Journal of Clinical Microbiology, Mar. 2007, p. 881–886 Vol. 45, No. 3
0095-1137/07/$08.00+0 doi:10.1128/JCM.02245-06
Copyright © 2007, American Society for Microbiology. All Rights Reserved.

Frequent In Vitro Recombination in Internal Transcribed Spacers 1 and 2 during Genotyping of *Pneumocystis jirovecii*▽

Jessica Beser, Per Hagblom, and Victor Fernandez*

Department of Parasitology, Mycology and Environmental Microbiology, Swedish Institute for Infectious Disease Control, Solna, and Department of Microbiology, Tumor and Cell Biology, Karolinska Institute, Stockholm, Sweden

Received 2 November 2006/Returned for modification 28 November 2006/Accepted 21 December 2006

Pneumocystis jirovecii **is the causative agent of** ***Pneumocystis*** **pneumonia (PCP) in immunocompromised persons. Knowledge of the transmission and epidemiology of PCP is still incipient, and investigations on these subjects are based exclusively on applications of molecular typing techniques. The polymorphic internal transcribed spacers ITS1 and ITS2 in the ribosomal DNA operon, which in the** ***P. jirovecii*** **genome exist as single-copy DNA, are commonly used as target loci for isolate typing. In the course of genotyping** ***P. jirovecii*** **in respiratory specimens from PCP patients by amplification and cloning of a large number of ITS sequences, we found mixed infections (two or more types) in 50% of the samples. In a majority of the specimens with mixed infections, we detected many ITS haplotypes (combinations of ITS1 and ITS2 types) that appeared to be products of recombination between globally common ITS haplotypes present in the same sample. Here we present results of a series of experiments showing that essentially all ITS recombinants are chimeras formed during the genotyping process. Under standard conditions, as many as 37% of the amplified sequences could be hybrid DNA artifacts. We show that by modifying PCR amplification conditions, ITS chimera formation could be largely abolished and the erroneous establishment of artifactual haplotypes avoided. The accurate assessment of genetic diversity is fundamental for a better understanding of the epidemiology and biology of** ***P. jirovecii*** **infections.**

The infectious fungus *Pneumocystis jirovecii* is the etiological agent of *Pneumocystis* pneumonia (PCP) in immunocompromised individuals. Since *P. jirovecii* cannot be cultivated in vitro, investigations of the transmission and epidemiology of this organism have been based on applications of molecular typing techniques (1). For *P. jirovecii*, sequence analysis of the internal transcribed spacer (ITS) regions in the nuclear ribosomal DNA (rDNA) gene complex, which in this fungus exists as a single-copy locus (2, 24), is the most informative epidemiological tool available at present. ITS1 is located between the coding regions of the 18S and 5.8S rRNA genes, and ITS2 is located between the 5.8S and 26S rRNA genes. The sequence diversity of ITS1 and ITS2 among different strains of *P. jirovecii* and the fact that these segments are removed during ribosomal biogenesis and should not be subjected to selective pressure make them suitable targets for genotyping (10). To date, more than 30 ITS1 genotypes and 40 ITS2 genotypes with more than 90 haplotypes (combinations of ITS1 and ITS2 types) have been described worldwide, based on two different but similar typing systems (8, 26). Some ITS haplotypes are globally more common, and coinfections with multiple types of *P. jirovecii* often occur (4, 8, 11, 15, 17, 25, 26). Although these typing systems have been useful, questions have been raised about the extensive polymorphism and stability of the ITS locus and its use for the genotyping of *P. jirovecii*.

During the genotyping of specimens from patients with *P. jirovecii* infections, we noticed the presence of recombinants between main ITS haplotypes in many of the specimens with mixed infections. Given the high frequency of this feature, we considered the possibility that some of these recombined haplotypes had been generated during the genotyping procedure. It is well established that recombination, or chimera formation, can occur during PCR when the template is a mixture of similar sequences rather than a single target (13, 19, 20). When this phenomenon takes place, in vitro-generated recombinant molecules will be present in the amplified product, which will result in overestimation of the diversity in the starting material (27, 28). Another potential complication is the introduction of point mutations during PCR amplification due to misincorporation by the DNA polymerase (16). In an effort to elucidate the origin of recombinants in the clinical material, we first performed experiments that showed the in vitro source of the recombinant sequences. Second, we attempted to resolve in what step and at what level the in vitro-generated recombinants were produced. Third, we made modifications to the standard protocol of Lee et al. (8) to reduce the number of erroneous haplotypes in samples with mixed infections. Furthermore, we used exons and introns of the thioredoxin reductase (*TRR1*) gene as reference DNA templates to evaluate the rate of point mutations and the performance of DNA polymerases in the ITS and 5.8S loci.

MATERIALS AND METHODS

Processing of specimens and DNA extraction. Respiratory specimens used in this study were collected and referred to the Swedish Institute for Infectious Disease Control, where they were analyzed by following routine procedures for the diagnosis of *Pneumocystis jirovecii*. Nine volumes of a bronchoalveolar lavage or sputum sample were mixed with 1 volume of 65 mM dithiothreitol and incubated for 15 min at 37°C. One volume of 50% ethanol was added for inactivation of human immunodeficiency virus. The suspension was centrifuged at 1,800 × *g* for 5 min, and the supernatant was discarded. DNA extractions were

* Corresponding author. Mailing address: Department of Parasitology, Mycology and Environmental Microbiology, Swedish Institute for Infectious Disease Control, SE-17182 Solna, Sweden. Phone: 46 8 4572553. Fax: 46 8 318450. E-mail: Victor.Fernandez@smi.ki.se.

▽ Published ahead of print on 3 January 2007.

performed with a commercial purification system, the QIAamp DNA minikit (QIAGEN, Hilden, Germany), according to the protocol supplied by the manufacturer.

Mixture of specimen DNA. Specimens 232, 790, 926, 1017, and 1874, containing a single ITS haplotype, and specimen 308, containing multiple haplotypes, as determined by genotyping carried out according to the work of Lee et al. (8), were used. *P. jirovecii* DNAs from two specimens were mixed to an equimolar ratio calculated by real-time PCR quantification of the gene coding for the large ribosomal subunit (LSU) of *P. jirovecii* (V. Fernandez et al., unpublished data). The mixtures of *P. jirovecii* DNA were amplified according to the protocols described below.

Mixture of PCR products. *P. jirovecii* DNAs from specimens 232 and 926 were amplified independently of each other according to the protocol for ITS amplification with *Taq* polymerase that is described below. The PCR products were purified with a QIAquick PCR purification kit (QIAGEN, Hilden, Germany) and spectrophotometrically quantified, and the PCR products were then mixed in a molar ratio of 1:1. After mixing, the PCR products were treated in two different ways: either cloned directly or pretreated before cloning. The pretreatment, which was designed to reproduce conditions favoring heteroduplex formation in later cycles of the regular genotyping PCR, was performed by taking 10 μl of the PCR mixture and combining it with 1× PCR buffer (Invitrogen, Carlsbad, CA), 3 mM $MgCl_2$, 200 μM each deoxynucleotide, and 20 pmol each of primers FX (5′-TTCCGTAGGTGAACCTGCG-3′) and RT2 (5′-CTGATT TGAGATTAAAATTCTTG-3′). The total volume was adjusted to 50 μl with MilliQ water, and DNA polymerase was not added. This mixture was then subjected to the following thermal cycling conditions: 5 min at 94°C; two cycles of 60 s at 94°C, 60 s at 56°C, and 60 s at 72°C; and 15 min at 72°C.

ITS amplification with *Taq* polymerase. *P. jirovecii* DNAs from individual specimens or in mixtures were amplified by a nested PCR targeting the ITS1–5.8S–ITS2 region as previously described (8). In brief, a first amplification reaction (PCR I) mixture containing 5 μl of template DNA mix or H_2O as a negative control, 2.5 U *Taq* polymerase (Invitrogen, Carlsbad, CA), 200 μM each deoxynucleotide, 3 mM $MgCl_2$, and 20 pmol each of primers 1724F2 (5′-AGTTG ATCAAATTTGGTCATTTAGAG-3′) and ITS2R (5′-CTCGGACGAGGATC CTCGCC-3′) in 1× PCR buffer (Invitrogen, Carlsbad, CA) was adjusted with MilliQ water to a total volume of 50 μl. The PCR was carried out under the following conditions: 3 min at 94°C; 5 cycles of 90 s at 94°C, 90 s at 62°C, and 90 s at 72°C; 30 cycles of 60 s at 94°C, 60 s at 60°C, and 60 s at 72°C; and 15 min at 72°C. A second amplification reaction (PCR II) containing 2 μl of the PCR I product, 2.5 U *Taq* polymerase, 200 μM each deoxynucleotide, 1.5 mM $MgCl_2$, 20 pmol each of primers FX (5′-TTCCGTAGGTGAACCTGCG-3′) and RT2 (5′-CTGATTTGAGATTAAAATTCTTG-3′), 1× PCR buffer, and MilliQ water in a total volume of 50 μl was set up. PCR II was performed under the following conditions: 3 min at 94°C; 5 cycles of 90 s at 94°C, 90 s at 58°C, and 90 s at 72°C; 30 cycles of 60 s at 94°C, 60 s at 56°C, and 60 s at 72°C; and 15 min at 72°C. The amplification products were separated by electrophoresis on a 1.2% agarose gel containing ethidium bromide (EtBr) and were visualized by UV light.

ITS amplification with proofreading *Tgo* polymerase. A first amplification of the ITS1–5.8S–ITS2 region was done with primers 1724F2 and ITS2R, i.e., the same primers as for amplifications using *Taq* polymerase. For this PCR, two master mixes were prepared as follows. Five microliters of a template DNA mixture (specimens 926 and 232) was added to master mix 1 containing 20 pmol of each primer, deoxynucleotides (final concentration, 200 μM each), and MilliQ water, to a final volume of 25 μl. In master mix 2, 1× Expand High Fidelity (Roche, Penzberg, Germany), 2.6 U Expand High Fidelity enzyme mix containing *Taq* polymerase and *Tgo* polymerase with proofreading activity (Roche, Penzberg, Germany), and MilliQ water were mixed to a final volume of 25 μl. The two master mixes were subsequently combined, and the PCR was conducted under the same conditions as for the first amplification with *Taq* polymerase. For the second amplification, forward primer FX and reverse primer RT2 were used. Two microliters of the first PCR product was added to a master mix 1 containing 20 pmol of each primer, deoxynucleotides (final concentration, 200 μM each), and MilliQ water, to a final volume of 25 μl. In master mix 2, 1× Expand High Fidelity buffer, 2.6 U Expand High Fidelity enzyme mix (both from Roche, Penzberg, Germany), and MilliQ water were mixed to a final volume of 25 μl. The two master mixes were subsequently combined, and the PCR was conducted under the same conditions as for the second amplification with *Taq* polymerase. The amplified products were separated by electrophoresis on a 1.2% agarose gel, stained with EtBr, and visualized under UV light. Prior to cloning, 1 μl ATP and 1 μl *Taq* polymerase were added to the PCR mix and incubated at 37°C for 30 min to produce A′ overhangs on the fragments.

ITS amplification with *Taq* polymerase under modified conditions. In one experiment, the DNA template mixture (specimens 926 and 232) was amplified by the same protocol used previously (8) but with longer extension times. PCR I consisted of 3 min at 94°C; 5 cycles of 90 s at 94°C, 90 s at 62°C, and 4.5 min at 72°C; 5 cycles of 60 s at 94°C, 60 s at 60°C, and 4.5 min at 72°C; 25 cycles of 60 s at 94°C, 60 s at 60°C, and 60 s at 72°C; and 15 min at 72°C. PCR II consisted of 3 min at 94°C; 5 cycles of 90 s at 94°C, 90 s at 58°C, and 4.5 min at 72°C; 5 cycles of 60 s at 94°C, 60 s at 56°C, and 4.5 min at 72°C; 25 cycles of 60 s at 94°C, 60 s at 56°C, and 60 s at 72°C; and 15 min at 72°C. In another experiment, the DNA mixture (specimens 926 and 232) was amplified by the same protocol but with a higher melting temperature, a longer extension time, and fewer cycles. PCR I consisted of 5 min at 96°C; 25 cycles of 60 s at 94°C, 60 s at 60°C, and 4.5 min at 72°C; and 7 min at 72°C. PCR II consisted of 5 min at 94°C; 20 cycles of 60 s at 94°C, 60 s at 56°C, and 4.5 min at 72°C; and 7 min at 72°C. In the third experiment, the DNA mixture (specimens 926 and 232) was amplified using a twofold primer concentration in the first PCR and a fivefold primer concentration in the second PCR but otherwise according to the same protocol as before (8), under the same modified conditions as just described, with higher denaturation temperatures, a longer extension time, and fewer cycles.

Reconditioning PCR. The *P. jirovecii* DNA mixture (specimens 926 and 232) was amplified with *Taq* polymerase under standard conditions, and 5 μl of PCR II products was transferred to a new master mix containing 2.5 U *Taq* polymerase (Invitrogen, Carlsbad, CA), 200 μM each deoxynucleotide, 1.5 mM $MgCl_2$, 20 pmol each of primers FX and RT2, 1× PCR buffer, and MilliQ water in a total volume of 50 μl. The reaction mixtures were subjected to 3 min at 94°C and three cycles of 60 s at 94°C, 60 s at 56°C, and 60 s at 72°C.

Amplification of the thioredoxin reductase (*TRR1*) gene. Primers TRR1F1 (5′-CTTGTTAATCTCTCTAGATCAACGTC-3′) and TRR1R1 (5′-TTAATTT GTCCCTCCTAACAAGTAG-3′) were designed to amplify an 888-bp segment of the *TRR1* gene, followed by nested primers TRR1F2 (5′-GAAGAAAAGAG CCTTTAATAGATAC-3′) and TRR1R2 (5′-GTAGATATACTTTAGTTGCA TATCTCG-3′), amplifying an 817-bp segment including 5 introns and 5 exons. For amplification with *Taq* polymerase (Invitrogen, Carlsbad, CA), the same protocol was used as for ITS but under the following conditions. For PCR I, conditions were 5 min at 95°C; 40 cycles of 30 s at 95°C, 60 s at 55°C, and 60 s at 72°C; and 7 min at 72°C. For PCR II, conditions were 3 min at 95°C; 35 cycles of 30 s at 95°C, 30 s at 53°C, and 60 s at 72°C; and 7 min at 72°C. For amplification with Expand High Fidelity enzyme mix (Roche, Penzberg, Germany), the same TRR1F1 and TRR1R1 primers were used for the first PCR, followed by TRRF2 and TRRR2 for the second amplification but with the protocol for proofreading polymerase described under "ITS amplification with proofreading *Tgo* polymerase" above. The same PCR conditions were used as described above. PCR I consisted of 5 min at 95°C; 40 cycles of 30 s at 95°C, 60 s at 55°C, and 60 s at 72°C; and 7 min at 72°C. PCR II consisted of 3 min at 95°C; 35 cycles of 30 s at 95°C, 30 s at 53°C, and 60 s at 72°C; and 7 min at 72°C.

Cloning and sequencing of PCR products. PCR II products were cloned into the pCR 2.1-TOPO plasmid vector using the Topo TA cloning kit (Invitrogen, Carlsbad, CA) according to the instructions provided by the manufacturer. Bacterial colonies were screened with M13 vector-specific primers in a PCR mixture containing 2.5 U *Taq* polymerase (Invitrogen, Carlsbad, CA), 200 μM each deoxynucleotide, 1.5 mM $MgCl_2$, 5 ng/μl of each primer, 1× PCR buffer, and MilliQ water to a final volume of 20 μl. Bacterial colonies were picked directly into the reaction mixture and amplified under the following conditions: 5 min at 94°C; 25 cycles of 30 s at 94°C, 30 s at 45°C, and 1 min at 72°C; and 7 min at 72°C. The PCR products were separated in a 1.2% agarose gel containing EtBr and were visually inspected to confirm the presence of cloned inserts. Recombinants were sequenced in both directions using M13 primers and dye terminator chemistry (Applied Biosystems, Warrington, United Kingdom). Sequences were edited and analyzed using the BioEdit Sequence Alignment Editor, version 7.0.4.1.

Statistical analysis and other computer analyses. Statistical analyses were performed using SigmaStat software, version 3.1. Mutation rate differences between loci were analyzed using the nonparametric Mann-Whitney rank sum test. Oligonucleotide primer secondary and tertiary structures were modeled with NetPrimer (PREMIER Biosoft) and MacVector (Accelrys Software, Inc.) software.

RESULTS

Recombinant ITS sequences in clinical specimens. Genotyping of *P. jirovecii* by amplification and cloning of 408 ITS sequences in respiratory specimens from 64 individuals infected with the fungus revealed the presence of only a single ITS haplotype in half of the cases (32/64), whereas the remain-

TABLE 1. *P. jirovecii* ITS haplotypes detected in PCR-amplified respiratory specimens of *P. jirovecii*-infected individuals[a]

Haplotype(s)[b]	No. of specimens	No. of clones[c]	Found as single-type infection
Eg	28	106	Yes
Ne	10	61	Yes
Bi, Eb	9	44, 38	Yes
Ai	6	24	Yes
Ea, Ec, Jf	5	23, 15, 24	Yes
Iv	4	17	Yes
Be, Ef, Ei, Ig, Ji	2	2–3	No
Af, Bg, Ce, Cg, Ci, Db, E100B, Ee, Eh, Eo, Ep, E6, Ev, Gg, Ja, Jc, Jg, Kc, Kf, Kg, Ki, Nf, Ng, Nh, Ni, Nj, No	1	1–3	No
Total (41 haplotypes)	64	408	

[a] Genotyping at the ITS1–5.8S–ITS2 locus of the rDNA was performed as described by Lee et al. (8).
[b] Combination of ITS1 and ITS2.
[c] Five to 13 clones were sequenced per specimen.

ing samples (32/64) contained two or more haplotypes. In 16 of the 32 specimens with mixed infections, we detected what appeared to be recombinant ITS sequences. This putative recombinant ITS offspring consisted of combinations of ITS1 and ITS2 seemingly generated from other ITS haplotypes also present in the sample. ITS haplotypes, as well as the number of specimens and the number of clones sequenced per haplotype, are shown in Table 1. A total of 41 different ITS haplotypes were identified, 9 of which (17 to 106 clones each) were found in the amplified DNA of four or more specimens. Each of these nine haplotypes was detected in at least one single-type specimen as verified by the analysis of five or more cloned sequences. Of the remaining haplotypes, 5 were found in two specimens (two to three clones each) and 27 in only one sample (one to three clones each). These 32 less frequent haplotypes were detected only in samples infected with multiple ITS types, which amounted to 25 specimens. In 23 of these 25 specimens, the ITS1 and/or ITS2 sequences that were detected in the less frequent haplotypes were the same as those found in the major haplotypes present in that particular sample. The major haplotypes found in these samples correspond to those that are frequently found in epidemiological studies of *P. jirovecii* infections. Therefore, the major dividing line in this material was between single infections, where only one ITS haplotype was found, and coinfections, where a multitude of recombination events were evident. It was not possible from this data set to draw any conclusions as to whether these recombinants were true in vivo phenomena or whether they were generated in vitro during the experimental procedure in the typing protocol.

Recombinant sequences are chimeras generated in vitro. To reconstitute a situation where *P. jirovecii* DNA from a coinfected specimen is typed according to the current protocol by Lee et al. (8), material from two single-type-infected specimens was mixed in equimolar amounts and then amplified, cloned, and sequenced with this methodology. For this purpose, the *P. jirovecii* DNAs in various samples were quantified using a real-time PCR assay targeting the mitochondrial LSU gene of the fungus (Fernandez et al., unpublished). Identical amounts of *P. jirovecii* DNA from specimen 926, with haplotype Eg, and specimen 232, with haplotype Bi, were mixed, and the standard ITS typing procedure was performed. Of 43 clones sequenced, 21 were of type Bi, 6 were of haplotype Eg, and 16 sequences were recombinants (9 Ei and 7 Bg), representing 37.2% of the total clones analyzed (Table 2). To confirm the finding of recombinants generated in vitro, a new haplotype mix was made with specimens 1874 and 790, containing haplotypes Ai and Iv, respectively (ITS2 v is a new genotype found in the Swedish population [J. Beser et al., submitted for publication]). Of 15 clones analyzed, 9 sequences were of haplotype Ai and 2 sequences were of type Iv. In addition to the parental haplotypes, two new sequences were detected, one the recombinant Ii and the other Iv/i, which is a chimeric product of a recombination between ITS2 types i and v. Unexpectedly, two Eg sequences appeared in the latter experiment; this might represent a contaminant or a real haplotype present in small amounts in either of the specimens (Table 2). A more complex mixture was also made with specimen 1017, of type Jf, and the multiple-type-infected specimen 308, containing haplotypes Bi and Ne. Sequence analysis of this

TABLE 2. In vitro generation of ITS chimeras during genotyping of mixed specimens[a]

Specimen ID	Total no. of clones sequenced	Haplotype(s)[b] (no. of clones)	Rate of chimera formation (%)
926	7	Eg (7)	
232	6	Bi (6)	
926 + 232[c]	43	Eg (6), Bi (21), **Ei** (9), **Bg** (7)	37.2
1874	34	Iv (34)	
790	5	Ai (5)	
1874 + 790[c]	15	Ai (9), Iv (2), **Ii** (1), **Iv/i**[d] (1), Eg (2)	13.3
1017	8	Jf (8)	
308	8	Bi (3), Ne (2), Be (2), Ni (1)	
1017 + 308[c]	40	Jf (12), Bi (6), Ne (7), Ni (2), **Nf** (2), **Bf** (2), **Ji** (5), **Je** (1), **Bh**[e] (1), **Nh**[e] (1), **Kh**[e,f] (1)	32.5

[a] Genotyping PCR conditions were those described by Lee et al. (8).
[b] PCR-generated recombinants are boldfaced.
[c] *P. jirovecii* LSU was quantified using a real-time PCR assay, and specimen DNAs were mixed 1:1 on a *P. jirovecii* LSU gene basis prior to PCR.
[d] The ITS2 is an in vitro intra-ITS2 recombination between ITS2 types v and i.
[e] ITS2 type h is identical to type f except for a missing T (most likely PCR generated) at position 71.
[f] The ITS1 type K sequence is presumably the product of an in vitro intra-ITS1 recombination between ITS1 types J and N or B.

TABLE 3. Suppression of chimera formation during ITS genotyping of *P. jirovecii*

Genotyping conditions	Mix of specimens 926 and 232[b]	No. of sequences detected					% Chimeras
		Total	Bi	Eg	Bg	Ei	
Standard[a]	Genomic DNA	43	21	6	7	9	37.2
PCR product mix 1[c]	PCR products	44	26	16	0	2	4.5
PCR product mix 2[d]	PCR products	35	9	24	2	0	5.7
Reconditioning PCR[e]	Genomic DNA	48	15	15	12	6	37.5
Proofreading *Tgo* polymerase	Genomic DNA	43	20	13	4	6	23.3
PCR modification 1[f]	Genomic DNA	46	22	9	8	7	32.6
PCR modification 2[g]	Genomic DNA	43	15	21	5	2	16.3
PCR modification 3[h]	Genomic DNA	36	17	18	0	1	2.8

[a] Genotyping performed according to the work of Lee et al. (8) using *Taq* polymerase. All subsequent modifications are described in Materials and Methods.
[b] *P. jirovecii* LSU was quantified using a real-time PCR assay, and specimen DNAs were mixed 1:1 on a *P. jirovecii* LSU gene basis prior to PCR.
[c] Specimen DNAs were PCR amplified separately and amplification products mixed prior to cloning.
[d] Specimen DNAs were amplified separately, PCR products mixed, and a step of denaturation/reannealing included before cloning to induce the formation of heteroduplexes.
[e] Before cloning, PCR products were diluted 10-fold and reamplified for three cycles in fresh master mix to eliminate heteroduplexes.
[f] Elongation time was increased to 4.5 min.
[g] Elongation time was increased to 4.5 min, the number of cycles was decreased from 35 to 25 in PCR I and 20 in PCR II, and the denaturing temperature was increased from 94°C to 96°C.
[h] Elongation time was increased to 4.5 min, the number of cycles was decreased, the denaturing temperature was increased from 94°C to 96°C, and the primer concentration was increased twofold in PCR I and fivefold in PCR II.

mixture revealed products of recombination between haplotypes and intra-ITS2 regions, confirming our assumption that an increased number of *P. jirovecii* DNA templates results in an even more varied population of recombinant PCR products.

ITS recombinants are generated during PCR amplification. To further investigate at what step in the genotyping protocol these in vitro artifacts were produced, experiments were designed to determine whether the recombinants were generated during PCR amplification or in the cloning into *Escherichia coli*. In the first set of these experiments, the PCR products from single-haplotype specimens 926 (Eg) and 232 (Bi) were purified, quantified, and mixed in equimolar amounts. To examine if recombination was taking place in the cloning step, this mixture was cloned directly; the analysis showed that 95.5% of the cloned sequences were of parental types (26 Bi and 16 Eg) and 4.5% were recombinants (2 Ei clones) (Table 3). When the mixture was subjected to two cycles of melting, annealing, and elongation, to allow for heteroduplex formation prior to cloning, analysis of 35 clones revealed that 94.3% of the sequences were identical to the parental haplotypes and 5.7% were recombinant Bg sequences. These findings suggested that the vast majority of the chimeras were generated during PCR amplification but also that a smaller part could have been produced later on, during the cloning step. To confirm that the recombinants were generated mainly during the elongation step and not as a consequence of heteroduplex formation in the annealing step of the PCR and subsequent DNA strand repair in *E. coli*, reconditioning PCR was performed. This was done by subjecting an equimolar mix of *P. jirovecii* DNAs from specimens 926 and 232 that had been amplified according to the ordinary protocol to three cycles of reamplification in a fresh master mix containing the full concentration of primers in order to minimize the presence of heteroduplexes. Forty-eight clones were sequenced, and in addition to the original haplotypes, 6 Ei and 12 Bg chimeric molecules were detected (37.5% [see Table 3]), showing that most of the recombinants were produced during the elongation step of the PCR.

Suppression of in vitro chimera formation. The following experiments were designed as an attempt to suppress the generation of artifacts during the PCR amplification of the ITS genotyping protocol. The standard genotyping method was used as a starting point from which the amplification protocol was changed in different ways. A modification that was tested consisted in the introduction of a proofreading polymerase (*Tgo*) in conjunction with the *Taq* polymerase. With this modification, the frequency of chimeric sequences decreased to 23.3% (Table 3), showing that the presence of a polymerase with proofreading activity in the amplification reaction had a considerable effect on the formation of recombinants. In another set of experiments, several stepwise alterations were made to the standard protocol. First, the elongation time was increased; this resulted in a minor reduction in the frequency of chimera formation, to 32.6% (Table 3). Second, a higher melting temperature and fewer cycles were used in addition to increased elongation time, resulting in a further reduction in the frequency of chimeras, to 16.3% in 43 clones analyzed. Finally, an increase in primer concentration was introduced together with the extended elongation time, elevated melting temperature, and fewer cycles. The primer concentration was increased twofold in the first PCR and fivefold in the second PCR of the nested protocol, and the combined effect of these measures was to drastically decrease the number of chimeras to 1 in 36 sequenced clones (2.8%). It was evident from this set of experiments that relatively simple modifications reduce the number of in vitro-generated artifacts to a degree that is almost at the level of detection.

PCR performance at the ITS and *TRR1* loci. An additional observation throughout this analysis was that the overall frequency of point mutations in the ITS1–5.8S–ITS2 DNA was higher than expected. This observation, in combination with the finding that in vitro generation of recombinants could be partially abolished by the addition of a proofreading polymerase to the reaction mixture, raised the question whether intrinsic features in this locus made it prone to nucleotide misincorporation during replication in vitro. To address this, we chose

intronic and exonic regions of the thioredoxin reductase gene (*TRR1*) as a reference locus with which to compare frequencies of nucleotide substitution under the standard PCR conditions for genotyping of *P. jirovecii* (8). This showed that by using *Taq* polymerase, a significantly higher rate ($P < 0.05$) of nucleotide substitution took place in the ITSs and the 5.8S DNA (2.6 ± 1.2 misincorporations/1,000 bp, with 12,558 bp sequenced) than in the introns and exons of the *TRR1* gene (1.5 ± 0.4 misincorporations/1,000 bp, with 30,600 bp sequenced). However, these differences disappeared when the proofreading DNA polymerase *Tgo* was added to the reaction mixture (0.9 ± 0.1 misincorporation/1,000 bp, with 16,905 bp sequenced, for ITS1–5.8S–ITS2 and 1.0 ± 0.1 misincorporation/1,000 bp, with 43,605 bp sequenced, for *TRR1*). These results further suggest that a proofreading enzyme should be included in PCR-based typing schemes for *P. jirovecii*.

DISCUSSION

In the course of genotyping *P. jirovecii* in respiratory samples by amplification and cloning of a large number of ITS sequences, we observed a high frequency of DNA recombinants in specimens infected with multiple strains of this fungus. Many of the less common ITS1–ITS2 combinations found in these specimens were evident recombinations of two globally common ITS haplotypes present in the same sample. The question that arose was whether these recombinants were a consequence of recombinations taking place in vivo or whether recombinant ITS haplotypes could have been generated in vitro during the genotyping procedure. To address this question, we mixed specimens of known single haplotypes, and after testing of these mixtures with the standard protocol, the results showed that ITS chimeras were generated in the test tube. These processes both generated new haplotypes, by combining different ITS1 and ITS2 genotypes, and created novel genotypes through recombination within the ITSs, even though the latter was observed infrequently. We therefore concluded that essentially all the recombinant ITS sequences observed in the clinical material could indeed be chimeras formed during the genotyping procedure.

The appearance of chimeric DNA molecules is a potential risk when mixed templates of related sequences are amplified by ordinary PCR protocols. In vitro recombination of rDNA and other DNA templates with high sequence similarity has been observed in several studies, including studies of bacterial diversity (7, 9), viral transcription factors (13), and genes coding for polymorphic antigens in protozoan parasites (23). These recombinant molecules can be created as a consequence of chimera formation during the elongation step of the PCR, or by heteroduplexes formed by base pairing between heterologous molecules during amplification, which in turn will be resolved into chimeras by the DNA repair systems of *E. coli* after transformation (6, 14, 18). Our data showed that with the ordinary protocol for genotyping of *P. jirovecii*, the chimeric ITS sequences are formed primarily during elongation. Under these conditions, more than one out of three amplified sequences could be a hybrid molecule. A factor that has an effect on the formation of PCR chimeras is the pausing or stalling of the DNA polymerase along the template, which may lead to the dissociation of the polymerase and the template (3), leaving a prematurely terminated extension product, which in turn may act as a primer on a closely related molecule in the next round of amplification. In this regard, the concentration of the primer in the PCR can be a critical variable influencing the rate of chimera formation. In the standard protocol used for ITS genotyping of *P. jirovecii*, the effective concentration of primer is possibly compromised, since hairpin and duplex formation in three of the primers used in the nested amplification is predicted by various software systems for primer design and sequence analysis. Furthermore, alternative target sites for the primers could be present in non-*Pneumocystis* DNA in the patient samples, since primers flanking the ITSs are directed to generally conserved sequences. The effect of increasing the primer concentration on the reduction of in vitro-generated artifacts is likely to reflect a more efficient priming in which the oligonucleotides outcompete the prematurely terminated elongation products in the annealing step.

The existence of a sexual replication cycle in the genus *Pneumocystis* is suggested by several pieces of indirect evidence, including the observation in early cysts of *Pneumocystis carinii* of structures resembling the synaptonemal complexes formed during meiotic recombination of sister chromatids (12); the ultrastructural morphology of binucleated trophic forms of *P. carinii*, reminiscent of nuclei and nucleus-associated organelles in conjugating *Saccharomyces cerevisiae* (5); and the identification in an expressed sequence tag library of meiosis-specific genes and other genes related to sexual reproduction in *P. carinii*, with a genomic organization similar to that of the mating locus described for *Cryptococcus neoformans* (22). In addition to these observations, ITS typing data have recently been presented as indicative of the occurrence of genetic recombination in *P. jirovecii* populations (17, 21). Although the results reported here cannot be used to draw specific conclusions about clonal or recombining modes of reproduction in *P. jirovecii*, they do suggest that meiotic recombination in the rDNA of this fungus is not a frequent event.

The findings described in this paper further suggest that the diversity in *P. jirovecii* populations has in all likelihood been overestimated by genotyping systems targeting the ITS locus. The modifications of the typing protocol described here drastically reduce the number of sequence artifacts that were produced by the original procedure. This enables a more correct estimation of the genetic diversity of *P. jirovecii* and a more accurate assessment of the epidemiology of *P. jirovecii* pneumonia. Since the phenomenon of in vitro-generated artifacts from mixed templates is common, the procedural modifications presented here could be applicable to other diagnostic and typing systems.

ACKNOWLEDGMENTS

This work was supported by grants from the Swedish International Development Cooperation Agency and the Swedish-South African Health Forum.

REFERENCES

1. **Beard, C. B., P. Roux, G. Nevez, P. M. Hauser, J. A. Kovacs, T. R. Unnasch, and B. Lundgren.** 2004. Strain typing methods and molecular epidemiology of *Pneumocystis* pneumonia. Emerg. Infect. Dis. **10:**1729–1735.
2. **Giuntoli, D., S. L. Stringer, and J. R. Stringer.** 1994. Extraordinarily low number of ribosomal RNA genes in *P. carinii*. J. Eukaryot. Microbiol. **41:**88S.
3. **Hacker, K. J., and B. M. Alberts.** 1994. The rapid dissociation of the T4 DNA polymerase holoenzyme when stopped by a DNA hairpin helix. A model for

polymerase release following the termination of each Okazaki fragment. J. Biol. Chem. **269:**24221–24228.

4. **Helweg-Larsen, J., C. H. Lee, S. Jin, J. Y. Hsueh, T. L. Benfield, J. Hansen, J. D. Lundgren, and B. Lundgren.** 2001. Clinical correlation of variations in the internal transcribed spacer regions of rRNA genes in *Pneumocystis carinii* f. sp. *hominis*. AIDS **15:**451–459.
5. **Itatani, C. A.** 1996. Ultrastructural morphology of intermediate forms and forms suggestive of conjugation in the life cycle of *Pneumocystis carinii*. J. Parasitol. **82:**163–171.
6. **Jensen, M. A., and N. Straus.** 1993. Effect of PCR conditions on the formation of heteroduplex and single-stranded DNA products in the amplification of bacterial ribosomal DNA spacer regions. PCR Methods Appl. **3:**186–194.
7. **Kopczynski, E. D., M. M. Bateson, and D. M. Ward.** 1994. Recognition of chimeric small-subunit ribosomal DNAs composed of genes from uncultivated microorganisms. Appl. Environ. Microbiol. **60:**746–748.
8. **Lee, C. H., J. Helweg-Larsen, X. Tang, S. Jin, B. Li, M. S. Bartlett, J. J. Lu, B. Lundgren, J. D. Lundgren, M. Olsson, S. B. Lucas, P. Roux, A. Cargnel, C. Atzori, O. Matos, and J. W. Smith.** 1998. Update on *Pneumocystis carinii* f. sp. *hominis* typing based on nucleotide sequence variations in internal transcribed spacer regions of rRNA genes. J. Clin. Microbiol. **36:**734–741.
9. **Liesack, W., S. Sela, H. Bercovier, C. Pitulle, and E. Stackebrandt.** 1991. Complete nucleotide sequence of the *Mycobacterium leprae* 23S and 5S rRNA genes plus flanking regions and their potential in designing diagnostic oligonucleotide probes. FEBS Lett. **281:**114–118.
10. **Lu, J. J., M. S. Bartlett, M. M. Shaw, S. F. Queener, J. W. Smith, M. Ortiz-Rivera, M. J. Leibowitz, and C. H. Lee.** 1994. Typing of *Pneumocystis carinii* strains that infect humans based on nucleotide sequence variations of internal transcribed spacers of rRNA genes. J. Clin. Microbiol. **32:**2904–2912.
11. **Matos, O., C. H. Lee, S. Jin, B. Li, M. C. Costa, L. Goncalves, and F. Antunes.** 2003. *Pneumocystis jiroveci* in Portuguese immunocompromised patients: association of specific ITS genotypes with treatment failure, bad clinical outcome and childhood. Infect. Genet. Evol. **3:**281–285.
12. **Matsumoto, Y., and Y. Yoshida.** 1984. Sporogony in *Pneumocystis carinii*: synaptonemal complexes and meiotic nuclear divisions observed in precysts. J. Protozool. **31:**420–428.
13. **Meyerhans, A., J. P. Vartanian, and S. Wain-Hobson.** 1990. DNA recombination during PCR. Nucleic Acids Res. **18:**1687–1691.
14. **Modrich, P.** 1987. DNA mismatch correction. Annu. Rev. Biochem. **56:**435–466.
15. **Nimri, L. F., I. N. Moura, L. Huang, C. del Rio, D. Rimland, J. S. Duchin, E. M. Dotson, and C. B. Beard.** 2002. Genetic diversity of *Pneumocystis carinii* f. sp. *hominis* based on variations in nucleotide sequences of internal transcribed spacers of rRNA genes. J. Clin. Microbiol. **40:**1146–1151.
16. **Pienaar, E., M. Theron, M. Nelson, and H. J. Viljoen.** 2006. A quantitative model of error accumulation during PCR amplification. Comput. Biol. Chem. **30:**102–111.
17. **Robberts, F. J., L. D. Liebowitz, and L. J. Chalkley.** 2004. Genotyping and coalescent phylogenetic analysis of *Pneumocystis jiroveci* from South Africa. J. Clin. Microbiol. **42:**1505–1510.
18. **Ruano, G., and K. K. Kidd.** 1992. Modeling of heteroduplex formation during PCR from mixtures of DNA templates. PCR Methods Appl. **2:**112–116.
19. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239:**487–491.
20. **Shuldiner, A. R., A. Nirula, and J. Roth.** 1989. Hybrid DNA artifact from PCR of closely related target sequences. Nucleic Acids Res. **17:**4409.
21. **Siripattanapipong, S., J. Worapong, M. Mungthin, S. Leelayoova, and P. Tan-ariya.** 2005. Genotypic study of *Pneumocystis jirovecii* in human immunodeficiency virus-positive patients in Thailand. J. Clin. Microbiol. **43:**2104–2110.
22. **Smulian, A. G., T. Sesterhenn, R. Tanaka, and M. T. Cushion.** 2001. The *ste3* pheromone receptor gene of *Pneumocystis carinii* is surrounded by a cluster of signal transduction genes. Genetics **157:**991–1002.
23. **Tanabe, K., N. Sakihama, A. Farnert, I. Rooth, A. Bjorkman, D. Walliker, and L. Ranford-Cartwright.** 2002. In vitro recombination during PCR of *Plasmodium falciparum* DNA: a potential pitfall in molecular population genetic analysis. Mol. Biochem. Parasitol. **122:**211–216.
24. **Tang, X., M. S. Bartlett, J. W. Smith, J. J. Lu, and C. H. Lee.** 1998. Determination of copy number of rRNA genes in *Pneumocystis carinii* f. sp. *hominis*. J. Clin. Microbiol. **36:**2491–2494.
25. **Totet, A., J. C. Pautard, C. Raccurt, P. Roux, and G. Nevez.** 2003. Genotypes at the internal transcribed spacers of the nuclear rRNA operon of *Pneumocystis jiroveci* in nonimmunosuppressed infants without severe pneumonia. J. Clin. Microbiol. **41:**1173–1180.
26. **Tsolaki, A. G., R. F. Miller, A. P. Underwood, S. Banerji, and A. E. Wakefield.** 1996. Genetic diversity at the internal transcribed spacer regions of the rRNA operon among isolates of *Pneumocystis carinii* from AIDS patients with recurrent pneumonia. J. Infect. Dis. **174:**141–156.
27. **von Wintzingerode, F., U. B. Gobel, and E. Stackebrandt.** 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol. Rev. **21:**213–229.
28. **Wang, G. C., and Y. Wang.** 1996. The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. Microbiology **142:**1107–1114.