Mycopathologia (2020) 185:503–514
https://doi.org/10.1007/s11046-020-00456-7

ORIGINAL ARTICLE

Multi-locus Sequence Typing and Whole Genome Sequence Analysis of *Cryptococcus neoformans* Isolated from Clinical Specimens in Vajira Hospital, Bangkok, Thailand

Thanwa Wongsuk · Anchalee Homkaew · Kiatichai Faksri · Chuphong Thongnak

Received: 18 February 2020 / Accepted: 12 May 2020 / Published online: 21 May 2020

Abstract The basidiomycete yeast *Cryptococcus neoformans* causes disease in immunocompromized patients. Whole genome sequencing (WGS) technology provides insights into the molecular epidemiology of *C. neoformans*. However, the number of such studies is limited. Here we used WGS and multilocus sequence typing (MLST) to determine the genetic diversity of *C. neoformans* isolates and genetic structures of their populations among patients admitted to a single hospital in Bangkok, Thailand. Seven isolates from six patients collected during 1 year were identified as *C. neoformans* sensu stricto according to colony morphology, microscopy, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and nucleotide sequence analysis of internal transcribed sequences. These isolates were sensitive to the antifungal drugs amphotericin B, fluconazole, 5-flucytosine, voriconazole, itraconazole and posaconazole and were mating type α and molecular type VNI. MLST analysis identified ST4, ST5 and ST6. We further employed WGS to determine the genetic diversity and relationships of *C. neoformans* isolated here combined with *C. neoformans* sequences data acquired from a public database (n = 42). We used the data to construct a phylogenetic tree. WGS provided additional genomics data and achieved high discriminatory power for identifying *C. neoformans* isolates isolated in Thailand. This report further demonstrates the applicability of WGS analysis for conducting molecular epidemiology and provides insight into the genetic diversity of *C. neoformans* isolates from one hospital in Thailand.

Keywords *Cryptococcus neoformans* · MLST · Whole genome sequence · MALDI-TOF MS

Handling Editor: Ferry Hagen.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11046-020-00456-7) contains supplementary material, which is available to authorized users.

T. Wongsuk · C. Thongnak (✉)
Department of Clinical Pathology, Faculty of Medicine Vajira Hospital, Navamindradhiraj University, 681 Samsen Road, Vajira District, Dusit, Bangkok 10300, Thailand
e-mail: chuphong@nmu.ac.th

A. Homkaew
Microbiology Laboratory, Department of Central Laboratory and Blood Bank, Faculty of Medicine Vajira Hospital, Navamindradhiraj University, Bangkok, Thailand

K. Faksri
Research and Diagnostic Center for Emerging Infectious Diseases, and Department of Microbiology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

Introduction

The basidiomycete *Cryptococcus neoformans* causes systemic disease and primarily affects the meninges of immunocompromized patients with underlying diseases, as well as healthy people [1, 2]. Risk factors for cryptococcosis include human immune deficiency virus (HIV) infection, as well as those with autoimmune disorders, malignancy and transplant-associated immunosuppression [3, 4].

Seven species have been recognized within the *Cryptococus neoformans*/gattii species complexes, namely *C. neoformans* (serotype A; genotype AFLP1/VNI, AFLP1A/VNB/VNII and AFLP1B/VNII), *C. deneoformans* (serotype D; genotype AFLP2/VNIV) and hybrids between both (serotype AD; genotype AFLP3/VNIII) form the *C. neoformans* species complex. *Cryptococcus gattii* s.s. (serotype B; genotype AFLP4/VGI), *C. deuterogattii* (serotype B; AFLP6; VGII), *C. bacillisporus* (serotype B and C; genotype AFLP5/VGIII), *C. tetragattii* (serotype C; genotype AFLP7/VGIV) and *C. decagattii* (serotype B; genotype AFLP10) [5–7].

Epidemiological surveys performed in Asia showed that *C. neoformans* causes the majority of cryptococcal infections [8–13]. In Thailand, the molecular epidemiology of *C. neoformans* isolated from human biological samples was studied using restriction fragment length polymorphism of *URA5* gene (*URA5*-RFLP), along with multilocus sequencing type analysis (MLST) [14–17]. Whole genome sequencing (WGS) has been applied to the study of the population structure of *C. neoformans* isolates in many countries [18–22]. However, only one study used WGS to analyze the population genetics based on WGS of *C. neoformans* in Thailand [18]. Thus, additional studies are required to comprehensively to determine the genetic diversity of *C. neoformans*. WGS analysis has been shown to have greater discriminatory power than MLST [23]. However, the number of studies comparing the performances of these two tools is low, particularly regarding genetic classifications.

To address this insufficiency in our knowledge of this important pathogen, here, we investigated the biological features and molecular characteristics of *C. neoformans* s.s isolated from patients admitted to our hospital. For this purpose, we employed nucleotide sequencing, in vitro drug-sensitivity profiling, mating type (MAT) analysis using multiplex PCR, and molecular typing using *URA5*-RFLP. Further, we employed MLST, WGS-based typing, and phylogenetic tree analysis to determine the genetic diversity and relationships among our isolates and those characterized by others worldwide.

Materials and Methods

Study Population

Seven isolates of *C. neoformans* from six clinical blood specimens and one sample of cerebrospinal fluid (CSF) were retrieved from the Microbiology Laboratory, Department of Central Laboratory and Blood Bank, Faculty of Medicine, Vajira Hospital, Navamindradhiraj University from September 1, 2018 to August 30, 2019. The isolates were from six patients (identified using our routine laboratory methods, including colony morphology, microscopy and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry [MALDI-TOF MS]) [24, 25]. These isolates were stored in Sabouraud dextrose broth with 20% glycerol at − 30 °C. Clinical data and samples were anonymized. The Ethical Review Board of the Faculty of Medicine, Vajira Hospital, Navamindradhiraj approved the study (COA119/61).

Growth on Caffeic Acid Ferric Citrate (CAFC) Test Agar

CAFC test agar was used to identify and then differentiate *C. neoformans* from other sibling yeast species by produces melanin or melanin-like pigments from an *o*-diphenol (caffeic acid) in the presence of ferric citrate [26–29]. Isolates were cultured on Sabouraud Dextrose Agar (SDA) and maintained at 37 °C for 48 h. Strains were adjusted to the 0.5 McFarland standard in 0.9% NaCl and then serially diluted fivefold (five dilutions). From each dilution, 5-µl aliquots were spotted on CAFC agar (Himedia, India) followed by incubation at 37 °C for 6 days (triplicate tests for each isolate).

Antifungal Sensitivity Testing

A Sensititre™ YeastONE YO10 plate (Thermo Scientific, USA) was used to test the sensitivity of amphotericin B (0.12–8 µg/ml), 5-flucytosine

(0.06–64 µg/ml), fluconazole (0.12–256 µg/ml), posaconazole (0.008–8 µg/ml), voriconazole (0.008–8 µg/ml) and itraconazole (0.015–16 µg/ml), following the manufacturer's guidelines and the guidelines of the Clinical Laboratory Standards Institutes (CLSI). The wells contained a colorimetric indicator, which improves end-point readability because of a change from blue to purple, which is used to calculate minimum inhibitory concentrations (MICs). A clinical breakpoint value is not available for *C. neoformans* in the CLSI guidelines. Therefore, the epidemiological cutoff value (ECV) was used as a reference to define a strain as wild-type [30–32]. *C. parapsilosis* ATCC 22,019 served as the control for the broth microdilution system.

Sample Preparation and Genomic DNA Extraction

Isolates were cultured on SDA and maintained at 37 °C for 48 h. Genomic DNA was extracted and modified according to a published method [14, 16]. In brief, a loop full of yeast culture fluid was transferred into a microtube and frozen at − 20 °C for 1 h. Lysis buffer (500 µL) was added with 0.5 g sodium dodecyl sulfate, 1.4 g NaCl, 0.73 g EDTA, 20 mL 1 M Tris–HCl per 100 mL lysis, and 5 µL of 2-mercaptoethanol, and the suspension was vortexed and incubated at 65 °C for 1 h. The lysate was extracted with 500 µL phenol–chloroform–isoamyl alcohol (25:24:1, v:v:v) and mixed by pipetting. The tubes were centrifuged for 5 min at 12,500 rpm at 4 °C. The aqueous phase was transferred to a new tube, and DNA was precipitated with 500 µL of isopropanol and incubated for 1 h at 4 °C. To pellet the DNA, the solution was centrifuged for 5 min at 12,500 rpm at 4 °C. The DNA pellet was washed with 70% ethanol, centrifuged for 5 min at 12,500 rpm, and air-dried. The DNA was resuspended in 100 µL of sterile deionized water and stored at − 20 °C.

MLST

All isolates were sequenced using primers specific for ITS5 and ITS4 [5.8S ribosomal DNA gene and internal transcribed spacers (ITS)] and the MLST scheme of the International Society for Human and Animal Mycology through the Fungal MLST Database (https://mlst.mycologylab.org/) [33]. The MLST scheme amplifies sequences at the genetic loci as follows: capsular associated protein (*CAP59*); glyceraldehyde-3-phosphate dehydrogenase (*GPD1*); laccase (*LAC1*); phospholipase (*PLB1*); Cu, Zn superoxide dismutase (*SOD1*); orotidine monophosphate pyrophosphorylase (*URA5*); ribosomal RNA intergenic spacer (IGS1). PCR was performed using the specific primer pairs listed in Table S1. Each PCR reaction mixture (25 µl final volume) contained 0.5 µM of each primer. KAPA 2G Fast HS ReadyMix PCR Kit with loading dye (Kapa Biosystems, USA), nuclease-free water, and genomic DNA. PCR amplifications were performed using a T100 Thermal Cycler (Bio-Rad, USA) as follows: initial step at 95 °C for 5 min, followed by 35 cycles of amplification at 95 °C for 1 min (at a gene-specific annealing temperature) (Table S1) for 45 s, 72 °C for 45 s, and a final extension at 72 °C for 5 min. The PCR products (5 µl) were electrophoresed through a 1.5% agarose gel in 1 × TBE buffer, stained with 1 µg/ml ethidium bromide, and photographed using a Gel Doc XR + system (Bio-Rad, USA).

Each PCR product was purified using a FavorPrep GEL/PCR Purification Mini Kit (Favorgen Biotech Corporation, Taiwan) and sequenced using gene-specific forward and reverse primers at Macrogen (South Korea) and Bio Basic Asia Pacific Pte Ltd (Singapore). The sequence files were analyzed using BioEdit software (https://www.mbio.ncsu.edu/bioedit/bioedit.html) and compared with sequences in GenBank using BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi). A single allele type (AT) number was assigned to each of the seven loci by comparing the consensus of DNA sequences with the Fungal MLST Database (https://mlst.mycologylab.org/cneoformans). The allelic profiles defined the corresponding Sequence Types (STs). Primers and PCR conditions used for the amplification of the MLST genes are listed in Table S1.

To assess evolutionary relationships among STs, the sequence of each locus was trimmed to the correct length from the start to the end of each gene. Sequences were concatenated in the order as follows: *CAP59*, *GDP1*, IGS1, *LAC1*, *PLB1*, *SOD1*, and *URA5*. The evolutionary history was inferred using the Maximum Likelihood method. The best models of evolution for the concatenated data set were selected from the Bayesian Information Criterion (BIC) of MEGA X [34]. The model with the lowest BIC score was chosen to construct a maximum likelihood

phylogenetic tree. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The initial tree for the heuristic search was automatically obtained by applying the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood approach, and we then selected the topology associated with a superior log likelihood value. The tree is drawn to scale, with branch lengths representing the number of substitutions per site. A bootstrap analysis was conducted using 10,000 replicates, and bootstrap values $\geq$ 50% are shown above branches.

Determination of MAT Using Multiplex PCR

MAT was performed following a published protocol [35]. The PCR reaction mixture (25 µl) contained 0.5 µM of each primer, KAPA 2G Fast HS ReadyMix PCR Kit with loading dye (Kapa Biosystems, USA), nuclease-free water, and genomic DNA. PCR amplifications were performed using a T100 Thermal Cycler (Bio-Rad, USA) as follows: initial step at 95 °C for 5 min, followed by 35 cycles of amplification at 95 °C for 1 min, 64 °C for 45 s, 72 °C for 45 s, and a final extension step at 72 °C for 5 min. The PCR products (5 µl) were electrophoresed through a 2% agarose gel containing SYBR™ Safe DNA Gel Stain (Thermo Fisher Scientific, USA) in 1 × TBE buffer, 80 V for 2 h. The images of the gels were acquired using a Gel Doc XR + system (Bio-Rad, USA). *C. neoformans* strains CBS132 and JEC21 (provided by Prof. Massimo Cogliati, Laboratorio di Micologia Medica, Dipartimento Scienze Biomediche per la Salute, Università degli Studi di Milano, Milano, Italy) were used to optimize PCR conditions.

Molecular Typing Using RFLP Analysis of *URA5* (*URA5*-RFLP)

We first performed bioinformatics analyses of the *URA5*-RFLP method. The *URA5* sequences acquired using MLST were used to predict the products of double-digests using HhaI and Cfr13I BioEdit v7.2.5. Wet-lab analysis of the *URA5*-RFLP technique was adapted from a published protocol [36]. PCR reaction mixtures (50 µl final volume) contained 0.5 µM of *URA5* primer pairs (Table S1), a KAPA 2G Fast HS ReadyMix PCR Kit with loading dye (Kapa Biosystems, USA), nuclease-free water and genomic DNA. PCR products were double-digested using HhaI (10 U/µl, Invitrogen™ Anza™ 59 HhaI; Thermo Fisher Scientific, USA) and Cfr13I (5 U/µl, Invitrogen™ Anza™ 116 Cfr13I; Thermo Fisher Scientific, USA) for 1 h, and the amplicons were separated using a 2% agarose gel containing SYBR™ Safe DNA Gel Stain (Thermo Fisher Scientific, USA) in 1 × TBE, 80 V for 2 h. RFLP patterns were visually determined by comparison with a published standard pattern [36].

WGS and In-Silico Analysis

WGS was performed using an Ion Proton System (Thermo Fisher Scientific, USA). Measurements of DNA concentrations and preparation of the whole genome library were performed according to the manufacturer's protocol (Medical Genome Company, Bangkok, Thailand). The FASTQ files were mapped against the H99 reference (GCF_000149245) using BWA-MEM (version 0.7.17-r1188) [37] with default parameters. To place our data into a Thai context, 40 whole genome sequences of Thai *C. neoformans* isolates from the EBI-ENA database were analyzed (BioProject accession numbers PRJEB27222 and PRJEB5282) [18]. Single nucleotide polymorphisms (SNPs) were simultaneously called from 49 samples using the GATK Unified Genotyper algorithm (version 3.3-0-g37228af1) [38] with the parameters as follows: (1) sample_ploidy 2 (Ploidy [number of chromosomes] per sample), (2) genotype_likelihoods_model SNP (genotype likelihoods calculation model employed), and (3) rf BadCigar (filters applied to reads before analysis). SNPs were filtered according to the following criteria: read depth > 5, genotype quality > 20, and mapping quality > 20, using BCFtools (version 1.9) (https://www.htslib.org/doc/bcftools.html). The variant call format files were converted to interleaved PHYLIP format using TASSEL (version 5.2.53) [39]. The maximum likelihood approach was performed using IQ-TREE (version 1.6.1) [40]. JEC21 (GCA_000091045) served as an outgroup to construct a phylogenetic tree with the following parameters: (1) Generalized time reversible model (GTR) + I + G (Standard model selection), where GTR is a substitution model, I (Invariable site), and G (discrete Gamma model) are rate-heterogeneity variables; (2) Ultrafast bootstrap (bb 10,000) [41]. A phylogenetic network depicting the association among whole genome sequences of Thai *C. neoformans*

isolates was performed using the neighbor-net algorithm of SplitsTree4 (https://www.splitstree.org/) [42].

Data Availability

Sequence reads were deposited in the NCBI Sequence Read Archive (SRA: Project Number PRJNA565894). The cultures of these isolates are available in the biobanking of Department of Clinical Pathology, Faculty of Medicine Vajira Hospital, Navamindradhiraj University and can be requested to the corresponding author. The MLST nucleotide sequences were deposited in GenBank under accession numbers MT310905 to MT310960.

Results

Population and Setting

The identification of each of the seven isolates as *C. neoformans* s. s. was achieved using routine laboratory identification, MALDI-TOF, PCR, and the sequences of the ITS regions. Six isolates were collected from blood cultures (VJR-CN001, VJR-CN002, VJR-CN003, VJR-CN005, VJR-CN0010, and VJR-CN0011), and only one isolate was collected from the CSF culture (VJR-CN009) from September 1, 2018 to August 30, 2019. The VJR-CN005 and VJR-CN009 strains were isolated from different specimens of the same patients.

We found that the seven isolates oxidized *o*-diphenols to melanin at variable rates from day 2 until day 6 after plating (Fig. 1). The highest rates detected after 48 h were achieved by the VJR-001, VJR-005, and VJR-009 strains.

In Vitro Antifungal Susceptibility

Voriconazole (MIC range, 0.03–0.06), posaconazole (MIC range 0.0–0.25), and itraconazole (MIC range, $\leq$ 0.015–0.12) had high antifungal activities. MICs ranged as follows: 0.5 μg/ml, amphotericin B; 2–4 μg/ml, 5-flucytosine; 2–8 μg/ml, fluconazole (Table 1).

MLST, MAT, and Molecular Typing

The seven isolates of *C. neoformans* were classified as ST4, ST5, or ST6 according to this scheme (Table 2). The isolates were identified as MATα when compared with the published pattern (320- and 400-bp) amplicons synthesized using the NAD4αAF and NAD4αAR primer pair (Fig. S1) [35]. *URA5*-RFLP identified all isolates as molecular type VNI (Fig. S2, Fig. S3, and Table 2).

The *URA5* sequences comprised two alleles (ATs) (Table 2). VJR-CN001 and VJR-CN002 were identified as AT1, and the other isolates were identified as AT5. In-silico analysis predicted the correct length of *URA5*-AT1 and *URA5*-AT5 (5′-terminal ATGTCCT to 3′-terminal GTCTTAA) as 777 bp, and when aligned, revealed variable sites between the two alleles at nucleotide positions 60 (C/T) and 504 (T/C), respectively (Fig. S2). We predicted the restriction map of the *URA5* sequence (777 bp) generated using double-digestion with HhaI and Cfr13I. HhaI did not cleave these two sequences, and Cfr13I cleaved at positions 213/214 and 274/275. The predicted, digested fragment lengths were 61 bp, 213 bp, and 503 bp (Fig. S2). These results were verified by wet-lab analysis. The RFLP profile is shown in Fig. S3. The lengths of the *URA5*-PCR products corresponded to 61 bp, 213 bp, and 503 bp, characteristic of molecular type VNI.

MLST and WGS Analyses of Population Structure and Genetic Diversity of Thai isolates of *C. neoformans*

The phylogenetic relationships among ST4, ST5, and ST6, as well as ST2 from *C. neoformans* H99, are shown in Fig. S4. The STs clustered into related branches with the same topology and separated into two branches. The first branch comprises ST4 and ST6, supported by bootstrap value (BV) = 65. The second branch comprises ST5. We used WGS to determine the genetic diversity and genomic relationships of the seven *C. neoformans* isolates, and the sequences were used in the phylogenetic analysis in comparison with sequences of other Thai isolates of *C. neoformans*. Reference-based phylogenetic analysis was performed, and the data were compared with the whole genomes of *C. neoformans* (ST4 and ST5) isolated in Thailand [18]. We performed WGS of 49

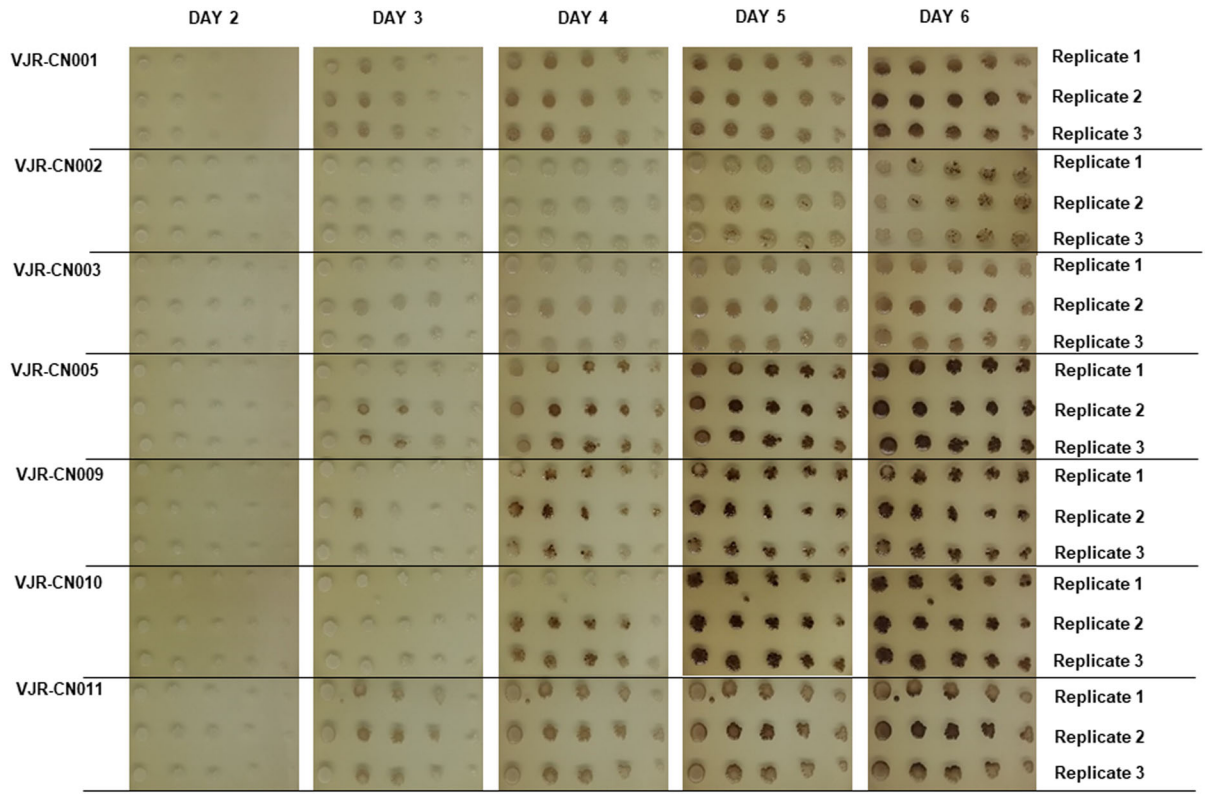

Fig. 1 Melanin production by *C. neoformans* cultured on CAFC agar. The initial spot is the final inoculum adjusted to the 0.5 McFarland standard, and subsequent spots are fivefold dilutions. Plates were incubated at 37 °C for 6 days

Table 1 Antifungal susceptibility profile

Strains	Antifungal drugs (MIC values; µg/ml)					
	AMB	5FC	FLC	VCZ	POZ	ITZ
VJR-CN001	0.5	4	8	0.03	0.06	0.03
VJR-CN002	0.5	4	4	0.03	0.12	0.06
VJR-CN003	0.5	2	4	0.03	0.25	0.06
VJR-CN005	0.5	4	8	0.06	0.06	0.12
VJR-CN009	0.5	4	8	0.06	0.06	0.12
VJR-CN010	0.5	2	8	0.06	0.12	0.12
VJR-CN011	0.5	4	2	0.03	0.03	≤ 0.015

AMB amphotericin B, *5FC* 5-flucytosine, *FLC* fluconazole, *VCZ* voriconazole, *POZ* posaconazole, *ITZ* itraconazole

isolates (seven isolates studied here and 40 previously published isolates) [18] to investigate their phylogenetic relationships. WGS analyses of *C. neoformans* H99 and *C. deneoformans* JEC21 are included (Fig. 2). All FASTQ sequences were mapped compared with the WGS data of the H99 strain (GCF_000149245). The WGS data of the JEC21 strain were used as an outgroup. The split-tree network analysis parallelogram is presented (Fig. 3). The characters in the split-tree network, which are shared by a set of species, group the isolates into the terminal branch (Fig. 3). The SNP data of the isolates based on WGS separated the isolates into two major clades (not compared with the WGS of *C. neoformans* H99; GCF 000149 and JEC21; GCA 000091) with BV = 19, which is supported by the split-tree network analysis. The first clade was the WGS of ST5-*C. neoformans*. VJR-CN001 and VJR-CN002 clustered together with *C. neoformans* ENA accession numbers ERR2624251, ERR2624103, ERR2624271 and ERR2624470. The second clade clustered together with the WGS data of ST4-*C. neoformans* and ST6-*C. neoformans* (BV = 24). This clade was divided into three subclades (Fig. 2). Among the WGS data of the ST4/ST6 clade, our ST4 isolates (VJR-CN003, VJR-CN005, and VJR-CN009) clustered together with our ST6 isolates

Table 2 *Cryptococcus neoformans* isolates

Strains	Time of isolation	Specimens	AT							ST	Mating type	RFLP profile
			CAP59	GPD1	IGS1	LAC1	PLB1	SOD1	URA5			
VJR-CN001	09-2018	Blood	1	3	1	5	2	1	1	5	α	VNI
VJR-CN002	11-2018	Blood	1	3	1	5	2	1	1	5	α	VNI
VJR-CN003	1-2019	Blood	1	1	1	4	2	1	5	4	α	VNI
VJR-CN005[a]	3-2019	Blood	1	1	1	4	2	1	5	4	α	VNI
VJR-CN009[a]	3-2019	CSF	1	1	1	4	2	1	5	4	α	VNI
VJR-CN010	6-2019	Blood	1	1	1	3	2	1	5	6	α	VNI
VJR-CN011	8-2019	Blood	1	1	1	3	2	1	5	6	α	VNI

[a]VJR-CN005 and VJR-CN009 were isolated from the same patient: allele type (AT): sequence type (ST)

(VJR-CN010, VJR-CN011), BV = 23, and separate from the other ST4-Thai isolates. From this WGS analysis, our ST4 and ST5 isolates were resolved at a greater resolution of intrataxa diversity in the same STs among other Thai strains when compared with the MLST data.

Discussion

The epidemiology of *Cryptococcus* species isolated from clinical samples was studied by several research groups in Thailand [14, 16–18]. The taxonomy of the *Cryptococcus* genus was revised [4–6]. Here we report the isolation of only *C. neoformans* strains from patients admitted to Vajira Hospital, Bangkok, Thailand. The isolates were identified using MALDI-TOF MS, and their ITS regions were sequenced using primers specific for ITS5 and ITS4. All isolates produced melanin, which is an important virulence factor of *C. neoformans*. Melanin protects *C. neoformans* from amphotericin B, macrophage-mediated phagocytosis, nitrogen- and oxygen-derived oxidants, microbicidal peptides, and ultraviolet light [43, 44].

The guidelines of the Infectious Diseases Society of America recommend therapy using amphotericin B, deoxycholate, and flucytosine in the induction regimen and fluconazole as an alternative in the consolidation and maintenance regimens [45]. In this study, antifungal susceptibility testing was performed using the Sensititre YeastOne kit, which achieves a good correlation with the broth microdilution method described by the CLSI [46–48]. The ECV was used as the reference to define a strain as wild-type (a clinical breakpoint for *C. neoformans* is not available from the CLSI) [30–32]. MIC values $\leq$ 0.5 μg/ml for amphotericin B, $\leq$ 8 μg/ml for 5-flucytosine, $\leq$ 8 μg/ml for fluconazole, and $\leq$ 0.25 μg/ml each for voriconazole, posaconazole, and itraconazole are considered characteristics of the wild-type strains [30–32]. Therefore, our in vitro susceptibility tests identified our isolates as wild-type. Compared with the results of other studies conducted in Thailand [9], all isolates of clinical isolates from the culture collections of the Mycology Research Unit, Faculty of Medicine, Chulalongkorn University, Bangkok (2002–2011) were also sensitive to amphotericin B, 5-flucytosine, and fluconazole.

The *URA5*-RFLP analysis groups *C. neoformans* s.s. into the major molecular types VNI and VNII. *C. deneoformans* is molecular type VNIV, and hybrids between *C. neoformans* s.s and *C. deneoformans* belong to molecular type VNIII [5, 6]. Our present study identified the seven isolates as molecular type VNI (Fig. S2, Fig. S3, and Table 2). No isolate was classified as molecular types VNII or VNB. Therefore, molecular type VNI of *C. neoformans* isolated from

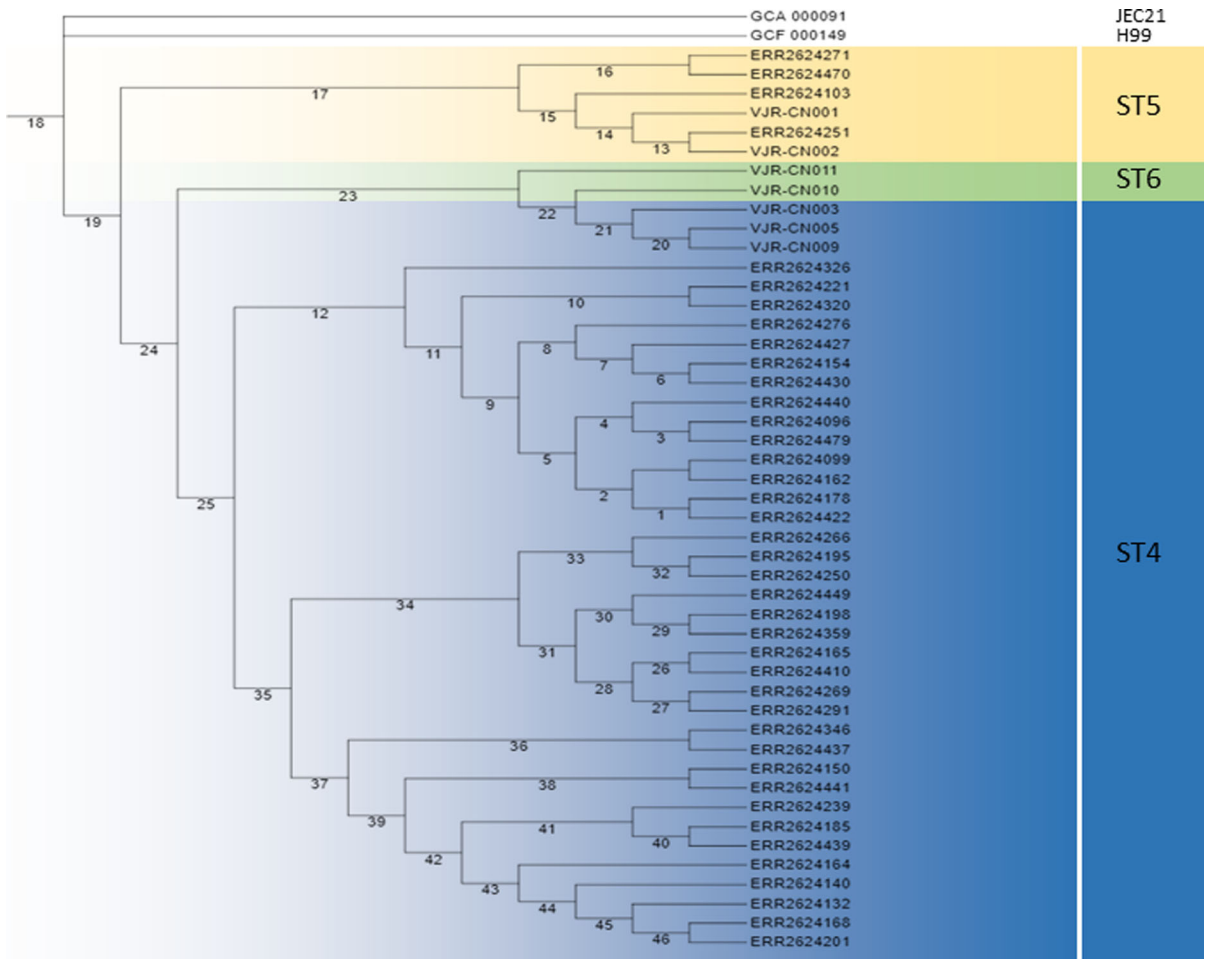

Fig. 2 Molecular phylogenetic maximum likelihood analysis of the whole genome sequences of 49 Thai strains

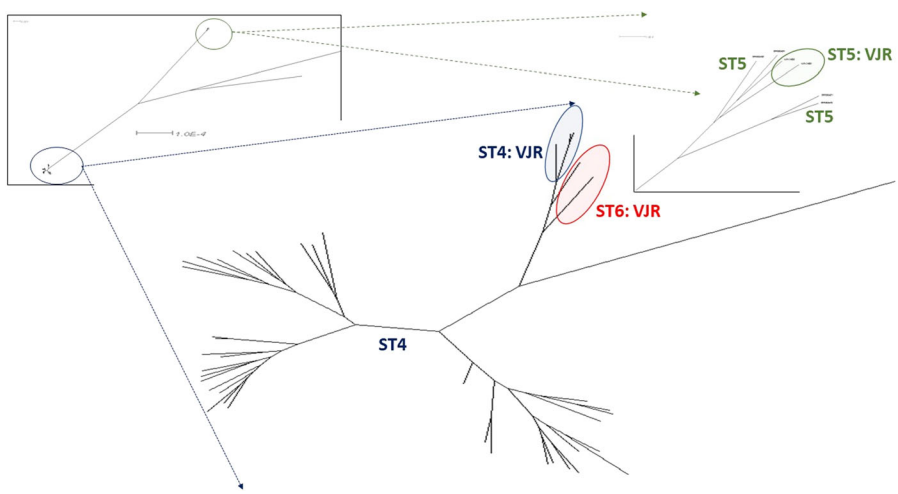

Fig. 3 Unrooted network tree depicting the associations among whole genome sequences of Thai *C. neoformans* isolates

our patients is consistent with other studies in Thailand that identified VNI as the most prevalent molecular type [14, 16]. Further, VNI is the most common molecular type of *C. neoformans* s.s associated with HIV-infected and immunocompromized (non-HIV) patients [49, 50].

The two *C. neoformans* mating types MATα and MATa are controlled by a single locus of the two-allele mating system [15, 51]. All isolates of the strains analyzed here were MATα. Moreover, clinical and environmental isolates of 183 Thai *C. neoformans* are MATα [15], and another study found that that all clinical isolates are MATα [14]. Therefore, MATα is the major MAT of circulating *C. neoformans* strains in Thailand. Notably, MATα strains are more virulent than MATa that infect the central nervous system. This may explain why MATα strains are isolated from patients in Thailand at high frequency and are more prevalent worldwide than MATa strains [52, 53].

MLST has the advantage of providing readily comparable results through access to databases that reside in the Cloud [54, 55]. In the first pipeline, our study used MLST as the genotyping tool to study the genetic diversity of *C. neoformans* isolated from clinical samples. We identified ST4, ST5, and ST6 among our seven isolates, which are the most commonly observed STs in Thailand [15, 17]. Further, *C. neoformans* isolated from clinical samples acquired from patients admitted to the Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok from 2012 to 2014 include ST4 (n = 23), ST6 (n = 18), ST5 (n = 8), ST93 (n = 1), and ST32 (n = 1) [14]. Moreover, the genetic diversity of 183 Thai clinical isolates of *C. neoformans* is lower compared with those found in Africa and the Americas [15].

Laotian isolates of *C. neoformans* are ST4 (n = 35), ST6 (n = 33), ST5 (n = 9), and ST93 (n = 2), in contrast to *C. neoformans* ST5, which predominates in Vietnam (n = 66), followed by ST4 (n = 32), ST93 (n = 8), and ST6 (n = 12) [10]. Thanh et al. [10], generated a geographical distribution of ST types of *C. neoformans* isolated from Vietnam and Laos, along with those from other Asian countries, including Thailand, as well as those isolated in North and South America, Europe, the Middle East, and Africa. A clonal complex was revealed using GoeBurst analysis, and the association between geographical location and STs were recoded [10]. For example, most Asian isolates are included in subgroup 4 (ST4/ST6) and subgroup 5 (ST5) [10], which is consistent with our present findings (Fig. 2). VNI-ST5 is the most frequently reported molecular and sequence type among Asian isolates [14, 56–59].

Limitations of the present study include the small numbers of isolates analyzed from one hospital, the short collection period, and the low diversity among the isolates. Thus, MLST captures only a small fraction of sequence diversity, which may provide incomplete data or inaccurate measures of species relationships [60, 61]. In the present study, MLST was, therefore, insufficient for analyzing genetic diversity.

A phylogenetic analysis based on WGS of 699 isolates (682 of *C. neoformans* sensu lalto, 12 of *C. gattii* sensu lato, and 5 of putative hybrids between *C. neoformans* sensu lato and *C. deneoformans*) from Vietnam, Laos, Thailand, Uganda, and Malawi clustered the isolates into the main subclades VNIa-4, VNIa-5, and VNIa-93 [18]. Further, 40 of 682 of *C. neoformans* isolates were from Thailand, and only *C. neoformans* ST4 and ST5 were isolated from 2013 through 2014 [18]. To gain insights into the genetic diversity of our *C. neoformans*, we constructed a phylogenetic tree using the whole genome sequences of our seven *C. neoformans* isolates and those of other Thai isolates (n = 40). The WGS-based-approach clearly discriminated the ST5 strains from each other and divided the ST4 strains into a subgroup.

WGS revealed the sequence diversities and relationships among the 47 *C. neoformans* Thai strains, attesting to the high discriminatory power of WGS compared with MLST. Moreover, other WGS-based approaches were used to characterize *Cryptococcus* species. For example, WGS analysis of 118 *C. gattii* s.l. isolates and a population genome analysis provides insights into species evolution and dispersal through the Pacific Northwest region of North America [62]. Here we show that the strains VJR-CN001 (VNI/MATα-ST4), VJR-CN005 (VNI/MATα-ST5), and VJR-CN009 (VNI/MATα-ST5) produced the highest level of melanin, and their MICs for fluconazole (8 µg/ml) were higher compared with those of other strains. The *LAC1* gene regulates melanin catalysis and associates with variable levels of melanin production [63]. Therefore, our finding indicates the relationship between this gene and melanin production.

In summary, we demonstrate here the usefulness of WGS analysis for genetic characterization of seven *C.*

neoformans isolated from six patients in Thailand. The value of WGS in the context of genetic diversity derives from its high discriminatory power for identifying *C. neoformans*. We further report the in vitro antifungal susceptibility profiles, mating types, ATs and STs of Thai isolates of *C. neoformans*, for the first time to our knowledge. Moreover, we present a phylogenetic analysis of the relationships among the whole genome sequences of their STs (ST4, ST5, and ST6) as well as other genomics data.

Acknowledgements This study was funded by a grant awarded by Navamindradhiraj University Research Fund (Grant Number worjor/sornorthor.26/2562) to TW, CT and AH. We thank Faculty of Medicine Vajira Hospital for the English editing supported and edited by the ENAGO company.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Statement The Ethical Review Board of the Faculty of Medicine, Vajira Hospital, Navamindradhiraj approved the study (COA119/61).

References

1. Aguiar P, Pedroso RDS, Borges AS, et al. The epidemiology of cryptococcosis and the characterization of *Cryptococcus neoformans* isolated in a Brazilian University Hospital. Rev Inst Med Trop Sao Paulo. 2017;59:e13.
2. Yan Z, Li X, Xu J. Geographic distribution of mating type alleles of *Cryptococcus neoformans* in four areas of the United States. J Clin Microbiol. 2002;40(3):965–72.
3. Sathirapanya P, Ekpitakdamrong N, Chusri S, et al. Predictors of hospital discharge outcome from the presenting clinical characteristics and the first cerebrospinal fluid analysis among the patients with cryptococcal meningitis. Clin Neurol Neurosurg. 2019;186:105539.
4. Quintero O, Trachuk P, Lerner MZ, et al. Risk factors of laryngeal cryptococcosis: a case report. Med Mycol Case Rep. 2019;24:82–5.
5. Munoz M, Camargo M, Ramirez JD. Estimating the intra-taxa diversity, population genetic structure, and evolutionary pathways of *Cryptococcus neoformans* and *Cryptococcus gattii*. Front Genet. 2018;9:148.
6. 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.
7. Kwon-Chung KJ, Bennett JE, Wickes BL, et al. The case for adopting the "Species Complex" nomenclature for the etiologic agents of Cryptococcosis. mSphere. 2017;2(1):00357.
8. Fang W, Fa Z, Liao W. Epidemiology of *Cryptococcus* and cryptococcosis in China. Fungal Genet Biol. 2015;78:7–15.
9. Worasilchai N, Tangwattanachuleeporn M, Meesilpavikkai K, et al. Diversity and antifungal drug susceptibility of *Cryptococcus* isolates in Thailand. Med Mycol. 2017;55(6):680–5.
10. Thanh LT, Phan TH, Rattanavong S, et al. Multilocus sequence typing of *Cryptococcus neoformans* var. *grubii* from Laos in a regional and global context. Med Mycol. 2018;57(5):557–65.
11. Okubo Y, Wakayama M, Ohno H, et al. Histopathological study of murine pulmonary cryptococcosis induced by *Cryptococcus gattii* and *Cryptococcus neoformans*. Jpn J Infect Dis. 2013;66(3):216–21.
12. Umeyama T, Ohno H, Minamoto F, et al. Determination of epidemiology of clinically isolated *Cryptococcus neoformans* strains in Japan by multilocus sequence typing. Jpn J Infect Dis. 2013;66(1):51–5.
13. Day JN, Qihui S, Thanh LT, 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(6):e0005628.
14. Hatthakaroon C, Pharkjaksu S, Chongtrakool P, et al. Molecular epidemiology of cryptococcal genotype VNIc/ST5 in Siriraj Hospital, Thailand. PLoS ONE. 2017;12(3):e0173744.
15. Simwami SP, Khayhan K, Henk DA, et al. Low diversity *Cryptococcus neoformans* variety grubii multilocus sequence types from Thailand are consistent with an ancestral African origin. PLoS Pathog. 2011;7(4):e1001343.
16. Kaocharoen S, Ngamskulrungroj P, Firacative C, et al. Molecular epidemiology reveals genetic diversity amongst isolates of the *Cryptococcus neoformans/C. gattii* species complex in Thailand. PLoS Negl Trop Dis. 2013;7(7):e2297.
17. Khayhan K, Hagen F, Pan W, 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(9):e72222.
18. Ashton PM, Thanh LT, Trieu PH, et al. Three phylogenetic groups have driven the recent population expansion of *Cryptococcus neoformans*. Nat Commun. 2019;10(1):2035.
19. Gerstein AC, Jackson KM, McDonald TR, et al. Identification of pathogen genomic differences that impact human immune response and disease during *Cryptococcus neoformans* infection. MBio. 2019;10(4):e01440.
20. Hua W, Vogan A, Xu J. Genotypic and phenotypic analyses of two "Isogenic" strains of the human fungal pathogen *Cryptococcus neoformans* var. *neoformans*. Mycopathologia. 2019;184(2):195–21212.
21. Rhodes J, Desjardins CA, Sykes SM, et al. Tracing genetic exchange and biogeography of *Cryptococcus neoformans* var. *grubii* at the global population level. Genetics. 2017;207(1):327–46.
22. Desjardins CA, Giamberardino C, Sykes SM, et al. Population genomics and the evolution of virulence in the fungal pathogen *Cryptococcus neoformans*. Genome Res. 2017;27(7):1207–19.
23. Rossen JWA, Friedrich AW, Moran-Gilad J, et al. Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. Clin Microbiol Infect. 2018;24(4):355–60.
24. McTaggart L, Richardson SE, Seah C, et al. Rapid identification of *Cryptococcus neoformans* var. *grubii*, *C.*

neoformans var. *neoformans*, and *C. gattii* by use of rapid biochemical tests, differential media, and DNA sequencing. J Clin Microbiol. 2011;49(7):2522–7.
25. McTaggart LR, Lei E, Richardson SE, et al. Rapid identification of *Cryptococcus neoformans* and *Cryptococcus gattii* by matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol. 2011;49(8):3050–3.
26. Wang HS, Zeimis RT, Roberts GD. Evaluation of a caffeic acid-ferric citrate test for rapid identification of *Cryptococcus neoformans*. J Clin Microbiol. 1977;6(5):445–9.
27. Hopfer RL, Blank F. Caffeic acid-containing medium for identification of *Cryptococcus neoformans*. J Clin Microbiol. 1976;2(2):115–20.
28. Chaskes S, Tyndall RL. Pigment production by *Cryptococcus neoformans* from para- and ortho-Diphenols: effect of the nitrogen source. J Clin Microbiol. 1975;1(6):509–14.
29. Pulverer G, Korth H. *Cryptococcus neoformans*: pigment formation on polyphenols. Med Microbiol Immunol. 1971;157(1):46–51.
30. Espinel-Ingroff A, Aller AI, Canton E, et al. *Cryptococcus neoformans-Cryptococcus gattii* species complex: an international study of wild-type susceptibility endpoint distributions and epidemiological cutoff values for fluconazole, itraconazole, posaconazole, and voriconazole. Antimicrob Agents Chemother. 2012;56(11):5898–906.
31. Espinel-Ingroff A, Chowdhary A, Cuenca-Estrella M, et al. *Cryptococcus neoformans-Cryptococcus gattii* species complex: an international study of wild-type susceptibility endpoint distributions and epidemiological cutoff values for amphotericin B and flucytosine. Antimicrob Agents Chemother. 2012;56(6):3107–13.
32. Pllana-Hajdari D, Cogliati M, Cicmak L, et al. First isolation, antifungal susceptibility, and molecular characterization of *Cryptococcus neoformans* from the environment in Croatia. J Fungi (Basel). 2019;5(4):99.
33. 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):561–70.
34. Kumar S, Stecher G, Li M, et al. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547–9.
35. Esposto MC, Cogliati M, Tortorano AM, et al. Determination of *Cryptococcus neoformans* var. *neoformans* mating type by multiplex PCR. Clin Microbiol Infect. 2004;10(12):1092–4.
36. Meyer W, Castaneda A, Jackson S, et al. Molecular typing of IberoAmerican *Cryptococcus neoformans* isolates. Emerg Infect Dis. 2003;9(2):189–95.
37. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXivorg %3e q-bio %3e arXiv:13033997; 2013.
38. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9): 1297–303.
39. Glaubitz JC, Casstevens TM, Lu F, et al. TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. PLoS ONE. 2014;9(2):e90346.
40. Nguyen LT, Schmidt HA, von Haeseler A, et al. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268–74.
41. Hoang DT, Chernomor O, von Haeseler A, et al. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35(2):518–22.
42. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 2006;23(2):254–67.
43. Hamilton AJ, Holdom MD. Antioxidant systems in the pathogenic fungi of man and their role in virulence. Med Mycol. 1999;37(6):375–89.
44. Nosanchuk JD, Rosas AL, Lee SC, et al. Melanisation of *Cryptococcus neoformans* in human brain tissue. Lancet. 2000;355(9220):2049–50.
45. Perfect JR, Dismukes WE, Dromer F, et al. Clinical practice guidelines for the management of cryptococcal disease: 2010 update by the infectious diseases society of america. Clin Infect Dis. 2010;50(3):291–32222.
46. Cuenca-Estrella M, Gomez-Lopez A, Alastruey-Izquierdo A, et al. Comparison of the Vitek 2 antifungal susceptibility system with the clinical and laboratory standards institute (CLSI) and European Committee on Antimicrobial Susceptibility Testing (EUCAST) broth microdilution reference methods and with the Sensititre YeastOne and Etest techniques for in vitro detection of antifungal resistance in yeast isolates. J Clin Microbiol. 2010;48(5):1782–6.
47. Espinel-Ingroff A, Pfaller M, Messer SA, et al. Multicenter comparison of the sensititre YeastOne colorimetric antifungal panel with the National Committee for Clinical Laboratory standards M27-A reference method for testing clinical isolates of common and emerging *Candida* spp., *Cryptococcus* spp., and other yeasts and yeast-like organisms. J Clin Microbiol. 1999;37(3):591–5.
48. Lee GA, Arthur I, Merritt A, et al. Molecular types of *Cryptococcus neoformans* and *Cryptococcus gattii* in Western Australia and correlation with antifungal susceptibility. Med Mycol. 2019;57(8):1004–100.
49. Bandalizadeh Z, Shokohi T, Badali H, et al. Molecular epidemiology and antifungal susceptibility profiles of clinical *Cryptococcus neoformans*/*Cryptococcus gattii* species complex. J Med Microbiol. 2020;69(1):72–81.
50. Ngamskulrungroj P, Khayhan K, Pitak-Arnnop P, et al. An association of *Cryptococcus neoformans*/*C. gattii* genotype and HIV status in Asia: a systematic review. Siriraj Med J. 2019;71(2):158–64.
51. Kwon-Chung KJ, Bennett JE. Distribution of alpha and alpha mating types of *Cryptococcus neoformans* among natural and clinical isolates. Am J Epidemiol. 1978;108(4):337–40.
52. Tomazin R, Matos T, Meis JF, et al. Molecular characterization and antifungal susceptibility testing of sequentially obtained clinical *Cryptococcus deneoformans* and *Cryptococcus neoformans* isolates from Ljubljana. Slov Mycopathol. 2018;183(2):371–80.
53. Nielsen K, Cox GM, Litvintseva AP, et al. Cryptococcus neoformans alpha strains preferentially disseminate to the central nervous system during coinfection. Infect Immun. 2005;73(8):4922–33.
54. Dodgson AR, Pujol C, Denning DW, et al. Multilocus sequence typing of *Candida glabrata* reveals geographically enriched clades. J Clin Microbiol. 2003;41(12): 5709–17.

55. Biswas C, Marcelino VR, Van Hal S, et al. Whole genome sequencing of Australian *Candida glabrata* isolates reveals genetic diversity and novel sequence types. Front Microbiol. 2018;9:2946.
56. Wu SY, Lei Y, Kang M, et al. Molecular characterisation of clinical *Cryptococcus neoformans* and *Cryptococcus gattii* isolates from Sichuan province. China Mycoses. 2015;58(5):280–7.
57. Fang LF, Zhang PP, Wang J, et al. Clinical and microbiological characteristics of cryptococcosis at an university hospital in China from 2013 to 2017. Braz J Infect Dis. 2020;24(1):7–12.
58. Park SH, Kim M, Joo SI, et al. Molecular epidemiology of clinical *Cryptococcus neoformans* isolates in Seoul. Korea Mycobiol. 2014;42(1):73–8.
59. Dou HT, Xu YC, Wang HZ, et al. 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(4):753–62.
60. Cuomo CA, Rhodes J, Desjardins CA. Advances in *Cryptococcus* genomics: insights into the evolution of pathogenesis. Mem Inst Oswaldo Cruz. 2018;113(7):e170473.
61. Beale MA, Sabiiti W, Robertson EJ, et al. Genotypic diversity is associated with clinical outcome and phenotype in cryptococcal meningitis across Southern Africa. PLoS Negl Trop Dis. 2015;9(6):e0003847.
62. Engelthaler DM, Hicks ND, Gillece JD, et al. *Cryptococcus gattii* in North American Pacific Northwest: whole-population genome analysis provides insights into species evolution and dispersal. MBio. 2014;5(4):e01464–e1514.
63. Samarasinghe H, Aceituno-Caicedo D, Cogliati M, et al. Genetic factors and genotype-environment interactions contribute to variation in melanin production in the fungal pathogen *Cryptococcus neoformans*. Sci Rep. 2018;8(1):9824.