

Identification of a repetitive sequence belonging to a PPE gene of *Mycobacterium tuberculosis* and its use in diagnosis of tuberculosis

Ranjana Srivastava,¹ D. Kumar,¹ M. N. Waskar,¹ Meera Sharma,²
V. M. Katoch³ and Brahm S. Srivastava¹

Correspondence

Ranjana Srivastava
drranjana@yahoo.com

¹Microbiology Division, Central Drug Research Institute, Lucknow 226001, India

²Medical Microbiology, Postgraduate Institute of Medical Education and Research, Chandigarh, India

³Central JALMA Institute for Leprosy, Agra, India

A repetitive sequence specific to *Mycobacterium tuberculosis* was isolated from a λ gt11 library of *M. tuberculosis* by DNA–DNA hybridization using genomic DNA of *M. tuberculosis* as probe followed by subtractive hybridization with a cocktail of other mycobacterial DNA. This led to identification of CD192, a 1291 bp fragment of *M. tuberculosis* containing repetitive sequences, which produced positive hybridization signals with *M. tuberculosis* DNA within 30 min. Nucleotide sequencing revealed the presence of several direct and inverted repeats within the 1291 bp fragment that belonged to a PPE family gene (Rv0355) of *M. tuberculosis*. The use of CD192 as a DNA probe for the identification of *M. tuberculosis* in culture and clinical samples was investigated. The 1291 bp sequence was present in *M. tuberculosis*, *Mycobacterium bovis* and *M. bovis* BCG, but was not present in many of the other mycobacterial strains tested, including *M. tuberculosis* H37Ra. More than 300 clinical isolates of *M. tuberculosis* were probed with CD192, and the presence of the 1291 bp sequence was observed in all the clinical strains, including those lacking IS6110. The sequence displayed RFLP among the clinical isolates. A PCR assay was developed which detected *M. tuberculosis* with 100% specificity from specimens of sputum, cerebrospinal fluid and pleural effusion from clinically diagnosed cases of tuberculosis.

Received 18 October 2005

Accepted 28 April 2006

INTRODUCTION

One-third of the world's population is estimated to be infected with *Mycobacterium tuberculosis*. Hence, the identification of infected individuals comes as a first priority in strategies for tuberculosis control. The culture of mycobacteria from clinical samples is considered to be the most reliable technique and provides for definitive diagnosis of tuberculosis. Although 100% specific, it takes 6–8 weeks, owing to the slow growth of the organisms and the need for further biochemical testing (Heifets & Good, 1994). On the other hand, nucleic acid probes coupled with amplification allow rapid and specific identification of *M. tuberculosis* in clinical samples (Pfyfer *et al.*, 1996). Some of the probes include insertion elements (IS6110, IS1081), genes for

immunodominant antigens (38 kDa, 85 protein complex, 30/32 kDa, MPB64) and ribosomal sequences (16S and 23S rRNA) (Eisenach, 1999). These DNA probes are genus and species specific and utilize a wide array of sequences from a single-copy sequence to repetitive DNA elements. The repetitive DNA elements offer the advantage of sensitivity, and hence are ideal as diagnostic markers for strains.

Several repetitive DNA elements/sequences have been reported in *M. tuberculosis*, such as IS6110, the direct repeat (DR) cluster, IS1081, major polymorphic tandem repeats (MPTRs), polymorphic GC-rich repetitive sequences (PGRSs) and IS-like elements (Poulet & Cole, 1995). IS6110 is typically present in multiple copies in *M. tuberculosis* and its species specificity, stability and RFLP make it a reliable diagnostic and epidemiological tool to fingerprint *M. tuberculosis* strains (Eisenach *et al.*, 1990; Kolk *et al.*, 1992). However, the discovery of several *M. tuberculosis* strains with one or no copy of IS6110 (Sahadevan *et al.*, 1995; Van Soolingen *et al.*, 1993) has shown that the repertoire of *M. tuberculosis* strains present all over the world may not be identified using a single repetitive element or probe. The

Abbreviations: ATT, antitubercular treatment; CSF, cerebrospinal fluid; DR, direct repeat; IAC, internal amplification control; MPTR, major polymorphic tandem repeat; NPV, negative predictive value; PGRS, polymorphic GC-rich repetitive sequence; PPV, positive predictive value; TBM, tuberculous meningitis.

The GenBank/EMBL/DBJ accession number for the CD192 sequence of *M. tuberculosis* is 810514.

search for novel DNA probes for *M. tuberculosis* remains a constant requirement.

The present investigation describes the identification of a repetitive sequence belonging to a PPE family gene of *M. tuberculosis* and its use in the molecular diagnosis of tuberculosis. The repetitive sequence was found useful in the identification of *M. tuberculosis* from culture and from clinical specimens of sputum, cerebrospinal fluid (CSF) and pleural effusion.

METHODS

Bacterial strains and clinical specimens. All mycobacterial and non-mycobacterial strains were obtained from the American Type Culture Collection and the Central JALMA Institute for Leprosy, Agra, India. Confirmed clinical isolates of *M. tuberculosis* were obtained from the All India Institute of Medical Sciences, New Delhi, the Postgraduate Institute of Medical Education and Research, Chandigarh, and the King George Medical College, Lucknow. Four clinical isolates of *M. tuberculosis* lacking IS6110 were obtained from the Tuberculosis Research Centre, Chennai. Clinical specimens (sputum, CSF and pleural effusion) were obtained from the King George Medical College, the Command Hospital and the Tuberculosis Hospital, all in Lucknow, and from the Ganesh Shanker Vidyarthi Medical College, Kanpur. Each specimen was streaked for culture on LJ slants and grown at 37 °C for isolation of DNA. To isolate DNA, the clinical materials were transferred into microcentrifuge tubes and incubated for 10 min at 80 °C to inactivate mycobacteria. Respiratory specimens, including sputum and pleural fluid, were digested and decontaminated by *N*-acetyl-L-cysteine (NALC)/4% NaOH and concentrated by centrifugation at 3000 g for 15 min. The sediments were resuspended in 50 ml phosphate buffer and recentrifuged at 3000 g for 15 min (Kent & Kubica, 1985). The CSF specimens were directly suspended in 50 ml phosphate buffer and centrifuged as described above. Part of the sample was streaked for culture on LJ slant and the rest of the sample was used for DNA extraction. The identification of *M. tuberculosis* and other non-tubercular mycobacteria was done by biochemical tests (Koneman *et al.*, 1997) and 16S rDNA sequencing (Kirschner *et al.*, 1993).

Extraction of genomic DNA from culture and clinical samples. DNA from mycobacterial and non-mycobacterial strains was isolated by the method of Marmur (1961), with slight modification for mycobacterial DNA as described by Connell & Ollar (1999). DNA from mycobacteria in clinical samples was isolated by the method described by Miyazaki *et al.* (1993).

DIG probe labelling and detection. Genomic DNA was passed through a 25 gauge needle 8–10 times to shear DNA. Sheared DNA was denatured in a boiling-water bath for 10 min at 95 °C followed by chilling in ice. Denatured DNA was labelled by the non-radioactive DIG DNA labelling and detection system (Boehringer Mannheim). The reaction was incubated at 37 °C for 20 h. Hybridized probes were immunodetected with anti-DIG-alkaline phosphatase and then visualized as blue spots with nitroblue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) substrates. Hybridization protocols described by Sambrook *et al.* (1989) were used.

Construction of genomic library and screening. A recombinant expression library of *M. tuberculosis* H37Rv DNA was constructed in the λ gt11 vector, essentially as described by Young *et al.* (1985). All DNA manipulations were carried out by standard protocols (Sambrook *et al.*, 1989). Briefly, genomic DNA was isolated from *M. tuberculosis* H37Rv, sheared to a mean length of 5 kb with serial passage through a 25 gauge needle, and treated with *Eco*RI methylase

and *S*-adenosyl methionine to methylate the *Eco*RI sites. The DNA ends were then made flush with T4 DNA polymerase, and the DNA was purified. *Eco*RI linkers were added to the flush ends of the genomic DNA, and *Eco*RI sites were generated by digestion with *Eco*RI restriction enzyme. The *Eco*RI fragments thus generated were ligated in a small volume with *Eco*RI-digested dephosphorylated λ gt11 DNA, and the library was packaged and amplified on *Escherichia coli* Y1088 at 42 °C, as described earlier (Young *et al.*, 1985). The library had a titre of 1×10^{10} p.f.u. ml⁻¹ and contained approximately 60% recombinants with a mean size of 4 kb.

The genomic library thus constructed was screened in two steps to select out *M. tuberculosis*-specific DNA sequences. The library was arrayed on a lawn of *E. coli* Y1088. The phage plaques were blotted onto a nitrocellulose membrane and probed first with a mixture of DIG-labelled genomic DNA of a clinical isolate of *M. tuberculosis* and the H37Rv strain. The plaques that produced hybridization signals intensely and rapidly within 30 min were picked and rescreened three times using the same genomic DNA probe. Finally, after tertiary screening, the plaques producing positive hybridization signals within 30 min were selected and subjected to subtractive screening in which the selected plaques were screened with a cocktail of mycobacterial DNA other than that of *M. tuberculosis*. This cocktail consisted of DIG-labelled genomic DNA of *Mycobacterium avium-intracellulare*, *Mycobacterium fortuitum*, *Mycobacterium chelonae*, *Mycobacterium smegmatis* and *Mycobacterium phlei*, mixed in equal concentration. After subtractive screening, the plaques hybridizing to *M. tuberculosis* but not to the cocktail of mycobacteria were chosen for further analysis.

PCR assay. The primers CD1 (5'-tggtgccgccgaagtcatta-3') and CD2 (5'-gcagtggaaacatcgagat-3') were designed from the 1291 bp CD192 fragment (GenBank accession no. 810514) using DNASTAR and OMIGA softwares. CD1 and CD2 primers amplified a 777 bp fragment. The PCR was performed in a 50 μ l reaction volume containing 15 mM MgCl₂, 5 μ l 10 \times PCR buffer (Promega; containing 100 mM Tris/HCl, pH 9.0 (25 °C), 500 mM KCl, 1% Triton X-100), 25 pmol each primer, 200 μ M each deoxynucleotide triphosphate, and 1.5 U *Taq* DNA polymerase (Promega). The temperature regimen consisted of a denaturation cycle of 94 °C for 5 min followed by 96 °C for 1 min, 60 °C for 1 min and 72 °C for 1 min, for 30 cycles. An elongation step of 72 °C for 5 min ended the PCR. The amplified product was electrophoresed on a 1.2% agarose gel in 1 \times Tris/borate/EDTA (TBE) running buffer, stained in ethidium bromide and photographed under UV light. The primers T4 (5'-ctcgcgagcgtaggcgtcgg-3') and T5 (5'-ctgtccagcgcgctcgg-3') were used for the detection of IS6110, which amplified a 123 bp product (Eisenach *et al.*, 1993). The PCR products were purified from the gel by a gel extraction kit (Qiagen) and the nucleotide sequence was determined by automated sequencing on an ABI PRISM 310 Genetic analyser (Perkin Elmer Applied Biosystems). The sequence analysis was done using the DNASTAR program.

Southern hybridization. For Southern hybridization, the DNA was digested with appropriate restriction enzymes and electrophoresed through an agarose gel. After the electrophoresis was complete, the gel was stained with ethidium bromide and photographed. The DNA from the agarose gel was transferred onto a nitrocellulose membrane by the capillary transfer method after denaturation and neutralization of DNA within the agarose gel. After the transfer, the membrane was air-dried and baked at 80 °C under vacuum for 2 h. The Southern blot after prehybridization at 42 °C for 2–3 h was hybridized with DIG-labelled probe at 42 °C for 8–16 h with continuous shaking, followed by detection. Hybridization conditions were essentially as described by Sambrook *et al.* (1989). DIG labelling and detection were done as described in the manufacturer's protocol (Boehringer Mannheim).

RFLP analysis. The genomic DNA from clinical isolates of *M. tuberculosis* and non-tubercular mycobacteria was digested with the *StuI* restriction enzyme. The restricted DNA fragments were electrophoresed through an agarose gel (0.8%, 0.5 × TBE buffer at ~1 V cm⁻¹). The DNA in the gel was denatured, and after neutralization was transferred onto a nitrocellulose membrane and hybridized with denatured DIG-labelled 1291 bp fragment.

Calculation of sensitivity and specificity. The sensitivity of the test was calculated as $[TP/(TP + FN)] \times 100\%$, the specificity was calculated as $[TN/(TN + FP)] \times 100\%$; the positive predictive value (PPV) was calculated as $[TP/(TP + FP)] \times 100\%$, and the negative predictive value (NPV) was calculated as $[TN/(TN + FN)] \times 100\%$, where TP is true positive, FN is false negative, TN is true negative and FP is false positive.

RESULTS AND DISCUSSION

Screening of the λ gt11 *M. tuberculosis* library with genomic DNA of *M. tuberculosis* as probe

The genomic λ expression library was plated to $\sim 6 \times 10^3$ p.f.u. per plate on *E. coli* Y1088. Approximately 60 000 recombinant plaques were screened, which represented the whole genomic library of *M. tuberculosis*. Fourteen plaques were selected as rapidly hybridizing clones after three successive screenings with DIG-labelled denatured chromosomal DNA of *M. tuberculosis* as probe. These 14 plaques produced hybridization signals within 30 min of detection. The clones were then subjected to subtractive screening with denatured chromosomal DNA from mycobacterial strains other than *M. tuberculosis* by DNA hybridization. This resulted in the elimination of six and the selection of eight recombinant plaques which hybridized with *M. tuberculosis* but not with denatured genomic DNA from *M. avium-intracellulare*, *M. fortuitum*, *M. chelonae*, *M. smegmatis* and *M. phlei*.

All eight clones contained inserts which were of different sizes. One of the clones, referred to as C8, produced hybridization signals within 10 min of detection. C8 contained an insert of 4.2 kb. The rapid-signal-producing sequences were mapped within a 1291 bp *StuI*-*StuI* fragment within the 4.2 kb insert (Srivastava *et al.*, 2000). The 1291 bp fragment was designated CD192.

Hybridization of CD192 with different mycobacterial strains and other pathogens

The genomic DNA from different mycobacteria and other pathogens was cleaved with the restriction enzyme *StuI*, blotted onto nitrocellulose paper and hybridized with DIG-labelled CD192 by Southern as well as dot-blot hybridization. CD192 did not hybridize to DNA of different mycobacterial and non-mycobacterial strains, which included *M. smegmatis*, *M. phlei*, *M. fortuitum*, *M. chelonae*, *Mycobacterium flavescens*, *Mycobacterium triviale*, *Mycobacterium duvali*, *Mycobacterium marinum*, *Mycobacterium gordonae*, *Mycobacterium kansasii*, *Mycobacterium avium*, *Mycobacterium intracellulare*, *Mycobacterium scrofulaceum*, *Mycobacterium xenopi*, *Mycobacterium aurum*, *Mycobacterium*

microti and *Mycobacterium szulgai*. Similarly, no hybridization was detected with *Salmonella typhimurium*, *Staphylococcus aureus*, *Proteus vulgaris*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, *Enterobacter alginii*, *Vibrio cholerae*, *Bacillus subtilis*, *E. coli*, salmon sperm DNA and human placental DNA.

Presence of CD192 in clinical isolates of *M. tuberculosis*

Mycobacterial cultures were raised from clinical samples and 366 cultures were obtained. Among these, 300 were identified as *M. tuberculosis*, while the rest were *M. flavescens* ($n=28$), *M. fortuitum* ($n=22$), *M. chelonae* ($n=8$), *M. triviale* ($n=6$) and *M. kansasii* ($n=2$). Genomic DNA from each isolate was digested with the restriction enzyme *StuI*, blotted onto a nitrocellulose membrane and hybridized with DIG-labelled CD192 probe DNA. Upon hybridization, the probe detected only *M. tuberculosis* and no other mycobacteria. Essentially, all clinical *M. tuberculosis* isolates consistently displayed three bands; however, in 30% of isolates, additional minor bands were obtained (Fig. 1). RFLP of *M. tuberculosis* clinical strains before and after prolonged passage in animals (mice) revealed identical patterns, suggesting that the element was quite stable *in vivo*. The 300 clinical isolates of *M. tuberculosis* were also analysed for the presence of IS6110 by PCR amplification of a 123 bp fragment (Eisenach *et al.*, 1993). Out of 300 isolates, 32 were devoid of IS6110 but contained CD192, as demonstrated by hybridization and later confirmed by PCR (data not shown).

Nucleotide sequence of the CD192 DNA fragment

The nucleotide sequence of CD192 was determined by the dideoxy chain-termination sequencing technique for both strands of CD192 (Biggins *et al.*, 1983). The sequence has

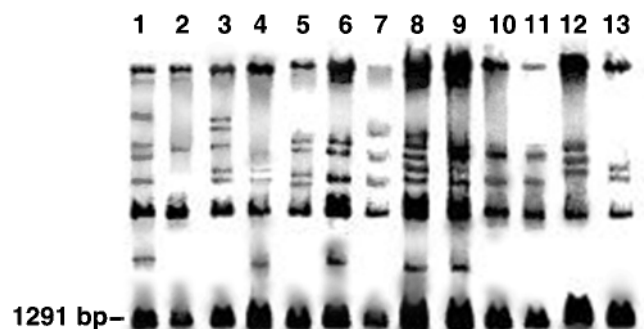


Fig. 1. RFLP analysis. The genomic DNA from *M. tuberculosis* H37Rv and clinical isolates of *M. tuberculosis* was isolated, digested with the restriction enzyme *StuI* and electrophoresed in a 0.8% agarose gel (1 × Tris/acetate/EDTA buffer). The fragments were electroblotted onto a nitrocellulose membrane and probed with DIG-labelled CD192 probe. Lane 2, *M. tuberculosis* H37Rv; lanes 1, 3–13, clinical isolates of *M. tuberculosis*.

been submitted to GenBank with accession number 810514. Computer-aided analysis of the nucleotide sequence and deduced protein sequence was performed using databases and programs provided by the National Institutes of Health (<http://www.ncbi.nlm.nih.gov/BLAST/BLAST>), as well as the programs of Chou & Fasman (1978) and Hopp & Woods (1981). The homology search of the CD192 sequence revealed it to be part of the PPE gene Rv0355c (424778–434677 bp) on the *M. tuberculosis* genome. However, a similar stretch on the *M. tuberculosis* genome is of 1286 bp (Cole *et al.*, 1998), whereas we have reported 1291 bp (Srivastava *et al.*, 2000). That the 1291 bp is part of Rv0355 was also confirmed by the sequencing of neighbouring sequences present in the original 4.2 kb insert of the C8 clone. Rv0355c is 9900 bp long and encodes a 327.03 kDa protein of no known function. This glycine/asparagine-rich protein is a member of the *M. tuberculosis* PPE family, and contains large numbers of DRs and MPTRs (Cole *et al.*, 1998). CD192 contained MPTRs (gccggtgtg or complement) characteristic of a PPE gene (Hermans *et al.*, 1992; Table 1).

Development of PCR-based assay

Different sets of primers were designed within the 1291 bp sequence of CD192 using the DNASTAR and OMIGA software packages and were subjected to PCR assay using target DNA of *M. tuberculosis*, atypical mycobacteria and non-mycobacterial species. The CD1 and CD2 primer pair was selected, which amplified 777 bp of DNA from *M. tuberculosis*, and this was visualized by agarose gel electrophoresis and DNA hybridization. A similar size fragment was amplified from *Mycobacterium bovis* and *M. bovis* BCG, but no amplification was observed with other mycobacterial and non-mycobacterial DNA, thus establishing the specificity of the primers for *M. tuberculosis*. An internal control of 461 bp was developed which could be amplified by the same CD1/CD2 primers for use as an internal amplification

control (IAC). CD192 has two *Sma*I sites at 227 and 407 bp, and one *Nhe*I site at 541 bp. The removal of the *Nhe*I–*Sma*I fragment resulted in the deletion of 316 bp from the 1291 bp fragment. When this deletion fragment served as template, CD1 and CD2 primers amplified a fragment of 461 bp which was easily distinguishable from the 777 bp fragment (Fig. 2). The deletion fragment could be seeded in the sample or used separately to rule out the presence of inhibitors in the sample in PCR reactions, which might otherwise lead to false-negative results.

Evaluation of CD1/CD2 primers on clinical samples

In this study, 200 clinical samples were selected to evaluate the CD192 probe, and were derived from patients who were clinically diagnosed cases of tuberculosis and responded to antitubercular treatment (ATT). The study protocol was as follows. Each collected sample was first streaked on LJ slants for culture and the rest was stored for PCR. These patients were immediately put on ATT. It is emphasized that the probe CD192 was evaluated on those samples whose donors were responding to ATT and that the samples were smear negative. Therefore, according to the study design, evaluation of the probe by PCR was carried out on clinically diagnosed culture-negative and culture-positive tuberculosis cases.

Sputum samples

Two hundred and fifty sputum samples were examined, which included samples from clinically diagnosed cases of tuberculosis ($n=105$), and controls from other non-mycobacterial respiratory diseases ($n=80$), mycobacteria other than tuberculosis (MOTT) ($n=20$) and healthy persons ($n=45$). Out of 105 samples, 65 were smear negative/culture positive and 40 were smear negative/culture negative. All culture-positive samples were PCR positive, and 30 out of 40 culture-negative samples were PCR

Table 1. Characteristics of DRs in the 1291 bp *Stu*I–*Stu*I DNA fragment

The US patent no. of the 1291 bp fragment is 6,114,514.

Serial no.	Length (bp)	Sequence	Position (bp)
1	26	CCG ATG TTG TTG TTG CCG GTG TTG GC	583, 613
2	10	TGA AGA AGC C	200, 1010, 1256
3	10	GCC CGT GTT G	312, 327, 357
4	10	TTG AGG ATG C	349, 379
5	9	CCG GTG TTG	208, 268, 298, 598
6	9	CCG ATG TTG	583, 613, 703, 733, 907
7	9	GTT GAA GTC	153, 333, 648, 864
8	9	GAA GAA GCC	201, 1011, 1161, 1257
9	9	GTT GCC GGT	264, 513, 594, 624
10	9	GCC GAC GTT	447, 672, 687, 807

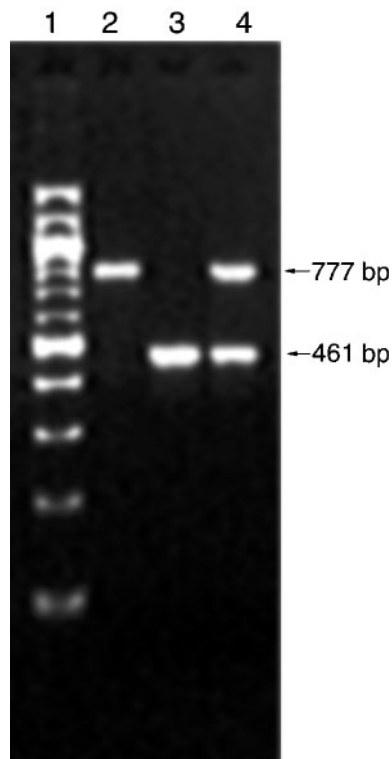


Fig. 2. PCR amplification of *M. tuberculosis* and IAC DNA by CD1 and CD2 primers. Lane 1, 100 bp ladder (MBI Fermenta); lane 2, 777 bp fragment amplified from *M. tuberculosis*; lane 3, 461 bp fragment of IAC DNA; lane 4, amplification of 777 and 461 bp fragments from sample containing both *M. tuberculosis* and IAC DNA.

positive. All controls were negative by PCR (Table 2). Hence, sensitivity and specificity for PCR were found to be 90.4 and 100 %, respectively. In a study reported earlier, the probe CD192 was evaluated with CD1 and CD2 primers on sputum samples, and 81 % sensitivity and 100 % specificity for PCR was reported (Prasad *et al.*, 2001).

Tuberculous meningitis (TBM)

One hundred and five CSF specimens were taken from 50 clinically diagnosed cases of TBM, and as a control, from 55 cases of non-tuberculous central nervous system (CNS) disease, including viral encephalitis and pyogenic meningitis. Patients with suspected TBM were those with fever and stiff neck for longer than 2 weeks. All CSF specimens from these patients had elevated leukocyte counts, polymorphonuclear pleocytosis, an elevated protein concentration ($>60 \text{ mg dl}^{-1}$) and CSF glucose concentration less than 60 % of the amount of glucose in the serum (Priyadarshi, 1999). The CSF samples were analysed for detection of *M. tuberculosis* DNA by PCR using CD1 and CD2 primers. A 777 bp fragment was amplified in 35 out of 50 CSF specimens from clinically diagnosed TBM cases. Out of these 35, five were positive by culture. All controls were PCR negative. Thus, the sensitivity of the probe for detection of *M. tuberculosis* in CSF was 70 %, with 100 % specificity (Table 2).

Pleural fluid specimens

The study was conducted on 45 cases of pleural effusion and 70 control subjects. Patients with pleural effusion presented fever, chest pain, dry cough and dyspnoea, with pleural fluid showing lymphocytosis, and the erythrocyte sedimentation rate (ESR) was raised. Test groups included clinically diagnosed cases of tuberculous pleural effusion ($n=45$); the age of patients varied between 12 and 71 years. Of 45 cases, 11 were culture positive and 34 PCR positive. All culture-positive specimens were PCR positive. In the control group of patients ($n=70$), none was culture or PCR positive. Thus the sensitivity of PCR was 77.7 %, with 100 % specificity (Table 2). In another isolated study, 22 specimens of pleural fluid were examined, out of which 15 were clinically suspected cases of pleural effusion and seven were of non-tuberculous aetiology (control). All 15 samples were culture negative, while 10 were PCR positive. In the control group ($n=7$), two were diagnosed cases of malignancy, being cases of adenocarcinoma confirmed by pleural fluid cytology; the

Table 2. Detection of *M. tuberculosis* in clinical specimens

Test	Tuberculosis samples*		Control samples		Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)
	Positive	Negative	Positive	Negative				
Sputum								
Culture	65	40	—	145	62.9	100	78.3	100
PCR	95	10	—	145	90.4	100	92.9	100
CSF								
Culture	5	45	—	55	10.0	100	55.0	100
PCR	35	15	—	55	70.0	100	78.5	100
Pleural fluid								
Culture	11	34	—	70	24.4	100	67.3	100
PCR	35	10	—	70	77.7	100	87.5	100

*Specimens from clinically diagnosed cases of tuberculosis; all responded to ATT.

remaining five were cases of pyogenic pleural effusion. One specimen of malignant pleural effusion gave a positive result by PCR, which is difficult to explain because the patient died shortly afterwards (Veerottam, 1999).

In this study, a new domain of repetitive sequence (CD192) has been identified within a PPE gene (Rv0355), as defined in the *M. tuberculosis* genome (Cole *et al.*, 1998). Rv0355 is a member of the PPE gene family, characterized by the presence of the motif proline-proline-glutamic acid (PPE) at positions 7 and 9 in a highly conserved N-terminal domain of 180 amino acids followed by a variable C-terminal region. The 68 members of the PPE family have been classified into three subfamilies; one of these subfamilies contains MPTR and PGRS sequences that are characterized by repeats of the motif AsnXGlyXGlyXAsnXGly. CD192 belongs to this family and contains several MPTRs (gccggtgttg), PGRSs (ttgccgccc), and several DRs, including a 26 bp sequence present in tandem copies separated by four bases. The MPTRs and PGRSs, originally described as non-coding repetitive sequences in the genome of *M. tuberculosis* (Hermans *et al.*, 1992; Poulet & Cole, 1995; Ross *et al.*, 1992), have been exploited as epidemiological tools to differentiate *M. tuberculosis* strains (Braden *et al.*, 1997; Ross *et al.*, 1992; Van Soolingen *et al.*, 1993).

Pulmonary tuberculosis is the most common presentation of tuberculosis, but it can affect any organ system, and clinical specimens other than sputum are known for the limited presence of bacilli. With the progression of disease, the presence of *M. tuberculosis*, the aetiological agent of tuberculosis, can be demonstrated microscopically in sputum, but the presence of bacilli in other forms of tuberculosis, such as TBM, pleural effusion or other extrapulmonary tuberculosis, is very difficult to detect because of low numbers of bacilli. The situation is aggravated by the low culture positivity of these samples. The high prevalence of infection in an otherwise clinically undiagnosed population is also a very serious problem. In the management of tuberculosis, an assay which is rapid, specific and can detect few bacilli is very important. Therefore, sequence-specific amplification by PCR becomes the method of choice. If the target which is amplified or the probe which is used for hybridization contains repetitive sequences, the probability of adequate sensitivity is enhanced several-fold.

CD192 appears to fulfil these criteria. It contains repetitive sequences that produce hybridization signals within 10–30 min and is specific to *M. tuberculosis*. To assay its potential in diagnosis, a PCR assay was developed, and primers were used to detect *M. tuberculosis* in pulmonary, CSF and pleural-fluid specimens. The latter two contain very few bacilli and are usually culture negative, and specific diagnosis is urgently needed, because if untreated, the disease follows a chronic course that causes significant morbidity and mortality.

There are few DNA probes available for the detection of *M. tuberculosis*. The commercially available PCR test for the

detection of *M. tuberculosis* complex marketed by Roche Diagnostic Systems amplifies a region of the 16S rDNA sequence that is genus specific and employs an *M. tuberculosis* complex-specific probe. The GenProbe amplified *M. tuberculosis* Direct (MTD) test employs the transcription-mediated amplification of 16S rRNA and is *M. tuberculosis* complex specific (Jonas *et al.*, 1993; Scarparo *et al.*, 2000). IS6110 has been used as a reliable, sensitive marker in the diagnosis of tuberculosis due to its high copy number, but has the limitation that it is not found in all *M. tuberculosis* clinical isolates (Sahadevan *et al.*, 1995). CD192 demonstrated good sensitivity and specificity in sputum, CSF and pleural effusion, and was present in *M. tuberculosis* strains lacking IS6110. Thus, CD192 may be recommended for use in the identification of *M. tuberculosis*. The results of this investigation are based on long-term evaluation of pulmonary (sputum) and extrapulmonary tuberculosis patients, and reports good sensitivity and 100% specificity with pulmonary and extrapulmonary samples.

ACKNOWLEDGEMENTS

Thanks are due to the Director, Central Drug Research Institute, for facilities, the Department of Biotechnology, India, for partial support and evaluation of probe, and to Professor P. N. Tandon, because without his constructive criticism and initiative, the potential of CD192 in diagnosis would not have been established. Thanks are also due to the King George Medical College, the Sanjay Gandhi Post Graduate Institute of Medical Sciences, the Command Hospital and the Tuberculosis Hospital, Lucknow, and the Ganesh Shanker Vidyarthi Medical College, Kanpur, for providing the clinical specimens and participating in the evaluation studies. This is communication number 6449 of this Institute.

REFERENCES

- Biggins, M. D., Gibson, T. J. & Hong, G. F. (1983). Buffer gradient gels and 35S label as an aid to rapid DNA sequence determination. *Proc Natl Acad Sci U S A* **80**, 3963–3965.
- Braden, C. R., Templeton, G. L., Cave, M. D. & 7 other authors (1997). Interpretation of restriction fragment length polymorphism. Analysis of *Mycobacterium tuberculosis* isolates from a state with a large rural population. *J Infect Dis* **175**, 1446–1452.
- Chou, P. Y. & Fasman, G. D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol* **47**, 45–148.
- Cole, S. T., Brosch, R., Parkhill, J. & 39 other authors (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544.
- Connell, N. D. & Ollar, R. A. (1999). The mycobacterial and nocardial genomes and their isolation and manipulation. In *Molecular Mycobacteriology: Techniques and Clinical Applications*, pp. 87–107. Edited by R. A. Ollar & N. D. Connell. New York: Marcel Dekker.
- Eisenach, K. D. (1999). Molecular diagnostics. In *Mycobacteria, Molecular Biology and Virulence*, pp. 161–179. Edited by C. Ratledge & J. Dale. Oxford, UK: Blackwell Science.
- Eisenach, K. D., Cave, M. D., Bates, J. H. & Crawford, J. T. (1990). Polymerase chain reaction amplification of a repetitive DNA sequence specific for *Mycobacterium tuberculosis*. *J Infect Dis* **161**, 977–981.

- Eisenach, K. D., Cave, M. D. & Crawford, J. T. (1993). PCR detection of *Mycobacterium tuberculosis*. In *Diagnostic Molecular Microbiology: Principles and Applications*, pp. 191–196. Edited by D. H. Persing, T. F. Smith, F. C. Tenover & T. J. White. Washington, DC: American Society for Microbiology.
- Heifets, L. B. & Good, R. C. (1994). Current laboratory methods for the diagnosis of tuberculosis. In *Tuberculosis, Pathogenesis, Protection and Control*, pp. 85–109. Edited by B. R. Bloom. Washington DC: American Society for Microbiology.
- Hermans, P. W. M., Van Soolingen, D. & Van Embden, J. D. A. (1992). Characterization of a major polymorphic tandem repeat in *Mycobacterium tuberculosis* and its potential use in the epidemiology of *Mycobacterium kansasii* and *Mycobacterium goodii*. *J Bacteriol* **174**, 4157–4165.
- Hopp, T. P. & Woods, K. P. (1981). Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci U S A* **78**, 3824–3828.
- Jonas, V., Alden, M. J., Curry, J. I., Kamisango, K., Knott, C. A., Lankford, R., Wolfe, J. M. & Moore, D. F. (1993). Detection and identification of *Mycobacterium tuberculosis* directly from sputum sediments by amplification of rRNA. *J Clin Microbiol* **31**, 2410–2416.
- Kent, P. T. & Kubica, G. P. (1985). *Public Health Mycobacteriology: a Guide for the Level III Laboratory*. Atlanta, GA: Centers for Disease Control.
- Kirschner, P., Springer, B., Vogel, U., Meier, A., Wrede, A., Kiekenbeck, M., Bange, F. C. & Bottger, E. C. (1993). Genotypic identification of mycobacteria by nucleic acid sequence determination. Report of a 2-year experience in a clinical laboratory. *J Clin Microbiol* **31**, 2882–2889.
- Kolk, A. H. J., Schuitema, A. R. J., Kuijper, S., van Leeuwen, J., Hermans, P. W. M., van Embden, J. D. A. & Hartskeerl, R. A. (1992). Detection of *Mycobacterium tuberculosis* in clinical samples by using polymerase chain reaction and the non radioactive detection system. *J Clin Microbiol* **30**, 2567–2575.
- Koneman, E. W., Allen, S. D., Janda, W. M., Schreckenberger, P. C. & Winn, W. C., Jr (1997). Mycobacteria. In *Color Atlas and Textbook of Diagnostic Microbiology*, 5th edn, pp. 893–952. Edited by E. W. Koneman, S. D. Allen, W. M. Janda, P. C. Schreckenberger & W. C. Winn, Jr. Philadelphia, PA: Lippincott-Raven.
- Marmur, J. (1961). A procedure for the isolation of deoxyribonucleic acid from microorganisms. *J Mol Biol* **3**, 208–218.
- Miyazaki, Y., Koga, H., Kohno, S. & Kaku, M. (1993). Nested polymerase chain reaction for detection of *Mycobacterium tuberculosis* in clinical samples. *J Clin Microbiol* **31**, 2228–2232.
- Pfyfer, G. E., Kisling, P., Jahn, E. M. I., Martin, H., Salfinger, W. M. & Weber, R. (1996). Diagnostic performance of amplified *Mycobacterium tuberculosis* direct test with cerebrospinal fluid, other nonrespiratory and respiratory specimens. *J Clin Microbiol* **34**, 834–841.
- Poulet, S. & Cole, S. T. (1995). Repeated DNA sequences in mycobacteria. *Arch Microbiol* **163**, 79–86.
- Prasad, R., Lath, S. K., Mukerji, P. K., Agarwal, S. K. & Srivastava, R. (2001). Clinical utility of polymerase chain reaction in patients of pulmonary tuberculosis. *Indian J Tuberc* **48**, 135–138.
- Priyadarshi, B. P. (1999). *Evaluation of CSF by polymerase chain reaction analysis in diagnosis of tuberculous meningitis*. MD thesis, GSVM Medical College, Kanpur.
- Ross, B. C., Raios, K. & Dwyer, B. (1992). Molecular cloning of a highly repeated DNA element from *Mycobacterium tuberculosis* and its use as an epidemiological tool. *J Clin Microbiol* **30**, 942–946.
- Sahadevan, R., Narayanan, S., Paramasivam, C. N., Prabhakar, R. & Narayanan, P. R. (1995). Restriction fragment length polymorphism typing of clinical isolates of *Mycobacterium tuberculosis* from patients with pulmonary tuberculosis in Madras, India, by use of direct-repeat probe. *J Clin Microbiol* **33**, 3037–3039.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Scarparo, C., Picoli, P., Rigon, A., Ruggiero, G., Scagnelli, M. & Piersimoni, C. (2000). Comparison of enhanced *Mycobacterium tuberculosis* amplified direct test with COBAS amplicor *Mycobacterium tuberculosis* complex in respiratory and extrapulmonary specimens. *J Clin Microbiol* **38**, 1559–1562.
- Srivastava, R., Kumar, D. & Srivastava, B. S. (2000). *A Mycobacterium tuberculosis specific DNA fragment*. US Patent, # 6,114,514.
- Van Soolingen, D., Dehass, P. E. W., Hermans, P. W. M., Groenen, P. M. A. & Van Embden, J. D. A. (1993). Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of *Mycobacterium tuberculosis*. *J Clin Microbiol* **31**, 1987–1995.
- Veerottam, P. (1999). *Value of polymerase chain reaction in pleural fluid in the diagnosis of tuberculous pleural effusion*. MD Thesis, GSVM Medical College, Kanpur.
- Young, R. A., Bloom, B. R., Grossinsky, C. M., Ivany, J., Thomas, D. & Davis, R. W. (1985). Dissection of *Mycobacterium tuberculosis* antigens using recombinant DNA. *Proc Natl Acad Sci USA* **82**, 2583–2587.