IS990, a new species-specific insertion-sequence-related element of the *Mycobacterium tuberculosis* complex

J. Dziadek,¹ A. Sajduda,² J. W. Dale³ and J. J. McFadden³

Author for correspondence: J. Dziadek. Tel: +48 42 491228. Fax: +48 42 491633.
e-mail: jdziadek@wirus.cmiwpan.lodz.pl

¹ Centre for Microbiology and Virology, Polish Academy of Sciences, Lodowa 106, 93-232 Łódź, Poland

² Department of Genetics of Microorganisms, University of Łódź, Banacha 12/16, 93-237 Łódź, Poland

³ Molecular Microbiology Group, School of Biological Sciences, University of Surrey, Guildford GU2 5XH, UK

The structure and distribution of IS990, a new *Mycobacterium tuberculosis* DNA sequence with homology to characterized insertion sequences (ISs), were investigated. IS990 was related to IS elements of the IS3 family and was present as a single copy in all 21 investigated *M. tuberculosis* strains, two *Mycobacterium bovis* strains and two *M. bovis* BCG strains. The sequence appears to be specific for the *M. tuberculosis* complex. The element carries two frameshift mutations and appears to be defective.

Keywords: *Mycobacterium tuberculosis* complex, mobile elements, IS990, IS3 family

INTRODUCTION

Mycobacterial mobile DNA elements have been widely used as target sequences for DNA probe applications both in diagnosis of mycobacterial disease (Vary *et al.*, 1990; Thierry *et al.*, 1990a; Brisson-Noel *et al.*, 1991; Noordhoek *et al.*, 1995) and for the study of the epidemiology of infectious disease (RFLP typing) (Eisenach *et al.*, 1988; Green *et al.*, 1989; Hermans *et al.*, 1990; McAdam *et al.*, 1990; Collins & Stephens, 1991; Kunze *et al.*, 1991, 1992; Fomukong *et al.*, 1992; van Embden *et al.*, 1993; Yuen *et al.*, 1993; Sola *et al.*, 1997). Mycobacterial insertion sequence (IS) elements (Collins & Stephens, 1991; Kunze *et al.*, 1991; Moss *et al.*, 1992; Hernandez Perez *et al.*, 1994) have also been successfully exploited for insertional mutagenesis (Cirillo *et al.*, 1991; Guilhot *et al.*, 1994; Billman-Jacobe *et al.*, 1996) and to integrate foreign genes in mycobacterial genomes (England *et al.*, 1991; Dellagostin *et al.*, 1993). The most commonly used IS element for RFLP typing in *Mycobacterium tuberculosis* is IS6110, found in *M. tuberculosis* and in *Mycobacterium bovis* (Hermans *et al.*, 1990; McAdam *et al.*, 1990; Thierry *et al.*, 1990b). Strains lacking IS6110 have been found in South-East Asia (Yuen *et al.*, 1993) and recently in Europe (McFadden & Sajduda, 1996). Every investigated *M. tuberculosis* strain carries the IS1081 element (Collins & Stephens, 1991). However, this element has limited use for epidemio-

logical studies and it demonstrates very little polymorphism among strains.

We have characterized a new IS from *M. tuberculosis* that was identified as part of the *M. tuberculosis* Genome Sequencing Project. It is a member of the IS3 family and is related to IS6110.

METHODS

Mycobacterial isolates. *Mycobacterium* strains used in this study are listed in Table 1. Culturing of the mycobacteria and chromosomal DNA isolation were performed as described previously (Kunze *et al.*, 1991; protocol according to D. van Soolingen, P. E. W. de Haas, P. W. M. Hermans & J. D. A. van Embden, unpublished).

PCR. On the basis of the Y49 cosmid sequence (EMBL/GenBank/DDBJ accession no. Z73966; B. G. Barrell, M. A. Rajandream & S. V. Walsh; *Mycobacterium tuberculosis* sequencing project, Sanger Centre, Hinxton Hall, Hinxton, Cambridge; Philipp *et al.*, 1996) the oligonucleotides were synthesized using a DNA synthesizer (Applied Biosystems). Nucleotide sequences of the primers and sizes of the amplification products are listed in Table 2. Amplification reactions were performed with 50 µl volumes containing 0.25 U *Taq* polymerase (Boehringer Mannheim), 1.25 mmol each deoxynucleotide triphosphate, 0.2 µmol primers and 0.015 µg chromosomal DNA in the commercial buffer (Boehringer Mannheim). The 30 amplification cycles (1 min at 65 °C; 2 min at 72 °C; 1 min at 95 °C) were preceded by a 1 min denaturation at 95 °C and followed by 2 min annealing at 65 °C and 5 min extension at 72 °C in a Hybaid thermocycler. Amplification products were purified using KRISTAL Gelex KX-50 DNA Extraction kit (Cambridge Molecular Technologies), ex-

Abbreviation: IS, insertion sequence.

Table 1. Mycobacterial strains used in the present study

Strain	Description	Source*
<i>M. tuberculosis</i> H37RV	Type strain	MGL, University of Surrey, UK
<i>M. tuberculosis</i> 518, 325, 370, 399, 444, 396, 403, 421, 104, 105, 110, 121, 127, 129, 138, 141, 142, 146, 153, 154, 164	Clinical isolates	Institute of Tuberculosis and Lung Diseases, Warsaw, Poland
<i>M. bovis</i> AF71, 2332		MGL, University of Surrey, UK
<i>M. bovis</i> BCG	Type strain	Bilthoven, the Netherlands
<i>M. bovis</i> BCG	Type strain	Copenhagen, Denmark
<i>M. avium</i> 19.7	Clinical isolate	AIDS patient, Australia
<i>M. avium</i> 19.2		MGL, University of Surrey, UK
<i>M. avium</i> 19.4	Serotype 1	MGL, University of Surrey, UK
<i>M. avium</i> 19.9		Africa
<i>M. avium</i> 20.5, 20.7	Clinical isolates	AIDS patients, Australia
<i>M. avium</i> 20.16, 20.17		MGL, University of Surrey, UK
<i>M. avium</i> 20.25	Environmental strain	Water soil, USA
MAIS 16.26	Environmental strain	Water, LA, USA
MAIS 16.29		Mouse
<i>M. avium</i> 16.50	Clinical isolate	AIDS patient, Africa
<i>M. avium</i> 20.18		MGL, University of Surrey, UK
<i>M. intracellulare</i> 19.6	ATCC 13950	MGL, University of Surrey, UK
<i>M. simiae</i> 19.10	Clinical isolate	MGL, University of Surrey, UK
<i>M. simiae</i> 19.11	Clinical isolate	MGL, University of Surrey, UK
<i>M. lufu</i> 20.32	Environmental strain	Water, Zaire
<i>M. lufu</i> 20.33	Environmental strain	Water
<i>M. malmoense</i> 16.27	Environmental strain	Water, USA
<i>M. malmoense</i> 16.38	Clinical isolate	AIDS patient, Africa
<i>M. gordonae</i> 16.47	Non-AIDS	Europe
<i>M. fortuitum</i>		MGL, University of Surrey, UK
<i>M. scrofulaceum</i>		MGL, University of Surrey, UK
<i>M. paratuberculosis</i>		MGL, University of Surrey, UK
<i>M. vaccae</i>	Environmental strain	Łódź, Poland
<i>M. phlei</i>	Environmental strain	Łódź, Poland
<i>M. smegmatis</i>	Environmental strain	Łódź, Poland
<i>M. microti</i> 19422		Jacksonville, FL, USA
<i>M. africanum</i> 11605, 1400-30056		Bilthoven, the Netherlands

* MGL, Molecular Genetics Laboratory.

Table 2. Primers used for amplification of the inner fragments of cosmid Y49

Primer	Sequence (5'–3')	Size of PCR product
49A	GCTGGAAGCGGCCGACTTCACC	310 bp (49A–49B)
49B	TCACATAAGGCGTTGTCGAGGG	
49D	TGTGGTGGCCGGCGTGTCGCC	
49E	AAGGCGCTGCCGCGGATCCGC	248 bp (49D–49E)
49C	CACGGTGAAGACGATCATGCGG	450 bp (49C–49B)

aminated by gel electrophoresis and applied in sequencing analysis, as well as DNA probes for hybridization experiments.

Southern hybridization. DNA–DNA hybridization experiments were performed by use of the DIG DNA Labelling and Detection kit according to the instructions of the manufacturer (Boehringer Mannheim). All hybridization reactions were

carried out at 68 °C. Then, the membrane was washed twice for 5 min at room temperature with 2 × SSC (1 × SSC is 0.15 M sodium chloride + 0.015 M sodium citrate), 0.1 % (w/v) SDS and twice for 15 min at 68 °C in 0.1 × SSC, 0.1 % (w/v) SDS.

DNA sequencing and analysis. Sequences of double-stranded PCR products were determined using the DNA Sequencing

Fig. 1. Composition of IS990 belonging to the IS3 family. IS990 is bordered by 16 bp inverted repeats (IRs). RS, ribosome slippage region; X, mutations in ORFb. Primers used for amplification of IS990 fragments are shown as arrows and designated 49A–E. Nucleotide positions of the start and end of IS990 are correlated with the Y49 cosmid.

kit, Dye Terminator Cycle Sequencing, according to the protocol of the manufacturer (Perkin Elmer). Similarities with IS990 were found in computer databases using GCG software (Oxford Molecular Group).

RESULTS

Analysis of the IS-related DNA sequence

The cosmid Y49 (accession no. Z73966; B. G. Barrell, M. A. Rajandream & S. V. Walsh) sequenced as part of the *M. tuberculosis* Genome Sequencing Project contains four adjacent ORFs (MTCY49.24–MTCY49.27) with a high degree of homology to IS6110 (McAdam *et al.*, 1990; Thierry *et al.*, 1990b) and other members of the IS3 family. We have designated this element IS990 (Fig. 1). IS990 is bounded by 16 bp inverted repeats. ISs of the IS3 family all contain a highly conserved three-amino-acid DD(35)E motif that is also found in retroviral integrases and retroelements (Khan *et al.*, 1990; Kulkosky *et al.*, 1992; Polard & Chandler, 1995). This motif was present within the IS990 sequence (Fig. 2). In common with other elements of the IS3 family, the IS990 sequence contains two ORFs (ORFa and ORFb) linked by a putative translational frameshifting site (–1) (Fayet *et al.*, 1990; Chandler & Fayet, 1993). ORFa starts with an ATG codon (position 24995 in the Y49 cosmid) and stops at a TGA codon (position 25291 in the Y49 cosmid). The stop codon of ORFa is preceded by a sequence displaying high similarity with the putative translational frameshifting sites of IS6110 and IS1137 (Fig. 3). ORFb appears downstream of ORFa and is ended by a TAG codon in position 26183 (Y49 cosmid). The amino acid sequence of ORFb was determined after detailed analysis of the homology level between ORFb of IS990 and other ORFs of ISs belonging to the IS3 family. ORFb of IS990 displayed 36.6% identity with ORFb of IS6110, 31.5% identity with ORF5 of IS3411 (Ishiguro & Sato, 1988) and 31% identity with ORF2 of IS629 (Matsutani & Ohtsubo, 1990) at the amino acid level. These analyses revealed two frameshift mutations within ORFb. The first mutation is localized in the region between positions 25616 and 25659, and the second mutation occurs between positions 25866 and 25942 in the Y49 cosmid.

```
IS3con  W--D-TY (52-61) HS D Q G S Q Y T S --Y--L--GI--S-S--G-S DNA--E S F-G-LK-E
IS6110  W--D-TY ( 60 ) H T D R G S Q Y T S --F--L--GI--S-G--G-SYD N A--E T I-G-YK-E
IS990   E--D-TY ( 60 ) H R D A G S Q Y T A --F--P--GL--S-G--G-ALD N A--E T T-G-HR-E
          * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
```

Fig. 2. Alignment of amino acids in the DD(35)E region of the ISs' consensus motif (IS3con), IS6110 of *M. tuberculosis* and IS990. The DDE motif is in bold. —, No real amino acid preferences. Asterisks indicate consensus amino acids.

```
IS6110      CGATTTTAAAGACCGCGTCGCTTCTTCGCGGCCGA-GCTCGACCGGCG
IS1137      CGATCTTGAAGCGCGCTCGGTTTTCTTCGCGCCCGA-GCTCGACCGGCG
IS990       AAATACTCAAGCTCGCAACGAGTTTCTTCGCG-GCGGAAGTGCAGCCGCG
          ** * * * * * * * * * * * * * * * * * * * * * * * *
```

Fig. 3. Alignment of ribosome slippage regions in IS6110 of *M. tuberculosis*, IS1137 of *M. smegmatis* and IS990. Putative sequences involved in translational frameshifting are shown in bold underlined letters. Repeated and inverted regions are underlined, and not bold. Conservative nucleotides are marked with asterisks.

To confirm that the frameshift mutations were real (not sequencing errors), regions of IS990 encoding ORFb were PCR-amplified from both *M. tuberculosis* H37Rv and *M. tuberculosis* 370 chromosomal DNA and sequenced. The sequences were found to be identical to the published sequence of Y49.

Determination of the copy number

The oligonucleotides (49A and 49B) amplifying a fragment of IS990 that showed the highest homology to IS6110 were used to amplify DNA from *M. tuberculosis* H37Rv chromosomal DNA and applied as a probe in Southern hybridization experiments. The copy number of the target sequence was initially studied in the chromosomal DNA of *M. tuberculosis* 370. DNA isolated from this strain was digested with *Pst*I, *Sal*I, *Pvu*II, *Kpn*I, *Bam*HI and *Xho*I restriction enzymes, separated in a 1% agarose gel, transferred to nylon membrane and hybridized with the 49A–49B probe (Fig.

Fig. 4. Southern hybridization analysis of the copy number of the IS990 sequence in genomic DNA from *M. tuberculosis* 370 digested with *Pst*I (lane 1), *Sal*I (lane 2), *Pvu*II (lane 3), *Kpn*I (lane 4), *Bam*HI (lane 5) and *Xho*I (lane 6). Lane 7 represents the molecular mass marker (bacteriophage λ DNA restricted by *Hind*III).

Fig. 5. Southern hybridization analysis of *Pvu*II-digested DNAs from 13 selected *M. tuberculosis* strains with IS990 as a probe (lanes 2–14). Lane 1 represents the molecular mass marker (bacteriophage λ DNA restricted by *Hind*III).

4). Identification of a single hybridization band, obtained after digestion with all enzymes, indicates that the IS990 element is present only in a single copy in the *M. tuberculosis* genome. The copy number of IS990 was confirmed by sequence homology searching of the *M. tuberculosis* H37Rv genome.

Distribution of IS990 within the *M. tuberculosis* complex

The presence of ISs is often restricted to a given species or group of micro-organisms. This allows the possibility of utilizing those elements as target sequences in PCR-based diagnosis of infections. This procedure requires the analysis of a number of strains from one species, as well as closely related bacteria. Twenty-one clinical strains of *M. tuberculosis* isolated from Polish tubercu-

Fig. 6. Species-specificity analysis of IS990 by Southern hybridization with *Pvu*II-digested DNAs from different mycobacterial species. (a) Lanes: 3, *M. paratuberculosis*; 4–16, *M. avium* (different strains—see Table 1). (b) Lanes: 3 and 4, *M. bovis* BCG; 5 and 6, *M. bovis*; 7, *M. scrofulaceum*; 8, *M. fortuitum*; 9, *M. gordonae*; 10 and 11, *M. malmoense*; 12 and 13, *M. lufu*; 14, *M. simiae* 19.10; 15, *M. simiae* 19.11; 16, *M. intracellulare*; 17, *M. avium*. In both panels, lane 1 represents the molecular mass marker (bacteriophage λ DNA restricted by *Hind*III) and lane 2 was loaded with the DNA of *M. tuberculosis* 370 (positive control).

losis patients, two strains of *M. bovis* and two strains of *M. bovis* BCG were chosen for further analysis. Thirteen *M. tuberculosis* strains were studied using fingerprinting analysis with the IS6110 element as a probe, which displayed a number of different hybridization patterns (data not shown). This indicates that all the investigated isolates belong to different epidemiological groups. The chromosomal DNA of all strains was digested with *Pvu*II endonuclease and hybridized with the 49A–49B probe. The expected 6.5 kb hybridization signals were detected in all samples studied (Fig. 5).

Species-specificity analysis

To determine the host range of IS990, the chromosomal DNA from different species of mycobacteria was extracted and analysed by Southern hybridization using the probe 49A–49B. Fourteen different strains of *Mycobacterium avium* isolated from various sources, two *Mycobacterium simiae* strains, two *Mycobacterium lufu* strains, two *Mycobacterium malmoense* strains, *Mycobacterium intracellulare*, *Mycobacterium gordonae*, *Mycobacterium fortuitum*, *Mycobacterium scrofulaceum* and *Mycobacterium paratuberculosis* were tested. None of these showed a hybridization signal when used with the investigated probe (Fig. 6). These data were confirmed by PCR-based analysis using 49D

and 49E primers (Table 2). The 248 bp amplification product was observed only in the case of *M. tuberculosis* complex bacteria (*M. tuberculosis*, *M. bovis* BCG, *Mycobacterium africanum* and *Mycobacterium microti*). None of the atypical mycobacterial species tested (*M. goodii*, *M. lufu*, *M. mageritense*, *M. simiae*, *M. avium*, *M. paratuberculosis*, *Mycobacterium vaccae*, *M. fortuitum*, *Mycobacterium phlei* and *Mycobacterium smegmatis*) showed a positive signal when primed with the investigated oligonucleotides (data not shown).

DISCUSSION

The IS990 sequence characterized here shows significant homology to the other insertion elements from the IS3 family. A characteristic of these elements is the involvement of ribosomal frameshifting in the generation of functional transposase by fusion of the ORFa and ORFb polypeptides (Chandler & Fayet, 1993). The two ORFs in IS990 have been shown to display significant homology to IS3 ORFa and ORFb sequences. The potential frameshift region overlaps exactly with the same region of the IS6110 element. The nucleotide homology of the region (Fig. 3) including 30 nucleotides overlap between IS6110 and IS990, is 70%, but the sequence characteristic for ribosome slippage of IS6110 (TTTAAAG) does not appear in IS990 (ACTAAG). However, this motif is not found in IS1137 either (Garcia *et al.*, 1994). Interestingly, IS6110, IS1137 and IS990 carry a T-rich sequence of 12 nucleotides downstream of the ribosome slippage site in IS6110, which is perfectly matched in these ISs. It is possible that also a T-rich sequence is involved in the ribosome slippage mechanism.

Two additional frameshift regions were found in IS990. Since these could have been specific for *M. tuberculosis* H37Rv (the subject of the Genome Sequencing Project), the DNA sequence of this region was determined for both *M. tuberculosis* H37Rv and an additional strain of *M. tuberculosis* (strain 370, a clinical isolate). The same nucleotide sequence was obtained for both *M. tuberculosis* H37Rv and *M. tuberculosis* 370, identical with data presented in the Y49 clone. This indicates that these frameshifts are real and likely to represent mutations in the transposase coding sequence. It is therefore likely that IS990 cannot code for a functional transposase. This may be the reason for the presence of only a single copy of the element in the *M. tuberculosis* chromosome – the element is likely to be defective.

ISs and other mobile elements usually exist in multiple copies in chromosomal DNA. The widely used *M. tuberculosis*-complex-specific IS6110 is present in a highly variable copy number, usually 5–20 copies per genome (van Soolingen *et al.*, 1991). However, some strains carry only one or two copies of the IS6110 element in a genome and the element is absent in some strains (Das *et al.*, 1995; McFadden & Sajduda, 1996). The presence of IS990 as a specific single-copy element in 21 different strains of *M. tuberculosis* strongly suggests its possible application in diagnosis of *M.*

tuberculosis infections. Further analysis using a larger number of strains isolated in other countries would, however, be required.

ACKNOWLEDGEMENTS

We thank Susan Wall for supplying the genomic DNA from MOTT strains and Stephen Dyer for sequencing. The work performed in Poland was supported by a grant from the State Committee for Scientific Research (KBN, contract no. 4PO-SD05412). A.S. was supported by a British Council Fellowship and J.D. by a British–Polish Joint Research Collaboration Programme at the University of Surrey, UK.

REFERENCES

- Billman-Jacobe, H., Sloan, J. & Coppel, R. L. (1996). Analysis of isoniazid-resistant transposon mutants of *Mycobacterium smegmatis*. *FEMS Microbiol Lett* **144**, 47–52.
- Brisson-Noel, A., Aznar, C., Chureau, C. & 7 other authors (1991). Diagnosis of tuberculosis by DNA amplification in clinical practice evaluation. *Lancet* **10**, 364–366.
- Chandler, M. & Fayet, O. (1993). Translational frameshifting in the control of the transposition in bacteria. *Mol Microbiol* **7**, 497–503.
- Cirillo, J. D., Barletta, R. G., Bloom, B. R. & Jacobs, W. R., Jr (1991). A novel transposon trap for mycobacteria: isolation and characterization of IS1096. *J Bacteriol* **173**, 7772–7780.
- Collins, D. M. & Stephens, D. M. (1991). Identification of an insertion sequence, IS1081, in *Mycobacterium bovis*. *FEMS Microbiol Lett* **83**, 11–16.
- Das, S., Paramasivan, C. N., Lowrie, D. B., Prabhakar, R. & Narayanan, P. R. (1995). IS6110 restriction fragment length polymorphism typing of clinical isolates of *Mycobacterium tuberculosis* from patients with pulmonary tuberculosis in Madras, South India. *Tubercle Lung Dis* **76**, 550–554.
- Dellagostin, O. A., Wall, S., Norman, E., O'Shaughnessy, T., Dale, J. W. & McFadden, J. (1993). Construction and use of integrative vectors to express foreign genes in mycobacteria. *Mol Microbiol* **10**, 983–993.
- Eisenach, K. D., Crawford, J. T. & Bates, J. H. (1988). Repetitive DNA sequences as probes for *Mycobacterium tuberculosis*. *J Clin Microbiol* **26**, 2240–2245.
- van Embden, J. D., Crawford, J. T., Dale, J. W., Gicquel, B., Hermans, P., McAdam, R., Shinnick, T. M. & Small, P. M. (1993). Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* **31**, 406–409.
- England, P. M., Wall, S. & McFadden, J. J. (1991). IS900-promoted stable integration of a foreign gene into mycobacteria. *Mol Microbiol* **5**, 2047–2052.
- Fayet, O., Ramond, P., Polard, P., Prere, M. F. & Chandler, M. (1990). Functional similarities between retroviruses and the IS3 family of bacterial insertion sequences. *Mol Microbiol* **4**, 1771–1777.
- Fomukong, N. G., Dale, J. W., Osborn, T. W. & Grange, J. M. (1992). Use of gene probes based on the insertion sequence IS986 to differentiate between BCG vaccine strains. *J Appl Bacteriol* **72**, 126–133.
- Garcia, M. J., Guilhot, C., Lathigra, R., Menendez, M. C., Domenech, P., Moreno, C., Gicquel, B. & Martin, C. (1994). Insertion sequence IS1137, a new IS3 family element from *Mycobacterium smegmatis*. *Microbiology* **140**, 2821–2828.

- Green, E. P., Tizard, M. L., Moss, M. T., Thompson, J., Winterbourne, D. J., McFadden, J. J. & Hermon-Taylor, J. (1989). Sequence and characteristics of IS900, an insertion element identified in a human Crohn's disease isolate of *Mycobacterium paratuberculosis*. *Nucleic Acids Res* 17, 9063–9073.
- Guilhot, C., Otal, I., van Rompaey, I., Martin, C. & Gicquel, B. (1994). Efficient transposition in mycobacteria: construction of *Mycobacterium smegmatis* insertional mutant libraries. *J Bacteriol* 176, 535–539.
- Hermans, W. M., van Soolingen, D., Dale, J. W., Schuitema, A. R., McAdam, R. A., Catty, D. & van Embden, J. D. A. (1990). Insertion element IS986 from *Mycobacterium tuberculosis*: a useful tool for diagnosis and epidemiology of tuberculosis. *J Clin Microbiol* 28, 2051–2058.
- Hernandez Perez, M., Fomukong, N. G., Hellyer, T., Brown, I. N. & Dale, J. W. (1994). Characterization of IS1110, a highly mobile genetic element from *Mycobacterium avium*. *Mol Microbiol* 12, 717–724.
- Ishiguro, N. & Sato, G. (1988). Nucleotide sequence of IS3411, which flanks the citrate utilization determinant of transposon Tn3411. *J Bacteriol* 170, 1902–1906.
- Khan, E., Mack, J. P. G., Katz, R. A., Kulkosky, J. & Skalka, A. M. (1990). Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res* 19, 851–860.
- Kulkosky, J., Jones, K. S., Katz, R. A., Mack, J. P. G. & Skalka, A. M. (1992). Residues critical for retroviral recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol Cell Biol* 12, 2331–2338.
- Kunze, Z. M., Wall, S., Appelberg, R., Silva, M. T., Portaels, F. & McFadden, J. J. (1991). IS901, a new member of a widespread class of atypical insertion sequences, is associated with pathogenicity in *Mycobacterium avium*. *Mol Microbiol* 5, 2265–2272.
- Kunze, Z. M., Portaels, F. & McFadden, J. J. (1992). Biologically distinct subtypes of *Mycobacterium avium* differ in possession of insertion sequence IS901. *J Clin Microbiol* 30, 2366–2372.
- McAdam, R. A., Hermans, P. W. M., van Soolingen, D., Zainuddin, Z. F., Catty, D., van Embden, J. D. A. & Dale, J. W. (1990). Characterization of a *Mycobacterium tuberculosis* insertion sequence belonging to the IS3 family. *Mol Microbiol* 4, 1607–1613.
- McFadden, J. & Sajduda, A. (1996). Application of molecular methods to the epidemiology of tuberculosis and other mycobacteriosis. *Med Microbiol Lett* 5, 394–399.
- Matsutani, S. & Ohtsubo, E. (1990). Complete sequence of IS629. *Nucleic Acids Res* 18, 1899.
- Moss, M. T., Malik, Z. P., Tizard, M. L. V., Green, E. P., Sanderson, J. D. & Hermon-Taylor, J. (1992). IS902, an insertion element of the chronic-enteritis-causing *Mycobacterium avium* subsp. *silvaticum*. *J Gen Microbiol* 138, 139–145.
- Noordhoek, G. T., Kaan, J. A., Mulder, S., Wilke, H. & Kolk, A. H. (1995). Routine application of the polymerase chain reaction for detection of *Mycobacterium tuberculosis* in clinical samples. *J Clin Pathol* 48, 810–814.
- Philipp, W. J., Poulet, S., Eiglmeier, K. & 7 other authors (1996). An integrated map of the genome of the tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*. *Proc Natl Acad Sci USA* 93, 3132–3137.
- Polard, P. & Chandler, M. (1995). Bacterial transposases and retroviral integrases. *Mol Microbiol* 15, 13–23.
- Sola, C., Horgen, L., Goh, K. S. & Rastogi, N. (1997). Molecular fingerprinting of *Mycobacterium tuberculosis* on a Caribbean island with IS6110 and DRr probes. *J Clin Microbiol* 35, 843–846.
- van Soolingen, D., Hermans, P. W. M., de Haas, P. E. W., Soll, D. R. & van Embden, J. D. A. (1991). Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J Clin Microbiol* 29, 2578–2586.
- Thierry, D., Brisson-Noel, A., Vincent-Levy-Frebault, V., Nguyen, S., Guesdon, J. L. & Gicquel, B. (1990a). Characterization of a *Mycobacterium tuberculosis* insertion sequence, IS6110, and its application in diagnosis. *J Clin Microbiol* 28, 2668–2673.
- Thierry, D., Cave, M. D., Eisenach, K. D., Crawford, J. T., Bates, J. H., Gicquel, B. & Guesdon, J. L. (1990b). IS6110, an IS-like element of *Mycobacterium tuberculosis*. *Nucleic Acids Res* 18, 188–202.
- Vary, P. H., Andersen, P. R., Green, E., Hermon-Taylor, J. & McFadden, J. J. (1990). Use of highly specific DNA probes and the polymerase chain reaction to detect *Mycobacterium paratuberculosis* in Johne's disease. *J Clin Microbiol* 28, 933–937.
- Yuen, L. K., Ross, B. C., Jackson, K. M. & Dwyer, B. (1993). Characterization of *Mycobacterium tuberculosis* strains from Vietnamese patients by Southern blot hybridization. *J Clin Microbiol* 31, 1615–1618.

Received 3 February 1998; revised 22 June 1998; accepted 12 August 1998.