Internal Transcribed Spacer Regions of rRNA Genes of *Pneumocystis carinii* from Monkeys

JOHN Y. C. HSUEH, ^1^ RUDOLF P. BOHM, JR., ^2^ PETER J. DIDIER, ^3^ XING TANG, ^1^ MARK E. LASBURY, ^1^ BAOZHENG LI, ^1^ SHAOLING JIN, ^1^ MARILYN S. BARTLETT, ^1^ JAMES W. SMITH, ^1^ AND CHAO-HUNG LEE^1^*

Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, Indiana 46202^1^; Division of Animal Resources, Yerkes Regional Primate Research Center, Atlanta, Georgia 30322^2^; and Unit of Clinical and Research Medicine, Tulane Regional Primate Research Center, Covington, Louisiana 70433^3^

Received 3 August 2000/Returned for modification 13 December 2000/Accepted 8 January 2001

Analysis of sequence variations among isolates of *Pneumocystis carinii* f. sp. *macacae* from 14 Indian rhesus monkeys (*Macaca mulatta*) at the internal transcribed spacer (ITS) regions of the nuclear rRNA gene was undertaken. Like those from *P. carinii* f. sp. *hominis*, the ITS sequences from various *P. carinii* f. sp. *macacae* isolates were not identical. Two major types of sequences were found. One type of sequence was shared by 13 isolates. These 13 sequences were homologous but not identical. Variations were found at 13 of the 180 positions in the ITS1 region and 28 of the 221 positions in the ITS2 region. These sequence variations were not random but exhibited definite patterns when the sequences were aligned. According to this sequence variation, ITS1 sequences were classified into three types and ITS2 sequences were classified into five types. The remaining specimen had ITS1 and ITS2 sequences substantially different from the others. Although some specimens had the same ITS1 or ITS2 sequence, all 14 samples exhibited a unique whole ITS sequence (ITS1 plus ITS2). The 5.8S rRNA gene sequences were also analyzed, and only two types of sequences that differ by only one base were found. Unlike *P. carinii* f. sp. *hominis* infections in humans, none of the monkey lung specimens examined in this study were found to be infected by more than one type of *P. carinii* f. sp. *macacae*. These results offer insights into the genetic differences between *P. carinii* organisms which infect distinct species.

Pneumocystis carinii is the major cause of pneumonia in AIDS patients and is one of the most common opportunistic pathogens in immunocompromised patients. *P. carinii* is also found in many different species of vertebrate animals. Although the morphologies of *P. carinii* from various animals are similar, nucleotide sequences of certain genes or genomic loci are variable. Examples of these genes include the major surface glycoproteins (22, 27, 28); beta-tubulin (4); alpha-tubulin (21); thymidylate synthase (18); arom (1); small (6) and large similar, nucleotide sequences of certain genes or genomic loci from different species of animals. The ITS sequences of *P. carinii* from both rats and humans have been determined (16). The ITS sequences of human *P. carinii* (*P. carinii* f. sp. *hominis*) have been found to vary among different isolates, and this sequence variation has been used to type *P. carinii* f. sp. *hominis* isolates (2, 7–14, 16, 17, 19, 24). In this study, we sequenced the ITS regions of *P. carinii* from 14 Indian rhesus monkeys (*Macaca mulatta*). It has been suggested that monkey *P. carinii* be named *P. carinii* f. sp. *macacae* (3). Like *P. carinii* f. sp. *hominis*, ITS sequence variations exist in different *P. carinii* f. sp. *macacae* isolates.

MATERIALS AND METHODS

Source of specimen. Fourteen lung tissues from 14 different SIV-infected Indian rhesus monkeys with *P. carinii* pneumonia (PCP) were obtained at necropsy. Diagnosis of PCP was based upon clinical signs including dyspnea, tachypnea, and development of interstitial pneumonia demonstrated by thoracic radiograph. All monkey lungs had moderate to severe interstitial infiltrates. *P. carinii* cysts were observed in methenamine-silver-nitrate-stained sections of all 14 lung tissue samples. These specimens were designated J637, C375, G762, G439, A349, G962, H468, J740, J308, H638, H423, I579, J129, and J203.

DNA isolation. Approximately 50 mm³ of each monkey lung tissue was homogenized and placed into a 50-ml conical tube containing 10 ml of Marmur’s solution (0.1 M NaCl, 0.04 M EDTA [pH 8.0]) and 5% sodium dodecyl sulfate (SDS) to lyse the cells. Protease K and RNase A were added to final concentrations of 100 µg/ml each to digest protein and RNA. After an overnight incubation at 37°C, a standard phenol-chloroform-isooamyl alcohol (25:24:1) extraction followed by a chloroform-isooamyl alcohol (24:1) extraction was performed to purify the DNA. The DNA in the aqueous phase of the extraction was precipitated with ethanol. The DNA pellet was washed with 70% ethanol, air dried, and resuspended in 250 µl of TE buffer (10 mM Tris·HCl, 1 mM EDTA [pH 8.0]).

PCR. The purified DNA was used as a template to amplify the ITS regions by nested PCR using primers 1724F2 (5'-AGTTGATCAAATTTGGTCATTTAG-3') and ITS2R (5'-TCTCGGACGAGGATCCTCGCC-3') for the first-step PCR and primers Fx (5'-TTCCGTAGGTGAACCTGCG-3') and Rm (5'-TGA TCTCCTGCTTATATGAGG-3') for the second-step PCR as described previously (12, 14, 15). The PCR mixture contained 200 ng of template DNA, PCR buffer (10 mM Tris·HCl [pH 8.3], 50 mM KCl, 0.001% gelatin), 20 pmol of each PCR primer, a 0.2 mM concentration of each deoxynucleoside triphosphate, 3 mM MgCl₂, and 1.5 U of Taq DNA polymerase in a total volume of 40 µl. For PCR with primers 1724F2 and ITS2R, the initial stage was a 3-min denaturation at 94°C; the second stage was 5 cycles of 1 min at 94°C, 1 min at 62°C, and 1 min at 72°C; the third stage was 35 cycles of 40 s at 94°C, 40 s at 60°C, and 1 min at 72°C. After amplification, the PCR products were subjected to electrophoresis on a 2% agarose gel and then stained with ethidium bromide. The amplified bands were excised from the gel and purified using a QIAEX II gel extraction kit (Qiagen). The purified DNA was sequenced using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) on an ABI 3700 automated sequencer (Applied Biosystems). The nucleotide sequences were compared with the National Center for Biotechnology Information database using the Basic Local Alignment Search Tool (Blast; http://www.ncbi.nlm.nih.gov/blast). The ITS sequences were aligned using Clustal W (17).
FIG. 1. Alignment of ITS sequences of *P. carinii* f. sp. *hominis*, *P. carinii* f. sp. *carinii*, and *P. carinii* f. sp. *macacae*. The DNA sequences were aligned using the program CLUSTAL W and then edited manually. The sequence of specimen J637, which was the first *P. carinii* f. sp. *macacae* ITS sequence determined in this study, is used as the reference. C375 is another monkey lung specimen containing *P. carinii* with ITS sequences that are very different from those of the other 13 specimens examined in this study. Bases that are the same as those of J637 are represented by asterisks, missing bases are denoted by hyphens, and those that are different from the reference sequence are indicated. The *P. carinii* f. sp. *hominis* ITS sequence used in this study is that of type Ee (14), and that of the rat *P. carinii* (P. carinii f. sp. carinii) sequence is as described previously (16).
72°C; and the final stage was a 2-min extension at 72°C. When the Fx-Rx primer set was used, primer annealing was done at 95°C and the MgCl₂ concentration was 2 mM in the PCR mix. All other PCR conditions were the same as described above. A positive specimen (bronchoalveolar lavage specimen obtained from a human patient with microscopically proven PCP and a negative control sample containing no template DNA were included in each run. The amplified products were electrophoresed on a 1.2% agarose gel. The expected amplified fragments were approximately 620 bp.

Cloning and sequencing. The PCR products were cloned into pCR2.1 (Invitrogen, Carlsbad, Calif.) according to the instructions provided by the manufacturer. The recombinant plasmid DNA used for nucleotide sequencing was isolated using the Wizard Plus SV kit (Promega, Madison, Wis.). Nucleotide sequencing was performed with the dRhodamine Terminator Cycle Sequencing kit (Perkin-Elmer, Foster City, Calif.), and the sequences were determined by the ABI 310 Genetic Analyzer (Perkin-Elmer).

Nucleotide sequence analysis. DNA sequences were aligned using the program CLUSTAL W (23) and edited manually. Phylogenetic analyses were performed using the maximum-likelihood analysis program DNAML of the PHYLIP software package (5) (version 3.57c; University of Washington, Seattle), and the phylogenetic trees were visualized using the Tree View program (Division of Environmental and Evolutionary Biology, Institute of Biomedical and Life Sciences, University of Glasgow, Scotland).

Nucleotide sequence accession numbers. Nucleotide sequences determined in this study have been assigned accession numbers as follows: ITS1 type A, AF288827; ITS1 type A1, AF288828; ITS1 type A2, AF288829; ITS1 type A3, AF288830; ITS1 type B, AF288831; ITS1 type B1, AF288832; ITS1 type C, AF288833; ITS1 type D, AF288834; ITS2 type a, AF288835; ITS2 type al, AF288836; ITS2 type a2, AF288837; ITS2 type a3, AF288838; ITS2 type b, AF288839; ITS2 type b1, AF288840; ITS2 type c, AF288841; ITS type c1, AF288842; ITS2 type d, AF288843; ITS2 type d1, AF288844; ITS2 type e, AF288845; ITS2 type f, AF288846; 5.8S type I, AF288847; and 5.8S type II, AF288848.

RESULTS

With primer sets 1724F2-ITS2R and Fx-Rx, a fragment of approximately 620 bp was amplified from specimen J637. The PCR product was cloned and sequenced. Four recombinant clones of the PCR product were sequenced. The sequences obtained from all four clones were found to be identical. The sequence was then compared with that of P. carinii f. sp. hominis type Ee (Fig. 1). The type Ee sequence was chosen because this study have been assigned accession numbers as follows: ITS1 type A, AF288827; ITS1 type A1, AF288828; ITS1 type A2, AF288829; ITS1 type A3, AF288830; ITS1 type B, AF288831; ITS1 type B1, AF288832; ITS1 type C, AF288833; ITS1 type D, AF288834; ITS2 type a, AF288835; ITS2 type al, AF288836; ITS2 type a2, AF288837; ITS2 type a3, AF288838; ITS2 type b, AF288839; ITS2 type b1, AF288840; ITS2 type c, AF288841; ITS type c1, AF288842; ITS2 type d, AF288843; ITS2 type d1, AF288844; ITS2 type e, AF288845; ITS2 type f, AF288846; 5.8S type I, AF288847; and 5.8S type II, AF288848.

To determine whether the sequences of the ITS regions were variable among different P. carinii f. sp. macacae isolates, 13 additional P. carinii-infected monkey lung specimens were examined. All 13 specimens generated the 620-bp PCR product. The PCR products were cloned and sequenced. Sequences of the extreme 3' end of the 18S rRNA genes of all 14 specimens were found to be identical, and only one base, the second base, of the first 14 bases of the 26S rRNA gene was found to be variable among the 14 sequences. Only one position was found to be variable within the 158-bp 5.8S rRNA gene. When these 14 sequences were aligned, 13 of them were found to be similar to each other, and one (specimen C375) was quite different (Fig. 1). Bases of these 13 sequences that were found to be variable were aligned as shown in Fig. 2, and a consensus sequence based on the majority of the sequences at each variable position was developed (Fig. 2, top line).

DISCUSSION

In this study, we have analyzed both ITS1 and ITS2 sequences of 14 P. carinii f. sp. macacae specimens. Similar to P. carinii f. sp. hominis, P. carinii f. sp. macacae isolates also exhibit nucleotide sequence variations in ITS regions. Excep for specimen C375 (Fig. 1), the sequence variations are not randomly distributed. They occur at certain positions. In the ITS1 region of the remaining 13 specimens, sequence variations are found at 13 of the 180 positions. Six (positions 125, 128, 130, 132, 179, and 189) of these 13 positions appear to be more critical. Based on sequence variations at these six posi-

![Table of nucleotide sequence variations](https://via.placeholder.com/150)

**FIG. 2.** Variable bases of the 13 homologous ITS sequences of P. carinii f. sp. macacae isolates. Locations of bases that exhibit sequence variation are shown above the sequence. The type of each sequence is indicated in the parentheses following the specimen identification number. The variation shown at the top is the consensus sequence derived from the majority of the sequence at each position. Bases that are the same as the consensus sequence are represented by asterisks, missing bases are denoted by hyphens, and those that are different from the reference sequence are indicated.
tions, the ITS1 sequences can be classified into three major types, designated A, B, and C (Fig. 2). The ITS1 sequence of specimen C375 is designated type D.

For type A, the bases at these six positions are A, G, missing, A, T, and A (AG-ATA), respectively. For type B, these positions are G, A, A, G, missing, and G (GAAG-G), respectively. For type C, these positions are G, G, missing, G, missing, and G (GG-G-G), respectively. Five specimens (J129, H423, J637, H468, and H638) belong to type A, and seven specimens belong to type B. Only one specimen (G962) is considered to be type C. Minor sequence variations are also found at other positions. At position 167, the A residue is replaced bya G in specimen J129. The same change is found at position 168 in specimen H423. Since these two specimens differ from the typical A type by only one base, they are designated types A1 (J129) and A2 (H423), respectively. Specimen H638 differs from the typical type A sequence by two bases and is classified as type A3. It is missing a base at position 68 and has an A instead of a G at position 270. In the type B isolates, specimens G762 and J308 are missing the two A residues at positions 180 and 181; otherwise, they are identical to typical type B. These two specimens are designated type B1.

Classification of ITS2 sequences is much more complicated. Sequence variations occur at 28 positions, and only three specimens (J129, H423, and J637) have the same ITS2 sequence. The sequences of the remaining 10 specimens are classified into five types (designated a, b, c, d, and e). Types a, b, and c are distinguished according to sequence variations at positions 569 to 571. Type a is missing the first two bases in this region. The sequence for type b is TAA and that for type c is TAT at this region. The type d sequence is quite distinct from the other types. It has a G residue at positions 441, 521, and 562 and a run of 3 Ts at positions 446 to 448. Type e is closer to type a except that it has two extra Ts at positions 446 and 447 and AT residues at positions 482 and 483. The ITS2 sequence of specimen C375 is designated type f.

There are also minor variations among type a isolates. Specimen G439, designated a1, has a T instead of an A at position 450; otherwise, it is identical to a typical type a sequence. Specimen J203 is missing bases at positions 389, 390, and 577 and is designated type a2. Specimen I579 has a T instead of a C residue at position 477 and is designated type a3. Among the two type b isolates, specimen G762 is missing a T residue at position 560 and has a G instead of an a residue at position 577. Specimen J308 is designated c1 because it differs from the typical type c (specimen G962) at position 540, where specimen J308 has an extra A. Three specimens (J129, H423, and J637) have identical sequences and are designated type d. Specimen H468 is designated type d1 because it differs from type d at positions 505, where the C residue is replaced by a T,
and at positions 558 to 560, where a run of three Ts is missing. This observation suggests that each monkey was infected by a different strain of P. carinii. Some isolates have the same ITS1 or ITS2 sequences, while other 13 isolates have similar sequences, but are different from those of the other 13 specimens. One isolate (C375) has very different ITS sequences. The other 13 isolates are quite different from those of the other 13 specimens.

Phylogenetic analyses were performed using the program DNAML in the PHYLIP software package (5). This program generates phylogenetic trees using the maximum-likelihood method. The results indicate that there are two major groups of P. carinii f. sp. macacae isolates. One group is represented by specimen C375; the other group consists of the other 13 specimens. Results are consistent with those obtained by visual examination. The phylogenetic analyses also reveal that C375 is more closely related than the other 13 isolates to P. carinii f. sp. hominis isotypes. Interestingly, these 13 isolates are closer to P. carinii f. sp. carinii (rat-derived P. carinii) than to P. carinii f. sp. hominis.

In summary, ITS sequences of 14 P. carinii f. sp. macacae isolates were examined. One isolate (C375) has very different ITS sequences. The other 13 isolates have similar sequences, with variations at certain nucleotide positions. These sequence variations exhibit certain patterns that allow classification of ITS1 into four types and classification of ITS2 sequences into six types. Some isolates have the same ITS1 or ITS2 sequences, but each isolate has a different combined ITS1 and ITS2 sequence. This observation suggests that each monkey was infected by a different strain of P. carinii f. sp. macacae. This is conceivable, since each monkey used in this study was caught in a different place. Whether P. carinii f. sp. macacae is transmissible among monkeys remains to be determined. The fact that each P. carinii f. sp. macacae isolate is different would make it feasible to develop a primate model to study P. carinii by transtracheal inoculation of P. carinii. It would be possible to distinguish the inoculated strain from the latent ones, if any, that become reactivated due to immunosuppression. At present, such studies as screening new drugs against P. carinii f. sp. hominis are mainly done using rodent models. The information derived from the rodent models may not be readily applicable to humans. A primate model would provide information more useful for the management of P. carinii infections in humans.
improving the sensitivity of progressive multiple sequence alignment through
sequence weighting, position-specific gap penalties and weight matrix

24. Tsolaki, A. G., R. F. Miller, A. P. Underwood, S. Banerji, and A. E. Wake-
field. 1996. Genetic diversity at the internal transcribed spacer regions of the
rRNA operon among isolates of *Pneumocystis carinii* from AIDS patients

25. Wakefield, A. E., C. C. Fritscher, A. S. Malin, L. Gwanzura, W. T. Hughes,
carinii* isolates from four geographical locations shown by analysis of mito-

26. Wakefield, A. E., C. C. Fritscher, A. S. Malin, L. Gwanzura, W. T. Hughes,
carinii* isolates from four geographical locations shown by analysis of mito-

27. Wright, T. W., P. J. Simpson-Haidaris, F. Gigliotti, A. G. Harmsen, and
C. G. Haidaris. 1994. Conserved sequence homology of cysteine-rich regions
in genes encoding glycoprotein A in *Pneumocystis carinii* derived from dif-

Cloning and characterization of a conserved region of human and rhesus
macaque *Pneumocystis carinii* gpA. Gene 167:185–189.