Mixed-Linker Polymerase Chain Reaction: A New Method for Rapid Fingerprinting of Isolates of the *Mycobacterium tuberculosis* Complex

WALTER H. HAAS,1,2* W. RAY BUTLER,1 CHARLES L. WOODLEY,1 AND JACK T. CRAWFORD1

1National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia 30333, and Department of General Pediatrics, University Children’s Hospital, Heidelberg University, Heidelberg 6900, Germany

Received 28 December 1992/Accepted 23 February 1993

Rapid recognition of multidrug-resistant strains of *Mycobacterium tuberculosis* is a desirable goal for treatment of patients and protection of health care workers. DNA fingerprints produced with the insertion sequence IS6110 generate restriction fragment length polymorphism (RFLP) patterns that reliably identify *M. tuberculosis* complex strains. This report describes a rapid technique for RFLP typing using the polymerase chain reaction. The method uses one primer specific for IS6110 and a second primer complementary to a linker ligated to the restricted genomic DNA. In one strand the linker contains uracil in place of thymidine, and specific amplification is obtained by elimination of this strand with uracil N-glycosylase. Mixed-linker fingerprinting clearly differentiated multidrug-resistant isolates from 12 outbreaks and unambiguously assigned them to 26 RFLP groups.

Tuberculosis has made a dramatic comeback in industrialized countries, in part because of the AIDS pandemic (17). In addition, the treatment of tuberculosis has been complicated by the appearance of drug-resistant strains spreading readily in the hospital setting (29). Since 1990 an increasing number of nosocomial outbreaks, especially among AIDS patients and including multidrug-resistant tuberculosis (MDRTB), have been investigated by the Centers for Disease Control (3). In several of these outbreaks, infection of health care workers demonstrated that MDRTB has become a threat to public health today (1, 20).

In more than 12 investigations of tuberculosis outbreaks, including 3 reported outbreaks of MDRTB (1, 6, 20), restriction fragment length polymorphism (RFLP) analysis using the repetitive element IS6110 has been shown to generate a strain-specific pattern consistent with epidemiologic data and antimicrobial susceptibility results. The RFLP pattern results from the variable location and copy number of this insertion sequence in the genomes of different strains. Although IS6110 is an active insertion element and is capable of transposition, the copy number and distribution of IS6110 have proved to be stable in a given outbreak situation (2, 13), resulting in the possibility of tracing transmission chains. However, to harvest sufficient genomic DNA for fingerprinting, the isolate has to be subcultured for 1 to 2 additional weeks after identification. In the setting of an outbreak, especially one of MDRTB, rapid identification of strains may enhance control efforts by detection and interruption of transmission chains.

The large amount of input DNA for genomic RFLP analysis and the need for more-sensitive techniques have led to the development of the polymerase chain reaction (PCR)-RFLP technique (22). In this method a specific region of the genome is amplified, and then restriction analysis of the PCR product is used to determine microheterogeneities of a single locus. In contrast, genomic RFLP analysis is based on the heterogeneous distribution of a genetic marker (e.g., a mobile genetic element) throughout the genome. This report describes a rapid new method to specifically amplify genomic RFLP fragments containing IS6110 by using PCR and its application to DNA fingerprinting of multidrug-resistant *Mycobacterium tuberculosis*.

**MATERIALS AND METHODS**

**Bacterial strains.** *M. tuberculosis* H37Rv was from laboratory stock, and *Mycobacterium bovis* 92-8122 was a recent patient isolate. Isolates of *M. tuberculosis* are listed with RFLP type and outbreak number in Table 1.

**RFLP analysis.** A 2- to 3-week-old subculture of the isolate in 7H9 broth was incubated overnight with cycloserine (1 mg/ml) at 37°C. The sedimented cells were transferred to a 1.5-ml microcentrifuge tube and incubated for 20 min at 80°C. The DNA was extracted with cetyltrimethylammonium bromide (Fisher Scientific, Pittsburgh, Pa.) as described previously (30). About 0.75 to 1.25 μg of genomic DNA was digested for 2 h with 10 U of *SspI* or *PvuII* (BRL Life Technologies Inc., Gaithersburg, Md.) and electrophoresed on a 1.0% agarose gel overnight. The gel was equilibrated in transfer solution (0.4 N NaOH, 1.0 M NaCl), and the restriction fragments were transferred by vacuum blotting to a Hybond-N+ nylon membrane (Amersham Corp., Arlington Heights, Ill.). Hybridization was carried out under stringent conditions with a 569-bp IS6110 PCR fragment (3' of the *SspI* restriction site), using the Amersham ECL direct labeling and detection kit (Amersham Corp.).

**Choice of the restriction enzyme.** A computer-supported restriction analysis of the IS6110 sequence (14, 25) was performed by using the PC/Gene restriction enzyme data bank and software (PC/Gene 6.50; Intelligenetics/Betagen Inc., Mountain View, Calif.).

**Oligonucleotides and linker construction.** All oligonucleotides were synthesized on a DNA synthesizer (model 381A; Applied Biosystems, Foster City, Calif.) by the Biotechnol-

* Corresponding author.
ology Core Facility, Centers for Disease Control and Prevention. Oligonucleotide sequences are displayed in Table 2. The mixed linker was constructed by annealing of the uracil-containing oligonucleotide IS04-92 with its complementary oligonucleotide IS08-92 containing thymidine. The oligonucleotides were mixed in a molar ratio of 1:1 in PCR buffer (10 mM Tris-HCl [pH 8.3], 50 mM KCl, 1.5 mM MgCl₂, 0.001% [wt/vol] gelatin). After the mixture was heated to 95°C, five cycles with 5 min at 60°C and 10 min at 45°C (melting temperature, 62°C) were carried out with a DNA thermal cycler (Perkin-Elmer, Branchburg, N.J.) to reduce the number of mismatched linker molecules. The annealing product was an unphosphorylated mixed linker with a GC overhang at the 3' end compatible with the HhaI-digested fragments.

**Mixed-linker PCR.** The general strategy for mixed-linker PCR is schematically displayed in Fig. 1. First (Fig. 1A), 100 pg of genomic DNA was digested for 1 h at 37°C with 10 U of HhaI. The same preparation was used for direct comparison of the traditional RFLP method and mixed-linker PCR. The sample was heated to 65°C for 5 min to dissociate spontaneously reannealed genomic fragments. One-tenth of the restriction product was transferred to a fresh microcentrifuge tube with a 1,000× molar excess of mixed linker, 0.5 mM ATP, and 1 U of T4 DNA ligase in 1× ligase reaction buffer (BRL Life Technologies Inc.). In the second step (Fig. 1B), ligation of the mixed linker to both ends of the restriction fragments was carried out overnight at 16°C. After addition of 10 U of HhaI, the samples were redigested for 30 min at 37°C. Next (Fig. 1C), half of the ligation reaction mixture was transferred to a fresh tube with 1 U of uracil-N-glycosylase (UNG) (Perkin-Elmer) in a total volume of 25 μl of PCR buffer (pH 8.3). The tubes were incubated for 20 min at 50°C and then for 15 min at 95°C (alkaline breakage of the apyrimidic sites) and then were cooled to 4°C. Finally (Fig. 1D), AmpliWax (Perkin-Elmer) was used to seal the reaction mixture, which was then overlaid with 75 μl of PCR reaction mix containing 0.5 U of Taq DNA polymerase (Perkin-Elmer), deoxynucleoside triphosphates (200 μM each), and oligonucleotide primers (IS04-92 and IS09-92, 1.0 μM each). The amplification profile consisted of an initial denaturation step at 94°C for 5 min, followed by 30 cycles with denaturation at 94°C for 2 min, primer annealing at 50°C for 2 min, and extension at 72°C for 2.5 min (5-s segment extension). For nested amplification, 1 μl of the PCR product was diluted 1:100, and 10 μl was amplified for 15 cycles with IS62-92 and IS09-92 as primers (Table 2). The temperature profile and reaction conditions remained the same. All experiments included negative controls processed together with the samples. Solutions were prepared in a DNA-free area separate from template preparation and amplification. Templates were added with positive-displacement pipettes to further decrease the risk of carryover contamination.

**Gel electrophoresis and hybridization.** An 18-μl aliquot of the PCR product was analyzed on a nondenaturating 8% polyacrylamide gel. The DNA was transferred to a Hybond-N+ nylon membrane (Amersham Corp.) by electroblotting. Fragments containing IS6110 were detected by hybridization of the membrane for 6 to 12 h at 42°C with the diagnostic oligonucleotide IS55-92 (Table 2). After stringent washing of the membrane (30 mM NaCl, 3 mM sodium citrate [pH 7.0], 0.1% sodium dodecyl sulfate), the 3'-fluorescein-dUTP-labeled oligonucleotide was developed by using the ECL 3' oligolabeling and detection system according to the recommendations of the manufacturer (Amersham Corp.).

**RESULTS**

The general strategy of the mixed-linker PCR approach is outlined in Fig. 1. Briefly, the genomic DNA is digested to completion with HhaI, and a double-stranded oligonucleotide linker is ligated to the ends of the restriction fragments. This linker is called a mixed linker because one of the two oligonucleotides constituting the linker was synthesized using uracil instead of thymidine. For efficient ligation, the mixed linker is added in a molar excess of 1:500 to 1:1,000. The uracil-containing strand is subsequently eliminated from the reaction by treatment with UNG (11) and heating to 95°C at alkaline pH. This step is followed by PCR with an IS6110-specific oligonucleotide primer and a linker primer. The sequence of the linker primer is identical to that of the thymidine-containing strand of the linker except that it lacks the two nucleotides complementary to the 3' overhang created by restriction with HhaI. In the first cycle of

---

**Table 1.** Correlation between mixed-linker fingerprints and PvuII RFLP type of outbreak strains of *M. tuberculosis*

<table>
<thead>
<tr>
<th>Outbreak no.</th>
<th>RFLP type</th>
<th>Strain no.</th>
<th>Lane</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fig. 3</td>
</tr>
<tr>
<td>001</td>
<td>0018114</td>
<td>91-8359</td>
<td>17</td>
</tr>
<tr>
<td>020</td>
<td>0202600</td>
<td>91-3370</td>
<td>6</td>
</tr>
<tr>
<td>021</td>
<td>0212072</td>
<td>92-8162</td>
<td>1</td>
</tr>
<tr>
<td>027</td>
<td>0212072</td>
<td>91-8334</td>
<td>12</td>
</tr>
<tr>
<td>028</td>
<td>0212072</td>
<td>91-3280</td>
<td>11</td>
</tr>
<tr>
<td>031</td>
<td>0318308</td>
<td>91-8309</td>
<td>5</td>
</tr>
<tr>
<td>036</td>
<td>0212072</td>
<td>92-8022</td>
<td>5</td>
</tr>
<tr>
<td>0368024</td>
<td>92-8039</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>0368027</td>
<td>92-8027</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>0368028</td>
<td>92-8028</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>0368033</td>
<td>92-8033</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>0368056</td>
<td>92-8056</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>0368057</td>
<td>92-8057</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>0368060</td>
<td>92-8031</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>052</td>
<td>0528274</td>
<td>92-8274</td>
<td>2</td>
</tr>
<tr>
<td>055</td>
<td>0212072a</td>
<td>92-8310</td>
<td>14</td>
</tr>
<tr>
<td>0558297</td>
<td>92-8297</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>0558300</td>
<td>92-8300</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.** Primer sequences used for linker construction and mixed-linker amplification

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS04-92</td>
<td>UGC GAG TCQ AGO UCA GUU CT</td>
</tr>
<tr>
<td>IS08-92</td>
<td>AGA ACT GTC AGC TCG GAC G</td>
</tr>
<tr>
<td>IS09-92</td>
<td>AGA ACT GTC AGC TCG CA</td>
</tr>
<tr>
<td>IS54-92</td>
<td>TCG ACT GCT TCA AGC ATG GCC G</td>
</tr>
<tr>
<td>IS62-92</td>
<td>AGC AGT GTC GCG AGC TC</td>
</tr>
<tr>
<td>IS55-92</td>
<td>TGT GAT CTG AGA CCT GAG C</td>
</tr>
</tbody>
</table>
A restriction

B Ligation

C uracil N-glycosylase

D PCR: first cycle

IS6110

mixed-linker

linker primer

1 2

IS6110 specific primer

FIG. 1. Flowchart of the general strategy for mixed-linker PCR. (A) Restriction of genomic DNA; (B) ligation of the synthetic linker to the restriction fragments; (C) treatment with UNG; (D) specific regeneration of the target sequence for the linker primer during the first PCR cycle. Detailed descriptions of the individual steps are in the text.
amplification, only restriction fragments containing the IS6110 sequence can serve as templates for the IS6110 oligonucleotide, and the annealing target for the linker oligonucleotide is regenerated by extension of these specifically primed sequences. After 30 cycles, the PCR products are reamplified in a seminested amplification with a second primer specific for IS6110. The amplified HhaI RFLP pattern could be analyzed directly by gel electrophoresis of the PCR products.

The mixed linker was constructed by annealing of two oligonucleotides to form a 3' overhang complementary to the HhaI-digested fragments, but without reconstitution of the HhaI restriction site after ligation. Treatment of the mixed linker with UNG demonstrated complete elimination of 100 ng (15 × 10^-12 mol) of the uracil-containing oligonucleotide after 5 min. After alkaline cleavage of the abasic sites, the short linker fragments did not participate in further reactions, thus eliminating the need for purification. Amplification without UNG treatment did not produce a reproducible RFLP pattern.

Genomic H37Rv DNA (100 µg/ml) was diluted stepwise to 1 fg/µl (less than one genome of M. tuberculosis), and 1 µl of each dilution was subjected to DNA fingerprinting by mixed-linker PCR to assess the sensitivity of this method. Figure 2 shows the polyacrylamide gel electrophoresis analysis of the PCR products after nesting, demonstrating an almost complete IS6110 fingerprint for as little as 1 pg of genomic DNA. This corresponds to 10^-6 x the amount of DNA needed for RFLP analysis without amplification.

Twelve M. tuberculosis strains isolated from different outbreaks of MDRTB and one isolate of M. bovis were analyzed by mixed-linker PCR, and the results were correlated to traditional RFLP analysis to demonstrate the specificity and reproducibility of this method (Table 1). The amplified fragments were visualized in the gel by ethidium bromide staining and verified to be specific for IS6110 by hybridization with the diagnostic oligonucleotide IS55-92 (Fig. 3). Four isolates from the first outbreak with the same PvuII RFLP pattern (0212072) exhibited identical fingerprints by mixed-linker PCR (Fig. 3, lanes 1 through 4). Two isolates from the second outbreak, with the same genomic PvuII RFLP pattern as the isolates of the first outbreak, were also indistinguishable by our method (Fig. 3, lanes 5 and 6). Two consecutive isolates from the same patient (Fig. 3, lanes 7 and 8), with PvuII RFLP pattern 0368024, had identical fingerprints with both methods. Lanes 9 through 11 of Fig. 3 show isolates from different patients, one exhibiting the same pattern as the isolates in lanes 7 and 8 and the other two exhibiting fewer copies of IS6110 by PCR as well as by traditional RFLP fingerprinting. Lanes 12 and 13 of Fig. 3 show patient isolates of M. tuberculosis and M. bovis, each with only one copy of IS6110 by RFLP analysis. In this example M. bovis could be identified by amplification of a 257-bp fragment. This fragment size corresponded to the distance from the nested IS6110 primer to the next HhaI restriction site (corrected by primer length) in the 3' flanking sequence in M. bovis (9).

To determine if fingerprinting of M. tuberculosis isolates by mixed-linker PCR could replace the traditional RFLP analysis, we examined 24 isolates from 10 different tuberculosis outbreaks. The mixed-linker fingerprints of 18 isolates from eight outbreaks grouped according to the outbreak and PvuII RFLP type are displayed in Fig. 4. For all 24 isolates, differentiation and grouping of the fingerprints by mixed-linker PCR were in full agreement with the RFLP results (Table 1). Most remarkably, the mixed-linker fingerprint of the isolate in lane 14 of Fig. 4 differed only by one extra band (arrow) from the pattern in lanes 11 through 13. A similar result was obtained by RFLP analysis of these isolates; they exhibited identical patterns except for one additional band for the isolate in lane 14 of Fig. 4 (fingerprint type 0212072a [Table 1]).
allowed efficient ligation of the mixed linker. *HhaI* was found to cut IS6110 157 bp from the 3' end and generated IS6110-containing fragments that are primarily 200 to 600 bp, with a maximum size of ~1,500 bp. This size range allowed efficient amplification of up to 20 restriction fragments containing IS6110 sequences. In addition, the use of a frequently cutting enzyme provided better resolution of IS6110 elements located close together in the genomic DNA than did restriction with *PvuII*, which was recently chosen for standardization of the traditional RFLP method (28).

Because most of the amplified restriction fragments are smaller than 1 kb, a rapid mechanical lysis of the mycobacterial cells can be used for DNA extraction for mixed-linker PCR without loss of specificity. Mechanical lysis of cells causes considerable shearing of the genomic DNA, whereas high-molecular-weight DNA is necessary for *PvuII* RFLP analysis. A comparison of mechanical lysis with DNA extracted by the cetyltrimethylammonium bromide procedure resulted in identical RFLP patterns (unpublished data). The fingerprint obtained by the mixed-linker technique sometimes showed a higher number of fragments containing the IS6110 sequence than did the hybridization pattern obtained by the traditional *PvuII* RFLP. This result could be explained by the higher resolution of the *HhaI* RFLP compared with that of *PvuII* (see above). Furthermore, amplification of the first PCR product may continue throughout the nested part because of carryover of primers from the first amplification. This mechanism was supported by the observation of a shadow band in the *M. bovis* fingerprint that differed only by 20 to 30 bp in size (data not shown). This difference was in good agreement with the distance of 27 bp between the two nested specific primers.

The main advantage of mixed-linker PCR fingerprinting over the traditional RFLP method is its independence from mycobacterial growth. Traditional DNA fingerprinting requires a 2- to 3-week subculture of the isolate or heavy growth on the original slant for IS6110 RFLP analysis (28). The high sensitivity of mixed-linker PCR allowed generation of an almost complete fingerprint with as little as 1 pg of genomic DNA, which corresponds to about 300 genomes of *H37Rv* (2.5 × 10^9 Da per haploid genome) (Fig. 2). One colony of the primary isolate on a Löffenstein-Jensen slant contains about 10^9 mycobacteria; therefore, a fingerprint could be obtained directly from a single colony of the primary isolate without further culture. IS6110 mixed-linker fingerprinting of single colonies could be used to investigate the frequency of simultaneous infections with more than one strain of *M. tuberculosis*.

The rapid in vitro amplification of RFLP fragments not only allows rapid analysis of isolates but also renders mixed-linker fingerprinting applicable to nonviable cells. This is important for outbreak investigations in which the original isolates fail to grow because they have been stored at room temperature or over an extended period of time. For the threat of MDRTB, this feature can also increase biosafety significantly by heat killing the organisms before they are sent to a central laboratory for fingerprinting.

It has been shown that the IS6110 RFLP pattern is independent of drug resistance, and acquisition of drug resistance does not change the fingerprint (19). Thus, in a general population, a multidrug-resistant strain and its drug-sensitive parent strain may coexist and could not be distinguished by RFLP typing. However, in an outbreak situation, a specific RFLP type should correlate well with the drug resistance pattern of the isolates, and rapid typing of the isolate might be used to design the treatment regimen.

**DISCUSSION**

Identification of *M. tuberculosis* strains by DNA fingerprinting has been proved to be an important tool for tuberculosis control (1, 4, 6, 20). We developed a rapid, highly sensitive and specific method for typing strains of the *M. tuberculosis* complex that is based on the IS6110 RFLP by using PCR. The key feature of our method, named mixed-linker PCR, is the ability to amplify multiple restriction fragments containing IS6110 sequences and variable sequences adjacent to the restriction site.

Three basic concepts have been developed for amplification of genomic DNA flanking known sequences by PCR: (i) inverse or circular PCR (18, 27), (ii) anchor PCR (7, 12, 16, 21), and (iii) single-site PCR (23). Similar to anchor PCR, mixed-linker PCR uses one primer specific for IS6110 and a second primer specific for a synthetic linker ligated to the restriction fragments. However, the mixed-linker approach allows the specific elimination of the target of the linker primer by treatment with UNG. Initial denaturation of the template in the hot start technique (5) prevented the unspecific re-creation of the complementary linker strand by filling in of the resulting 3' recessed ends by *Taq* polymerase. By using this approach, the efficient ligation of a double-stranded linker, as demonstrated for cloning purposes or whole genome PCR (10), could be combined with the high specificity of a single-stranded single-site PCR (15).

Parallel amplification of multiple DNA fragments depends on their size distribution, because laboratory experience shows that small fragments are more efficiently amplified by PCR than large fragments. Endonuclease *HhaI* has a 4-bp recognition sequence and produces a 2-bp 3' overhang that
Preliminary results show that the mixed-linker technique can be applied directly to fingerprinting of smear-positive clinical specimens. Therefore, early aggressive treatment could be targeted to the known drug-resistant strains, long before susceptibility data become available.

RFLP analysis has wide applications ranging from typing of bacterial strains (26) and identification of eukaryotic cell lines (24) to prenatal diagnosis of inherited diseases in human genetics (9). The sensitivity and reproducibility of the mixed-linker amplification demonstrated for fingerprinting of multidrug-resistant M. tuberculosis suggest that this concept could facilitate rapid RFLP analysis in other areas of biomedical research.

ACKNOWLEDGMENTS

This work was supported in part by research grant HA 1921/1-1 from the Deutsche Forschungsgemeinschaft.

REFERENCES


