

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

Tania Kurbessoian,<sup>1</sup> Daniel Murante,<sup>2</sup> Alex Crocker,<sup>2</sup> Deborah A. Hogan,<sup>2,\*</sup> Jason E. Stajich <sup>1,\*</sup>

<sup>1</sup>Department of Microbiology and Plant Pathology and Institute of Integrative Genome Biology, University of California Riverside, Riverside, CA 92521, USA

<sup>2</sup>Department of Microbiology and Immunology, Geisel School of Medicine, Dartmouth College, Hanover, NH 03755, USA

\*Corresponding author: Department of Microbiology and Plant Pathology and Institute of Integrative Genome Biology, University of California Riverside, 900 University Avenue, Riverside, CA 92521, USA. Email: [jason.stajich@ucr.edu](mailto:jason.stajich@ucr.edu); \*Corresponding author: Department of Microbiology and Immunology, Geisel School of Medicine, Dartmouth College, Hanover, NH 03755, USA. Email: [Deborah.A.Hogan@dartmouth.edu](mailto:Deborah.A.Hogan@dartmouth.edu)

## 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 2 years apart. One isolate genome was sequenced using long-read Nanopore technology as an in-population reference to use in comparative single nucleotide polymorphism and insertion–deletion variant analyses of 23 isolates. We then used population genomics and phylo-genomics to compare the isolates to each other as well as the reference genome 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, and 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.

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

## Introduction

Cystic fibrosis (CF) is an autosomal recessive disorder caused by mutations in the cystic fibrosis transmembrane regulator (CFTR) gene that impairs 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; Horr  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) lusitanae* (Kusenbach et al. 1992; Cimon et al. 2000; Defontaine et al.

2002; Horr  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*, which has been previously named *Hormiscium dermatitidis* and *Wangiella dermatitidis*, is taxonomically classified in Phylum Ascomycota, Order Chaetothyriales, Family Herpotrichiellaceae. To date, about 40 species in the *Exophiala* genus have been identified, 17 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 (Matos et al. 2002; Revankar 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 isolate of *E. dermatitidis* from a CF patient was obtained in 1990 by culturing from sputum

Received: September 26, 2022. Accepted: May 18, 2023

  The Author(s) 2023. Published by Oxford University Press on behalf of The Genetics Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

(Haase et al. 1990; Kusenbach et al. 1992). Multiple documented cases of *E. dermatitidis* in CF patients have followed (Rath et al. 1997; Diemert et al. 2001; Horré et al. 2004; Griffard et al. 2010; Packeu et al. 2012), some of which have developed into late-stage mycosis (Sudfeld et al. 2010; Kondori et al. 2011; Song et al. 2017; Grenouillet et al. 2018), with some infections benefiting from treatment (Packeu et al. 2012; Mukai et al. 2014) while some disease can lead to death (Klasinc et al. 2019).

*Exophiala* species are morphologically grouped in the so-called “black yeasts”, which are not a taxonomic group but are classified by three defining features. Black yeasts produce melanin through 1–8 dihydroxynaphthalene (DHN) biosynthesis pathway, exhibit morphological plasticity or meristematic growth (yeast cells, hyphae, or even pseudohyphae), and have membrane-associated carotenoids and 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 of melanin 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 fungal success in growth in mammalian hosts and their ability to cause disease in susceptible hosts. To better study the molecular biology and genetics of this species, the *E. dermatitidis* strain NIH/UT8656 (ATCC 34100, CBS 525.76) genome was sequenced and assembled into 11 scaffolds presumed to correspond to chromosomes (Robertson et al. 2012; Chen et al. 2014; Schultzhaus et al. 2020; Malo et al. 2021). This genomic resource has enabled evolutionary and comparative genomics studies of the species, and identification of genes that 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). *Exophiala dermatitidis* are isolated from many man-made substrates such as humidifiers (Nishimura and Miyaji 1982), saunas (Matos et al. 2002), and dishwashers (Zupančič et al. 2016; Babič et al. 2018). 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. *Exophiala dermatitidis* and related species are also found in broad environmental niches including wasp nests (Conti-Díaz et al. 1977), healthy bats (Reiss and Mok 1979), lesions of toads (Frank and Roester 1970), and rotten wood (Dixon et al. 1980), which suggests human patients could acquire infections from environmental exposure (Sudhadham et al. 2011).

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. 2021; 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 2 years. This dataset 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 medium as described previously (Grahil et al. 2018). Individual isolates were obtained and

banked; isolate identifiers are listed in [Supplementary Table 1](#). Isolates were stored as frozen stocks with 25% glycerol and maintained at  $-80^{\circ}\text{C}$  and subcultured on solid YPD. Strains were grown in liquid YPD overnights for 18 h prior to use in experiments.

### DNA extraction and sequencing

*Exophiala* isolates were grown in yeast peptone dextrose media (YPD) for ~24 h in 5-ml roller-drum cultures at  $37^{\circ}\text{C}$ . Cells were spun down for 5 min at 5,000 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 ~20 ng/ $\mu\text{l}$ . DNA extractions were sent to Novogene, (Novogene Corporation Inc., Sacramento, California) for  $2 \times 150$  bp sequencing on an Illumina NovoSeq 6000. DNA from isolate DCF04 was extracted and sequenced on Oxford Nanopore (ONT) platform with library preparation and sequencing following the manufacturer's directions (Oxford Nanopore, Oxford United Kingdom). Flow cells with R9.1.4 chemistry on a MinION followed base-calling with Guppy (v. 3.4.4 + a296cb) run with High Accuracy (HAC) model (Wick et al. 2019).

### Genome assembly and annotation

Genome assemblies were constructed for the 23 *E. dermatitidis* isolates from Illumina sequencing. One isolate, Ex4, was also re-sequenced using ONT technology. All genomes were de novo assembled with AATF 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 = 200,000,000. 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 collinearity of these isolates' genomes. Assembly summary statistics were calculated with the “assess” tool in AATF and genome completeness by BUSCO (v5.4.4) (Manni et al. 2021) with the eurotiomycetes\_odb10 database of 290 marker genes.

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 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* Sequence Read Archive accession SRR12590377 under BioProject PRJNA660011. 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) (Brüna et al. 2020), and GlimmerHMM (v.3.0.4) (Majoros et al. 2004) 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. Nonprotein-coding tRNA genes were predicted by tRNAscan-SE (v.2.0.9) (Lowe and Chan 2016). Ribosomal rRNA genes were predicted using RepeatMasker (v.4.1-1) (Smit 2004).

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.

Copy number variation (CNV) was examined by plotting 10-kbp 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 analysis followed previous work done by Metin et al. (2019) was identified through searching for HMG-box homologs of the MAT genes HMG-box *HMPREF1120\_08862* (MAT1-2) and *HMPREF1120\_05727* (APN2), in this study's 23 *E. dermatitidis* isolate genomes with cblaster (Gilchrist et al. 2021). A homothallic black yeast, *Capronia coronata* CBS 617.96 (AMWN000000000.1) (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 that contained the locus.

Identification of telomeric repeat sequences was performed using FindTelomeres.py script (Hiltunen et al. 2021). Briefly, this searches for chromosomal assembly with a regular expression pattern for telomeric sequences at the 5' and 3' end of each scaffold.

## 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 (INDELs). Illumina paired-end reads were aligned to the isolate DCF04 genome 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) tools. 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 (>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 1,000 quality score threshold, where any individual isolate had a "no call", and where the standard deviation in read depth 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 23 CF patient population isolates with reads aligned to the DCF04 isolate and one using the *E. dermatitidis* strain NIH/UT8656 reference genome. 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 an 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: (Pairwise SNP Count)/(Adjusted Genome Length)/Pair/Year. The time between isolated collections was 22 months. The value used for the "Adjusted Genome Length" was collected from genome assembly statistics of DCF04 (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.

A comparison of the DNA similarity of DCF04 and NIH/UT8656 genomes was visualized with a dot-plot constructed with D-GENIES (Cabanettes and Klopp 2018). The tool was run using minimap2 and default parameters on the tool's website (<https://dgenies.toulouse.inra.fr>).

Ortholog and paralog assignment of genes between the annotated DCF04 and NIH/UT8656 genomes was calculated with Orthofinder (Emms and Kelly 2019). First, genes with alternatively spliced isoforms were reduced to the single longest isoform. The pairwise similarity of proteins was computed by DIAMOND with ultra-sensitive parameters (Buchfink et al. 2015). Assessment of shared prediction protein-coding genes was extracted from the Orthology tables generated by Orthofinder and a custom Python script. Singleton proteins generated from Orthofinder results were assessed from both DCF04 and NIH/UT8656 using the proteins as a query and e-value set to 1e-5 with TBLASTN from the NCBI-BLAST toolkit and miniprot (Li 2023).

## Phylogenetic 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 to accurately model the ascertainment bias of a SNP-only dataset [-m GTR + ASC]. The nucleotide substitution model GTR + F + ASC selected based on Bayesian information criteria runs within IQTree. Statistical support for the tree nodes was evaluated from 1,000 bootstrap replicates using UFBoot ultra-fast bootstrapping approximation

(Hoang *et al.* 2018). The tree was visualized using iTOL (Letunic and Bork 2016).

### Minimum inhibitory concentration assays and spot titers

*Exophiala* isolates were struck from  $-80^{\circ}\text{C}$  glycerol stocks onto yeast extract peptone dextrose (YPD) plates (2% glucose, 2% yeast extract, 1% peptone) and allowed to grow for 48 h at  $37^{\circ}\text{C}$ . Overnight cultures were started from YPD patches inoculated into 5 ml of liquid YPD and grown for  $\sim 24$  h in a roller drum at  $30^{\circ}\text{C}$ . For minimum inhibitory concentration (MIC) assays, cultures were spun down for 5 min at 5,000 RCF and washed thrice in deionized water. Cells were counted on a hemocytometer and added to a final concentration of  $1 \times 10^3$  cells per well in a 96-well flat-bottom dish, then grown at  $37^{\circ}\text{C}$  for 72 h before determining the final MIC through visual inspection for lack of growth. For spot titers on rich and minimal media with and without the supplementation of 2% casamino acids, cells were CFU equilibrated to  $1 \times 10^6$  CFU/ml, and serial dilutions (1:10) were spotted in 5- $\mu\text{l}$  volumes. Plates were incubated for 48 h at  $37^{\circ}\text{C}$  prior to imaging.

### Plate-based melanization assessment

*Exophiala* isolates were initially cultured as described above onto YPD plates and arranged in order of isolate number. Plates were incubated at  $37^{\circ}\text{C}$  for 5 days, to permit for sufficient growth and melanin production. Top-down images of each plate were captured using a Canon EOS Rebel T6i with an 18- to 15-mm lens.

### Microscopy

*Exophiala* isolates were initially cultured onto solid media and liquid overnights as described above. From confluent overnights, isolates were inoculated at 1,000 CFU/ml into RPMI-1640 in glass-bottom 24-well culture dishes and incubated at  $37^{\circ}\text{C}$  in 5% CO for 18 h. Cultures were then imaged at 63X magnification, DIC with a Zeiss Axiovert 200 M inverted microscope.

## Results

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). *Exophiala dermatitidis* isolates were recovered from banked sputum samples, collected 2 years apart. *Staphylococcus aureus* and *Candida albicans* were also identified in clinical cultures from the patient in the intervening years between the two time points (Fig. 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 (11 from the early time point and 12 from the late time point) 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 23 isolates

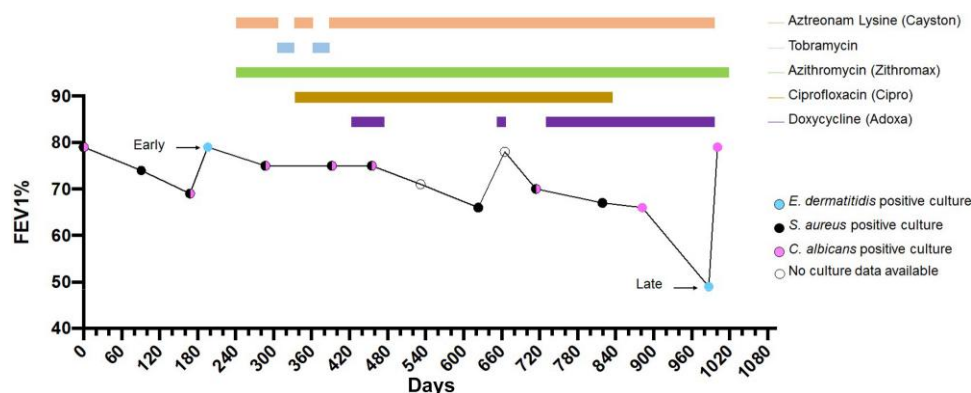
(Supplementary Table 2). The depth of coverage ranged from 11 to 47x coverage across all 23 Illumina sequenced samples. Benchmarking Universal Single-Copy Orthologs (BUSCO) was used to test genome assembly completeness; within our dataset, it ranged 99–99.3% complete, 98.3–99.3% single copies present, 0–1% for duplicates, and 0–0.3% missing BUSCOs. Contig counts spanned 44–1,422 and the average genome assembly size is 26.7 Mbp, N50 of 1,078,277, and L50 of 9.21. Repeat content for all 23 strains ranged from 3.8% to 4.5%, while strain NIH/UT8656 was 3.5%, considerably lower than the CF lung isolates.

To further examine the fine-scale variation within the CF isolates, we sought to generate a within-population high-quality reference genome. DNA sequence reads for isolate DCF04 were generated using ONT sequenced with an average genome depth of coverage of 11x. A hybrid genome assembly was constructed from the combined ONT and Illumina reads (Supplementary Table 2) to produce a 26.6-Mbp assembly. Only 3.19% of the genome was identified as repetitive elements. The candidate telomeric repeat units “TTTAGGG/CCCTAA” were identified as arrays of repeats at both ends of five of the scaffolds and as repeats on only one end of the remaining three scaffolds. (Supplementary Table 3). A total of 9,599 genes were predicted comprising 64 tRNA genes and 9,535 protein-coding genes, of which 426 have at least two alternative isoforms predicted based on RNA-seq evidence. A search for secondary metabolite gene clusters with antiSMASH identified 15 secondary metabolite clusters, of which four are classified as terpenes, seven as NRPS containing, two T1PKS, 1 T3PKS, and one betalactone cluster. RepeatMasker identified rDNA copy numbers vary from 18 to 22 found in 23 *E. dermatitidis* CF isolates and 22 copies in strain NIH/UT8656.

### Comparing assembly quality and gene content of DCF04 to NIH/UT8656

We sought to examine how similar the assembled genome content and gene content is between two fungal isolates cultured from different environments and sequenced with different technologies. The isolate DCF04 was isolated from the lung sputum of a CF patient and sequenced with Illumina and ONT technology and the NIH/UT8656 isolate, likely cultured from a skin infection, and the genome was sequenced with Sanger sequencing technology (BioProject: PRJNA225511, Assembly: GCF\_000230625.1). To assess the differences between the two genome assemblies, we compiled summary statistics (Supplementary 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 11 scaffolds. The DCF04 scaffolds were derived by a comparative assembly against the NIH/UT8656 assembly to achieve the best assembly after first testing for any evidence of rearrangements. The total length of the genome assemblies is nearly the same where DCF04 is 26.6 Mb and NIH/UT8656 is 26.4 Mb. The summary statistics for scaffold L50 and N50 are also nearly identical for DCF04 at 4 and 3.7 Mb and NIH/UT8656 were 4 and 3.6 Mb, respectively. The BUSCO completeness statistics search with ascomycota\_odb10 (1,706 models) identified DCF04 as 99.6% complete and NIH/UT8656 as 99.4% complete. A dot-plot comparing the two genome assemblies revealed minimal rearrangements or discontinuity suggesting high similarity of the two isolated genomes (Supplementary Figure 1).

The gene content was further compared using OrthoFinder (Emms and Kelly 2015; Emms and Kelly 2019) (Supplementary Table 4), which identified 8,256 orthologous groups comprising 17,640 orthologous protein-coding genes between DCF04 and



**Fig. 1.** FEV1 pulmonary function data, antibiotic use, and sputum collection time points. FEV1 (forced expiratory volume in one second), an indication of pulmonary function, was measured over 3 years. Sputum samples 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. Horizontal bars indicate the duration of treatment for each listed antibiotic given during the study period.

NIH/UT8656. A total of 1,180 singleton genes were detected through OrthoFinder and can be found in the Othroggroups\_UnassignedGenes.tsv file. Seven hundred and five singleton genes were identified in the DCF04 strain and 475 in NIH/UT8656 strain. In addition, some gene families were multicopy, but lineage-specific to a strain as found in the Orthologs.tsv file where only a single strain was found. Ten families were multicopy in DCF04 and 5 in NIH/UT8656.

To test the hypothesis that genes classified as lineage-specific were truly unique to a strain or a product of missing annotation in the other isolate, we performed translated searches of the predicted proteins with TBLASTN and miniprot against each of the isolate genome assemblies.

This search found that 683/705 (96.9%) of the putative DCF04 lineage-specific loci had identifiable similarities in the NIH/UT8656 strain genome assembly so that only 22 genes are truly DCF04 lineage-specific genes (Supplementary Table 4). The reciprocal analysis found that 468/475 (98.5%) of the putative NIH/UT8656 lineage-specific loci had identifiable similarities in the CF isolate genomes. Only seven genes were found to be NIH/UT8656 lineage-specific genes and were missing in all 23 isolates of *E. dermatitidis* from this study. Overall, we find the majority of the gene content differences reflect variation in predictions produced by gene annotation pipelines consistent with observations in other systems (Weisman et al. 2022).

### All Ex CF isolates are MAT1-1 mating type

To examine and genotype the mating-type locus in the isolates, we searched for homologs of the ascomycete MAT genes with the tool cblaster and visualized their shared microsynteny with clinker. The genomes of all the CF isolates including DCF04/Ex4 *E. dermatitidis* isolate encoded a MAT1-1 locus (Fig. 2, Supplementary Figure 2, and Supplementary Table 5). Consistent with previous MAT loci described in *E. dermatitidis* (Metin et al. 2019), DCF04 had both MAT1-1-4 and MAT1-1-1 genes (Fig. 2). The NIH/UT8656 encodes a MAT1-2 mating type in contrast to the uniform MAT1-1 type found in all CF clinical isolates.

Examination of genomes of additional *E. dermatitidis* isolates and *Capronia* species identified homologs of SLA2, a SRC-like adapter protein, and APN2, apurinic-apyrimidinic endonuclease 2, flanking the MAT loci in all isolates. Within the MAT locus, a predicted gene with no known function was found between SLA2 and MAT1-1-4, a common feature present across all the

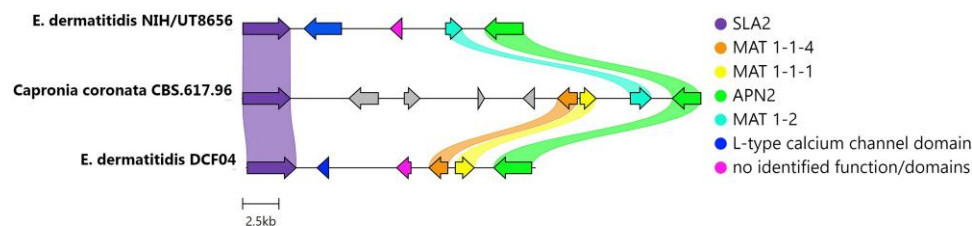
*E. dermatitidis* isolates. The outgroup *Capronia coronata* CBS617.96 is homothallic and its genome contains both MAT1-1 and MAT1-2 genes (Teixeira et al. 2017).

### Chromosome copy number variation across *E. dermatitidis* isolates

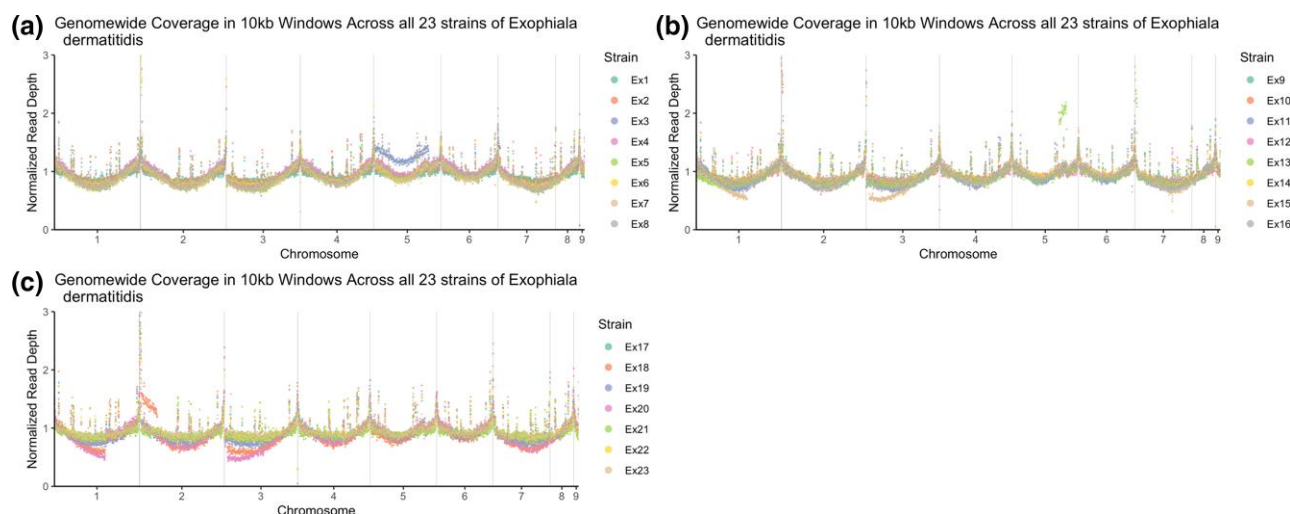
To test for variation in partial or complete chromosome copy numbers, we calculated sequence read coverage depth from each *E. dermatitidis* isolate in 10-kb sliding windows along the nine reference genome chromosomes (Fig. 3). Visual scanning of the plots identified an anomaly of at least 1.5 $\times$  higher coverage on chromosome 5 in Ex3 (Fig. 3a) encompassing about five genes or 11,600 bp (hypothetical protein—HRR75\_002843, hypothetical protein—HRR75\_002844, hypothetical protein—HRR75\_002845, hypothetical protein—HRR75\_002846, and hypothetical protein—HRR75\_002847). A similar but much smaller region (2,600 bp) of chromosome 5 appears to have 2–2.25 $\times$  coverage and may be duplicated in Ex13. Additional partial 1.25–1.5 $\times$  coverage on the left arm of chromosome 2 (about 1,750 bp) in isolate Ex18 is also observed.

Some chromosome regions showed reduced coverage in the analysis. The plots of isolates Ex15, Ex18, and Ex20 indicate a normalized read depth of 0.5 $\times$  on chromosomes 1 and 3 (Fig. 3b and c), a possible sign of a segmental chromosomal aneuploidy event. Although these isolates are haploid, this reduced coverage could indicate a possible chromosome fusion of chromosomes 1 and 3. One interesting feature of these plots is the continuity in coverage patterns on chromosomes 1 and 3, which might indicate they are in fact part of the same in these isolates. Haploid organisms, like *E. 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 Ex3's potential partial aneuploidy on chromosome 5, an isolate from one of the early time points, it may be that this genome copy may contribute to an adaptive mechanism to support colonization and persistence in the human host.

In addition, the genes encoded in the variable copy number region in isolate Ex3 include multiple rDNA genes. Fungi generally have larger numbers of ribosomal genes ranging from 14 to 1,442 copies (Lofgren et al. 2019), while these *E. dermatitidis* isolates have 22–23 rDNA copies and Ex3 has 3 complete rDNA copies in chromosome 5. Other isolates sharing potential aneuploidies (Ex15, Ex18, and Ex20) are from the late time point and also



**Fig. 2.** Mating-type determination of four 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 and APN2. The MAT 1-1 genes, MAT 1-1-4, and MAT 1-1-1 are observed in DCF04 isolate, while the reference strain NIH/UT8656 has a MAT 1-2 gene. In addition, two genes are predicted in the interval between SLA2 and the MAT genes. One is an L-type calcium channel domain, which is only found in *E. dermatitidis*, and a gene with no identified function or domains, which is syntentically adjacent to the MAT genes in all isolates and among all species.



**Fig. 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) Ex1-8, b) Ex9-16, and c) Ex17-23.

have the highest mutation rate among this population as observed from the phylogenetic tree and mutation rate calculation (Figs. 4 and 6).

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

Using the DCF04/Ex4 isolate as a reference for variant identification, we identified 441 SNP variants among the 22 isolates. A phylogenetic tree was generated using these filtered SNPs, rooted with the group containing Ex1, Ex2, Ex12, and Ex14 (Fig. 4). In an additional analysis using the out-population strain NIH/UT8656 as the genome reference, we found 10,344 SNPs, with most of these being fixed differences between the outgroup and the CF isolates. The phylogenetic tree for this dataset was also generated to resolve the relative differences and root the population relationships with the NIH/UT8656 outgroup (Supplementary Figure 3).

Three clades (Clade I, Clade II, and Clade III) were identified based on the trees. Clade I is composed of two early collected (Ex1 and Ex2) and two late collected isolates (Ex12 and Ex14) of *E. dermatitidis*. Clade II is composed of three subgroups, one of which contains only early isolates and a second group with both early isolates and late isolates. A 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 (Fig. 3) showed these three isolates,

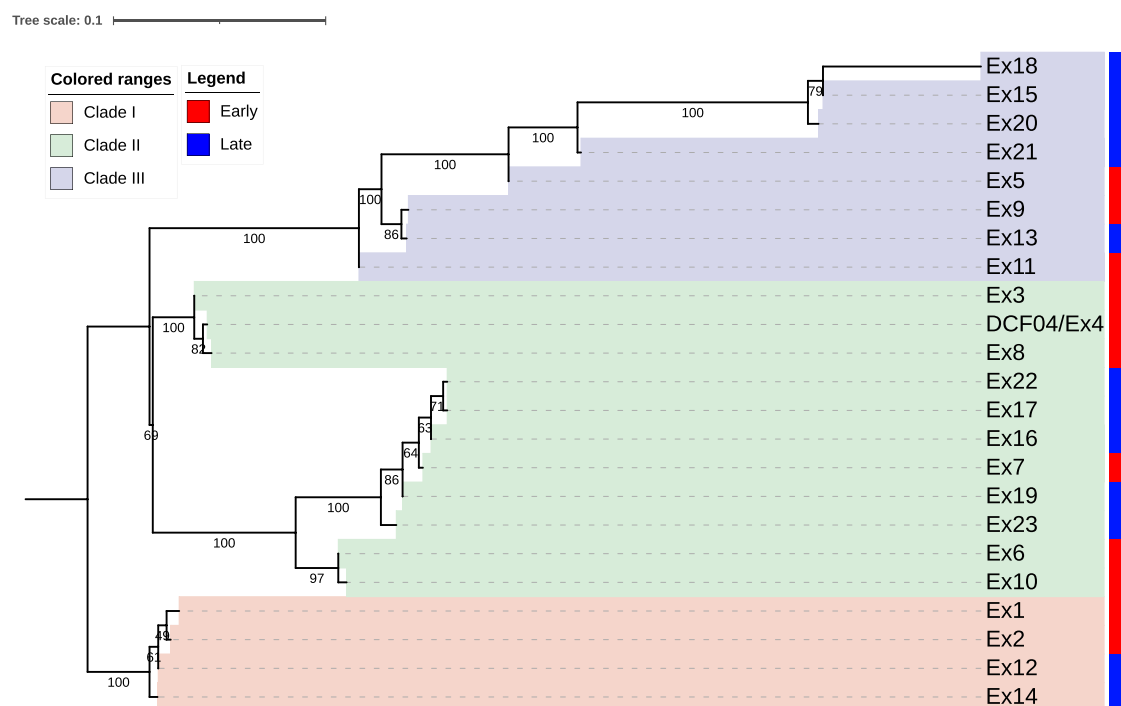
Ex15, Ex20, and Ex18 contained similar CNV differences in chromosomes 1 and 3 when compared to the other isolates.

### Nonsynonymous 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 (Fig. 5a). Our analysis further confirmed consistency in mutation patterns where the number of SNP differences correlated with the number of INDELs detected in pairwise analyses (Fig. 5b). Within Clade I and Clade II, isolates had few overall variants, indicating the evolutionary distance between them was small. The total variant counts within Clade III were much higher indicating a higher divergence within this group of isolates as compared to other groups. Grouping the matrix using results from Fig. 4 provided a way to visualize incongruence, especially amongst SNPs that distinguish Ex7/Ex19 (169), and INDELs between Ex5/Ex17 (6).

### Nonsynonymous SNP differences seen between clades

Examination of predicted variant impacts and genes produced by snpEff combined with functional gene annotation provides an overview of the types of changes (Supplementary Table 6). Global comparisons between early collected Ex1-Ex11 isolates



**Fig. 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. Strains numbered from 1-11 are early collected isolates, while strains numbered from 12-23 are late isolates. Values in tree branches indicate bootstrap values calculated through iqtree2.

and the second time point isolates Ex12–Ex23 documented 441 variants, with a breakdown of 379 nonsynonymous to 62 synonymous mutations. Breaking down these variant numbers for all 23 isolates when compared to the reference DCF04 include 227 intergenic mutations (single point nucleotide changes), 21 frameshift mutations (affecting read count for resulting genes), 10 disruptive in-frame insertion/deletions (leading to disrupted proteins), 104 missense mutations (producing alternative amino acids), and 4 variants that fall within splice regions.

Variants generated through a gain or loss of a stop codon (nonsense mutations) were also considered in this analysis. Nine stop codon gained variants were determined, where six were described with annotated hypothetical genes, and the remaining three had descriptive annotated functions including, XPT1 xanthine phosphoribosyltransferase 1 (HRR76\_002280), SIP3 SNF1-interacting protein (HRR76\_004004), and ARP8 actin-like protein (HRR76\_006707). The stop codon found in the XPT1 and SIP3 variants was only found in species *E. dermatitidis* Ex5, while ARP8 was found in isolates Ex16, Ex17, Ex19, Ex22, Ex23, and Ex7, which are a small branch of Clade II species further indicating these isolates may be all descendants of Ex7. One “loss of a stop codon” was determined in our variant analysis and was annotated as a hypothetical protein. Only about 10 variants resulted in a codon gained or lost in our 23 isolates of *E. dermatitidis*.

Further stratifying the genetic differences among the early and late isolates provided a more in-depth understanding of variants; ~95 nonsynonymous to 20 synonymous early mutations were detected, while among the later isolates, there were 97 nonsynonymous and 13 synonymous segregating mutations shared among isolates as well as singleton mutations. We found that the lung environment did not generate a uniform general shift or a common variant found in all later isolates among a particular shared gene.

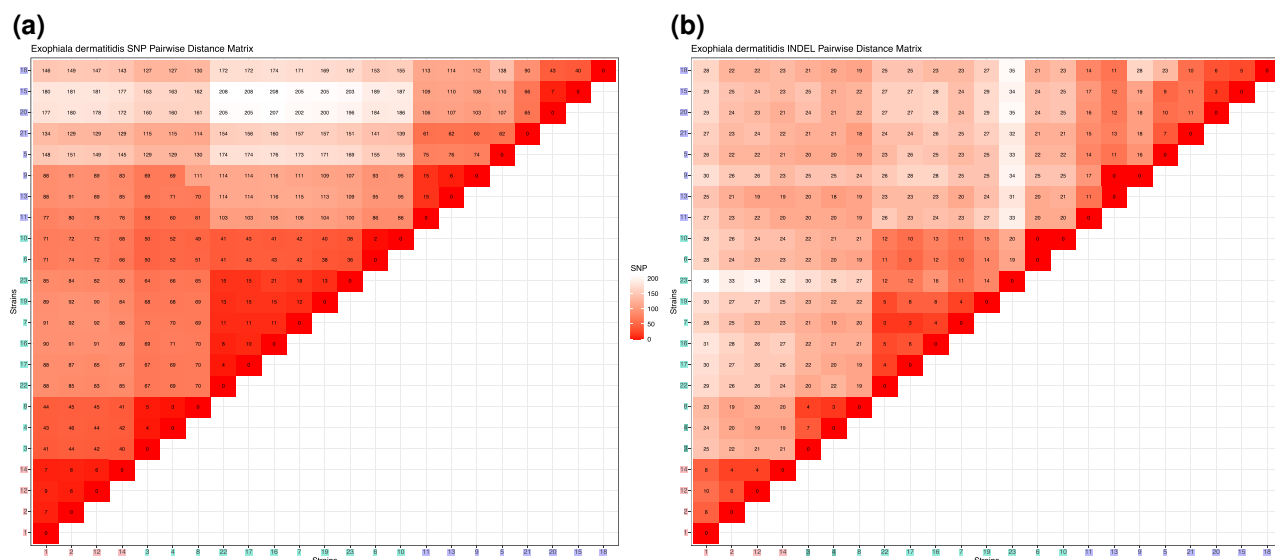
Instead, we observed variants unequally distributed among the Clades with isolates Ex15, Ex18, Ex20, and Ex21 (Clade III) containing 58 variants. Examining the variants found only in early isolates, we counted 60 intergenic mutations, 8 frameshift mutations, 1 disruptive in-frame insertion/deletion, 21 missense mutations, and 4 variants that fall within splice regions. Among late isolates, we found 55 intergenic mutations, 5 frameshift mutations, 5 disruptive in-frame insertions/deletions, and 28 missense mutations as compared to DCF04 (an early isolate). We were expecting an overall larger number of all types of mutations would be seen in the late collected isolates, but that was not the case other than the missense mutation numbers. Otherwise, there seemed to be more variety of variants seen among the early collected isolates.

### Analysis of SNP differences between clades

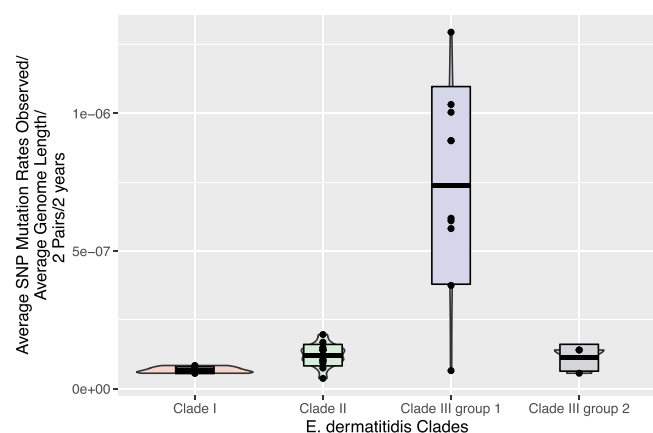
A single nonsynonymous mutation was detected in a gene orthologous to *Saccharomyces cerevisiae* MRS4 (HMPREF1120\_06597), a mitochondrial iron transporter. One allele of the MRS4 allele encoded a protein that was identical in sequence to the allele in the reference strain NIH/UT8656, present in the isolates in Clade I (Ex1, Ex2, Ex12, and Ex14), and a subclade of clade II (Ex3, Ex4/DCF04, and Ex8). The remainder of the isolates had a second allele with a nonsynonymous mutation in the 40th amino acid position, converting a glutamic acid residue to glycine. The functional consequences and differences of these MRS4 alleles are described elsewhere (Murante et al. 2023).

### Testing for enrichment of evolutionary patterns within clades

A pairwise comparison of the synonymous and nonsynonymous SNP differences was performed on 23 pairs of isolates identified as early and late members of the same lineage. A comprehensive



**Fig. 5.** Matrix of *E. dermatitidis* CF isolates SNP and INDEL pairwise dissimilarities. The number of SNPs (a) and (b) INDELs that differs among pairs of isolates. The more similar isolate pairs have lower numbers and the number of variants increases 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.



**Fig. 6.** Pairwise mutation rates observed among clades. Pairwise mutation rates were calculated for all isolates in a clade and within the two subgroups of Clade III. The boxplots show median (center line in boxes) and standard deviation (box top and bottom) mutation rate for all isolate pairs in a group or clade.

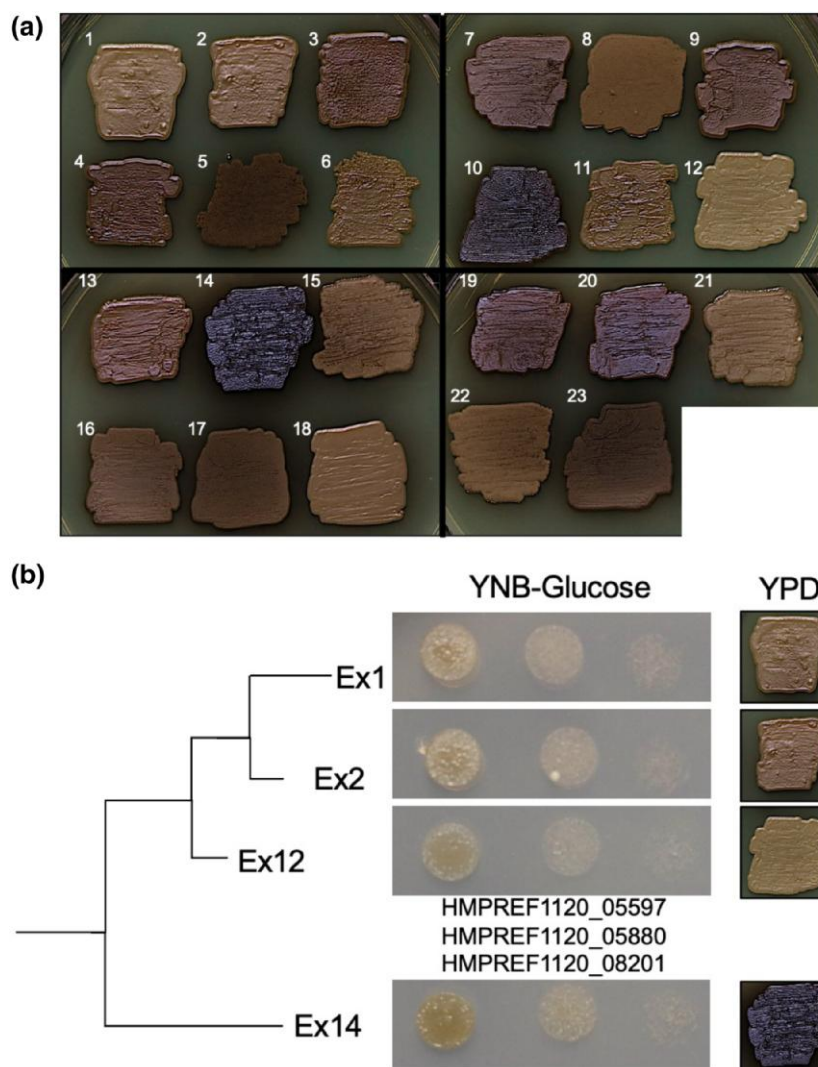
list of gene differences among all pairwise comparisons is in [Supplementary Table 7](#) and includes hypothetical proteins with no identified function. This analysis tested for differences between early and late isolates found in the same clade to focus on changes that may have occurred within the host.

To better understand the pattern of variation that arose through the time points, we focused on within-clades comparison so that each early isolate within a clade was compared to the later isolate(s). Clade I is composed of four isolates: two early (Ex1 and Ex2) and two isolates from the late time point (Ex12 and Ex14). These isolates show very little genetic differentiation (about 6–9 variant SNPs and 6–10 variant INDELs). 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 G4 quadruplex nucleic acid binding proteins in synthetic biology applications. The helical complex is formed through guanine-rich nucleic acid sequences and is found at the telomeric regions of chromosomes. These variants are an interesting observation and warrant extra analysis in the future.

Clade II contains 11 isolates, six from the early time point and five from the later time point, which fall into three different subgroups. Ex3, Ex4/DCF04, and Ex8 fall into one group, Ex22, Ex17, Ex16, Ex7, Ex19, and Ex23 fall into the second group, and Ex6 and Ex10 form a third group. Group One and Group Three both contain only early isolates, while the second group contains the majority of the later isolates. 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 isolate was not sampled. A variant of note, NAD-dependent histone deacetylase SIR2, is involved with chromosomal remodeling specifically with phenotype transcription modification (Freire-Benítez et al. 2016).

Clade III, which contains eight isolates, is 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. Comparison of Clade III group 1 member Ex5 (early) to Ex15, 18, 20, and 21 identified the highest number of variants as compared to the other clades (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 and II (6–15 variant SNPs and 0–17 variant INDELs). Four genes within the second group (Ex9, Ex13, and Ex11) 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



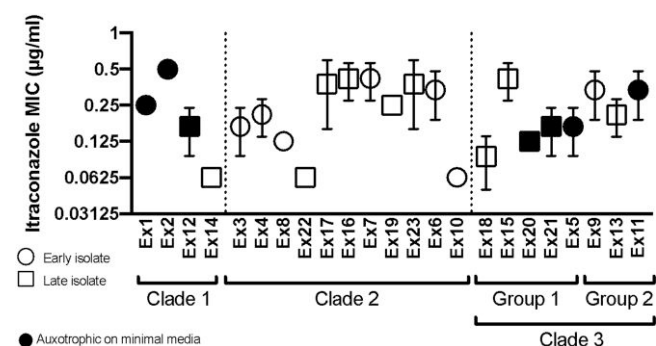
**Fig. 7.** Melanin production phenotypes in the *E. dermatitidis* CF isolates. a) Melanization differences between *E. dermatitidis* CF isolates grown for 120 h on YPD at 37°C. b) Melanization differences between closely isolates in Clade I were observed, but only a small number of SNPs differentiated low melanin producers (Ex1, Ex2, and Ex12) from a high melanin producer (Ex14). The SNPs that distinguish Ex14 from the three low production isolates are in the upstream intergenic regions of three genes: a monoamine oxidase HRR76\_005273; a putative biphenyl-2,3-dioxygenase, HRR76\_005538, and a gene of no known function, HRR6\_007729. The three loci identifiers correspond to HMPREF1120\_05597, HMPREF1120\_05880, and HMPREF1120\_08201 respectively.

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 much more significance than the first group, consisting of 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 12 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* has been found to impact phenotypes and adjusts the melanin synthesis pathway, resistance or susceptibility to antifungals, and extreme environment adaptability (Paolo et al. 2006).

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 did not identify any nonsynonymous variants in this gene across the CF lung isolates. Though, the presence of SNP variants in iron-binding and siderophore transporters indicates there may be selective pressure for *E. dermatitidis* to obtain iron in the CF lung environment, in which iron is strictly regulated. 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 by taking the mean and standard deviation of all the pairwise comparisons with a clade and visualized the results using a violin plot with all datapoints visible (Fig. 6). The mean Clade I mutation rate



**Fig. 8.** Itraconazole sensitivity testing of 23 CF isolates. Average minimum inhibitory concentration (MIC) of CF isolates grown with Itraconazole. Standard deviation of three biological replicates is indicated by error bars. The auxotrophies of some isolates, indicated by filled symbols, was identified by growth on minimal medium (YNB) without casamino acids. Symbol shape indicates whether isolates were recovered from the early or late sputum samples.

based on the four pairwise calculations for all combinations of early and late isolates was  $6.76 \times 10^{-8}$ . The average mutation rate for the five comparisons of early to late isolates in Clade II was  $1.22 \times 10^{-7}$ . The average mutation rate for the six comparisons of early to late isolates in Clade III (Group 1) values was averaged and determined to be  $7.39 \times 10^{-7}$ , while the second group values were averaged and determined to be  $1.13 \times 10^{-7}$ . 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 [Supplementary Table 8](#).

## Heterogeneous phenotypes observed in the *E. dermatitidis* isolates

The 23 *E. dermatitidis* isolates were heterogeneous for traits in several ways including melanization, antifungal sensitivity, and auxotrophy. We noted differences in melanin pigmentation across isolates (Fig. 7a) 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 (Fig. 7b). In a search for genetic determinants responsible for this phenotypic range, we identified three candidate intergenic mutations, which differentiated Ex14 from the other members of the clade.

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 an ~10-fold range (MIC from 0.0625 to 0.5 µg/ml) (Fig. 8). Heterogeneity in amino acid auxotrophy, scored as no growth on a minimal medium that was rescued by supplementation of amino acids (Fig. 8), was also observed. Lastly, there were stable differences in filamentation across isolates in two of the three clades; closely related isolates did not have similar morphologies (Fig. 9).

## Discussion

### Fungus-dominant population

While opportunistic fungal pathogens are often culturable from the sputum of patients with CF, 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 these uncommon fungi 2 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.

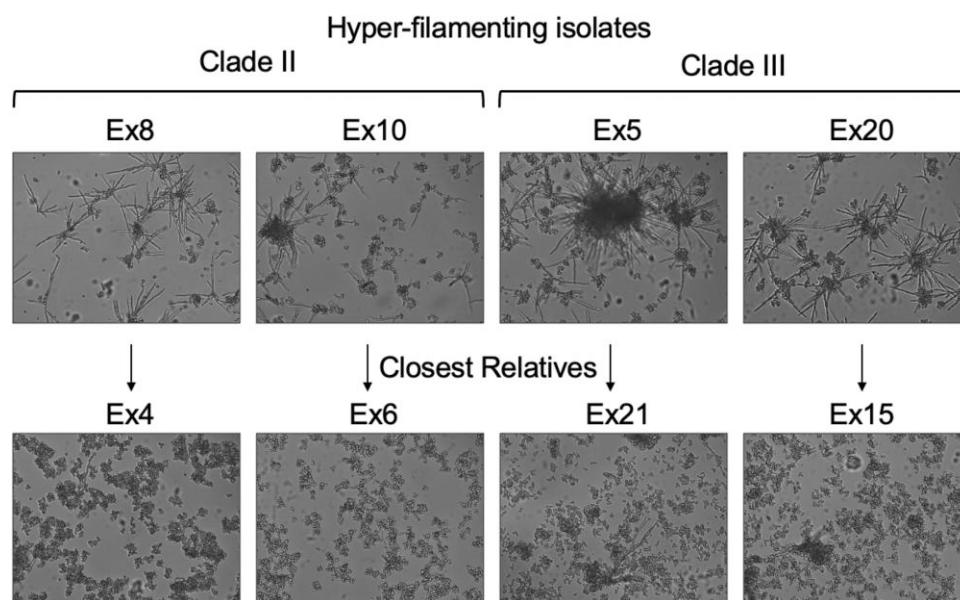
### Phenotypic heterogeneity, melanin production, amino acid metabolism, filamentation, and drug resistance

Developing methods for evaluating population structure and heterogeneity in infecting microbes will aid 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. Melanin production in fungi is associated with increased survival in stress conditions, virulence, dissemination, and host immune response (Eisenman and Casadevall 2012). *Exophiala dermatitidis* produce DHN-melanin from 1,8 dihydroxynaphthalene polymerization, the more common and better-characterized pathway in fungi, compared to L-DOPA, produced DOPA-melanin, or pyomelanin produced from the degradation of tyrosine (Jia et al. 2021). DHN synthesis initiates from endogenous acetyl CoA or melanyl CoA, which is converted to 1,3,6,8 tetrahydroxynaphthalene by polyketide synthase, and converted to DHN-melanin through a series of reductions and polymerizations. While we observed stable differences in melanin production in closely related isolates with few mutations differentiating them, future experiments are required to determine whether these SNPs play a role in melanin synthesis, cell wall localization, or other melanin-associated biological processes. While melanin, which comprises part of the cell wall in black yeasts, can influence antifungal resistance, there were no obvious differences in sensitivity to Itraconazole and Voriconazole that correlated with melanin production (Fig. 8). Other characteristics, such as growth rate or auxotrophy also showed no correlation with levels of melanization.

The discovery of amino acid auxotrophs in the population was consistent with reports of auxotrophs among bacterial CF isolates and likely reflects the presence of amino acids in the lumen of the CF lung (Barth and Pitt 1995; Thomas et al. 2000). Auxotrophy was not monophyletic, and neither was growth as filaments in RPMI with 5% CO<sub>2</sub>. The filamentation phenotype in phylogenetically diverse isolates may indicate a nonmutation-based difference between isolates such as “switching” or phase variation, epigenetic regulation, or strong repeated selection for the phenotype.

### Siderophore gene mutations and iron acquisition

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, *P. aeruginosa* produces the fluorescent siderophores, pyoverdine, and pyochelin, which are used to sequester iron



**Fig. 9.** Cell morphology by microscopy. A hyper-filamentous cell morphology was found among members of two of the clades (top row). For each of these isolates, closely related isolates from the same clade grew in the typical yeast-like cellular morphology (second row). All isolates were grown in RPMI-1640 medium for 18 h at 37°C in a 3-ml culture 6-well dish format. Culture aliquots were imaged with a Zeiss Axiovert 200M inverted microscope.

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). A study by Zajc et al. (2019) indicated 10 predicted siderophores in four *E. 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) encodes the nonribosomal peptide synthase SidC necessary for siderophore synthesis (Zajc et al. 2019; Malo et al. 2021). Polymicrobial infections persist in the CF lung environment through the 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 and Noverr Mairi 2022). 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 2-year timeline (Warren et al. 2011).

### 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 ONT long-read technology sequencing to improve 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 Fig. 4, chromosomal aneuploidies seen in Fig. 3, the very high SNP and INDEL counts seen in Fig. 5, the increased mutation rates seen in Fig. 6, and hyper-filamentous phenotypes seen in Fig. 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 2-year span. Are there physical separations preventing 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. The iron acquisition can be a possible virulence method *E. dermatitidis* could be used to persist in the lung environment as seen in *Aspergillus fumigatus* and *P. 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 nonsynonymous mutation, MRS3/4 a mitochondrial iron transporter, may also help point to the pathogenicity and virulence (Murante et al. 2023). 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 iron molecules from its environment. Further analysis of these variants will point to supporting phenotypic differences allowing for effective colonization of CF lung environments. It is 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, and no mating was detected as all isolates were haploid, thus indicating the low likelihood that even given the opportunity by proximity, they are unable to mate.

## Conclusions

FAs it has become increasingly clear, collecting a single isolate and using it as a metric to assess an environment or moment in time is likely to misrepresent the population; a population-level approach will likely provide more insight into microbes present in infections (Demers et al. 2018). Population-level studies greatly benefit from having a closely related reference strain for the purposes of assessing variants. 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 [Supplementary 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 the Github repository [https://github.com/tania-k/CF\\_Exophiala\\_dermatitidis](https://github.com/tania-k/CF_Exophiala_dermatitidis) and archived in Zenodo (Kurbessoian 2022).

[Supplemental material](#) available at G3 online.

## Funding

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 NIH grant 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.

## Conflicts of interest

The author(s) declare no conflict of interest.

## Literature cited

- Alonge M, Soyk S, Ramakrishnan S, Wang X, Goodwin S, Sedlazeck FJ, Lippman ZB, Schatz MC. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. *Genome Biol.* 2019;20(1):224. doi: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. Scavenging iron: a novel mechanism of plant immunity activation by microbial siderophores. *Plant Physiol.* 2014;164(4):2167–2183. doi:10.1104/pp.113.233585.
- Babič MN, Zupančič J, Gunde-Cimerman N, de Hoog S, Zalar P. Ecology of the human opportunistic black yeast *Exophiala dermatitidis* indicates preference for human-made habitats. *Mycopathologia.* 2018;183(1):201–212. doi:10.1007/s11046-017-0134-8.
- Banin E, Vasil ML, Peter Greenberg E. Iron and *Pseudomonas aeruginosa* biofilm formation. *Proc Natl Acad Sci USA.* 2005;102(31):11076–11081. doi:10.1073/pnas.0504266102.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012;19(5):455–477. doi:10.1089/cmb.2012.0021.
- Bao W, Kojima KK, Kohany O. Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob DNA.* 2015;6(1):11. doi:10.1186/s13100-015-0041-9.
- Barth AL, Pitt TL. Auxotrophic variants of *Pseudomonas aeruginosa* are selected from prototrophic wild-type strains in respiratory infections in patients with cystic fibrosis. *J Clin Microbiol.* 1995;33(1):37–40. doi:10.1128/jcm.33.1.37-40.1995.
- Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 2021;49(W1):W29–W35. doi:10.1093/nar/gkab335.
- Brūna T, Lomsadze A, Borodovsky M. GeneMark-EP: eukaryotic gene prediction with self-training in the space of genes and proteins. *NAR Genom Bioinform.* 2020;2(2):lqaa026. doi:10.1093/nargab/lqaa026.
- Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods.* 2015;12(1):59–60. doi:10.1038/nmeth.3176.
- Burns JL, Emerson J, Stapp JR, Yim DL, Krzewinski J, Loudon L, Ramsey BW, Clausen CR. Microbiology of sputum from patients at cystic fibrosis centers in the United States. *Clin Infect Dis.* 1998;27(1):158–163. doi:10.1086/514631.
- Bushnell B. BMap: A Fast, Accurate, Splice-Aware Aligner. Berkeley, CA (United States): Lawrence Berkeley National Lab (LBNL); 2014.
- Cabanettes F, Klopp C. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ.* 2018;6:e4958. doi:10.7717/peerj.4958.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;10:421. doi:10.1186/1471-2105-10-421.
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4(1):7. doi:10.1186/s13742-015-0047-8.
- Chen M, Kondori N, Deng S, van den Ende AHGG, Lackner M, Liao W, de Hoog GS. Direct detection of *Exophiala* and *Scedosporium* species

- in sputa of patients with cystic fibrosis. *Med Mycol.* 2018;56(6): 695–702. doi:10.1093/mmy/myx108.
- Chen Z, Martinez DA, Gujja S, Sykes SM, Zeng Q, Szaniszló PJ, Wang Z, Cuomo CA. Comparative genomic and transcriptomic analysis of *Wangiella dermatitidis*, a major cause of phaeohyphomycosis and a model black yeast human pathogen. *G3 (Bethesda)*. 2014; 4(4):561–578. doi:10.1534/g3.113.009241.
- Cimon B, Carrère J, Vinatier JF, Chazalette JP, Chabasse D, Bouchara JP. Clinical significance of *Scedosporium apiospermum* in patients with cystic fibrosis. *Eur J Clin Microbiol Infect Dis.* 2000;19(1): 53–56. doi:10.1007/s100960050011.
- Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)*. 2012;6(2):80–92. doi:10.4161/fly.19695.
- Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, et al. Biopython: freely available python tools for computational molecular biology and bioinformatics. *Bioinformatics.* 2009;25(11):1422–1423. doi:10.1093/bioinformatics/btp163.
- Conti-Díaz IA, Mackinnon JE, Civilia E. Isolation and identification of black yeasts from the external environment in Uruguay. *Pan Amer Health Org Sci Publ.* 1977;356:109–114.
- Crowley DE, Wang YC, Reid CPP, Szaniszló PJ. 1991. Mechanisms of iron acquisition from siderophores by microorganisms and plants. In: Chen Y, Hadar Y, editors. *Iron Nutrition and Interactions in Plants*, 11–17 June 1989, Jerusalem, Israel, 1989. Dordrecht: Springer Netherlands. p. 213–232.
- Damveld RA, Arentshorst M, Franken A, van Kuyk PA, Klis FM, van den Hondel CAMJJ, Ram AFJ. The *Aspergillus Niger* MADS-box transcription factor RlmA is required for cell wall reinforcement in response to cell wall stress. *Mol Microbiol.* 2005;58(1):305–319. doi:10.1111/j.1365-2958.2005.04827.x.
- Davis PB. Cystic fibrosis since 1938. *Am J Respir Crit Care Med.* 2006; 173(5):475–482. doi: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. Genotyping study of *Scedosporium apiospermum* isolates from patients with cystic fibrosis. *J Clin Microbiol.* 2002;40(6):2108–2114. doi:10.1128/jcm.40.6.2108-2114.2002.
- De Hoog GS, Queiroz-Telles F, Haase G, Fernandez-Zeppenfeldt G, Angelis DA, Gerrits van den Ende AH, Matos T, Peltroche-Llacsahuanga H, Pizzirani-Kleiner AA, Rainer J, et al. 2000. Black fungi: clinical and pathogenic approaches. *Med Mycol.* 38(1):243–250. doi:10.1080/mmy.38.s1.243.250.
- De Jong CCM, de Jong CCM, Slabbers L, Engel TGP, Yntema JB, van Westreenen M, Croughe PD, Roeleveld N, Brimicombe R, Verweij PE, et al. 2020. Clinical relevance of *Scedosporium* spp. and *Exophiala dermatitidis* in patients with cystic fibrosis: a nationwide study. *Med Mycol.* 58(7):859–866. doi:10.1093/mmy/myaa003.
- Demers EG, Biermann AR, Masonjones S, Crocker AW, Ashare A, Stajich JE, Hogan DA. Evolution of drug resistance in an antifungal-naïve chronic *Candida lusitanae* infection. *Proc Natl Acad Sci USA.* 2018;115(47):12040–12045. doi:10.1073/pnas.1807698115.
- Diemert D, Kunimoto D, Sand C, Rennie R. Sputum isolation of *Wangiella dermatitidis* in patients with cystic fibrosis. *Scand J Infect Dis.* 2001;33(10):777–779. doi:10.1080/003655401317074644.
- Dixon DM, Shadomy HJ, Shadomy S. Dematiaceous fungal pathogens isolated from nature. *Mycopathologia.* 1980;70(3):153–161. doi:10.1007/BF00443026.
- Eisenman HC, Casadevall A. 2012. Synthesis and assembly of fungal melanin. *Appl Microbiol Biotechnol.* 93(3):931–940. doi:10.1007/s00253-011-3777-2. <http://dx.doi.org/10.1007/s00253-011-3777-2>.
- Emms DM, Kelly S. Orthofinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 2015;16(1):157. doi:10.1186/s13059-015-0721-2.
- Emms DM, Kelly S. Orthofinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 2019;20(1):238. doi:10.1186/s13059-019-1832-y.
- Ferec C, Cutting GR. Assessing the disease-liability of mutations in CFTR. *Cold Spring Harb Perspect Med.* 2012;2(12):a009480. doi: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. Pfam: the protein families database. *Nucleic Acids Res.* 2014;42(D1):D222–D230. doi:10.1093/nar/gkt1223.
- Fothergill AW, Rinaldi MG, Sutton DA. Antifungal susceptibility testing of *Exophiala* spp.: a head-to-head comparison of amphotericin B, itraconazole, posaconazole and voriconazole. *Med Mycol.* 2009; 47(1):41–43. doi:10.1080/13693780802512451.
- Frank W, Roester U. Amphibia as carriers of *Hormiscium (Hormodendrum) dermatitidis* a causative agent of chromoblastomycosis (chromomycosis) in man. *Z Tropenmed Parasitol.* 1970;21(1):93–108.
- Franke KR, Crowgey EL. Accelerating next generation sequencing data analysis: an evaluation of optimized best practices for genome analysis toolkit algorithms. *Genomics Inform.* 2020;18(1): e10. doi:10.5808/GI.2020.18.1.e10.
- Freire-Benítez V, Gourlay S, Berman J, Buscaino A. Sir2 regulates stability of repetitive domains differentially in the human fungal pathogen *Candida albicans*. *Nucleic Acids Res.* 2016;44(19): 9166–9179. doi:10.1093/nar/gkw594.
- Gilchrist CLM, Booth TJ, van Werssch B, van Grieken L, Medema MH, Chooi Y-H. Cblaster: a remote search tool for rapid identification and visualization of homologous gene clusters. *Bioinform Adv.* 2021;1(1):vbab016. doi:10.1093/bioadv/vbab016.
- Gilchrist CLM, Chooi Y-H. Clinker & clustermap.js: automatic generation of gene cluster comparison figures. *Bioinformatics.* 2021; 37(16):2473–2475. doi:10.1093/bioinformatics/btab007.
- Goldman GH, McGuire SL, Harris SD. The DNA damage response in filamentous fungi. *Fungal Genet Biol.* 2002;35(3):183–195. doi:10.1006/fgbi.2002.1344.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011;29(7):644–652. doi:10.1038/nbt.1883.
- Grahl N, Dolben EL, Filkins LM, Crocker AW, Willger SD, Morrison HG, Sogin ML, Ashare A, Gifford AH, Jacobs NJ, et al. Profiling of bacterial and fungal microbial communities in cystic fibrosis sputum using RNA. *mSphere.* 2018;3(4):e00292-18. doi:10.1128/mSphere.00292-18.
- Grenouillet F, Cimon B, Pana-Katatali H, Person C, Gainet-Brun M, Malinge M-C, Le Govic Y, Richaud-Thiriez B, Bouchara J-P. *Exophiala dermatitidis* revealing cystic fibrosis in adult patients with chronic pulmonary disease. *Mycopathologia.* 2018;183(1): 71–79. doi:10.1007/s11046-017-0218-5.
- Griffard EA, Guajardo JR, Cooperstock MS, Scoville CL. Isolation of *Exophiala dermatitidis* from pigmented sputum in a cystic fibrosis patient. *Pediatr Pulmonol.* 2010;45(5):508–510. doi:10.1002/ppul.21187.
- Haas H. Fungal siderophore metabolism with a focus on *Aspergillus fumigatus*. *Nat Prod Rep.* 2014;31(10):1266–1276. doi:10.1039/c4np00071d.
- Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the program to assemble spliced alignments. *Genome Biol.* 2008;9(1):R7. doi:10.1186/gb-2008-9-1-r7.

- Haase G, Skopnik H, Kusenbach G. *Exophiala dermatitidis* infection in cystic fibrosis. *Lancet*. 1990;336(8708):188–189. doi:10.1016/0140-6736(90)91721-1.
- Hiltunen M, Ament-Velásquez SL, Johannesson H. The assembled and annotated genome of the fairy-ring fungus *marasmius oreades*. *Genome Biol Evol*. 2021;13(7):evab126. doi:10.1093/gbe/evab126.
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol*. 2018;35(2):518–522. doi:10.1093/molbev/msx281.
- Horré R, Marklein G, Siekmeier R, Nidermajer S, Reiffert SM. Selective isolation of *Pseudallescheria* and *Scedosporium* species from respiratory tract specimens of cystic fibrosis patients. *Respiration*. 2009;77(3):320–324. doi:10.1159/000167419.
- Horré R, Schaal KP, Siekmeier R, Sterzik B, de Hoog GS, Schnitzler N. Isolation of fungi, especially *Exophiala dermatitidis*, in patients suffering from cystic fibrosis. *Respiration*. 2004;71(4):360–366. doi:10.1159/000079640.
- Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res*. 2019;47(D1):D309–D314. doi:10.1093/nar/gky1085.
- Institute. “Picard Toolkit.” Broad Institute, GitHub Repository. Picard Toolkit. <https://github.com/broadinstitute/picard>.
- Jia W, Li H, Wang Q, Zheng K, Lin H, Li X, Huang J, Xu L, Dong W, Shu Z. Screening of perhydrolases to optimize glucose oxidase-perhydrolase-in situ chemical oxidation cascade reaction system and its application in melanin decolorization. *J Biotechnol*. 2021;328:106–114. doi:10.1016/j.jbiotec.2021.01.013.
- Johnson L. Iron and siderophores in fungal–host interactions. *Mycol Res*. 2008;112(2):170–183. doi:10.1016/j.mycres.2007.11.012.
- Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. Interproscan 5: genome-scale protein function classification. *Bioinformatics*. 2014;30(9):1236–1240. doi:10.1093/bioinformatics/btu031.
- Jones JT, Liu K-W, Wang X, Kowalski CH, Ross BS, Mills KAM, Kerkaert JD, Hohl TM, Lofgren LA, Stajich JE, et al. *Aspergillus fumigatus* strain-specific conidia lung persistence causes an allergic broncho-pulmonary aspergillosis-like disease phenotype. *mSphere*. 2021;6(1):e01250-20. doi:10.1128/msphere.01250-20.
- Kim DH. Bacterial siderophores promote animal host iron acquisition and growth. *Cell*. 2018;175(2):311–312. doi:10.1016/j.cell.2018.09.020.
- Kim SH, Clark ST, Surendra A, Copeland JK, Wang PW, Ammar R, Collins C, Tullis DE, Nislow C, Hwang DM, et al. Global analysis of the fungal microbiome in cystic fibrosis patients reveals loss of function of the transcriptional repressor *nrg1* as a mechanism of pathogen adaptation. *PLoS Pathog*. 2015;11(11):e1005308. doi:10.1371/journal.ppat.1005308.
- Kirchhoff L, Olsowski M, Rath P-M, Steinmann J. *Exophiala dermatitidis*: key issues of an opportunistic fungal pathogen. *Virulence*. 2019;10(1):984–998. doi:10.1080/21505594.2019.1596504.
- Klasinc R, Riesenhuber M, Bacher A, Willinger B. Invasive fungal infection caused by *Exophiala dermatitidis* in a patient after lung transplantation: case report and literature review. *Mycopathologia*. 2019;184(1):107–113. doi:10.1007/s11046-018-0275-4.
- Kondori N, Gilljam M, Lindblad A, Jönsson B, Moore ERB, Wennerås C. High rate of *Exophiala dermatitidis* recovery in the airways of patients with cystic fibrosis is associated with pancreatic insufficiency. *J Clin Microbiol*. 2011;49(3):1004–1009. doi:10.1128/JCM.01899-10.
- Kondori N, Lindblad A, Welinder-Olsson C, Wennerås C, Gilljam M. Development of IgG antibodies to *Exophiala dermatitidis* is associated with inflammatory responses in patients with cystic fibrosis. *J Cyst Fibros*. 2014;13(4):391–399. doi:10.1016/j.jcf.2013.12.007.
- Korf I. Gene finding in novel genomes. *BMC Bioinformatics*. 2004;5(1):59. doi:10.1186/1471-2105-5-59.
- Krogh BO, Symington LS. Recombination proteins in yeast. *Annu Rev Genet*. 2004;38(1):233–271. doi:10.1146/annurev.genet.38.072902.091500.
- Kurbessoian T. tania-k/CF\_Exophiala\_dermatitidis: Version 6 update 2022. <https://zenodo.org/record/7106110>.
- Kusenbach G, Skopnik H, Haase G, Friedrichs F, Döhmen H. *Exophiala dermatitidis* pneumonia in cystic fibrosis. *Eur J Pediatr*. 1992;151(5):344–346. doi:10.1007/BF02113255.
- Lavrin T, Konte T, Kostanjšek R, Sitar S, Sepčić K, Prpar Mihevc S, Žagar E, Župunski V, Lenassi M, Rogelj B, et al. The neurotropic black yeast *Exophiala dermatitidis* induces neurocytotoxicity in neuroblastoma cells and progressive cell death. *Cells*. 2020;9(4):963. doi:10.3390/cells9040963.
- Legrand M, Jaitly P, Feri A, d’Enfert C, Sanyal K. *Candida albicans*: an emerging yeast model to study eukaryotic genome plasticity. *Trends Genet*. 2019;35(4):292–307. doi:10.1016/j.tig.2019.01.005.
- Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res*. 2016;44(W1):W242–W245. doi:10.1093/nar/gkw290.
- Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34(18):3094–3100. doi:10.1093/bioinformatics/bty191.
- Li H. Protein-to-genome alignment with minimap. *Bioinformatics*. 2023;39(1):btad014. doi:10.1093/bioinformatics/btad014.
- Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2010;26(5):589–595. doi:10.1093/bioinformatics/btp698.
- Li H, Handsaker B, Danecek P, McCarthy S, Marshall J. BCFtools 2019. <https://github.com/samtools/bcftools>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–2079. doi:10.1093/bioinformatics/btp352.
- Lofgren LA, Uehling JK, Branco S, Bruns TD, Martin F, Kennedy PG. Genome-based estimates of fungal rDNA copy number variation across phylogenetic scales and ecological lifestyles. *Mol Ecol*. 2019;28(4):721–730. doi:10.1111/mec.14995.
- Lowe TM, Chan PP. tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res*. 2016;44(W1):W54–W57. doi:10.1093/nar/gkw413.
- Majoros WH, Pertea M, Salzberg SL. Tigrscan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics*. 2004;20(16):2878–2879. doi:10.1093/bioinformatics/bth315.
- Malo ME, Schultzhause Z, Frank C, Romsdahl J, Wang Z, Dadachova E. Transcriptomic and genomic changes associated with radioadaptation in *Exophiala dermatitidis*. *Comput Struct Biotechnol J*. 2021;19:196–205. doi:10.1016/j.csbj.2020.12.013.
- Manni M, Berkeley MR, Seppay M, Simão FA, Zdobnov EM. BUSCO Update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 2021;38(10):4647–4654. doi:10.1093/molbev/msab199.
- Mariani-Kurkdjian P, Bingen E. Bactéries pathogènes dans la mucoviscidose. *Archives de Pédiatrie*. 2003;10(Suppl 2):S342–S346. doi:10.1016/S0929-693X(03)90050-9.
- Matilla MA, Ramos JL, Duque E, de Dios Alché J, Espinosa-Urgel M, Ramos-González MI. Temperature and pyoverdine-mediated

- iron acquisition control surface motility of *Pseudomonas putida*. *Environ Microbiol.* 2007;9(7):1842–1850. doi:10.1111/j.1462-2920.2007.01286.x.
- Matos T, de Hoog GS, de Boer AG, de Crom I, Haase G. High prevalence of the neurotrope *Exophiala dermatitidis* and related oligotrophic black yeasts in sauna facilities. *Mycoses.* 2002;45(9–10):373–377. doi:10.1046/j.1439-0507.2002.00779.x.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297–1303. doi:10.1101/gr.107524.110.
- Metin B, Döğen A, Yıldırım E, de Hoog GS, Heitman J, Ilkit M. Mating type (MAT) locus and possible sexuality of the opportunistic pathogen *Exophiala dermatitidis*. *Fungal Genet Biol.* 2019;124:29–38. doi:10.1016/j.fgb.2018.12.011.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol.* 2020;37(5):1530–1534. doi:10.1093/molbev/msaa015.
- Mochochoko BM, Ezeokoli OT, Sebolai O, Albertyn J, Pohl CH. Role of the high-affinity reductive iron acquisition pathway of *Candida albicans* in prostaglandin E2 production, virulence, and interaction with *Pseudomonas aeruginosa*. *Med Mycol.* 2021;59(9):869–881. doi:10.1093/mmy/myab015.
- Mossialos D, Amoutzias GD. Role of siderophores in cystic fibrosis pathogenesis: foes or friends? *Int J Med Microbiol.* 2009;299(2):87–98. doi:10.1016/j.ijmm.2008.06.008.
- Mukai Y, Nureki S-I, Hata M, Shigenaga T, Tokimatsu I, Miyazaki E, Kadota J-I, Yaritha K, Kamei K. *Exophiala dermatitidis* pneumonia successfully treated with long-term itraconazole therapy. *J Infect Chemother.* 2014;20(7):446–449. doi:10.1016/j.jiac.2014.02.006.
- Murante D, Demers EG, Kurbessoian T, Ruzic M, Ashare A, Stajich JE, Hogan DA. Mrs4 loss of function in fungi during adaptation to the cystic fibrosis lung. *bioRxiv.* 2023. doi:10.1101/2023.04.05.535776.
- Nagano Y, Millar BC, Johnson E, Goldsmith CE, Elborn JS, Rendall J, Moore JE. Fungal infections in patients with cystic fibrosis. *Rev Med Microbiol.* 2007;18(1):11–16. doi:10.1097/mrm.0b013e3282e1c70a.
- Neilands JB. Siderophores: structure and function of microbial iron transport compounds. *J Biol Chem.* 1995;270(45):26723–26726. doi:10.1074/jbc.270.45.26723.
- Nishimura K, Miyaji M. Studies on a saprophyte of *Exophiala dermatitidis* isolated from a humidifier. *Mycopathologia.* 1982;77(3):173–181. doi:10.1007/bf00518803.
- Nosanchuk JD, Casadevall A. The contribution of melanin to microbial pathogenesis. *Cell Microbiol.* 2003;5(4):203–223. doi:10.1046/j.1462-5814.2003.00268.x.
- Nosanchuk JD, Casadevall A. Impact of melanin on microbial virulence and clinical resistance to antimicrobial compounds. *Antimicrob Agents Chemother.* 2006;50(11):3519–3528. doi:10.1128/aac.00545-06.
- Packeu A, Lebecque P, Rodriguez-Villalobos H, Boeras A, Hendrickx M, Bouchara J-P, Symoens F. Molecular typing and antifungal susceptibility of *Exophiala* isolates from patients with cystic fibrosis. *J Med Microbiol.* 2012;61(9):1226–1233. doi:10.1099/jmm.0.042317-0.
- Palmer JM, Stajich JE. Funannotate v1.8.1: Eukaryotic genome annotation 2020. <https://zenodo.org/record/4054262>. doi: 10.5281/zenodo.1134477.
- Palmer JM, Stajich JE. Automatic assembly for the fungi (AAFTF): genome assembly pipeline 2022. <https://zenodo.org/record/6326242>. doi: 10.5281/zenodo.1620526.
- Paolo WF Jr, Dadachova E, Mandal P, Casadevall A, Szaniszló PJ, Nosanchuk JD. Effects of disrupting the polyketide synthase gene WdPKS1 in *Wangiella [Exophiala] dermatitidis* on melanin production and resistance to killing by antifungal compounds, enzymatic degradation, and extremes in temperature. *BMC Microbiol.* 2006;6(1):55. doi:10.1186/1471-2180-6-55.
- Parize P, Billaud S, Bienvenu AL, Bourdy S, le Pogam MA, Reix P, Picot S, Robert R, Lortholary O, Bouchara J-P, et al. Impact of *Scedosporium apiospermum* complex seroprevalence in patients with cystic fibrosis. *J Cyst Fibros.* 2014;13(6):667–673. doi:10.1016/j.jcf.2014.01.011.
- Pedersen BS, Quinlan AR. Mosdepth: quick coverage calculation for genomes and exomes. *Bioinformatics.* 2018;34(5):867–868. doi:10.1093/bioinformatics/btx699.
- Pihet M, Carrere J, Cimon B, Chabasse D, Delhaes L, Symoens F, Bouchara J-P, Pihet M, Carrere J, Cimon B, et al. Occurrence and relevance of filamentous fungi in respiratory secretions of patients with cystic fibrosis—a review. *Med Mycol.* 2009;47(4):387–397. doi:10.1080/13693780802609604.
- Pohl CH. Competition for iron during polymicrobial infections may increase antifungal drug susceptibility—how will it impact treatment options? *Infect Immun.* 2022;90(4):e0005722. doi:10.1128/iai.00057-22.
- Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER Web server: 2018 update. *Nucleic Acids Res.* 2018;46(W1):W200–W204. doi: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. MADS-box transcription factor SsMADS is involved in regulating growth and virulence in *Sclerotinia sclerotiorum*. *Int J Mol Sci.* 2014;15(5):8049–8062. doi:10.3390/ijms15058049.
- Rath P-M, Muller K-D, Dermoumi H, Ansorg R. A comparison of methods of phenotypic and genotypic fingerprinting of *Exophiala dermatitidis* isolated from sputum samples of patients with cystic fibrosis. *J Med Microbiol.* 1997;46(9):757–762. doi:10.1099/00222615-46-9-757.
- Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* 2018;46(D1):D624–D632. doi:10.1093/nar/gkx1134.
- Reiss NR, Mok WY. *Wangiella dermatitidis* isolated from bats in Manaus Brazil. *Sabouraudia.* 1979;17(3):213–218. doi:10.1080/00362177985380321.
- Revankar SG, Patterson JE, Sutton DA, Pullen R, Rinaldi MG. Disseminated phaeohyphomycosis: review of an emerging mycosis. *Clin Infect Dis.* 2002;34(4):467–476. doi:10.1086/338636.
- Revankar SG, Sutton DA. Melanized fungi in human disease. *Clin Microbiol Rev.* 2010;23(4):884–928. doi:10.1128/CMR.00019-10.
- Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavski N, Chou JL. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science.* 1989;245(4922):1066–1073. doi:10.1126/science.2475911.
- Robertson KL, Mostaghim A, Cuomo CA, Soto CM, Lebedev N, Bailey RF, Wang Z. Adaptation of the black yeast *Wangiella dermatitidis* to ionizing radiation: molecular and cellular mechanisms. *PLoS One.* 2012;7(11):e48674. doi:10.1371/journal.pone.0048674.
- Ross BS, Lofgren LA, Ashare A, Stajich JE, Cramer RA. *Aspergillus fumigatus* in-host HOG pathway mutation for cystic fibrosis lung microenvironment persistence. *MBio.* 2021;12(4):e0215321. doi:10.1128/mBio.02153-21.
- Sass G, Ansari SR, Dietl A-M, Déziel E, Haas H, Stevens DA. Intermicrobial interaction: *aspergillus fumigatus* siderophores protect against competition by *Pseudomonas aeruginosa*. *PLoS One.* 2019;14(5):e0216085. doi:10.1371/journal.pone.0216085.
- Saunte DM, Tarazooie B, Arendrup MC, de Hoog GS. Black yeast-like fungi in skin and nail: it probably matters. *Mycoses.* 2012;55(2):161–167. doi:10.1111/j.1439-0507.2011.02055.x.

- Schrettl M, Bignell E, Kragl C, Sabaia Y, Loss O, Eisendle M, Wallner A, Arst HN, Haynes K, Haas H. Distinct roles for intra- and extracellular siderophores during *Aspergillus fumigatus* infection. *PLoS Pathog*. 2007;3(9):e128. doi:10.1371/journal.ppat.0030128.
- Schultzhaus Z, Romsdahl J, Chen A, Tschirhart T, Kim S, Leary D, Wang Z. The response of the melanized yeast *Exophiala dermatitidis* to gamma radiation exposure. *Environ Microbiol*. 2020;22(4):1310–1326. doi:10.1111/1462-2920.14936.
- Selmecki A, Forche A, Berman J. Genomic plasticity of the human fungal pathogen *Candida albicans*. *Eukaryot Cell*. 2010;9(7):991–1008. doi:10.1128/EC.00060-10.
- Seyedmousavi S, Netea MG, Mouton JW, Melchers WJG, Verweij PE, de Hoog GS. Black yeasts and their filamentous relatives: principles of pathogenesis and host defense. *Clin Microbiol Rev*. 2014;27(3):527–542. doi:10.1128/CMR.00093-13.
- Shore P, Sharrocks AD. The MADS-box family of transcription factors. *Eur J Biochem*. 1995;229(1):1–13. doi:10.1111/j.1432-1033.1995.00011.x.
- Slater GSC, Birney E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinform*. 2005;6(1):31. doi:10.1186/1471-2105-6-31.
- Smit AFA. 2004. Repeat-Masker Open-3.0. <http://www.repeatmasker.org>. <https://ci.nii.ac.jp/naid/10029514778/>.
- Smit AFA, Hubley R. 2008. RepeatModeler Open-1.0. <https://ci.nii.ac.jp/naid/10029514778/>.
- Smith DFQ, Casadevall A. The role of melanin in fungal pathogenesis for animal hosts. *Curr Top Microbiol Immunol*. 2019;422:1–30. doi:10.1007/82\_2019\_173.
- Song Y, de Sande WWJL, Moreno LF, van den Ende BG, Li R, de Hoog S. Comparative ecology of capsular *exophiala* Species causing disseminated infection in humans. *Front Microbiol*. 2017;8:2514. doi:10.3389/fmicb.2017.02514.
- Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res*. 2006;34(Web Server):W435–W439. doi:10.1093/nar/gkl200.
- Steinkamp G, Wiedemann B, Rietschel E, Krahel A, Gielen J, Bärmeier H, Ratjen F. Prospective evaluation of emerging bacteria in cystic fibrosis. *J Cyst Fibros*. 2005;4(1):41–48. doi:10.1016/j.jcf.2004.10.002.
- Sudfeld CR, Dasenbrook EC, Merz WG, Carroll KC, Boyle MP. Prevalence and risk factors for recovery of filamentous fungi in individuals with cystic fibrosis. *J Cyst Fibros*. 2010;9(2):110–116. doi:10.1016/j.jcf.2009.11.010.
- Sudhadham M, Gerrits van den Ende AHG, Sihanonth P, Sivichai S, Chaiyarat R, Menken SBJ, van Belkum A, de Hoog GS. Elucidation of distribution patterns and possible infection routes of the neurotropic black yeast *Exophiala dermatitidis* using AFLP. *Fungal Biol*. 2011;115(10):1051–1065. doi:10.1016/j.funbio.2010.07.004.
- Sudhadham M, Prakitsin S, Sivichai S, Chaiyarat R, Dorrestein GM, Menken SBJ, de Hoog GS. The neurotropic black yeast *Exophiala dermatitidis* has a possible origin in the tropical rainforest. *Stud Mycol*. 2008;61:145–155. doi:10.3114/sim.2008.61.15.
- Teixeira MM, Moreno LF, Stielow BJ, Muszewska A, Hainaut M, Gonzaga L, Abouelleil A, Patané JSL, Priest M, Souza R, et al. Exploring the genomic diversity of black yeasts and relatives (Chaetothyriales, Ascomycota). *Stud Mycol*. 2017;86(1):1–28. doi:10.1016/j.simyco.2017.01.001.
- Thomas SR, Ray A, Hodson ME, Pitt TL. Increased sputum amino acid concentrations and auxotrophy of *Pseudomonas aeruginosa* in severe cystic fibrosis lung disease. *Thorax*. 2000;55(9):795–797. doi:10.1136/thorax.55.9.795.
- Tunney MM, Field TR, Moriarty TF, Patrick S, Doering G, Muhlebach MS, Wolfgang MC, Boucher R, Gilpin DF, McDowell A, et al. Detection of anaerobic Bacteria in high numbers in sputum from patients with cystic fibrosis. *Am J Respir Crit Care Med*. 2008;177(9):995–1001. doi:10.1164/rccm.200708-1151c.
- Turcios NL. Cystic fibrosis lung disease: an overview. *Respir Care*. 2020;65(2):233–251. doi:10.4187/respcare.06697.
- Tyrrell J, Callaghan M. Iron acquisition in the cystic fibrosis lung and potential for novel therapeutic strategies. *Microbiology*. 2016;162(2):191–205. doi:10.1099/mic.0.000220.
- Uijthof MJM, Hoog GS, Cock AWA, Takeo K, Nishimura K. Pathogenicity of strains of the black yeast *Exophiala (Wangiella) dermatitidis*: an evaluation based on polymerase chain reaction. *Mycoses*. 2009;37(7–8):235–242. doi:10.1111/j.1439-0507.1994.tb00419.x.
- Warren AE, Boulianne-Larsen CM, Chandler CB, Chiotti K, Kroll E, Miller SR, Taddei F, Sermet-Gaudelus I, Ferroni A, McInnerney K, et al. Genotypic and phenotypic variation in *Pseudomonas aeruginosa* reveals signatures of secondary infection and mutator activity in certain cystic fibrosis patients with chronic lung infections. *Infect Immun*. 2011;79(12):4802–4818. doi:10.1128/IAI.05282-11.
- Weisman CM, Murray AW, Eddy SR. Mixing genome annotation methods in a comparative analysis inflates the apparent number of lineage-specific genes. *Curr Biol*. 2022;32(12):2632–2639.e2. doi:10.1016/j.cub.2022.04.085.
- Wick RR, Judd LM, Holt KE. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol*. 2019;20(1):129. doi:10.1186/s13059-019-1727-y.
- Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag; 2016.
- Xiong D, Wang Y, Tian L, Tian C. MADS-Box Transcription factor VdMcm1 regulates conidiation, microsclerotia formation, pathogenicity, and secondary metabolism of *Verticillium dahliae*. *Front Microbiol*. 2016;7:1192. doi:10.3389/fmicb.2016.01192.
- Yan K, Yin H, Wang J, Cai Y. Subtle relationships between *Pseudomonas aeruginosa* and fungi in patients with cystic fibrosis. *Acta Clin Belg*. 2022;77(2):425–435. doi:10.1080/17843286.2020.1852850.
- Zajc J, Gostinčar C, Černoša A, Gunde-Cimerman N. Stress-Tolerant yeasts: opportunistic pathogenicity versus biocontrol potential. *Genes (Basel)*. 2019;10(1):42. doi:10.3390/genes10010042.
- Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res*. 2018;46(W1):W95–W101. doi:10.1093/nar/gky418.
- Zhao J, Schloss PD, Kalikin LM, Carmody LA, Foster BK, Petrosino JF, Cavalcoli JD, VanDevanter DR, Murray S, Li JZ, et al. Decade-long bacterial community dynamics in cystic fibrosis airways. *Proc Natl Acad Sci USA*. 2012;109(15):5809–5814. doi:10.1073/pnas.1120577109.
- Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. The MaSuRCA genome assembler. *Bioinformatics*. 2013;29(21):2669–2677. doi:10.1093/bioinformatics/btt476.
- Zupančič J, Babič MN, Zalar P, Gunde-Cimerman N. The black yeast *Exophiala dermatitidis* and other selected opportunistic human fungal pathogens spread from dishwashers to kitchens. *PLoS One*. 2016;11(2):e0148166. doi:10.1371/journal.pone.0148166.

Editor: M. Nowrousian