

Microsatellite Analysis of Three Phylogenetic Species of *Paracoccidioides brasiliensis*†

Daniel R. Matute,^{1‡} Victoria E. Sepulveda,^{1‡} Lina M. Quesada,² Gustavo H. Goldman,³
John W. Taylor,⁴ Angela Restrepo,^{1,5} and Juan G. McEwen^{1,6*}

Corporación para Investigaciones Biológicas (CIB), Medellín, Colombia¹; Laboratorio de Micología y Fitopatología Uniandes, Universidad de los Andes, Bogotá, Colombia²; Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil³; Department of Plant and Microbial Biology, 111 Koshland Hall, University of California, Berkeley, California 94720-3102⁴; Facultad de Medicina, Universidad Pontificia Bolivariana, Medellín, Colombia⁵; and Facultad de Medicina, Universidad de Antioquia, Medellín, Colombia⁶

Received 6 December 2005/Returned for modification 27 January 2006/Accepted 27 March 2006

***Paracoccidioides brasiliensis* is the etiological agent of paracoccidioidomycosis, an important human systemic mycosis in Latin America. Recently, the existence of three different phylogenetic species (S1, PS2, and PS3) of *P. brasiliensis* was demonstrated. Despite being genetically isolated, all three species were capable of inducing disease in both humans and animals, although lower virulence has been found with the PS2 species. The available molecular methods developed to characterize and type strains have not been useful for assigning isolates to the described species, creating the need for molecular markers capable of distinguishing genetically isolated groups. Here, we describe a PCR and sequencing-based microsatellite marker system that is stable, easy to assay, adaptable to large series of isolates, and discriminatory enough to be used as a typing system in identifying the three proposed species of *P. brasiliensis*. In addition, this system provides an unambiguous tool for strain discrimination between two (S1 and PS2) of the three phylogenetic species.**

Paracoccidioides brasiliensis is the etiological agent of paracoccidioidomycosis, an important human systemic mycosis in Latin America (21). Paracoccidioidomycosis is endemic to an area extending from Mexico to Argentina, and it is estimated that approximately 10 million people are infected with the fungus (2). The annual incidence rate in Brazil, the country with the highest endemicity, is 10 to 30 per million inhabitants, and the mean mortality rate is 1.4 per million per year (4). This fungus has shown extensive genetic variability as determined by molecular tools such as restriction fragment length polymorphism (RFLP) analysis (20) and electrophoretic karyotyping (17). Recently, the existence of three different phylogenetic species of *P. brasiliensis* has been demonstrated (16). They have been identified as S1 (species 1; 38 isolates), PS2 (phylogenetic species 2; 6 isolates), and PS3 (phylogenetic species 3; 21 isolates). Variations in the virulence and gene expression of antigenic proteins have been found between *P. brasiliensis* isolates now known to belong to species S1 and PS2 (11). Despite the differences reported among them, all three species are capable of inducing disease in both humans and armadillos (3), and consequently, the ability to cause disease is not a characteristic for species identification. In addition, none of the molecular methods developed to characterize and type strains, such as karyotyping and RFLP analysis, is useful for assigning strains to genetically isolated groups due to the low number of strains tested. Thus, there is a need for molecular markers capable of distinguishing among phylogenetic species. Once a

suitable classification system has been developed, well-grounded detection and differentiation protocols should be devised in order to promptly characterize new isolates and pursue specific biological questions (15).

Microsatellites, or simple sequence repeats, are tandem repeated segments of DNA composed of one to six base pair units. Length polymorphism in microsatellites arises from variations in the number of repeated units (7, 22). Indeed, these simple, tandem repeated oligonucleotides are highly polymorphic and amenable to precise, robust, and highly reproducible allelic characterization through the sizing or sequencing of PCR products (23), and it has been discovered that they can be used to identify several pathogenic fungi (1, 8, 9, 13). Recently, the existence of microsatellite markers from a collection of unique random sequence tags that are representative of the genome composition has been described for *P. brasiliensis* (19). Here, we describe the microsatellite patterns observed and the utility of some of them as molecular markers to discriminate among *P. brasiliensis* isolates in a search for correlations between the genetic backgrounds of the strains and the species boundaries previously proposed.

MATERIALS AND METHODS

Selection of isolates. The 64 *P. brasiliensis* isolates used in this study have been described previously (16) and are listed in Table S1 in the supplemental material. The entire sample encompassed isolates from the three phylogenetic species described for *P. brasiliensis*, namely, S1 ($n = 37$), PS2 ($n = 6$), and PS3 ($n = 21$), and represented six areas of paracoccidioidomycosis endemicity. The corresponding cultures were grown as reported previously (6). Total DNA was extracted from the yeast culture with protocols using glass beads (25) or maceration of frozen cells (18).

Microsatellite loci. OLIGO 4.0 (National Biosystems, Plymouth, Minn.) was used to design the oligonucleotide primers (Operon Technologies, Inc., Alameda, California) for selection of sequences of 10 nonrepetitive DNA regions

* Corresponding author. Mailing address: Carrera 72A 78B-141, Medellín, Colombia. Phone: (574)-441-08 55. Fax: (574)-441-55 14. E-mail: mcewen@epm.net.co.

† Supplemental material for this article may be found at <http://jcm.asm.org/>.

‡ D.R.M. and V.E.S. contributed equally to this work.

TABLE 1. Characteristics of the five polymorphic regions tested, including perfect microsatellites

Marker	Upper primer (5'-3')	Lower primer (5'-3')	Repeat sequence	PCR mean product size (bp)	GenBank accession no.	T_m (°C) ^a
PBMS1	CCACCCTGGCGAGGAAGACCCC	GCTTCCACGTTGGATCAGGTGC	T	530.09	DQ313963–DQ314026	59
PBMS2	TGCCCCAAGCAGCCCCCGGG	GAGAAAGTGAGTTGGTTTACG	TAAA	198.12	DQ314027–DQ314090	58
PBMS3	TGTCTAAGAGGACCGACCACG	GTGCCAGTGCATCATGAACAG	GAAA	388.94	DQ313835–DQ313898	58.5
PBMS4	TCGTGCAACTCTTGTGAAGG	TGAGGGAAGCACGCAGAACGC	AC	368.28	DQ313899–DQ313962	57.4
PBMS5	TTTGCTACACTTCCCTCTCCC	CTTCCCCATTCTGATTCTCG	AT	395.58	DQ314091–DQ314154	61

^a T_m , annealing temperature.

with embedded microsatellites. The primers and their respective annealing temperatures for PCR amplification are described in Table 1 (19).

The cyclic reaction termination method using a Big-Dye terminator cycle sequencing kit (Applied Biosystems, Foster City, California) was performed to define the nucleotide sequences. Sequence data were collected from both strands and examined with Sequence Navigator v. 1.0.1 (Applied Biosystems). The sequences were aligned manually, and the repeats were counted by visual inspection. Only perfect microsatellites (those composed only of identical repeats) were selected for the analysis.

Data analysis. A search for single microsatellites that could recognize all three phylogenetic species was performed, but no single microsatellite sufficed. A

graphic approach was used to determine which of the microsatellite pairs could differentiate among the phylogenetic species, using the number of repeats in each microsatellite. The number of repeats from all microsatellite pairs formed was treated as a coordinate, and the number of repeats of one of the microsatellites was plotted against the number of repeats of the other one in a Cartesian system.

Discriminatory power (D) was calculated using the Simpson index of diversity, utilizing the following formula:

$$D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^s X_j(X_j - 1)$$

TABLE 2. Allelic frequencies of the five markers for the phylogenetic species S1, PS3, and PS2 of *P. brasiliensis*

Marker	S1			PS3			PS2		
	Allele size (no. of repeats)	No. of isolates	Allelic frequency (%)	Allele size (no. of repeats)	No. of isolates	Allelic frequency (%)	Allele size (no. of repeats)	No. of isolates	Allelic frequency (%)
PBMS1	9	1	2.70	9	0	0	9	1	16.66
	10	6	16.21	10	21	100	10	3	50
	11	3	8.10	11	0	0	11	0	0
	12	18	48.64	12	0	0	12	2	33.33
	13	11	29.72	13	0	0	13	0	0
PBMS2	4	6	16.21	4	0	0	4	0	0
	5	1	2.70	5	0	0	5	6	100
	9	1	2.70	9	20	95.23	9	0	0
	10	18	48.64	10	1	4.76	10	0	0
	11	6	16.21	11	0	0	11	0	0
	14	1	2.70	14	0	0	14	0	0
PBMS3	7	0	0	7	1	4.76	7	3	50
	8	0	0	8	0	0	8	2	33.33
	9	15	40.54	9	0	0	9	1	16.66
	10	2	5.40	10	1	4.76	10	0	0
	11	4	10.81	11	2	9.52	11	0	0
	12	3	8.10	12	16	76.19	12	0	0
	13	12	32.43	13	0	0	13	0	0
	14	0	0	14	1	4.76	14	0	0
PBMS4	10	1	2.70	10	0	0	10	3	50
	11	0	0	11	0	0	11	3	50
	12	10	27.02	12	0	0	12	0	0
	13	1	2.70	13	0	0	13	0	0
	14	23	62.16	14	0	0	14	0	0
	15	2	5.40	15	1	4.76	15	0	0
	17	0	0	17	19	90.47	17	0	0
PBMS5	8	2	5.40	8	0	0	8	0	0
	11	2	5.40	11	1	4.76	11	0	0
	12	8	21.62	12	19	90.47	12	3	50
	13	11	29.72	13	1	4.76	13	2	33.33
	14	13	35.13	14	0	0	14	1	16.66
	18	1	2.70	18	0	0	18	0	0

where N is the number of strains tested, S is the number of different types, and X_j is the number of strains exhibiting the j th microsatellite genotype (14).

Statistical analysis. Each microsatellite was evaluated by means of one-way analysis of variance with SPSS v. 11 for Mac OS X (SPSS, Inc. Chicago, Ill.) to determine if there were significant differences in the distribution of its repeats among the species. A P value of less than 0.01 was considered significant for rejecting the null hypothesis of the lack of differences in the number of repeats among the species. The box plots associated with this analysis were done with Statistix 8.0 (Tallahassee, FL).

Nucleotide sequence accession numbers. The sequences obtained during this study have been deposited in the GenBank database under the accession numbers listed in Table 1.

RESULTS

Five of the nine microsatellites originally sequenced were perfect microsatellites (56%). All microsatellites suffer from length homoplasy; i.e., alleles identical by length may not be identical by descent. This homoplasy can be recognized in imperfect microsatellites, so they were discarded from further analysis (5, 10).

The stability of the markers in *P. brasiliensis* reference strains ATCC 60855 and ATCC 76533 was checked after 30 subcultures of the mycelial phase over a 3-year period. In all these experiments, repeated genotyping of an isolate always resulted in the same profile, indicating that the assay was reproducible and that the selected microsatellite loci were stable.

Each locus appeared polymorphic, with five, seven, nine, eight, and six alleles for microsatellites PBMS1, PBMS2, PBMS3, PBMS4, and PBMS5, respectively (Table 2).

The discriminatory powers of the individual microsatellite loci were as follows: for PBMS1, 0.68; for PBMS2, 0.81; for PBMS3, 0.81; for PBMS4, 0.76; and for PBMS5, 0.69. Moreover, the distributions of the repeats for each locus showed extensive overlap among the three phylogenetic species (see Fig. S1 in the supplemental material).

Overall, a total of 43 distinct multilocus genotypes could be identified using the five markers in combination. One genotype corresponding to PS3 was overrepresented; it was found in 12 of the isolates (18.8% of all the isolates, 57.2% of the PS3 isolates) (Table 3). The index of discrimination for the whole sample was 0.96 according to the Simpson index of diversity (14). The diversity index for each species was 0.98 for S1, 0.93 for PS2, and 0.69 for PS3. Of the 10 possible pairs of microsatellites, 3 pairs provided good resolution for differentiating among the three species, and all of these pairs involved PBMS4 in combination with either PBMS1, PBMS3, or PBMS5 (Fig. 1). Although other microsatellite pairs could be used to recognize one of the three species, they failed to recognize the other two (see Fig. S2 in the supplemental material). The description of the patterns obtained from these analyses for the recognition of species is described in Table S1 in the supplemental material.

DISCUSSION

The main objective of this study was to develop a microsatellite marker system to recognize the three phylogenetic species of the pathogenic, dimorphic fungus *P. brasiliensis*. This typing system could have clinical relevance, because variations in virulence and gene expression have been de-

TABLE 3. Characteristics and frequency of the 43 multilocus genotypes

Genotype ^a	No. of isolates ^b	% of total isolates	Allele size (no. of repeats) for:				
			PBMS1	PBMS2	PBMS3	PBMS4	PBMS5
1	1	1.56	9	5	8	11	13
2	1	1.56	12	5	8	11	13
3	1	1.56	12	5	7	10	12
4	2	3.12	10	5	7	10	12
5	1	1.56	10	5	9	11	14
6	1	1.56	13	11	9	14	13
7	1	1.56	10	14	12	14	12
8	2	3.12	10	15	11	14	12
9	1	1.56	10	15	11	14	13
10	1	1.56	9	15	11	14	12
11	1	1.56	13	9	9	14	13
12	2	3.12	13	10	9	12	13
13	4	6.25	13	10	9	14	13
14	1	1.56	12	10	9	14	13
15	2	3.12	13	10	9	14	14
16	1	1.56	13	11	13	12	14
17	1	1.56	12	11	13	12	12
18	1	1.56	12	11	13	12	14
19	1	1.56	12	11	9	12	14
20	1	1.56	12	10	9	12	13
21	1	1.56	12	11	9	12	11
22	1	1.56	12	10	10	14	14
23	3	4.68	12	10	13	14	14
24	1	1.56	12	10	12	14	12
25	1	1.56	11	10	9	12	8
26	1	1.56	11	10	10	12	18
27	1	1.56	11	10	13	14	8
28	2	3.12	12	4	13	15	14
29	1	1.56	12	4	13	14	14
30	1	1.56	12	4	13	14	11
31	1	1.56	12	4	15	14	14
32	1	1.56	10	4	13	13	12
33	1	1.56	10	5	12	10	12
34	1	1.56	10	9	10	17	12
35	1	1.56	10	9	7	17	12
36	12	18.75	10	9	12	17	12
37	1	1.56	10	9	12	17	13
38	1	1.56	10	10	12	17	12
39	1	1.56	10	9	14	17	12
40	1	1.56	10	9	12	17	11
41	1	1.56	10	9	11	15	12
42	1	1.56	10	9	12	18	12
43	1	1.56	10	9	11	17	12

^a The numbers assigned to the multilocus genotypes are arbitrary.

^b The total number of isolates, representing all the isolates from the three species, was 64.

tected among *P. brasiliensis* isolates belonging to species S1 and PS2 (3, 11).

The microsatellite marker system provides a highly effective method for the DNA fingerprinting of several organisms and has been successfully used for typing fungi, such as *Saccharomyces cerevisiae* (13), *Aspergillus fumigatus* (1), and *Candida* spp. (9), and also for assigning unknown strains of *Coccidioides* to the corresponding species and population (8). The method has proven to be discriminative, reproducible, and easy to perform. Furthermore, it has been demonstrated that the microsatellites are stable markers over several generations in many fungi, such as *Candida* (9) and *A. fumigatus* (1).

Although microsatellites have been previously reported for *P. brasiliensis*, this is the first report of their use for recognizing

7. **Ellegren, H.** 2004. Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* **5**:435–445.
8. **Fisher, M. C., B. Rannala, V. Chaturvedi, and J. W. Taylor.** 2002. Disease surveillance in recombining pathogens: multilocus genotypes identify sources of human *Coccidioides* infections. *Proc. Natl. Acad. Sci. USA* **99**:9067–9071.
9. **Foulet, F., N. Nicolas, O. Eloy, F. Botterel, J.-C. Gantier, J.-M. Costa, and S. Bretagne.** 2005. Microsatellite marker analysis as a typing system for *Candida glabrata*. *J. Clin. Microbiol.* **43**:4574–4579.
10. **Goldstein, D. B., and C. Schlötterer (ed.).** 1999. Microsatellites: evolution and applications. Oxford University Press, Oxford, United Kingdom.
11. **Hebeler-Barbosa, F., M. R. Montenegro, and E. Bagagli.** 2003. Virulence profiles of ten *Paracoccidioides brasiliensis* isolates obtained from armadillos (*Dasypus novemcinctus*). *Med. Mycol.* **41**:89–96.
12. **Henderson, S. T., and T. D. Petes.** 1992. Instability of simple sequence DNA in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12**:2749–2757.
13. **Hennequin, C., A. Thierry, G. F. Richard, G. Lecointre, H. V. Nguyen, C. Gaillardin, and B. Dujon.** 2001. Microsatellite typing as a new tool for identification of *Saccharomyces cerevisiae* strains. *J. Clin. Microbiol.* **39**:551–559.
14. **Hunter, P. R.** 1990. Reproducibility and indices of discriminatory power of microbial typing methods. *J. Clin. Microbiol.* **28**:1903–1905.
15. **Louws, F., J. Rademaker, and F. de Bruijn.** 1999. The three Ds of PCR-based genomic analysis of phyto bacteria: diversity, detection, and disease diagnosis. *Annu. Rev. Phytopathol.* **37**:81–125.
16. **Matute, D. R., J. G. McEwen, R. Puccia, B. A. Montes, G. San-Blas, E. Bagagli, J. T. Rauscher, A. Restrepo, F. Morais, G. Nino-Vega, and J. W. Taylor.** 2006. Cryptic speciation and recombination in the fungus *Paracoccidioides brasiliensis* as revealed by gene genealogies. *Mol. Biol. Evol.* **23**:65–73.
17. **Montoya, A. E., A. L. Alvarez, M. N. Moreno, A. Restrepo, and J. G. McEwen.** 1999. Electrophoretic karyotype of environmental isolates of *Paracoccidioides brasiliensis*. *Med. Mycol.* **37**:219–222.
18. **Morais, F. V., T. F. Barros, M. K. Fukada, P. S. Cisalpino, and R. Puccia.** 2000. Polymorphism in the gene coding for the immunodominant antigen gp43 from the pathogenic fungus *Paracoccidioides brasiliensis*. *J. Clin. Microbiol.* **38**:3960–3966.
19. **Nascimento, E., R. Martinez, A. R. Lopes, L. A. de Souza Bernardes, C. P. Barco, M. H. S. Goldman, J. W. Taylor, J. G. McEwen, M. P. Nobrega, F. G. Nobrega, and G. H. Goldman.** 2004. Detection and selection of microsatellites in the genome of *Paracoccidioides brasiliensis* as molecular markers for clinical and epidemiological studies. *J. Clin. Microbiol.* **42**:5007–5014.
20. **Nino-Vega, G. A., A. M. Calcagno, G. San-Blas, F. San-Blas, G. W. Gooday, and N. A. Gow.** 2000. RFLP analysis reveals marked geographical isolation between strains of *Paracoccidioides brasiliensis*. *Med. Mycol.* **38**:437–441.
21. **Restrepo-Moreno, A.** 2003. Paracoccidioidomycosis, p. 328–345. *In* W. E. Dismukes, P. G. Pappas, and J. D. Sobel (ed.), *Clinical mycology*. Oxford University Press, Oxford, United Kingdom.
22. **Schlötterer, C.** 1998. Genome evolution: are microsatellites really simple sequences? *Curr. Biol.* **8**:R132–R134.
23. **Squirrel, J., P. M. Hollingsworth, M. Woodhead, J. Russell, A. J. Lowe, M. Gibby, and W. Powell.** 2003. How much effort is required to isolate nuclear microsatellites from plants? *Mol. Ecol.* **12**:1339–1348.
24. **Taylor, J. W., and M. C. Fisher.** 2003. Fungal multilocus sequence typing—it's not just for bacteria. *Curr. Opin. Microbiol.* **6**:351–356.
25. **Van Burik, J. A., R. W. Schreckhise, T. C. White, R. A. Bowden, and D. Myerson.** 1998. Comparison of six extraction techniques for isolation of DNA from filamentous fungi. *Med. Mycol.* **36**:299–303.