ELSEVIER

FEMS Yeast Research 5 (2005) 1129–1140

FEMS Yeast Research

www.fems-microbiology.org

Comparative analysis of the intergenic spacer regions and population structure of the species complex of the pathogenic yeast *Cryptococcus neoformans*

Mara R. Diaz [a,*], Teun Boekhout [b], Traci Kiesling [a,1], Jack W. Fell [a]

[a] *Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Key Biscayne, FL 33149, USA*
[b] *Centraalbureau voor Schimmelcultures, Yeast Division, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands*

Received 26 January 2005; received in revised form 18 April 2005; accepted 3 May 2005

First published online 1 July 2005

Abstract

Cryptococcus neoformans is an opportunistic basidiomycete responsible for the high incidence of cryptococcosis in patients with AIDS and in other immune-compromised individuals. This study, which focused on the molecular structure and genetic variability of the two varieties in the *C. neoformans* and *Cryptococcus gattii* species complex, employed sequence analysis of the intergenic spacer regions, IGSI and IGSII. The IGS region is the most rapidly evolving region of the rDNA families. The IGSI displayed the most genetic variability represented by nucleotide base substitutions and the presence of long insertions/deletions (indels). In contrast, the IGSII region exhibited less heterogeneity and the indels were not as extensive as those displayed in the IGSI region. Both intergenic spacers contained short, interspersed repeat motifs, which can be related to length polymorphisms observed between sequences. Phylogenetic analysis undertaken in the IGSI, IGSII and IGSI + 5S rRNA + IGSII regions revealed the presence of six major phylogenetic lineages, some of which segregated into subgroups. The major lineages are represented by genotypes 1 (*C. neoformans* var. *grubii*), genotype 2 (*C. neoformans* var. *neoformans*), and genotypes 3, 4, 5 and 6 represented by *C. gattii.* Genotype 6 is a newly described IGS genotypic group within the *C. neoformans* species complex. With the inclusion of IGS subgenotypic groups, our sequence analysis distinguished 12 different lineages. Sequencing of clones, which was performed to determine the presence of multiple alleles at the IGS locus in several hybrid strains, yielded a single IGS sequence type per isolate, thus suggesting that the selected group of cloned strains was mono-allelic at this locus. IGS sequence analyses proved to be a powerful technique for the delineation of the varieties of *C. neoformans* and *C. gattii* at genotypic and subgenotypic levels.

Keywords: Yeast; Intergenic spacer (IGS); Length polymorphism; Repeated motifs; *Cryptococcus neoformans*

1. Introduction

The basidiomycetous, encapsulated yeast *Cryptococcus neoformans* (Sanfelice) Vuillemin is a prevalent clinical opportunistic pathogen, which causes life-threatening infections of the central nervous system in immunocompromised and immunocompetent hosts [1,2]. Estimates indicate that cryptococcosis occurs in about 6% to 10% of AIDS patients in the United States and Western Europe, and in 15% to 30% of patients in sub-Saharan Africa [1–4].

Although *C. neoformans* var. *grubii* is the most common cause of fungal meningitis in AIDS and immunocompromised individuals [1], *C. neoformans* var.

* Corresponding author. Tel.: +1 305 421 4879; fax: +1 305 421 4600.
E-mail address: mdiaz@rsmas.miami.edu (M.R. Diaz).
[1] Current address: Department of Biochemistry, Georgetown University Medical Center, 4000 Reservoir Rd NW, Washington, DC 20057, USA.

1567-1356/$22.00
doi:10.1016/j.femsyr.2005.05.005

neoformans and *Cryptococcus gattii* have been reported as ethiological agents of cryptococcosis in HIV-infected patients [5–8]. *C. neoformans* var. *neoformans* and *C. neoformans* var. *grubii* occur worldwide and are frequently isolated from bird droppings, but also in trees, soil, house dust, and domestic animals, e.g., cats and cows [1,9–11], whereas *C. gattii* is mainly found in tropical and subtropical areas (Australia, Asia, South America, South California and Southern Europe). The geographic boundary of this species has been expanded with the recent outbreak of *C. gattii* in the Vancouver Islands, British Columbia [12,13]. According to a recent proposal, which raised the status of *C. neoformans* var. *gattii* to species status, *C. neoformans* represents a species complex comprising two species: *C. gattii* (serotypes B and C) and *C. neoformans* with variety *grubii* (serotypes A and AD) and variety *neoformans* (serotype D) [14,15]. The two species differ in their capsule polysaccharide structure, antigenic structure, molecular and morphological characteristics, epidemiology, virulence characteristics, ecology, and geography [1].

Genetic heterogeneity has been observed within the species complex [16–20]. AFLP analyses and partial sequence analyses of the intergenic spacer region I (IGSI) has shown considerable genetic divergence between the species and varieties [17,19]. Support for distinct genetic lineages for *C. neoformans* var. *neoformans* and *C. neoformans* var. *grubii* have also been based on URA5 sequences and restriction fragment length polymorphism (RFLP) patterns employing a CNRE 1 probe [7,21].

The scope of the present study is to elucidate and compare the genetic diversity and phylogeny within the *C. neoformans* species complex based on a comparative study involving complete sequence analysis of IGSI and an analysis of the 609–675-bp region of the IGSII region. This region has been frequently used as a tool for species identification and phylogenetic studies [19,22–25]. A detailed analysis of the molecular structure organization of the two intergenic spacers is presented.

2. Methods

2.1. Isolates

One hundred and seven clinical and environmental DNA isolates originating from different geographic areas were analyzed. The source of isolation, genotypes, and serotypes are described in Table 1. Data were obtained from the CBS collection, Boekhout et al. [17], or by information provided by the depositors of the isolates.

2.2. DNA isolation and PCR reaction

DNA was isolated from culture cells as described by Fell et al. [26] using lysing enzyme and QIAmp Tissue kit (Qiagen, Valencia, CA, USA) or by the CTAB method[27].

The DNA amplification of the IGS regions employed a primer walking technique using different sets of universal and specific primers yielding amplicon sizes ranging ~1.5 to 2 kb. The set of primers, which involved a forward and a reverse primer, used to generate amplicons in the IGSI region, were: (a) Lr12F: 5′CTGAACGCCTCTAAGTCAGAA (universal forward primer located in 28SrDNA) and 5SR: 5′GCACCCTGCCCCGTCCGATCC (reverse primer located at position 44–25 of the 5SrDNA gene; (b) Lr12F and IG3R: 5′TGATTCAGCTAGCCAGTAA (reverse primer located at position 619–600 of the IGSII region); (c) Lr12F and IG4R: 5′GTCGCACCCAGTCGCACCTC (reverse primer located at position 820–801 of the IGSII region). Analysis of the IGSII region used the primer combinations: (a) IG1F: 5′CAGACGACTTGAATGGGAACG (forward primer located at position 3613–3633 of the LrRNA region) and IG4R; (b) IG1F and NS1R: 5′GAGACAAGCATATGACTAC (reverse primer located in the18SrDNA); (c) IG2F: 5′CAACAGCTTTCTATGCA (forward primer located at position 777–793 of the IGSI and IG4R; (d) IG2F and NS1R. The strain, CBS 882, was used as reference strain for primer positions, except for the IG4R primer, which employed ATCC 24067.

The PCR reaction for sequence analysis was carried out in microtubes containing a master mix with a final volume of 100 μl. The master mix contained: Target DNA; 10 mM Tris HCl (pH 9); 5 mM KCl; 0.1% Triton X-100; 2 mM $MgCl_2$; 100 nM forward and reverse primers; 2.5 U of AmpliTaq DNA polymerase; dNTPs containing 200 μM each of dGTP, dCTP, dGTP and dATP. The PCR reaction was performed for 40 cycles in a MJ Research PTC 100 thermocyler (GMI, Ramsey, MN, USA) as follows: 2 min denaturating step at 94 °C, 1 min annealing at 57 °C and 3 min extension at 72 °C, followed by a final extension 7 min at 72 °C.

2.3. Cloning

PCR products generated by the primer combination IG1F and IG2R (5′ATG CAT AGA AAG CTG TTG G) were separated and eluted with QIAquick PCR purification kit from Qiagen, and cloned into the pCR 2.1 vector kit according to the manufacturer's instructions (Invitrogen-Carlsbad, CA, USA). Of each ligation, 2 μl were transformed into INVα F′ competent cells following the TA cloning kit procedure (Invitrogen-Carlsbad, CA, USA). Twenty-four white colonies were isolated from each cloned strain and grown overnight in LB liquid media. Plasmid preparations were performed using the Wizard Plus SV Miniprep DNA Purification System from Promega (Madison, WI, USA). All 24 plasmids were sequenced with an Applied Biosystems 3730 DNA Analyzer

Table 1
List of experimental strains

Strain	Source of isolation	Serotype	IGS genotype	AFLP Type	GenBank #
C. neoformans var. *grubii*					
CBS 879	Ulcerated cheek	A	1a	1	=DQ007972
CBS 886	Unknown	A	1a	1	DQ007978
CBS 916	Unknown	A	1a	1	=DQ007972
CBS 935	Unknown	A	1a	1	=DQ007978
CBS 1143	Cerebrospinal fluid	A	1a	1	DQ007974
CBS 1144	Cerebrospinal fluid	A	1a	1	DQ007973
CBS 1931	Soil	A	1a	1	DQ007972
CBS 1932	Soil	A	1a	1A	DQ007977
CBS 1933	Mastitic cow, USA	A	1a	1	=DQ007979
CBS 4572	Cerebrospinal fluid	A	1a	1	=DQ007972
CBS 4868	Sputum, The Netherlands	A	1a	1	DQ007983
CBS 5756	Unknown	A	1a	1	=DQ007978
CBS 6961	Man, Oklahoma, USA	A	1a	1	DQ007975
CBS 7779	Urease negative isolate from AIDS patient, Argentina	A	1a	1	=DQ007972
CBS 7812	Cerebrospinal fluid, USA	A	1a	1	DQ007982
RV 46115	Plants, India	A	1b	1	=DQ007985
RV 52733	Pigeon droppings, Belgium	D	1a	3	=DQ007972
RV 53794	Canary bird droppings, Belgium	D	1a	3	=DQ007978
RV 55446	House dust, Zaire	A	1a	1	=DQ007979
RV 55447	Air inside house, Zaire	A	1a	?	=DQ007979
RV 55451	Cockroach, Zaire	A	1b	1	DQ007984
RV 58145	Wood, Zaire	A	1a	1	=DQ007979
RV 58146	Wood, Zaire	A	1c	1A	DQ007986
RV 59351	Parrot dropping, Belgium	A	1b	1	=DQ007984
RV 59369	Parrot dropping, Belgium	A	1b	1	=DQ007984
RV 60074	Skin, cryptococcosis, Belgium	A	1a	1	DQ007981
RV 61756	Man, Belgium (visited Zaire)	AD	1a	1A	=DQ007972
RV 62210	Cerebrospinal fluid, Belgium fluid from AIDS patient, Belgium	A	1b	1	=DQ007984
RV 64610	AIDS patient, Rwanda	A	1a	1A	=DQ007980
RV 65662	Man, Portugal (visited Venezuela)	A	1a	1A	DQ007980
RV 66025	Cryptococcoma, Belgium	A	1a	1	DQ007979
RDA 1335 AVB0	AIDS patient, The Netherlands	A	1b	1	=DQ007984
RDA 1371 AVB2	AIDS patient, The Netherlands	A	1a	1	=DQ007978
RDA 1369 AVB3	AIDS patient, The Netherlands	A	1a	1	=DQ007978
RDA N/A AVB5	AIDS patient, The Netherlands	A	1a	1	=DQ007972
RDA 1549 AVB7	AIDS patient, The Netherlands	A	1a	1	=DQ007978
RDA 4092 AVB10	AIDS patient, The Netherlands	A	1a	1	=DQ007979
RDA 4094 AVB11	AIDS patient, The Netherlands	A	1b	1	=DQ007984
RDA 4054 AVB12	AIDS patient, The Netherlands	A	1a	1	=DQ007978
RDA 4091 AVB13	AIDS patient, The Netherlands	A	1b	1	=DQ007985
BD2	AIDS patient, France	A	1a	1	DQ007976
WM 164*	Pigeon droppings, Australia	A	1c	1A	DQ007987
WM 553*	House dust, Brazil	A	1c	1A	=DQ007987
WM 554	Dust from pigeon, Brazil	A	1b	1	DQ007985
WM 555	Dust from pigeon, Brazil	A	1a	1	=DQ007972
WM 715	Pine needles	A	1a	1	=DQ007972
WM 716	Woody debris *Eucalyptus camaldulensis*, Australia	A	1a	1	=DQ007978
Hamdan 214L*	AIDS patient, Brazil	A	1c	1A	=DQ007987
C. neoformans var. *neoformans*					
CBS 131	Institut Pasteur, France	AD	2c	3	=DQ007943
CBS 132	Fermenting fruit juice, Italy	D	2c	3	DQ007944
CBS 464	Laboratoire de Parasitologie, Paris,France	A	2c	3	DQ007945
CBS 882	Nasal tumor of horse, type strain of *Torula nasalis*, USA	D	2a	2	=DQ007940
CBS 888	Unknown	D	2a	2	DQ007940
CBS 918	Dead white mouse, The Netherlands	D	2c	2	=DQ007943
CBS 939	Unknown	D	2c	3	=DQ007943
CBS 950	Tumor	AD	2c	3	=DQ007943
CBS 1584	Unknown	A	2c	?	=DQ007943
CBS 4194	Spleen, Germany	D	2a	2	DQ007942
CBS 5467	Milk from mastitic cow, Switzerland	D	2c	2	=DQ009483
CBS 5474	Mastitic cow	D	2c	2	=DQ007943

(*continued on next page*)

Table 1 (*continued*)

Strain	Source of isolation	Serotype	IGS genotype	AFLP Type	GenBank #
CBS 5728	Nonmeningitic cellulitis and osteomyelitis, USA	D	2c	2	=DQ007943
CBS 6885	Lesion on bone in man, USA	D	2c	2	DQ007943
CBS 6886	Dropping of pigeon, Denmark	D	2a	2	DQ007941
CBS 6900	Genetic offspring of CBS 6885 × CBS 7000	D	2b	2	DQ007946
CBS 6901	Genetic offspring of CBS 6885 × CBS 7000	D	2b	2	DQ007947
CBS 7814	Air, Belgium	D	2c	2	=DQ007943
CBS 7815	Pigeon droppings, Czechoslovakia	D	2c	2	=DQ007943
CBS 7816	Cuckoo dropping, Thailand	D	2b	2	DQ007950
CBS 7824	Unknown	D	2b	N/A	DQ007948
CBS 7826	Unknown	AD	2b	N/A	DQ007949
RV 52755	Cerebrospinal fluid, Belgium	D	2c	3	=DQ007943
RV 62692	Skin cryptococcosis, Belgium	D	2c	2	DQ007951
BA1	AIDS patient, France	D	2c	3	=DQ007943
BA3	AIDS patient, France	AD	2c	3	=DQ007943
BA4	AIDS patient, France	AD	2c	3	=DQ007943
C. gattii					
CBS 919	Meningoencephalic lesion, type strain *Torulopsis neoformans*	B	4b	4	DQ007958
CBS 1930	Sick goat, Aruba	B	3	6	DQ007952
CBS 5757	Unknown	B	4a	4	DQ007955
CBS 5758	Unknown	C	5	5	DQ007964
CBS 6290	Man, Zaire	B	4c	4	DQ007959
CBS 6289	Subculture of type strain RV 20186	B	4c	4	DQ007963
CBS 6955	Spinal fluid, type strain of *Filobasidiella bacillispora*, USA	C	5	5	DQ007966
CBS 6956	Sputum, USA	B	3	6	DQ007954
CBS 6992	Man	B	4a	4	=DQ007955
CBS 6994	Cerebrospinal fluid, USA	C	5	5	=DQ007966
CBS 6996	Man	B	5	5	DQ007965
CBS 6998	Cerebrospinal fluid, Thailand	B	4a	4	DQ007956
CBS 7748	Air in hollow, *Eucalyptus camaldulensis*, Australia	B	4b	4	DQ007957
CBS 7750	Bark debris of *E. camaldulensis*, USA	B	3	6	=DQ007954
RV 5265	Cerebrospinal fluid, Zaire	B	4c	4	DQ007962
NIH 139	Patient, USA	C	5	5	=DQ007966
NIH 178	Patient, USA	C	5	5	=DQ007966
IMH 1658 CBS 8684	Nest of wasp, Uruguay	B	3	6	DQ007953
48A	Lung of a goat, Spain	B	4c	4	DQ007960
52A	Brain of a goat, Spain	B	4c	4	=DQ007960
56A	Gut of a goat, Spain,	B	4c	4	=DQ007960
59A	Lung of a goat, Spain	B	4c	4	DQ007961
60A	Lung of a goat, Spain	B	4c	4	=DQ007961
CGBMA1*	Pink shower tree, Brazil	B	3	6	DQ007971
CGBMA6*	Pink shower tree, Brazil	B	3	6	DQ007970
CGBMA15*	Pink shower tree, Brazil	B	3	6	=DQ007971
OITIGYPI10*	Pottery tree, Brazil	BC	3	6	DQ007969
OITIGYPI15*	Pottery tree, Brazil	B	3	6	=DQ007969
WM 161*	Unknown	B	5	5	DQ007968
WM 176	Eucalyptus tree, USA	B	4b	4	=DQ007958
WM 779	Human, cerebrospinal fluid, India	C	6	7	=DQ007967
B-5742	Eucalyptus tree, USA	C	6	7	DQ007967

CBS, from Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands; RV, from collection previously held at the Laboratory of Mycology, Institute of Tropical Medicine, Antwerp, Belgium, presently in the Scientific Institute of Public Health, Brussels, Belgium; RDA, from Erasmus Medical Center Rotterdam, The Netherlands; BD, BA, from Institut Pasteur, Paris, France; WM, from Wieland Meyer, Westmead Hospital, Sydney, Australia; NIH and B-5742, from National Institutes of Health, Bethesda, MD, USA; IMH, from Instituto de Hygiene, Montevideo, Uruguay; 48A, 52A, 56A, 59A 60A, from Dr. J. Torres, Barcelona, Spain; CGBMA & OITIGYP from M. Lazera, Instituto Oswaldo Cruz, Rio de Janeiro, Brazil.; GenBank#: GenBank accession number; *: IGSII sequence unavailable; =Identical sequence; N/A= not available.

(Foster City, CA, USA) using a standard manufacturer protocol. Sequence analysis was performed using the SeqMan program from DNASTAR (Madison, WI, USA).

2.4. Cycle sequence analysis and data analysis

For cycle sequence analysis, the IGSI region was divided in two segments: IGSI(a) and IGSI(b). Cycle

sequence primers for IGSI(a) employed the forward primer, IG1F and the reverse primer, IG2R. IGSI(a) sequences of this segment had been deposited in EMBL [19]. IGSI(b) segment employed IG2F as a forward primer and 5SR as the reverse primer. The IGII region sequences were obtained using the forward primer 5SF and the reverse primer IG3R. For Genbank accession numbers we refer to Table 1.

The sequences were obtained with a Li-Cor (Lincoln, NE, USA) NEN Global IR^2 Automated Sequencer using the manufacturer's protocol. Nucleotide sequences were read using Base image IR (Li-Cor) software and edited using AlignIR (Li-Cor). Sequence alignment was made with MegaAlign (DNA Star, Inc, Madison) and by manual adjustment. The phylogenetic trees were computed with PAUP*4.0 using Parsimony analyses (heuristic search, stepwise addition, random addition, nearest neighbor, 100 maximum tree). Gaps were represented as missing data. Each character was treated as an independent, unordered, multiple character of equal weight. The reliability of the clusters was calculated using bootstrap analysis with 500 replications.

3. Results

3.1. IGS sequence analysis

The sequence strategy used three sets of primers, which yielded three different segments varying in sizes: IGSI(a): (657–822 bp), IGSI(b): (436–599 bp) and IGSII (609–657 bp). Our partial IGSII sequence analyses lacked ~460 characters at the 3′ end. This segment of the IGSII region was a fairly homogenous area with approximately 14-bp base pair differences between *C. neoformans* var. *neoformans* and *C. gattii*. Complete sequences of the IGSI region were obtained combining the segments, IGSI(a) and IGSI(b). When all three segments (IGSI(a), IGSI(b) and IGSII) were pooled, sequences containing over 2 kb were generated (including 118 bp from the 5S rRNA gene) and were denominated as IGSI + 5S rRNA + IGSII.

A map depicting the organization of the intergenic spacer regions is illustrated in Fig. 1. The first 55 bp of the IGSI(a) region, which is located between the 28SLrDNA and the 5S RNA gene, showed nearly identical sequences for all the genotypes. Beyond this point, divergence in sequences was pronounced between the genotypes. These divergences were the result of nucleotide substitutions and the presence of short indels. Details about the sequence of the IGSI(a) region have been published elsewhere [19]. The first segment of the IGI(b) (~355 bp in length), which starts at position 877 bp in our alignment, displayed random nucleotide substitutions and few indel areas of one or two bp, except at position 1049, where a 7-bp indel was located. Several long indel areas (location: 1232–1467), consisting of up to 118 bp, were also observed and were the most prominent feature of the region. At the 3′ end of the IGSI, a modified TATA box (position 1500) with a TATAAT consensus sequence was identified flanking the 5S RNA gene of IGS genotypes 1 and 2 (Fig. 1). A high similarity in sequences was found in the 5S rRNA gene, which divided the IGS region into two regions (Fig. 1). Immediately following the 3′ end of the 5S RNA, there was a CT- and T-rich area containing two consecutive blocks of seven T motifs separated by GC residues. This area, which represents the 5′ end of the IGSII, has been described as the site where the RNA polymerase III terminator binds [28]. About 30 bases downstream, an area comprising over 80 nucleotides downstream, contained mostly purine residues, G + A, arranged in GAGA, GAAAA or GAAA tandem repeats. The incidence of GAGA motifs, which extended from position 1733 to 1765, was more frequent in genotypes 1 and 2. The GAGA repeats also occured in the IGI(a) region [19]. Another feature of the IGSII was the presence of consecutive modules of TATA and ATAATA, extending from position 1938 to 2010.

The longest consensus of repeats consisted of a 19-bp motif: GTCATGGGGGACTTGGGAG. This motif was found at five locations within the IGSI(a) region. Variations of the motif between locations and genotypes are displayed in Fig. 1. Genotypes 3–4–5–6 of *C. gattii* displayed the most divergent sequences at position 432, with a similarity value of 15.8%. Low identity values (26.3–31.6%) were also observed at position 421 for genotypes 1 and 2 of *C. neoformans*. In contrast, at location 341 genotypes 1 and 2 displayed similarity values ranging from 95% to 100%. Other short repeat and imperfect copies were identified, among them CAAAAAATT occurred four times in IGSI(a), four times in IGSI(b) and three times in IGSII (Fig. 1). Another repeat, GTTTTT, occurred once in IGSI(a) and nine times in IGSII (Fig. 1). Beginning at position 1783, genotype 1c displayed five consecutive copies of the repeat, followed by genotypes 1a and 1b, with three copies each. In contrast, genotypes 2, 3, 4, 5 and 6 showed nucleotide variations and/or deletions of the repeat. Overall, most of the motifs were found as consecutive tandem arrays and/or interspersed throughout the sequence.

3.2. Identity values

The percentage of identity, which was calculated by sequence comparisons using the DNA Star Program (Clustal W alignment with manual adjustments), within and between the members of the main phylogenetic clusters in IGSI + 5S rRNA + IGSII regions are shown in Table 2. High identity or similarity values were obtained within each genotypic group. The levels of similarity values within the main genotypes were as

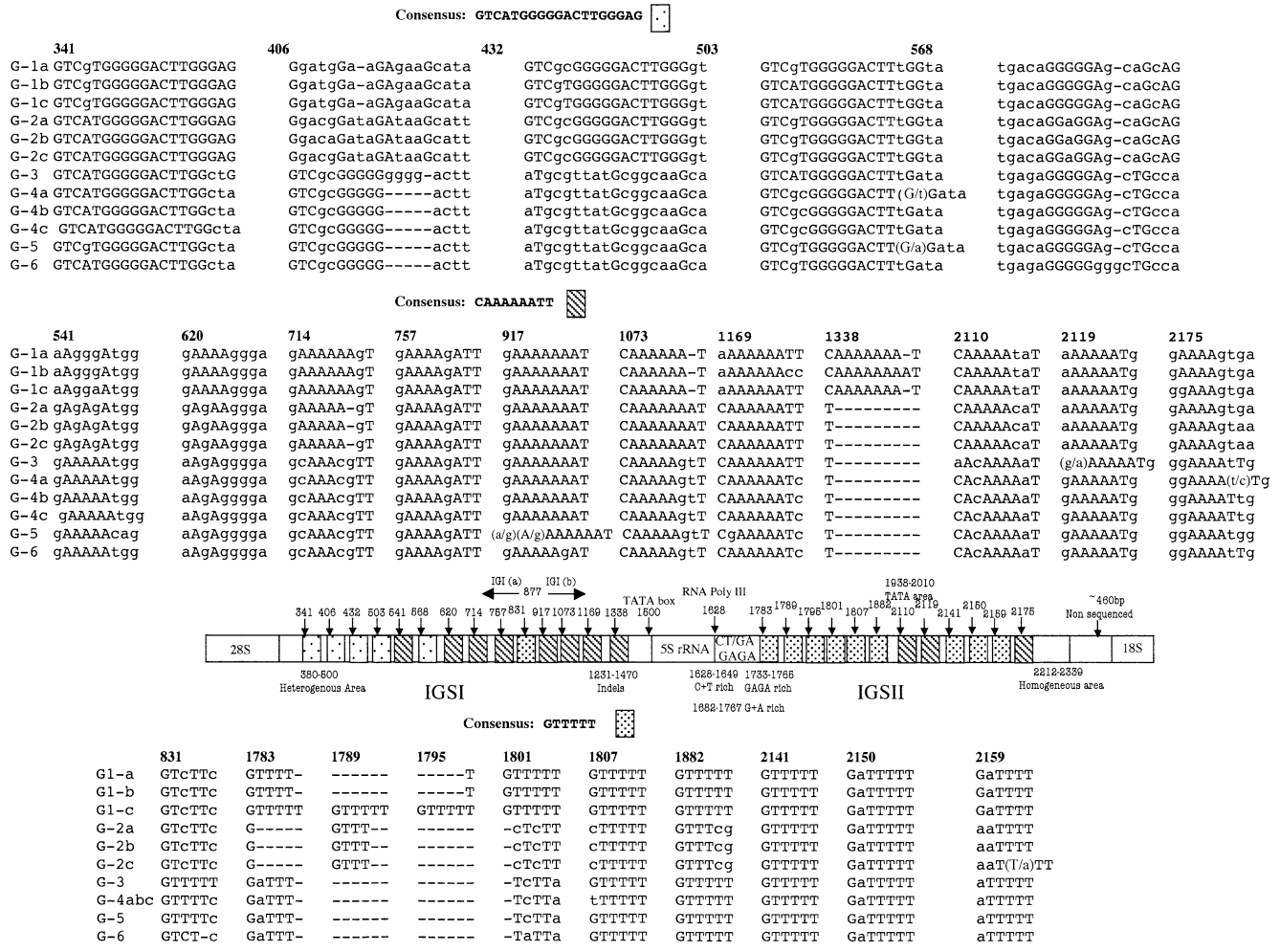

Fig. 1. Diagrammatic presentation of the repeat motifs, locations and their variations in the IGSI and IGSII regions of *C. neoformans* species complex. Numbers indicate location.

Table 2
Sequence identities among and within the genotypes of the *C. neoformans* species complex

	Genotype 1	Genotype 2	Genotype 3	Genotype 4	Genotype 5	Genotype 6
Genotypes						
G-1	97.3–100	–	–	–	–	–
G-2	78.4–79.6	98.0–100	–	–	–	–
G-3	66.9–67.6	68.7–69.4	99.5–100	–	–	–
G-4	66.0–67.7	68.4–69.5	91.7–92.9	98.4–100	–	–
G-5	66.6–67.5	68.1–69.2	92.3–92.6	95.0–96.0	99.7–100	–
G-6	66.3–67.0	68.2–69.0	90.9–91.1	93.0–94.0	93.3–94.4	100

Percent identity values were determined with Clustal W alogorithm alignment generated by MegAlign. IGSI + 5SrRN + IGSII.

follows: (a) genotype 1: 97.3–100%; (b) genotype 2: 98.0–100%; (c) genotype 3: 99.5–100%; (d) genotype 4: 99–100%; (e) genotype 5: 99.7–100% and genotype 6: 100%. Due to the sequence variability between the genotypes, lower sequence similarity values were obtained. For example, some strains within genotype 1 differed by a maximum of 78.5%, 66.9%, 66.0%, 66.6% and 66.3% from genotypes 2, 3, 4, 5 and 6, respectively (Table 2).

3.3. Length polymorphism

The length heterogeneity between the genotypes in the IGSI region was highly variable, with genotype 1

differing by an average of 185, 146, 152, 151 and 145 bp from genotypes 2, 3, 4, 5 and 6. The highest variation was found between some members of genotypes 1b and 2a, which differed by approx. 194 bp. The IGSII region displayed less length heterogeneity. The highest variation in the IGSII region (48 bp) was observed between genotypes 2b and 3. When both intergenic spacers where combined, including the 5S rRNA gene, (IGSI–5S rRNA–IGSII), length polymorphism among and within the genotypes were: genotype 1: (2157–2167 bp); genotype 2: (1963–1983 bp); genotype 3: (2046–2047 bp); genotype 4: (2006–2024 bp); genotype 5: (2017–2018 bp) and genotype 6: (2034 bp). A maximum length variation of 203 bp was documented between genotypes 1b and 2b.

3.4. Phylogenetic analysis

Phylogenetic analysis of each intergenic spacer (IGSI, IGSII) recovered the same major genotypic distribution as the combined sequence IGSI + 5S rRNA + IGSII. Therefore, to avoid redundancy, we present the tree derived from IGSI + 5S rRNA + IGSII sequence analyses (Fig. 2). The tree topology displayed six distinct

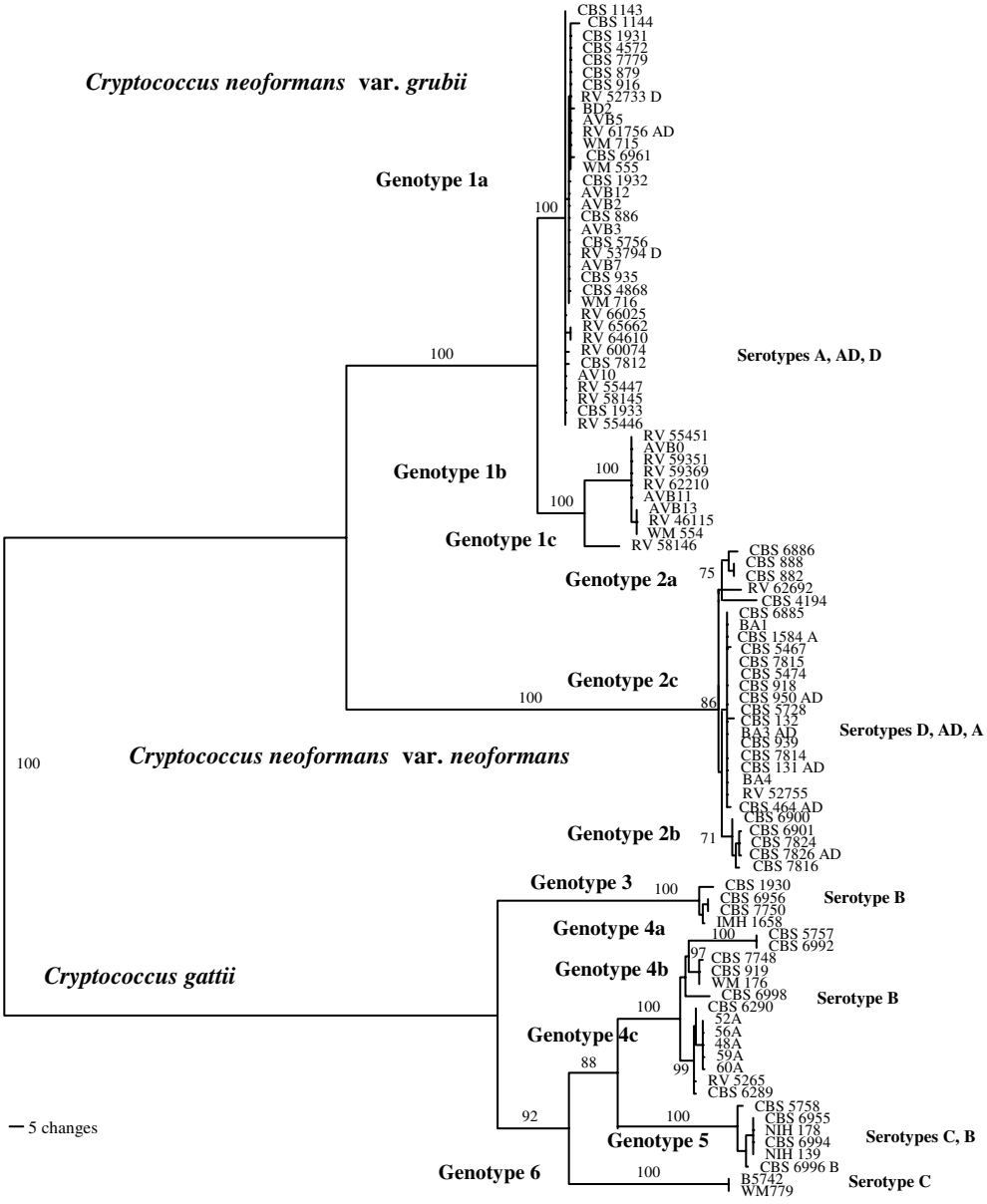

Fig. 2. Phylogenetic tree of *C. neoformans* species complex derived from IGSI + 5S rRNAt + IGSII sequence data. Presented is one of 100 most-parsimonious tree (length 948 bp; CI: 0.91; RI: 0.995) computed with PAUP*4 (heuristic search, stepwise addition, random addition sequence, nearest neighbor interchange, 100 maximum trees). Data consisted of 2350 characters (constant characters: 1582; uninformative variable characters: 46; informative characters: 722). Gaps were treated as missing data. Numbers indicate bootstrap values of 100 replicates.

genotypes representing the three varieties as three distinct phylogenetic lineages. The first clade or IGS genotype 1 was mostly represented by serotype A isolates belonging to *C. neoformans* var. *grubii*. Few isolates, which belong to serotypes D and AD, namely RV 52733 (D), RV 61756 (AD), and RV 53794 (D), also occurred in this phylogenetic group. Genotype 2, represented by *C. neoformans* var. *neoformans,* included serotype D and few isolates with serotypes AD and A (CBS 1584 (A), CBS 950 (AD), CBS 131 (AD), CBS 464 (AD), and CBS 7826 (AD). Both genotypes segregated in three distinct subgroups. Overall, genotype 1 subgroups were phylogenetically more distantly related from each other than genotype 2 sub clusters (Fig. 2). The other main phylogenetic group comprised all *C. gattii* isolates (serotype B and C) and consisted of four distinct clades represented by genotype 3 (serotype B), genotype 4 with sub-genotypes 4a, 4b and 4c (serotype B), genotype 5 (serotypes B and C) and genotype 6 (serotypes C). Genotype 6 is a newly discovered IGS genotypic group comprising two isolates, namely WM 779 and B 5742. All trees showed topological congruency, but they differed in the level of diversity among the three taxa. For example, the sequence diversity inferred from IGSI showed that *C. neoformans* var. *grubii* differed from *C. neoformans* var. *neoformans* and *C. gattii* by a maximum of 178 and 376 bp, respectively, whereas in the IGSII region, *C. neoformans* var. *grubii* differed by a maximum of 80 and 138 bp, respectively. Analysis of IGSI + 5S rRNA + IGSII showed that *C. gattii* diverged from var. *neoformans* by 548 and 506 bp from *C. neoformans* var. *grubii*. Considerable genetic diversity was also observed between the four genotypes of *C. gattii*. For instance, analysis from the IGSI + 5S rRNA + IGSII region showed that genotypes 4, 5 and 6 differed from genotype 3 by a maximum of 171, 173 and 164 bp, respectively, whereas genotype 4 and 5 differed from genotype 6 by over 125 and 127 bp.

A bootstrap confidence test, which is one of the most commonly used tests of reliability of an inferred tree, demonstrated that the six main clades were well supported in all trees (Fig. 2). Similar bootstrap values, ranging from 97 to 100, were obtained among the subgenotypes of *C. neoformans* var. *grubii* (genotypes 1a, 1b and 1c) and genotype 4 of *C. gattii* (genotypes 4a, 4b and 4c) (Fig. 2). In contrast, lower bootstrap values, ranging from 71% to 86%, were observed among some of the sub-clades representing genotypes 2a, 2b and 2c of *C. neoformans* var. *neoformans* (Fig. 2).

3.5. Cloning analysis

To investigate the hybrid nature of some strains and the possible presence of multiple alleles of the IGS locus, amplicons from serotype AD strains, e.g., RV 61756 and BA4, and IGS genotype 3 isolates, e.g., CBS 1930 (serotype B) and IMH1658 (serotype B), were cloned. In addition, cloning analyses were undertaken with discordant serotype strains that did not follow the typical serotype boundaries of *C. neoformans* var. *neoformans, C. neoformans* var. *grubii* and *C. gattii*, e.g., CBS 464 (IGS 2c – serotype A), RV 52733 (IGS 1a – serotype D); CBS 6956 (IGS 3 – serotype B), CBS 6996 (IGS 5 – serotype B), CBS 5756 (IGS 1a – serotype A), CBS 7750 (IGS 3 – serotype B). The cloned sequences were compared with sequences derived from clones of non-hybrid origin, e.g., CBS 1143 (IGS 1a – serotype A), CBS 5474 (IGS 2c – serotype D), 48A (IGS 4c – serotype B) and CBS 6955 (IGS 5 – serotype C).

Cloning sequences derived by the primer set IGF and IG2R yielded a single band with amplicon sizes ranging from ~910 (RV 52733, RV 61756, CBS 1143; CBS 464, CBS 5474) to ~430 bp (CBS 6956, CBS 6996, CBS 1930, CBS 7750, IMH 1658, CBS 6955, 48A). All cloned sequences recovered from each strain were identical and displayed identical genotypic sequences of the archived strains.

4. Discussion

4.1. Molecular structural organization and characterization of the IGS region

The overall molecular structure of the rDNA IGS of *C. neoformans* species complex follows the same characteristic features as those reported in other fungi and plants [29]. As in most basidiomycetes, the intergenic spacers of *C. neoformans* species complex are separated by the 5S rRNA gene, whose transcription occurs in the same direction as that of subunits [30,31]. Both regions consist of several repeated motifs, indels and variable regions with nucleotide substitutions. Besides the common occurrence of point mutations and short indels involving few nucleotides, the prominent feature of IGSI is the presence of extensive indel areas, which were more frequent than those in the IGSII. The different indel lengths can result from different operating mechanisms. For example, short indels can be mostly explained by processes of DNA replication such as slipped strand miss-pairing, whereas long insertions and deletions might be explained by unequal crossing-over or by DNA transposition [32,33]. Although, the two regions showed a similar phylogenetic pattern, the IGSI region contained more phylogenetic informative characters and more sequence diversity than the IGSII. The degree of sequence heterogeneity in the IGSI region suggests that this region possibly functions as a hot spot for unequal crossing-over and miss-pairing events, which is one of the multiple functions attributed to intergenic spacers to maintain homogeneity of the rDNA repeats [34]. Even though, the IGSII region revealed several

areas that can act as hot spots, this region is relatively more conserved and, therefore, can accommodate more functions related to rRNA production or processing [35].

Length variation in spacer size was found between the different IGS genotypes. Most of the observed length polymorphisms can be attributed to the presence of indels associated with repeats. The incidence of the commonly occurring GTTTTT and TATA motifs are examples of repeats that contributed to the presence of length polymorphisms between the genotypes. Also, inserts unrelated to repeat motifs are responsible for the differential size variation seen within members of the same genotypic group. The presence of repeats has also been reported in fungi like *Pleurotus cornucopiae* and *Clavispora opuntiae* [36,37].

The intergenic spacer region contains sequences that are essential for the initiation of transcription, RNA processing, transcription termination and replication processes of ribosomal DNA [35,38]. The repeat motifs, which are among the prominent features of the IGS region, are found in the transcribed region or upstream of transcription start [35,39,40]. They function as promoters, enhancers and regulators of transcription of rRNA, and possibly have originated from processes involving duplication and amplification of short sequences and from slippage replication mechanisms [40,41]. Many of the repeat motifs have been described as highly conserved in different species. For example, the repeat CAAAAA, which is a shorter version of our motif CAAAAAAT, has been described as a conservative motif in the promoter region of crucifers such as *Brassica* spp. [42], *Raphanus* spp. [43] and *Arabidopsis* [44]. Although the role of these repeats is yet to be determined, some functional significance related to transcriptional regulation is suspected. Other common repeat motifs, e.g., TATA, which are believed to be involved in the assembly of the pre-initiation complex and selection of the transcription site [45], were present in all six genotypes and have been reported to occur as a common element in various fungi such as *Schizophyllum commune* [46], *Laccaria bicolor* [47] and *Neurospora* species [48].

4.2. Phylogeny, genetic diversity and geographic substructure

Our phylogenetic analyses expanded on our earlier findings, which were based on partial sequence analysis of the IGSI region [19]. In addition, we included a new IGS genotypic group (IGS genotype 6) based on unique sequences and a discrete clustering pattern of two serotype-C isolates originating from India and South Africa. These isolates have recently been reported to belong to a new PCR-fingerprinting molecular type VGIV, and AFLP type 7 [12,49]. Isolates from this molecular type have also been found in Colombia, Mexico and Canada [12,16]. Based on our phylogenetic analyses which were supported by high bootstrap values, the *C. neoformans* species complex consists of six major IGS types. Besides these genotypic groups, our sequence analyses of IGSI + 5S rRNA + IGSII revealed the presence of additional sub-genotypes, e.g., genotypes 4a, 4b and 4c (Fig. 2). As in our previous study [19], our IGS data were not fully concordant with the traditional classification based on serotypes, which described serotype A isolates as *C. neoformans* var. *grubii* [15] and serotype D as *C. neoformans* var. *neoformans* [1].

The level of divergence between the genotypes increased when both IGS regions were included in the phylogenetic analysis. Overall, our IGS genetic lineages are in agreement with the genotypic classification as inferred by AFLP [17], M13/URA 5 analysis [16], PCR-fingerprinting/RAPD analysis [50], and ITS/RAPD profiles [49]. Boekhout et al. [17] have described six molecular AFLP types, which correlate with our current classification. The equivalent types are: IGS genotype 1 = AFLP 1 + AFLP 1A; IGS genotype 2 = AFLP 2; IGS genotype 3 = AFLP 6; IGS genotype 4 = AFLP 4; IGS genotype 5 = AFLP 5 and IGS genotype 6 = AFLP 7 [12]. AFLP analysis recognized the presence of an additional phylogenetic lineage (i.e., AFLP 3), comprising isolates of hybrid origin between serotype A and D [17]. In contrast, our phylogenetic analysis placed these isolates (i.e., CBS 131, CBS 132, CBS 464, CBS 939, CBS 950, RV 52733, RV 52755, BA1, BA3, BA4 and RV 53794) as members of IGS genotypes 1 or 2 (Fig. 2). Even though, our IGS genotypic classification correlated with major AFLP groups, differences in subgroup classification were evident. For example, AFLP analysis failed to discern the IGS sub-genotypes 1b and 1c. Similarly, AFLP type 2 did not segregate into subtypes, which contrasts with our findings on the IGS genotype 2. In addition, isolates comprising the molecular subtypes AFLP 4B and AFLP 4A clustered randomly in our IGS genotypes 4a, 4b and 4c. This apparent lack of correlation among the subtype-related phylogenetic patterns may be attributed to inherent differences in the nature of both techniques. Based on sequence analysis of the ITS region, seven genotypes have also been described within the species complex. Via combinations of eight bp differences at various locations in the ITS1 and ITS2 regions, Katsu et al. [49] have described a genotyping technique that agreed with our IGS classification. Other fingerprinting techniques, such as PCR fingerprint using minisatellite (M13), URA 5-RFLP typing and PCR–RFLP of the PLB1 genes, grouped the varieties of *C. neoformans* into eight major molecular types [16,51]. However, if we include our IGS subtypes as independent lineages, our phylogeny could delineate 12 molecular subtypes. The further discrimination into 12 molecular sub-genotypes was possible since the IGS region is a fast-evolving region, which portrays

the highest amount of sequence variation within the rDNA [25]. The observed degree of genetic heterogeneity among the IGS genotypes and sub-genotypes can be a valuable tool to understand the global epidemiology of cryptococcosis.

The segregation of *C. neoformans* var. *grubii* and *C. neoformans* var. *neoformans* into two distinct clades has been confirmed by multiple gene analysis, i.e., IGS, ITS, laccase gene, mitochondrial large ribosomal-subunit RNA (mtLrRNA), topoisomerase (*TOPI*), Cap59, 26S LrRNA and URA5 gene [7,19,20,52,53]. Some of these studies have reported the presence of considerable sequence divergence between these varieties, namely up to 6% for the URA5 gene [7], 3.6% for the topoisomerase I gene [1] and 8% for a partial analysis of the IGSI region [19]. When both intergenic spacers (IGSII + 5S rRNA + IGSII) were combined, the sequence divergence increased to nearly 12%. Other studies based on fingerprinting-PCR techniques, i.e., UT-4p probe, CNRE1 probe and multilocus enzyme electrophoresis profile, confirmed the separation of the serotypes A and D [16,54,55]. In view of these significant differences, Franzot et al. [15], have proposed variety *grubii* for serotype A isolates. We consider it likely that IGS genotype 2 (=AFLP genotype 2) represents an individual species, but multilocus sequence typing data are needed to corroborate this hypothesis. The different phylogenetic placement of serotype-AD isolates, and the switching pattern of some isolates from serotype AD to serotype A or D, further complicates the current classification, as these isolates may represent intervarietal or even interspecies hybrids [17,19,52,56].

Lack of concordance between serotype and genotype delineation has also been documented for *C. gattii*. Our phylogenetic studies demonstrated that *C. gattii* showed more genetic diversity than *C. neoformans.* This was demonstrated by the presence of four phylogenetic lineages (i.e., genotypes 3–6), which were all supported by high bootstrap values. This genetic diversity among *C. gattii* isolates has also been documented by Latouche et al. [51], who reported five distinct lineages employing PCR–RFLP of the *PLBI gene*. The phylogenetic clustering based on IGSI + 5S rRNA + IGSII indicated that IGS genotype 3 is closer related to genotype 4 than to genotypes 5 and 6, whereas genotype 6 was phylogenetically closer related to genotypes 4 and 5 than to genotype 3. Genotype 6 differed from genotypes 3, 4 and 5 by ~6%, 4.2%, 4.9% nucleotide substitutions, respectively. Based on rate of nucleotide substitutions in the ITS region, Katsu et al. [49] have shown that ITS type 6, which corresponds to IGS genotype 6, was phylogenetically closer related to ITS type 5 (=IGS genotype 5) and ITS type 4 (=IGS genotype 3).

Overall, the IGS of *C. gattii* differed by up to 28.4% and 30.1% from *C. neoformans* var. *neoformans* and *C. neoformans* var. *grubii*, respectively. In contrast, the URA5 gene revealed only 8% sequence divergence between *C. gattii* and the varieties of *C. neoformans* [57], and nearly 100% similarity has been reported for the other ribosomal genes (16SrDNA: two differences; 2SrDNA: ten differences; 5.8S and 5S rDNA: no differences). Our data support the previous recognition of the occurrence of at least two species within the *C. neoformans* species complex [14].

Sequence analysis of cloned PCR amplicons of the IGS did not indicate the presence of hybridization events, contrary to results obtained by PCR analysis employing specific primers for *CNA1, CLA4* and *GPA1* genes [58], PCR-fingerprinting with $(GACA)_4$ primers [17,59,60], AFLP analysis [17,60], RAPD analysis of the PLB1 gene [51, and cloning analysis of the ITS region [49]. Our IGS cloning analysis revealed that the two serotype-AD strains RV61756 and BA4 appear to have a single allele or that there is a strong selection for a single locus. However, more samples need to be analyzed to elucidate the allelic structure of serotype-AD isolates at the IGS locus. Although serotype-AD strains are known to be diploids [17,49,51,58,59,61,62], there are many AD isolates reported as haploids [62–65] and aneuploids [7]. For example, the existence of a single *STE20* allele and one mating-type allele has been reported to occur in serotype-AD strains [63,64].

Hybridization events may also explain part of the genetic heterogeneity displayed within members of *C. gattii* isolates. For instance, Sugita et al. [52] have found that serotype B exhibited higher genetic divergence than serotype C isolates. Although the latter observation is supported by our and previous studies [17,19], the inclusion of IGS genotype 6 (serotype C) as a new phylogenetic group within *C. gattii* indicates that serotype C isolates are phylogenetically diverse. Contrary to previous suggestions [17], our cloning analysis with *C. gattii* isolates (IGS genotype 3 = AFLP 6) indicated that these strains are not heterozygotous diploids for the IGS locus. Similar findings based on cloning analysis of *PLB1* fragments have been reported by Latouche et al. [51]. IGS genotype 3, which was previously thought to be restricted to tropical and subtropical America [17,19], has been reported in Brazil [60], Thailand, Australia [49], and Canada [12,13].

Overall, the apparent lack of geographic concordance with phylogeny suggests that *C. neoformans* has undergone recent global dispersal [20]. This recent dispersal model might explain why IGS sequences did not discriminate among isolates from different geographic locations.

This study provided further evidence on the genetic diversity among the *C. neoformans* species complex and further supports the recognition of at least two, and probably three species within the complex. The present study confirms the utility of the IGS region as a powerful tool for species or varietal identification. The divergences, which are mostly represented by nucleotide variations and by deletion/insertion areas, can

serve as a platform to design specific probes for rapid and concise identification of the species, varieties and genotypes of *C. neoformans* and *C. gattii* [66] which may be very useful for a better understanding of the causes of infection and the epidemiology of cryptococcosis.

Acknowledgments

Our sincere gratitude to B. Theelen and A. Statzell-Tallman for their assistance in DNA isolation and to Drs. M. Lazera, F. Coenjaerts, K.J. Kwon-Chung, S. Hamdan, J. Torres-Rodriguez, D. Swinne, W. Meyer, A. van Belkum, M.T. Barreto de Oliviera, and F. Frommer for providing strains. This work was funded by National Institutes of Health Grant 1-UO1 AI53879-01.

References

[1] Casadevall, A. and Perfect, J.R. (1998) *Cryptococcus neoformans*. ASM Press, Washington, DC.
[2] White, M., Cirrincione, C., Blevins, A. and Armstrong, D. (1992) Cryptococcal meningitis with AIDS and patients with neoplastic disease. J. Infect. Dis. 165, 960–966.
[3] Levitz, S.M. (1999) Macrophage interactions with *Cryptococcus neoformans*. ASM News 64, 693–698.
[4] Powderly, W.G. (1993) Cryptococcal meningitis and AIDS. Clin. Infect. Dis. 17, 837–842.
[5] Banerjee, U., Datta, K. and Casadevall, A. (2004) Serotype distribution of *Cryptococcus neoformans* in patients in a tertiary center in India. Med. Mycol. 42, 181–186.
[6] Baro, T., Torres-Rodriguez, J.M., Morera, Y., Alia, C., Lopez, O. and Mendez, R. (1999) Serotyping of *Cryptococcus neoformans* isolates from clinical and environmental sources in Spain. J. Clin. Microbiol. 37, 1170–1172.
[7] Franzot, S.P., Fries, B.C., Cleare, W. and Casadevall, A. (1998) Genetic relationship between *Cryptococcus neoformans* var. *neoformans* strains of serotypes A and D. J. Clin. Microbiol. 36, 2200–2204.
[8] Rozenbaum, R., Gonçalves, A.J., Wanke, B. and Viera, W. (1990) *Cryptococcus neoformans* var. *gattii* in a Brazilian AIDS patient. Mycopathologia 112, 33–34.
[9] O'Brien, C.R., Krockenberger, M.B., Wigney, D.I., Martin, P. and Malik, R. (2004) Retrospective study of feline and canine cryptococcosis in Australia from 1981 to 2001: 195 cases. Med. Mycol. 42, 449–460.
[10] Lazera, M.S., Pires, F.D.A., Camiloo-Coura, L., Nishikawa, M.M., Bezerra, C.C.F., Trilles, L. and Wanke, B. (1996) Natural habitat of *Cryptococcus neoformans* var. *neoformans* in decaying wood forming hollows in living trees. J. Med. Vet. Mycol. 34, 127–131.
[11] Lazera, M.S., Wanke, B. and Nishikawa, N.M. (1993) Isolation of the two varieties of *Cryptococcus neoformans* from saprophytic sources in the city of Rio de Janeiro, Brazil. J. Med. Vet. Mycol. 31, 449–454.
[12] Kidd, S.E, Hagen, F., Tscharke, R.L., Huynh, M., Barlett, K.H, Fyfe, M., Macdougall, L., Boekhout, T., Kwong-Chung, K.J. and Meyer, W. (2004) A rare genotype of *Cryptococcus gattii* caused the cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). Proc. Natl. Acad. Sci. USA 101, 17258–17263.
[13] Stephen, C, Lester, S., Black, W., Fyfe, M. and Raverty, S. (2002) Multispecies outbreak of cryptococcosis on southern Vancouver Island, British Columbia. Can. Vet. J. 43, 792–794.
[14] Kwong-Chung, K.J, Boekhout, T., Fell, J.W. and Diaz, M. (2002) Proposal to conserve the name *Cryptococcus gattii* against *C. hondurianus* and *C. bacillisporus* (Basidiomycota, Hymenomycetes, Tremellomycetidae. Taxon 51, 804–806.
[15] Franzot, S.P., Salkin, I.F. and Casadevall, A. (1999) *Cryptococcus neoformans* var. *grubii* separate varietal status for *Cryptococcus neoformans* serotype A isolates. J. Clin. Microbiol. 37, 838–840.
[16] Meyer, W., Castaneda, A., Jackson, S., Huynh, M. and Castaneda, E. (2003) Molecular typing of IberoAmerican *Cryptococcus neoformans* isolates. Emerg. Infect. Dis. 2, 189–195.
[17] Boekhout, T., Theelen, B., Diaz, M., Fell, J.W., Wim, C.J., Edwin, C.A., Albeln, E.C., Dromer, F. and Meyer, W. (2001) Hybrid genotypes in the pathogenic yeast *Cryptococcus neoformans*. Microbiology 147, 891–907.
[18] Boekhout, T., Van Belkum, A., Leenders, A.C., Verbrugh, H.A., Mukamurangwa, P., Swinne, D. and Scheffers, W.A. (1997) Molecular typing of *Cryptococcus neoformans*: taxonomic and epidemiological aspects. Int. J. Syst. Bacteriol. 47, 432–442.
[19] Diaz, M.R., Fell, J.W, Boekhout, T. and Theelen, B. (2000) Molecular sequence analyses of the intergenic spacer (IGS) associated with rDNA of the two varieties of the pathogenic yeast, *Cryptococcus neoformans*. Syst. Appl. Microbiol. 23, 535–545.
[20] Xu, J., Vilgalys, R. and Mitchell, T.G. (2000) Multiple gene genealogies reveal recent dispersion and hybridization in the human pathogenic fungus *Cryptococcus neoformans*. Mol. Ecol. 9, 1471–1481.
[21] Chen, F., Currie, B.P, Chen, L.C., Spitzer, E.D. and Casadevall, A. (1995) Genetic relatedness of *Cryptococcus neoformans* clinical isolates grouped with the repetitive DNA probe CNRE-I. J. Clin. Microbiol. 33, 2818–2822.
[22] Diaz, M.R. and Fell, J.W. (2004) High-throughput detection of pathogenic yeasts in the genus *Trichosporon*. J. Clin. Microbiol. 42, 3696–3706.
[23] Diaz, M.R. and Fell, J.W. (2000) Molecular analyses of the IGS and ITS regions rDNA of the psychrophilic yeasts in the genus *Mrakia*. Antonie van Leeuwenhoek 77, 7–12.
[24] Fell, J.W. and Blatt, G. (1999) Separation of strains of the yeasts *Xanthophyllomyces dendrorhous* and *Phaffia rhodozyma* based on rDNA IGS and ITS sequence analysis. J. Indust. Microbiol. Biotech. 21, 677–681.
[25] Polanco, C. and Perez de la Vega, M. (1994) The structure of the rDNA intergenic spacer of *Avena sativa* L.: a comparative study. Plant Mol. Biol. 25, 751–756.
[26] Fell, J.W., Boekhout, T., Fonseca, A., Scorzetti, G. and Statzell-Tallman, A. (2000) Biodiversity and systematics of basidiomycetous yeasts as determined by large subunit rDNA D1/D2/domain sequence analysis. Int. J. Syst. Bacteriol. 50, 1351–1371.
[27] O'Donnell, K., Cigelnik, E., Weber, N.S. and Trappe, J.M. (1997) Phylogenetic relationships among ascomycetous truffles and the true and false morels inferred from 18S and 28S rDNA sequence analysis. Mycologia 89, 48–65.
[28] Clark, C.G. (1987) On the evolution of ribosomal RNA. J. Mol. Evol. 25, 343–350.
[29] Cordesse, F., Cooke, R., Tremousaygue, D., Grellet, F. and Delseney, M. (1993) Fine structure of the rDNA intergenic spacer in rice and other cereals. J. Mol. Evol. 36, 369–379.
[30] Fan, M., Chen, L.C., Ragan, M.A., Gutell, R.R., Warner, J.R., Currie, B.P. and Casadevall, A. (1995) The 5S rRNA and the rRNA intergenic spacer of the two varieties of *Cryptococcus neoformans*. J. Med. Vet. Mycol. 33, 215–221.
[31] Kwon-Chung, K.J. and Chang, Y.C. (1993) Gene arrangement and sequence of the 5S rRNA in *Filobasidiella neoformans* (*Cryptococcus neoformans*) as a phylogenetic indicator. Int. J. Syst. Biol. 44, 209–213.

[32] Hancock, J.M. and Dover, G.A. (1990) Compensatory slippage in the evolution of ribosomal RNA genes. Nucleic Acids Res. 20, 5949–5954.
[33] Jemtland, R, Maehhlum, E., Gabrielsen, O.S and Oyen, T.E. (1986) Regular distribution of length heterogeneities within non-transcribed spacer regions of cloned and genomic rDNA of *Saccharomyces cerevisiae*. Nucleic Acids Res. 14, 5145–5158.
[34] Dover, G. (1982) Molecular drive: A cohesive mode of species evolution. Nature 299, 111–117.
[35] Fernandez, M., Polanco, C., Ruiz, M.L. and Perez de la Vega, M. (2000) A comparative study of the structure of the rDNA intergenic spacer of *Lens culinaris* Medik. and other legume species. Genome 43, 597–602.
[36] Lachance, M.A., Starmer, W.T., Bowles, J.M., Phaff, H.J. and Rosa, C.A. (2000) Ribosomal DNA, species structure, and biogeography of the cactophilic yeast *Clavispora opuntiae*. Can. J. Microbiol. 46, 195–210.
[37] Iraçabal, B. and Labarére, J. (1994) Restriction site and length polymorphism of the rDNA unit in the cultivated basidiomycete *Pleurotus cornucopiae*. Theor. Appl. Genet. 88, 824–830.
[38] Van't Hof, J. and Lamm, S.S. (1991) Single stranded replication intermediates of ribosomal DNA replicons of pea. EMBO J. 10, 1949–1953.
[39] Grellet, F., Delcasso-Tremousaygue, D. and Delseney, M. (1989) Isolation and characterization of an unusual repeated sequence from the ribosomal intergenic spacer of the crucifer *Sisymbrium irio*. Plant Mol. Biol. 12, 695–706.
[40] Barker, R.F., Harberd, N.P., Harvis, M.G. and Flavell, R.B. (1988) Structure and evolution of the intergenic region of a ribosomal DNA repeat unit of wheat. J. Mol. Biol. 201, 1–17.
[41] Stupar, R.M, Song, J., Tek, A.L., Cheng, Z.K., Dong, F. and Jiang, J. (2002) Highly condensed potato pericentromeric heterochromatin contains rDNA-related tandem repeats. Genetics 162, 1435–1444.
[42] Bhatia, S., Negi, M.S. and Lakshmikumaran, M. (1996) Structural analysis of the rDNA intergenic spacer of *Brassica nigra*: Evolutionary divergence of the spacers of the three diploid *Brassica* species. J. Mol. Evol. 43, 460–468.
[43] Delcasso-Tremousaygue, D., Grellet, F., Panabieres, F., Ananiev, E.D. and Delseney, M. (1988) Structural and transcriptional characterization of the external spacer of a ribosomal RNA nuclear gene from a higher plant. Eur. J. Biochem. 172, 767–776.
[44] Gruendler, P., Unfried, I., Pointner, R. and Schweizer, C. (1991) Nucleotide sequence of the 25S–18S ribosomal gene spacer from *Arabidopsis thaliana*. Nucleic Acids Res. 17, 6395–6396.
[45] Melané, A., Vivier, M.A. and Pretorius, I.S. (1998) Identification of a functional TATA element in the *STA2* glucoamylase gene promoter from *Saccharomyces cerevisiae*. Curr. Genet. 33, 10–15.
[46] James, T.Y., Moncalvo, J.M., Li, S. and Vilgalys, R. (2001) Polymorphism at the ribosomal DNA spacers and its relation to breeding structure of the widespread mushroom *Schizophyllum commune*. Genetics 157, 149–161.
[47] Martin, F., Selosse, M.A. and Le Tacon, F. (1999) The nuclear rDNA intergenic spacer of the ectomycorrhizal basidiomycete *Laccaria bicolor*: structural analysis and allelic polymorphism. Microbiology 145, 1605–1611.
[48] Selker, E.U., Morzycka-Wroblewska, E., Stevens, J.N. and Metzenberg, R.L. (1986) An upstream signal is required for in vitro transcription of *Neurospora* 5S RNA genes. Mol. Gen. Genet. 205, 189–192.
[49] Katsu, M., Kidd, S., Ando, A., Moretti-Branchini, M.L., Mikami, Y., Nishimura, K. and Meyer, W. (2004) The internal transcribed spacers and 5.8S rRNA gene show extensive diversity among isolates of the *Cryptococcus neoformans* species complex. FEMS Yeast Res. 4, 377–388.
[50] Meyer, W., Marszewska, K., Amirmostofian, M., Igreja, R.P., Hardtke, C., Methling, K., Viviani, M.A., Chindamporn, A., Sukroongreung, T.C., John, M.A., Ellis, D.H. and Sorrell, T.C. (1999) Molecular typing of global isolates of *Cryptococcus neoformans* var. *neoformans* by polymerase chain reaction fingerprinting and randomly amplified polymorphic DNA – a pilot study to standardize techniques on which to base a detailed epidemiological survey. Electrophoresis 20, 1790–1799.
[51] Latouche, G.N., Huynh, M., Sorrel, T.C and Meyer, W. (2003) PCR–restriction fragment length polymorphism analysis B (*PLB1*) gene for subtyping of *Cryptococcus neoformans* isolates. Appl. Environ. Microbiol. 69, 2080–2086.
[52] Sugita, T., Ikeda, R. and Shinoda, T. (2001) Diversity among strains of *Cryptococcus neoformans* var. *gattii* as revealed by a sequence analysis of multiple genes and a chemotype analysis of capsular polysaccharide. Microbiol. Immunol. 45, 757–768.
[53] Guého, E., Improvisi, L., Christen, R. and de Hoog, G.S. (1993) Phylogenetic relationship of *Cryptococcus neoformans* and some related basidiomycetous yeasts determined from partial large subunit rRNA sequences. Antonie van Leeuwenhoek 63, 175–189.
[54] Brandt, M.E., Bragg, S.L. and Pinner, R.W. (1993) Multilocus enzyme typing of *Cryptococcus neoformans*. J. Clin. Microbiol. 31, 2819–2823.
[55] Varma, A. and Kwon-Chung, K.J. (1992) DNA probe for strain typing of *Cryptococcus neoformans*. J. Clin. Microbiol. 30, 2960–2967.
[56] Xu, J., Luo, G., Vilgalys, R., Brandt, M.E. and Mitchell, T.G. (2002) Multiple origins of hybrid strains of *Cryptococcus neoformans* with serotype AD. Microbiology 148, 203–212.
[57] Casadevall, A. and Fan, M. (1992) URA 5 gene of *Cryptococcus neoformans* var. *gattii*. Evidence for a close phylogenetic relationship between *Cryptococcus neoformans* var. *gattii* and *Cryptococcus neoformans* var. *neoformans*. J. Gen. Appl. Microbiol. 38, 491–495.
[58] Lengeler, K.B., Cox, G.M. and Heitman, J. (2001) Serotype AD strains of *Cryptococcus neoformans* are diploid or aneuploid and are heterozygous at the mating-type locus. Infect. Immun. 69, 115–122.
[59] Cogliati, M., Esposto, M.C., Clarke, D.L., Wickes, B.L. and Viviani, M.A. (2001) Origin of *Cryptococcus neoformans* var. *neoformans* diploid strains. J. Clin. Microbiol. 39, 3889–3894.
[60] Barreto de Oliviera, M.T., Boekhout, T., Theelen, B., Hagen, F., Baroni, A., Lazera, M.S., Lengeler, K.B., Heitman, J., Rivera, I.N.G. and Paula, C.R. (2004) *Cryptococcus neoformans* shows a remarkable genotypic diversity in Brazil. J. Clin. Microbiol. 42, 1356–1359.
[61] Tanaka, R., Nishimura, K., Imanishi, Y., Takahashi, I., Hata, Y. and Miyaji, M. (2003) Analysis of serotype AD strains from F1 progenies between urease positive and negative strains of *Cryptococcus neoformans*. Nippon Ishinkin Gakkai Zasshi 44, 293–297.
[62] Takeo, K., Tanaka, R., Taguchi, H. and Nishimura, K. (1993) Analysis of ploidy and sexual characteristics of natural isolates of *Cryptococcus neoformans*. Can. J. Microbiol. 39, 363–958.
[63] Chaturverdi, V., Fan, J., Stein, B., Behr, M.J., Samsonoff, W.A., Wickes, B.L. and Chaturverdi, S. (2002) Molecular genetic analyses of mating pheromones reveal intervariety mating or hybridization in *Cryptococcus neoformans*. Infect. Immun. 70, 5225–5235.
[64] Yan, Z., Li, X. and Xu, J. (2002) Geographic distribution of mating type alleles of *Cryptococcus neoformans* in four areas of the United States. J. Clin. Microbiol. 40, 965–972.
[65] Tanaka, R., Nishimura, K. and Miyaji, M. (1999) Ploidy of serotype AD strains of *Cryptococcus neoformans*. Nippon Ishinkin Gakkai Zasshi 40, 31–34.
[66] Diaz, M.R. and Fell, J.W. (2005). Use of a suspension array for rapid identification of the varieties and genotypes of *Cryptococcus neoformans* species complex. J. Clin. Microbiol., in press.