

Use of Microsatellite Markers and Gene Dosage To Quantify Gene Copy Numbers in *Candida albicans*

J.-M. Costa,^{1,2} O. Eloy,³ F. Botterel,¹ G. Janbon,⁴ and S. Bretagne^{1*}

Laboratoire de Parasitologie-Mycologie, Hôpital Henri Mondor-APHP and UMR BIPAR 956, Créteil,¹ Laboratoire de Biologie Moléculaire, Hôpital Américain, Neuilly,² Laboratoire de Parasitologie, Hôpital Mignot, Versailles,³ and Unité de Mycologie Moléculaire, Institut Pasteur, Paris,⁴ France

Received 2 April 2004/Returned for modification 22 April 2004/Accepted 19 October 2004

With microsatellite marker typing, the number of alleles must be known for calculation of allelic frequencies in the diploid *Candida albicans* for a given locus. We describe a gene dosage with a double real-time PCR. Such a dosage should also be useful in exploring the loss of heterozygosity in *C. albicans*.

Genetic relationships among clinical isolates of *Candida albicans*, the most prevalent yeast in human pathology, can be studied using fingerprinting with DNA-based methods (3, 14). Among them, the microsatellite markers have the advantage of simplicity, reproducibility, and high throughput (2, 5, 6, 11, 13). Microsatellites are tandemly repetitive stretches of DNA in which a short motif, from 1 to 5 nucleotides long, is repeated several times. They are common, easy to identify from genomic databases, and extremely polymorphic in eucaryotes. Microsatellite alleles generally refer to DNA fragments of different sizes or electromorphs obtained after amplification with primers flanking the microsatellite region.

The microsatellite markers have the additional advantage of easily identifying the heterozygosity of diploid eucaryotes. Since *C. albicans* is diploid, when we observe electromorphs with one signal for a given locus, we infer that the isolates are homozygous for this locus (2, 4). However, distinction between a single allele and two homozygous alleles is difficult. A single allele could be due to the loss of one chromosome, deletion of the locus, or the lack of amplification of one of the two alleles. Knowing the number of copies of a given allele has consequences for the calculation of allelic frequencies in genetics of the *C. albicans* population. Moreover, a technique to quantify gene copy number may have some use in exploring the loss of heterozygosity observed in *C. albicans* (10).

Therefore, we sought to take advantage of quantitative real-time PCR to distinguish between one or two copies of the EF3 locus, for which we regularly observe electromorphs with one or two signals after amplification (2, 4). To validate our technique, we analyzed the Sor19 strain, monosomic for chromosome 5 and disomic for chromosome 1, and its original diploid CAF4-2 strain (8).

In a first step, two microsatellite markers known to be on chromosomes 5 and 1 were studied: the upstream sequence of the elongation factor 3 (EF3) gene, located on chromosome 5 (12), and the downstream sequence of the cell division cycle protein gene (CDC3), located on chromosome 1 (7). The primers, probes, PCR assays, and loci are subsequently referred to

as EF3 and CDC3. The PCR and analysis were performed as previously described (2). The microsatellite marker analysis confirmed that the Sor19 strain had a single band of 123 bp at the EF3 locus in contrast to its original CAF4-2 strain, which had two bands at 123 and 132 bp, while both were heterozygous for the CDC3 locus with the same alleles being 117 and 125 bp in length (Fig. 1).

The quantification was performed with a LightCycler instrument (Roche Molecular Biochemicals, Meylan, France) with SYBR Green I for online measurement of PCR products. The method involves amplification of the EF3 locus for which the copy number is to be determined; this is then compared to the reference gene locus CDC3, which has a known copy number. Two real-time PCR assays were developed for the EF3 (GenBank accession number Z11484) and CDC3 (GenBank accession number Z25869) genes. The primers were chosen close to but outside of the microsatellite to avoid differences in amplification efficiency due to the length of the microsatellite. For the EF3 gene dosage, the forward and reverse primers were 5'-CCGTAAAGCTGGTGACAAA-3' and 5'-CAGCTCTC AAGATGGCAACT-3', respectively. For the CDC3 gene dosage, the forward and reverse primers were 5'-CATTGGTGA AAGTGGATTGG-3' and 5'-CACCATCTTCTTCAATTCT GC-3', respectively. The PCRs were set up in a final volume of 20 μ l with the Fast DNA Master Hybridization SYBR Green Kit (Roche Biochemicals), with 0.5 μ M (each) primer, 2 mM magnesium chloride, and 5 μ l of extracted DNA. A first denaturation step of 8 min at 95°C according to the manufacturer's instructions was followed by an amplification performed for 40 cycles of denaturation (95°C, 10 s, ramping rate of 20°C/s), annealing (56°C, 10 s, ramping rate of 20°C/s), and extension (72°C, 15 s, ramping rate of 2°C/s). A single fluorescence reading for each sample was taken at the elongation step. Each amplification was performed at least three times in duplicate.

To generate a standard curve for each gene, we amplified serial dilutions of DNA of the CAF4-2 strain, diploid for both genes (Fig. 2). The second derivative maximum method automatically calculates the fractional cycle numbers, or crossing point, where the fluorescence rises above background. Crossing point values for a set of standards are expressed as a fractional cycle number that can be plotted against log con-

* Corresponding author. Mailing address: Laboratoire de Parasitologie-Mycologie, Hôpital H. Mondor, 51 avenue du Maréchal De-Lattre de Tassigny, 94010 Créteil, France. Phone: 33 1 49 81 36 41. Fax: 33 1 49 81 36 01. E-mail: bretagne@univ-paris12.fr.

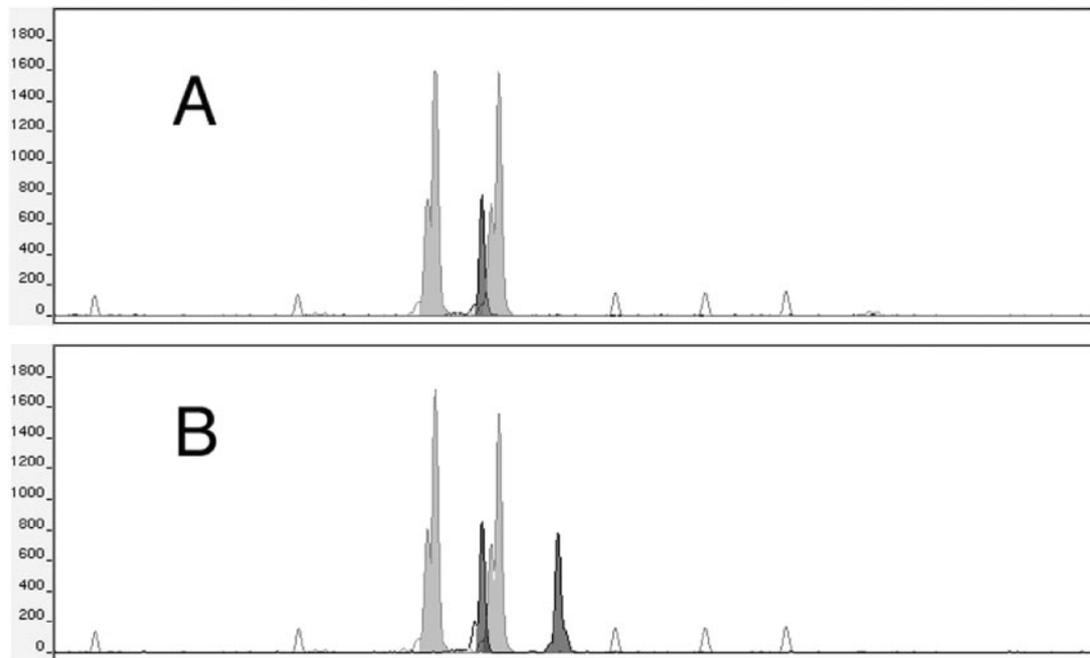


FIG. 1. Electromorphs of the EF3 (black peaks) and CDC3 (gray peaks) microsatellite marker analysis of the Sor19 strain (A) and its original diploid CAF4-2 strain (B). The Sor19 strain monosomic for chromosome 5 has a single EF3 peak at 123 bp in contrast with the diploid original strain, which has two EF3 peaks at 123 and 132 bp. The empty peaks represent the internal size standards (from left to right: 75, 100, 139, 150, and 160 bp). The y axis shows fluorescence intensities in arbitrary units.

centration to give standard curves. These standard curves were used to calculate the relative starting template concentration of the unknown samples for both genes. Therefore, the results were expressed as a ratio of the EF3 to the CDC3 gene quan-

tification, regardless of the precise amounts of input DNA. We obtained a crossing point value for the two loci. Then, we calculated the ratio $N = \text{starting copy number of EF3 gene} / \text{starting copy number of CDC3 gene}$. In this method, a two-

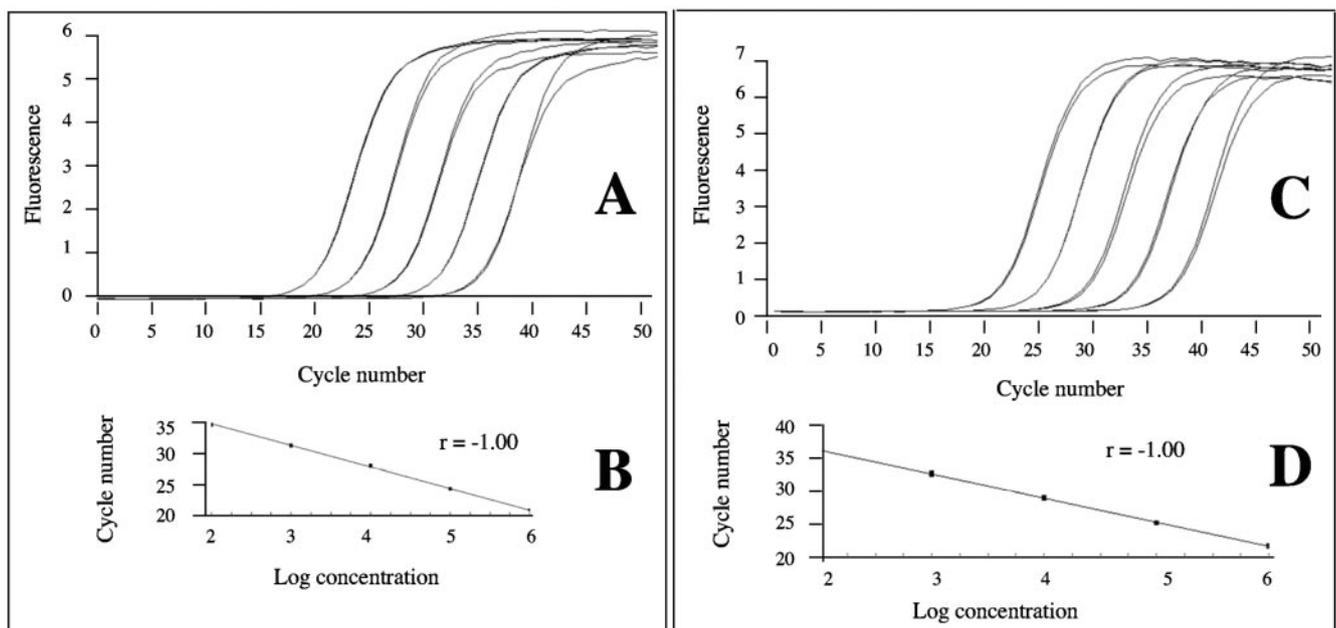


FIG. 2. Amplification plots and standard curves obtained for EF3 gene dosage (A and B) and for CDC3 gene dosage (C and D) with a LightCycler instrument (Roche Molecular Biochemicals) with the use of SYBR Green I for online measurement of PCR products. (A and C) Serial 10-fold dilutions in water of CAF4-2 *C. albicans* strain DNA, diploid for both genes; the amplification curves shift to the right as the input target quantity is reduced. (B and D) Standard curves obtained by plotting the cycle number against the input target quantity. For a given *C. albicans* isolate, these standard curves were used to calculate the ratio $N = \text{starting copy number of EF3 gene} / \text{starting copy number of CDC3 gene}$.

copy gene for both loci is expected to yield a ratio of $N = 1$, as opposed to $N = 0.5$ when the EF3 gene is single copy. With three independent assays, the Sor19 strain had $N = 0.44 \pm 0.04$, confirming its haploid feature for chromosome 5, whereas the original CAF4-2 strain had $N = 0.98 \pm 0.05$.

Secondarily, we investigated 10 clinical *C. albicans* isolates known to have a single signal for EF3 and a double signal for CDC3. We observed each time a ratio N close to 1, confirming that these clinical samples had two copies of the EF3 gene. Interestingly, we had the opportunity to study 20 colonies of a single stool sample, among which one had a single signal for EF3 and the others had two signals. For this specific colony, the ratio N was 0.99 ± 0.05 , confirming that there were two copies of the allele, as for the other 19 colonies. The more probable hypothesis for this finding is the loss of one chromosome followed by the duplication of the remaining one. This mechanism has been previously described for chromosome 5 (8). However, this would be the first description of this chromosome loss and duplication phenomenon in a patient.

Here we propose for the first time two real-time RCR assays to assess the gene copy number of the diploid yeast *C. albicans*, as recently proposed for human genetics (9, 16). We targeted the gene close to the microsatellite for which the analysis showed a single signal. Therefore, the method does not depend on the polymorphisms of the microsatellite itself. The method is also independent of differences in the yield of the amplification reactions or in the quantity of DNA used for amplification, as the copy number of the allele is calculated as a ratio (N), which erases these differences. When N is close to 0.5, the allele is single copy; when N is close to 1, the allele has a double copy. Until now, we have observed only two copies of the EF3 allele when a single signal was observed, confirming that it is correct for allelic frequency calculation to consider the isolate homozygous and count two alleles. The loss of one allele in a given clinical specimen, as we observed in testing 20 CFU from a stool specimen, warrants additional studies. In particular, this phenomenon can be chromosome dependent. This phenomenon is also probably dependent on the rate of multiplication of the yeast and the pressure of antifungal drugs. When performing serial testing, we usually observed the same genotype in a given patient (15). This chromosomal loss and duplication may not be selected over time. This needs an extensive follow-up of patients but should give some clues to a better understanding of the reproductive system of *C. albicans* (1).

The real-time fluorescent PCR gene dosage assays that we have developed are an accurate, nonradioactive, relatively cheap and fast method for testing gene copy number in *C.*

albicans. This method could be useful to elucidate some genetic evolution in *C. albicans* and to help accurate evaluation in epidemiological studies.

We are indebted to Catherine Costa for technical support.

REFERENCES

1. Balloux, F., L. Lehmann, and T. de Meus. 2003. The population genetics of clonal and partially clonal diploids. *Genetics* **164**:1635–1644.
2. Botterel, F., C. Desterke, C. Costa, and S. Bretagne. 2001. Analysis of microsatellite markers of *Candida albicans* used for rapid typing. *J. Clin. Microbiol.* **39**:4076–4081.
3. Bougnoux, M. E., A. Tavanti, C. Bouchier, N. A. Gow, A. Magnier, A. D. Davidson, M. C. Maiden, C. D'Enfert, and F. C. Odds. 2003. Collaborative consensus for optimized multilocus sequence typing of *Candida albicans*. *J. Clin. Microbiol.* **41**:5265–5266.
4. Bretagne, S., J. M. Costa, C. Besmond, R. Carsique, and R. Calderone. 1997. Microsatellite polymorphism in the promoter sequence of the elongation factor 3 gene of *Candida albicans* as the basis for a typing system. *J. Clin. Microbiol.* **35**:1777–1780.
5. Dalle, F., L. Dumont, N. Franco, D. Mesmacque, D. Caillot, P. Bonnin, C. Moiroux, O. Vagner, B. Cuisenier, S. Lizard, and A. Bonnin. 2003. Genotyping of *Candida albicans* oral strains from healthy individuals by polymorphic microsatellite locus analysis. *J. Clin. Microbiol.* **41**:2203–2205.
6. Dalle, F., N. Franco, J. Lopez, O. Vagner, D. Caillot, P. Chavanet, B. Cuisenier, S. Aho, S. Lizard, and A. Bonnin. 2000. Comparative genotyping of *Candida albicans* bloodstream and nonbloodstream isolates at a polymorphic microsatellite locus. *J. Clin. Microbiol.* **38**:4554–4559.
7. DiDomenico, B. J., N. H. Brown, J. Lupisella, J. R. Greene, M. Yanko, and Y. Koltin. 1994. Homologs of the yeast neck filament associated genes: isolation and sequence analysis of *Candida albicans* CDC3 and CDC10. *Mol. Gen. Genet.* **242**:689–698.
8. Janbon, G., F. Sherman, and E. Rustchenko. 1998. Monosomy of a specific chromosome determines L-sorbose utilization: a novel regulatory mechanism in *Candida albicans*. *Proc. Natl. Acad. Sci. USA* **95**:5150–5155.
9. Laurendeau, I., M. Bahuau, N. Vodovar, C. Larramendy, M. Olivi, I. Bieche, M. Vidaud, and D. Vidaud. 1999. TaqMan PCR-based gene dosage assay for predictive testing in individuals from a cancer family with INK4 locus haploinsufficiency. *Clin. Chem.* **45**:982–986.
10. Legrand, M., P. Lephart, A. Forche, F. M. Mueller, T. Walsh, P. T. Magee, and B. B. Magee. 2004. Homozygosity at the MTL locus in clinical strains of *Candida albicans*: karyotypic rearrangements and tetraploid formation. *Mol. Microbiol.* **52**:1451–1462.
11. Lott, T. J., R. E. Fundyga, M. E. Brandt, L. H. Harrison, A. N. Sofair, R. A. Hajjeh, and D. W. Warnock. 2003. Stability of allelic frequencies and distributions of *Candida albicans* microsatellite loci from U.S. population-based surveillance isolates. *J. Clin. Microbiol.* **41**:1316–1321.
12. Myers, K. K., W. A. Fonzi, and P. S. Sypheer. 1992. Isolation and sequence analysis of the gene for translation elongation factor 3 from *Candida albicans*. *Nucleic Acids Res.* **20**:1705–1710.
13. Sampaio, P., L. Gusmao, C. Alves, C. Pina-Vaz, A. Amorim, and C. Pais. 2003. Highly polymorphic microsatellite for identification of *Candida albicans* strains. *J. Clin. Microbiol.* **41**:552–557.
14. Soll, D. R. 2000. The ins and outs of DNA fingerprinting the infectious fungi. *Clin. Microbiol. Rev.* **13**:332–370.
15. Stephan, F., M. S. Bah, C. Desterke, S. Rezaiguia-Delclaux, F. Foulet, P. Duvaldestin, and S. Bretagne. 2002. Molecular diversity and routes of colonization of *Candida albicans* in a surgical intensive care unit, as studied using microsatellite markers. *Clin. Infect. Dis.* **35**:1477–1483.
16. Wilke, K., B. Duman, and J. Horst. 2000. Diagnosis of haploidy and triploidy based on measurement of gene copy number by real-time PCR. *Hum. Mutat.* **16**:431–436.