Clinical Relevance of Multiple Single-Nucleotide Polymorphisms in *Pneumocystis jirovecii* Pneumonia: Development of a Multiplex PCR-Single-Base-Extension Methodology[▽]

F. Esteves,¹ J. Gaspar,² B. De Sousa,³ F. Antunes,⁴ K. Mansinho,⁵ and O. Matos^{1*}

Unidade de Protozoários Oportunistas/VIH e Outras Protozooses–CMDT, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisbon, Portugal¹; Departamento de Genética, Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisbon, Portugal²; Unidade de Epidemiologia e Bioestatística–CMDT, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisbon, Portugal³; Clínica das Doenças Infecciosas, Hospital de Santa Maria, Lisbon, Portugal⁴; and Serviço de Doenças Infecciosas, Hospital de Egas Moniz, Lisbon, Portugal⁵

Received 15 November 2010/Returned for modification 30 December 2010/Accepted 25 February 2011

Pneumocystis jirovecii pneumonia (PcP) is a major cause of respiratory illness in patients with AIDS. The identification of multiple single-nucleotide polymorphisms (SNPs) at three distinct *P. jirovecii* loci encoding dihydrofolate reductase (*DHFR*), mitochondrial large-subunit rRNA (*mtLSU rRNA*), and superoxide dismutase (*SOD*) was achieved using multiplex-PCR (MPCR) followed by direct sequencing and two single-base extension (SBE) techniques. Four SNPs (*DHFR312*, *mt85*, *SOD215*, and *SOD110*), correlated previously with parameters of disease, were amplified and genotyped simultaneously. The concordance of results between the standard sequencing technique (direct sequencing) and SBE analysis was 96.9% for the acrylamide gel electrophoresis and 98.4% for the capillary electrophoresis. The cross-genetic analysis established several statistical associations among the SNPs studied: *mt85C-SOD110T*, *SOD110T-SOD215C*, and *SOD110C-SOD215T*. These results were confirmed by cluster analysis. Data showed that among the isolates with low to moderate parasite burden, the highest percentages of *DHFR312C*, *mt85C*, *SOD110T*, and *SOD215C* were detected, whereas for high parasite burden cases the highest frequencies were observed among isolates with *DHFR312T*, *mt85T*, *SOD110C*, and *SOD215T*. The polymorphisms studied were shown to be suitable genetic targets potentially correlated with PcP clinical data that can be used as predictors of outcome in further studies to help clinical decision-making in the management of PcP. The MPCR/SBE protocol described for the first time in the present study was shown to be a rapid, highly accurate method for genotyping *P. jirovecii* SNPs encoded by different loci that could be used for epidemiological studies and as an additional procedure for the prognostic classification and diagnosis of PcP.

Pneumocystis jirovecii is an opportunistic pathogen capable of causing fatal pneumonia (PcP), which is frequently the first serious illness encountered by human immunodeficiency virus (HIV)-infected persons (1, 25, 29). One major problem in *P. jirovecii* research is the absence of an *in vitro* culture system. Molecular typing, particularly PCR-based approaches, have been used to describe differences between organisms (1, 24, 37). The advances in the understanding of the genetic diversity of *P. jirovecii* raised important questions, such as whether some organisms are more pathogenic than others and whether resistant or more virulent genotypes spread between hosts (5, 7, 8, 13–15, 24, 30).

A correlation between *P. jirovecii* subtypes and clinical parameters of PcP was described previously in the literature, revealing that the failure of anti-PcP prophylaxis may be associated with a specific *P. jirovecii* multilocus genotype (MLG). These data suggested that drug resistance could be detected by

molecular typing with the use of specific genetic markers (12, 14). Recently, a multilocus report on *P. jirovecii* isolates showed that specific single-nucleotide polymorphisms (SNPs) were associated with particular parameters of infection. This finding has led to the speculation that potential *P. jirovecii* haplotypes are related to clinical data and the outcome of PcP (7, 24).

The aims of the present study were to: (i) optimize a multiplex PCR (MPCR) protocol for the simultaneous amplification of the dihydrofolate reductase (*DHFR*), mitochondrial large-subunit rRNA (*mtLSU rRNA*), and superoxide dismutase (*SOD*) *P. jirovecii* genes associated previously with clinical parameters of infection; (ii) develop a single-base extension (SBE) technique to genotype four candidate SNPs (*DHFR312*, *mt85*, *SOD110*, and *SOD215*); and (iii) demonstrate that this new MPCR/SBE methodology is an accurate method to be applied in the characterization of multiple *P. jirovecii* SNPs located at different loci.

* Corresponding author. Mailing address: Unidade de Protozoários Oportunistas/VIH e Outras Protozooses–CMDT, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, 1349-008 Lisboa, Portugal. Phone: (351) 21 365 26 00. Fax: (351) 21 363 21 05. E-mail: omatos@ihmt.unl.pt.

[▽] Published ahead of print on 9 March 2011.

MATERIALS AND METHODS

Subjects and data. Thirty-two pulmonary specimens (eight induced sputa and 24 bronchoalveolar lavage fluids) tested previously and found to be positive for *P. jirovecii* (2000 to 2006) were included in the study. The pulmonary specimens were obtained from 32 randomly selected HIV-infected adult patients living in

TABLE 1. Nucleotide sequences and main characteristics of the SBE-TAG probes used for the SBE reaction of the SNPs at bases *DHFR312*, *mt85*, *SOD110*, and *SOD215*^a

SNP	SBE-TAG probe sequence (5'→3')	nt	%CG	T _m
<i>DHFR312</i>	GACATGTTGATTTCCTATCTTGCATGGTACGCTCC	35	46	54.8
<i>mt85</i>	GTTGTGCTGAATTAAGCGAATACCGTTGGCAAATTGTTTATTCCTCT	47	38	54.5
<i>SOD215</i>	GCAACGATGTGCGCGATTATCGCCTGGTTTCATTCTGAGTTAAAAAGATTTCATACGAGTTGCTAA	64	44	54.5
<i>SOD110</i>	GATGTTGCTGCGCTCGATGCAAAATACACGAAGGAGTTAGCTGATGCGAGGTAATACTTTTCT TTGTTTAAAG	75	40	53.5

^a Boldface sequences correspond to the SBE-TAGs, and the underlined sequences correspond to the annealing SBE primers (nt, nucleotides; %CG, cytosine-guanine percentage; T_m, melting temperature of the SBE primer).

Lisbon, Portugal, and were collected at two major hospitals for diagnostic purposes. The mean patient age was 40 years (range, 26 to 62 years), and most were men (78%).

All specimens had been subjected to DNA extraction by a Mini-BeadBeater/guanidinium thiocyanate-silica method (4). The detection of *P. jirovecii* organisms was performed by the amplification of the *mtLSU rRNA* gene by nested PCR (4, 23). Parasite burden was quantified by scoring the number of cysts observed by applying indirect immunofluorescence with monoclonal antibodies (MonoFluo kit *P. jirovecii*; Bio-Rad, France) and designated low (no cysts identified but positive by PCR), moderate (1 to 3 cysts in 30 fields at ×1,000), and high (4 to 30 or more cysts in 1 field at ×1,000) (7).

The 32 pulmonary specimens were grouped as described previously (8): 23 (72%) were obtained from patients with confirmed AIDS-related PCP, 3 (9%) from patients with atypical PCP presentation, and 6 (19%) from colonized patients. Data on anti-PCP prophylaxis were available for 30 patients: 6 were receiving trimethoprim sulfamethoxazole (TMP-SMZ) at a prophylactic dose, and 24 were not on a prophylaxis regimen. Data on anti-PCP therapy were available for 29 patients (23 were receiving TMP-SMZ at a therapeutic dose, and 6 were not receiving therapy).

DNA amplification by MPCR. An MPCR protocol was optimized to simultaneously amplify three genetic hot spots (*DHFR*, *mtLSU rRNA*, and *SOD*) that previously were correlated with parameters of disease (7, 22, 24). The selected polymorphic loci were analyzed in separate PCRs and then were simultaneously amplified using an MPCR reaction with the following steps: 95°C for 10 min, followed by 40 cycles at 95°C for 1 min, 61.5°C for 1 min, and 72°C for 1 min, and a final extension at 72°C for 10 min (T1 Thermocycler; Biometra) using 4 µl extracted DNA, 2.5 U AmpliTaq Gold DNA polymerase (Applied Biosystems), 1.5× reaction buffer (75 mM KCl, 15 mM Tris-HCl [pH 8.3]; Applied Biosystems), 0.5 mM deoxynucleoside triphosphates (dNTPs) (Applied Biosystems), 4 mM MgCl₂ (Applied Biosystems), 0.01 µg/µl bovine serum albumin (BSA) (Sigma), and 0.75 µl dimethylsulfoxide (DMSO) (Sigma) in a 50-µl reaction volume. The simultaneous amplification of *mtLSU rRNA*, *SOD*, and *DHFR* was achieved by using three pairs of primers (MWG Biotech), pAZ102-X and pAZ102-E (0.6 µM), MnSODFw and MnSODRw2 (1 µM), and FR 208 and FR 1018 (1.2 µM), respectively. Except for MnSODRw2 (5'-CACTTCCTTGAATCCAGACAA-3'), all primers were described previously (8, 21, 35, 37).

Direct sequencing. To confirm the presence of the three *P. jirovecii* amplicons and to analyze their sequences, portions of the MPCR products were separated by electrophoresis (1.5% agarose), purified individually using the Jetquick PCR purification spin kit (Genomed), and then subjected to direct sequencing from both strands as described previously (8).

SBE genotyping. SBE oligonucleotide sequences (SBE primers) were designed using Primer Express (version 3.0; Applied Biosystems). This software makes use of specific algorithms to calculate the melting temperature (T_m) based on oligonucleotide length and GC content, as well as the probability of hairpin and dimer formation. The online bioinformatics software BLAST (version 2.2), accessible at the National Centre for Biotechnology Information's website (www.ncbi.nlm.nih.gov), was applied for sequence alignments to achieve the desired complementarities between the designed oligonucleotides and the *P. jirovecii* sequences to guarantee the absence of homology within the designed sequences and to ensure no homology between the designed sequences and nontarget genomic regions of *P. jirovecii* or other organisms that could be present in the pulmonary specimens (eg., human cells, fungi, or bacteria). Four SBE primers (Table 1) were designed with melting temperatures around 55°C and to be located 1 nucleotide 5' upstream of the polymorphic spots (*DHFR312*, *mt85*, *SOD215*, and *SOD110*). Oligonucleotide TAGs with a length of 16 to 47 nucleotides were selectively paired within the SBE primers to generate hybrid SBE-TAG probes of various lengths, enabling further length discrimination (Table 1). These TAGs

were designed on the basis of the bacteriophage lambda genome. BLAST resources were used to guarantee the absence of complementarity with the *P. jirovecii* genome or with other organisms that could be detected in the pulmonary specimens. TAGs used for genotyping *mt85*, *SOD215*, and *SOD110* were described previously by Faber et al., while the *DHFR312* TAG was developed during the present study (10, 11, 16).

SBE preparation. To avoid carryovers of unincorporated dNTPs and MPCR primers into the SBE reaction, 10 µl of MPCR product was directly incubated with 2 U of shrimp alkaline phosphatase (SAP) (USB Corporation) and 4 U of exonuclease I (EXO) (USB Corporation) at 37°C for 1 h in a 20-µl reaction volume. The enzymes were inactivated at 96°C for 15 min (10, 17).

SBE analysis in acrylamide gel electrophoresis. Twenty µl of treated MPCR product was added to an optimized 40-µl reaction mix containing 2 U of ThermoSequenase (Amersham Biosciences); 1× reaction buffer (26 mM Tris-HCl [pH 9.5], 6.5 mM MgCl₂; Amersham Biosciences); 0.05 µM each dideoxynucleotide (ddNTP) (ddUTP-JOE, ddATP-FAM, ddGTP-ROX, and ddCTP-TAMRA; Jena Bioscience); and 4, 1, 6, and 5 µM *DHFR312*, *mt85*, *SOD110*, and *SOD215* SBE-TAG probes (MWG Biotech), respectively. The SBE reaction consisted of an initial denaturation step at 95°C for 3 min, followed by 55 cycles at 95°C for 30 s and 55°C for 30 s. SBE products (5 µl) were incubated with 0.5 U of SAP at 37°C for 1 h to eliminate excess ddNTPs, followed by enzyme inactivation at 96°C for 15 min. Treated SBE products were analyzed by electrophoretic separation (TBE buffer, pH 8.0) using a 10% denaturing polyacrylamide gel containing 0.5% urea in a gene scanner system 373 with Genescan 672 software (Applied Biosystems) as described previously (10, 32, 36).

SBE analysis in capillary electrophoresis. Reaction volumes of 15 µl were prepared with 5 µl of treated MPCR product; 4 µl reaction mix (SNPstart Master Mix, GenomeLab SNPstart primer extension kit; Beckman Coulter); and 5.6, 0.45, 2.7, and 2.25 µM *DHFR312*, *mt85*, *SOD110*, and *SOD215* SBE-TAG probes (MWG Biotech), respectively. The SBE reaction and the SBE product treatment were completed as described above for the SBE analysis by acrylamide gel electrophoresis. SBE products were analyzed by capillary electrophoretic separation using a CEQ 8000-XL (Beckman Coulter).

Data analysis. A descriptive statistical analysis was performed for qualitative data. Due to the failure of the requirements for the application of the Chi-square test, Fisher's exact test was applied to investigate associations between qualitative variables at a significance level of 0.05 using SPSS v.18.0 software (SPSS Inc., Chicago, IL). Linkage disequilibrium was studied using the online bioinformatics software SNPstats, which is accessible at the website of the Biostatistics and Bioinformatics Unit, Catalan Institute of Oncology (<http://bioinfo.iconcologia.net/index.php>). Additionally, a cluster hierarchical analysis of the multiple SNPs is shown in which complete linkage and Jaccard similarity measurements were used. Binary logistic regression with multiple independent variables was not possible in this study due to the limitation of the small sample size of 32.

RESULTS

DNA amplification by MPCR. In this work, an MPCR procedure was optimized for three *P. jirovecii* genetic targets (*DHFR*, *mtLSU rRNA*, and *SOD*). The primers were tested in separate PCRs for the 32 pulmonary specimens. The three loci then were amplified simultaneously. The validation of the MPCR protocol was achieved by the successful simultaneous amplification of the three genetic targets (*mtLSU rRNA*, 304 bp; *SOD*, 448 bp; and *DHFR*, 831 bp) in the 32 pulmonary

TABLE 2. *Pneumocystis jirovecii* SNPs detected at the three studied loci

Locus	SNP	Nucleotide (amino acid) position: identity	No. (%) of isolates
<i>DHFR</i>	<i>DHFR312T</i>	312 (117): T (Gly)	28 (88)
	<i>DHFR312C</i>	312 (117): C (Gly)	3 (9)
	<i>DHFR312C-DHFR312T^a</i>	312 (117): C/T (Gly)	1 (3)
<i>mtLSU rRNA</i>	<i>mt85C</i>	85: C	14 (44)
	<i>mt85A</i>	85: A	9 (28)
	<i>mt85T</i>	85: T	3 (9)
	<i>mt85C-mt85A^a</i>	85: C/A	3 (9)
	<i>mt85C-mt85T^a</i>	85: C/T	1 (3)
	<i>mt85A-mt85T^a</i>	85: A/T	1 (3)
	<i>mt85C-mt85T-mt85A^a</i>	85: C/T/A	1 (3)
<i>SOD</i>	<i>SOD110C</i>	110: C	16 (50)
	<i>SOD110T</i>	110: T	14 (44)
	<i>SOD110C-SOD110T^a</i>	110: C/T	2 (6)
	<i>SOD215C</i>	215 (41): C (Asp)	14 (44)
	<i>SOD215T</i>	215 (41): T (Asp)	11 (34)
	<i>SOD215C-SOD110T^a</i>	215 (41): C/T (Asp)	7 (22)

^a More than one nucleotide was found at a single base.

specimens studied. All MPCR products were subjected to direct sequencing, which confirmed the nature of the amplified fragments.

Genotyping. SBE primers were designed specifically to genotype the selected SNPs (*DHFR312*, *mt85*, *SOD110*, and *SOD215*). MPCR products were used as templates for SBE reactions. The SBE procedure was optimized with hybrid SBE-TAG probes that allowed length discrimination beyond the fluorescence identity of the incorporated dye-labeled ddNTPs. The visualization of the SBE products by acrylamide gel electrophoresis or by capillary electrophoresis showed the presence of four distinct fragments corresponding to the four SBE-TAG probes used in the experiments (Table 1). The 32 *P. jirovecii* isolates were tested in triplicate for each technique. In total, 128 SNPs were analyzed using direct sequencing (standard technique) and SBE. The concordance of results between the standard sequencing method and both SBE techniques was 96.9% for the acrylamide gel electrophoresis and 98.4% for the capillary electrophoresis. Of the 32 isolates, 29 (91%) demonstrated concordance of genotypes between direct sequencing and SBE analysis in acrylamide gel electrophoresis, and 3 (9%) showed a divergence of results: isolate Pj753 showed *mt85A* instead of *mt85C*, isolate Pj949 showed *SOD110C* instead of *SOD110T*, and isolate Pj1247 showed *DHFR312T-mt85C* instead of *DHFR312C-mt85A*. Thirty (94%) isolates demonstrated the concordance of genotypes between direct sequencing and SBE analysis in capillary electrophoresis, while 2 (6%) showed a divergence of results: isolate Pj753 showed *mt85A* instead of *mt85C*, and Pj1121 showed *mt85A* instead of *mt85T*. Table 2 summarizes the genotyping results obtained with the most sensitive method for the visualization of the SBE products (capillary electrophoresis).

Theoretically, multilocus genotyping at the four selected SNPs enables the discrimination of a total of 24 multilocus

TABLE 3. *P. jirovecii* MLGs identified

MLG	Nucleotide at each SNP base ^a				No. (%) of isolates sharing a genotype
	<i>DHFR312</i>	<i>mt85</i>	<i>SOD110</i>	<i>SOD215</i>	
Pj A	T	C	T	C	13 (23)
Pj B	T	A	C	T	10 (17)
Pj C	T	C	C	T	6 (11)
Pj D	T	T	C	T	5 (9)
Pj E	T	C	C	C	4 (7)
Pj F	T	A	C	C	4 (7)
Pj G	T	A	T	C	3 (5)
Pj H	T	T	C	C	2 (4)
Pj I	T	C	T	T	2 (4)
Pj J	T	T	T	C	2 (4)
Pj K	T	T	T	T	1 (2)
Pj L	C	C	C	T	1 (2)
Pj M	C	C	T	C	1 (2)
Pj N	C	A	T	C	1 (2)
Pj O	C	A	C	T	1 (2)

^a For further details, see Table 2.

genotypes (MLGs). In this study, 20 specimens produced four simple sequence patterns and thus corresponded to infections with single *P. jirovecii* MLGs. The other 12 specimens produced a complex sequence pattern made of two or three alleles for at least one of the four SNPs, revealing the presence of two or more putative coinfecting MLGs. Overall, 15 different *P. jirovecii* MLGs were observed. The frequency of each MLG ranged from 2 to 23% of all 32 isolates analyzed (Table 3).

Data analysis. Descriptive statistical analysis for qualitative data was performed. The isolates containing more than one nucleotide at a single base contributed two or more putative genotypes for the associated variations and cross-genetic statistical analysis.

The cross-genetic analysis showed the presence of some statistically significant associations among isolates with and without the SNPs in our study. The results pertaining to the presence of certain SNPs are the following: *mt85C* was more frequent among isolates with *SOD110T* (81%; $P = 0.025$) and less common in isolates with *SOD110C* (39%; $P = 0.025$) and *SOD215T* (39%; $P = 0.025$); *mt85A* was less common among isolates with *SOD110T* (25%; $P = 0.013$); *SOD110T* was more frequent among isolates with *SOD215C* (73%; $P < 0.001$); and *SOD110C* was more frequent among isolates with *SOD215T* (100%; $P < 0.001$). The results from SNPStats software demonstrated that the SNPs at bases 110 and 215 of the *SOD* gene were at linkage disequilibrium ($P < 0.001$). These results were confirmed through cluster analysis. In Fig. 1 two distinct groups can be seen clearly: *mt85C*, *SOD110T*, *SOD215C*, and *DHFR312C* and *mt85A/mt85T*, *SOD110C*, *SOD215T*, and *DHFR312T* (the most closely related SNPs within the two main clusters are emphasize in boldface).

The uniform distribution of the SNPs was observed during the experimental period, and no association was found between SNPs and the year or season in which the episodes of *P. jirovecii* infection occurred. No association was observed between SNPs and patient's clinical groups (AIDS-related PcP cases, atypical PcP episodes, or colonized patients), gender, anti-*P. jirovecii* prophylaxis, or therapy. However, several SNPs and MLGs were statistically associated (Fisher's exact test)

FIG. 1. Dendrogram constructed on the basis of a cluster hierarchical analysis, where complete linkage and Jaccard similarity measure were used, showing the relationship between the multiple *P. jirovecii* SNPs studied. SNPs were clustered in two major groups: (i) *mt85C*, *SOD110T*, *SOD215C*, and *DHFR312C* and (ii) *mt85A*/*mt85T*, *SOD110C*, *SOD215T*, and *DHFR312T*.

with specific age groups and parasite burden levels (Table 4). A lack of high parasite burden among isolates with *mt85C* was observed together with an excess of high parasite burden among isolates without *mt85C* ($P = 0.022$). For isolates with *mt85T*, an excess of high parasite burden was observed, violating the null hypothesis of no association among the variables *mt85T* and parasite burden ($P = 0.038$). Continuing this analysis, with respect to the patient's age the rejection of the null hypothesis was due mainly to a lower number of cases ob-

served for isolates with *SOD110C*, *SOD215T*, and *SOD110C-SOD215T* for patients in age group A, together with an excess and lack of isolates without those genotypes for patients in age group A and B, respectively ($P = 0.019$). Isolates presenting *SOD110T* and *SOD110T-SOD215C* were less frequent among patient age group B and more frequent among age group A, which is what would have been expected if the null hypothesis was not rejected. The results were reversed for isolates not presenting with *SOD110T* and *SOD110T-SOD215C* in age groups A and B ($P = 0.021$).

DISCUSSION

SNPs are the most common source of DNA sequence variation in *P. jirovecii*, contributing largely to its genetic variability. These variations may be important molecular shifts influencing the pathogenesis of the organism and the clinical outcome of the PcP cases (5, 7, 14, 15, 24, 25). Typing *P. jirovecii* SNPs should meet certain specific selection criteria, which include their frequency, location, and type of change (10, 19). In this study, the clinical relevance of four SNPs formerly associated with parameters of the disease was evaluated (7, 8). To our knowledge, the present report is the first molecular study of *P. jirovecii* in which the MPCR/SBE methodology was applied. The multiplex amplification of genomic DNA associated with SBE genotyping was shown to be a suitable high-throughput methodology for *P. jirovecii* SNP screening.

Three polymorphic loci (*mtLSU rRNA*, *SOD*, and *DHFR*) were simultaneously amplified by MPCR using specific primers, which have been tested in separate PCRs. The multiplex SBE was designed using four detection SBE-TAG probes that specifically annealed to the target nucleic acid sequences immediately adjacent to the variable nucleotide positions

TABLE 4. Fisher's exact test for patient age groups and parasite burden with respect to the four SNPs studied and to the most relevant MLGs

SNP and genotype	No. (%) in age group ^a :			<i>P</i> ^b	No. (%) with parasite burden:		<i>P</i> ^b
	A	B	C		Low to moderate	High	
<i>DHFR312T</i>	3 (10)	15 (52)	11 (38)	NS	20 (69)	9 (31)	NS
<i>DHFR312C</i>	0	2 (50)	2 (50)	NS	3 (75)	1 (25)	NS
<i>mt85C</i>	2 (11)	9 (50)	7 (39)	NS	16 (89)	2 (11)	0.022
<i>mt85A</i>	0	10 (66)	5 (33)	NS	11 (73)	4 (27)	NS
<i>mt85T</i>	1 (17)	3 (50)	2 (33)	NS	2 (34)	4 (66)	0.038
<i>SOD110C</i>	0	13 (72)	5 (28)	0.019	11 (61)	7 (39)	NS
<i>SOD110T</i>	3 (19)	5 (31)	8 (50)	0.021	13 (81)	3 (19)	NS
<i>SOD215C</i>	3 (14)	11 (50)	8 (36)	NS	17 (77)	5 (23)	NS
<i>SOD215T</i>	0	13 (72)	5 (28)	0.019	11 (61)	7 (39)	NS
<i>SOD110T-SOD215C</i>	3 (19)	5 (31)	8 (50)	0.021	13 (81)	2 (19)	NS
<i>SOD110C-SOD215T</i>	0	13 (72)	5 (28)	0.019	11 (61)	7 (39)	NS
MLG							
Pj A	2 (15)	4 (31)	7 (54)	NS	11 (85)	2 (15)	NS
Pj B	0	8 (80)	2 (20)	NS	6 (60)	4 (40)	NS
Pj C	0	5 (83)	1 (17)	NS	5 (83)	1 (17)	NS
Pj D	0	3 (60)	2 (40)	NS	2 (40)	3 (60)	NS

^a Age groups were defined as follows: A (patient age, <30 years), B (patient age, between 30 and 39 years), and C (patient age, ≥40 years).

^b Results are based on Fisher's exact test. NS, no statistically significant association ($P > 0.05$). Statistically significant associations ($P < 0.05$) are emphasized in boldface.

(*DHFR312*, *mt85*, *SOD110*, and *SOD215*). The 3' end of each SBE-TAG probe was extended with a dye-labeled ddNTP complementary to the nucleotides at the SNP sites. The identification of the different SBE products was achieved through length discrimination provided by the SBE-TAG probes, and the identity of the SNPs was revealed by the fluorescence of the incorporated ddNTP (10, 11, 36).

The differences observed between the direct sequencing method and the SBE techniques showed that in one isolate, the genotyping results of both SBE techniques were identical but were dissimilar from the results obtained by direct sequencing (for *Pj753*, *mt85A* was found instead of *mt85C*). These data suggest that SBE techniques, especially the SBE analysis in capillary electrophoresis, had increased sensitivity compared to that of direct sequencing.

TMP-SMZ was the key agent for the treatment and prophylaxis of PcP in the studied subjects. TMP inhibits DHFR, an enzyme of the folic acid pathway, in which specific genetic polymorphisms were shown to be potentially associated with resistance to TMP-SMZ (2, 3, 7, 8, 24, 30). DHFR converts dihydrofolate into tetrahydrofolate during the biosynthetic pathway, providing thymidylc acid, several amino acids, and a methyl group for the *de novo* synthesis of purines (33). New accurate high-throughput molecular techniques, such as SBE, may be helpful tools for the detection of *DHFR* polymorphisms potentially associated with resistance to TMP-SMZ. The allelic frequency distribution analysis revealed that *DHFR312T* was the most prevalent *DHFR* genotype, as has been found in previous studies, in which identical immunocompromised populations presented similar *DHFR* allelic distribution patterns (2, 3, 7, 8). No significant variation of the *P. jirovecii* *DHFR* distribution pattern was observed in the last 15 years in Portugal, although the present study showed a lower incidence of *DHFR312C* (12%) than that detected in those previous reports (range, 21 to 28%).

The *mtLSU rRNA* gene, which is involved in basic metabolic mechanisms during translation, providing peptidyl transferase activity to the mitochondrial ribosome, is considered to be a highly informative marker for discerning associations between specific genotypes and geographical or clinical data, and it has been used most widely for PCR-based studies of *P. jirovecii* (8, 24, 31, 35, 37). The *mtLSU rRNA* sequencing results were similar to those reported in recent studies on Portuguese subjects (7–9). Other reports from different geographical regions (e.g., the United States and United Kingdom) (1, 26) also have revealed an *mtLSU rRNA* pattern identical to that of the Portuguese samples. A different allelic distribution pattern was detected persistently in previous studies on Spanish subjects, whereas genotypes with *mt85T* were the second most frequent and genotypes with *mt85A* were the third most common (9, 28). These differences demonstrate a variation in the geographical distribution of *P. jirovecii* genotypes. Epidemiological factors inherent to different geographical locations may influence the circulation and transmission of genetically distinct *P. jirovecii* organisms (9, 13).

P. jirovecii is exposed to oxidative bursts from alveolar macrophages or neutrophils. Also, reactive oxygen intermediates can be produced endogenously during mitochondrial ATP synthesis. Superoxide dismutases are ubiquitous enzymes that protect against free oxygen radicals. Polymorphisms at bases 110

and 215 of the *P. jirovecii* *SOD* gene have been reported as frequent genetic changes (6–8, 27). The distribution pattern of the SNP at base 110 of the *SOD* locus (*SOD110C*, 50%; *SOD110T*, 44%) was identical to that observed in recent studies on Portuguese subjects (7, 8) but dissimilar from that reported for different geographical regions (e.g., the United Kingdom and Zimbabwe), in which *SOD110T* was the prevalent allele (27). On the contrary, the distribution pattern of the SNP at base 215 of the *SOD* gene (*SOD215C*, 44%; *SOD215T*, 34%) was different from that reported in previous Portuguese studies (7, 8), in which *SOD215T* was the most prevalent allele. The previous report on subjects from the United Kingdom and Zimbabwe also has revealed an *SOD215* distribution pattern identical to that found in the present study (27). As suggested in a previous study (8), the detection of isolates presenting genotypes *SOD110C-SOD215C* and *SOD110T-SOD215T* confirmed the different allelic distribution of the *SOD* locus between the Portuguese subjects and the subjects from the United Kingdom and Zimbabwe, since these genotypes were not detected in those geographical regions. Again, this allelic variation demonstrates the different geographical distributions of *P. jirovecii* genotypes.

The generated data (Table 4) showed that, among the cases with low to moderate parasite burden, the highest percentages were obtained for isolates presenting *DHFR312C*, *mt85C*, *SOD110T*, and *SOD215C*, whereas for high parasite burden cases the highest frequencies were detected in isolates with *DHFR312T*, *mt85T*, *SOD110C*, and *SOD215T*. These results were identical to those found in a previous report on *P. jirovecii* isolates collected from HIV-infected patients (7), suggesting that genetic markers for parasite burden in PcP episodes exist. The synonymous polymorphisms *DHFR312* and *SOD215*, as well as the SNP *SOD110* (located at an intron), do not modify amino acid coding, and therefore it is not expected to change the function of the corresponding proteins. However, it is well known that synonymous polymorphisms may interfere with protein activity as a result of an association with haplotypes involving other functional nonsynonymous polymorphisms, or as a result of allele-specific differences in mRNA folding that could influence the splicing process or translational control and regulation (7, 20, 34). Also, data showed that the association observed for isolates *mt85C* and *mt85T* (Table 4) was due mainly to a lack of high parasite burden for *mt85C* cases and an excess of high parasite burden among isolates with *mt85T*. It has been reported that nucleotide variation in rRNA may affect ribosomal conformation and activity (31). These results suggest that the SNP at position 85 of the *mtLSU rRNA* gene of *P. jirovecii* is somehow related to its pathogenic capacity and virulence. Additionally, these associations may be caused by the existence of *P. jirovecii* haplotypes, which may confer different replication rates to genetically distinct organisms. The SNP at base 85 of the *mtLSU rRNA* locus should be tested as a genetic marker of parasite burden in further studies.

The cross-genetic analysis established several statistical associations among the SNPs studied: *mt85C-SOD110T*, *SOD110T-SOD215C*, and *SOD110C-SOD215T*. Previous studies revealed that *SOD110T* was persistently associated with *SOD215C*, just as *SOD110C* was associated with *SOD215C* (7, 8, 27). In the present study the linkage disequilibrium between the SNPs at bases 110 and 215 of the *SOD* locus was confirmed,

suggesting the existence of distinct *P. jirovecii* haplotypes composed of the SNP combinations *SOD215C-SOD110T* and *SOD215T-SOD110C*. These results are consistent with the hierarchical cluster analysis performed in this study.

Statistical analysis demonstrated that the MLG *SOD110C-SOD215T* was less frequent in patients in age group A, while the MLG *SOD110T-SOD215C* was more frequent in that age group and less frequent among patients in age group B. These results were identical to those found in a recent study (8) and suggest that the majority of HIV/*P. jirovecii*-infected Portuguese patients between 30 and 39 years old are infected with *P. jirovecii* organisms presenting genotypes with the SNP combination *SOD110C-SOD215T*. Further research with a larger population and the respective demographic and clinical information should be conducted to clarify these associations.

Compared to conventional molecular techniques, the MPCR/SBE protocol, described for the first time in the study of *P. jirovecii*, was demonstrated to be a high-sensitivity and -fidelity method for genotyping multiple *P. jirovecii* SNPs in a rapid two-step procedure that could be used for molecular epidemiology, prognostic classification, and the molecular diagnosis of PcP. Also, the present data suggest that this method is a helpful tool in the study of other microorganisms, such as bacteria (e.g., *Escherichia coli*), or even of humans (10, 18). The polymorphisms studied that were shown to be suitable genetic targets for use in further epidemiological investigation focused on the identification and characterization of *P. jirovecii* haplotypes potentially correlated with PcP clinical data and outcome. After confirming the importance of these specific molecular markers, MPCR/SBE may be used as a promising method for the prognosis and diagnosis of PcP. This technique may be used to rapidly identify PcP cases presenting more pathogenic or resistant *P. jirovecii* genotypes to help clinical decision-making in the management of PcP. Additional studies using high-throughput methodologies, such as DNA pooling, in association with MPCR/SBE are required to investigate these potential associations in a large-scale format.

ACKNOWLEDGMENTS

The present study was partially supported by Associação para a Investigação e Desenvolvimento da Faculdade de Medicina de Lisboa.

F. Esteves was supported by a Ph.D. grant (SFRH/BD/22193/2005) from Fundação para a Ciência e a Tecnologia.

REFERENCES

1. Beard, C. B., et al. 2000. Genetic variation in *Pneumocystis carinii* isolates from different geographic regions: implications for transmission. *Emerg. Infect. Dis.* **6**:265–272.
2. Costa, M. C., F. Esteves, F. Antunes, and O. Matos. 2006. Genetic characterization of the dihydrofolate reductase gene of *Pneumocystis jirovecii* isolates from Portugal. *J. Antimicrob. Chemother.* **58**:1246–1249.
3. Costa, M. C., F. Esteves, F. Antunes, and O. Matos. 2006. Multilocus genotyping of *Pneumocystis jirovecii* in immunocompromised patients: preliminary results. *J. Eukaryot. Microbiol.* **53**(Suppl. 1):S104–S105.
4. Costa, M. C., et al. 2005. Detection of *Pneumocystis jirovecii* dihydropteroate synthase polymorphisms in patients with *Pneumocystis pneumonia*. *Scand. J. Infect. Dis.* **37**:766–771.
5. Crothers, K., et al. 2005. Severity and outcome of HIV-associated *Pneumocystis pneumonia* containing *Pneumocystis jirovecii* dihydropteroate synthase gene mutations. *AIDS* **19**:801–805.
6. Denis, C. M., et al. 1998. Molecular cloning and characterization of a superoxide dismutase (sod) gene in *Pneumocystis carinii*. *J. Eukaryot. Microbiol.* **45**:475–483.
7. Esteves, F., et al. 2010. Identification of relevant single-nucleotide polymorphisms in *Pneumocystis jirovecii*: relationship with clinical data. *Clin. Microbiol. Infect.* **16**:878–884.
8. Esteves, F., et al. 2010. Population structure of *Pneumocystis jirovecii* isolated from immunodeficiency virus-positive patients. *Infect. Genet. Evol.* **10**:192–199.
9. Esteves, F., et al. 2008. *Pneumocystis jirovecii* multilocus genotyping profiles in patients from Portugal and Spain. *Clin. Microbiol. Infect.* **14**:356–362.
10. Faber, A., G. N. Ferreira, S. N. Silva, J. Rueff, and J. Gaspar. 2005. Multiplex PCR-single-base extension genotyping of multiple glutathione S-transferase polymorphisms. *Biotechnol. Appl. Biochem.* **41**:9–15.
11. Fan, J. B., et al. 2000. Parallel genotyping of human SNPs using generic high-density oligonucleotide tag arrays. *Genome Res.* **10**:853–860.
12. Hauser, P. M., et al. 2001. Genetic diversity of *Pneumocystis carinii* in HIV-positive and -negative patients as revealed by PCR-SSCP typing. *AIDS* **15**:461–466.
13. Hauser, P. M., et al. 2010. Interhuman transmission as a potential key parameter for geographical variation in the prevalence of *Pneumocystis jirovecii* dihydropteroate synthase mutations. *Clin. Infect. Dis.* **51**:e28–e33.
14. Hauser, P. M., P. Sudre, A. Nahimana, and P. Francioli. 2001. Prophylaxis failure is associated with a specific *Pneumocystis carinii* genotype. *Clin. Infect. Dis.* **33**:1080–1082.
15. Helweg-Larsen, J., T. L. Benfield, J. Eugen-Olsen, J. D. Lundgren, and B. Lundgren. 1999. Effects of mutations in *Pneumocystis carinii* dihydropteroate synthase gene on outcome of AIDS-associated *P. carinii* pneumonia. *Lancet* **354**:1347–1351.
16. Hirschhorn, J. N., et al. 2000. SBE-TAGS: an array-based method for efficient single-nucleotide polymorphism genotyping. *Proc. Natl. Acad. Sci. U. S. A.* **97**:12164–12169.
17. Hirschhorn, L. R., et al. 2006. Gender differences in quality of HIV care in Ryan White CARE Act-funded clinics. *Women's Health Issues* **16**:104–112.
18. Hommais, F., S. Pereira, C. Acquaviva, P. Escobar-Paramo, and E. Denamur. 2005. Single-nucleotide polymorphism phylotyping of *Escherichia coli*. *Appl. Environ. Microbiol.* **71**:4784–4792.
19. Kim, S., and A. Misra. 2007. SNP genotyping: technologies and biomedical applications. *Annu. Rev. Biomed. Eng.* **9**:289–320.
20. Kimchi-Sarfaty, C., et al. 2007. A “silent” polymorphism in the MDR1 gene changes substrate specificity. *Science* **315**:525–528.
21. Ma, L., L. Borio, H. Masur, and J. A. Kovacs. 1999. *Pneumocystis carinii* dihydropteroate synthase but not dihydrofolate reductase gene mutations correlate with prior trimethoprim-sulfamethoxazole or dapsone use. *J. Infect. Dis.* **180**:1969–1978.
22. Markoulatos, P., N. Siafakas, and M. Moncany. 2002. Multiplex polymerase chain reaction: a practical approach. *J. Clin. Lab. Anal.* **16**:47–51.
23. Matos, O., et al. 2001. Effect of oral washes on the diagnosis of *Pneumocystis carinii* pneumonia with a low parasite burden and on detection of organisms in subclinical infections. *Eur. J. Clin. Microbiol. Infect. Dis.* **20**:573–575.
24. Matos, O., and F. Esteves. 2010. *Pneumocystis jirovecii* multilocus gene sequencing: findings and implications. *Future Microbiol.* **5**:1257–1267.
25. Matos, O., and F. Esteves. 2010. Epidemiology and clinical relevance of *Pneumocystis jirovecii* Frenkel, 1976 dihydropteroate synthase gene mutations. *Parasite* **17**:219–232.
26. Miller, R. F., et al. 2005. Genotypic variation in *Pneumocystis jirovecii* isolates in Britain. *Thorax* **60**:679–682.
27. Miller, R. F., A. R. Lindley, A. S. Mallin, H. E. Ambrose, and A. E. Wakefield. 2005. Isolates of *Pneumocystis jirovecii* from Harare show high genotypic similarity to isolates from London at the superoxide dismutase locus. *Trans. R. Soc. Trop. Med. Hyg.* **99**:202–206.
28. Montes-Cano, M. A., et al. 2004. *Pneumocystis jirovecii* genotypes in the Spanish population. *Clin. Infect. Dis.* **39**:123–128.
29. Morris, A., et al. 2004. Current epidemiology of *Pneumocystis pneumonia*. *Emerg. Infect. Dis.* **10**:1713–1720.
30. Nahimana, A., M. Rabodonirina, J. Bille, P. Francioli, and P. M. Hauser. 2004. Mutations of *Pneumocystis jirovecii* dihydrofolate reductase associated with failure of prophylaxis. *Antimicrob. Agents Chemother.* **48**:4301–4305.
31. Noller, H. F. 1984. Structure of ribosomal RNA. *Annu. Rev. Biochem.* **53**:119–162.
32. Pastinen, T., A. Kurg, A. Metspalu, L. Peltonen, and A. C. Syvanen. 1997. Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res.* **7**:606–614.
33. Schnell, J. R., H. J. Dyson, and P. E. Wright. 2004. Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu. Rev. Biophys. Biomol. Struct.* **33**:119–140.
34. Shen, L. X., J. P. Babilion, and V. P. Stanton, Jr. 1999. Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proc. Natl. Acad. Sci. U. S. A.* **96**:7871–7876.
35. Tsolaki, A. G., P. Beckers, and A. E. Wakefield. 1998. Pre-AIDS era isolates of *Pneumocystis carinii* f. sp. hominis: high genotype similarity with contemporary isolates. *J. Clin. Microbiol.* **36**:90–93.
36. Vreeland, W. N., R. J. Meagher, and A. E. Barron. 2002. Multiplexed, high-throughput genotyping by single-base extension and end-labeled free-solution electrophoresis. *Anal. Chem.* **74**:4328–4333.
37. Wakefield, A. E., et al. 1990. Amplification of mitochondrial ribosomal RNA sequences from *Pneumocystis carinii* DNA of rat and human origin. *Mol. Biochem. Parasitol.* **43**:69–76.