MYCOBIOLOGY
2023, VOL. 51, NO. 6, 452–462
https://doi.org/10.1080/12298093.2023.2272380

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Determining Potential Link between Environmental and Clinical Isolates of *Cryptococcus neoformans/Cryptococcus gattii* Species Complexes Using Phenotypic and Genotypic Characterisation

Kenosi Kebabonye[a], Mosimanegape Jongman[b,c], Daniel Loeto[b], Sikhulile Moyo[a,c,d,e,f], Wonderful Choga[c] and Ishmael Kasvosve[a]

[a]School of Health Allied Professions, Faculty of Health Sciences, University of Botswana, Gaborone, Botswana; [b]Department of Biological Sciences, Faculty of Science, University of Botswana, Gaborone, Botswana; [c]Research Laboratory, Botswana Harvard AIDS Institute Partnership, Gaborone, Botswana; [d]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [e]Division of Medical Virology, Stellenbosch University, Cape Town, South Africa; [f]School of Health Systems of Public Health, University of Pretoria, Pretoria, South Africa

ABSTRACT

Opportunistic infections due to *Cryptococcus neoformans* and *C. gattii* species complexes continue to rise unabated among HIV/AIDS patients, despite improved antifungal therapies. Here, we collected a total of 20 environmental and 25 presumptive clinical cryptococcal isolates from cerebrospinal fluid (CSF) samples of 175 patients enrolled in an ongoing clinical trial Ambition 1 Project (Botswana-Harvard Partnership). Identity confirmation of the isolates was done using MALDI-TOF MS and PCR. We describe the diversity of the isolates by PCR fingerprinting and sequencing (Oxford Nanopore Technology) of the intergenic spacer region. Mating types of the isolates were determined by amplification of the *MAT* locus. We report an unusual prevalence of 42.1% of *C. neoformans* x *C. deneoformans* hybrids Serotype AD ($n = 16$), followed by 39.5% of *C. neoformans* Serotype A ($n = 15$), 5.3% of *C. deneoformans*, *Serotype* D ($n = 2$), 7.9% of *C. gattii* ($n = 3$), and 5.3% of *C. tetragattii* ($n = 2$) in 38 representative isolates that have been characterized. Mating type-specific PCR performed on 38 representative environmental and clinical isolates revealed that 16 (42.1%) were *MAT**a**/MAT*α hybrids, 17 (44.7%) were *MAT*α, and five (13.2%) possessed *MAT**a*** mating type. We used conventional and NGS platforms to demonstrate a potential link between environmental and clinical isolates and lay a foundation to further describe mating patterns/history in Botswana.

ARTICLE HISTORY
Received 16 February 2023
Revised 10 October 2023
Accepted 10 October 2023

KEYWORDS
Cryptococcus gattii; *Cryptococcus neoformans*; genetic diversity; intergenic spacer region; Oxford nanopore sequencing

1. Introduction

Cryptococcosis is a life-threatening opportunistic infection that can kill adults living with HIV in Sub-Saharan Africa. It is estimated that about 130,000 individuals die annually due to cryptococcal meningitis in this region [1]. Human cryptococcal infection is attributed to the *Cryptococcus neoformans/C. gattii* (*CnCg*) species complexes which have been divided into two sibling species namely; *C. neoformans* and *C. gattii* [2]. Both yeasts are well-known human pathogens, have broad geographic distributions and can grow in diverse environments such as soil, pigeon droppings and tree sources that include mopane trees (*Colophospermum mopane*) [3,4]. Pulmonary infection typically occur through inhalation of infectious propagules or desiccated yeasts spores from the environmental sources establishing a latent, asymptomatic or minimal symptomatic disease [5].

Initial studies predominantly associated occurrence of *C. gattii* species complex with tropical and subtropical regions, nonetheless, the spread of this yeast into temperate climate regions has been reported too [6]. Remarkably, varying strains belonging to *C. gattii* species complex have emerged globally as a primary pathogen and cause of infection in both immunocompetent hosts and those with a compromised immune system [3]. Unlike the sibling species, *C. neoformans* is diverse and commonly cause disease in immunocompromised individuals [3]. Findings on autopsy analysis revealed that the central nervous system and the lungs frequently get affected by disseminated infections [7] which may also affect other various organs especially in HIV-infected patients [8]. In recent molecular phylogenetic studies, significant diversity and genetic divergence within the *CnCg* species complexes have been identified. A study by Hagen et al. proposed that these species complexes be divided into seven sibling species, namely, *C. neoformans*, *C. deneoformans*, and

CONTACT Jongman Mosimanegape jongmanm@ub.ac.bw

C. neoformans × *C. deneoformans* hybrids (for *C. neoformans* species complex), and *C. gattii*, *C. deuterogattii*, *C. bacillisporus*, *C. decagattii* and *C. tetragattii* (for *C. gattii* species complex) [9]. In addition, a study by Farrer et al. [10] identified a new lineage of the *C. gattii* species complex designated *C. decagattii* from isolates collected from a Hyrax in Zambia. However, in light of the revised taxonomy, this new lineage can be regarded as a separate species upon formal description by the authors.

Traditionally, several techniques such as staining with India ink and antigen testing have been employed in primary detection of these species [11]. These techniques are relatively fast but cannot distinguish the pathogenic yeasts [12]. Detection methods such as matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), and next generation sequencing (NGS) technology have been used for the identification of the species with capacity to distinguish between all affiliated genotypes [13]. The International Society for Human and Animal Mycology (ISHAM) accepted multi-locus sequence typing (MLST) as a gold standard for genotyping isolates of *CnCg* [14]. This technology uses a consensus set of seven to eleven primers to determine the global genotypes of these pathogenic yeasts [14].

Botswana, located at the epicenter of the cryptococcosis and HIV/AIDS epidemic, has limited resources that make sequencing multiple loci too expensive. This is especially concerning as 1 in 4 adults in Botswana are HIV positive. In addition, strains of the CnCg species complexes found in Botswana are genetically diverse when compared to global strains. Therefore, it is crucial to continuously monitor and track the source of these pathogenic yeasts, particularly in the management and prevention of cryptococcal disease in HIV/AIDS patients. To address this, we utilized Oxford Nanopore Technology (ONT) to obtain complete intergenic spacer region (IGS) sequences from clinical and environmental isolates. Our analysis revealed globally prevalent species of *C. neoformans*, *C. deneoformans*, *C. neoformans* x *C. deneoformans hybrids*, *C. gattii*, *C. bacillisporus*, *C. deuterogattii* and *C. tetragattii*.

We utilized MALDI-TOF MS to quickly and accurately identify the isolates up to the species level. This allowed us to create a collection of protein masses from clinical and environmental isolates in Botswana. Our analysis of mating types showed a higher-than-usual occurrence of hybrids, and we used M13 microsatellites to produce genetic fingerprints. This helped us understand the genetic makeup and potential connections between clinical and environmental isolates gathered in the central region of Botswana.

2. Materials and methods

2.1. Study site and sample collection

To investigate the occurrence of *CnCg* species, sampling was conducted at ten areas (urban/rural) in the central district of Botswana (Figure 1). In total, 1500 samples were collected between August and December 2020 from hollows of mopane trees within the vicinity of human settlements across the sampling sites. Sterile transport swabs containing Amies medium without charcoal (LabocareTM, Johannesburg, South Africa) were used to swab probable areas of cryptococcal habitation e.g., tree hollows. In addition, a total of 25 presumptive clinical cryptococcal isolates were recovered from cerebrospinal fluid (CSF) samples of 175 patients enrolled on an ongoing clinical trial Ambition 1 Project (Botswana-Harvard Partnership). Formal consent for environmental and clinical isolates and patient information followed institutional guidelines of the University of Botswana, and a research permit was obtained from Ministry of Health and Wellness (Permit number: HPDME 13/18/1).

2.2. Isolation and identification of the CnCg species complexes

For primary isolation of *Cryptococcus* species from clinical and environmental samples, Staib media [15] was modified by replacing *Guizotia abyssinica* seeds with mopane seeds. Briefly, reconstituted media containing pulverized mopane seeds (200 g) were suspended in distilled water (1 L) and allowed to boil for 30 min. Solid seed material was removed and distilled water was added accordingly to obtain 1 L of mopane seed extract solution [16]. Thereafter, the solution was allowed to cool prior to adding dextrose (Sigma Aldrich) (1 g) and agar (Sigma-Aldrich) (20 g), followed by autoclaving of the medium for 30 min. Swabs were thoroughly rubbed on the plates and incubated at 25 °C for 72 hrs. Presumptive brown yeast colonies that formed were selected, grown in pure culture on mopane reconstituted agar plates without antibiotics, confirmed to be *CnCg* species by standard morphological and physiological criteria, and maintained on yeast extract-peptone-dextrose (YPD) agar (Difco, Baltimore, MD) at 30 °C.

2.3. Identity confirmation, mating type determination and minisatellite fingerprinting

Following Matrix-Assisted Laser Desorption/Ionization–Time-of-Flight Mass Spectrometry (MALDI-TOF MS) (Bruker Daltonics, Inc., Germany) confirmation, the MALDI Biotyper

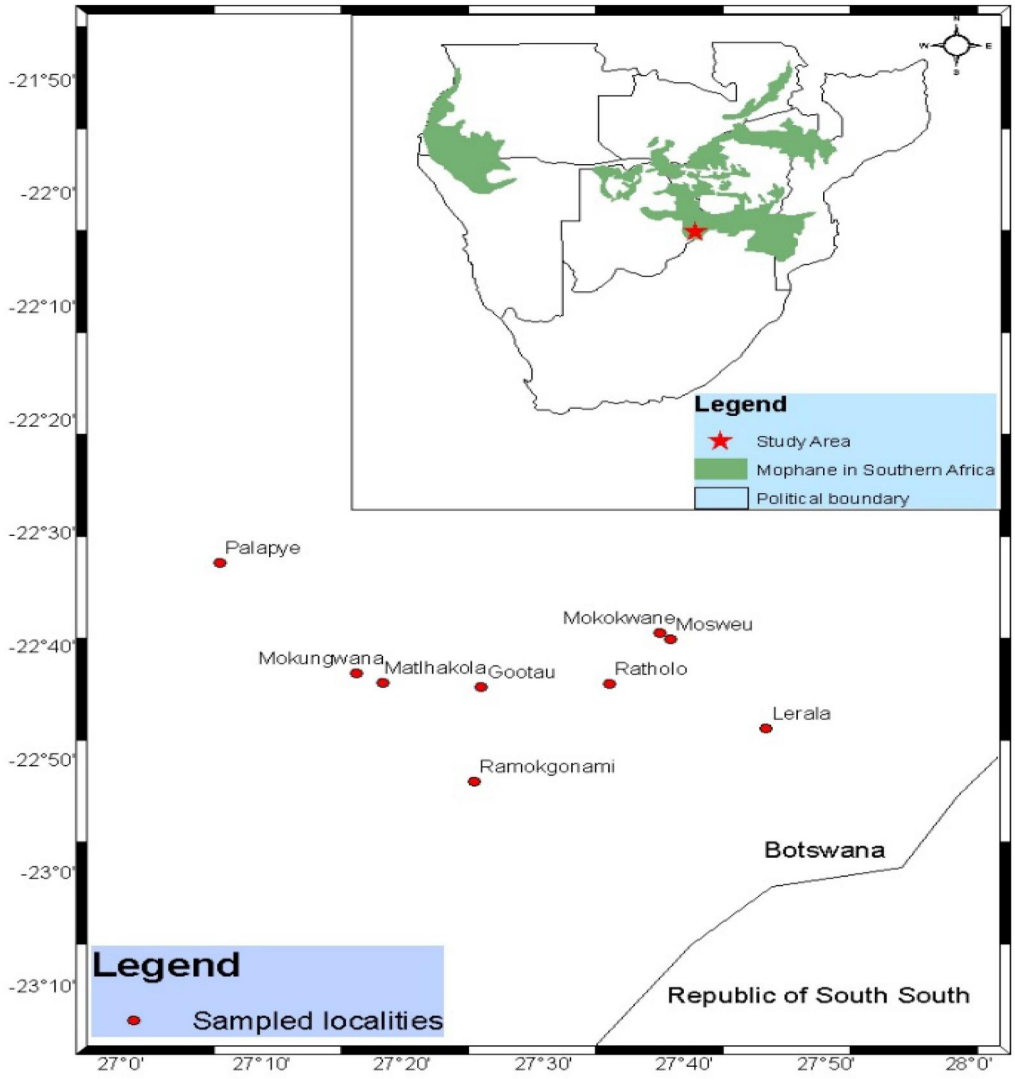

Figure 1. Map showing the geographic distribution of localities where environmental samples were collected in the central region of Botswana.

3.0 SOFTWARE (Bruker Daltonics, Germany) was used to visualize the mass spectra and, generate dendrograms using seuclidean as distance measure. For all confirmed isolates, genomic DNA was extracted using manufactures' instructions, followed by DNA quantification (Thermo Fisher Scientific, Inc). Thereafter, PCR was performed using specific primers for *CgIGS* and *CnIGS* genes [5], mating type determination using *STE12α* and *STE20a* [17] and genotyping using the minisatellite-specific core of the wild-type M13 phage [18]. All visible fingerprint bands were included in the analysis despite their intensity. Visual PCR fingerprints were compared with those of the reference standards amplified in parallel for genotypic characterization of each isolate. Genetic relatedness of the strains was conducted by analyses of the electrophoretic banding patterns using BIONUMERICS VERSION 8.0 SOFTWARE (Applied Math, Kortrijk, Belgium). Bands tolerance positions were defined manually at 0.8%. Similarity coefficients were calculated using the Dice algorithm, and cluster analyses were performed by the Unweighted Pair Group Method for Arithmetic averages (UPGMA) [17].

2.4. Amplification of the IGS region and sequencing

The IGS amplicons to be sequenced were prepared using 2 µl of extracted DNA as PCR template. The final PCR mixture consisted of OneTaq quick load (New England Biolabs (NEB)), 2X MM with standard buffer M0486S, genomic DNA and 30 ppmol of both forward and reverse primers (Table 1) [19]. The PCR conditions were as follows: 95 °C for 5 min, followed by 35 cycles of 94 °C for 45 s, 63 °C for 1 min and 72 °C for 1 min. After confirming the targeted size of amplicons using gel electrophoresis, amplicons were purified using QIAquick PCR purification kit (Qiagen, Hilden, Germany) according to manufacturer's instructions.

Oxford nanopore sequencing libraries were prepared by multiplexing the DNA from 29 isolates per flow cell using the Rapid Barcoding Sequencing kit

Table 1. A and B show total number of environmental and clinical samples collected, as well as *Cryptococcus neoformans* and *C. gattii* isolates acquired from central region of Botswana.

A	Positive sample		
Location	*C. gattii* (%)	*C. neoformans* (%)	Number of trees swabbed
Palapye	0	4 (2.29)	175
Makoro	0	2 (1.6)	125
Mokungwane	0	2 (1.38)	145
Matlhakola	0	2 (1.29)	155
Gootau	0	2 (1.33)	150
Ratholo	0	3 (2)	150
Majwaneng	0	1 (0.74)	135
Lerala	0	2 (1.21)	165
Mokokwana	0	1 (0.83)	120
Mosweu	0	1 (0.56)	180
Total	**0**	**20 (1.33)**	**1500**

B	Positive sample		
Location	*C. gattii* (%)	*C. neoformans* (%)	Number of patients enrolled
P/Marina Hospital	5 (2.85)	20 (11.4)	175
Total	**5 (2.85)**	**20 (11.4)**	**175**

(SQK-RBK110.96; ONT, Oxford, UK), following protocol version MRT_9127_v110_revH_14Jul2021, as per the instructions of the manufacturer. Sequencing libraries were loaded onto a FLO-MIN106 R9.4.1 SpotON flow cell and sequenced in the GridION X5 Mk1 sequencing device. Primary acquisition of data and real-time basecalling was carried out using the graphical user interface MinKNOW v2.0 and Guppy basecaller v3.0.6 (both from ONT). The demultiplexing of barcodes and quality control of the reads were also accomplished in real-time using EPI2ME platform. All quality reads (quality score above 7) were extracted after 20 h of the sequencing run for downstream analysis.

2.5. Bioinformatics and identification of the isolates

An open source, cloud-based metagenomics pipeline CZID was used for identification of the isolates from raw sequencing reads [20]. In this pipeline, raw sequencing files are uploaded to CZID (https://idseq.net), which performs several processing steps prior to analysis of non-host data. The first step in the pipeline is an alignment to the human genome to remove host sequences. The subsequent dataset contained sequences from various pathogens including *CnCg* species. The subsequent steps in the CZID pipeline (i.e., removal of low-quality reads, duplicate reads, low complexity reads, and additional human sequence filtering) further reduce contaminating and uninformative sequences from the nonhuman dataset. The remaining sequences were assembled using an indexed version of the NCBI's GenBank database to identify the source of nonhuman sequences in the datasets.

Geneious Assembled contigs were queried against the NCBI nucleotide database using BLAST to identify the closest related taxa. GenBank annotation files from genome sequence records corresponding to the highest scoring alignments were used to identify potential features within the *de novo* assembled genomes. Geneious v10.3.2 was used to annotate newly assembled genomes. For multiple sequence alignments and phylogenetic analyses, we downloaded reference strains which were used in a study performed by Morrison et al. [21] from NCBI, and consist of JEC21 (*C. deneoformans*), CC-100-5 (AD hybrid), H99 (*C. neoformans*), WM1850 (*C. deuterogattii*), WM178 (*C. deuterogattii*), WM-11.139 (*C. bacillisporus*), and WM-179 (*C. gattii*) [21]. Multiple sequence alignments were generated using ClustalW in MEGA v6.0 and the Geneious aligner in Geneious v10.3.2.

Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 11 [22]. The *CnCg* species complexes and reference sequences were aligned using Mega 11. A maximum-likelihood tree topology was developed using following parameters: a General Time Reversible model of nucleotide substitution and a total of 500 bootstrap replicates [23]. The resulting maximum-likelihood tree topology was transformed into a time-calibrated phylogeny in which branches along the tree were scaled in calendar time using TreeTime [23]. The resulting tree was then visualized and annotated in ggtree in R [23].

2.6. Statistical analysis

Statistical analysis including geometric means, standard deviations, ranges, and medians of the recovered isolates were performed using SAS VERSION 9.4 (SAS Institute Inc, Cary, NC, USA, SAS, 2010). Analysis of variance (ANOVA) was performed and least significance difference (LSD) test was used to derive statistical differences ($P < 0.05$) of the data.

3. Results

3.1. Recovery of Cryptococcus species from environmental and clinical settings

We sampled from ten areas in the central region of Botswana along the *Co. mopane* ecoregion: Palapye ($n = 175$), Makoro ($n = 125$), Mokungwane ($n = 145$), Matlhakola ($n = 155$), Gootau ($n = 150$), Ratholo ($n = 150$), Majwaneng ($n = 135$), Lerala ($n = 165$), Mokokwana ($n = 120$), Mosweu ($n = 180$). A total of 20 and 25 presumptive isolates of the *CnCg* species complexes were recovered from environmental (Table 1A) and clinical (Table 1B)

Table 2. Phenotypic and genotypic characterization of clinical and environmental isolates of the *Cryptococcus neoforman/C. gattii* species complexes.

Isolate ID	Bruker ID (Score)	Pigmentation	Eco region	*CnCg* species	Mating type
Cl 007	*C. gattii* (1.848)	Dark brown	P/Marina	*C. deuterogattii*	*MATα*
Env 137	*C. neoformans* (2.050)	Dark brown	Palapye	*C. neoformans*	***MATa***
Env 1001	*C. neoformans* (1.957)	Dark brown	Palapye	*C. neoformans*	*MATa/MATα*
Cl 001	*C. neoformans* (1.929)	Dark brown	P/marina	*C. neoformans*	*MATα*
Env 850	*C. neoformans* (1.936)	Dark brown	Lerala	*C. neoformans*	***MATa***
Cl 0921	*C. neoformans* (2.046)	Dark brown	P/Marina	*C. neoformans*	*MATa/MATα*
Env 202	*C. neoformans* (2.006)	Dark brown	Makoro	*C. neoformans*	*MATa/MATα*
Cl 317	*C. gattii* (1.916)	Dark brown	P/Marina	*C. deuterogattii*	*MATα*
Env 02/25	*C. neoformans* (2.050)	Dark brown	Mokungwane	*C. neoformans*	*MATa/MATα*
Cl 003	*C. neoformans* (1.976)	Dark brown	P/Marina	*C. neoformans*	*MATa/MATα*
Cl 173	*C. gattii* (1.943)	Dark brown	P/Marina	*C. gattii*	*MATα*
Cl 950	*C. gattii* (1.864)	Dark brown	P/Marina	*C. gattii*	*MATα*
Env 030	*C. neoformans* (1.872)	Dark brown	Gootau	*C. neoformans*	*MATa/MATα*
Cl 02/20	*C. neoformans* (1.865)	Dark brown	P/Marina	*C. neoformans*	*MATα*
Cl 0700	*C. neoformans* (1.929)	Dark brown	P/Marina	*C. neoformans*	*MATα*
Env 010	*C. neoformans* (1.979)	Dark brown	Matlhakola	*C. neoformans*	*MATa/MATα*
Cl 011	*C. neoformans* (2.106)	Dark brown	P/Marina	*C. neoformans*	*MATα*
Cl 01/024	*C. neoformans* (2.017)	Dark brown	P/Marina	*C. neoformans*	*MATa/MATα*
Env 820	*C. neoformans* (1.776)	Dark brown	Ratholo	*C. neoformans*	*MATa/MATα*
Env 02/023	*C. neoformans* (2.039)	Dark brown	Ratholo	*C. neoformans*	*MATa/MATα*
Cl 018	*C. neoformans* (2.004)	Dark brown	P/Marina	*C. neoformans*	*MATa/MATα*
Cl 013	*C. neoformans* (2.075)	Dark brown	P/Marina	*C. neoformans*	*MATα*
Cl 0921	*C. neoformans* (1.888)	Dark brown	P/Marina	*C. neoformans*	*MATa/MATα*
Env 09	*C. neoformans* (1.927)	Dark brown	Ratholo	*C. neoformans*	***MATa***
Env 020	*C. neoformans* (1.838)	Dark brown	Mokokwana	AD Hybrid	*MATa/MATα*
Cl 023	*C. neoformans* (2.094)	Dark brown	P/Marina	AD Hybrid	*MATa/MATα*
Env 093	*C. neoformans* (2.020)	Dark brown	Majwaneng	AD Hybrid	*MATa/MATα*
Cl 042	*C. neoformans* (1.863)	Dark brown	P/Marina	AD Hybrid	*MATa/MATα*
Cl 018	*C. gattii* (1.921)	Dark brown	P/Marina	*C. gattii*	*MATα*
Cl 010	*C. neoformans* (2.001)	Dark brown	P/Marina	*C. neoformans*	*MATα*
Env 950	*C. neoformans* (2.123)	Dark brown	Mosweu	*C. neoformans*	***MATa***
Cl 0061	*C. neoformans* (1.890)	Dark brown	P/Marina	*C. neoformans*	*MATα*
Env 02/25	*C. neoformans* (2.012)	Dark brown	Makoro	*C. neoformans*	***MATa***
Cl 0122	*C. neoformans* (1.923)	Dark brown	P/Marina	*C. neoformans*	*MATα*
Cl 015	*C. neoformans* (2.003)	Dark brown	P/Marina	*C. neoformans*	*MATα*
Cl 0070	*C. neoformans* (1.912)	Dark brown	P/Marina	*C. neoformans*	*MATα*
Cl 260	*C. neoformans* (1.986)	Dark brown	P/Marina	*C. neoformans*	*MATα*
Cl 189	*C. neoformans* (2.001)	Dark brown	P/Marina	*C. neoformans*	*MATα*

**ND* asterisks indicate that the test has not been performed on the isolate; P/Marina = Princess Marina Hospital.

samples, respectively. Recovery rate of the *CnCg* species complexes was 1.33% in environmental samples and 14.3% in clinical samples. The highest prevalence of environmental *C. neoformans* species complex was observed in Palapye 2.28%). There were no positive outcomes for *C. gattii* species complex in our environmental samples, but this yeast was recovered in clinical samples. The majority of the isolates 80% recovered from clinical samples belonged to *C. neoformans* species complex.

3.2. Identification of the CnCg species complexes

Using MALDI-TOF analysis, representative 38 (100%) of the presumptive isolates of *Cryptococcus* species were correctly identified to the species level as *C. neoformans* (33) and *C. gattii* (5) with identification scores ranging between 1.7 and 2.0 [9]. The resulting dendrograms for *CnCg* isolates clustered corresponding to the five molecular types which are prevalent in Southern Africa. Amplification and presence of the IGS region (*CgIGS1* and *CnIGS1*) of the rDNA subsequently further affirmed identities to the species level [19].

3.3. Molecular genotyping, genetic diversity and mating type

The minisatellite-specific primer M13 and Oxford nanopore sequencing revealed intra-species variability among the isolates. The cryptococal isolates were successfully grouped into the major molecular types of C. *neoformans, C. deneoformans, C. neoformans x C. deneoformans* hybrids (AD hybrids), *C. gattii* and *C. tetragattii*. (Table 2). The highest prevalence of *C. neoformans* species complex was observed in clinical samples (22/38) compared to environmental samples (11/38). Except for isolates from Palapye ($n = 4$) and Ratholo ($n = 3$), we observed a relatively equal distribution of *C. neoformans* (≤ 2 isolates) across all other sites (Table 2). All isolates belonging to the *C. gattii* species complex were acquired from the clinical isolates.

Mating type specific PCRs performed on the environmental and clinical isolates revealed mating type **a** (*MAT**a***), mating type α (*MATα*), and *MAT**a**/MATα* hybrids. Among all the isolates, 16 (42.1%) were *MAT**a**/MATα* hybrids, 17 (44.7%) were *MAT α*, and five (13.2%) possessed *MAT**a*** mating types. Using singleplex PCR, all isolates belonging to *C. gattii* and *C. tetragattii* were *MATα*. The isolates

belonging to *C. neoformans* comprised four environmental isolates which possessed *MAT**a*** mating type, whereas 11 clinical *C. neoformans* isolates belonged to *MATα* mating types. There were 16 *MAT**a**/MATα* hybrids, from which nine were environmental and seven were clinical. With respect to *C. deneoformans*, two isolates, of which one was environmental with *MAT**a*** mating type and the other one was *MATα* clinical isolate were identified.

Oxford Nanopore Sequencing platform identified the *CnCg* species complexes specific reads in 38/38 (100%) clinical and environmental samples. The reads of the *C. neoformans* species complex were detected in 38 samples, of which 35 isolates were also positive for *C. gattii*, 34 isolates were positive for *C. deuterogattii*, 33 for *C. bacillisporus* and 15 isolates of *C. tetragattii*. The reads of *C. tetragattii* were not detected in environmental isolates (Figure 2A and B). In none of the samples were the *C. neoformans* species complex reads recorded in the absence of *C. gattii* species complex reads. All identified *C. neoformans* species complex reads were attributed to *C. neoformans, C. deneoformans* and *C. neoformans x C. deneoformans* hybrids. The molecular genotypes of the *CnCg* species complexes resolved well into subclades of strains with the same genotypes as well as species designation for each genotypes as per identification using CZID platform (Figure 2A and B).

Phylogenetic analysis of the clinical and environmental datasets using maximum likelihood method revealed three distinct clusters within the *C. neoformans* species complex. Molecular groups of *C. neoformans, C. deneoformans* and *C. neoformans x C. deneoformans* hybrids clustered separately in different groups, and *C. neoformans* H99 was used as a reference strain (Figure 3A). Cluster I of the *C. deneoformans* population includes five different strains all of which were clinical isolates. Cluster II is composed of six different strains of *C. neoformans x C. deneoformans* hybrids with clinical isolates ($n = 4$) constituting the majority of the group. Cluster III is composed of eighteen strains of *C. neoformans* which constitute large number of the group. Majority of the isolates within this group came from the environmental samples ($n = 12$) followed by clinical isolates ($n = 6$). In all the cluster groups, isolates of both clinical and environmental isolates clustered randomly amongst themselves.

With respect to phylogenetic analysis of *C. gattii* species complex, Maximum-likelihood of the strains revealed three clusters (Figure 3B). Cluster I, Cluster II and Cluster III constitute 13 strains belonging to *C. deuterogattii*, 10 strains of *C. bacillisporus* species, and 15 strains belonging to *C. gattii* molecular groups, respectively. Majority of the isolates were recovered from clinical samples.

4. Discussions

Cryptococcosis is an opportunistic infection in HIV/AIDS patients in sub-Saharan Africa with reported annual deaths of 75% in infected people [1]. Despite achievements of the UNAIDS 95-95-95 targets, cryptococcosis remains a significant challenge for HIV positive individuals in Botswana [24]. Thus, continued surveillance of the *CnCg* species complexes and rapid treatment of cryptococcosis is necessary. In fact, recent findings suggest that effective HIV associated cryptococcosis antifungal treatment even without antiretroviral therapy can reverse CSF HIV escape [24].

Previous findings reported genetically unique populations of *Cryptococcus* species that are primarily associated with trees in Southern Africa [16,25]. Here, we provide a comprehensive analysis of clinical and environmental isolates to demonstrate a potential link between these populations. Interestingly, our study observed a possible disturbing emergence of *C. neoformans x C. deneoformans* hybrids (AD hybrids) among clinical and environmental isolates from the central region of Botswana. The reasons for this unusual dominance of hybrid genotypes among natural populations in Botswana are unknown, but call for urgent and extensive investigations.

The results of our environmental sampling confirmed previous studies regarding the high prevalence of *C. neoformans* on mopane trees [16,26]. Mopane trees located in and around households were sampled to determine the possibility of local acquisition of cryptococcal infections. More significantly, our results indicated close similarities between genotypic profiles of *C. neoformans* recovered from clinical and mopane trees at homesteads. These results heighten the potential risk posed to people living with HIV/AIDS who are exposed to mopane trees or allied products. Forest communities visited herein are highly dominated by *Co. mopane* trees, a culturally and economically important tree species used for medicinal and construction purposes [27]. Mopane tree is also the sole substrate for mopane worms which forms part of diet and culinary delicacy across Southern African countries [28] that are heavily burdened by HIV/AIDS.

Population genetic analysis on clinical and environmental isolates determined significant ($p < 0.05$) diversity among the isolates by revealing the occurrence of molecular genotypes *C. neoformans, C. deneoformans*, as well as *C. neoformans X C. deneoformans* hybrids. The *C. gattii* species complex

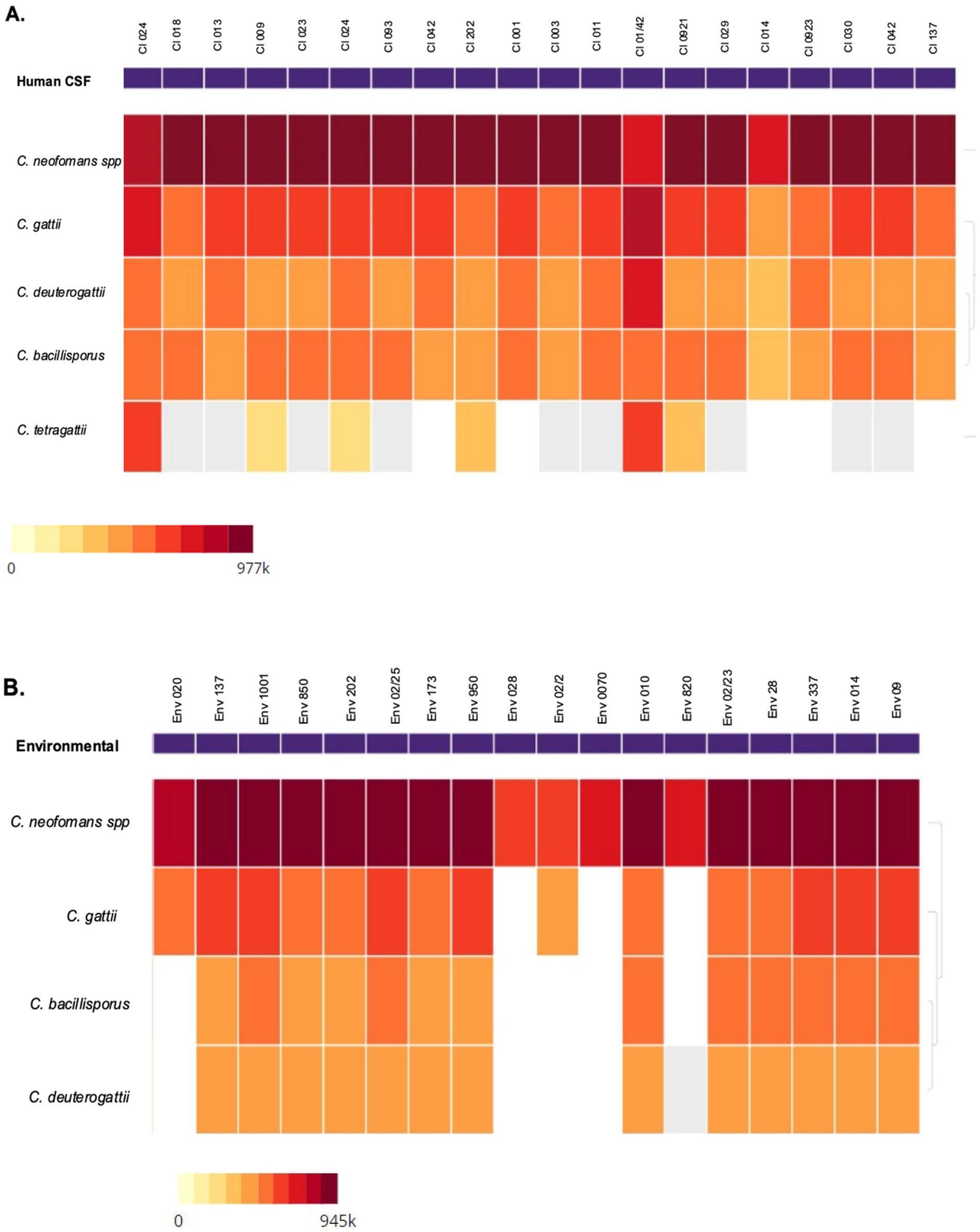

Figure 2. Heat mapping tree indicating abundance of clinical (A) and environmental (B) isolates of the *Cryptococcus neoforman/C. gattii* species complexes. Dark colors indicate more abundance while light colors indicate more abundance while light colors indicate less abundance of taxa.

identified by ONT sequencing in the clinical and environmental settings were attributed to *C. gattii, C. deuterogattii,* and *C. bacillisporus*, while *C. tetragattii* was observed only in isolates acquired from CSF. Of the total isolates sequenced in this study, only 18, five, and six were grouped as *C. neoformans*, *C. deneoformans*, and *C. neoformans* x *C. deneoformans* hybrids, respectively. The ISHAM community reached a consensus that MLST is the gold standard method used in genotyping of *Cryptococcus* species, but this platform requires amplification and sequencing of seven different loci, making analysis of number of isolates very expensive, for resource-limited settings such as Botswana. Hence, sequencing of the IGS region using the Oxford nanopore technology provided an affordable

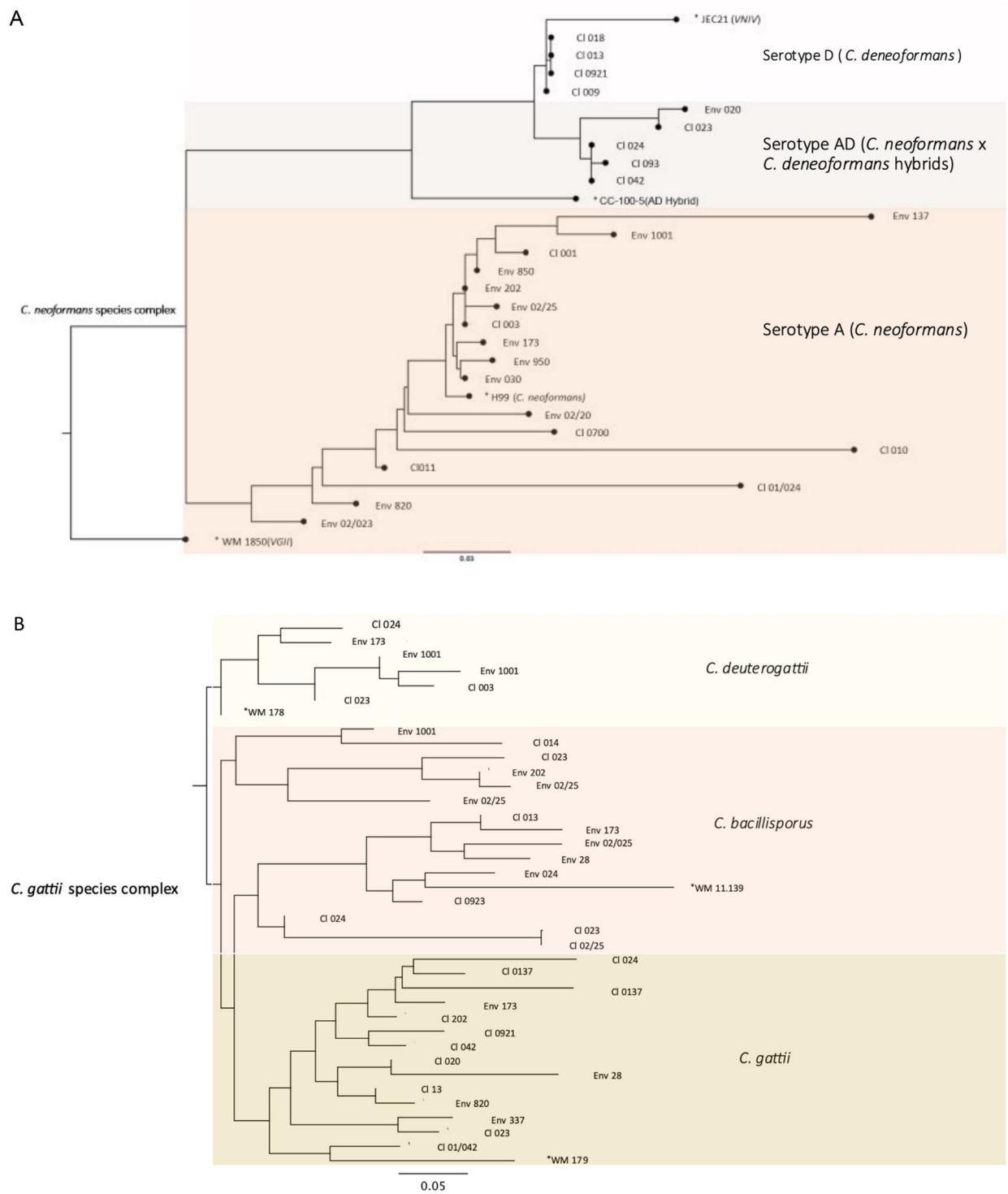

Figure 3. Phylogenetic relationships of the *Cryptococcus neoformans* (A) and *C. gattii* (B) isolates inferred from maximum likelihood analysis of IGS sequence using the Oxford nanopore sequencing. *asterisks represent the reference strains, env = environmental isolates, Cl = clinical isolates.

alternative of rapid diagnosis and profiling of these pathogenic yeasts [21,29,30].

Interestingly, our study observed co-occurrence of *C. neoformans* and *C. gattii* species complexes in both the clinical and environmental isolates, with *C. gattii* and *C. deuterogattii* dominating other molecular genotypes, while pathogenic *C. neoformans* dominated other molecular genotypes. Co-infection with these pathogenic molecular genotypes could account for high morbidity and mortality among patients with depleted immune system as well as healthy individuals. Previous researchers have identified *C. neoformans* as the major cause of cryptococcosis among individuals who are immunocompromised [31–33]. Among the four molecular types of the *C. gattii* species complex identified, *C. bacillisporus* has been previously reported as the most virulent and fertile, as was the case in the

Pacific Northwest in Canada [34]. Furthermore, it has been well documented that *C. gattii* and *C. bacillisporus* have widespread distributions worldwide, while *C. tetragattii* and *C. deuterogattii* are reported as less common [35].

This is in concordance with our findings on *C. deuterogattii*, which was not recovered in environmental samples and was less prevalent in clinical settings. Previous studies showed *C. neoformans* as dominating cryptococcal infections among AIDS patients [36–38], however, we observed a probable epidemiological shift within the environmental populations in Botswana, where hybrids are now becoming prevalent among patients infected with HIV as well as in the environment. The clinical and environmental AD hybrids clustered randomly among themselves in the phylogenetic tree. We identified seven *MAT**a***/*MAT*α hybrids in clinical isolates and nine in environmental isolates. The presence of cryptococcal hybrids in clinical and environmental settings is a serious public health concern. It has been well documented that sexual recombination produce infectious propagules [39,40], i.e., basidiospores [41], that may be inhaled by individuals with HIV or immunosuppressed individuals while harvesting mopane worms, or firewood potentially leading to cryptococcal infection.

Previous studies of mating type surveys in Botswana utilized comparatively smaller sample sizes [42] and only focused on *C. neoformans* to reveal evidence of recombination among isolates of the yeast [42]. Here, we used a larger number of samples, and the results indicate a diverse and complex mating history among Botswana populations. We observed both *MAT**a*** and *MAT*α haploids, and *MAT**a***/*MAT*α hybrids in environmental and clinical isolates. Only 10.5% of the total environmental *C. neoformans* isolates possessed *MAT**a*** mating type allele, while 28.9% possessed *MAT*α allele. We identified 44.9% of *MAT**a***/*MAT*α allele. There was an equal proportions and distribution of *MAT**a*** and *MAT*α mating types among *C. deneoformans* in clinical and environmental settings. Majority of clinical isolates (31.6%) possessed *MAT*α allele, and only 2.63% possessed *MAT**a*** allele. These findings are in concordance with findings of Chen et al. [36], who reported a higher frequency of *MAT*α in clinical and environmental populations from various locations in Botswana.

No clinical *C. neoformans* isolates with the *MAT**a*** mating type were recovered but were present in environmental samples, in addition to *MAT*α. The presence of *MAT**a*** mating type among *C. neoformans* is significant because this represents initial reporting of this mating type in mopane trees, although previously reported on other trees at a public park in South Africa [25]. In addition, the observation of *MAT**a*** and *MAT*α mating types in our results is not surprising but the prevalence of diploid hybrids represents unique findings compared to previous environmental investigations of mopane associated *CnCg* species complexes [42]. We reiterate that future studies focus on *in vitro* mating to determine intricate sexual recombination of the species complexes and elucidate how this natural process influences virulence and genetic structure. In addition, studies on how sexual recombination may have possibly led to the predomination of hybrid genotypes as observed in this study are indispensable.

5. Conclusions

This study employed a plethora of molecular techniques to establish a link between clinical and environmental isolates. Our investigation addressed several aspects of genetic diversity, ecological niche, population structure, and mode of reproduction of these pathogenic yeasts. The results show the presence of seven genetically isolated groups within the *CnCg* species complexes. These findings support previous findings that associated *C. neoformans* with *Co. mopane*. We used conventional and NGS platforms to demonstrate a link between environmental and clinical isolates and lay a foundation to further describe mating patterns/history in Botswana. Using NGS technology enabled us to establish co-occurrence of *C. neoformans* and *C. gattii* species in the environment and among HIV positive patients. To the best of our best knowledge, this is the first time this platform has been used to unequivocally identify cryptococcal isolates in Botswana. Given the observed emergence of hybrids among patients living with HIV, our immediate focus is to fully determine the geographic scope and detailed mating history of these pathogenic yeasts.

The results presented in our study will allow local laboratories to carry out cost effective molecular typing in Botswana, using sequencing under standardized conditions, as well as characterizing autochthonous isolates to establish Botswana reference strains. The ultimate objective is to employ highly reproducible and specific platforms to reveal similarities and differences among the clinical and environmental isolates, identification of exact mechanisms of human infections, and enlarging global biogeographical knowledge of these pathogenic yeasts. Hence, the need for more environmental, clinical and sentinel-oriented studies in Botswana to monitor and diagnose any emerging trends of this species complex cannot be underestimated.

Acknowledgements

The authors would like to express sincere gratitude to Professor Joe Jarvis and Ms Kwana Lechile who availed clinical isolates from their Ambition 1 Project. The MALDI-TOF MS for species identification was performed in Prof Lise Korsten' s laboratory in University of Pretoria, and the biotyper was acquired from the National Research Foundation (NRF) of South Africa. We would like to send sincere gratitude to Prof. Wieland Meyer (Molecular Mycology Research Laboratory, Westmead Hospital, Australia) as the reference isolates used in our study originated from their laboratory.

Ethics approval

Formal consent for environmental and clinical isolates and patient information followed institutional guidelines of the University of Botswana, and Ministry of Health and Wellness. Demographic and clinical data of patients are confidential.

Consent for publication

The authors declare that they have no conflicts of interest concerning this article.

Authors' contributions

Conceptualization, Kenosi Kebabonye, Mosimanegape Jongman; Supervision, Mosimanegape Jongman, Ishmael Kasvosve, Daniel Loeto; Validation, Mosimanegape Jongman; Writing—original draft, Kenosi Kebabonye and Mosimanegape Jongman; Writing—review & editing, Kenosi Kebabonye, Mosimanegape Jongman, Ishmael Kasvosve, Daniel Loeto, Sikhulile Moyo, Tatenda Choga.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The research was funded by University of Botswana through Office of Research and Development (ORD), under ID: 501100022089 (Grant number: UBR/RES/ROUND OF 39).

ORCID

Kenosi Kebabonye http://orcid.org/0000-0001-9636-0278
Mosimanegape Jongman http://orcid.org/0000-0002-6454-3173
Daniel Loeto http://orcid.org/0000-0001-5645-8828
Sikhulile Moyo http://orcid.org/0000-0003-3821-4592
Wonderful Choga http://orcid.org/0000-0001-7606-0569
Ishmael Kasvosve http://orcid.org/0000-0002-2046-545X

References

[1] Rajasingham R, Smith RM, Park BJ, et al. Global burden of disease of HIV-associated cryptococcal meningitis: an updated analysis. Lancet Infect Dis. 2017;17(8):873–881. doi: 10.1016/S1473-3099(17)30243-8.

[2] You M, Xu J. The effects of environmental and genetic factors on the germination of basidiospores in the *Cryptococcus gattii* species complex. Sci Rep. 2018;8(1):15260. doi: 10.1038/s41598-018-33679-2.

[3] Tavares ER, Gionco B, Morguette AEB, et al. Phenotypic characteristics and transcriptome profile of *Cryptococcus gattii* biofilm. Sci Rep. 2019;9(1):6438. doi: 10.1038/s41598-019-42896-2.

[4] Trilles L, Lazéra MdS, Wanke B, et al. Regional pattern of the molecular types of *Cryptococcus neoformans* and *Cryptococcus gattii* in Brazil. Mem Inst Oswaldo Cruz. 2008;103(5):455–462. doi: 10.1590/s0074-02762008000500008.

[5] Firacative C, Trilles L, Meyer W. MALDI-TOF MS enables the rapid identification of the major molecular types within the *Cryptococcus neoformans/C. gattii* species complex. PLoS One. 2012;7(5):e37566. doi: 10.1371/journal.pone.0037566.

[6] Herkert P, Hagen F, Pinheiro R, et al. Ecoepidemiology of *Cryptococcus gattii* in developing countries. JoF. 2017;3(4):62. doi: 10.3390/jof3040062.

[7] Hurtado JC, Castillo P, Fernandes F, et al. Mortality due to *Cryptococcus neoformans* and *Cryptococcus gattii* in low-income settings: an autopsy study. Sci Rep. 2019;9(1):7493. doi: 10.1038/s41598-019-43941-w.

[8] Torres RG, Etchebehere RM, Adad SJ, et al. Cryptococcosis in acquired immunodeficiency syndrome patients clinically confirmed and/or diagnosed at necropsy in a teaching hospital in Brazil. Am J Trop Med Hyg. 2016;95(4):781–785. doi: 10.4269/ajtmh.16-0148.

[9] Hagen F, Khayhan K, Theelen B, et al. Recognition of seven species in the *Cryptococcus gattii/Cryptococcus neoformans* species complex. Fungal Genet Biol. 2015;78:16–48. doi: 10.1016/j.fgb.2015.02.009.

[10] Farrer RA, Desjardins CA, Sakthikumar S, et al. Genome evolution and innovation across the four major lineages of *Cryptococcus gattii*. MBio. 2015;6(5):e00868–e00815. doi: 10.1128/mBio.00868-15.

[11] Hu S, Gu F, Chen M, et al. A novel method for identifying and distinguishing *Cryptococcus neoformans* and *Cryptococcus gattii* by surface-enhanced Raman scattering using positively charged silver nanoparticles. Sci Rep. 2020;10(1):12480. doi: 10.1038/s41598-020-68978-0.

[12] Maziarz EK, Perfect JR. Cryptococcosis. Infect Dis Clin North Am. 2016;30(1):179–206. doi: 10.1016/j.idc.2015.10.006.

[13] Clark AE, Kaleta EJ, Arora A, et al. Matrix-assisted laser desorption ionization–time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. Clin Microbiol Rev. 2013;26(3):547–603. doi: 10.1128/CMR.00072-12.

[14] Meyer W, Aanensen DM, Boekhout T, et al. Consensus multi-locus sequence typing scheme for *Cryptococcus neoformans* and *Cryptococcus gattii*. Med Mycol. 2009;47(6):561–570. doi: 10.1080/13693780902953886.

[15] Staib F. *Cryptococcus neoformans* and *Guizotia abyssinica* (syn. *G. oleifera* DC).(colour reaction for *Cr. neoformans*). Zeitschr f Hygiene. 1962; 148(5):466–475. doi: 10.1007/BF02156784.

[16] Litvintseva AP, Carbone I, Rossouw J, et al. Evidence that the human pathogenic fungus *Cryptococcus neoformans* var. *grubii* may have evolved in Africa. PLoS One. 2011;6(5):e19688. doi: 10.1371/journal.pone.0019688.

[17] Campbell LT, Fraser JA, Nichols CB, et al. Clinical and environmental isolates of *Cryptococcus gattii* from Australia that retain sexual fecundity. Eukaryot Cell. 2005;4(8):1410–1419. doi: 10.1128/EC.4.8.1410-1419.2005.

[18] Meyer W, Marszewska K, Amirmostofian M, et al. Molecular typing of global isolates of *Cryptococcus neoformans* var. *neoformans* by polymerase chain reaction fingerprinting and randomly amplified polymorphic DNA—a pilot study to standardize techniques on which to base a detailed epidemiological survey. ELECTROPHORESIS: Int J. 1999;20(8):1790–1799. doi: 10.1002/(SICI)1522-2683(19990101)20:8<1790::AID-ELPS1790>3.0.CO;2-2.

[19] Tavares ER, Azevedo CS, Panagio LA, et al. Accurate and sensitive real-time PCR assays using intergenic spacer 1 region to differentiate *Cryptococcus gattii* sensu lato and *Cryptococcus neoformans* sensu lato. Med Mycol. 2016;54(1):89–96. doi: 10.1093/mmy/myv078.

[20] Bohl JA, Lay S, Chea S, et al. Discovering disease-causing pathogens in resource-scarce southeast asia using a global metagenomic pathogen monitoring system. Proc Natl Acad Sci USA. 2022; 119(11):e2115285119. doi: 10.1073/pnas.2115285119.

[21] Morrison GA, Fu J, Lee GC, et al. Nanopore sequencing of the fungal intergenic spacer sequence as a potential rapid diagnostic assay. J Clin Microbiol. 2020;58(12):e01972–20. doi: 10.1128/JCM.01972-20.

[22] Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38(7):3022–3027. doi: 10.1093/molbev/msab120.

[23] Sagulenko P, Puller V, Neher RA. TreeTime: maximum-likelihood phylodynamic analysis. Virus Evol. 2018;4(1):vex042. doi: 10.1093/ve/vex042.

[24] Kelentse N, Moyo S, Molebatsi K, et al. Reversal of CSF HIV-1 escape during treatment of HIV-Associated cryptococcal meningitis in Botswana. Biomedicines. 2022;10(6):1399. doi: 10.3390/biomedicines10061399.

[25] Vreulink J-M, Khayhan K, Hagen F, et al. Presence of pathogenic cryptococci on trees situated in two recreational areas in South Africa. Fungal Ecol. 2017;30:101–111. doi: 10.1016/j.funeco.2017.09.005.

[26] Litvintseva AP, Xu J, Mitchell TG. Population structure and ecology of *Cryptococcus neoformans* and *Cryptococcus gattii*. Cryptococcus: from Human Pathogen to Model Yeast. 2010;2:97–111.

[27] Mashabane L, Wessels D, Potgieter M. The utilisation of *Colophospermum mopane* by the Vatsonga in the Gazankulu region (eastern Northern Province, South Africa). South Afr J Bot. 2001;67(2): 199–205. doi: 10.1016/S0254-6299(15) 31120-0.

[28] Vogel G. For more protein, filet of cricket. Science. 2010;327(5967):811–811. doi: 10.1126/science.327.5967.811.

[29] Morrison VA. Echinocandin antifungals: review and update. Expert Rev anti Infect Ther. 2006;4(2): 325–342. doi: 10.1586/14787210.4.2.325.

[30] Nicholls SM, Quick JC, Tang S, et al. Ultra-deep, long-read nanopore sequencing of mock microbial community standards. GigaScience. 2019;8(5): giz043. doi: 10.1093/gigascience/giz043.

[31] Chau TT, Mai NH, Phu NH, et al. Research article a prospective descriptive study of *Cryptococcal meningitis* in HIV uninfected patients in Vietnam-high prevalence of *Cryptococcus neoformans* var. *grubii* in the absence of underlying disease. BMC Infect Dis. 2010;10:199. doi: 10.1186/1471-2334-10-199.

[32] Jain N, Cook E, Xess I, et al. Isolation and characterization of senescent *Cryptococcus neoformans* and implications for phenotypic switching and pathogenesis in chronic cryptococcosis. Eukaryot Cell. 2009;8(6):858–866. doi: 10.1128/EC.00017-09.

[33] Meyer W, et al. Molecular typing of the *Cryptococcus neoformans*/*Cryptococcus gattii* species complex. Cryptococcus: from Human Pathogen to Model Yeast. 2010;:327–357.

[34] Kidd SE, Hagen F, Tscharke RL, et al. A rare genotype of *Cryptococcus gattii* caused the cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). Proc Natl Acad Sci USA. 2004;101(49): 17258–17263. doi: 10.1073/pnas.0402981101.

[35] Acheson ES, Galanis E, Bartlett K, et al. Searching for clues for eighteen years: deciphering the ecological determinants of *Cryptococcus gattii* on Vancouver Island, British Columbia. Med Mycol. 2018;56(2):129–144. doi: 10.1093/mmy/myx037.

[36] Chen Y, Litvintseva AP, Frazzitta AE, et al. Comparative analyses of clinical and environmental populations of *Cryptococcus neoformans* in Botswana. Mol Ecol. 2015;24(14):3559–3571. doi: 10.1111/mec.13260.

[37] Litvintseva AP, Thakur R, Reller LB, et al. Prevalence of clinical isolates of *Cryptococcus gattii* serotype C among patients with AIDS in Sub-Saharan Africa. J Infect Dis. 2005;192(5):888–892. doi: 10.1086/432486.

[38] Litvintseva AP, Thakur R, Vilgalys R, et al. 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.

[39] Sukroongreung S, Kitiniyom K, Nilakul C, et al. Pathogenicity of basidiospores of *Filobasidiella neoformans* var. *neoformans*. Med Mycol. 1998;36(6): 419–424. doi: 10.1046/j.1365-280X.1998.00181.x.

[40] Velagapudi R, Hsueh Y-P, Geunes-Boyer S, et al. Spores as infectious propagules of *Cryptococcus neoformans*. Infect Immun. 2009;77(10):4345–4355. doi: 10.1128/IAI.00542-09.

[41] Kwon-Chung K. A new genus, filobasidiella, the perfect state of *Cryptococcus neoformans*. Mycologia. 1975;67(6):1197–1200. doi: 10.2307/3758842.

[42] Chen Z-F, Ying G-G. Occurrence, fate and ecological risk of five typical azole fungicides as therapeutic and personal care products in the environment: a review. Environ Int. 2015;84:142–153. doi: 10.1016/j.envint.2015.07.022.