

Infrequent Genetic Exchange and Recombination in the Mitochondrial Genome of *Candida albicans*

JAMES B. ANDERSON,^{1*} CLAIRE WICKENS,¹ MUSTAFA KHAN,¹ LEAH E. COWEN,¹
NANCY FEDERSPIEL,² TED JONES,² AND LINDA M. KOHN¹

Department of Botany, University of Toronto, Mississauga, Ontario, Canada L5L 1C6,¹ and Stanford DNA Sequencing and Technology Center, Palo Alto, California 94304²

Received 25 July 2000/Accepted 2 November 2000

Previous analyses of diploid nuclear genotypes have concluded that recombination has occurred in populations of the yeast *Candida albicans*. To address the possibilities of clonality and recombination in an effectively haploid genome, we sequenced seven regions of mitochondrial DNA (mtDNA) in 45 strains of *C. albicans* from human immunodeficiency virus-positive patients in Toronto, Canada, and 3 standard reference isolates of *C. albicans*, CA, CAI4, and WO-1. Among a total of 2,553 nucleotides in the seven regions, 62 polymorphic nucleotide sites and seven indels defined nine distinct mtDNA haplotypes among the 48 strains. Five of these haplotypes occurred in more than one strain, indicating clonal proliferation of mtDNA. Phylogenetic analysis of mtDNA haplotypes resulted in one most-parsimonious tree. Most of the nucleotide sites undergoing parallel change in this tree were clustered in blocks that corresponded to sequenced regions. Because of the existence of these blocks, the apparent homoplasmy can be attributed to infrequent, past genetic exchange and recombination between individuals and cannot be attributed to parallel mutation. Among strains sharing the same mtDNA haplotypes, multilocus nuclear genotypes were more similar than expected from a random comparison of nuclear DNA genotypes, suggesting that clonal proliferation of the mitochondrial genome was accompanied by clonal proliferation of the nuclear genome.

Candida albicans, a widespread commensal and important opportunistic pathogen of humans, has no known sexual cycle (15). Although there is genetic evidence for recombination in populations of this fungus (7, 13), there is no indication that genetic exchange between individuals is occurring presently or has ever been frequent. In diploid fungi such as *C. albicans*, inference of sexual genetic exchange based on analysis of nuclear loci is complicated by the occurrence of mitotic recombination between the two component haploid genomes. Mitotic recombination results in loss of heterozygosity both through gene conversion, which affects short tracts of DNA, and through mitotic crossing over, which affects entire chromosome arms distal to the point of exchange (6). Such loss of heterozygosity in diploid fungi is well documented (6) and has even been observed in experimental populations of *C. albicans* over relatively short time spans of a few hundred cell generations (4). Starting with a small number of heterozygous ancestors, repeated events of mitotic recombination in different parts of the genome could easily produce the multitude of genotypes now seen in populations of *C. albicans* in the complete absence of sexual genetic exchange.

How might one or more such highly heterozygous ancestral genotypes of *C. albicans* have arisen? One possibility is that one or more heterozygous diploid strains from the sexual ancestral population founded the present *C. albicans* population and simultaneously lost their propensity to mate and/or to complete meiosis. That only two alleles have been observed at most nuclear loci assayed (5) is entirely consistent with a small

number of founders of present populations of this fungus. Although the putative ancestral population of *C. albicans* is not known, this hypothesis is not in conflict with the recent observations that genes very similar to functional mating types in *Saccharomyces cerevisiae* exist in *C. albicans* (9) and that strains made homo- or hemizygous for mating-type-like genes are capable of mating (10, 12). Retention of the ability to mate under specialized laboratory conditions, which include strong selection for the products of mating, does not imply that mating has necessarily occurred in “natural” populations of *C. albicans* in human hosts.

The purpose of this study was to test for the signature of genetic exchange and recombination in the mitochondrial genome of *C. albicans*. An advantage of the mitochondrial genome over the diploid nuclear genome is that it is effectively haploid, with no heterozygosity and hence no potential for loss of heterozygosity during asexual propagation. Recombination between mitochondrial DNA (mtDNA) loci can therefore be detected only after genetic exchange between individuals. In *S. cerevisiae* and other fungi, the mitochondrial genome is highly recombinogenic; genetic exchange and recombination in mtDNA have been observed both in artificial matings (6) and in natural populations (14).

In this study, we screened 48 isolates of *C. albicans* for sequence variation in seven regions of mtDNA, six of which were unique and one of which occurred within the inverted repeat. We found clear evidence for a small number of past events of genetic exchange and recombination against a pattern of predominantly clonal reproduction. The clonal nature of the lineages characterized by the mtDNA haplotypes was reinforced by the observation that strains harboring identical mtDNA haplotypes carried nuclear genotypes that were sig-

* Corresponding author. Mailing address: Department of Botany, 3359 Mississauga Rd. North, University of Toronto at Mississauga, Mississauga, Ontario, Canada L5L 1C6. Phone: (905) 828-5362. Fax: (905) 828-3792. E-mail: janderso@credit.erin.utoronto.ca.

TABLE 1. mtDNA regions, sizes, positions in complete *C. albicans* mtDNA sequence, and primers

Region	Length of consensus sequence (bp) ^a	Positions in complete mtDNA sequence	Primers flanking the sequenced region
1	373	4939–4567	GAGCGGTTGATCTAATACCACC ^b TCTTCTATCGTGCATATGCTCC ^b
2	440	1507–1068	ACGCCAATTATACGTCAAGG GAACGTGTAACAACACACTAATCG
3	352	35511–35160 (10449–10800) ^c	GGTGATATAGTAGGGATACCTACTGG TAGGTAACCTGGGAATGTCC
4	201	30934–30734	TTATGGTCTTAGCGATTGGTAT TATGACCCTATGAGGCTTTATT ^b
5	493	24219–23728	AGAAGAATGAGGCACCATTAGC CACACGATAAGGATAGACGTGG
6	325	21399–21079	ATGTGCATCATTGGATAGGAGG ACTATTGGTGCCTTACCTGG
7	369	16332–15991	GTGCTACTAATTGATGATAGAGGTGG CCATTACATTAATGCTCTAACGCG ^b

^a The consensus sequences for regions 5, 6, and 7 are larger than those for the corresponding regions in the complete mtDNA sequence (haplotype 1) due to gaps in haplotype 1 (see Fig. 1).

^b Perfect matches to the complete *C. albicans* mtDNA sequence would have been as follows (differences from the respective primers are underlined): region 1, GAGCGGTTGATCTAATACTACC and TCTTCTATAGTGATATGCTCC; region 4, TATGACTATATGAGGCTTTATT; and region 7, CAATTATTAATGCTCTAAAGCG.

^c Region 3 is located within the inverted repeat in *C. albicans* mtDNA. No heterogeneity between copies of the repeat was detected in any of the haplotypes under sequencing conditions that routinely detect such heterogeneity in nuclear DNA regions heterozygous at specific nucleotide sites in *C. albicans*. Numbers in parentheses are the coordinates of region 3 in the alternate copy of the repeat.

nificantly more similar to one another than were nuclear genotypes drawn randomly from the general population.

MATERIALS AND METHODS

Strains. There were two main considerations in selecting the present sample of *C. albicans* strains for analysis of mtDNA sequences. The first was that the sample be taken from what is consistent with one population of *C. albicans* strains with no geographic or temporal subdivision; admixture of highly unrelated strains would complicate interpretations of clonality versus recombination. The second was that the size of the sample be sufficiently limited so that all of the variations in multiple loci could be detected by a modest nucleotide sequencing effort. Included in the sample were WO-1, CA, and CAI4, commonly studied reference strains supplied by P. T. Magee and B. B. Magee (University of Minnesota), and 45 strains from the sample of 81 strains of *C. albicans* from superficial infections of human immunodeficiency virus-positive patients attending two different hospitals in Toronto, Ontario, Canada (5). The 45 strains studied here were selected from the 81 strains in a random draw.

mtDNA loci, sequencing, and analysis. Our goal was to locate multiple non-contiguous regions of mtDNA with nucleotide sequence variations, without regard to whether the regions to be sequenced were coding or noncoding. First, primers (Table 1) for regions 1, 4, 5, and 7 were designed from sequences available through the Candida Information website (<http://alces.med.umn.edu/Candida.html>), and primers for region 6 were designed on the basis of the GenBank accession no. AF080074 sequence. Subsequently, primers for regions 2 and 3 were designed from the complete nucleotide sequence of the *C. albicans* mitochondrial genome (strain SC5314) when it became available; this sequence is now designated GenBank accession no. AF285261. The coordinates of the seven regions relative to the complete mtDNA sequence are provided in Table 1. In four cases, the primers designed earlier and actually used in this study did not perfectly match the complete *C. albicans* mtDNA sequence; these cases are noted in Table 1 along with the primer sequences that would have matched perfectly.

The conditions used for PCR and sequencing were exactly as described by Cowen et al. (4). Sequencing of the second strand was done for at least three representatives of each type of variant sequence. Raw sequences were trimmed

to include only areas of high-quality sequence from all 48 strains and were aligned with Sequencher (version 3.1.1; Gene Codes Corporation). Parsimony analysis was done with PAUP version 4.0b3a. Distances between nuclear genotypes were calculated as described by Cowen et al. (5) as the number of allelic differences. Randomization tests of association between the three most common mtDNA types and nuclear DNA genotypes were done with random numbers, data sorting, and macros written in Excel 98.

RESULTS

Among a total of 2,553 nucleotides in seven mtDNA loci, 62 polymorphic nucleotide sites and seven indels (Fig. 1) defined nine distinct mtDNA haplotypes among the 48 strains of *C. albicans*. mtDNA haplotypes 1, 3, 4, and 5 included strains from both Toronto hospitals from which samples were obtained. In addition, reference isolates CAI4 and CA belonged to mtDNA haplotypes 1 and 5, respectively, and the complete mtDNA sequence from *C. albicans* strain SC5314 belonged to haplotype 1.

Phylogenetic analysis resulted in one most-parsimonious tree of mtDNA haplotypes, with a consistency index of 0.68 (Fig. 2a). This most-parsimonious tree contained homoplasy, evident as several parallel changes in character state. The cause of the apparent homoplasy was found in pairs of nucleotide sites for which all four possible combinations of nucleotide states were observed (Fig. 3). These pairs are incompatible in the sense of Hudson and Kaplan (8). With no recombination, at least one parallel mutation is required to generate the four genotypes starting from any two-locus genotype arbitrarily designated as ancestral; with recombination, the four genotypes can be generated without parallel mutation and each

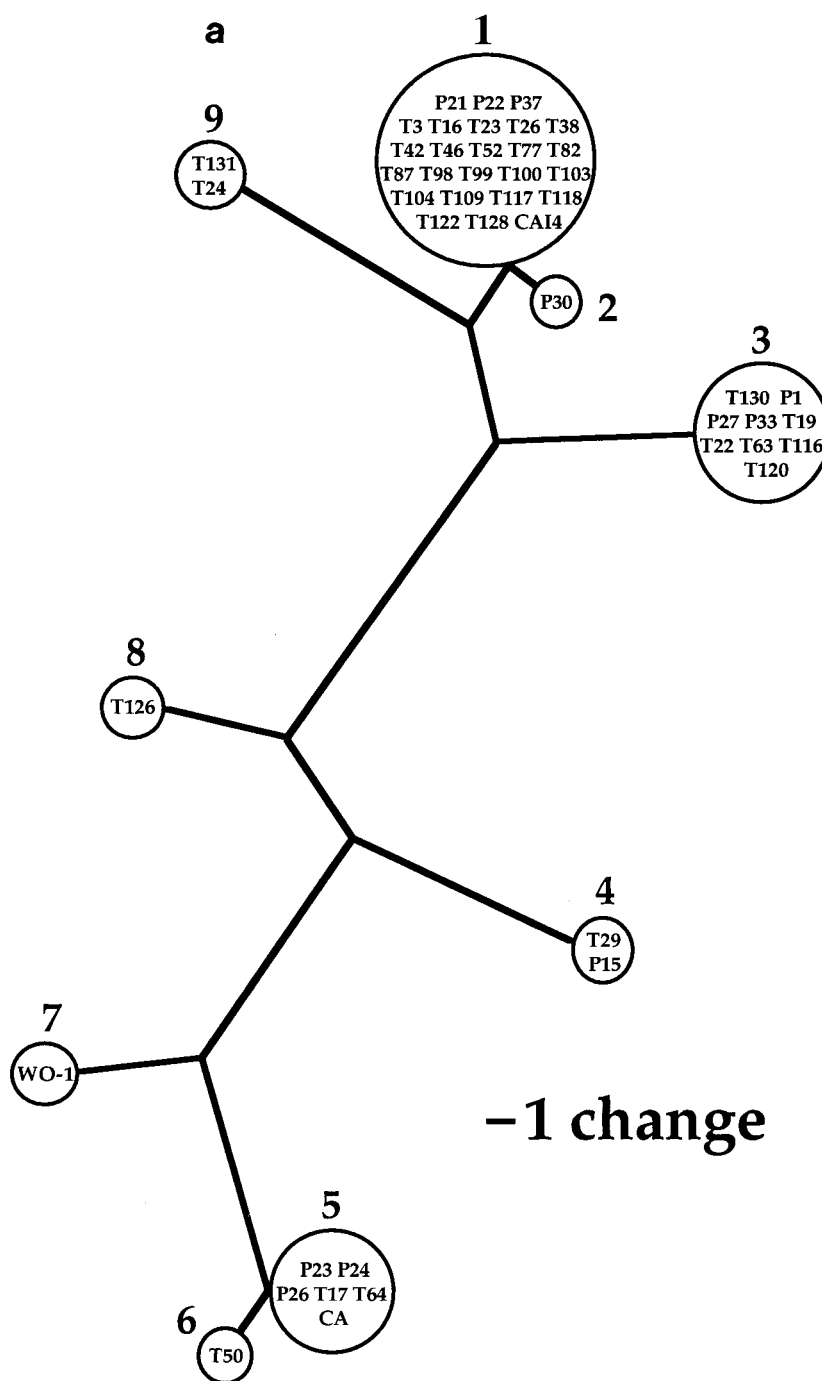


FIG. 2. (a) Single most-parsimonious tree of mtDNA haplotypes. Designations within circles represent strains listed by Cowen et al. (5). Numbers outside the circles represent mtDNA haplotype designations. Branch length is proportional to the number of character state changes (nucleotide substitutions), and scale is provided. (b) Same tree as that in panel a, except that reticulation is added as broken lines whose connections to branches are made with solid circles. The character numbers for the state changes (Fig. 1) associated with the reticulation in this tree appear on the broken lines. Although indels are not represented in this tree, only one (region 5, position 322 in the consensus sequence) showed reticulation, connecting haplotypes 4 and 7.

stands in sharp contrast to a previous phylogenetic analysis of nuclear DNA genotypes of *C. albicans* (5), in which the number of most-parsimonious trees was extremely large and all trees had very low internal consistency. The inability to find a plausible phylogeny of nuclear DNA genotypes in this presum-

ably asexual organism was attributed at least in part to parallel loss of heterozygosity through mitotic recombination, which could explain much of the apparent homoplasy in all of the phylogenetic trees of nuclear DNA genotypes.

Unlike the diploid nuclear genome of *C. albicans*, the mito-

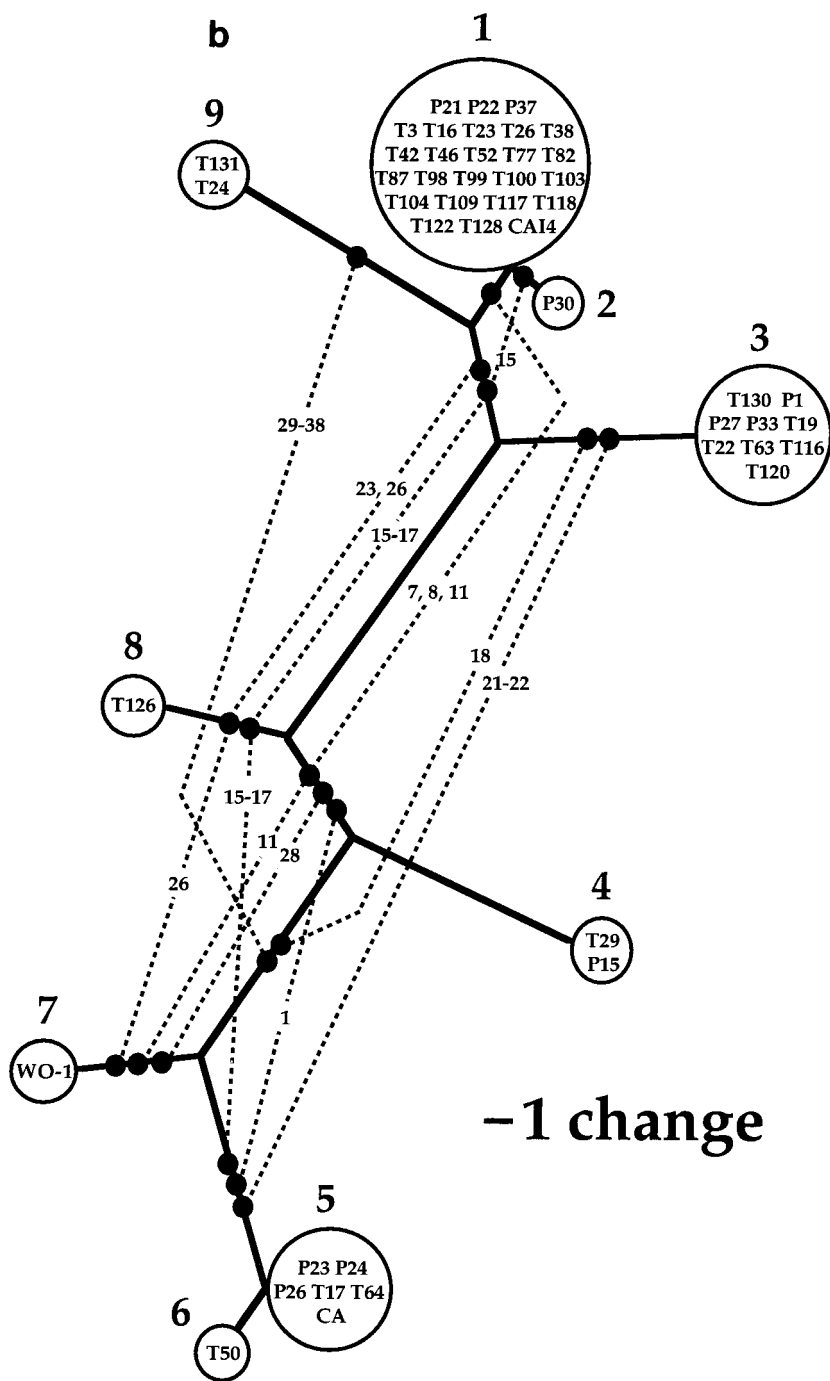


FIG. 2—Continued.

chondrial genome is effectively haploid; recombination between any two mtDNA loci must therefore be due to past genetic exchange between individuals. In many fungi, and especially in isogamous yeasts, the cytoplasmic mixing that occurs during mating is accompanied by the opportunity for mtDNA recombination (6). Wherever physical exchange occurs between two mtDNA molecules that are polymorphic and not identical by descent, recombinant genotypes will be detected after the sorting of mtDNA lineages that occurs during repeated cycles of cell budding and growth. Although physical

exchange might occur among the multiple mtDNA molecules within an individual, this process would not result in a recombinant genotype, since all of the molecules are identical by descent.

Although the limited number of haplotypes related in a single most-parsimonious tree is consistent with a largely clonal mode of reproduction, there were several examples of apparent homoplasy (Fig. 2b). There are two reasons why these examples can be the result only of genetic exchange and recombination (reticulation), rather than parallel mutation in

		Regions, character number					
		1	3	4	5	6	7
		0	001	111	1111122222	2333333333	444555555555666
		1	781	234	5678912368	9012345678	7890123456789012
1	01	-	--X	---	--X-----	-----	-----
	07		---	---	XXXX-XXXX-	XXXXXXXXXX	XXXXXXXXXXXXXXXXXX
3	08		--	---	XXXX-XXXX-	XXXXXXXXXX	XXXXXXXXXXXXXXXXXX
	09		-	---	XXXX-XXXX-	XXXXXXXXXX	XXXXXXXXXXXXXXXXXX
4	12			---	-----	-----	-----
	13			--	-----	-----	-----
	14			-	-----	-----	-----
	15				-----	-----	-----
	16				--X-XX-XX	XXXXXXXXXX	XXX-XXXXXXXXXXXX
	17				-X-XX-XX	XXXXXXXXXX	XXX-XXXXXXXXXXXX
	18				-----XX	XXXXXXXXXX	XXX-XXXXXXXXXXXX
5	21				----X	XXXXXXXXXX	XXX-XXXXXXXXXXXX
	22				----X	XXXXXXXXXX	XXX-XXXXXXXXXXXX
	23				---	XXXXXXXXXX	XXX-XXXXXXXXXXXX
	26				--	XXXXXXXXXX	XXX-XXXXXXXXXXXX
	28				-	XXXXXXXXXX	-----
	29					-----	XXXXXXXXXXXXXXXXXX
	30					-----	XXXXXXXXXXXXXXXXXX
	31					-----	XXXXXXXXXXXXXXXXXX
	32					-----	XXXXXXXXXXXXXXXXXX
6	33					-----	XXXXXXXXXXXXXXXXXX
	34					-----	XXXXXXXXXXXXXXXXXX
	35					-----	XXXXXXXXXXXXXXXXXX
	36					---	XXXXXXXXXXXXXXXXXX
	37					--	XXXXXXXXXXXXXXXXXX
	38					-	XXXXXXXXXXXXXXXXXX
	47						-----
	48						-----
	49						-----
	50						-----
	51						-----
	52						-----
7	53						-----
	54						-----
	55						-----
	56						-----
	57						-----
	58						-----
	59						---
	60						--
	61						-
	62						

FIG. 3. Compatibility matrix of nucleotide substitutions that were phylogenetically informative among the nine mtDNA haplotypes of *C. albicans*. X, incompatible (i.e., all four possible combinations of 2 nucleotides at the two variable positions are present in Fig. 1); -, compatible (i.e., fewer than all four possible combinations of 2 nucleotides at the two variable positions are present in Fig. 1). Region 2 was not included because none of the polymorphisms in this region was phylogenetically informative among the haplotypes. (The variation in region 2 only distinguishes haplotype 3 from all of the rest. Therefore, sites in region 2 cannot be incompatible with sites in other regions.)

the absence of recombination. First, there are only two nucleotides present for each of the 62 variable sites in the total of 2,553 nucleotides. Were these sites affected by repeated mutation, we might expect to have observed more than 2 nucleotides at at least some of the variable sites (3). It is not plausible that the 62 variable sites are the only ones of the 2,553 sites assayed that are available as potential targets for mutation. In the majority of instances, identity by nucleotide state can therefore be interpreted as identity by descent. The second reason why reticulation must be due to recombination rather than parallel mutation is that the majority of incompatibilities in Fig. 3 were among, rather than within, blocks of nucleotide sites grouped as regions in the mtDNA. To explain this pattern of incompatibility by mutation alone would require parallel changes in a highly nonrandom pattern. The likely signature of parallel mutation would instead be incompatibilities among

nucleotide sites that were randomly distributed, without the appearance of blocks of incompatibilities equivalent to those in Fig. 3.

Because of the positions of the reticulations in the most-parsimonious tree (Fig. 2b), it can be concluded that all of the recombination events occurred before the clonal proliferation of mtDNA haplotypes 1, 3, 4, 5, and 9. Exactly when these recombination events occurred, however, cannot be inferred from the present data because the times of the clonal proliferations are not known. Clonal proliferation of *C. albicans*, even between widely separated geographic locations (16), may be recent, possibly within the past few decades, due to the movement of humans.

Although several past events of mtDNA recombination are clearly evident, the data are not consistent with high-frequency recombination, either recently or in the distant past. Under frequent recombination, the distinct blocks of adjacent nucleotides within which no recombination was observed would become smaller and less distinct; the nonrandom pattern of incompatibilities would become progressively more random with time (2). In comparison, a pattern of more frequent mtDNA recombination prevails in the outcrossing basidiomycete fungus *Armillaria gallica*, in which no distinct blocks of sites and no linkage disequilibrium have been detected (14).

The entire 40.4-kb mtDNA represents less than 1% of the genome of *C. albicans*. When considering questions of clonality versus recombination, it is therefore important to determine if there is any association between mtDNA haplotypes and the rest of the genome. Since mtDNA haplotypes might be transmitted among individuals during mating independently of the nuclear genome, one possibility is that mtDNA haplotypes show no association with the nuclear genotype. By analysis of nuclear DNA genotypes accompanying each of the three mtDNA haplotypes with the largest numbers of strains, this possibility was rejected; mtDNA haplotypes and nuclear genotypes were strongly associated (Table 2). This relationship can be seen graphically in Fig. 4 in the unweighted pair-group method with arithmetic mean (UPGMA) tree representing similarity among the nuclear genotypes determined previously by Cowen et al. (5) and on which the mtDNA haplotypes are indicated. With only three exceptions, the mtDNA haplotypes corresponded to distinct clusters of nuclear genotypes. The exceptions, strains T22 and T19 of haplotype 3 and T104 of haplotype 1, may represent rare past events in which mitochondrial types reassociated with nuclear types. The predominant pattern, however, is that the clonal proliferation of mtDNA haplotypes was associated with what appears to be clonal proliferation of nuclear genotypes. The nuclear genotypic diversity within each of these clones is consistent with mitotic recombination and its attendant loss of heterozygosity, although genetic exchange between strains is not ruled out.

In this study, there were several examples of recombination in mtDNA that occurred before clonal proliferation of the mtDNA haplotypes of *C. albicans*. Because of the haploid nature of mtDNA, these recombination events can be attributed only to genetic exchange between individuals and not to recombination during asexual propagation. The mechanism of this genetic exchange between individuals is not known. Regardless of whether mtDNA exchange occurred through cytoplasmic mixing associated with some unknown pathway of so-

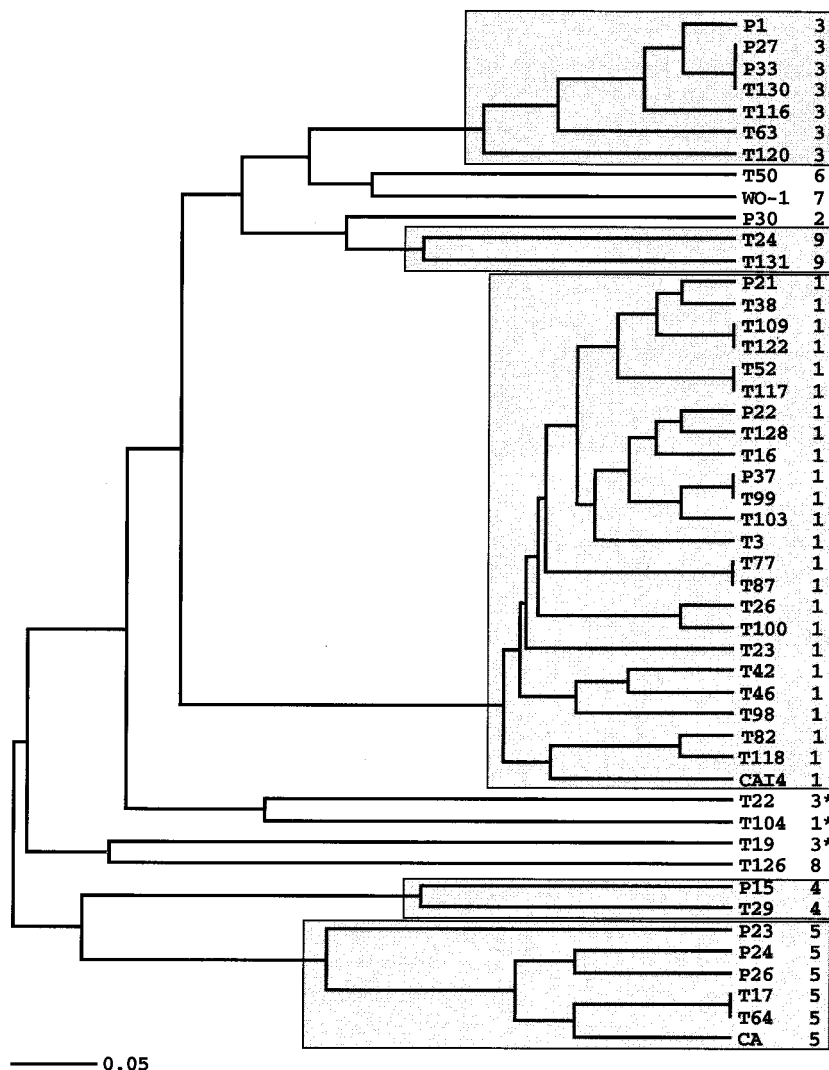


FIG. 4. UPGMA tree of similarity in nuclear DNA genotypes (5). The scale is the fraction dissimilar from zero to one. Boxes enclose distinct clusters of nuclear genotypes that also corresponded to mtDNA haplotypes. Asterisks designate strains of haplotypes 1 and 3 with exceptional nuclear genotypes.

matic fusion or with mating, the nuclear genome also would have been subject to recombination at the same time as the mtDNA genome. Although rare, these past events of genetic exchange may be of evolutionary significance in present-day populations. Even if *C. albicans*, as it now occurs in human hosts, has completely lost the abilities for mating and meiosis, populations of this fungus may still carry an adaptive benefit (11) from past episodes of recombination.

ACKNOWLEDGMENTS

We thank P. T. Magee and B. B. Magee for isolates CA, CAI4, and WO-1.

This work was supported by research grants from the Natural Sciences and Engineering Research Council of Canada to J.B.A. and to L.M.K. and by a grant-in-aid from Pfizer Canada, Inc.

REFERENCES

1. Anderson, J. B., and L. M. Kohn. 1998. Genotyping, gene genealogies, and genomics bring fungal population genetics above ground. *Trends Ecol. Evol.* **13**:444-449.

2. Awadalla, P., A. Eyre-Walker, and J. M. Smith. 1999. Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science* **286**:2524-2525.

3. Burt, A., D. A. Carter, G. L. Koenig, T. J. White, and J. W. Taylor. 1996. Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *Proc. Natl. Acad. Sci. USA* **93**:770-773.

4. Cowen, L. E., D. Sanglard, D. Calabrese, C. Sirjusingh, J. B. Anderson, and L. M. Kohn. 2000. Evolution of drug resistance in experimental populations of *Candida albicans*. *J. Bacteriol.* **182**:1515-1522.

5. Cowen, L. E., C. Sirjusingh, R. C. Summerbell, S. Walmsley, S. Richardson, L. M. Kohn, and J. B. Anderson. 1999. Multilocus genotypes and DNA fingerprints do not predict variation in azole resistance among clinical isolates of *Candida albicans*. *Antimicrob. Agents Chemother.* **43**:2930-2938.

6. Fincham, J.R.S., P. R. Day, and A. Radford. 1979. *Fungal genetics*, 4th ed. University of California Press, Berkeley.

7. Gräser, Y., M. Volovsek, J. Arrington, G. Schöniak, W. Presber, T. G. Mitchell, and R. Vilgalys. 1996. Molecular markers reveal that population structure of the human pathogen *Candida albicans* exhibits both clonality and recombination. *Proc. Natl. Acad. Sci. USA* **93**:12473-12477.

8. Hudson, R. R., and N. L. Kaplan. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**:147-164.

9. Hull, C. M., and A. D. Johnson. 1999. Identification of a mating type-like locus in the asexual pathogenic yeast *Candida albicans*. *Science* **285**:1271-1275.

10. Hull, C. M., R. M. Raisner, and A. D. Johnson. 2000. Evidence for mating of

- the "asexual" yeast *Candida albicans* in a mammalian host. *Science* **289**:307–310.
11. **Kondrashov, A. S.** 1993. Classification of hypotheses on the advantage of amphimixis. *J. Hered.* **84**:372–387.
 12. **Magee, B. B., and P. T. Magee.** 2000. Induction of mating in *Candida albicans* by construction of MTL α and MTL α strains. *Science* **289**:310–313.
 13. **Pujol, C., J. Reynes, F. Renaud, M. Raymond, M. Tibayrenc, F. J. Ayala, F. Janbon, M. Mallie, and J. M. Bastide.** 1993. The yeast *Candida albicans* has a clonal mode of reproduction in a population of infected human immunodeficiency virus-positive patients. *Proc. Natl. Acad. Sci. USA* **90**:9456–9459.
 14. **Saville, B. J., Y. Kohli, and J. B. Anderson.** 1998. mtDNA recombination in a natural population. *Proc. Natl. Acad. Sci. USA* **95**:1331–1335.
 15. **Scherer, S., and P. T. Magee.** 1990. Genetics of *Candida albicans*. *Microbiol. Rev.* **54**:226–241.
 16. **Xu, J. P., R. Vilgalys, and T. G. Mitchell.** 1999. Lack of genetic differentiation between two geographically diverse samples of *Candida albicans* isolated from patients infected with human immunodeficiency virus. *J. Bacteriol.* **181**:1369–1373.