a minimal medium that was rescued by supplementation of amino acids (**Figure 7**), was also observed. Lastly, there were stable differences in filamentation across isolates in two of the three clades; closely related strains did not have similar morphologies (**Figure 8**).

Discussion

Fungus-dominant population

While opportunistic fungal pathogens are often culturable from the sputum of patients with cystic fibrosis, they do not commonly present as the dominant microbe or as a risk for infection. In this paper we have presented a clinical case where a population of the black yeast *E. dermatitidis* was the predominant microbe concomitant with a lung exacerbation event. Other reported clinical cases have shown that *E. dermatitidis* has an underappreciated role as a CF pathogen (Haase *et al.* 1990; Kusenbach *et al.* 1992; Kondori *et al.* 2014). Previous sputum isolates revealed the presence of this uncommon fungi two years prior, indicating that it persisted in low concentrations before having the opportunity to dominate the lung microbiome. Because *E. dermatitidis* is relatively slow growing (Rath *et al.* 1997; Sudhadham *et al.* 2011; Malo *et al.* 2021), it may be easily missed during routine clinical microbiological identification. To better understand the CF disease and the fungal infections associated with a long-term disease such as this one, it is important to consider testing for slow growing microorganisms such as black yeasts.

Phenotypic heterogeneity; Melanin production and drug resistance

Identification of key evolutionary strategies is crucial to the understanding of microbial pathogenesis in clinically relevant settings. Developing methods for evaluating population structure and heterogeneity for these human-associated microbes will aid in future studies. Initial observations of these isolates indicated a striking amount of heterogeneity in melanin production, which propelled an investigation into the genotypic diversity of the population. Population heterogeneity is an important factor to consider in treatment, and determining adaptive responses through genomics can help identify selective stimuli in chronic infections (Demers *et al.* 2018). We expected that differing levels of melanin, which comprises part of the cell wall in black yeasts, would have an effect on antifungal resistance. While drug resistance to clinically relevant antifungals Itraconazole and Voriconazole varied slightly between isolates, there was no consistent pattern of evolution that corresponded either with time of isolation or melanin production (**Figure 7**). We found amino acid auxotrophs in the population, and the isolation of

auxotrophs in CF infections has been previously described in bacterial isolates, such as *P. aeruginosa* (Barth and Pitt 1995; Thomas *et al.* 2000). This repeated occurrence of filamenting phenotypes in phylogenetically diverse isolates may indicate evolution towards hyper-filamentation or a non-mutation-based difference between isolates such as "switching" or phase variation, or epigenetic regulation.

Short and long read sequencing, in-population vs outside reference

While short-read sequencing provided us with the basic and high quality read depth for each CF isolate, using the Oxford Nanopore long-read technology sequencing to build on top of a short read assembly provided us with a well assembled genome. This genome was then used as an in-population reference allowing for a better recovery of variants due to the read mapping to a closer isolate. When using our in-population reference, compared to the NIH/UT8656 strain (Robertson *et al.* 2012; Chen *et al.* 2014; Schultzhaus *et al.* 2020; Malo *et al.* 2021), our quantification of population specific number of variants is higher (due to higher sensitivity) while still maintaining relevance in our study system.

Temporal resolution of variation accumulation in a fungal CF lung population

Contrasting mutations accumulating in isolates collected from two time points allowed comparison in mutation rate differences among three different genetic lineages of *E. dermatitidis* and evaluated if any changes could suggest adaptations that enabled lineages to survive the lung environment. Average mutation rates within a Clade tended to be similar indicating common diversification events. Statistical comparisons of mutation rates between Clades indicate significance seen between Clade I and Clade III group 1, Clade II and Clade III group 1 and Clade III group 1 and group 2. The main difference here is the greatly increased diversification in Clade III group 1 which is shown with the phylogenetic tree in Figure 4, chromosomal aneuploidies seen in Figure 3, the very high SNP and INDEL counts seen in Figure 5, the increased mutation rates seen in Table 4 and hyper-filamentous phenotypes see in Figure 8. We propose the mutations found in *RAD50* may have contributed to the increased diversification seen in this group and future studies will test this hypothesis (Goldman *et al.* 2002; Krogh and Symington 2004).

These clades persisted over a two-year span. Are there physical separations disallowing for SNPs in these sub-populations to mix? When observing the functional assessment of variants unique to each clade, we identified alleles in genes that encode for transporters, cytochrome P450 oxidoreductases, iron acquisition and DNA repair processes. Iron acquisition can be a possible virulence method *E. dermatitidis* could be using to persist in the lung environment as seen in *Aspergillus fumigatus* and *Pseudomonas aeruginosa* (Neilands 1995; Matilla *et al.* 2007; Schrettl *et al.* 2007). Mutations in transporters could be an evolutionary move to adapt to antibiotics or antifungal treatments. The finding of a non-synonymous mutation, *MRS3/4* a mitochondrial iron transporter, may also help point into the pathogenicity and virulence (Murante, 2022, In preparation). Though, the presence of certain SNP variants in iron-binding and siderophore transporters suggests that there could be other means by which *E. dermatitidis* is obtaining the elusive iron molecules from its environment. It would be important to observe the type of variants (e.g. stop codons or a loss of function) produced in these iron-related genes to begin searching for other virulence genes. Further analysis of these variants will point to supporting phenotypic differences allowing for effective colonization of CF lung environments. It's clear that this mutation is not found in the rapidly diverging Clade III and only in the Clade I cluster

with very few variants. The isolates we have sequenced only contain one mating type, indicating the low likelihood that even if the isolates have close proximity to each other they are unable to mate. As these isolates were collected from sputum plates, it could be certain isolates have co-localized within different lobes of the lung and could have nutritional or other ecological partitioning that could be driving these clades to diverge and remain diverged. Our results suggest that there may have been a diversification event that occurred early perhaps during the initial colonization. The clades then stabilized once co-localizing into their respective niches, or there may have been a second 'inoculation' event from a common, stably heterogeneous environmental source at some point in the two year timeline (Warren *et al.* 2011).

Conclusions

As it has become increasingly clear, collecting a single strain and using it as a metric to assess a single environment and moment of time is inaccurate and a more population approach must be used to best assess microbial infection (Demers *et al.* 2018). Having a closer reference strain to assess variants is also necessary to observe true variants and not a result of years of strain formation. Our results indicate the CF lung environment supports stably diverged populations of clonally derived yeasts.

Data availability

Sequence data generated for isolates with Illumina and Oxford Nanopore technology are deposited in NCBI Sequence Read Archive linked under BioProject PRJNA628510. The assembled genomes of each CF isolate (Ex1-23) are available under accessions listed in **Supplemental Table 2**. The assembled and annotated genome of the in-population reference isolate (DCF04) is available at accession JAJGCF000000000. All analysis pipelines, custom scripts used for data analysis, and raw variant data in the variant call format are available in Github repository https://github.com/tania-k/CF_Exophiala_dermatitidis and archived in Zenodo (Kurbessoian 2022).

Acknowledgments

This work was supported by NIH grants R01 AI127548 (D.A.H. and J.E.S.) and T32 AI007519 (D.R.M.). JES was partially supported by R01 AI130128. This project was also supported by DartCF (NIH grant P30-DK117469) and the Cystic Fibrosis Foundation Research Development Program (STANTO19R0). Computational analyses were performed at the University of California-Riverside HPCC supported by grants from the National Science Foundation (MRI-1429826) and NIH (S10OD016290). JES is a CIFAR Fellow in the Fungal Kingdom: Threats and Opportunities program.

Figure Legends

Figure 1. FEV1 pulmonary function data. FEV1 pulmonary function data was tested over the course of three years, and sputum cultures were acquired at each indicated time point and assessed for the presence of pathogens other than the mixed bacteria that compose normal upper respiratory flora. Colored bars indicate the duration of treatment for each listed antimicrobial given during infection.

Figure 2. Mating-type determination of 4 clinical isolates of *E. dermatitidis.* Gene content, order and orientation of the MAT locus from *E. dermatitidis* DCF04 and NIH/UT8656 strains and the Chaetothyriales black yeast *Capronia coronata CBS 617.98*. The locus is flanked by two genes SLA2 (purple) and APN2 (green). The MAT 1-1 genes, MAT 1-1-4 (orange) and MAT 1-1-1 (yellow) are observed in DCF04 strains while the reference strain NIH/UT8656 has a MAT 1-2 (teal) gene. In addition, two genes are predicted in the interval between SLA2 and the MAT genes. One is a L-type calcium channel domain (blue), which is only found in the *E. dermatitidis* strains, and a gene with no identified function or domains (pink) which is syntentically adjacent to the MAT genes in all strains and species.

Figure 3. Genome sequencing depth coverage visualization to test for copy number variation across isolates. Visualization of depth of coverage was generated by plotting a normalized read depth across chromosomes for isolates **A)** Ex 1-8, **B)** Ex9-16, and **C)** Ex17-23.

Figure 4: Phylogenetic tree of 23 isolates. A Maximum-Likelihood phylogenetic tree constructed from the Single Nucleotide Variants by IQTREE2 identified from the isolate resequencing data. Tree is rooted with the Clade I branch based on additional analyses that included NIH/UT8656 as an outgroup. Isolates are labeled as one of three clades based on the phylogenetic relationships and the isolation time point is indicated with a red (early) or blue (late) colored box.

Figure 5. Matrix of *E. dermatitidis* CF isolates SNP and INDEL pairwise dissimilarities. The number of SNPs **(A)** and **(B)** INDELs that differ among pairs of isolates. The more similar isolate pairs have lower (darker red) numbers and the color approaches white indicating more dissimilar isolates. Clade I and Clade II isolates generally differ by very few SNP and INDELs (noting a few exceptions) consistent with their inferred near phylogenetic relationships. Within Clade III isolate pairs differ by more SNPs which may indicate a higher mutation rate within these isolates.

Figure 6. *E. dermatitidis* clades and their comparisons to melanin phenotype and MIC. (A) Isolates were struck from freezer stock onto YPD plates and imaged at 120 h at 37˚C to compare melanin accumulation. **(B)** Strong differences in melanin production between related isolates may be due to intergenic mutations in the 5' UTR of a biphenyl-2,3-dioxygenase encoding gene.

Figure 7. Itraconazole MIC testing on 23 isolates of *E. dermatitidis,* while comparing clades. Isolates were grown for ~24 hours in liquid YPD at 30˚C in a rolling barrel culture, then inoculated into 96-well flat bottom plates at a concentration of 1000 CFU/well, and allowed to grow for 72 hours at 37˚ before determining MIC, which is represented on the figure as the mean of three biological replicates. For determination of auxotrophy, cells were grown as described above, CFU equilibrated, and spotted onto rich media and minimal media (YNB) with and without the addition of casamino acids. isolates with decreased growth on YNB relative to YPD, and could be rescued with the addition of amino acids are noted as auxotrophic.

Figure 8. Cell phenotype microscopy. Isolates were grown for 18 h in RPMI, at 37˚ in 5% CO, and imaged at 63X using DIC microscopy$_2$. The first row of images depicts isolates displaying hyper-filamentous attributes, and are distributed throughout their respective clades. The second row depicts the closest related isolate to each respective isolate in the first row, which lack similar phenotypes.

Table 1. Genome assembly summary statistics for reference isolate *E. dermatitidis* DCF04. Table indicates assembly summary statistics of the hybrid assembly using Illumina and Nanopore sequencing of isolate DCF04. Summary statistics of contig and scaffold lengths are presented with genome completeness data calculated with BUSCO using the ascomycota_odb9 database.

Table 2. Genes with SNP variants found stratified by clade. Isolates collected in the early (red) and late (blue) time points are labeled in the comparisons. The functions of genes of interest found to have nonsynonymous mutations among entire clade of comparisons are reported with the total count of nonsynonymous and synonymous (NonSyn/Syn) variants found in each clade to identify the relative frequency of these changes and general functional differences in the types of genes with variants across the population clades.

Table 3. Genes with nonsynonymous SNP variants of interest observed in pairwise comparisons of *E. dermatitidis* CF lung isolates. Genes with notable non-synonymous differences identified in pairwise comparisons of early and late time point isolates. Full results of all pairwise differences are in **Supplemental Table 7**. For continuity, gene locus names are in parentheses and refer to the locus names in the reference strain *E. dermatitidis* NIH/UT8656.

Table 4. Average mutation rates for each clade/group. Mutation rates calculated for each clade of isolates and two sub-groups of Clade III. Calculations were summarized as the median of all pairs of isolates within a group or clade.

Supplemental Figure 1. Genome dot plot of *E. dermatitidis* DCF04 and *E. dermatitidis* NIH/UT8656. D-GENIES web application was used to generate a dot plot representation to compare genome assembly content and test for structural changes or rearrangements. Both iterations of the plot indicate a majority of the genomes are nearly identical with each other.

Supplemental Figure 2. Mating-type determination of 23 clinical isolates of *E. dermatitidis.* Gene content, order and orientation of the MAT locus from 23 CF isolated *E. dermatitidis*, NIH/UT8656 strain and the Chaetothyriales black yeast *Capronia coronata CBS 617.98*. The locus is flanked by two genes SLA2 (purple) and APN2 (green). The MAT 1-1 genes, MAT 1-1-4 (orange) and MAT 1-1-1 (yellow) are observed in all 23 CF strains while the reference strain NIH/UT8656 has a MAT 1-2 (teal) gene. In addition, two genes are predicted in the interval between SLA2 and the MAT genes. One is a L-type calcium channel domain (blue), which is only found in the *E. dermatitidis* strains, and a gene with no identified function or domains (pink) which is syntentically adjacent to the MAT genes in some strains.

Supplemental Figure 3. Phylogenetic tree of 24 isolates. A Maximum-Likelihood phylogenetic tree constructed from the Single Nucleotide Variants by IQTREE2 identified from the isolate resequencing data. Tree is rooted with NIH/UT8656 as an outgroup. Isolates are labeled as one of three clades based on the phylogenetic relationships and the isolation time point is indicated with a red (early) or blue (late) colored box.

Supplemental Table 1. Collection and MIC values for CF patient derived *E. dermatitidis* isolates. Information on the 23 CF isolates cultured from one patient sputum across three years. The table summarizes phylogenetic clade designation, itraconazole MIC, date of collection, and classification as an Early or Late.

Supplemental Table 2. Genome assembly statistics for 23 CF isolates. Assembly statistics for scaffolded assemblies of isolates and NIH/UT8656 previously published genome. Table indicates assembly summary statistics of the assembly using Illumina sequencing of all 23 CF isolates. Summary statistics of scaffold lengths are presented with genome completeness data calculated with BUSCO using the ascomycota_odb9 database.

Supplemental Table 3. Telomere Recovery for DCF04. Table depicting telomere recovery results for DCF04 CF *E. dermatitidis*. Candidate telomeric repeat units "TTTAGGG/CCCTAA" were identified as repeat arrays at both ends of five scaffolds, but also found as single pairs in the remaining 4 scaffolds, as would be expected for 9 complete chromosomes.

Supplemental Table 4. OrthoFinder summary comparing DCF04 and NIH/UT8656. Comparison of the shared and unique orthogroups found in the annotated proteomes of *E. dermatitidis* strains NIH/UT8656 and DCF04. A majority of orthogroups (8,256; 99%) had members from both strains, of these 15 were single-copy orthogroups containing a single protein-coding gene from each strain. There were 10 orthogroups unique to DCF04 encompassing 34 protein-coding genes and 5 orthogroups unique to NIH/UT8656 made up of 24 protein-coding genes. DCF04 contained 705 unassigned genes, while NIH/UT8656 had 475.

Supplemental Table 5. Mating-type determination locus name descriptions of *E. dermatitidis*. This table describes all the proteins depicted in Figure 2 of the gene content, order and orientation of MAT locus from *E. dermatitidis* isolates in this study, strain NIH/UT8656 along with the Chaetothyriales black yeast *Capronia coronata CBS 617.98*.

Supplemental Table 6. Functional impact of identified variants. Filtered human readable snpEff tabular results for all CF 23 *E. dermatitidis*. Annotations have been added to the final list to better help assess the function of each protein while also observing variants detected from GATK.

Supplemental Table 7. All functional SNP and INDEL results for early and late pairs. Results indicated in this table include all hypothetical or undescribed results along with results described in Tables 2 and 3. Proteins listed include Protein ID numbers to better facilitate identification.

Supplemental Table 8. All 23 mutation rates calculated. Mutation rates for each 23 CF *E. dermatitidis* isolates calculated using formula described in methods. One-way ANOVA was run to detect significance (p-value= 0.00104, F-value = 7.0711) along with Tukey multiple comparisons of means indicating Clade III had the highest significance among the six pairwise comparisons.

References

Alonge M, Soyk S, Ramakrishnan S, Wang X, Goodwin S, Sedlazeck FJ, Lippman ZB, Schatz MC. 2019. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. Genome Biol. 20(1):224. doi:10.1186/s13059-019-1829-6. http://dx.doi.org/10.1186/s13059-019-1829-6.

Aznar A, Chen NWG, Rigault M, Riache N, Joseph D, Desmaële D, Mouille G, Boutet S, Soubigou-Taconnat L, Renou J-P, et al. 2014. Scavenging iron: a novel mechanism of plant immunity activation by microbial siderophores. Plant Physiol. 164(4):2167–2183. doi:10.1104/pp.113.233585. http://dx.doi.org/10.1104/pp.113.233585.

Babič MN, Zupančič J, Gunde-Cimerman N, de Hoog S, Zalar P. 2018. Ecology of the Human Opportunistic Black Yeast *Exophiala dermatitidis* Indicates Preference for Human-Made Habitats. Mycopathologia. 183(1):201–212. doi:10.1007/s11046-017-0134-8. http://dx.doi.org/10.1007/s11046-017-0134-8.

829 Banin E, Vasil ML, Peter Greenberg E. 2005. Iron and *Pseudomonas aeruginosa* biofilm formation.
830 Proceedings of the National Academy of Sciences. 102(31):11076–11081.
831 doi:10.1073/pnas.0504266102. http://dx.doi.org/10.1073/pnas.0504266102.

832 Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI,
833 Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its
834 applications to single-cell sequencing. J Comput Biol. 19(5):455–477.
835 doi:10.1089/cmb.2012.0021. http://dx.doi.org/10.1089/cmb.2012.0021.

836 Bao W, Kojima KK, Kohany O. 2015. Repbase Update, a database of repetitive elements in eukaryotic
837 genomes. Mob DNA. 6:11. doi:10.1186/s13100-015-0041-9. http://dx.doi.org/10.1186/s13100-
838 015-0041-9.

839 Barth AL, Pitt TL. 1995. Auxotrophic variants of *Pseudomonas aeruginosa* are selected from
840 prototrophic wild-type strains in respiratory infections in patients with cystic fibrosis. Journal of
841 Clinical Microbiology. 33(1):37–40. doi:10.1128/jcm.33.1.37-40.1995.
842 http://dx.doi.org/10.1128/jcm.33.1.37-40.1995.

843 Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. 2021.
844 antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res.
845 49(W1):W29–W35. doi:10.1093/nar/gkab335. http://dx.doi.org/10.1093/nar/gkab335.

846 Brůna T, Lomsadze A, Borodovsky M. 2020. GeneMark-EP : eukaryotic gene prediction with self-
847 training in the space of genes and proteins. NAR Genomics and Bioinformatics. 2(2).
848 doi:10.1093/nargab/lqaa026. http://dx.doi.org/10.1093/nargab/lqaa026.

849 Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat
850 Methods. 12(1):59–60. doi:10.1038/nmeth.3176. http://dx.doi.org/10.1038/nmeth.3176.

851 Burns JL, Emerson J, Stapp JR, Yim DL, Krzewinski J, Louden L, Ramsey BW, Clausen CR. 1998.
852 Microbiology of sputum from patients at cystic fibrosis centers in the United States. Clin Infect
853 Dis. 27(1):158–163. doi:10.1086/514631. http://dx.doi.org/10.1086/514631.

854 Bushnell B. 2014. BBMap: A fast, accurate, splice-aware aligner. Lawrence Berkeley National Lab.
855 (LBNL), Berkeley, CA (United States) Report No.: LBNL-7065E.
856 https://www.osti.gov/biblio/1241166.

857 Cabanettes F, Klopp C. 2018. D-GENIES: dot plot large genomes in an interactive, efficient and simple
858 way. PeerJ. 6:e4958. doi:10.7717/peerj.4958. http://dx.doi.org/10.7717/peerj.4958.

859 Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK:
860 rising to the challenge of larger and richer datasets. Gigascience. 4:7. doi:10.1186/s13742-015-
861 0047-8. http://dx.doi.org/10.1186/s13742-015-0047-8.

862 Chen M, Kondori N, Deng S, van den Ende AHGG, Lackner M, Liao W, de Hoog GS. 2018. Direct
863 detection of *Exophiala* and *Scedosporium* species in sputa of patients with cystic fibrosis.
864 Medical Mycology. 56(6):695–702. doi:10.1093/mmy/myx108.
865 http://dx.doi.org/10.1093/mmy/myx108.

866 Chen Z, Martinez DA, Gujja S, Sykes SM, Zeng Q, Szaniszlo PJ, Wang Z, Cuomo CA. 2014.
867 Comparative genomic and transcriptomic analysis of *Wangiella dermatitidis*, a major cause of
868 phaeohyphomycosis and a model black yeast human pathogen. G3. 4(4):561–578.
869 doi:10.1534/g3.113.009241. http://dx.doi.org/10.1534/g3.113.009241.

870 Cimon B, Carrère J, Vinatier JF, Chazalette JP, Chabasse D, Bouchara JP. 2000. Clinical Significance of
871 *Scedosporium apiospermum* in Patients with Cystic Fibrosis. European Journal of Clinical
872 Microbiology & Infectious Diseases. 19(1):53–56. doi:10.1007/s100960050011.
873 http://dx.doi.org/10.1007/s100960050011.

874 Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A
875 program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff.
876 Fly 6: 80--92. doi:10.4161/fly.19695. http://dx.doi.org/10.4161/fly.19695.

877 Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F,
878 Wilczynski B, et al. 2009. Biopython: freely available Python tools for computational molecular
879 biology and bioinformatics. Bioinformatics. 25(11):1422–1423.
880 doi:10.1093/bioinformatics/btp163. http://dx.doi.org/10.1093/bioinformatics/btp163.

881 Conti-Díaz IA, Mackinnon JE, Civila E. 1977. Isolation and identification of black yeasts from the
882 external environment in Uruguay. Pan Amer Health Org Sci Publ. 356:109–114.
883 https://books.google.com/books?hl=en&lr=&id=nogTAQAAMAAJ&oi=fnd&pg=RA1-
884 PA109&dq=Isolation+and+identification+of+black+yeasts+from+the+external+environment+in+
885 Uruguay&ots=PQwiT-q-DM&sig=2i5R398cOJG6e0u7824a0whsG0U.

886 Crowley DE, Wang YC, Reid CPP, Szaniszlo PJ. 1991. Mechanisms of iron acquisition from
887 siderophores by microorganisms and plants. In: Chen Y, Hadar Y, editors. Iron Nutrition and
888 Interactions in Plants, 11–17 June 1989, Jerusalem, Israel, 1989. Dordrecht: Springer
889 Netherlands. p. 213–232. https://doi.org/10.1007/978-94-011-3294-7_27.

890 Damveld RA, Arentshorst M, Franken A, vanKuyk PA, Klis FM, van den Hondel CAMJJ, Ram AFJ.
891 2005. The *Aspergillus niger* MADS-box transcription factor RlmA is required for cell wall
892 reinforcement in response to cell wall stress. Mol Microbiol. 58(1):305–319. doi:10.1111/j.1365-
893 2958.2005.04827.x. http://dx.doi.org/10.1111/j.1365-2958.2005.04827.x.

894 Davis PB. 2006. Cystic Fibrosis Since 1938. American Journal of Respiratory and Critical Care
895 Medicine. 173(5):475–482. doi:10.1164/rccm.200505-840oe.
896 http://dx.doi.org/10.1164/rccm.200505-840oe.

Defontaine A, Zouhair R, Cimon B, Carrère J, Bailly E, Symoens F, Diouri M, Hallet J-N, Bouchara J-P. 2002. Genotyping Study of *Scedosporium apiospermum* Isolates from Patients with Cystic Fibrosis. Journal of Clinical Microbiology. 40(6):2108–2114. doi:10.1128/jcm.40.6.2108-2114.2002. http://dx.doi.org/10.1128/jcm.40.6.2108-2114.2002.

Demers EG, Biermann AR, Masonjones S, Crocker AW, Ashare A, Stajich JE, Hogan DA. 2018. Evolution of drug resistance in an antifungal-naive chronic *Candida lusitaniae* infection. Proceedings of the National Academy of Sciences. 115(47):12040–12045. doi:10.1073/pnas.1807698115. http://dx.doi.org/10.1073/pnas.1807698115.

Diemert D, Kunimoto D, Sand C, Rennie R. 2001. Sputum isolation of *Wangiella dermatitidis* in patients with cystic fibrosis. Scand J Infect Dis. 33(10):777–779. doi:10.1080/003655401317074644. http://dx.doi.org/10.1080/003655401317074644.

Dixon DM, Shadomy HJ, Shadomy S. 1980. Dematiaceous fungal pathogens isolated from nature. Mycopathologia. 70(3):153–161. doi:10.1007/BF00443026. http://dx.doi.org/10.1007/BF00443026.

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20(1):238. doi:10.1186/s13059-019-1832-y. http://dx.doi.org/10.1186/s13059-019-1832-y.

Ferec C, Cutting GR. 2012. Assessing the Disease-Liability of Mutations in CFTR. Cold Spring Harb Perspect Med. 2(12):a009480. doi:10.1101/cshperspect.a009480. http://dx.doi.org/10.1101/cshperspect.a009480.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, et al. 2014. Pfam: the protein families database. Nucleic Acids Res. 42(Database issue):D222–30. doi:10.1093/nar/gkt1223. http://dx.doi.org/10.1093/nar/gkt1223.

Fothergill AW, Rinaldi MG, Sutton DA. 2009. Antifungal susceptibility testing of *Exophiala spp*.: a head-to-head comparison of amphotericin B, itraconazole, posaconazole and voriconazole. Med Mycol. 47(1):41–43. doi:10.1080/13693780802512451. http://dx.doi.org/10.1080/13693780802512451.

Frank W, Roester U, Others. 1970. Amphibia as carriers of *Hormiscium (Hormodendrum) dermatitidis* a causative agent of chromo-blastomycosis (chromomycosis) in man. Z Tropenmed Parasitol. 21(1):93–108. https://www.cabdirect.org/cabdirect/abstract/19702902020.

Franke KR, Crowgey EL. 2020. Accelerating next generation sequencing data analysis: an evaluation of optimized best practices for Genome Analysis Toolkit algorithms. Genomics Inform. 18(1):e10. doi:10.5808/GI.2020.18.1.e10. http://dx.doi.org/10.5808/GI.2020.18.1.e10.

Freire-Benéitez V, Gourlay S, Berman J, Buscaino A. 2016. Sir2 regulates stability of repetitive domains differentially in the human fungal pathogen Candida albicans. Nucleic Acids Res. 44(19):9166–9179. doi:10.1093/nar/gkw594. https://academic.oup.com/nar/article-abstract/44/19/9166/2468381.

Gilchrist CLM, Booth TJ, van Wersch B, van Grieken L, Medema MH, Chooi Y-H. 2021. cblaster: a remote search tool for rapid identification and visualization of homologous gene clusters. Bioinformatics Advances. 1(1):vbab016. doi:10.1093/bioadv/vbab016. https://academic.oup.com/bioinformaticsadvances/article-abstract/1/1/vbab016/6342405.

Gilchrist CLM, Chooi Y-H. 2021 Jan 18. Clinker & clustermap.js: Automatic generation of gene cluster comparison figures. Bioinformatics. doi:10.1093/bioinformatics/btab007. http://dx.doi.org/10.1093/bioinformatics/btab007.

941 Goldman GH, McGuire SL, Harris SD. 2002. The DNA damage response in filamentous fungi. Fungal
942 Genet Biol. 35(3):183–195. doi:10.1006/fgbi.2002.1344.
943 http://dx.doi.org/10.1006/fgbi.2002.1344.

944 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury
945 R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a
946 reference genome. Nat Biotechnol. 29(7):644–652. doi:10.1038/nbt.1883.
947 http://dx.doi.org/10.1038/nbt.1883.

948 Grahl N, Dolben EL, Filkins LM, Crocker AW, Willger SD, Morrison HG, Sogin ML, Ashare A, Gifford
949 AH, Jacobs NJ, et al. 2018. Profiling of Bacterial and Fungal Microbial Communities in Cystic
950 Fibrosis Sputum Using RNA. mSphere. 3(4). doi:10.1128/mSphere.00292-18.
951 http://dx.doi.org/10.1128/mSphere.00292-18.

952 Grenouillet F, Cimon B, Pana-Katatali H, Person C, Gainet-Brun M, Malinge M-C, Le Govic Y,
953 Richaud-Thiriez B, Bouchara J-P. 2018. *Exophiala dermatitidis* Revealing Cystic Fibrosis in
954 Adult Patients with Chronic Pulmonary Disease. Mycopathologia. 183(1):71–79.
955 doi:10.1007/s11046-017-0218-5. http://dx.doi.org/10.1007/s11046-017-0218-5.

956 Griffard EA, Guajardo JR, Cooperstock MS, Scoville CL. 2010. Isolation of *Exophiala dermatitidis* from
957 pigmented sputum in a cystic fibrosis patient. Pediatr Pulmonol. 45(5):508–510.
958 doi:10.1002/ppul.21187. http://dx.doi.org/10.1002/ppul.21187.

959 Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008.
960 Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to
961 Assemble Spliced Alignments. Genome Biol. 9(1):R7. doi:10.1186/gb-2008-9-1-r7.
962 http://dx.doi.org/10.1186/gb-2008-9-1-r7.

963 Haas H. 2014. Fungal siderophore metabolism with a focus on *Aspergillus fumigatus*. Nat Prod Rep.
964 31(10):1266–1276. doi:10.1039/c4np00071d. http://dx.doi.org/10.1039/c4np00071d.

965 Haase G, Skopnik H, Kusenbach G. 1990. *Exophiala dermatitidis* infection in cystic fibrosis. Lancet.
966 336(8708):188–189. doi:10.1016/0140-6736(90)91721-l. http://dx.doi.org/10.1016/0140-
967 6736(90)91721-l.

968 Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the Ultrafast
969 Bootstrap Approximation. Mol Biol Evol. 35(2):518–522. doi:10.1093/molbev/msx281.
970 http://dx.doi.org/10.1093/molbev/msx281.

971 Hoog GS de, Queiroz-Telles F, Haase G, Fernandez-Zeppenfeldt G, Angelis DA, H. G. Gerrits van den
972 Ende A, Matos T, Peltroche-Llacsahuanga H, Pizzirani-Kleiner AA, Rainer J, et al. 2000. Black
973 fungi: clinical and pathogenic approaches. Med Mycol. 38(sup1):243–250.
974 doi:10.1080/mmy.38.s1.243.250.
975 https://www.tandfonline.com/doi/abs/10.1080/mmy.38.s1.243.250.

976 Horré R, Marklein G, Siekmeier R, Nidermajer S, Reiffert SM. 2009. Selective isolation of
977 *Pseudallescheria* and *Scedosporium* species from respiratory tract specimens of cystic fibrosis
978 patients. Respiration. 77(3):320–324. doi:10.1159/000167419.
979 http://dx.doi.org/10.1159/000167419.

980 Horré R, Schaal KP, Siekmeier R, Sterzik B, de Hoog GS, Schnitzler N. 2004. Isolation of Fungi,
981 Especially *Exophiala dermatitidis*, in Patients Suffering from Cystic Fibrosis. Respiration.
982 71(4):360–366. doi:10.1159/000079640. http://dx.doi.org/10.1159/000079640.

983 Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic
984 I, Rattei T, Jensen LJ, et al. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically

985 annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res.
986 47(D1):D309–D314. doi:10.1093/nar/gky1085. http://dx.doi.org/10.1093/nar/gky1085.

987 Institute. “Picard Toolkit.” Broad Institute, GitHub Repository. Picard Toolkit.
988 https://github.com/broadinstitute/picard

989 Johnson L. 2008. Iron and siderophores in fungal–host interactions. Mycol Res. 112(2):170–183.
990 doi:10.1016/j.mycres.2007.11.012.
991 https://www.sciencedirect.com/science/article/pii/S0953756207002948.

992 Jones JT, Liu K-W, Wang X, Kowalski CH, Ross BS, Mills KAM, Kerkaert JD, Hohl TM, Lofgren LA,
993 Stajich JE, et al. 2020. Strain specific persistence in the murine lung of *Aspergillus fumigatus*
994 conidia causes an Allergic Broncho-Pulmonary Aspergillosis-like disease phenotype. bioRxiv.
995 doi:10.1101/2020.12.04.412726. https://www.biorxiv.org/content/10.1101/2020.12.04.412726.

996 Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka
997 G, et al. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics.
998 30(9):1236–1240. doi:10.1093/bioinformatics/btu031.
999 http://dx.doi.org/10.1093/bioinformatics/btu031.

1000 Jong CCM de, de Jong CCM, Slabbers L, Engel TGP, Yntema JB, van Westreenen M, Croughs PD,
1001 Roeleveld N, Brimicombe R, Verweij PE, et al. 2020. Clinical relevance of Scedosporium spp.
1002 and *Exophiala dermatitidis* in patients with cystic fibrosis: A nationwide study. Medical
1003 Mycology. 58(7):859–866. doi:10.1093/mmy/myaa003. http://dx.doi.org/10.1093/mmy/myaa003.

1004 Kim DH. 2018. Bacterial Siderophores Promote Animal Host Iron Acquisition and Growth. Cell.
1005 175(2):311–312. doi:10.1016/j.cell.2018.09.020. http://dx.doi.org/10.1016/j.cell.2018.09.020.

1006 Kim SH, Clark ST, Surendra A, Copeland JK, Wang PW, Ammar R, Collins C, Tullis DE, Nislow C,
1007 Hwang DM, et al. 2015. Global Analysis of the Fungal Microbiome in Cystic Fibrosis Patients
1008 Reveals Loss of Function of the Transcriptional Repressor Nrg1 as a Mechanism of Pathogen
1009 Adaptation. PLoS Pathog. 11(11):e1005308. doi:10.1371/journal.ppat.1005308.
1010 http://dx.doi.org/10.1371/journal.ppat.1005308.

1011 Kirchhoff L, Olsowski M, Rath P-M, Steinmann J. 2019. *Exophiala dermatitidis*: Key issues of an
1012 opportunistic fungal pathogen. Virulence. 10(1):984–998. doi:10.1080/21505594.2019.1596504.
1013 http://dx.doi.org/10.1080/21505594.2019.1596504.

1014 Klasinc R, Riesenhuber M, Bacher A, Willinger B. 2019. Invasive Fungal Infection Caused by *Exophiala*
1015 *dermatitidis* in a Patient After Lung Transplantation: Case Report and Literature Review.
1016 Mycopathologia. 184(1):107–113. doi:10.1007/s11046-018-0275-4.
1017 http://dx.doi.org/10.1007/s11046-018-0275-4.

1018 Kondori N, Gilljam M, Lindblad A, Jönsson B, Moore ERB, Wennerås C. 2011. High rate of *Exophiala*
1019 *dermatitidis* recovery in the airways of patients with cystic fibrosis is associated with pancreatic
1020 insufficiency. J Clin Microbiol. 49(3):1004–1009. doi:10.1128/JCM.01899-10.
1021 https://journals.asm.org/doi/10.1128/JCM.01899-10.

1022 Kondori N, Lindblad A, Welinder-Olsson C, Wennerås C, Gilljam M. 2014. Development of IgG
1023 antibodies to *Exophiala dermatitidis* is associated with inflammatory responses in patients with
1024 cystic fibrosis. J Cyst Fibros. 13(4):391–399. doi:10.1016/j.jcf.2013.12.007.
1025 http://dx.doi.org/10.1016/j.jcf.2013.12.007.

1026 Korf I. 2004. Gene finding in novel genomes. BMC Bioinformatics. 5:59. doi:10.1186/1471-2105-5-59.
1027 http://dx.doi.org/10.1186/1471-2105-5-59.

1028 Krogh BO, Symington LS. 2004. Recombination proteins in yeast. Annu Rev Genet. 38:233–271.
1029 doi:10.1146/annurev.genet.38.072902.091500.
1030 http://dx.doi.org/10.1146/annurev.genet.38.072902.091500.

1031 Kurbessoian T. 2022. tania-k/CF_Exophiala_dermatitidis: Version 6 update.
1032 https://zenodo.org/record/7106110.

1033 Kusenbach G, Skopnik H, Haase G, Friedrichs F, Döhmen H. 1992. *Exophiala dermatitidis* pneumonia in
1034 cystic fibrosis. Eur J Pediatr. 151(5):344–346. doi:10.1007/BF02113255.
1035 http://dx.doi.org/10.1007/BF02113255.

1036 Lavrin T, Konte T, Kostanjšek R, Sitar S, Sepčič K, Prpar Mihevc S, Žagar E, Župunski V, Lenassi M,
1037 Rogelj B, et al. 2020. The Neurotropic Black Yeast *Exophiala dermatitidis* Induces
1038 Neurocytotoxicity in Neuroblastoma Cells and Progressive Cell Death. Cells. 9(4).
1039 doi:10.3390/cells9040963. http://dx.doi.org/10.3390/cells9040963.

1040 Legrand M, Jaitly P, Feri A, d'Enfert C, Sanyal K. 2019. *Candida albicans*: An Emerging Yeast Model to
1041 Study Eukaryotic Genome Plasticity. Trends Genet. 35(4):292–307.
1042 doi:10.1016/j.tig.2019.01.005. http://dx.doi.org/10.1016/j.tig.2019.01.005.

1043 Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of
1044 phylogenetic and other trees. Nucleic Acids Res. 44(W1):W242–5. doi:10.1093/nar/gkw290.
1045 http://dx.doi.org/10.1093/nar/gkw290.

1046 Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform.
1047 Bioinformatics. 26(5):589–595. doi:10.1093/bioinformatics/btp698.
1048 https://academic.oup.com/bioinformatics/article-abstract/26/5/589/211735.

1049 Li H, Handsaker B, Danecek P, McCarthy S, Marshall J. 2019. BCFtools.
1050 https://github.com/samtools/bcftools

1051 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000
1052 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and
1053 SAMtools. Bioinformatics. 25(16):2078–2079. doi:10.1093/bioinformatics/btp352.
1054 http://dx.doi.org/10.1093/bioinformatics/btp352.

1055 Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer
1056 RNA genes. Nucleic Acids Res. 44(W1):W54–7. doi:10.1093/nar/gkw413.
1057 http://dx.doi.org/10.1093/nar/gkw413.

1058 Majoros WH, Pertea M, Salzberg SL. 2004. TigrScan and GlimmerHMM: two open source ab initio
1059 eukaryotic gene-finders. Bioinformatics. 20(16):2878–2879. doi:10.1093/bioinformatics/bth315.
1060 http://dx.doi.org/10.1093/bioinformatics/bth315.

1061 Malo ME, Schultzhaus Z, Frank C, Romsdahl J, Wang Z, Dadachova E. 2021. Transcriptomic and
1062 genomic changes associated with radio-adaptation in *Exophiala dermatitidis*. Comput Struct
1063 Biotechnol J. 19:196–205. doi:10.1016/j.csbj.2020.12.013.
1064 http://dx.doi.org/10.1016/j.csbj.2020.12.013.

1065 Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO Update: Novel and
1066 Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of
1067 Eukaryotic, Prokaryotic, and Viral Genomes. Mol Biol Evol. 38(10):4647–4654.
1068 doi:10.1093/molbev/msab199. http://dx.doi.org/10.1093/molbev/msab199.

1069 Mariani-Kurkdjian P, Bingen E. 2003. Bactéries pathogènes dans la mucoviscidose. Archives de
1070 Pédiatrie. 10:S342–S346. doi:10.1016/S0929-693X(03)90050-9.
1071 https://www.sciencedirect.com/science/article/pii/S0929693X03900509.

Matilla MA, Ramos JL, Duque E, de Dios Alché J, Espinosa-Urgel M, Ramos-González MI. 2007. Temperature and pyoverdine-mediated iron acquisition control surface motility of *Pseudomonas putida*. Environmental Microbiology. 9(7):1842–1850. doi:10.1111/j.1462-2920.2007.01286.x. http://dx.doi.org/10.1111/j.1462-2920.2007.01286.x.

Matos T, de Hoog GS, de Boer AG, de Crom I, Haase G. 2002. High prevalence of the neurotrope *Exophiala dermatitidis* and related oligotrophic black yeasts in sauna facilities. Mycoses. 45(9-10):373–377. doi:10.1046/j.1439-0507.2002.00779.x. http://dx.doi.org/10.1046/j.1439-0507.2002.00779.x.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20(9):1297–1303. doi:10.1101/gr.107524.110. http://dx.doi.org/10.1101/gr.107524.110.

Metin B, Döğen A, Yıldırım E, de Hoog GS, Heitman J, Ilkit M. 2019. Mating type (MAT) locus and possible sexuality of the opportunistic pathogen *Exophiala dermatitidis*. Fungal Genetics and Biology. 124:29–38. doi:10.1016/j.fgb.2018.12.011. http://dx.doi.org/10.1016/j.fgb.2018.12.011.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 37(5):1530–1534. doi:10.1093/molbev/msaa015. http://dx.doi.org/10.1093/molbev/msaa015.

Mochochoko BM, Ezeokoli OT, Sebolai O, Albertyn J, Pohl CH. 2021. Role of the high-affinity reductive iron acquisition pathway of *Candida albicans* in prostaglandin E2 production, virulence, and interaction with *Pseudomonas aeruginosa*. Medical Mycology. 59(9):869–881. doi:10.1093/mmy/myab015. http://dx.doi.org/10.1093/mmy/myab015.

1095 Mossialos D, Amoutzias GD. 2009. Role of siderophores in cystic fibrosis pathogenesis: foes or friends?
1096 Int J Med Microbiol. 299(2):87–98. doi:10.1016/j.ijmm.2008.06.008.
1097 http://dx.doi.org/10.1016/j.ijmm.2008.06.008.

1098 Mukai Y, Nureki S-I, Hata M, Shigenaga T, Tokimatsu I, Miyazaki E, Kadota J-I, Yarita K, Kamei K.
1099 2014. *Exophiala dermatitidis* pneumonia successfully treated with long-term itraconazole
1100 therapy. J Infect Chemother. 20(7):446–449. doi:10.1016/j.jiac.2014.02.006.
1101 http://dx.doi.org/10.1016/j.jiac.2014.02.006.

1102 Nagano Y, Cherie Millar B, Johnson E, Goldsmith CE, Stuart Elborn J, Rendall J, Moore JE. 2007.
1103 Fungal infections in patients with cystic fibrosis. Reviews in Medical Microbiology. 18(1):11–16.
1104 doi:10.1097/mrm.0b013e3282e1c70a. http://dx.doi.org/10.1097/mrm.0b013e3282e1c70a.

1105 Neilands JB. 1995. Siderophores: structure and function of microbial iron transport compounds. J Biol
1106 Chem. 270(45):26723–26726. doi:10.1074/jbc.270.45.26723.
1107 http://dx.doi.org/10.1074/jbc.270.45.26723.

1108 Nishimura K, Miyaji M. 1982. Studies on a saprophyte of *Exophiala dermatitidis* isolated from a
1109 humidifier. Mycopathologia. 77(3):173–181. doi:10.1007/bf00518803.
1110 http://dx.doi.org/10.1007/bf00518803.

1111 Nosanchuk JD, Casadevall A. 2003. The contribution of melanin to microbial pathogenesis. Cell
1112 Microbiol. 5(4):203–223. doi:10.1046/j.1462-5814.2003.00268.x.
1113 http://dx.doi.org/10.1046/j.1462-5814.2003.00268.x.

1114 Nosanchuk JD, Casadevall A. 2006. Impact of Melanin on Microbial Virulence and Clinical Resistance to
1115 Antimicrobial Compounds. Antimicrobial Agents and Chemotherapy. 50(11):3519–3528.
1116 doi:10.1128/aac.00545-06. http://dx.doi.org/10.1128/aac.00545-06.

1117 Packeu A, Lebecque P, Rodriguez-Villalobos H, Boeras A, Hendrickx M, Bouchara J-P, Symoens F.
1118 2012. Molecular typing and antifungal susceptibility of *Exophiala* isolates from patients with
1119 cystic fibrosis. J Med Microbiol. 61(Pt 9):1226–1233. doi:10.1099/jmm.0.042317-0.
1120 http://dx.doi.org/10.1099/jmm.0.042317-0.

1121 Palmer JM, Stajich J. 2020. Funannotate v1.8.1: Eukaryotic genome annotation.
1122 https://zenodo.org/record/4054262.

1123 Palmer JM, Stajich JE. 2022. Automatic assembly for the fungi (AAFTF): genome assembly pipeline.
1124 https://zenodo.org/record/6326242.

1125 Paolo WF Jr, Dadachova E, Mandal P, Casadevall A, Szaniszlo PJ, Nosanchuk JD. 2006. Effects of
1126 disrupting the polyketide synthase gene WdPKS1 in *Wangiella [Exophiala] dermatitidis* on
1127 melanin production and resistance to killing by antifungal compounds, enzymatic degradation,
1128 and extremes in temperature. BMC Microbiol. 6:55. doi:10.1186/1471-2180-6-55.
1129 http://dx.doi.org/10.1186/1471-2180-6-55.

1130 Parize P, Billaud S, Bienvenu AL, Bourdy S, le Pogam MA, Reix P, Picot S, Robert R, Lortholary O,
1131 Bouchara J-P, et al. 2014. Impact of *Scedosporium apiospermum* complex seroprevalence in
1132 patients with cystic fibrosis. Journal of Cystic Fibrosis. 13(6):667–673.
1133 doi:10.1016/j.jcf.2014.01.011. http://dx.doi.org/10.1016/j.jcf.2014.01.011.

1134 Pedersen BS, Quinlan AR. 2018. Mosdepth: quick coverage calculation for genomes and exomes.
1135 Bioinformatics. 34(5):867–868. doi:10.1093/bioinformatics/btx699.
1136 http://dx.doi.org/10.1093/bioinformatics/btx699.

1137 Pihet M, Carrere J, Cimon B, Chabasse D, Delhaes L, Symoens F, Bouchara J-P, Pihet M, Carrere J,
1138 Cimon B, et al. 2009. Occurrence and relevance of filamentous fungi in respiratory secretions of

patients with cystic fibrosis – a review. Medical Mycology. 47(4):387–397. doi:10.1080/13693780802609604. http://dx.doi.org/10.1080/13693780802609604.

Pohl Carolina H., Noverr Mairi C. Competition for Iron during Polymicrobial Infections May Increase Antifungal Drug Susceptibility—How Will It Impact Treatment Options? Infect Immun. 0(0):e00057–22. doi:10.1128/iai.00057-22. https://doi.org/10.1128/iai.00057-22.

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER web server: 2018 update. Nucleic Acids Res. 46(W1):W200–W204. doi:10.1093/nar/gky448. http://dx.doi.org/10.1093/nar/gky448.

Qu X, Yu B, Liu J, Zhang X, Li G, Zhang D, Li L, Wang X, Wang L, Chen J, et al. 2014. MADS-box transcription factor SsMADS is involved in regulating growth and virulence in *Sclerotinia sclerotiorum*. Int J Mol Sci. 15(5):8049–8062. doi:10.3390/ijms15058049. http://dx.doi.org/10.3390/ijms15058049.

Rath P-M, Muller K-D, Dermoumi H, Ansorg R. 1997. A comparison of methods of phenotypic and genotypic fingerprinting of *Exophiala dermatitidis* isolated from sputum samples of patients with cystic fibrosis. Journal of Medical Microbiology. 46(9):757–762. doi:10.1099/00222615-46-9-757. http://dx.doi.org/10.1099/00222615-46-9-757.

Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res. 46(D1):D624–D632. doi:10.1093/nar/gkx1134. http://dx.doi.org/10.1093/nar/gkx1134.

Reiss NR, Mok WY. 1979. *Wangiella dermatitidis* isolated from bats in Manaus Brazil. Sabouraudia. 17(3):213–218. https://www.ncbi.nlm.nih.gov/pubmed/531715.

1161 Revankar SG, Patterson JE, Sutton DA, Pullen R, Rinaldi MG. 2002. Disseminated phaeohyphomycosis:
1162 review of an emerging mycosis. Clin Infect Dis. 34(4):467–476. doi:10.1086/338636.
1163 http://dx.doi.org/10.1086/338636.

1164 Revankar SG, Sutton DA. 2010. Melanized fungi in human disease. Clin Microbiol Rev. 23(4):884–928.
1165 doi:10.1128/CMR.00019-10. http://dx.doi.org/10.1128/CMR.00019-10.

1166 Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N,
1167 Chou JL. 1989. Identification of the cystic fibrosis gene: cloning and characterization of
1168 complementary DNA. Science. 245(4922):1066–1073. doi:10.1126/science.2475911.
1169 http://dx.doi.org/10.1126/science.2475911.

1170 Robertson KL, Mostaghim A, Cuomo CA, Soto CM, Lebedev N, Bailey RF, Wang Z. 2012. Adaptation
1171 of the black yeast *Wangiella dermatitidis* to ionizing radiation: molecular and cellular
1172 mechanisms. PLoS One. 7(11):e48674. doi:10.1371/journal.pone.0048674.
1173 http://dx.doi.org/10.1371/journal.pone.0048674.

1174 Ross BS, Lofgren LA, Ashare A, Stajich JE, Cramer RA. 2021. *Aspergillus fumigatus* In-Host HOG
1175 Pathway Mutation for Cystic Fibrosis Lung Microenvironment Persistence. MBio.
1176 12(4):e0215321. doi:10.1128/mBio.02153-21. http://dx.doi.org/10.1128/mBio.02153-21.

1177 Sass G, Ansari SR, Dietl A-M, Déziel E, Haas H, Stevens DA. 2019. Intermicrobial interaction:
1178 *Aspergillus fumigatus* siderophores protect against competition by *Pseudomonas aeruginosa*.
1179 PLoS One. 14(5):e0216085. doi:10.1371/journal.pone.0216085.
1180 http://dx.doi.org/10.1371/journal.pone.0216085.

1181 Saunte DM, Tarazooie B, Arendrup MC, de Hoog GS. 2012. Black yeast-like fungi in skin and nail: it
1182 probably matters. Mycoses. 55(2):161–167. doi:10.1111/j.1439-0507.2011.02055.x.
1183 https://onlinelibrary.wiley.com/doi/10.1111/j.1439-0507.2011.02055.x.

1184 Schrettl M, Bignell E, Kragl C, Sabiha Y, Loss O, Eisendle M, Wallner A, Arst HN, Haynes K, Haas H.
1185 2007. Distinct Roles for Intra- and Extracellular Siderophores during *Aspergillus fumigatus*
1186 Infection. PLoS Pathogens. 3(9):e128. doi:10.1371/journal.ppat.0030128.
1187 http://dx.doi.org/10.1371/journal.ppat.0030128.

1188 Schultzhaus Z, Romsdahl J, Chen A, Tschirhart T, Kim S, Leary D, Wang Z. 2020. The response of the
1189 melanized yeast *Exophiala dermatitidis* to gamma radiation exposure. Environmental
1190 Microbiology. 22(4):1310–1326. doi:10.1111/1462-2920.14936. http://dx.doi.org/10.1111/1462-
1191 2920.14936.

1192 Selmecki A, Forche A, Berman J. 2010. Genomic plasticity of the human fungal pathogen *Candida*
1193 *albicans*. Eukaryot Cell. 9(7):991–1008. doi:10.1128/EC.00060-10.
1194 http://dx.doi.org/10.1128/EC.00060-10.

1195 Seyedmousavi S, Netea MG, Mouton JW, Melchers WJG, Verweij PE, de Hoog GS. 2014. Black yeasts
1196 and their filamentous relatives: principles of pathogenesis and host defense. Clin Microbiol Rev.
1197 27(3):527–542. doi:10.1128/CMR.00093-13. https://journals.asm.org/doi/10.1128/CMR.00093-
1198 13.

1199 Shore P, Sharrocks AD. 1995. The MADS-box family of transcription factors. Eur J Biochem. 229(1):1–
1200 13. doi:10.1111/j.1432-1033.1995.0001l.x. https://doi.org/10.1111/j.1432-
1201 1033.1995.tb20430.x?locatt=mode:legacy.

1202 Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison.
1203 BMC Bioinformatics. 6:31. doi:10.1186/1471-2105-6-31. http://dx.doi.org/10.1186/1471-2105-6-
1204 31.

1205 Smit A. F. A. 2004. Repeat-Masker Open-3.0. http://www.repeatmasker.org.
1206 https://ci.nii.ac.jp/naid/10029514778/.

1207 Smit AFA, Hubley R. 2008. RepeatModeler Open-1.0. https://ci.nii.ac.jp/naid/10029514778/.

1208 Smith DFQ, Casadevall A. 2019. The Role of Melanin in Fungal Pathogenesis for Animal Hosts. Curr
1209 Top Microbiol Immunol. 422:1–30. doi:10.1007/82_2019_173.
1210 http://dx.doi.org/10.1007/82_2019_173.

1211 Sofi MY, Shafi A, Masoodi KZ. 2022. NCBI BLAST. Bioinformatics for Everyone.:95–102.
1212 doi:10.1016/b978-0-323-91128-3.00021-5. http://dx.doi.org/10.1016/b978-0-323-91128-3.00021-
1213 5.

1214 Song Y, de Sande WWJL, Moreno LF, van den Ende BG, Li R, de Hoog S. 2017. Comparative Ecology
1215 of Capsular Exophiala Species Causing Disseminated Infection in Humans. Frontiers in
1216 Microbiology. 8. doi:10.3389/fmicb.2017.02514. http://dx.doi.org/10.3389/fmicb.2017.02514.

1217 Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: ab initio
1218 prediction of alternative transcripts. Nucleic Acids Res. 34(Web Server issue):W435–9.
1219 doi:10.1093/nar/gkl200. http://dx.doi.org/10.1093/nar/gkl200.

1220 Steinkamp G, Wiedemann B, Rietschel E, Krahl A, Gielen J, Bärmeier H, Ratjen F. 2005. Prospective
1221 evaluation of emerging bacteria in cystic fibrosis. Journal of Cystic Fibrosis. 4(1):41–48.
1222 doi:10.1016/j.jcf.2004.10.002. http://dx.doi.org/10.1016/j.jcf.2004.10.002.

1223 Sudfeld CR, Dasenbrook EC, Merz WG, Carroll KC, Boyle MP. 2010. Prevalence and risk factors for
1224 recovery of filamentous fungi in individuals with cystic fibrosis. Journal of Cystic Fibrosis.
1225 9(2):110–116. doi:10.1016/j.jcf.2009.11.010. http://dx.doi.org/10.1016/j.jcf.2009.11.010.

1226 Sudhadham M, Gerrits van den Ende AHG, Sihanonth P, Sivichai S, Chaiyarat R, Menken SBJ, van
1227 Belkum A, de Hoog GS. 2011. Elucidation of distribution patterns and possible infection routes

1228 of the neurotropic black yeast *Exophiala dermatitidis* using AFLP. Fungal Biol. 115(10):1051–
1229 1065. doi:10.1016/j.funbio.2010.07.004. http://dx.doi.org/10.1016/j.funbio.2010.07.004.

1230 Sudhadham M, Prakitsin S, Sivichai S, Chaiyarat R, Dorrestein GM, Menken SBJ, de Hoog GS. 2008.
1231 The neurotropic black yeast *Exophiala dermatitidi*s has a possible origin in the tropical rainforest.
1232 Studies in Mycology. 61:145–155. doi:10.3114/sim.2008.61.15.
1233 http://dx.doi.org/10.3114/sim.2008.61.15.

1234 Teixeira MM, Moreno LF, Stielow BJ, Muszewska A, Hainaut M, Gonzaga L, Abouelleil A, Patané JSL,
1235 Priest M, Souza R, et al. 2017. Exploring the genomic diversity of black yeasts and relatives
1236 (Chaetothyriales, Ascomycota). Stud Mycol. 86:1–28. doi:10.1016/j.simyco.2017.01.001.
1237 http://dx.doi.org/10.1016/j.simyco.2017.01.001.

1238 Thomas SR, Ray A, Hodson ME, Pitt TL. 2000. Increased sputum amino acid concentrations and
1239 auxotrophy of *Pseudomonas aeruginosa* in severe cystic fibrosis lung disease. Thorax.
1240 55(9):795–797. doi:10.1136/thorax.55.9.795. http://dx.doi.org/10.1136/thorax.55.9.795.

1241 Tunney MM, Field TR, Moriarty TF, Patrick S, Doering G, Muhlebach MS, Wolfgang MC, Boucher R,
1242 Gilpin DF, McDowell A, et al. 2008. Detection of Anaerobic Bacteria in High Numbers in
1243 Sputum from Patients with Cystic Fibrosis. American Journal of Respiratory and Critical Care
1244 Medicine. 177(9):995–1001. doi:10.1164/rccm.200708-1151oc.
1245 http://dx.doi.org/10.1164/rccm.200708-1151oc.

1246 Turcios NL. 2020. Cystic Fibrosis Lung Disease: An Overview. Respiratory Care. 65(2):233–251.
1247 doi:10.4187/respcare.06697. http://dx.doi.org/10.4187/respcare.06697.

1248 Tyrrell J, Callaghan M. 2016. Iron acquisition in the cystic fibrosis lung and potential for novel
1249 therapeutic strategies. Microbiology. 162(2):191–205. doi:10.1099/mic.0.000220.
1250 http://dx.doi.org/10.1099/mic.0.000220.

1251 Uijthof JMJ, Hoog GS, Cock AWA, Takeo K, Nishimura K. 2009. Pathogenicity of strains of the black
1252 yeast *Exophiala (Wangiella) dermatitidis*: an evaluation based on polymerase chain reaction.
1253 Mycoses. 37(7-8):235–242. doi:10.1111/j.1439-0507.1994.tb00419.x.
1254 http://dx.doi.org/10.1111/j.1439-0507.1994.tb00419.x.

1255 Warren AE, Boulianne-Larsen CM, Chandler CB, Chiotti K, Kroll E, Miller SR, Taddei F, Sermet-
1256 Gaudelus I, Ferroni A, McInnerney K, et al. 2011. Genotypic and phenotypic variation in
1257 *Pseudomonas aeruginosa* reveals signatures of secondary infection and mutator activity in certain
1258 cystic fibrosis patients with chronic lung infections. Infect Immun. 79(12):4802–4818.
1259 doi:10.1128/IAI.05282-11. http://dx.doi.org/10.1128/IAI.05282-11.

1260 Weisman CM, Murray AW, Eddy SR. 2022. Mixing genome annotation methods in a comparative
1261 analysis inflates the apparent number of lineage-specific genes. Curr Biol. 32(12):2632–2639.e2.
1262 doi:10.1016/j.cub.2022.04.085. http://dx.doi.org/10.1016/j.cub.2022.04.085.

1263 Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford
1264 Nanopore sequencing. Genome Biol. 20(1):129. doi:10.1186/s13059-019-1727-y.
1265 http://dx.doi.org/10.1186/s13059-019-1727-y.

1266 Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer.
1267 https://play.google.com/store/books/details?id=XgFkDAAAQBAJ.

1268 Xiong D, Wang Y, Tian L, Tian C. 2016. MADS-Box Transcription Factor VdMcm1 Regulates
1269 Conidiation, Microsclerotia Formation, Pathogenicity, and Secondary Metabolism of Verticillium
1270 dahliae. Front Microbiol. 7:1192. doi:10.3389/fmicb.2016.01192.
1271 http://dx.doi.org/10.3389/fmicb.2016.01192.

1272 Yan K, Yin H, Wang J, Cai Y. 2022. Subtle relationships between *Pseudomonas aeruginosa* and fungi in
1273 patients with cystic fibrosis. Acta Clinica Belgica. 77(2):425–435.
1274 doi:10.1080/17843286.2020.1852850. http://dx.doi.org/10.1080/17843286.2020.1852850.

1275 Zajc J, Gostinčar C, Černoša A, Gunde-Cimerman N. 2019. Stress-Tolerant Yeasts: Opportunistic
1276 Pathogenicity Versus Biocontrol Potential. Genes. 10(1):42. doi:10.3390/genes10010042.
1277 http://dx.doi.org/10.3390/genes10010042.

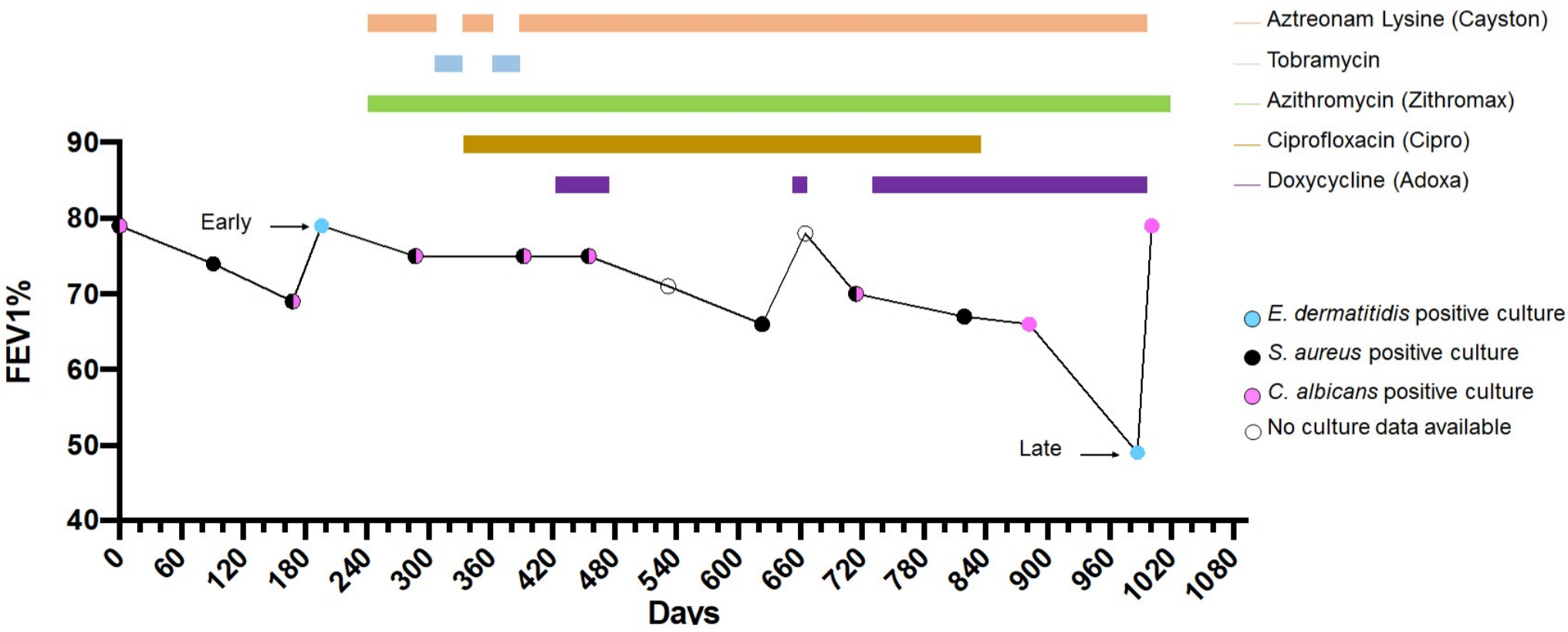
1278 Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. 2018. dbCAN2: a meta
1279 server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 46(W1):W95–
1280 W101. doi:10.1093/nar/gky418. http://dx.doi.org/10.1093/nar/gky418.

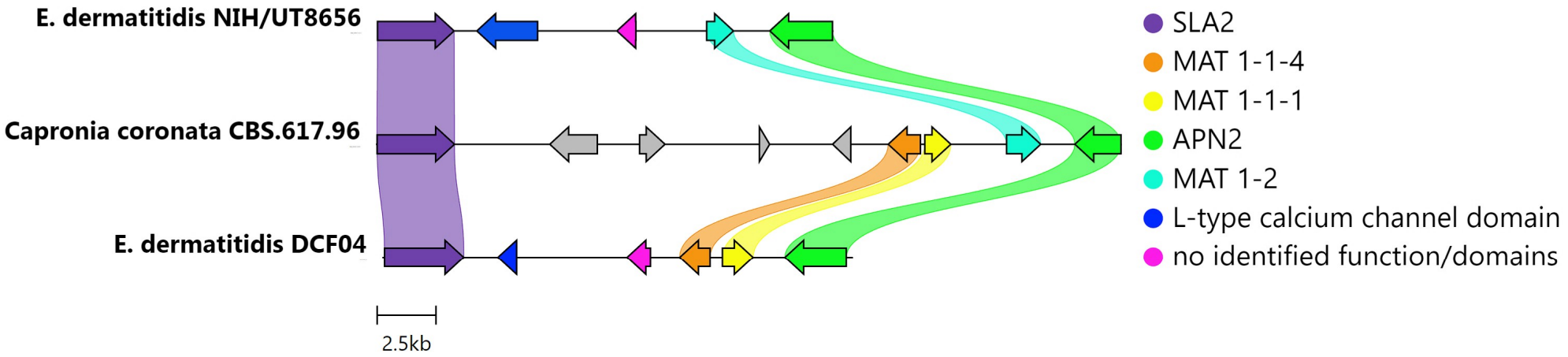
1281 Zhao J, Schloss PD, Kalikin LM, Carmody LA, Foster BK, Petrosino JF, Cavalcoli JD, VanDevanter DR,
1282 Murray S, Li JZ, et al. 2012. Decade-long bacterial community dynamics in cystic fibrosis
1283 airways. Proceedings of the National Academy of Sciences. 109(15):5809–5814.
1284 doi:10.1073/pnas.1120577109. http://dx.doi.org/10.1073/pnas.1120577109.

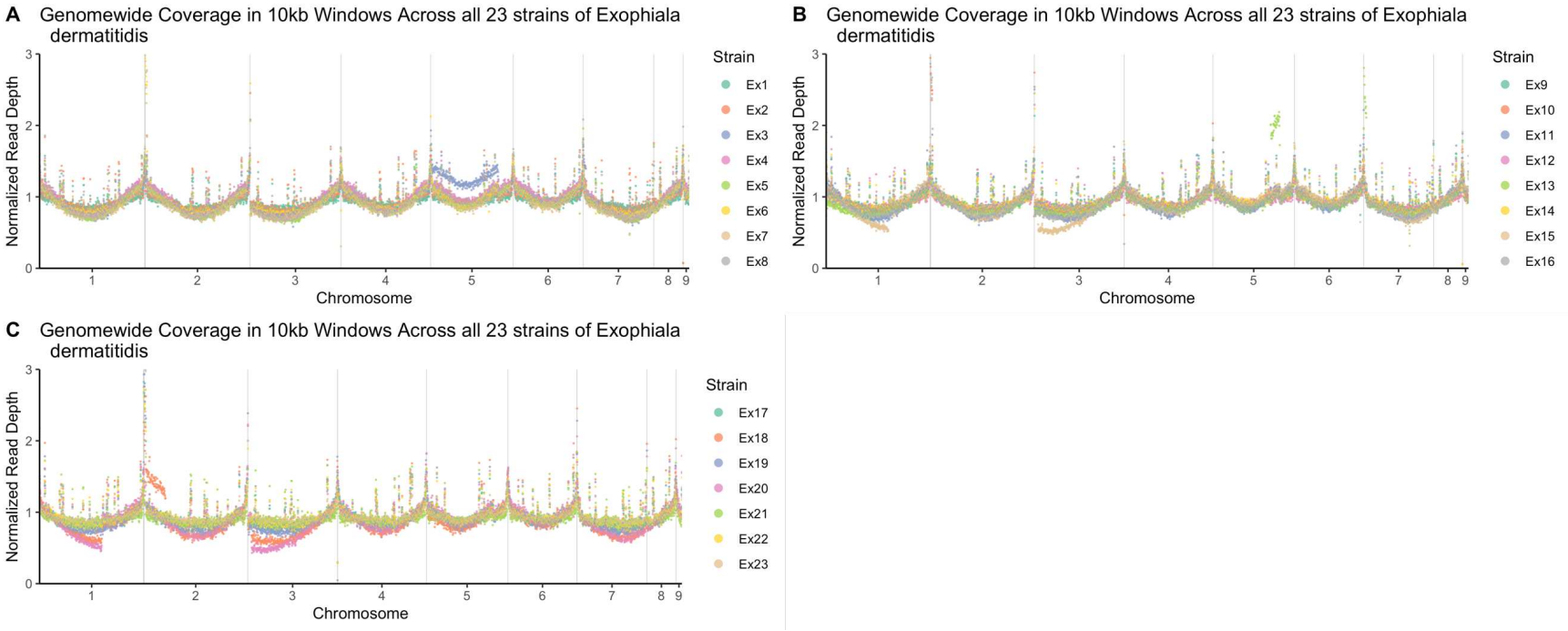
1285 Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome
1286 assembler. Bioinformatics. 29(21):2669–2677. doi:10.1093/bioinformatics/btt476.
1287 http://dx.doi.org/10.1093/bioinformatics/btt476.

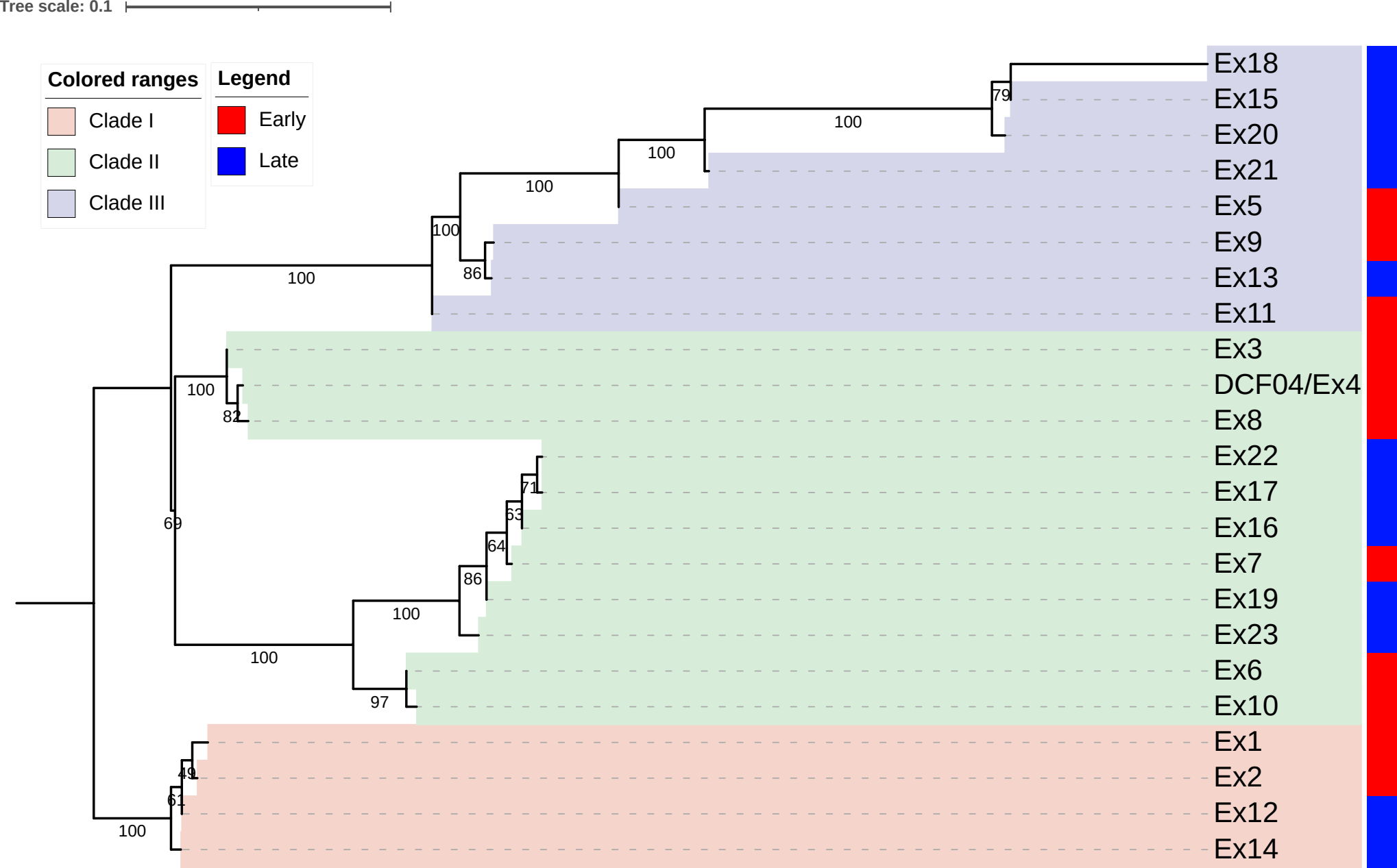
1288 Zupančič J, Babič MN, Zalar P, Gunde-Cimerman N. 2016. The Black Yeast *Exophiala dermatitidis* and
1289 Other Selected Opportunistic Human Fungal Pathogens Spread from Dishwashers to Kitchens.
1290 PLOS ONE. 11(2):e0148166. doi:10.1371/journal.pone.0148166.
1291 http://dx.doi.org/10.1371/journal.pone.0148166.

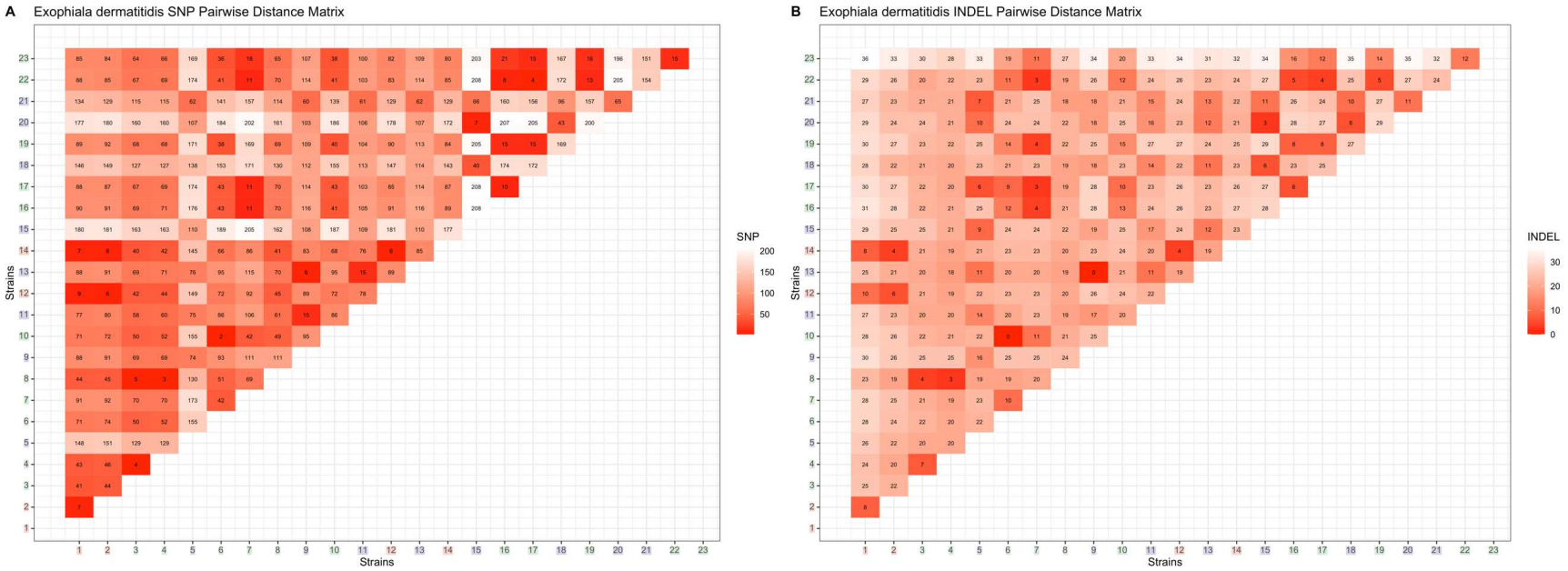
1292

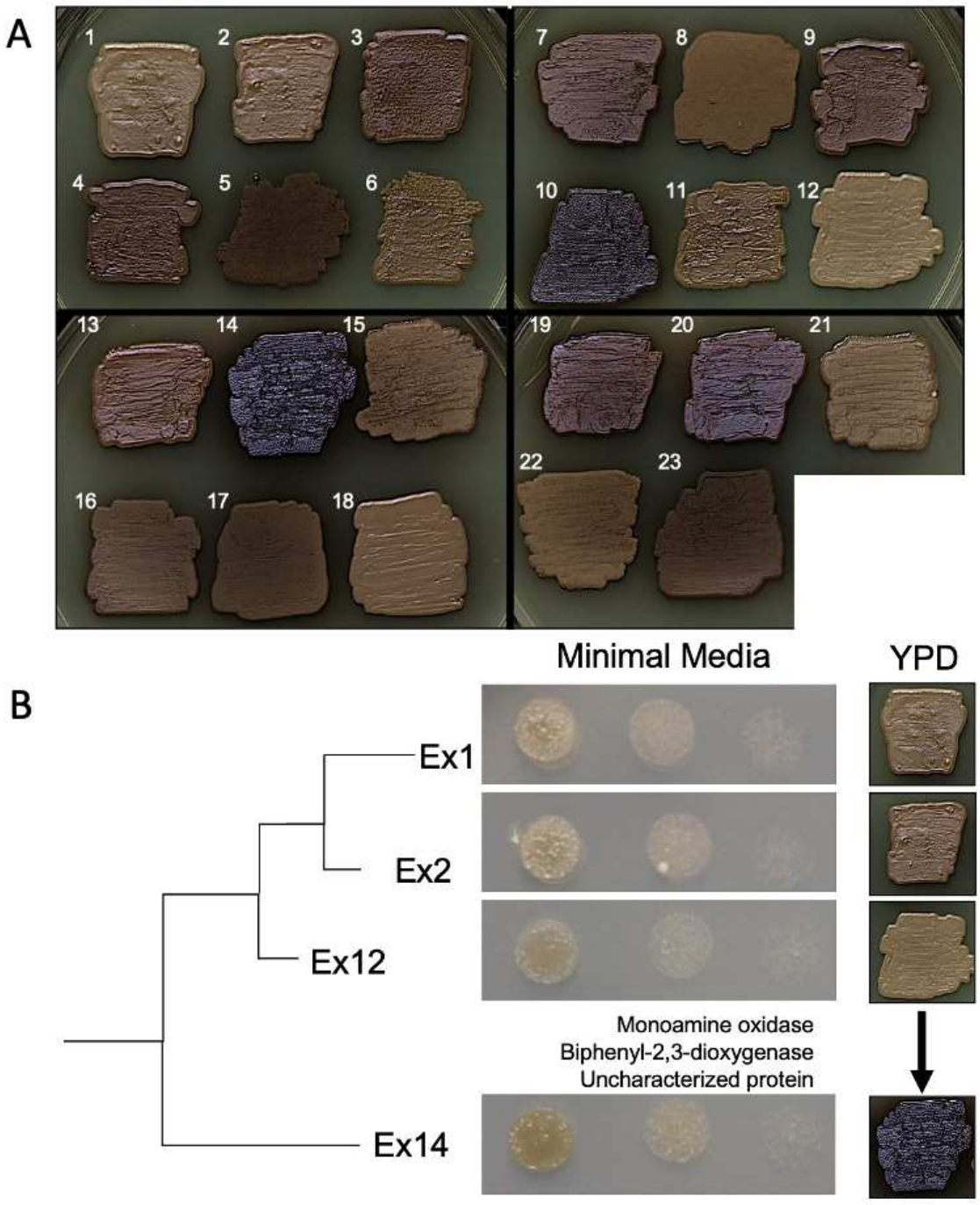

Aztreonam Lysine (Cayston)
Tobramycin
Azithromycin (Zithromax)
Ciprofloxacin (Cipro)
Doxycycline (Adoxa)
E. dermatitidis positive culture
S. aureus positive culture
C. albicans positive culture
No culture data available
Early
Late
FEV1%
90
80
70
60
50
40
0
60
120
180
240
300
360
420
480
540
600
660
720
780
840
900
960
1020
1080
Days

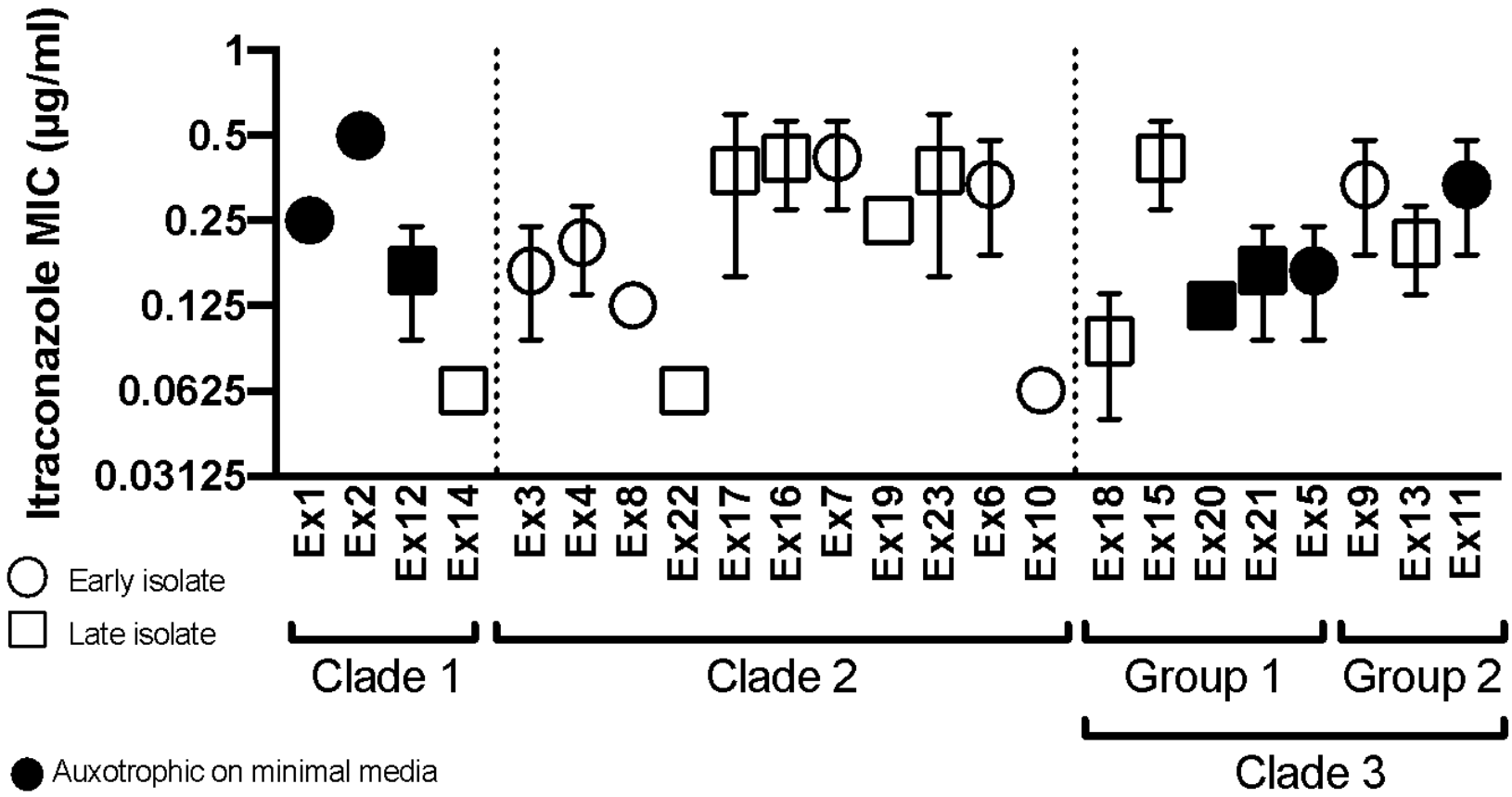

E. dermatitidis NIH/UT8656
Capronia coronata CBS.617.96
E. dermatitidis DCF04
2.5kb
SLA2
MAT 1-1-4
MAT 1-1-1
APN2
MAT 1-2
L-type calcium channel domain
no identified function/domains

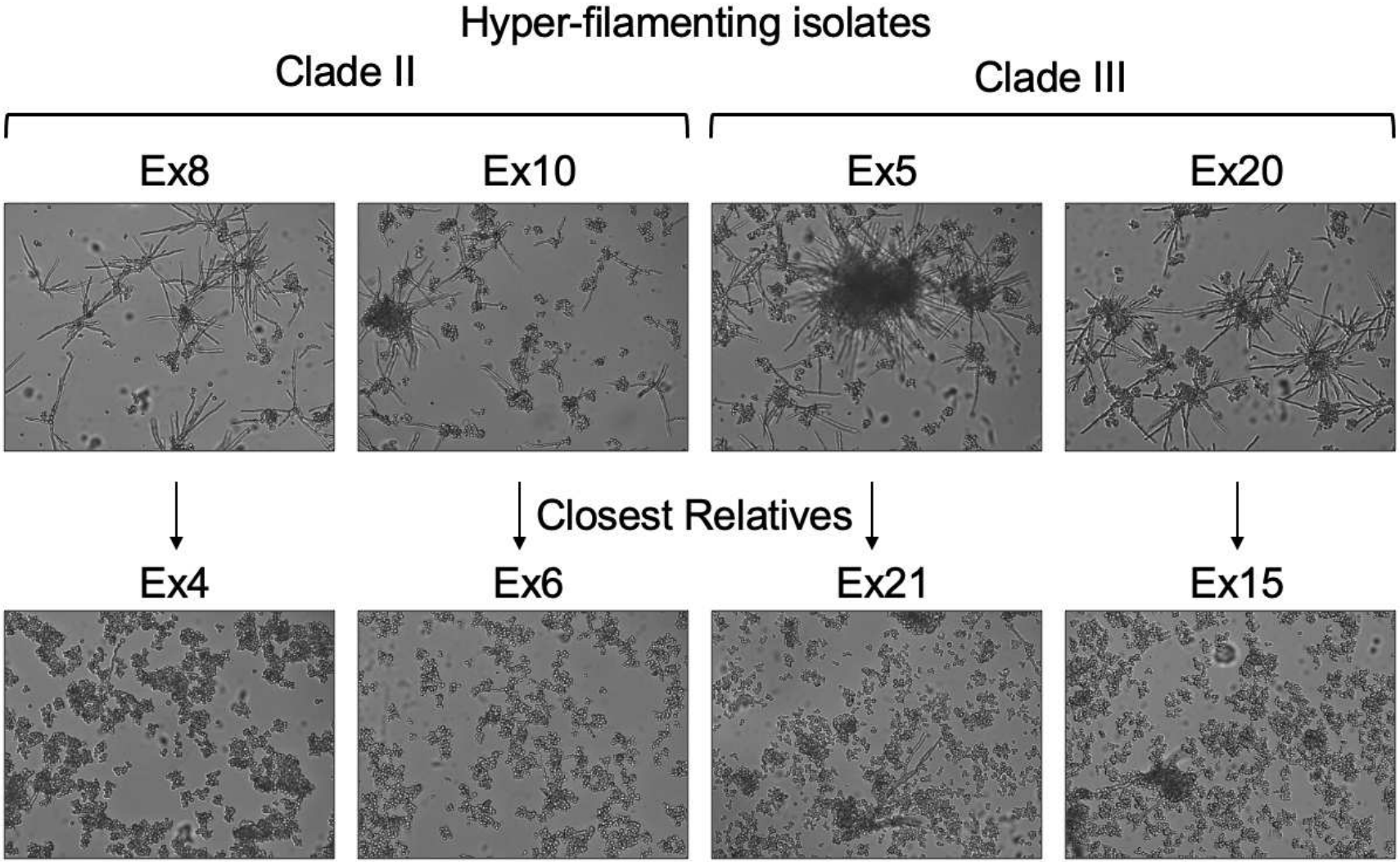

A
Genomewide Coverage in 10kb Windows Across all 23 strains of Exophiala dermatitidis
Normalized Read Depth
Chromosome
Strain
Ex1
Ex2
Ex3
Ex4
Ex5
Ex6
Ex7
Ex8
B
Genomewide Coverage in 10kb Windows Across all 23 strains of Exophiala dermatitidis
Normalized Read Depth
Chromosome
Strain
Ex9
Ex10
Ex11
Ex12
Ex13
Ex14
Ex15
Ex16
C
Genomewide Coverage in 10kb Windows Across all 23 strains of Exophiala dermatitidis
Normalized Read Depth
Chromosome
Strain
Ex17
Ex18
Ex19
Ex20
Ex21
Ex22
Ex23


Tree scale: 0.1
Colored ranges
Clade I
Clade II
Clade III
Legend
Early
Late
Ex18
Ex15
Ex20
Ex21
Ex5
Ex9
Ex13
Ex11
Ex3
DCF04/Ex4
Ex8
Ex22
Ex17
Ex16
Ex7
Ex19
Ex23
Ex6
Ex10
Ex1
Ex2
Ex12
Ex14
79
100
100
100
100
86
100
100
82
69
71
63
64
86
100
100
97
49
61
100


A
Exophiala dermatitidis SNP Pairwise Distance Matrix
Strains
Strains
SNP
200
150
100
50
B
Exophiala dermatitidis INDEL Pairwise Distance Matrix
Strains
Strains
INDEL
30
20
10
0




Itraconazole MIC (µg/ml)
1
0.5
0.25
0.125
0.0625
0.03125
Ex1
Ex2
Ex12
Ex14
Ex3
Ex4
Ex8
Ex22
Ex17
Ex16
Ex7
Ex19
Ex23
Ex6
Ex10
Ex18
Ex15
Ex20
Ex21
Ex5
Ex9
Ex13
Ex11
Early isolate
Late isolate
Clade 1
Clade 2
Group 1
Group 2
Clade 3
Auxotrophic on minimal media


Hyper-filamenting isolates
Clade II
Clade III
Ex8
Ex10
Ex5
Ex20
Closest Relatives
Ex4
Ex6
Ex21
Ex15