Analyses of the Global Multilocus Genotypes of the Human Pathogenic Yeast *Cryptococcus neoformans* Species Complex

Megan Hitchcock and Jianping Xu *

Department of Biology, McMaster University, Hamilton, ON L8S 4K1, Canada
* Correspondence: jpxu@mcmaster.ca

Abstract: *Cryptococcus neoformans* species complex (CNSC) is a globally distributed human opportunistic yeast pathogen consisting of five major molecular types (VNI, VNII, VNB, VNIII and VNIV) belonging to two species, *C. neoformans* (VNI, VNII and VNB, collectively called serotype A) and *C. deneoformans* (VNIV, commonly called serotype D), and their hybrids (VNIII, serotype AD). Over the years, many studies have analyzed the geographical distribution and genetic diversity of CNSC. However, the global population structure and mode of reproduction remain incompletely described. In this study, we analyze the published multilocus sequence data at seven loci for CNSC. The combined sequences at the seven loci identified a total of 657 multilocus sequence types (STs), including 296 STs with known geographic information, representing 4200 non-redundant isolates from 31 countries and four continents. Among the 296 STs, 78 and 52 were shared among countries and continents, respectively, representing 3643 of the 4200 isolates. Except for the clone-corrected serotype D sample among countries, our analysis of the molecular variance of the 4200 isolates revealed significant genetic differentiations among countries and continents in populations of CNSC, serotype A, and serotype D. Phylogenetic analyses of the concatenated sequences of all 657 STs revealed several large clusters corresponding to the major molecular types. However, several rare but distinct STs were also found, representing potentially novel molecular types and/or hybrids of existing molecular types. Phylogenetic incompatibility analyses revealed evidence for recombination within all four major molecular types—VNI, VNII, VNIV and VNB—as well as within two VNB subclades, VNBI and VNBII, and two ST clusters around the most common STs, ST5 and ST93. However, linkage disequilibrium analyses rejected the hypothesis of random recombination across most samples. Together, our results suggest evidence for historical differentiation, frequent recent gene flow, clonal expansion and recombination within and between lineages of the global CNSC population.

Keywords: *Cryptococcus*; *Cryptococcus neoformans*; multilocus sequence typing; sequence type; yeast; geographical distribution; recombination

1. Introduction

The human pathogenic *Cryptococcus* (HPC) is a group of globally distributed basidiomycete yeasts. These yeasts are opportunistic pathogens to humans and other mammals. They are commonly found in soil, avian excretion and rotting tree barks [1–4]. HPC consists of two species complexes, the *Cryptococcus neoformans* species complex (CNSC) and the *Cryptococcus gatti* species complex (CGSC). Globally, most human infections are caused by strains of CNSC. CNSC tends to infect immunocompromised hosts and is a leading cause of death in HIV patients [5,6]. Infections can lead to systemic cryptococcosis, with the most common and detrimental form being cryptococcal meningitis. In 2014, there were ~223,100 recorded cases of cryptococcal meningitis resulting in ~180,100 deaths worldwide [5,6]. With cases on the rise over the past five decades due to increasing populations of immunocompromised hosts [6,7], it is important that we improve our understanding of the global distribution and genetic diversity of *C. neoformans*.

CNSC is a highly heterogeneous group of organisms, with divergent lineages showing >10% nucleotide sequence divergence [1]. Over the last 50 years, a variety of molecular markers have been used to identify strains of CNSC [8]. These markers have revealed divergent lineages within CNSC. The current emerging consensus separates CNSC into two species, *C. neoformans* (serotype A) and *C. deneoformans* (serotype D). *C. neoformans* (Serotype A) is further divided into three major molecular types, VNI, VNB and VNII, while *C. deneoformans* (serotype D) corresponds to the molecular type VNIV. In addition to these four major molecular types, VNB was further divided into two subtypes, VNBI and VNBII, and diploid/aneuploid hybrids have been observed in nature and are referred to as VNIII or serotype AD hybrids [1,8–10].

To help standardize the genotyping system and make it easy to share information among labs, in 2007, the International Society for Human and Animal Mycology (ISHAM) established a committee to set up a multi-locus sequence typing (MLST) method. The recommended system was published in 2009 and included partial DNA sequences at the following seven loci: *CAP59, GPD1, LAC1, PLB1, SOD1, URA5* and *IGS1* [11]. Subsequently, an online international fungal multi-locus serotyping database (IFMLST) was established for storing and comparing the MLST data. This data repository is comprised of allelic profiles for each recorded sequence type (ST) and the nucleotide sequence for each determined allele type (AT) at each of the seven loci [12]. Newly genotyped strains can be compared to the database to determine their AT and ST profiles. The MLST system is a great tool for the consistent and efficient comparison of strain genotypes across labs. However, little analysis has been conducted utilizing the publicly available dataset.

Considering the high global burden of infections by CNSC, it is important to understand the global population genetic variation of this species complex. In this paper, we investigate the global genotype distribution and population structure by analyzing the 4200 CNSC isolates with MLST data published in 41 studies. We aim to estimate the overall geographical pattern of genetic variation and determine if recombination plays a role in shaping the diversity observed within individual species and individual molecular types of CNSC. We hypothesize that the genetic variations within CNSC are geographically structured, and recombination plays an important role during the evolution of this species complex.

2. Materials and Methods

2.1. Strains, Genotypes and Metadata

We conducted a literature search through PubMed using the search term "Cryptococcus MLST". All retrieved papers were scanned for isolates, with MLST data belonging to *C. neoformans*, *C. neoformans* var. *grubii*, *C. neoformans* var. *neoformans*, *C. neoformans* serotype A and *C. neoformans* serotype D. The reported MLST data on the isolates of CNSC from these studies were extracted and compiled. For each sequence type, its allelic profile and DNA sequence data at the seven loci were retrieved from the publicly available International Fungal Multi Locus Sequence Typing Database. All MLST data for strains of CNSC published by January 2022 were included. A total of 4200 isolates from 41 studies were included in our population genetic analyses. The detailed descriptions of these isolates and their genotype data, including their protocols for DNA extraction, amplification and sequencing and allelic profiling can be found in the original reports [13–53].

2.2. Phylogenetic Distribution of Strains and Genotypes

To determine the relationships among STs within CNSC, DNA sequences at the seven MLST loci were concatenated based on the allelic profile for each sequence type. The concatenated sequences were then imported into MEGA X and aligned through Muscle [54,55]. A phylogenetic tree was reconstructed using a neighbor-joining method based on the K2P distance model (Kimura 2-parameter model) [54]. Associated molecular type classifications from the IFMLST database, as well as geographical and ecological source data based on the

published literature, were added to each ST to show their distributions on the phylogenetic tree using the iTOL software [56].

2.3. Population Genetic Analyses

To investigate if geographic populations of CNSC were genetically subdivided, strains were separated into different countries and continents based on their metadata and analyzed using the GenAlEx V.6.5 program [57]. In our population genetic analyses, three taxonomy group-based samples were analyzed: the entire CNSC sample, the *C. neoformans* (serotype A) sample and the *C. deneoformans* (serotype D) sample. For each of the three taxonomic samples, two types of datasets were analyzed: non-clone-corrected (NCC), and clone-corrected (CC). The NCC datasets included all isolates extracted from the literature. For the CC dataset, only one representative strain of each ST from each country was included in our analyses.

For each of the six datasets (three taxonomic NCC datasets and three taxonomic CC datasets), we separately conducted the analyses of molecular variance (AMOVA) at the country and continental levels. In addition, Wright's F_{ST} values were obtained between pairs of geographic samples. To reduce potential biases due to small samples sizes, individual subpopulations with <five isolates were excluded from the country analysis. Statistical significance for each test was obtained by comparing the observed with the distributions of 1000 permuted datasets generated based on a null hypothesis of no genetic differentiations within each analyzed dataset.

The MLST dataset including all 657 STs was also used to identify potential evidence for recombination within individual species (serotypes), molecular types and selected phylogenetic clusters. For this test, only the clone-corrected allelic profiles at the seven loci were analyzed. Specifically, phylogenetic compatibility and linkage disequilibrium analyses were performed with 1000 randomizations for the total clone-corrected CNSC dataset, the serotype A subset, the serotype D subset, the three molecular types (VNI, VNII and VNB) of serotype A, two subtypes of VNB (VNBI and VNBII) as well as phylogenetic clusters associated with the two most common STs to investigate potential evidence for recombination across CNSC. Details regarding the underlying principles of these tests and how these tests were conducted can be found in the MultiLocus V1.3 manual [58].

3. Results

As of January 2022, there were 657 total sequence types (STs) deposited into the *Cryptococcus* MLST database for CNSC, with associated DNA sequence data for all seven loci. Of the 657 STs, the geographical location information was documented for 296 (45%), while the remaining 361 (55%) STs had unknown geographic information. Our population genetic analyses focused on the 296 STs. The 296 STs represented 4200 CNSC isolates, as extracted from 41 published reports. The metadata for all isolates were retrieved from these published reports. Below, we summarize the retrieved data on the 4200 isolates and present the results of our analyses.

3.1. Geographical and Ecological Distributions

The geographic distribution of the non-redundant 4200 CNSC isolates is presented in Table 1. These isolates were from 31 countries and five continents, with the majority being found in Asia (61.9%), followed by Africa (12.6%), Europe (14.3%), South America (10.9%) and North America (0.3%). At the country level, the highest number of isolates in this dataset came from China (1216 isolates; 28.95%), while the lowest came from the Congo, the Dominican Republic and the Democratic Republic of Congo (with one isolate each). In between these two extremes, the second largest national population of CNSC in the retrieved dataset was from Thailand (524 isolates; 12.27%), followed by India (380; 9.05%), Brazil (318 isolates; 7.57%), South Africa (268 isolates; 6.38%), Uganda (241 isolates; 5.74%), France (226; 5.38%), Italy (151 isolates; 3.60%), Germany (145 isolates, 3.45%), Vietnam (136 isolates; 3.24%) and Japan (119 isolates; 2.83%). The remaining

20 countries each had <100 isolates analyzed, and, together, they contributed 476 isolates to the analyzed dataset. The geographic associations of the 296 STs are presented in Supplementary Table S1. Among the 296 STs, ST5 was the most abundant; it was found across 18 countries on four continents (Supplementary Table S1).

Table 1. Summary of the geographic distributions of isolates of the *C. neoformans* species complex genotyped using the ISHAM multilocus sequence typing scheme.

Region	*n*	%	Region	*n*	%	Region	*n*	%	Region	*n*	%
Asia	2600	61.9%	Africa	529	12.6%	Europe	601	14.3%	South America	457	10.9%
China	1216	28.95%	South Africa	268	6.38%	France	226	5.38%	Brazil	318	7.57%
Thailand	524	12.48%	Uganda	241	5.74%	Italy	151	3.6%	Colombia	88	2.10%
India	380	9.05%	Nigeria	12	0.29%	Germany	145	3.45	Peru	51	1.21%
Vietnam	136	3.24%	Libya	3	0.07%	Turkey	41	0.98%			
Japan	119	2.83%	Tanzania	3	0.07%	Greece	24	0.57%			
Saudi Arabia	75	1.79%	Congo	1	0.02%	Spain	10	0.24%	North America	13	0.3%
South Korea	50	1.19%	DRC	1	0.02%	Cyprus	4	0.10%	USA	12	0.29%
Indonesia	44	1.05%							D. Republic	1	0.02%
Iran	30	0.71%									
Kuwait	15	0.36%									
Qatar	8	0.19%									
Malaysia	3	0.07%									

n: total isolates, %: percent of isolates of the total recorded in CNSC.

Of the 657 STs in the MLST database, only 284 had ecological niche information (Supplementary Table S2; Table 2). These 284 STs represented a total of 4064 isolates, while the remaining 373 STs had no ecological niche/source data. Here, we broadly categorized the isolates into three ecological sources: clinical, environmental and veterinary. The majority of the 4200 isolates were collected from clinical sources (3370 isolates; 80.24%), followed by environmental (648 isolates; 15.43%), and veterinary (46 isolates; 1.10%) sources, leaving 3.24% of isolates with unknown source information. The ecological distributions of the individual STs are shown in Supplementary Table S2. A total of 14 STs were found in all three niches; 3 STs were found from both clinical and veterinary sources only; 32 STs were found in both clinical and environmental sources only; and no ST was shared between only environmental and veterinary sources. The remaining 235 STs with ecological niche information were each found in only one of the three ecological niches (Table 2).

Table 2 summarizes the geographic and ecological distributions of the 296 STs in the published MLST literature for CNSC. Geographically, among the 296 STs, 15 STs (representing a total of 2675 isolates) were found in all four continents, 9 STs (representing a total of 656 isolates) were found in three of the four continents, 28 STs (representing a total of 312 isolates) were found in two of the four continents and 244 (representing 557 isolates) were found in only one of the continents (Table 2). Among the 244 STs, 176 were each represented by only one isolate in the database. Ecologically, among the 296 STs, 284 STs, including 4064 isolates, had ecological niche data. Of these 284 STs, 14 (representing 2550 isolates) were found in all three ecological niches, 35 STs (representing 1034 isolates) were found in two of the three ecological niches and 235 STs (representing 480 isolates) were found in one niche only (Table 2). The detailed geographic and ecological distributions for each of the 296 STs are shown in Supplementary Tables S1 and S2. At the country level, 78 (26%) sequence types were reported from two or more countries each (Supplementary Table S3).

Table 2. Summary distribution of 296 sequence types of the *C. neoformans* species complex among continents and ecological niches.

Distribution Patterns	Specific Continent(s)/Ecological Niche(s)	Number of Sequence Types	Number of Isolates
Geographic			
In all four continents		15	2675
In three continents only			
	Asia + Africa + Europe	1	207
	Asia + Africa + America	4	353
	Africa + America + Europe	1	4
	Asia + Europe + America	3	92
In two continents only			
	Asia + Africa	8	24
	Asia + Europe	6	197
	Asia + America	3	28
	Africa + Europe	3	23
	Africa + America	4	13
	Europe + America	4	27
In one continent only			
	Asia	50	190
	Africa	73	107
	America	19	41
	Europe	102	219
Ecological niches			
In all three niches		14	2550
In two niches only			
	Clinical + Veterinary	3	8
	Clinical + Environmental	32	1026
	Veterinary + Environmental	0	0
In one niche only			
	Clinical	184	380
	Veterinary	11	16
	Environmental	40	84

3.2. DNA Sequence Variation

The allelic profiles of each ST, including the allele type (AT) number at each of the seven MLST loci (*CAP59, GPD1, LAC1, IGS1, PLB1, SOD1, URA5*), were retrieved for all 657 STs in the database. Summaries of the allele types across the total sample, the serotype A sample and the serotype D sample are shown in Table 3. The differences in length of bp per allele type for each gene range from 0 difference for *CAP59* to 45bp difference for *IGS1*. Among the seven loci, in the total sample, *PLB1* had the fewest number of alleles (44) while *IGS1* had the most (93). A largely similar pattern was observed for the serotype A and serotype D samples, where *IGS1* had the highest allele number in both. However, *GPD1* had the lowest allele number in the serotype D sample. The range of occurrence of each allele type in each of the samples is shown in Supplementary Table S3.

3.3. Phylogenetic Analysis

The *C. neoformans* species complex is commonly grouped into five broad molecular types: VNI, VNII, VNIII, VNB and VNIV. The strains of VNI, VNII and VNB belonged to *C. neoformans* (serotype A); the strains of VNIV belonged to *C. deneoformans* (serotype D); and the strains of VNIII (serotype AD) represented hybrids of serotypes A and D. For some VNB strains, they were further classified into VNBI and VNBII [8,10]. The molecular type designations were mostly based on the restriction enzyme digest pattern of the *URA5* sequence, amplified fragment length polymorphisms or PCR fingerprinting [8]. Analyses of the concatenated sequences at the seven MLST loci showed a largely consistent clustering of STs into their original molecular type designations, with VNIV being the most distant

from VNI, VNII and VNB (Figure 1). Similarly, except for one ST (ST434) that was originally assigned to VNBII but was clustered more closely with VNBI strains, all other STs originally assigned to VNBI and VNBII were clearly separated into two groups (Figure 1). However, there were several other notable inconsistencies. Specifically, six STs originally assigned to VNIV (ST521, ST254, ST266, ST355, ST489 and ST538) showed a closer relationship with the VNI clade. In contrast, 14 STs originally assigned to VNI (ST210, ST224, ST225, ST249, ST259, ST263, ST326, ST345, ST353, ST354, ST358, ST365, ST366 and ST651) and three STs originally assigned to VNII (ST221, ST222 and ST363) showed intermediate phylogenetic placing between the major serotypes A and D genotypes. Interestingly, 15 of the above 23 STs contained alleles with mixed clustering patterns, where some of the alleles belonged to the serotype A cluster, while others belonged to the serotype D allele cluster (Table 4; Supplementary Figures S1–S7). In addition, multiple STs originally assigned to molecular types VNI, VNII and VNB showed ambiguous placements within serotype A, often showing large distances from the three main molecular types (Figure 1). In contrast, 51 STs with previously undefined molecular type assignments were grouped into various species/molecular types (Figure 1). Phylogenic trees showing relationships among allele sequences for each of the seven genes can be seen in Supplementary Figures S1–S7.

Table 3. Allelic variation among the seven loci used for the multilocus sequence typing of the *C. neoformans* species complex. The number of overlapping alleles at each locus between the serotype A and serotype D strains was calculated based on the assigned molecular types by the MLST database.

Gene	Gene Name	Chromosome Location	Length (bp)	Total Allele Number in CNSC	Allele Number in Serotype A	Allele Number in Serotype D	Number of Overlapping Alleles between A and D
CAP59	Capsular-associated protein	1	560	55	46	18	9
GPD1	Glyceraldehyde-3-phosphate dehydrogenase	7	543–546	45	38	15	8
IGS1	Ribosomal RNA intergenic spacer	2	685–730	93	75	32	14
LAC1	Laccase	8	469–473	50	40	22	12
PLB1	Phospholipase	12	517–543	44	38	18	12
SOD1	Cu, Zn superoxide dismutase	5	526–543	68	58	19	9
URA5	Orotidine monophosphate pyrophosphorylase	8	636–652	57	46	24	13

Table 4. Allelic profiles of STs with intermediate and inconsistent molecular type assignments between the original assignments in the MLST database and those based on the seven-gene phylogeny in Figure 1. Fifteen STs with mixed background colors are likely hybrids of serotypes A and D.

VNIV	ST	*CAP59*	*GPD1*	*IGS1*	*LAC1*	*PLB1*	*SOD1*	*URA5*
	254	9	1	14	6	4	1	16
	266	5	22	34	19	10	23	18
	355	22	5	1	19	3	17	20
	489	1	1	1	1	1	27	1
	521	27	13	12	6	9	8	13
	538	54	42	1	5	38	60	55
VNI								
	210	4	11	56	6	6	1	3
	224	9	11	57	6	6	33	2
	225	13	1	55	11	4	36	1
	249	34	1	55	27	4	7	1
	259	24	23	28	3	2	1	2
	263	9	11	57	6	6	27	2
	326	9	23	60	23	16	1	40

Table 4. *Cont.*

VNIV	ST	CAP59	GPD1	IGS1	LAC1	PLB1	SOD1	URA5
	345	5	22	34	6	10	23	3
	353	22	5	32	19	3	1	1
	354	37	5	32	19	3	1	1
	358	24	1	32	3	3	1	1
	365	22	6	32	19	3	1	1
	366	25	2	60	23	16	1	1
	651	1	1	60	3	14	1	34
VNII								
	221	8	10	58	8	2	3	11
	222	8	10	58	8	12	3	11
	363	9	22	24	19	19	1	33

The first column represents the original molecular type assignment found in the ISHAM MLST database. Alleles highlighted in blue represent those in the serotype A cluster, while those highlighted in yellow represent alleles in the serotype D cluster.
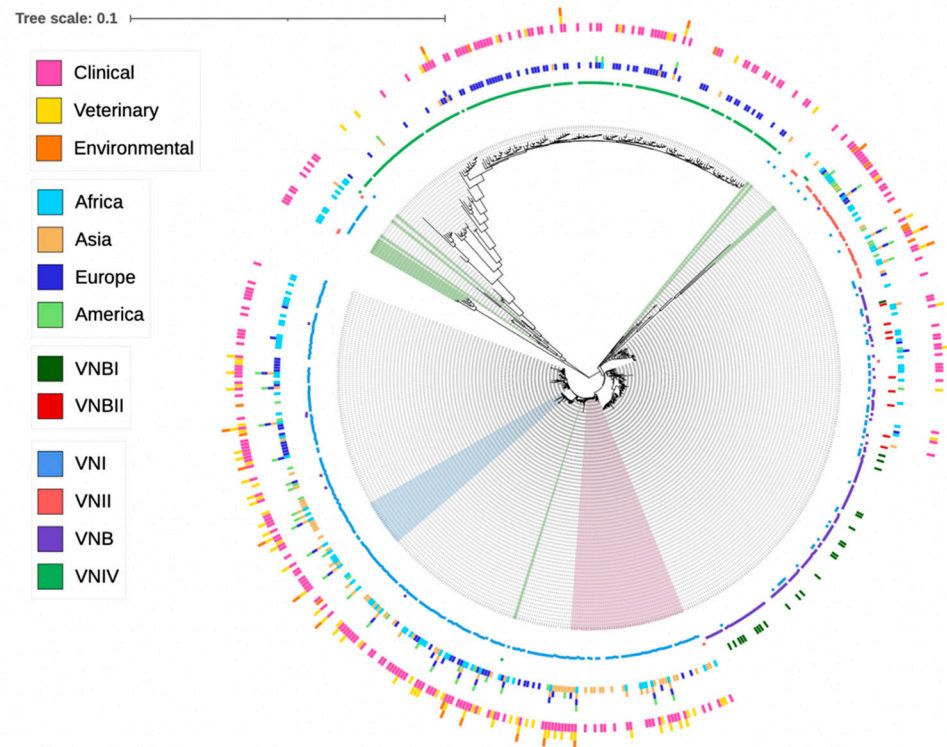

Figure 1. Phylogenic tree showing the relationships among 657 sequence types (ST) of the *C. neoformans* species complex. A total of 3 large circles and 10 small circles were used to display the metadata associated with each ST. The outermost large circle containing three small circles shows ecological niches, with each colored bar corresponding to one ecological niche. The middle large circle containing four small circles shows the continental distributions of each ST, with each colored bar corresponding to one continent. The next circle inside containing two colored bars corresponds to strains that were identified as VNBI (green) and VNBII (red) by authors who reported these STs. The next four small circles correspond to four original molecular types (VNI, VNII, VNB and VNIV) assigned for each ST, with each small circle corresponding to one molecular type and the presence of the ST in the molecular type indicated by a colored bar stick. Legends on the left indicate the correspondence between each color and the metadata of each ST. STs shaded in green are those showing significant deviations from their original four main molecular type assignments. STs shaded in pink and blue correspond to those closely related to ST5 and ST93, respectively, which were separately analyzed for evidence of recombination.

3.4. AMOVA

Because of the highly skewed population sizes among countries, with seven countries each having fewer than five isolates represented, our AMOVA was conducted separately at the continental and country levels, instead of through a two-level hierarchical analysis. At the country level, only those with more than five isolates represented are included. The overall objective of our AMOVA was to assess how much geographic separations contributed to the total genetic variation. Below, we briefly summarize the results.

At the continental level analyses, in the none-clone-corrected sample, genetic variations within continents contributed 72%, 78% and 84% of the total observed genetic variations in the total CNSC population, the serotype A population and the serotype D population, respectively. The remaining 28%, 22% and 16% were attributed among continents. The within-continent and among-continent contributions for each of the three taxonomic populations were statistically significant at the $p < 0.001$ level (Table 5). In the three clone-corrected samples, genetic variations within continents contributed 96%, 98% and 96% of the total observed genetic variations in the entire CNSC population, the serotype A population and the serotype D population, respectively. These percentages were significantly greater than those without clone corrections. The remaining 4%, 2% and 4% were attributed among continents. Despite the smaller percentages of contributions, the among-continent contributions for two of the three population types were statistically significant at $p < 0.001$, while the serotype D population was significant at $p = 0.03$ (Table 5). The pairwise comparisons between continents for the three taxonomic samples are shown in Table 6.

Table 5. Analysis of molecular variance at the continental level.

	None-Clone-Corrected				**Clone-Corrected**			
	df	**MS**	**Est.Var**	**%**	**df**	**MS**	**Est.Var**	**%**
Total Population								
Among Continents	3	513.8	0.64	28% ***	3	14.5	0.1	4% ***
Within Continents	4196	1.7	1.65	72% ***	382	2.8	2.77	96% ***
Total	4199	2	2.3	100%	385	2.9	2.90	100%
Serotype A								
Among Continents	3	298.5	0.42	22% ***	3	5.2	0.04	2% ***
Within Continents	3895	1.5	1.5	78% ***	275	2.5	2.51	98% ***
Total	3898	1.7	2.93	100%	278	2.6	2.54	100%
Serotype D								
Among Continents	3	15.6	0.35	16% ***	3	3.5	0.12	4% *
Within Continents	250	1.9	1.91	84% ***	75	2.6	2.65	96% ***
Total	253	2	2.3	100%	78	2.7	2.8	100%

df: degrees of freedom; MS: mean square; Est.Var: estimated variance; %: percentage of variance; * $p < 0.05$; *** $p < 0.001$.

At the country level, in the non-clone-corrected sample analyses, genetic variations within countries contributed 39%, 41% and 74% of the total observed genetic variations in the total CNSC sample, the serotype A sample and the serotype D sample, respectively. The remaining 61%, 59% and 26% were attributed among countries. The within-country and among-country contributions for each of the three taxonomic sample types were statistically significant at the $p < 0.001$ level (Table 7). In the three clone-corrected samples, genetic variations within countries contributed 83%, 79% and 99% of the total observed genetic variations in the total CNSC sample, the serotype A sample and the serotype D sample, respectively. Similar to those observed at the continental level, these percentages by within-countries in the clone-corrected samples were significantly greater than those without clone corrections. The remaining 17%, 21% and 1% were attributed among countries. Despite the smaller percentages of contributions, except for the serotype D sample, the remaining two among-country contributions for the three sample types were statistically significant at the $p < 0.001$ level (Table 7). The pairwise comparisons among countries for the three taxonomic samples are shown in Supplementary Tables S4–S6.

Table 6. Pairwise population comparison at the continental level.

	Non-Clone-Corrected			Clone-Corrected		
	Africa	**Asia**	**Europe**	**Africa**	**Asia**	**Europe**
Total						
Africa						
Asia	0.089 ***			0.101 ***		
Europe	0.512 ***	0.347 ***		0.019 ***	0.050 ***	
America	0.226 ***	0.165 ***	0.254 ***	0.052 ***	0.013 ***	0.012 ***
Serotype A						
Africa						
Asia	0.088 ***			0.032 ***		
Europe	0.576 ***	0.345 ***		0.015 ***	0.019 ***	
America	0.263 ***	0.136 ***	0.15 ***	0.012 ***	0.018 ***	0.000 ***
Serotype D						
Africa						
Asia	0.217 ***			0.058 *		
Europe	0.241 ***	0.151 ***		0.011 *	0.023 *	
America	0.223 ***	0.103 ***	0.197 ***	0.223 *	0.163 *	0.066 *

$* \; p < 0.05; *** \; p < 0.001.$

Table 7. Analysis of molecular variance at the country level.

	None-Clone-Corrected				Clone-Corrected			
	df	**MS**	**Est.Var**	**%**	**df**	**MS**	**Est.Var**	**%**
Total Population								
Among Countries	23	210.8	1.3	61% ***	19	14.2	0.5	17% ***
Within Countries	34,159	0.86	0.86	39% ***	495	2.3	2.3	83% ***
Total	4182	2	2.2	100%	514	2.7	2.8	100%
Serotype A								
Among Countries	22	170.1	1.2	59% ***	17	12.9	0.52	21% ***
Within Countries	3862	0.77	0.77	41% ***	378	1.9	21.9	79% ***
Total	3884	1.8	1.9	100%	395	2.4	2.4	100%
Serotype D								
Among Countries	4	25.4	0.6	26% ***	2	3.3	0.04	1%
Within Countries	234	1.6	1.6	74% ***	68	2.6	2.6	99% ***
Total	338	2	2.1	100%	70	2.6	2.6	100%

df: degrees of freedom; MS: mean square; Est.Var: estimated variance; %: percentage of variance; *** $p < 0.001$.

3.5. Recombination & Linkage Disequilibrium

We investigated the potential signatures of recombination among different samples of CNSC using two common indicators: phylogenetic incompatibility and linkage equilibrium. Here, aside from the three large taxonomic samples (the total CNSC, serotype A and serotype D), we also separately analyzed the three major molecular types (VNI, VNII and VNB) within serotype A, two subclades (VNBI and VNBII) within VNB as well as the two genotype clusters closely related to the two most dominant STs (ST5 and ST93) in the global sample.

In the phylogenetic incompatibility test, we found that none of the 10 samples showed 100% phylogenetic compatibility (Table 8). Specifically, 4 (the total CNSC, serotype A, serotype D and clade VNI) of the 10 analyzed datasets showed no phylogenetic compatibility among the seven loci, which was consistent with the evidence of recombination among all 21 pairs of loci within each of the four samples. For VNII-, VNBI- and ST5-associated genotype groups, 16 of the 21 pairwise loci combinations were phylogenetically incompatible, with only five pairs (23.8%) being phylogenetically compatible. For the VNB and VNBII datasets, 19 of the 21 pairwise loci were phylogenetically incompatible. For the

ST93-associated genotype group, 5 of the 21 pairs showed phylogenetic incompatibility, which was also consistent with the evidence for recombination in this sample.

Table 8. Summary of genotypic diversity and phylogenetic incompatibility based on the sequence type and molecular type information assigned by the MLST database. ** $p < 0.01$; *** $p < 0.001$.

	Number	Phylogenetic Compatibility (% of 21 Pairs)	Index of Association
Total	657	0	0.75 ***
Serotype A	441	0	0.69 ***
VNI	296	0	0.83 ***
ST5-associated	49	23.8%	0.37 **
ST93-associated	20	76.2%	0.19
VNII	43	23.8%	0.44 ***
VNB	102 *	9.5%	0.58 ***
VNBI	93 *	23.8%	0.19 ***
VNBII	59 *	9.5%	0.99 ***
Serotype D	158	0	0.30 ***

* The number of STs analyzed here for VNB (102) was based on the molecular type assignment in the MLST database. In the MLST database, only 24 and 9 STs were assigned to VNBI and VNBII, respectively. The numbers are too small for effective analyses. Instead, in our recombination analyses for VNBI and VNBII, we extracted the numbers of STs for VNBI (93) and VNBII (59) from our phylogenetic results of the concatenated genes in Figure 1.

Linkage disequilibrium analyses revealed that in nine of the ten samples, the null hypothesis of random recombination was rejected (Table 8). The only exception was the ST93-associated genotype group, where the null hypothesis of random recombination was not rejected, likely due to the small sample size and the lack of statistical power to reject the null hypothesis. However, variable numbers of pairs of loci within each of the ten samples showed no significant deviation from those expected under the random recombination hypothesis (Supplementary Table S7). For example, in the VNII sample, 4 of the 21 loci pairs had observed genotype frequencies not significantly different from random recombination, with all 4 involving the *IGS1* locus. In the VNB sample, 9 of the 21 loci pairs had observed genotype frequencies not significantly different from random recombination. Interestingly, while no evidence for linkage equilibrium across all 21 loci pairs was observed in the ST93-associated genotype cluster, there was abundant evidence for linkage equilibrium between pairs of loci within the ST5-associated genotype cluster (Table 9). The complete allelic profiles for these clusters are shown in Supplementary Tables S8 and S9.

Table 9. Linkage disequilibria between pairs of loci in two sequence type clusters of *C. neoformans*.

ST93 Cluster	Loci	N (# of Alleles)	CAP59	GPD1	IGS1	LAC1	PLB1	SOD1	URA5
	CAP59	(1)		<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	GPD1	(6)	N/A		<0.001	<0.001	<0.001	<0.001	<0.001
	IGS1	(4)	N/A	0.0363		<0.001	<0.001	<0.001	<0.001
	LAC1	(5)	N/A	0.0501	0.0336		<0.001	<0.001	<0.001
	PLB1	(5)	N/A	0.0973	0.2239	0.2198			<0.001
	SOD1	(3)	N/A	0.0347	0.0058	0.1750	0.0445		<0.001
	URA5	(5)	N/A	0.0942	0.2193	0.2754	0.2378	0.1907	
ST5 cluster									
	CAP59	(7)		0.402	0.402	0.402	0.402	0.402	0.402
	GPD1	(6)	0.0158		0.938	0.938	0.938	0.938	0.938
	IGS1	(11)	0.1506	0.0844		0.002	0.002	0.002	0.002
	LAC1	(7)	0.3829	0.3525	0.0265		0.013	0.013	0.013
	PLB1	(8)	0.2611	0.2486	0.0770	0.5576		0.768	0.768
	SOD1	(7)	0.0650	0.0705	0.0173	0.1190	0.1314		0.024
	URA5	(8)	0.2110	0.0329	0.2518	0.2640	0.1588	0.1501	

n = number of allele types recorded for each gene for the two analyzed samples. Numbers in the bottom left part of the table indicate the linkage disequilibrium D values between pairs of loci. Numbers in the top right part of the table indicate corresponding p values. $p < 0.05$ indicates a rejection of the null hypothesis that the two loci are in linkage equilibrium.

4. Discussion

This study analyzed the genetic structure of geographic populations of the *C. neoformans* complex based on published multilocus sequence data. We focused on 296 STs with available geographical data. Our analyses included a robust global population of CNSC, with 4200 isolates originating across 31 countries and four continents (South and North America combined as one continent). Of the 296 sequence types with geographical data, 244 (82%) were sampled from a single continent, with 24 of these 244 STs found across multiple countries within a continent. The remaining 18% STs were distributed across multiple continents, with 28 STs (9%) found in two continents, 9 STs (3%) among three continents and 15 STs (5%) across all four continents. The broad distributions of multiple sequence types across multiple continents and countries are consistent with the recent and frequent gene flow in CNSC. Contemporary factors such as wind, animal and human migrations and other anthropogenic activities could all have facilitated the dispersals of genes and genotypes, causing wide distributions of certain genotypes [8,59–61].

However, our population genetic analyses revealed statistically significant differentiations among continental and national populations of CNSC. At the whole CNSC level, the observed genetic differentiations were contributed by differences in the distributions of the four molecular types and by the localized clonal expansion of specific sequence types. Indeed, evidence for the clonal expansion of specific genotypes was found for all molecular types. For example, the most abundantly collected ST in the analyzed data was ST5, represented by 1332 isolates. The second most abundant was ST93, represented by 460 isolates. Even though both were found in all four continents, ST5 and ST93 were mainly found in Asia (1211 of the 1332 isolates) and the Americas (224 of the 460 isolates), respectively. Among the serotype D (VNIV) isolates, ST160 was the most recorded sequence type, representing 78 isolates across three continents. Localized clonal expansion can significantly skew the allele and genotype frequencies and contribute to observed genetic differences among geographic populations [62]. Thus, we analyzed clone-corrected samples where only one representative from each country was included for analyses. Using the clone-corrected samples, the amount of contribution due to geographic separation to the total genetic variance reduced by 70–93% at the continental level and by 60–93% at the country level for the total CNSC, the serotype A and the serotype D samples, respectively. This result is consistent with the presence of indigenous genetic variations within most national and continental populations, likely due to historical differentiations.

Interestingly, out of the 529 isolates representing 109 STs collected within Africa, only 2 isolates belonged to serotype D (VNIV). Our meta-analysis was consistent with an earlier study that analyzed the molecular types of 505 isolates from Africa and found that none of them belonged to molecular type VNIV [61]. However, these results do not mean that VNIV is unimportant in Africa. For example, one study that analyzed 252 isolates from South Africa identified 5 cases of molecular type VNIV [36]. However, those isolates were not genotyped using MLST. At present, Africa accounts for the greatest global burden of HPC infection [63,64] and contains the most genetically diverse population of serotype A [36], including the relatively frequent distributions of both mating types. The high genetic diversity of serotype A strains in Africa has led to the "Out of Africa" hypothesis for the origin and spread of serotype A [65]. Interestingly, four STs from Africa that were originally identified as belonging to VNI were clustered to the basal clade of VNIV (Figure 1). These STs likely represent some of the ancestor genotypes of VNIV or recent hybrids between the VNI and VNIV strains.

Among the 72 serotype D (VNIV) STs, 65 (90%) were represented in Europe. In comparison, only 23% of serotype A STs were found in Europe. The results are consistent with multiple studies reporting a relatively high prevalence of serotype D within Europe [66]. The relatively broad distributions of both serotype A and serotype D strains in Europe are likely the main contributors to the frequent observations of serotype AD hybrids within Europe [66].

Although there is an abundant *C. neoformans* population in North America [67,68], only 12 isolates have been analyzed using the ISHAM MLST scheme. Such a lack of MLST data from North America was not due to the lack of samples for analyses. Indeed, between 1992 and 1994, the US CDC conducted a large-scale surveillance of the agents of cryptococcosis [69]. Those isolates were genotyped using random amplified polymorphic DNA and/or multilocus enzyme electrophoresis, the commonly used molecular markers at that time, revealing abundant genetic variations, including at least three independent hybridizing events between serotypes A and D [8,59,69,70]. However, those strains, as well as many strains isolated afterwards from North America, have not been genotyped using the ISHAM MLST scheme, which was published in 2009. It would be very interesting to analyze the North American population of *C. neoformans* and to compare them with those from other parts of the world. In contrast to the lack of MLST data from North America, there is a large representation of *C. neoformans* from China, making that population of CNSC one of the best for understanding fine-scale spatial and temporal structures of CNSC.

In this study, we performed a phylogenetic analysis of all STs based on their concatenated DNA sequences. Overall, the phylogenetic results were consistent with the molecular type designations for most isolates based on PCR fingerprinting, AFLP and/or PCR-RFLP of the URA5 gene fragment. However, we found that the placements of a small number of STs in the phylogenetic trees were inconsistent with their original molecular type designations (Figure 1). Such inconsistencies were found among molecular types within serotype A as well as between serotype A and serotype D. Among these inconsistently placed STs, 15 contained mixtures of alleles from serotype A and serotype D (Table 4). These STs were likely recombinants derived from the hybridization between serotypes A and D strains. After hybridization, either meiotic or mitotic recombination could have led to a loss of heterozygosity (LOH) to generate the haploid recombinant genotypes observed here. Indeed, LOH in serotype AD hybrids has been observed through both meiosis and mitosis, with environmental stress facilitating LOH [71,72]. The observed recombination is also consistent with the results from nuclear and mitochondrial genome phylogenetic comparisons, where the mitochondrial genome-based phylogeny showed several differences with that based on nuclear genome-based phylogeny within CNSC [73]. Furthermore, some of these STs had distinct alleles and showed significant divergence with the main serotype A and the main serotype D genotypes and clades, a result suggesting the existence of distinct novel lineages within CNSC (Figure 1, Supplementary Figures S1–S7). Finally, our phylogenetic analyses successfully placed 51 previously undefined STs into the phylogenetic framework and revealed their possible origins. The presence of these hybrids, as well as many intermediate STs with ambiguous molecular type assignments in the phylogeny, supports the continued use of CNSC for this group of fungal pathogens [9].

CNSC has been shown to reproduce predominantly asexually in nature, but evidence for recombination has been reported for both the serotype A and serotype D populations [28,74–76]. Sexual reproduction can accelerate adaptation to diverse environments and remove deleterious mutations more effectively than asexual reproduction [77]. Across all four major molecular types (VNI, VNII, VNB and VNIV) as well as two subtypes of VNB (VNBI and VNBII), we found clear evidence of recombination within each. Significantly, evidence for recombination was also found in two presumed "clonal ST clusters" within VNI. The observed results suggest that, even in geographic populations of CNSC dominated by a few STs, recombination is still possible. Such recombination could be achieved through same-sex mating or opposite-sex mating, as suggested previously when regional populations of CNSC were found to contain evidence of recombination, despite only one mating type (MATα) being found in those analyzed samples [74,75].

In conclusion, analyses of the published MLST data for CNSC allowed us to quantify the genetic diversity within and among geographic populations of this important human pathogenic yeast. Our analyses revealed evidence for historical geographic differentiations of CNSC, both at the whole CNSC level as well as within the populations of serotypes A and D. Not surprisingly, evidence for the clonal expansion of many STs was found at

both the local population level as well as across countries and continents, suggesting the potentially important roles of recent anthropogenic activities in the dispersals of alleles and genotypes of CNSC. Importantly, we found evidence for recombination within all molecular types, including at least two presumed "clonal ST clusters" within VNI. The results indicate the diverse methods of CNSC reproduction in nature. While a large number of isolates were analyzed in this study for population genetic patterns, about 55% of the 657 STs in the ISHAM MLST database were not included for geographic structure analyses due to the lack of geographic location information associated with these 361 STs. In the future, authors should be required to submit the metadata associated with each isolate and each sequence type that they publish in their MLST study. Additional information on these STs, as well as more MLST data, especially from under-reported regions such as North America, will provide a more comprehensive understanding of the global population structure of CNSC and help develop more realistic models of global cryptococcal threat predictions and management strategies against cryptococcosis [78].

References

1. Hagen, F.; Khayhan, K.; Theelen, B.; Kolecka, A.; Polacheck, I.; Sionov, E.; Falk, R.; Parnmen, S.; Lumbsch, H.T.; Boekhout, T. Recognition of seven species in the *Cryptococcus gattii/Cryptococcus neoformans* species complex. *Fungal Genet. Biol.* **2015**, *78*, 16–48. [CrossRef] [PubMed]
2. Lazera, M.S.; Cavalcanti, M.A.S.; Londero, A.T.; Trilles, L.; Nishikawa, M.M.; Wanke, B. Possible primary ecological niche of *Cryptococcus neoformans*. *Med. Mycol.* **2000**, *38*, 379–383. [CrossRef] [PubMed]
3. Xu, J.; Manosuthi, W.; Banerjee, U.; Zhu, L.-P.; Chen, J.H.; Kohno, S.; Izumikawa, K.; Chen, Y.H.; Sungkanuparph, S.; Harrison, T.S.; et al. Cryptococcosis in Asia. In *Cryptococcus: From Human Pathogen to Model Organism*; Heitman, J., Kwon-Chung, J., Perfect, J., Casadevall, A., Eds.; ASM Press: Washington, DC, USA, 2011; pp. 287–298.

4.　Randhawa, H.S.; Kowshik, T.; Chowdhary, A.; Preeti Sinha, K.; Khan, Z.U.; Sun, S.; Xu, J. The expanding host tree species spectrum of *Cryptococcus gattii and Cryptococcus neoformans* and their isolations from surrounding soil in India. *Med. Mycol.* **2008**, *46*, 823–833. [CrossRef] [PubMed]

5.　Negroni, R. Cryptococcosis. *Clin. Dermatol.* **2012**, *30*, 599–609. [CrossRef]

6.　Rajasingham, R.; Smith, R.M.; Park, B.J.; Jarvis, J.N.; Govender, N.P.; Chiller, T.M.; Denning, D.W.; Loyse, A.; Boulware, D.R. Global burden of disease of HIV-associated cryptococcal meningitis: An updated analysis. *Lancet Infect. Dis.* **2017**, *17*, 873–881. [CrossRef]

7.　Gushiken, C.A.; Saharia, K.K.; Baddley, W.J. Cryptococcosis. *Infect. Dis. Clin. N. Am.* **2021**, *35*, 493–514. [CrossRef]

8.　Hong, N.; Chen, M.; Xu, J. Molecular markers reveal epidemiological patterns and evolutionary histories of the human pathogenic *Cryptococcus*. *Front. Cell. Infect. Microbiol.* **2021**, *11*, 398. [CrossRef]

9.　Kwon-Chung, K.J.; Bennett, J.E.; Wickes, B.L.; Meyer, W.; Cuomo, C.A.; Wollenburg, K.R.; Bicanic, T.A.; Castañeda, E.; Chang, Y.C.; Chen, J.; et al. The Case for Adopting the "Species Complex" Nomenclature for the Etiologic Agents of Cryptococcosis. *mSphere* **2017**, *2*, e00357–16. [CrossRef]

10.　Fernandes, K.E.; Brockway, A.; Haverkamp, M.; Cuomo, C.A.; Ogtrop, F.V.; Perfect, J.R.; Carter, D.A. Phenotypic Variability Correlates with Clinical Outcome in *Cryptococcus* Isolates Obtained from Botswanan HIV/AIDS Patients. *mBio* **2018**, *9*, e02016–e02018. [CrossRef]

11.　Meyer, W.; Aanensen, D.M.; Boekhout, T.; Cogliati, M.; Diaz, M.R.; Esposto, M.C.; Fisher, M.; Gilgado, F.; Hagen, F.; Kaocharoen, S.; et al. Consensus multi-locus sequence typing scheme for *Cryptococcus neoformans* and *Cryptococcus gattii*. *Med. Mycol.* **2009**, *47*, 561–570. [CrossRef]

12.　Meyer, W. International Fungal Multi Locus Sequence Typing Database. Available online: https://mlst.mycologylab.org/page/Home1 (accessed on 31 January 2022).

13.　Ma, Y. Molecular epidemiology and antifungal susceptibilities of *Cryptococcus* species isolates from HIV and non-HIV patients in Southwest China. *Eur. J. Clin. Microbiol. Infect. Dis.* **2021**, *40*, 287–295. [CrossRef]

14.　Andrade-Silva, L.E.; Ferreira-Paim, K.; Ferreira, T.B.; Vilas-Boas, A.; Mora, D.J.; Manzato, V.M.; Fonseca, F.M.; Buosi, K.; Andrade-Silva, J.; Prudente, B.D.S.; et al. Genotypic analysis of clinical and environmental *Cryptococcus neoformans* isolates from Brazil reveals the presence of VNB isolates and a correlation with biological factors. *PLoS ONE* **2018**, *13*, e0193237. [CrossRef] [PubMed]

15.　Beale, M.A.; Sabiiti, W.; Robertson, E.J.; Fuentes-Cabrejo, K.M.; O'Hanlon, S.J.; Jarvis, J.N.; Loyse, A.; Meintjes, G.; Harrison, T.S.; May, R.C.; et al. Genotypic Diversity Is Associated with Clinical Outcome and Phenotype in Cryptococcal Meningitis across Southern Africa. *PLoS Negl. Trop. Dis.* **2015**, *9*, e0003847. [CrossRef]

16.　Brito-Santos, F.; Trilles, L.; Firacative, C.; Wanke, B.; Carvalho-Costa, F.A.; Nishikawa, M.M.; Campos, J.P.; Junqueira, A.C.V.; Souza, A.C.; Lazra, M.D.S.; et al. Indoor Dust as a Source of Virulent Strains of the Agents of Cryptococcosis in the Rio Negro Micro-Region of the Brazilian Amazon. *Microorganisms* **2020**, *8*, 682. [CrossRef] [PubMed]

17.　Chen, M.; Wang, Y.; Li, Y.; Hong, N.; Zhu, X.; Pan, W.; Liao, W.; Xu, J.; Du, J.; Chen, J. Genotypic diversity and antifungal susceptibility of environmental isolates of *Cryptococcus neoformans* from the Yangtze River Delta region of East China. *Med. Mycol.* **2021**, *59*, 653–663. [CrossRef] [PubMed]

18.　Chidebelu, P.E.; Nweze, E.I.; Meis, J.F.; Cogliati, M.; Hagen, F. Multi-locus sequence typing reveals genotypic similarity in Nigerian *Cryptococcus neoformans* AFLP1/VNI of environmental and clinical origin. *J. Med. Microbiol.* **2021**, *70*, 001440. [CrossRef]

19.　Cogliati, M.; Desnos-Ollivier, M.; McCormick-Smith, I.; Rickerts, V.; Ferreira-Paim, K.; Meyer, W.; Boekhout, T.; Hagen, F.; Theelen, B.; Incio, J.; et al. Genotypes and population genetics of *Cryptococcus neoformans* and *Cryptococcus gattii* species complexes in Europe and the mediterranean area. *Fungal Genet. Biol.* **2019**, *129*, 16–29. [CrossRef]

20.　Cogliati, M.; Zamfirova, R.R.; Tortorano, A.M.; Viviani, M.A.; Network, T.F.C. Molecular epidemiology of Italian clinical *Cryptococcus neoformans var. grubii* isolates. *Med. Mycol.* **2013**, *51*, 499–506. [CrossRef]

21.　Danesi, P.; Firacative, C.; Cogliati, M.; Otranto, D.; Capelli, G.; Meyer, W. Multilocus sequence typing (MLST) and M13 PCR fingerprinting revealed heterogeneity amongst *Cryptococcus* species obtained from Italian veterinary isolates. *FEMS Yeast Res.* **2014**, *14*, 897–909. [CrossRef]

22.　Day, J.N.; Qihui, S.; Thanh, L.T.; Trieu, P.H.; Van, A.D.; Thu, N.H.; Chau, T.T.H.; Lan, N.P.H.; Chau, N.V.V.; Ashton, P.M.; et al. Comparative genomics of *Cryptococcus neoformans* var. *grubii* associated with meningitis in HIV infected and uninfected patients in Vietnam. *PLoS Negl. Trop. Dis.* **2017**, *11*, e0005628. [CrossRef]

23.　Desnos-Ollivier, M.; Patel, S.; Raoux-Barbot, D.; Heitman, J.; Dromer, F. Cryptococcosis Serotypes Impact Outcome and Provide Evidence of *Cryptococcus neoformans* Speciation. *mBio* **2015**, *6*, e00311. [CrossRef] [PubMed]

24.　Dou, H.T.; Xu, Y.C.; Wang, H.Z.; Li, T.S. Molecular epidemiology of *Cryptococcus neoformans* and *Cryptococcus gattii* in China between 2007 and 2013 using multilocus sequence typing and the DiversiLab system. *Eur. J. Clin. Microbiol. Infect. Dis.* **2015**, *34*, 753–762. [CrossRef] [PubMed]

25.　Dou, H.; Wang, H.; Xie, S.; Chen, X.; Xu, Z.; Xu, Y. Molecular characterization of *Cryptococcus neoformans* isolated from the environment in Beijing, China. *Med. Mycol.* **2017**, *55*, 737–747. [CrossRef] [PubMed]

26.　Fan, X.; Xiao, M.; Chen, S.; Kong, F.; Dou, H.T.; Wang, H.; Xiao, Y.L.; Kang, M.; Sun, Z.Y.; Hu, Z.D.; et al. Predominance of *Cryptococcus neoformans var. grubii* multilocus sequence type 5 and emergence of isolates with non-wild-type minimum inhibitory concentrations to fluconazole: A multi-centre study in China. *Clin. Microbiol. Infect.* **2016**, *22*, P887.E1–P887.E9. [CrossRef]

27. Fang, L.F.; Zhang, P.P.; Wang, J.; Yang, Q.; Qu, T.T. Clinical and microbiological characteristics of cryptococcosis at an university hospital in China from 2013 to 2017. *Braz. J. Infect. Dis.* **2020**, *24*, 7–12. [CrossRef] [PubMed]

28. Ferreira-Paim, K.; Andrade-Silva, L.; Fonseca, F.M.; Ferreira, T.B.; Mora, D.J.; Andrade-Silva, J.; Khan, A.; Dao, A.; Reis, E.C.; Almeida, M.T.; et al. MLST-Based Population Genetic Analysis in a Global Context Reveals Clonality amongst *Cryptococcus neoformans var. grubii* VNI Isolates from HIV Patients in Southeastern Brazil. *PLoS Negl. Trop. Dis.* **2017**, *11*, e0005223. [CrossRef]

29. Fu, Y.; Xu, M.; Zhou, H.; Yao, Y.; Zhou, J.; Pan, Z. Microbiological and clinical characteristics of cryptococcemia: A retrospective analysis of 85 cases in a Chinese hospital. *Med. Mycol.* **2020**, *58*, 478–484. [CrossRef]

30. Hatthakaroon, C.; Pharkjaksu, S.; Chongtrakool, P.; Suwannakarn, K.; Kiratisin, P.; Ngamskulrungroj, P. Molecular epidemiology of cryptococcal genotype VNIc/ST5 in Siriraj Hospital, Thailand. *PLoS ONE* **2017**, *12*, e0173744. [CrossRef]

31. Hong, N.; Chen, M.; Xu, N.; Al-Hatmi, A.M.S.; Zhang, C.; Pan, W.H.; Hagen, F.; Boekhout, T.; Xu, J.; Zou, X.B.; et al. Genotypic diversity and antifungal susceptibility of *Cryptococcus neoformans* isolates from paediatric patients in China. *Mycoses* **2019**, *62*, 171–180. [CrossRef]

32. Kaocharoen, S.; Ngamskulrungroj, P.; Firacative, C.; Trilles, L.; Piyabongkarn, D.; Banlunara, W.; Poonwan, N.; Chaiprasert, A.; Meyer, W.; Chindamporn, A. Molecular epidemiology reveals genetic diversity amongst isolates of the *Cryptococcus neoformans/C. gattii* species complex in Thailand. *PLoS Negl. Trop. Dis.* **2013**, *7*, e2297. [CrossRef]

33. Khayhan, K.; Hagen, F.; Pan, W.; Simwami, S.; Fisher, M.C.; Wahyuningsih, R.; Chakrabarti, A.; Chowdhary, A.; Ikeda, R.; Taj-Aldeen, S.J.; et al. Geographically structured populations of *Cryptococcus neoformans* Variety grubii in Asia correlate with HIV status and show a clonal population structure. *PLoS ONE* **2013**, *8*, e72222. [CrossRef] [PubMed]

34. Mihara, T.; Izumikawa, K.; Kakeya, H.; Ngamskulrungroj, P.; Umeyama, T.; Takazono, T.; Tashiro, M.; Nakamura, S.; Imamura, Y.; Miyazaki, T.; et al. Multilocus sequence typing of *Cryptococcus neoformans* in non-HIV associated cryptococcosis in Nagasaki, Japan. *Med. Mycol.* **2013**, *51*, 252–260. [CrossRef] [PubMed]

35. Moslem, M.; Fatahinia, M.; Kiasat, N.; Mahmoudabadi, A.Z. Genotypic diversity of Iranian *Cryptococcus neoformans* using multilocus sequence typing (MLST) and susceptibility to antifungals. *Mol. Biol. Rep.* **2021**, *48*, 4201–4208. [CrossRef] [PubMed]

36. Naicker, S.D.; Magobo, R.E.; Maphanga, T.G.; Firacative, C.; van Schalkwyk, E.; Monroy-Nieto, J.; Bowers, J.; Engelthaler, D.M.; Shuping, L.; Meyer, W.; et al. Genotype, Antifungal Susceptibility, and Virulence of Clinical South African *Cryptococcus neoformans* Strains from National Surveillance, 2005–2009. *J. Fungi* **2021**, *7*, 338. [CrossRef]

37. Park, S.H.; Kim, M.; Joo, S.I.; Hwang, S.M. Molecular Epidemiology of Clinical *Cryptococcus neoformans* Isolates in Seoul, Korea. *Mycobiology* **2014**, *42*, 73–78. [CrossRef]

38. Prakash, A.; Sundar, G.; Sharma, B.; Hagen, F.; Meis, J.F.; Chowdhary, A. Genotypic diversity in clinical and environmental isolates of *Cryptococcus neoformans* from India using multilocus microsatellite and multilocus sequence typing. *Mycoses* **2020**, *63*, 284–293. [CrossRef]

39. Reis, R.S.; Bonna, I.C.F.; Antonio, I.; Pereira, S.A.; Nascimento, C.; Ferraris, F.K.; Brito-Santos, F.; Ferreira Gremi√£o, I.D.; Trilles, L. *Cryptococcus neoformans* VNII as the Main Cause of Cryptococcosis in Domestic Cats from Rio de Janeiro, Brazil. *J. Fungi* **2021**, *7*, 980. [CrossRef]

40. Rocha, D.F.S.; Cruz, K.S.; Santos, C.; Menescal, L.S.F.; Neto, J.; Pinheiro, S.B.; Silva, L.M.; Trilles, L.; Braga de Souza, J.V. MLST reveals a clonal population structure for *Cryptococcus neoformans* molecular type VNI isolates from clinical sources in Amazonas, Northern-Brazil. *PLoS ONE* **2018**, *13*, e0197841. [CrossRef]

41. Samarasinghe, H.; Aljohani, R.; Jimenez, C.; Xu, J. Fantastic yeasts and where to find them: The discovery of a predominantly clonal *Cryptococcus deneoformans* population in Saudi Arabian soils. *FEMS Microbiol. Ecol.* **2019**, *95*, fiz122. [CrossRef]

42. Selb, R.; Fuchs, V.; Graf, B.; Hamprecht, A.; Hogardt, M.; Sedlacek, L.; Schwarz, R.; Idelevich, E.A.; Becker, S.L.; Held, J.; et al. Molecular typing and in vitro resistance of *Cryptococcus neoformans* clinical isolates obtained in Germany between 2011 and 2017. *Int. J. Med. Microbiol.* **2019**, *309*, 151336. [CrossRef]

43. Silva, L.M.; Ferreira, W.A.; Filho, R.; Lacerda, M.V.G.; Ferreira, G.M.A.; Saunier, M.N.; Macedo, M.M.; Cristo, D.A.; Alves, M.J.; Jackisch-Matsuura, A.B.; et al. New ST623 of *Cryptococcus neoformans* isolated from a patient with non-Hodgkin's lymphoma in the Brazilian Amazon. *Ann. Clin. Microbiol. Antimicrob.* **2020**, *19*, 20. [CrossRef] [PubMed]

44. Simwami, S.P.; Khayhan, K.; Henk, D.A.; Aanensen, D.M.; Boekhout, T.; Hagen, F.; Brouwer, A.E.; Harrison, T.S.; Donnelly, C.A.; Fisher, M.C. Low diversity *Cryptococcus neoformans* variety *grubii* multilocus sequence types from Thailand are consistent with an ancestral African origin. *PLoS Pathog.* **2011**, *7*, e1001343. [CrossRef] [PubMed]

45. Takahashi, Y.; Osawa, R.; Kubota, Y.; Fujii, M.; Matsuda, N.; Watanabe, N.; Watari, T.; Otsuka, Y.; Hosokawa, N. Early diagnosis of *Cryptococcus neoformans var. grubii* meningitis using multiplex PCR assay in an immunocompetent patient. *J. Infect. Chemother.* **2021**, *27*, 1765–1768. [CrossRef] [PubMed]

46. Umeyama, T.; Ohno, H.; Minamoto, F.; Takagi, T.; Tanamachi, C.; Tanabe, K.; Kaneko, Y.; Yamagoe, S.; Kishi, K.; Fujii, T.; et al. Determination of epidemiology of clinically isolated *Cryptococcus neoformans* strains in Japan by multilocus sequence typing. *Jpn. J. Infect. Dis.* **2013**, *66*, 51–55. [CrossRef] [PubMed]

47. van de Wiele, N.; Neyra, E.; Firacative, C.; Gilgado, F.; Serena, C.; Bustamante, B.; Meyer, W. Molecular Epidemiology Reveals Low Genetic Diversity among *Cryptococcus neoformans* Isolates from People Living with HIV in Lima, Peru, during the Pre-HAART Era. *Pathogens* **2020**, *9*, 665. [CrossRef]

48. Velez, N.; Escandon, P. Multilocus sequence typing (MLST) of clinical and environmental isolates of *Cryptococcus neoformans* and *Cryptococcus gattii* in six departments of Colombia reveals high genetic diversity. *Rev. Soc. Bras. Med. Trop.* **2020**, *53*, e20190422. [CrossRef]

49. Wongsuk, T.; Homkaew, A.; Faksri, K.; Thongnak, C. Multi-locus Sequence Typing and Whole Genome Sequence Analysis of *Cryptococcus neoformans* Isolated from Clinical Specimens in Vajira Hospital, Bangkok, Thailand. *Mycopathologia* **2020**, *185*, 503–514. [CrossRef]

50. Wu, S.Y.; Lei, Y.; Kang, M.; Xiao, Y.L.; Chen, Z.X. Molecular characterisation of clinical *Cryptococcus neoformans* and *Cryptococcus gattii* isolates from Sichuan province, China. *Mycoses* **2015**, *58*, 280–287. [CrossRef]

51. Xess, I.; Pandey, M.; Dabas, Y.; Agarwal, R.; Das, S.; Srivastava, P.M.V.; Thakur, R.; Sharma, S.; Mani, P.; Biswas, A.; et al. Multilocus Sequence Typing of Clinical Isolates of *Cryptococcus* from India. *Mycopathologia* **2021**, *186*, 199–211. [CrossRef]

52. Xu, X.; Du, P.; Wang, H.; Yang, X.; Liu, T.; Zhang, Y.; Wang, Y. Clinical characteristics, *Cryptococcus neoformans* genotypes, antifungal susceptibility, and outcomes in human immunodeficiency virus-positive patients in Beijing, China. *J. Int. Med. Res.* **2021**, *49*, 3000605211016197. [CrossRef]

53. Yang, C.; Bian, Z.; Blechert, O.; Deng, F.; Chen, H.; Li, Y.; Yang, Y.; Chen, M.; Zhan, P. High Prevalence of HIV-Related Cryptococcosis and Increased Resistance to Fluconazole of the *Cryptococcus neoformans* Complex in Jiangxi Province, South Central China. *Front. Cell Infect. Microbiol.* **2021**, *11*, 723251. [CrossRef] [PubMed]

54. Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027. [CrossRef] [PubMed]

55. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef]

56. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [CrossRef]

57. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research–An update. *Bioinformatics* **2012**, *28*, 2537–2539. [CrossRef] [PubMed]

58. Agapow, P.-M.; Burt, A. Indices of multilocus linkage disequilibrium. *Mol. Ecol. Notes* **2001**, *1*, 101–102. [CrossRef]

59. Xu, J.; Vilgalys, R.; Mitchell, T.G. Multiple gene genealogies reveal recent dispersion and hybridization in the human pathogenic fungus *Cryptococcus neoformans*. *Mol. Ecol.* **2000**, *9*, 1471–1481. [CrossRef] [PubMed]

60. Chen, Y.H.; Yu, F.; Bian, Z.Y.; Hong, J.M.; Zhang, N.; Zhong, Q.S.; Hang, Y.P.; Xu, J.; Hu, L.H. Multilocus Sequence Typing Reveals both Shared and Unique Genotypes of *Cryptococcus neoformans* in Jiangxi Province, China. *Sci. Rep.* **2018**, *8*, 1495. [CrossRef]

61. Cogliati, M. Global Molecular Epidemiology of *Cryptococcus neoformans* and *Cryptococcus gattii:* An Atlas of the Molecular Types. *Scientifica* **2013**, *2013*, 675213. [CrossRef]

62. Xu, J. Fundamentals of Fungal Molecular Population Genetic Analyses. *Curr. Issues Mol. Biol.* **2006**, *8*, 75–90.

63. Hakim, J.G.; Gangaidzo, I.T.; Heyderman, R.S.; Mielke, J.; Mushangi, E.; Taziwa, A.; Robertson, V.J.; Musvaire, P.; Mason, P.R. Impact of HIV infection on meningitis in Harare, Zimbabwe: A prospective study of 406 predominantly adult patients. *AIDS* **2000**, *14*, 1401–1407. [CrossRef] [PubMed]

64. Holmes, C.B.; Losina, E.; Walensky, R.P.; Yazdanpanah, Y.; Freedberg, K.A. Review of Human Immunodeficiency Virus Type 1-Related Opportunistic Infections in Sub-Saharan Africa. *Clin. Infect. Dis.* **2003**, *36*, 652–662. [CrossRef] [PubMed]

65. Litvintseva, A.P.; Carbone, I.; Rossouw, J.; Thakur, R.; Govender, N.P.; Mitchell, T.G. Evidence that the Human Pathogenic Fungus *Cryptococcus neoformans* var. *grubii* May Have Evolved in Africa. *PLoS ONE* **2011**, *6*, e19688. [CrossRef]

66. Cogliati, M.; Roger, F.; Meyer, W.; Robert, V.; Bertout, S. New multilocus sequence typing primers to enable genotyping of AD hybrids within the *Cryptococcus neoformans* species complex. *Med. Mycol.* **2020**, *58*, 1005–1009. [CrossRef]

67. Litvintseva, A.P.; Kestenbaum, L.; Vilgalys, R.; Mitchell, T.G. Comparative Analysis of Environmental and Clinical Populations of *Cryptococcus neoformans*. *J. Clin. Microbiol.* **2005**, *43*, 556–564. [CrossRef]

68. Mitchell, T.G.; Perfect, J.R. Cryptococcosis in the era of AIDS–100 years after the discovery of *Cryptococcus neoformans*. *Clin. Microbiol. Rev.* **1995**, *8*, 515–548. [CrossRef]

69. Brandt, M.E.; Hutwagner, L.C.; Klug, L.A.; Baughman, W.S.; Rimland, D.; Graviss, E.A.; Hamill, R.J.; Thomas, C.; Pappas, P.G.; Reingold, A.L.; et al. and The Cryptococcal Disease Active Surveillance Group.1996. Molecular subtype distribution of *Cryptococcus neoformans* in four areas of the United States. *J. Clin. Microbiol.* **1996**, *34*, 912–917. [CrossRef]

70. Xu, J.; Luo, G.; Vilgalys, R.; Brandt, M.E.; Mitchell, T.G. Multiple origins of hybrid strains of *Cryptococcus neoformans* with serotype AD. *Microbiology* **2002**, *148*, 203–212. [CrossRef]

71. Dong, K.; You, M.; Xu, J. Genetic changes in experimental populations of a hybrid in the *Cryptococcus neoformans* species complex. *Pathogens* **2020**, *9*, 3. [CrossRef]

72. Samarasinghe, Y.A.H.; Vogan, A.A.; Pum, N.; Xu, J. Patterns of allele distribution in a hybrid population of the *Cryptococcus neoformans* species complex. *Mycoses* **2020**, *63*, 275–283. [CrossRef]

73. Wang, Y.; Xu, J. Mitochondrial Genome Polymorphisms in the Human Pathogenic Fungus *Cryptococcus neoformans*. *Front. Microbiol.* **2020**, *11*, 706. [CrossRef] [PubMed]

74. Xu, J.; Mitchell, T.G. Comparative gene genealogical analyses of strains of serotype AD identify recombination in populations of serotypes A and D in the human pathogenic yeast *Cryptococcus neoformans*. *Microbiology* **2003**, *149*, 2147–2154. [CrossRef] [PubMed]

75. Hiremath, S.S.; Chowdhary, A.; Kowshik, T.; Randhawa, H.S.; Sun, S.; Xu, J. Long-distance dispersal and recombination in environmental populations of *Cryptococcus neoformans* var. *grubii* from India. *Microbiol. SGM* **2008**, *154*, 1513–1524. [CrossRef] [PubMed]

76. Halliday, C.L.; Carter, D.A. Clonal Reproduction and Limited Dispersal in an Environmental Population of *Cryptococcus neoformans* var. *gattii* Isolates from Australia. *J. Clin. Microbiol.* **2003**, *41*, 703–711. [CrossRef] [PubMed]

77. Xu, J. The prevalence and evolution of sex in microorganisms. *Genome* **2004**, *47*, 775–780. [CrossRef]

78. Xu, J. Assessing global fungal threats to humans. *mLife* **2022**, *1*, 223–240. [CrossRef]