Genetic diversity and microevolution in clinical *Cryptococcus* isolates from Cameroon

Poppy Sephton-Clark[1], Elvis Temfack[2,3], Jennifer L. Tenor[4], Dena L. Toffaletti[4], Angela Loyse[5,6], Síle F. Molloy[5], John R. Perfect[4], Tihana Bicanic[5,6], Thomas S. Harrison[5,7], Olivier Lortholary[8,9†], Charles Kouanfack[10,11,12†], and Christina A. Cuomo[1†*]

1. Infectious Disease and Microbiome Program, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.
2. Internal Medicine Unit, Douala General Hospital, Cameroon.
3. Institut Pasteur, Molecular Mycology Unit, CNRS UMR 2000, 75015, Paris, France.
4. Division of Infectious Diseases, Department of Medicine, Duke University School of Medicine, Durham, North Carolina, USA.
5. Institute of Infection and Immunity, St George's University of London, London, UK.
6. Clinical Academic Group in Infection, St George's University Hospital, London, UK.
7. MRC Centre for Medical Mycology, University of Exeter, Exeter, UK.
8. Paris Cité University, Department of Infectious Diseases and Tropical Medicine, Necker-Enfants Malades Hospital, AP-HP, IHU Imagine, 149 Rue de Sèvres, 75015, Paris, France.
9. Institut Pasteur, Mycology Department and National Reference Center for Invasive Mycoses and Antifungals, 75015, Paris, France.
10. Department of Public Health, Faculty of Medicine and Pharmaceutical Sciences, University of Dschang, Cameroon.
11. Day Hospital, Hospital Central Yaoundé, Cameroon.
12. Research Center for Emerging and Re-emerging Diseases, Cameroon Baptist Convention Health Services (CBCHS), Yaoundé, Cameroon.

†Authors contributed equally

*Corresponding author: Christina A. Cuomo (cuomo@broadinstitute.org, 1-617-714-7904)

Keywords: Cryptococcus, genome sequencing, intrahost diversity, GWAS, phylogeography

Abstract

Cryptococcal meningitis is the second most common cause of death in people living with HIV/AIDS, yet we have a limited understanding of how cryptococcal isolates change over the course of infection. Cryptococcal infections are environmentally acquired, and the genetic diversity of these infecting isolates can also be geographically linked. Here, we employ whole genome sequences for 372 clinical *Cryptococcus* isolates from 341 patients with HIV-associated cryptococcal meningitis obtained via a large clinical trial, across both Malawi and Cameroon, to enable population genetic comparisons of isolates between countries. We see that isolates from Cameroon are highly clonal, when compared to those from Malawi, with differential rates of disruptive variants in genes with roles in DNA binding and energy use. For a subset of patients (22) from Cameroon, we leverage longitudinal sampling, with samples taken at days 7 and 14 post enrolment, to interrogate the genetic changes that arise over the course of infection, and the genetic diversity of isolates within patients. We see disruptive variants arising over the course of infection in several genes, including the phagocytosis regulating transcription factor GAT204. In addition, in 13% of patients sampled longitudinally we see evidence for mixed infections. This approach identifies geographically linked genetic variation, signatures of microevolution, and evidence for mixed infections across a clinical cohort of patients affected by cryptococcal meningitis in Central Africa.

Lay Summary

Cryptococcal meningitis, caused by *Cryptococcus*, results in approximately half a million deaths per year globally. We compare clinical *Cryptococcus* samples from Cameroon and Malawi to explore genetic diversity of these isolates. We find instances of mixed-strain infections and identify genetic variants arising in *Cryptococcus* over disease.

Introduction

Cryptococcal infections represent a major threat to global health. These infections account for 15-19% of deaths in those living with HIV/AIDS; the majority of fatal cases of cryptococcal meningitis (between 63 to 75% per year) occur in those living in sub-Saharan Africa[1]. The predominant species responsible for disease is *Cryptococcus neoformans*[2]. *C. neoformans* isolates representing all 3 major lineages, VNI, VNB, and the rarer VNII, have been isolated from countries across sub-Saharan Africa[3,4]. While VNI and VNII isolates are found globally, commonly sampled from clinical settings and found environmentally in pigeon guano, VNB isolates are primarily detected in sub-Saharan Africa from soil and tree sources[4–6] and only

rarely isolated in other regions[7,8]. Environmental VNB isolates from Botswana display a strong clade-like structure, with two non-recombining lineages encompassing VNB isolates, VNBI and VNBII, defined through phylogenetic analysis[6]. While VNB isolates are also capable of causing disease, the majority of clinical *C. neoformans* isolates obtained from Uganda, Malawi, Botswana, Zambia, Laos, Thailand, and Vietnam, across multiple clinical trials, belong to the VNI lineage[4,6,9–11]. Sub-lineages termed VNIa, VNIb, and VNIc describe the population structure observed for this lineage[6].

Cryptococcal infections are environmentally acquired, most likely through inhalation of *Cryptococcus* yeast cells or spores[12,13]. Cryptococcosis can also arise through reactivation of a latent infection[14–16]. This in turn raises the question of whether cryptococcal infections are the result of single or multiple genetically distinct isolates. Longitudinal sampling has shown that the majority (89-100%) of recurrent infections are caused by a single isolate, based on high genome-wide identity between the incident and recurrent samples[17,18]. However, studies that analyzed sequence data for multiple colonies from patients at a single time point have shown that mixed infections may be common and are observed in up to 18-30% of patients when multiple colonies are screened[19,20].

To assess lineage, population structure, and within-patient relatedness, we sequenced 86 clinical isolates obtained from patients enrolled at the Cameroon site for the multi-country phase III non-inferiority trial, Advancing Cryptococcal meningitis Treatment in Africa (ACTA)[21]. Based on longitudinal sample analysis we assess the likelihood of patient infections arising from multiple infecting strains. We compare isolates from Cameroon to those obtained from Malawi in the same clinical trial and across the same timeframe to assess country linked population structure and identify variants significantly associated with isolate country of origin.

Materials and Methods

Sample preparation and sequencing

Cryptococcal isolates (86) were obtained from 57 participants enrolled in the Advancing Cryptococcal meningitis Treatment in Africa (ACTA) trial in Cameroon between 2013 and 2016[21]. 22 participants underwent longitudinal sampling, with samples collected at days 7 and 14, in addition to the baseline day 1 sample. For each timepoint per patient, a single colony was selected for follow up sequencing. Collected samples were stored at -80°C and grown for two days in 3.5 mL of nutrient rich yeast peptone dextrose (YPD) (2% yeast extract, 4% peptone,

4% glucose) media at 30°C and 225 rpm. Genomic DNA was then extracted from a single colony per sample for sequencing with the MasterPure Yeast DNA Purification Kit, as described by Desjardins et al.[6]. DNA was sheared to 250 bp using a Covaris LE instrument and adapted for Illumina sequencing as described by Fisher et al.[22]. Libraries were sequenced on a HiSeq X10, generating 150 bp paired reads.

Data processing and variant calling

To identify genomic variants, reads were aligned to the *Cryptococcus neoformans* H99 reference genome (GCA_000149245.3)[23] with BWA-MEM version 0.7.17[24]. GATK version 4 variant calling [25] was carried out as documented in our publicly available cloud-based pipeline (https://github.com/broadinstitute/fungal-wdl/tree/master/gatk4)[26]. Post calling, variants were filtered on the following parameters: QD < 2.0, QUAL < 30.0, SOR > 3.0, FS > 60.0 (indels > 200), MQ < 40.0, GQ < 50, alternate allele percentage = 0.8, DP < 10. Variants were annotated with SNPeff, version 4.3t[27]. This annotated variant call file (VCF) was used for genome-wide analysis, with further filtering as described below.

Population genomic analysis

A maximum likelihood phylogeny was estimated using the segregating SNP sites present in one or more isolates (VCF sites), allowing ambiguity in a maximum of 10% of samples, with RAxML version 8.2.12 with GTRCAT rapid bootstrapping[28], rooted to VNII isolates as in previous population studies [6,7]. Isolate lineage was identified based on phylogenetic comparison to isolates previously assigned to lineages from Malawi, which was based on a larger phylogenetic analysis with isolates of known sequence types[6,11]. In this phylogeny (Fig. 1), the Malawi isolates grouped into three major clades consistent with their previously assigned lineages (VNI, VNII, and VNB); lineages were assigned to Cameroon isolates based on co-occurrence in these clades. PopGenome (R version 3.5.0, PopGenome version 2.7.5) was used to calculate nucleotide diversity, per chromosome, in 5 kb windows[29].

Genome-wide association studies (GWAS)

Association analysis between isolate origin and variants was carried out using PLINK version 1.08p formatted files and Gemma version 0.94.1[30] (options: centered relatedness matrix gk 1, linear mixed model), as previously described[6]. Rare variants (present in < 5% of the population) were collapsed by gene.

Results

To assess population structure in clinical isolates obtained from patients in Cameroon enrolled in the ACTA clinical trial, which evaluated the efficacy of two new treatment strategies for cryptococcal meningitis when compared to the international standard, whole genome sequence data for 86 baseline and longitudinal isolates from Cameroon were combined with data for 284 isolates from Malawi, also obtained from patients enrolled in the ACTA trial[11,21]. Sequences were aligned to the *Cryptococcus neoformans* (H99) reference genome[23], and variants were called and used to infer a maximum likelihood phylogeny (**Fig. 1a**). Isolates from Cameroon (N=86) were obtained from 57 patients, with 22 of these patients undergoing longitudinal isolate collection at days 7 and 14. Overall, there were 81 *C. neoformans* isolates that belonged to the VNI lineage, 2 that belonged to the VNII lineage, and all *C. neoformans* isolates from Cameroon possessed the MAT alpha locus. Three isolates from Cameroon were excluded from this analysis as they were identified as either AD hybrids (JPC6840 and JPC6884) based on read alignment to all chromosomes of the H99 and JEC21 (GCF000091045.1) genomes or as non-*Cryptococcus* species (JPC6896). Isolates from Cameroon appear highly clonal, with an average terminal branch length across VNI isolates of 0.00070, compared with 0.00085 for VNI isolates (N=266) from Malawi, and a lower average nucleotide diversity across Cameroon VNI isolates of 0.00153, compared to 0.00176 for VNI isolates from Malawi. A single phylogenetic cluster of isolates (**Table 1**, **Fig. 1b**) are a major contributing factor to this clonality, as these isolates obtained from different patients appear highly related to one another, with fewer than 20 SNP differences between isolate pairs in this cluster. There does not appear to be a temporal spike to the clonality between these isolates, as these highly related samples were collected between 2014-2016, however, this might suggest the presence of a single clone within the environment at this time.

Whilst samples in cluster I appear highly related, this cluster includes isolates from both Malawi and Cameroon. We see that the isolates from Cameroon appear well distributed throughout the phylogeny, despite the geographical separation of these two countries, and independent patient population isolate origins. To assess whether there might be distinct genetic markers of isolates from Cameroon or Malawi, we performed a genome-wide association study (GWAS) analysis to assess variants significantly associated with isolates from either country. Significantly associated with isolates from Cameroon is a missense variant in CNAG_00241, a E3 ubiquitin ligase (Gemma score test p=5.70x10-5). This variant is present in 8 Day 1 isolates from Cameroon, in addition to 3 corresponding Day 7 and 3 corresponding Day 14 isolates from

Cameroon, and only 1 isolate from Malawi, and may impact the regulation of protein degradation in these isolates. Significantly associated with isolates from Malawi are two missense variants, one in CNAG_03546, a KOG predicted ATP-dependent DNA helicase, that is present in 55 isolates from Malawi and only 2 isolates from Cameroon (Gemma score test p=2.49x10-12). The other is a missense in CNAG_05511, an ATPase, that is present in 17 isolates from Malawi and 4 isolates from Cameroon (Gemma score test p=3.81x10-6). These variants may impact DNA binding and energy use in these isolates.

In addition to identifying variants associated with isolates from these two countries, we also interrogated variants arising over the course of infection through variant analysis of the longitudinal isolates from Cameroon. We identified 3 variants arising over the course of infection, in 2 patients, that were predicted to impact gene function (**Table 2**). Of these, a conservative in-frame insertion in CNAG_03106, a predicted phosphatidylethanolamine-binding protein, arose over 7 days in one patient. Isolates from another single patient developed two variants, a missense variant in CNAG_02166, a DNA-directed RNA polymerase II subunit *RPB1*, and a stop gain in CNAG_06762 (GAT204). This suggests that genes with roles in the regulation of gene expression, including the transcription factor Gat204 that has been implicated in the capsule-independent inhibition of phagocytosis through synergism with Blp1[31], may be a source of in host microevolution.

To identify potential instances of mixed infection, we calculated SNP differences between isolates obtained longitudinally from the same patient, and identified three sets of longitudinal isolates that appear to be unrelated to their corresponding primary patient isolates (**Table 3**). For one patient, the day 7 isolate had 35,854 SNP differences when compared to the day 1 isolate. In another patient, the day 7 isolate has 22,613 SNP differences when compared to the day 1 isolate, however, the day 14 isolate appears highly related to the day 1 isolate (0 SNP differences), suggesting the original isolate was maintained but not sampled at day 7, potentially the result of a mixed infection. In patient C the day 14 isolate has 44,364 SNP differences when compared to the day 1 isolate, however, the day 7 isolate appears highly similar to the day 1 isolate. These SNP differences identified between longitudinally collected isolates indicate that these may be an instances of mixed strain infections that can be captured through the isolation and sequencing of multiple colonies per patient over the course of an infection.

Discussion

Phylogeographic substructure of fungal populations have been observed for multiple fungal pathogens[32,33]. Here, we compared two cohorts of isolates obtained across the same time period, from different countries and geographic regions. We see isolates from Central Africa are highly clonal, when compared to isolates obtained from southeastern Africa, and this clonality is predominantly driven by a single phylogenetic cluster. Of note, the highly clonal groupings observed consist of isolates from Cameroon and Malawi, suggesting that highly clonal *Cryptococcus neoformans* isolates of the VNI lineage may be present across multiple countries in Africa[34]. We also identify multiple variants significantly associated with either country; given the mixed population structure, these variants are more likely to be linked to country of origin than driven by population sub-structure. Finally, we identify coding variants arising in clinically relevant genes over the course of infection and explore the possibility of multi-strain infections through assessment of longitudinally collected patient samples.

While isolates of the VNI lineage are less genetically diverse than those of the VNB lineage, we observed distinct differences in the levels of nucleotide diversity between VNI isolates from Cameroon and Malawi. Based on measures of population nucleotide diversity, VNI isolates from Malawi appear to be more genetically diverse than those from Cameroon. Previous studies have calculated nucleotide diversity in VNI isolates from a diverse set of countries to be 0.002[6,7], this is similar to the values we see for VNI isolates obtained from Malawi, and to a lesser extent those values for isolates from Cameroon. The presence of highly clonal isolates in Cameroon is not unlike the trends observed in other African countries, for example, isolates from Uganda predominantly belong to a single sequence type (ST93)[9,34,35]. While we observe country linked differences in diversity here, more extensive sampling by region, and from the environment, is needed to further assess how populations of *Cryptococcus* vary by geographic origin.

Missense and loss-of-function variants were identified in samples from patients after 7 days of infection. These variants are predicted to alter the functions of a phosphatidylethanolamine binding protein, a DNA-directed RNA polymerase, GAT204, and IDH3. GAT204 has been implicated in virulence, and was shown to be involved in the evasion of phagocytosis[31]. While variants in these specific genes have not been previously reported, variants in genes with similar functions, including inositol transport and DNA binding, and genes in close proximity to those identified here (CNAG_05330), have been shown to arise over the course of infection[36].

Multiple studies have identified copy number variation and variants arising in virulence implicated genes over the course of infection or relapse[17,18,36,37], suggesting that this might also be a mechanism of adaptation to the host that sampling of longitudinal isolates offers an important and unique insight into.

The study of mixed infections has been somewhat limited by the practice of sequencing a single colony for follow up study per patient, with the assumption that an infection may generally be caused by an isolate of a single genotype. The specific sequencing of multiple samples per patient due to longitudinal sampling offers the unique opportunity to assess within-patient isolate heterogeneity. Previous studies that have set out to assess the frequency of mixed strain infections have reported mixed infection rates of 18-30%, with mixed strains representing different species, serotypes, mating types, and genotypes[19,20]. While we report a lower proportion of likely mixed strain infections (13%), it is anticipated that the sequencing of multiple colonies per patient, per time point, is needed to better capture the heterogeneity of patient infections. Multiple colony sampling per patient also aids in the evaluation of true mixed infections vs potential sample swaps. The likelihood of sample swaps occurring here was deemed to be low, given that the phylogenetic analysis did not reveal evidence for sample swaps between longitudinal isolates.

Through the comparison of isolates across two countries, with well-controlled temporal and collection methods, we can assess population structure, diversity, and geographically linked genetic components. Further studies using long-read sequencing could target the additional types of sequence variants missed by short read sequencing approaches, however the application of such approaches to population studies is still limited by cost considersations. Future studies leveraging data from multi-country clinical trials will allow for a finer scale look into geographical population structure, as well as assessing whether there are specific genotypes associated with region that may be associated with diverse patient outcomes. In addition, studies that include the collection and sequencing of multiple isolates per patient will assess the true state of intra-host pathogen heterogeneity. Together, such studies, linked to clinical metadata, will enable a better understanding of genotype diversity and its contribution to patient outcomes.

Data Availability

Isolate sequence data can be accessed via NCBI with accession number PRJNA1006382 and PRJNA764746.

Acknowledgements

We thank all the patients and their families, as well as the staff at all of the sites not directly involved in the ACTA trial. We thank Marie Desnos-Ollivier for their administrative assistance in transferring the isolates sequenced in this study. We thank the Broad Institute Genomics Platform for generating the sequence data for this study.

Funding

Public Health Service Grants AI73896, AI93257(JRP, JLT). This work was supported by the NIAID grant U19AI110818 to the Broad Institute (CAC).

Disclosure of Conflicts of Interest

No Conflict of interest.

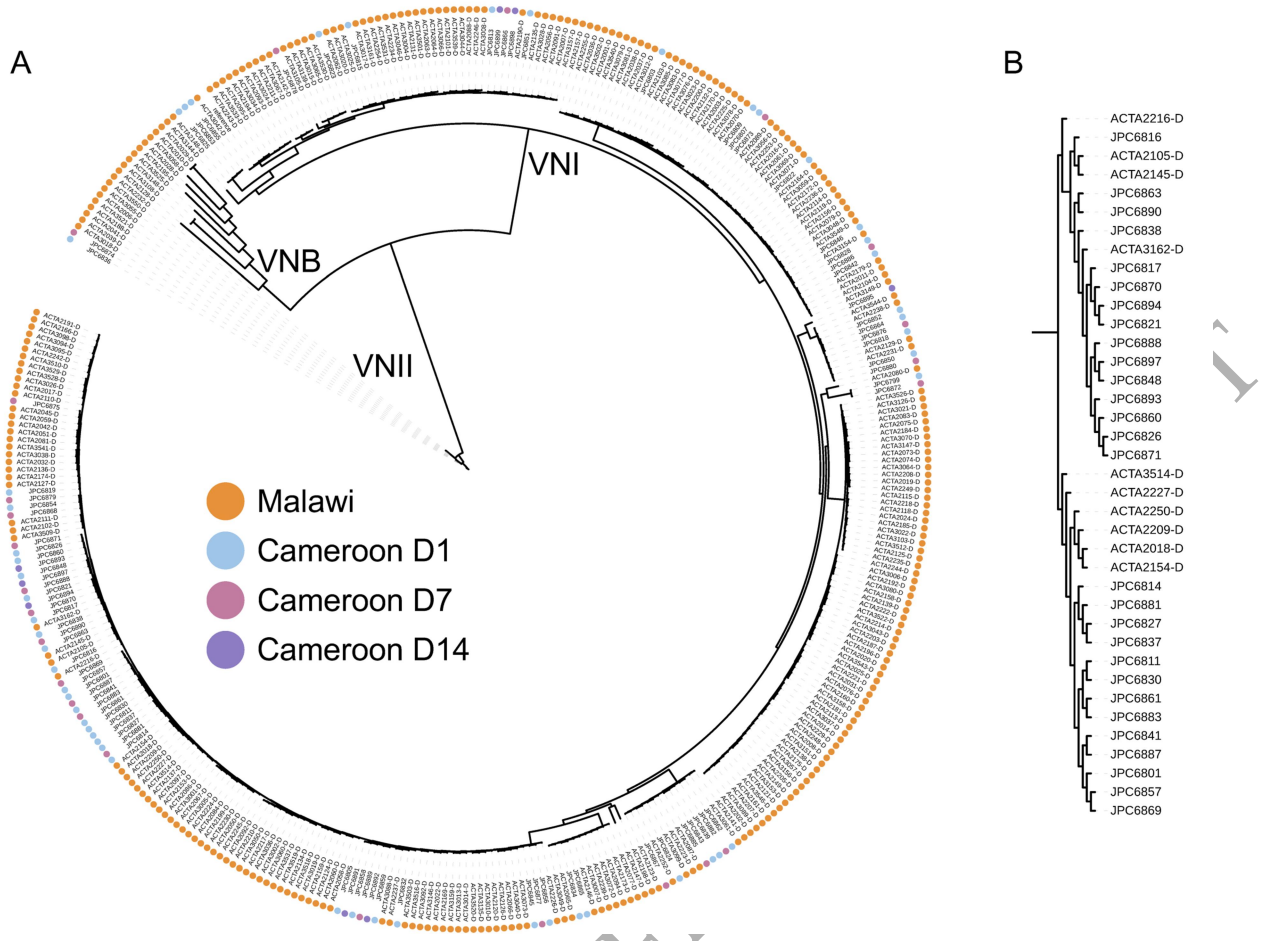

Figure 1. Maximum likelihood phylogeny of isolates from Cameroon and Malawi, rooted to the VNII lineage. A) Phylogeny estimated from segregating SNP sites, with isolates collected on the first day of enrollment from patients in Malawi and Cameroon highlighted in orange and blue, respectively. Isolates collected from patients in Cameroon on day 7 and 14 post enrollment are highlighted in pink and purple, respectively. B) Phylogenetic cluster I containing highly related isolates obtained from different patients.

References

1. Rajasingham R, Govender NP, Jordan A, et al. The global burden of HIV-associated cryptococcal infection in adults in 2020: a modelling analysis. *The Lancet Infectious Diseases*. 2022;22(12):1748-1755. doi:10.1016/S1473-3099(22)00499-6

2. Chayakulkeeree M, Perfect JR. Cryptococcosis. *Infect Dis Clin North Am*. 2006;20(3):507-544, v-vi. doi:10.1016/j.idc.2006.07.001

3. Litvintseva AP, Mitchell TG. Population Genetic Analyses Reveal the African Origin and Strain Variation of *Cryptococcus neoformans* var. *grubii*. *PLoS Pathog*. 2012;8(2):e1002495. doi:10.1371/journal.ppat.1002495

4. Vanhove M, Beale MA, Rhodes J, et al. Genomic epidemiology of *Cryptococcus* yeasts identifies adaptation to environmental niches underpinning infection across an African HIV/AIDS cohort. *Molecular Ecology*. 2017;26(7):1991-2005. doi:https://doi.org/10.1111/mec.13891

5. Litvintseva AP, Thakur R, Vilgalys R, Mitchell TG. Multilocus sequence typing reveals three genetic subpopulations of *Cryptococcus neoformans* var. *grubii* (serotype A), including a unique population in Botswana. *Genetics*. 2006;172(4):2223-2238. doi:10.1534/genetics.105.046672

6. Desjardins CA, Giamberardino C, Sykes SM, et al. Population genomics and the evolution of virulence in the fungal pathogen *Cryptococcus neoformans*. *Genome research*. 2017;27(7):1207-1219. doi:10.1101/gr.218727.116

7. Rhodes J, Desjardins CA, Sykes SM, et al. Tracing Genetic Exchange and Biogeography of *Cryptococcus neoformans* var. *grubii* at the Global Population Level. *Genetics*. 2017;207(1):327-346. doi:10.1534/genetics.117.203836

8. Ergin Ç, Şengül M, Aksoy L, et al. *Cryptococcus neoformans* Recovered From Olive Trees (Olea europaea) in Turkey Reveal Allopatry With African and South American Lineages. *Front Cell Infect Microbiol*. 2019;9:384. doi:10.3389/fcimb.2019.00384

9. Gerstein AC, Jackson KM, McDonald TR, et al. Identification of pathogen genomic differences that impact human immune response and disease during *Cryptococcus neoformans* infection. *mBio*. 2019;10(4). doi:10.1128/mBio.01440-19

10. Ashton PM, Thanh LT, Trieu PH, et al. Three phylogenetic groups have driven the recent population expansion of *Cryptococcus neoformans*. *Nature Communications*. 2019;10(1):12. doi:10.1038/s41467-019-10092-5

11. Sephton-Clark P, Tenor JL, Toffaletti DL, et al. Genomic Variation across a Clinical *Cryptococcus* Population Linked to Disease Outcome. *mBio*. 2022;13(6):e02626-22. doi:10.1128/mbio.02626-22

12. Velagapudi R, Hsueh YP, Geunes-Boyer S, Wright JR, Heitman J. Spores as infectious propagules of *Cryptococcus neoformans*. *Infection and Immunity*. 2009;77(10):4345-4355. doi:10.1128/IAI.00542-09

13. Walsh NM, Botts MR, McDermott AJ, et al. Infectious particle identity determines dissemination and disease outcome for the inhaled human fungal pathogen *Cryptococcus*. *PLoS Pathog*. 2019;15(6):e1007777. doi:10.1371/journal.ppat.1007777

14. Goldman DL, Khine H, Abadi J, et al. Serologic evidence for *Cryptococcus neoformans* infection in early childhood. *Pediatrics*. 2001;107(5):E66. doi:10.1542/peds.107.5.e66

15. Alanio A. Dormancy in *Cryptococcus neoformans*: 60 years of accumulating evidence. *J Clin Invest*. 2020;130(7):3353-3360. doi:10.1172/JCI136223

16. Hagen F, Colom MF, Swinne D, et al. Autochthonous and dormant *Cryptococcus gattii* infections in Europe. *Emerg Infect Dis*. 2012;18(10):1618-1624. doi:10.3201/eid1810.120068

17. Chen Y, Farrer RA, Giamberardino C, et al. Microevolution of serial clinical isolates of *Cryptococcus neoformans* var. *grubii* and *C. gattii*. *mBio*. 2017;8(2). doi:10.1128/mBio.00166-17

18. Rhodes J, Beale MA, Vanhove M, et al. A population genomics approach to assessing the genetic basis of within-host microevolution underlying recurrent cryptococcal meningitis infection. *G3: Genes, Genomes, Genetics*. 2017;7(4):1165-1176. doi:10.1534/g3.116.037499

19. Desnos-Ollivier M, Patel S, Spaulding AR, et al. Mixed infections and In Vivo evolution in the human fungal pathogen *Cryptococcus neoformans*. *mBio*. 2010;1(1):10.1128/mBio.00091-10. doi:10.1128/mBio.00091-10 [doi]

20. Kassi FK, Drakulovski P, Bellet V, et al. *Cryptococcus* genetic diversity and mixed infections in Ivorian HIV patients: A follow up study. *PLOS Neglected Tropical Diseases*. 2019;13(11):e0007812. doi:10.1371/journal.pntd.0007812

21. Molloy SF, Kanyama C, Heyderman RS, et al. Antifungal Combinations for Treatment of Cryptococcal Meningitis in Africa. *N Engl J Med*. 2018;378(11):1004-1017. doi:10.1056/NEJMoa1710922

22. Fisher S, Barry A, Abreu J, et al. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biology*. 2011;12(1):R1. doi:10.1186/gb-2011-12-1-r1

23. Janbon G, Ormerod KL, Paulet D, et al. Analysis of the Genome and Transcriptome of *Cryptococcus neoformans* var. *grubii* Reveals Complex RNA Expression and Microevolution Leading to Virulence Attenuation. *PLoS Genetics*. 2014;10(4). doi:10.1371/journal.pgen.1004261

24. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:13033997 [q-bio]*. Published online May 26, 2013. Accessed January 19, 2021. http://arxiv.org/abs/1303.3997

25. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43:11.10.1-11.10.33. doi:10.1002/0471250953.bi1110s43

26. Martinez-Zurita A, Cuomo CA. Genome-Wide Identification of Variants Associated with Antifungal Drug Resistance. *Methods Mol Biol*. 2023;2658:81-103. doi:10.1007/978-1-0716-3155-3_7

27. Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w ^{1118} ; iso-2; iso-3. *Fly*. 2012;6(2):80-92. doi:10.4161/fly.19695

28. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9):1312-1313. doi:10.1093/bioinformatics/btu033

29. Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol Biol Evol*. 2014;31(7):1929-1936. doi:10.1093/molbev/msu136

30. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*. 2012;44(7):821-824. doi:10.1038/ng.2310

31. Chun CD, Brown JCS, Madhani HD. A major role for capsule-independent phagocytosis-inhibitory mechanisms in mammalian infection by *Cryptococcus neoformans*. *Cell Host Microbe*. 2011;9(3):243-251. doi:10.1016/j.chom.2011.02.003

32. Chow NA, Muñoz JF, Gade L, et al. Tracing the evolutionary history and global expansion of *Candida auris* using population genomic analyses. *mBio*. 2020;11(2). doi:10.1128/mBio.03364-19

33. Sephton-Clark P, Nguyen T, Hoa NT, et al. Impact of pathogen genetics on clinical phenotypes in a population of *Talaromyces marneffei* from Vietnam. *Genetics*. Published online May 25, 2023:iyad100. doi:10.1093/genetics/iyad100

34. Wiesner DL, Moskalenko O, Corcoran JM, et al. Cryptococcal genotype influences immunologic response and human clinical outcome after meningitis. *mBio*. 2012;3(5):e00196-12. doi:10.1128/mBio.00196-12

35. Smith KD, Achan B, Hullsiek KH, et al. Increased Antifungal Drug Resistance in Clinical Isolates of *Cryptococcus neoformans* in Uganda. *Antimicrob Agents Chemother*. 2015;59(12):7197-7204. doi:10.1128/AAC.01299-15

36. Stone NRH, Rhodes J, Fisher MC, et al. Dynamic ploidy changes drive fluconazole resistance in human cryptococcal meningitis. *J Clin Invest*. 2019;129(3):999-1014. doi:10.1172/JCI124516

37. Ormerod KL, Morrow CA, Chow EWL, et al. Comparative genomics of serial isolates of *Cryptococcus neoformans* reveals gene associated with carbon utilization and virulence. *G3: Genes, Genomes, Genetics*. 2013;3(4):675-686. doi:10.1534/g3.113.005660

Sample 1	Sample 2	SNP Differences	Lineage
JPC6866	JPC6898	6	VNI
JPC6888	**JPC6817**	8	VNI
JPC6827	**JPC6881**	9	VNI
JPC6870	**JPC6888**	9	VNI
JPC6888	**JPC6893**	9	VNI
JPC6821	**JPC6888**	10	VNI
JPC6826	**JPC6888**	10	VNI
JPC6860	**JPC6888**	10	VNI
JPC6871	**JPC6888**	11	VNI
JPC6888	**JPC6894**	11	VNI
JPC6889	JPC6891	11	VNI
JPC6898	JPC6899	11	VNI
JPC6858	JPC6889	12	VNI
JPC6826	**JPC6860**	19	VNI
JPC6826	**JPC6871**	19	VNI

Table 1. Highly related non-longitudinal isolates. Isolates from the same phylogenetic cluster (cluster I) are bolded.

Sample Pairs	Sample Days	Gene	Gene Function	Variant Type
JPC6837/JPC6881	1/7	CNAG_05329	IDH3, myo-inositol 2-dehydrogenase. Likely non-function due to position (385/389 AA)	Disruptive inframe deletion
JPC6837/JPC6881	1/7	CNAG_03106	Phosphatidylethanolamine-binding protein	Conservative inframe insertion
JPC6862/JPC6882	1/7	CNAG_02166	DNA-directed RNA polymerase II subunit RPB1	Missense variant
JPC6862/JPC6882	1/7	CNAG_06762	GAT204 Transcription factor	Stop gained

Table 2. Missense and loss-of-function variants arising in longitudinal isolates.

Sample	Patient	Day Comparison	SNP Differences
JPC6827	A	1 vs 14	35,854
JPC6858	B	1 vs 7	22,613
JPC6878	B	7 vs 14	22,602
JPC6891	B	1 vs 14	0
JPC6857	C	1 vs 14	44,364
JPC6869	C	7 vs 14	44,712
JPC6898	C	1 vs 7	0

Table 3. Longitudinal isolates with high numbers of within-patient SNP differences.