A Genetic Linkage Map of Cryptococcus neoformans variety neoformans Serotype D (Filobasidiella neoformans)

Robert E. Marra,*Johnny C. Huang,* Eula Fung,† Kirsten Nielsen,* Joseph Heitman,* Rytas Vilgalys*,† and Thomas G. Mitchell*

*Department of Molecular Genetics and Microbiology and †Howard Hughes Medical Institute, Duke University Medical Center, Durham, North Carolina 27710, ‡Department of Biology, Duke University, Durham, North Carolina 27710 and
§Stanford Genome Technology Center, Palo Alto, California 94304

Manuscript received October 21, 2003
Accepted for publication January 16, 2004

ABSTRACT

To construct a genetic linkage map of the heterothallic yeast, Cryptococcus neoformans (Filobasidiella neoformans), we crossed two mating-compatible strains and analyzed 94 progeny for the segregation of 301 polymorphic markers, consisting of 228 restriction site polymorphisms, 63 microsatellites, two indels, and eight mating-type (MAT)-associated markers. All but six markers showed no significant (P < 0.05) segregation distortion. At a minimum LOD score of 6.0 and a maximum recombination frequency of 0.30, 20 linkage groups were resolved, resulting in a map length of ~1500 cM. Average marker density is 5.4 cM (range 1–28.7 cM). Hybridization of selected markers to blots of electrophoretic karyotypes unambiguously assigned all linkage groups to chromosomes and led us to conclude that the C. neoformans genome is ~20.2 Mb, comprising 14 chromosomes ranging in size from 0.8 to 2.3 Mb, with a ratio of ~13.2 kb/cM averaged across the genome. However, only 2 of 12 ungrouped markers hybridized to chromosome 10. The hybridizations revealed at least one possible reciprocal translocation involving chromosomes 8, 9, and 12. This map has been critical to genome sequence assembly and will be essential for future studies of quantitative trait inheritance.

The fungus Cryptococcus neoformans (teleomorph is Filobasidiella neoformans) is an opportunistic pathogen of worldwide distribution and the leading cause of fungal infections of the central nervous systems in humans and other mammals (Casadevall and Perfect 1998; Stephen et al. 2002). Infections are acquired by inhalation of desiccated yeast cells or basidiospores (Sukroongreung et al. 1998), which may subsequently disseminate from the lungs to the central nervous system, the skin, the prostate, the eyes, and elsewhere (Mitchell and Perfect 1995). In addition to its broad host range, C. neoformans is cosmopolitan and nearly ubiquitous in nature, having been isolated from trees, soil, insects, and most frequently, the excreta of birds, especially pigeons. Whether these habitats represent the primary environmental sources of infectious propagules is not yet known; indeed, an understanding of the natural history of C. neoformans remains incomplete.

C. neoformans exists normally in its anamorphic or asexual state, as a haploid yeast (Casadevall and Perfect 1998). Low genetic diversity, the widespread occurrence of identical genotypes, and tests for recombination all support the hypothesis that the mode of reproduction is primarily clonal (Brandt et al. 1996; Franzot et al. 1997; Litvintseva et al. 2003). However, phylogenetic data have revealed a low frequency of current or historical recombination (Franzot et al. 1997; Xu et al. 2000; Litvintseva et al. 2003; Xu and Mitchell 2003). Although not observed in nature, sexual reproduction has been demonstrated in the laboratory. The teleomorph or sexual state is designated F. neoformans (Kwon-Chung 1975, 1976) and is characterized by dikaryotic filamentous hyphae with clamp connections and terminal basidia that produce long chains of basidiospores. Sexual self-incompatibility in C. neoformans is bipolar: mating compatibility requires opposite alleles, MATa and MATa, at the mating-type locus (Kwon-Chung and Hill 1981). However, most environmental and clinical strains possess only the MATa allele (Kwon-Chung and Bennett 1978; Kwon-Chung et al. 1992; Yan et al. 2002; Litvintseva et al. 2003). The prevalence of MATa isolates may be due to “haploid fruiting” (Wickes et al. 1996), a form of asexual reproduction characteristic of some MATa strains in response to starvation conditions or the proximity of mating partners (Wickes et al. 1996; Wang et al. 2000). In the plant pathogen Microbotryum violaceum, mating-type bias is due to recessive lethal mutations that are tightly linked to the MATa allele (Kaltz and Shykoff 1997; Oudemans et al. 1998), although no evidence for this type of mating-type bias has been identified in C. neoformans.

Phylogenetic studies place C. neoformans in the order

1Corresponding author: Department of Plant Pathology and Ecology, Connecticut Agricultural Experiment Station, PO Box 1106, New Haven, CT 06504. E-mail: robert.marra@po.state.ct.us

Genetics 167: 619–631 (June 2004)
Tremellales, a lineage that diverged early from other orders within the class Hymenomycetes (Fell et al. 2000). The Tremellales bridge the Hymenomycetes with its sibling classes Ustilaginomycetes and Uredinomycetes, the plant-pathogenic smuts and rusts (Swann and Taylor 1995). The polysaccharide capsule of C. neoformans, an important virulence factor (Dykstra et al. 1977; Chang and Kwon-Chung 1998; Bose et al. 2003), expresses one of five serotypes—A, D, AD, B, or C—as well as a sixth designation, “untypeable.” Additional phenotypic, epidemiological, ecological, physiological, and genetic data classify isolates into one of three varieties, C. neoformans var. gattii (composed of serotypes B and C), var. neoformans (serotype D), and var. grubii (serotype A), which causes most human infections (Mitchell and Perfect 1995; Casadevall and Perfect 1998; Hull and Heitman 2002). Phylogenetic evidence suggests that var. gattii diverged from the other two varieties ~37 million years ago and that var. grubii and var. neoformans diverged ~19 million years ago (Xu et al. 2000).

C. neoformans is a globally important human pathogen, largely due to its high prevalence among patients with AIDS. It has become a model for fungal pathogenesis and is among several fungal species that have been chosen for complete genome sequencing (http://www-genome.wi.mit.edu/annotation/fungi/C.neoformans/). Although most clinical strains belong to the serotype A group, an isolate of serotype D was the first strain chosen for sequencing because at the time it was the only serotype with a defined sexual cycle and tractable teleomorph. The availability of congeneric mating strains in serotype D enabled molecular and genetic studies that until recently were not possible in serotype A. The sequencing effort for serotype D MATα strain JEC21 and the related serotype D MATα strain B-3501 (Heitman et al. 1999) was undertaken by two groups, The Institute for Genomic Research (TIGR; http://www.tigr.org/) and the Stanford Genome Technology Center (SGTC; http://www.sequence.stanford.edu/group/C.neoformans). Shotgun sequencing of both strains has been completed, and closure, assembly, and annotation of the genomes are currently underway. In addition, a physical map based on bacterial artificial chromosome fingerprinting (Schein et al. 2002) is providing a framework for sequence assembly.

A genetic linkage map is an important component of the C. neoformans genome project. The map provides independent data on genome organization, which is essential where repetitive elements in the sequence complicate or preclude definitive assembly. Indeed, our map has accelerated genome sequence assembly at TIGR and SGTC. A preliminary genetic map of C. neoformans was constructed largely from amplified fragment length polymorphisms (AFLPs; Vos et al. 1995) of progeny derived from a cross between a clone of the MATα strain B-3501 and its sibling MATα strain B-3502 (Forche et al. 2000). However, subsequent analyses of these progeny revealed that many either were infertile or had ambiguous mating types; in addition, the progeny population contained a substantial number of clones (our unpublished results). Upon inspection of the AFLP markers, we found that many did not reflect true genotypic polymorphisms, but likely were artifacts of the AFLP procedure (our unpublished results). The goals of the current study were to prepare a reliable and unambiguous set of single-locus genetic markers, produce a new set of progeny, create a dense linkage map with an average recombinational distance of 10 cM, and correlate linkage groups to chromosomes by hybridizing markers to electrophoretic karyotypes.

MATERIALS AND METHODS

Mapping population: We used a mapping population of 94 progeny, generated by a cross between the MATα strain B-3501 and MATα strain B-3502. B-3501 and B-3502 are siblings from a cross between MATα strain NIH433, isolated from pigeon guano in Denmark, and MATα strain NIH12, isolated from a human patient in the United States (Heitman et al. 1999). The use of strains B-3501 and B-3502 as parents of the mapping population allowed us to take full advantage of DNA sequence information to identify polymorphisms: strain B-3501 has been sequenced by SGTC, and we have discovered that strain B-3502 is isogenic with strain JEC20/B-4476 and the TIGR-sequenced strain JEC21, except at the MAT locus (Heitman et al. 1999; our unpublished results).

The crosses were performed as described by Nielsen et al. (2003). Briefly, strains B-3501 and B-3502 were grown individually for 2 days on YEPD (1% yeast extract, 2% dextrose, 2% Bacto-peptone, 1.5% agar). Small portions of each colony were admixed and patched onto V8 (pH 5) agar (Casadevall and Perfect 1998). Plates were incubated at room temperature in the dark for ~7 days until basidiospores formed. To isolate single basidiospores, areas showing basidiospore formation upon microscopic examination were excised and transferred to a fresh YEPD plate, where spores were dissected by micromanipulation and cultured. A total of 103 single-basidiospore cultures were obtained. Mating types were determined using strains JEC20 (MATα) and JEC21 (MATα) as mating-type tester strains (Kwon-Chung et al. 1992) and confirmed using MAT-specific PCR primers (see Marker development). Subsequent genotyping of progeny numbered 1–94 revealed that none were clones of the parents. However, as genotyping progressed, several progeny appeared to have identical genotypes. We identified four such genotypes that were represented by two or more clones, removed the redundant genotypes from the population, and replaced them with progeny numbered 95–103.

DNA preparation: All cultures were grown on YEPD at 30°C. DNA was extracted using the Genomic Miniprep DNA extraction kit (Whatman Bioscience, Cambridge, United Kingdom) with the following preparatory steps: ~100 μl of cells from freshly streaked plates was mixed with 700 μl of solution A plus 0.3 g of 425- to 600-μm glass beads (Sigma-Aldrich, St. Louis), followed by vortexing at high speed for 1 min and then proceeding with the kit protocol. Quality and quantity of DNA were assessed by electrophoresing aliquots of samples in 0.8% agarose gels in 0.5× Tris-borate-EDTA (TBE; Sambrook and Russell 2001). The gels were stained with 0.5 μg/ml ethidium bromide and visualized and photographed using ultraviolet light on a Gel-Doc imager (Bio-Rad, Richmond, CA). DNA concentrations were determined by comparison with bacteriophage lambda DNA. Concentrations were
adjusted to ~10 ng/μl and then diluted 10-fold and 100-fold for PCRs in 10 mM Tris-Cl, pH 8.0.

Marker development: Microsatellites: To detect microsatellites, we performed a BLAST search of the C. neoformans genome sequence database at the SGTC (http://www.sequence.stanford.edu/group/C.neoformans) for a minimum of five repeats of all possible di- and trinucleotide microsatellite motifs (Table 1). For each microsatellite locus identified, we designed PCR primers to amplify 100–300 bp of sequence flanking the locus. To determine whether the locus was polymorphic, PCR amplicons from the parents, B-3501 and B-3502, were evaluated using polyacrylamide gel electrophoresis (PAGE). Briefly, they were electrophoresed in 6.7% polyacrylamide in 1X TBE in a 30 × 38-cm Bio-Rad Sequi-Gen GT vertical gel apparatus, according to manufacturer’s instructions. Subsequent to electrophoresis, gels were silver stained according to BASSAM et al. (1991) and then exposed for 1–2 sec with white light onto X-ray duplication film (Kane X-Ray, Crofton, MD) and developed according to standard procedures. Polymorphic microsatellite markers that were used in the map and their primer sequences are provided as supplemental data (see Table S1 at http://www.genetics.org/supplemental/).

RFLPs: Because B-3502 is nearly isogenic with JEC21 (HEITMAN et al. 1999; our unpublished results), we were able to identify restriction sites that should be polymorphic between the mapping parents B-3501 and B-3502 by comparing the B-3501 and JEC21 versions of c neo011005 contig sequences. Twenty-four restriction enzyme recognition sequences (9 4-bp and 15 6-bp sequences) were queried in this manner (Table 2). Primers for PCR were designed to flank each restriction site and generate PCR amplicons of ~800 bp. To validate putative polymorphisms, PCRs using B-3501 and B-3502 DNAs were digested with the appropriate restriction enzyme and electrophoretically separated in 2% agarose gels in 0.5X TBE. Gels were visualized with UV light and photographed as described above. Polymorphic RFLP markers that were used in the map and their primer sequences are shown in the supplemental data (see Table S1 at http://www.genetics.org/supplemental/).

Additional markers: Two length polymorphisms due to insertions or deletions (indels) were identified in a manner similar to that described for RFLPs, except that polymorphisms were validated using PAGE. We also used two MATa-specific markers, NCP1a (LENGELER et al. 2002) and STE20a (LENGELER et al. 2001); a MATα-specific marker, STE20a (LENGELER et al. 2001); the mating-type phenotype (see below); and four “dominant” PCR markers flanking the MAT locus: KL1 and KL7 amplify flanking sequences unique to B-3502; KL2 and KL8 amplify flanking sequences unique to B-3501 (LENGELER et al. 2002). For markers and primer sequences see Table S1 at http://www.genetics.org/supplemental/.

Genotyping: Mating type: The mating type was identified by two methods. First, isolates were tested for mating compatibility against MATa and MATα tester strains (JEC20 and JEC21, respectively), as described above. A second, 10-fold dilution of each progeny DNA was tested by PCR using MAT-specific primers for STE20α and STE20α (see Table S1 at http://www.genetics.org/supplemental/; LENGELER et al. 2001).

Microsatellites and indels: Dilutions (100-fold) of progeny DNA from microsatellite and indel PCRs were electrophoresed in 6.7% polyacrylamide gels and silver stained as described above.

RFLPs: PCRs were performed as described above, using 1:99 dilutions of progeny DNA. PCR products (2 μl) were added to 12-μl restriction digestion reactions containing the manufacturer’s recommended units of restriction enzyme (Pro-mega, Madison, WI) and 1X restriction enzyme buffer and incubated at 37° for a minimum of 2 hr. A 5-μl sample of each restriction digest was electrophoresed, visualized, and photographed as described above.

Genotypes were scored visually and entered into a spreadsheet for segregation and linkage analysis.

Data analysis and map construction: Genetic mapping of all markers was performed using JoinMap 3.0 software (OOSTJEN and VOORRIPS 2001) on a Windows-based PC. Data in spreadsheet format were converted to text files, imported, and formatted according to the software instructions. This program assembles pairwise marker associations on the basis of likelihood of odds (LOD) scores of recombination frequencies and then constructs maps of these groupings utilizing a modified least-squares method (STAM 1993), in which the squares of the LOD scores are used as weights (OOSTJEN and VOORRIPS 2001). Prior to weighted least-squares map estimation, JoinMap used the Kosambi mapping function (KOSAMBI 1944) to translate recombination frequency into map distance. Of the several population types that JoinMap accommodates, the most appropriate for this population was “hapl1,” which is defined by JoinMap as “a haploid population derived from the F1 of a cross between two fully homozygous diploid parents” (OOSTJEN and VOORRIPS 2001). Markers were treated as codominant with phase (parental genotypes) known. Segregation ratios of markers were tested for deviations from the expected Mendelian 1:1 ratio using the chi-square test (α = 0.05) and one degree of freedom. Linkage was concluded using a recombination frequency (φ) of 0.30 as the threshold “swept radius” (CARTER and FALCONER 1951; SILVER 1995), which represents the maximum recombinational “distance” for a mapping population of 94 progeny over which linkage between two loci can be demonstrated at a significance level of 95% (SILVER 1995). We initially required LOD scores of 6.0 or more to conclude linkage and then subsequently tested LODs of 5.0 and 4.0 to determine whether linkage groups would combine in a manner that was consistent with results from karyotype hybridizations (see below).

Electrophoretic karyotyping: Parental isolates B-3501 and B-3502 were grown to an OD600 of 0.5 in YEPD broth at 30°C on a rotary shaker at 250 rpm. Spheroplasts were prepared according to LENGELER et al. (2000), with modifications based on WICKES et al. (1994). Plugs containing ~5 × 106 cells (~1 μg DNA) were loaded into a 13 × 14-cm 1% agarose gel (Chromosomal-Grade; Bio-Rad) in 0.5X TBE. Chromosomes were separated in 0.5X TBE using contour-clamped homogeneous electric field (CHEF) on a CHEF DR-II apparatus (Bio-Rad) under the following conditions: block 1, 75- to 150-sec switch, 4 V/cm, 12°C for 30 hr; block 2, 150- to 300-sec switch, 4 V/cm, 12°C for 54 hr. Gels were stained in 0.5 μg/ml ethidium bromide for 0.5–1.0 hr and visualized and photographed as above. To separate isomorphic chromosomes, electrophoretic conditions were modified and a longer gel (13 × 21 cm) was used. To separate the B-3502 7/8 doublet, the gel was electrophoresed for 240 hr at 9°C, 4 V/cm, with a 200- to 240-sec switch time. Chromosome sizes of C. neoformans were estimated by comparison with commercially prepared chromosomes of Hansenula wingei and Saccharomyces cerevisiae (Bio-Rad).

Chromosomes were transferred from gels onto positively charged nylon membranes (Roche, Mannheim, Germany) using the denaturing downward transfer method described in SAMBROOK and RUSSELL (2001) with the following modifications: after acid depurination, gels were neutralized and denatured with gentle shaking on orbital platforms; they were treated twice for 15 min in 0.5 N NaOH, 1.5 M NaCl, followed by twice for 30 min in 1 M ammonium acetate, 20 mM NaOH; the latter buffer served as the denaturing transfer buffer.
TABLE 1

Microsatellite motifs searched for polymorphisms
between B-3501 and B-3502

<table>
<thead>
<tr>
<th>Motif</th>
<th>No. screened</th>
<th>No. polymorphic</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>AG</td>
<td>27</td>
<td>5</td>
</tr>
<tr>
<td>AT</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>CT</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>CG</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>GT</td>
<td>18</td>
<td>3</td>
</tr>
<tr>
<td>AAC</td>
<td>41</td>
<td>7</td>
</tr>
<tr>
<td>AAG</td>
<td>76</td>
<td>2</td>
</tr>
<tr>
<td>AAT</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>ACC</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ACG</td>
<td>29</td>
<td>5</td>
</tr>
<tr>
<td>AGC</td>
<td>92</td>
<td>7</td>
</tr>
<tr>
<td>AGT</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ATG</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ATT</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CGG</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>CGT</td>
<td>21</td>
<td>2</td>
</tr>
<tr>
<td>CTT</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>GCC</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>GGT</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>CACT</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CATC</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>426</td>
<td>63</td>
</tr>
</tbody>
</table>

Microsatellite loci were identified using a BLAST search of the SGTC sequence database.

*Length polymorphisms between B-3501 and B-3502, as determined by PAGE.

Transfers proceeded for 24–48 hr, using a total of 2 liters of transfer buffer per gel. Membranes were rinsed and neutralized in 2x SSC for 5 min before UV crosslinking in a Stratagene UV crosslinker (Stratagene, La Jolla, CA) and then air dried.

Physical mapping—assignment of linkage groups to chromosomes: Selected markers were labeled and used as hybridization probes of Southern blots of electrophoretically separated chromosomes. Probes were labeled with digoxigenin (DIG), and hybridizations were performed according to kit instructions supplied with the Roche DIG luminescent detection system. To assign linkage groups to separated chromosomes, the blots were probed with a minimum of two markers from opposite ends of each linkage group. In addition to markers from the linkage map, we also probed karyotypes with ribosomal RNA gene sequences (Kwon-Chung and Chang 1994) and 4 of the 12 ungrouped markers.

RESULTS

Identification of 301 polymorphic markers: Of the 426 microsatellite loci that were identified, 63 were polymorphic between strains B-3501 and B-3502 and were used for genotyping the progeny and linkage analysis (Table 1 and supplementary materials at http://www.genetics.org/supplemental/). On the basis of comparisons of the B-3501 and B-3502 sequences, we identified a total of 13,041 restriction endonuclease sites that were polymorphic between them (Table 2 and supplementary materials at http://www.genetics.org/supplemental/). We selected RFLP markers that would permit, if possible, all SGTC sequence contigs to be represented by at least 2 markers, with larger contigs (>100,000 bp) represented by 3 or more markers. After each round of linkage mapping, additional markers were chosen to extend the ends of linkage groups and/or to join smaller linkage groups. We utilized a total of 228 RFLP markers for this linkage map. In addition to microsatellite and RFLP markers, we used 2 indel markers (Indel1, Indel2), the MATa-specific markers NCPla and STE20a, the MATa-specific marker STE20a, the markers flanking MAT (KL1, KL2, KL7, and KL8), and the mating-type phenotype (MATa vs. MATa) for a total of 301 markers. Primer sequences for PCR amplification of all markers used in the linkage map are available as supplementary data (Table S1 at http://www.genetics.org/supplemental/).
The genetic linkage map is composed of 20 linkage groups: At a LOD threshold of 6.0 and a swept radius of 0.30, JoinMap 3.0 distributed 289 markers into 20 linkage groups (Lynch and Walsh 1998; Ooijen and Voorrips 2001), but 12 markers (Alu2, AT4, Bam6, Eco15, Hind31, Mbo2, Pst13, Pst17, Pst26, Xho6, Xho9, and Xho13) did not group to any of the linkage groups (Figure 1). The MAT-associated markers (NCPIa, STE20a, STE20q, KL1, KL2, KL7, and KL8) mapped with or near the mating-type phenotype (MAT) and in the pattern predicted by the published sequence (Lengeler et al. 2002). Segregation data for all markers are available as supplemental data (Table S2 at http://www.genetics.org/supplemental/).

The linkage groups ranged in size from 3.8 cM (LG15) to 157.5 cM (LG1). The genome size, estimated by summing all linkage groups, is ~1296 cM. This underestimates the actual size of the genome because gaps between linkage groups assigned to the same chromosome can be as large as 30 cM each. There are eight such gaps in the linkage maps of both B-3501 and B-3502, and therefore the maximum possible genome size is 240 cM larger, or 1536 cM.

Segregation analysis reveals little segregation distortion: Of the 301 markers used, 295 (98.7%) showed 1:1 Mendelian segregation ratios in the progeny mapping population (Figure 2). Six markers showed segregation distortion; 4 were significant at the 95% significance level, and 2 were significant at the 99.5% significance level (Table 3). The MAT markers (MAT, NCP1a, STE20a, STE20q, KL1, KL2, KL7, and KL8) showed segregation distortion that was significant at the 90% level, with slight skewing toward the B-3502 (MATa) parent. However, there was an overall even distribution of parental alleles (i.e., not skewed toward either parent). Among the 94 progeny, at all 301 loci, there were 13,829...
B-3501 alleles and 13,821 B-3502 alleles. Nonetheless, several of the linkage groups show a slight, though insignificant, bias toward one or the other parent (Figure 2): LGs 9 and 18 are slightly skewed toward parent B-3501, while LGs 3, 5, 13, 14, 16, and 19 are slightly skewed toward parent B-3502.

The *C. neoformans* genome comprises 14 chromosomes: The electrophoretic karyotypes of strains B-3501 and B-3502 are shown in Figure 3. The genome of strain B-3501 separated into 10 bands, and the genome of strain B-3502, into 11 bands. On the basis of relative intensities, multiple chromosomes appear to be comigrating in bands 1, 4, and 7 of B-3501 and in bands 2, 7, and 11 of B-3502. The differential intensities of the larger bands are more apparent in the original images and in other chromosomal blots. Assuming that the chromosomes are homologous, bands 2, 4–6, and bands 8 and 9 of B-3502 should correspond to a doublet at band 1, a triplet at band 4, and a doublet at band 7, respectively, of B-3501. Similarly, bands 1, 5, and 6 of B-3501 should correspond to a doublet of chromosomes at bands 2 and 7 of B-3502, respectively, and bands 9 and 10 of B-3501 should correspond to a doublet at band 11 of B-3502. On the basis of this interpretation, we conclude that the *C. neoformans* genome comprises 14 chromosomes, which are numbered from largest to smallest in Figure 3. The B-3502 8/9 doublet was separated by altering the electrophoretic conditions and using a longer (21 vs. 13 cm) gel (Figure 4). The 1/2 and 10/11 doublets and 5/6/7 triplet of B-3501 and the 2/3 and 13/14 doublets of B-3502 were not separated in the longer gel.

We estimated the physical sizes (in base pairs) of chromosomes using *S. cerevisiae* and *H. wingei* chromosome size standards and generated standard curves using point-to-point semilog transformations to linearize the data. On the basis of these standard curves, the *C. neoformans* chromosomes ranged in size from 766 kb (chromosome 14 of both strains) to 2.37 Mb (chromosomes 1/2 of B-3501), from which we estimated total genome sizes of 20.2 Mb for B-3501 and 20.14 Mb for B-3502. These sizes are ~1 Mb larger than those proposed by SGTC and TIGR, on the basis of sequence analysis that does not include the ribosomal DNA repeat sequences.
in the size estimation. Differences in the electrophoretic mobility of *C. neoformans* chromosomes with GC content different from the size standards likely account for this ~5% difference in genome size estimates by the two approaches.

**All linkage groups are physically mapped to chromosomes:** Using at least 2 markers per linkage group as hybridization probes (indicated by asterisks in Figure 1), we assigned all 20 linkage groups to unique bands on CHEF blots (Figures 1, 4, and 5). Vertical alignment of linkage groups in Figure 1 (e.g., LG12 and LG13) indicated that marker probes from these linkage groups hybridized to the same chromosome, although the true orientation of linkage groups within a chromosome is not known from these data. Assignment of linkage groups to chromosomes is also depicted in Figures 4 and 5. With the exception of an irresolvable triplet in band 4 of B-3501, all linkage groups associated unambiguously with single chromosomes. Chromosomes 1, 2, 3, 4, 5, 6, 7, 11, 13, and 14 appear to be homologous between parent strains B-3501 and B-3502. Linkage group 2 hybridizes to chromosome 8 of B-3501, but to chromosome 9 of B-3502. None of the linkage groups hybridized to chromosome 10 of either B-3501 or B-3502. However, Alu2 and Pst13, which were among the 12 ungrouped markers, hybridized to chromosome 10. The remaining 10 ungrouped markers have all been assigned by hybridization to chromosomes, as shown in Table 4. The ribosomal DNA (rDNA) gene sequence hybridized to band 1 of B-3501 and band 2 of B-3502, which is consistent with the interpretation that band 1 of B-3501 represents the comigration of chromosomes 1 (LGs 12 and 13) and 2 (LGs 8 and 19) and that band 2 of B-3502 contains chromosomes 2 (LGs 8 and 19) and 3 (LG 6). The difference in size and migration of chromosome 2 most likely reflects a larger number of rDNA repeats.

**Evidence for reciprocal translocations:** Four linkage groups did not hybridize to the same chromosomes of strains B-3501 and B-3502, and the pattern suggested reciprocal translocations. LG4 hybridized to chromosome 9 of B-3501 and chromosome 12 of B-3502, whereas LG18 and LG14 hybridized to chromosome 12 of B-3501 and chromosome 8 of B-3502 (Figures 4 and

**Figure 1.—Continued.**
Figure 2.—Segregation ratios for LGs 1–20. Boundaries for the 95% confidence interval around the null hypothesis of Mendelian segregation (1:1) are indicated by dotted lines at 0.399 and 0.601. Five markers showed significant \( (P < 0.05) \) segregation distortion: CTG12, GTT9, Hind23, and Xba19 belong to LG9, and AG4 belongs to LG6. Linkage groups showing segregation ratios that are skewed toward parental strain B-3501 (LGs 9 and 18) or parental strain B-3502 (LGs 3, 5, 13, 14, 16, and 19) are indicated by bold lines and are highlighted in color as shown in the legend.

It is also possible that these apparent translocations are attributable to inconsistent migration of these chromosomes because of the electrophoretic conditions and/or differential expansion of subtelomeric repeats in homologous chromosomes.

**DISCUSSION**

This *C. neoformans* var. *neoformans* (serotype D) genetic linkage map used 289 markers to identify 20 linkage groups. Our estimate of chromosome number \( (n = 14) \) based on electrophoretic karyotyping of the parental strains B-3501 and B-3502 is in general agreement with the chromosome number proposed by Forche et al. (2000). The genome size estimated from these karyotypes is \( \sim 20.2 \) Mb, slightly larger than, but still consistent with, estimates from other reports (Perfect et al. 1989; Wickes et al. 1994; Forche et al. 2000), as well as the sequence-based estimate of the SGTC of 18.7 Mb (http://valefor.stanford.edu/group/Cneoformans/overview.html). By hybridizing representative markers to Southern blots of electrophoretically separated chromosomes, all 20 linkage groups were assigned to chromosomes. However, none of the linkage groups and only ungrouped markers Alu2 and Pst13 hybridized to chromosome 10.

The linkage map covers a minimum genetic distance

<table>
<thead>
<tr>
<th>Marker loci with significant segregation distortion among mapping progeny</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locus</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>CTG12</td>
</tr>
<tr>
<td>GTT9</td>
</tr>
<tr>
<td>AG4</td>
</tr>
<tr>
<td>Hind23</td>
</tr>
<tr>
<td>Xba19</td>
</tr>
</tbody>
</table>
C. neoformans Genetic Linkage Map

Figure 3.—Electrophoretically separated chromosomes of C. neoformans parental strains B-3501 and B-3502 using CHEF conditions as described in MATERIALS AND METHODS. Band numbers and chromosome designations are discussed in the text. Sizes of the chromosomes were estimated using commercially prepared chromosomes from H. wingei and S. cerevisiae as size standards and run simultaneously with B-3501 and B-3502.

The thresholds we used for declaring linkage (viz., recombination ≤0.30 and LOD ≥ 6.0) are stringent, but were chosen to minimize the likelihood of creating false linkages (Ooijen and Voorrips 2001). With a mapping population of 94 progeny, a swept radius of 0.33 is the boundary whereby only recombination values ≤0.33 are considered significant at a confidence level of ≥95% (Shields 1982). To confirm the importance of...
Assignment of ungrouped markers to chromosomes by hybridization to CHEF blots

<table>
<thead>
<tr>
<th>Ungrouped marker</th>
<th>B-3501</th>
<th>B-3502</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT4</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Mbo2</td>
<td>1/2</td>
<td>1</td>
</tr>
<tr>
<td>Xho6</td>
<td>1/2</td>
<td>1</td>
</tr>
<tr>
<td>Eco15</td>
<td>1/2</td>
<td>1</td>
</tr>
<tr>
<td>Hind31</td>
<td>1/2</td>
<td>1</td>
</tr>
<tr>
<td>Pst17</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Alu2</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Pst13</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Pst26</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Xho13</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Bam6</td>
<td>1/2</td>
<td>2/3</td>
</tr>
<tr>
<td>Xho9</td>
<td>3</td>
<td>2/3</td>
</tr>
</tbody>
</table>

Figure 5.—Resolution of the B-3502 band 7 doublet into chromosomes 8 and 9. Hybridizations of linkage group markers are indicated, showing at least partial identity between chromosome pairs as follows: chromosome 8 from B-3501 and chromosome 9 from B-3502; chromosome 9 from B-3501 and chromosome 12 from B-3502; chromosome 12 from B-3501 and chromosome 8 from B-3502. Conditions for separating these chromosomes included the use of a longer gel (21 vs. 12 cm), a narrower range of switch times (200–240 sec), and a longer run time (240 hr). Hybridizations to chromosome 12, which electrophoresed off this gel, are based on CHEF blots in Figure 4.

The small percentage of unlinked markers (4%) would suggest that we have mapped, or at least marked, a significant portion of the genome. Nonetheless, an entire chromosome—chromosome 10—is not represented on this linkage map, although 2 of the ungrouped markers hybridized to it. Chromosome 10 is ~1.1 Mb in length or 6% of the genome. In contrast, chromosome 11, which is similar in size to chromosome 10, is represented by a linkage group composed of 27 polymorphic markers. We do not know for certain the reason for this result. With an average map density of 12 kb between markers, the failure to map chromosome 10 suggests that an additional 92 cM is not covered by the map. It is unlikely that our suite of markers would by chance miss an entire chromosome of this size. A comparison of B-3501 and JEC21 sequence data confirm the relative paucity of variability in chromosome 10 (B. Loftus, personal communication). Therefore, it is reasonable to conclude that chromosome 10 is highly or entirely monomorphic between the parental strains, which is logical given that B-3501 and B-3502 are siblings. The lack of variability between B-3501 and B-3502 at chromosome 10 can be explained by either of two hypotheses. The first is that both B-3501 and B-3502 inherited, either by selection or by chance, a nonrecombiant chromosome 10 from the same parent (NIH433 or NIH12). This is entirely within expectation, as any large sample of meiotic products will include nonrecombinant as well as recombinant chromosomes. Analysis of the haplotypes at three of the larger linkage groups (LG1, LG6, and LG11) of the 94 progeny revealed a similar and otherwise unremarkable distribution of recombinant and nonrecombinant haplotypes (results not shown). Haplotype analysis of progeny from a cross used for a linkage map of Gibberella zeae (Jurgenson et al. 2002) also revealed parental haplotypes of entire linkage groups. However, the presence of two polymorphisms on chromosome 10 (ungrouped markers Alu2 and Pst13) argues against this hypothesis. Therefore, the more likely hypothesis is that the parental strains, NIH433 and NIH12, were similarly monomorphic at chromosome 9. Considering the degree of variability apparent at all other chromosomes, as well as the disparate origins of NIH433 and NIH12 (Heitman et al. 1999), this unexpected finding warrants further study of chromosome 10.

Compared with most other fungal linkage maps, we observed a remarkably uniform distribution of parental alleles among the progeny, with only 5 loci (1.7%) show-
ing segregation ratios sufficiently distorted from the null hypothesis of 1:1 to be considered significant at the 95% confidence level (Table 3). By comparison, in mapping the A. bisporus genome, KERRIGAN et al. (1993) found that 21 of 64 polymorphic loci (32.8%) showed segregation distortion (P < 0.05), which they suggested may have been caused by differential viability or growth rates among the progeny. The maize pathogen, C. heterostrophus, has been genetically mapped by TZENG et al. (1992), who found segregation distortion among 21 of 132 polymorphic loci (15.9%) and suggested that several were distorted because they were linked to the ToxI locus. In creating a genetic map of the related C. sativus, ZHONG et al. (2002) found segregation distortion among 20 of 102 polymorphic loci (19.6%). Using 468 polymorphic loci, a genetic map of G. zeae was created by JURGENS-SON et al. (2002); they identified nine linkage groups and found highly distorted segregation ratios (significance values not reported) across the entirety of five linkage groups, which they attributed to the selection of only one class of recombinants. However, observations of segregation distortion may also be due to sampling error. Using our data as an example, at a significance level of 95%, under the null hypothesis of no segregation distortion among the 297 loci, significant deviations would be expected in 15 loci (5% of 297). Yet five of the six markers showing distortion belonged to the same linkage group (LG 9), which overall was slightly skewed toward B-3501 (Figure 3). Seven other linkage groups were slightly skewed in parental contribution (Figure 3), with six being slightly skewed toward B-3502. However, apart from the aforementioned 5 loci, the segregation ratios of the loci of these slightly skewed linkage groups are not significant at the 95% confidence level.

The variability in electrophoretic karyotypes of B-3501 and B-3502, as shown in Figures 3 and 4, is not unexpected. Karyotypic variability in C. neoformans has been observed by others (KOWN-CHUNG et al. 1992; WICKEs et al. 1994) and has been used for taxonomic and epidemiologic studies (BOEKHOUT et al. 1997). In addition, karyotypic variability has been documented in other fungi, including Ustilago maydis (KINSCHERF and LEONG 1988), S. cerevisiae (ONO and ISHINOARAO 1988), and Candida species (MAGEE and MAGEE 1987).

By successfully separating the 8/9 doublets from B-3501 and B-3502 and probing them with markers that had already hybridized to the unresolved doublets, we discovered that chromosome 8 from B-3502 is homologous with chromosome 9 from B-3501. In both strains, chromosomes 8 and 9 are so similar in size that their separation required unusually lengthy electrophoresis times (10 days).

In assigning linkage groups to karyotypes, we discovered hybridization patterns that suggest reciprocal translocations (Figure 4). Linkage group 4 hybridized to chromosome 9 of B-3501 and chromosome 12 of B-3502, whereas the linkage group pair, LG14/LG18, hybridized to chromosome 12 of B-3501 and chromosome 8 of B-3502 (Figures 4 and 5). However, it is unlikely that LG 4 represents the entire chromosome 9 in B-3501 or that LG14/LG18 represents the entire chromosome 8 in B-3502; chromosome 9 is ~250 kb larger than chromosome 12, which at the genome-wide ratio of 13.2 kb/cM would be 18.9 cM. The reciprocating segment of DNA probably includes sequences that are monomorphic between B-3501 and B-3502 and, therefore, not represented by the linkage map. However, it is also possible that chromosomes 9 and 12 of B-3501 are respectively homologous with chromosomes 12 and 8 of B-3502, but that the electrophoretic migrations of these chromosomes are not entirely stable or well correlated with their sizes because of the prolonged conditions of electrophoresis or size differences in telomeric repeats. In this putative translocation, the involved linkage groups show normal segregation ratios and haplotype distributions (results not shown), suggesting that during prophase, synthesis is unhindered between homologous sections of the otherwise nonhomologous chromosomes. Synaptic pairing between reciprocal translocation chromosomes typically results in two types of meiotic segregants: those with duplicated and deleted segments, which are usually nonviable (also known as “adjacent-1”), and those with the normal, translocated chromosomes (also known as “alternate”; SUZUKI et al. 1989). Among the basidiospores from the cross used for this map, ~50% were nonviable. Further studies will analyze karyotypes and physical maps of B-3501/B-3502 parental isolates NIH433 and NIH12, as well as a random sample of the mapping progeny. In addition, our collaborations with the genome sequencing and assembly efforts underway at SGTC and TIGR should contribute to a better understanding of these anomalous regions.

Although this is the first report of a possible translocation in C. neoformans, translocations have been identified in other fungi, including the pathogenic yeast Candida dubliniensis, in which repeat sequence clusters mediate recombination between nonhomologous chromosomes (JOLY et al. 2002). A reciprocal translocation was also identified in the maize pathogen, C. heterostrophus (TZENG et al. 1992). Chromosomal translocations, like other chromosomal rearrangements, have been postulated as postzygotic reproductive isolating mechanisms that may be important, if not essential, agents of speciation (WHITE 1978). Evidence to support this hypothesis has been convincingly demonstrated among the Saccharomyces “sensu stricto” yeasts, where the number of reciprocal translocations among sibling species correlates with spore viability from hybrid crosses (FISCHER et al. 2000; DELNERI et al. 2003). The authors concluded that these chromosomal translocations are not likely to drive, but probably do contribute to, speciation through reproductive isolation. A better understanding of chromo-
LITERATURE CITED


Lengeler, K. B., G. M. Cox and J. Heitman, 2001 Serotype AD strains of Cryptococcus neoformans are diploid and aneuploid and are heterozygous at the mating-type locus. Infect. Immun. 69: 115-122.


Ooijen, J. W. V., and R. E. Voorrips, 2001 JoinMap 3.0: Software for the
C. neoformans Genetic Linkage Map


