J Clin Pathol: Mol Pathol 1995;48:M124–M132

## Papers

# Rapid identification of mycobacteria from AIDS patients by capillary electrophoretic profiling of amplified SOD gene

T J Bull, D C Shanson, L C Archard

Medical Microbiology Department, Charing Cross and Westminster Medical School, London W6 8RF
T J Bull
D C Shanson

Department of Biochemistry, Charing Cross Hospital, London W6 8RF
L C Archard

Correspondence to:
Dr Tim Bull, Medical Microbiology Department, Charing Cross Hospital, London W6 8RF, UK.

Accepted for publication
20 March 1995

**Abstract**

***Aim*—Rapid differentiation of mycobacterial species at the genomic level.**

***Methods*—The manganese superoxide dismutase (SOD) gene (464 bp) and 16SrRNA (353 bp) from 104 isolates (18 species) of mycobacteria were amplified using polymerase chain reaction (PCR). Products were sequenced and a phenogram of SOD sequences derived. PCR products of SOD gene were digested with HaeIII, and restriction fragment profiles visualised using capillary electrophoresis.**

***Results*—Novel SOD sequences were found for *M szulgai, M marinum, M phlei, M smegmatis, M chelonei, M paratuberculosis, M malmoense, M intracellulare* serotype 7, *M intracellulare* serotype 18, and *M celatum* types 1, 2, and 3. Phylogenetic analysis indicated that 18 of 19 species studied had 8–29% interspecies and <6% intraspecies sequence diversity in the SOD gene. No consistent differences were detected between AIDS and non-AIDS isolates. *M paratuberculosis* showed a unique SOD sequence with a 1·1% (SD 0·5%) diversity from *M avium*. Capillary electrophoresis profiles were able to differentiate 16 of 18 species within 24 hours.**

***Conclusions*—A phenogram of SOD sequences clearly delineated all mycobacterial species and showed two distinct clusters, fast growing species, and the *M avium* complex (MAC). Within the MAC, *M avium* (five types), *M intracellulare* (five types), *M scrofulaceum* (two types), and *M paratuberculosis* (one type) could be demonstrated. Phylogenetic diversity of *M celatum* from MAC, previously suggested by 16SrRNA data, was confirmed. This simple and rapid method for DNA extraction, in conjunction with capillary electrophoresis of SOD restriction fragments, allows rapid identification of mycobacterial isolates.**

(*J Clin Pathol: Mol Pathol* 1995;**48**:M124–M132)

Keywords: Mycobacteria, superoxide dismutase, rapid identification.

Until recently the taxonomic classification and identification of bacteria has been restricted to phenotypic systems based upon biochemical tests.[1] The advent of molecular based techniques, which can rapidly detect and identify bacteria in clinical samples,[2-4] has highlighted the need for a genetically based classification of bacteria. The relatively slow growth of mycobacteria in culture media (1–12 weeks) has emphasised this need, with particular reference to the rapid differential diagnosis of *M tuberculosis* and *M avium* complex (MAC) infection in AIDS patients.[5-7] Studies have shown the short variable region of the 16SrRNA gene to be suitable for phylogenetic comparisons[8] and a detailed taxonomic system has been proposed for the *Mycobacteriaceae* based upon these data.[9] Confirmation of these relationships in other genes has not been made; however, recent studies of the manganese superoxide dismutase (SOD) gene have indicated a significant sequence diversity,[10] suitable for phylogenetic analysis.

## Methods

### CULTURES

The 104 isolates used for this study were all identified by conventional biochemical means at the Regional Centre for Tuberculosis Bacteriology, PHLS, Dulwich Hospital, London, and consisted of: (1) 52 isolates of MAC from 36 patients (this group contained 35 isolates of *M avium* from 23 patients; 10 identical isolates from eight patients described here as *M celatum* type 3; four isolates of *M intracellulare*, serotype unknown, from two patients; and three isolates of *M intracellulare* from two patients consistent with serotype 18), 10 isolates of *M kansasii* from eight patients, nine isolates of *M tuberculosis* from nine patients, one isolate of *M scrofulaceum*, and one isolate of *M simiae*, all from patients with AIDS; (2) 10 isolates of MAC from 10 patients, one isolate of *M malmoense*, two isolates of *M kansasii*, all from patients without AIDS; (3) 18 type culture mycobacterial reference strains: *M fortuitum* (NCTC 10394), *M smegmatis* (NCTC 10265), *M phlei* (NCTC 8151), *M*

*szulgai* (NCTC 25932), *M tuberculosis* H37Rv (NCTC 7416), *M malmoense* (NCTC 11298), *M avium* (NCTC 8559), *M intracellulare* (NCTC 10682), *M chelonei* (NCTC 0946), *M xenopi* (NCTC 10042), *M marinum* (NCTC 2275), *M gordonae* (NCTC 10267), *M kansasii* (NCTC 10268), *M scrofulaceum* (NCTC 10803), *M paratuberculosis* (NCTC 8578), *M intracellulare* serotype 7 (ATCC 35847), *M celatum* type 1 (ATCC 51131), and *M celatum* type 2 (ATCC 51130).

#### DNA EXTRACTION

All cultures were grown to purity on Middlebrooks 7H10 for 4–6 weeks. Two or three colonies were emulsified in 500 μl TE buffer (50 mM Tris (pH 8·2), 1 mM EDTA) in a screw capped reaction tube and centrifuged at 13 000 rpm for five minutes. The supernatant was removed and the pellet resuspended in 25 μl TE. This suspension was frozen rapidly for 10 seconds in liquid nitrogen and then sonicated in a water bath (Perkin Elmer) for 30 seconds at room temperature. This procedure was repeated three times and the suspension heated for five minutes at 98°C in a heating block, then rapidly cooled for 10 seconds in liquid nitrogen and sonicated again for one minute. DNA preparations were stored at −20°C before amplification.

#### PRIMERS

*For SOD*

Primers were designed from homologous regions of previously published SOD gene sequences from *M tuberculosis*[7] and *M leprae*[11] to give a 464 base pair (bp) product. These were as follows:

SF1: ACATCTCGGGTCAGATCAACGACG
SR1: GACGTTCTTGTACTGCAGGTA

*For 16SrRNA*

Primers were as previously described[3] and gave a 353 bp product. These were as follows:

16SpA: AGAGTTTGATCCTGGCTCAG
16SPC*: CCCACTGCTGCCTCCCGTAG

Primers were synthesised on an Applied Biosystems 380B automated synthesiser and desalted using NAP 10 columns (Pharmacia). Primer quality was assessed following capillary electrophoresis using Microgel capillaries (Applied Biosystems) on an Applied Biosystems 270A capillary electrophoresis unit.

#### PCR CONDITIONS AND SEQUENCING

PCR reaction mixes (100 μl) for each of the primer pairs were prepared from each of the test isolates using the following conditions: 1 μl of extracted chromosomal DNA was amplified in a Gene Amp 9600 thermocycler (Perkin Elmer) with 10 μl 10X polymerase chain reaction (PCR) buffer (100 mM Tris (pH 8·3), 500 mM KCl, 15 mM $MgCl_2$, 0·015% gelatin). 16 μl deoxyribonucleotide mix (1·25 μmol per dNTP), 5 μl of each primer (20 μM), and 5 U Taq polymerase made to 100 μl with water. The cycling profile consisted 30 cycles of one minute at 94°C, one minute at 58°C, and one minute at 72°C, followed by a final incubation at 72°C for five minutes. Relevant products were then isolated by agarose gel electrophoresis of 70 μl amplified reaction product, excised by use of a sterile scalpel, extracted with phenol-chloroform-isoamylalcohol, concentrated by ethanol precipitation, and final concentration estimated for sequencing using a further agarose gel electrophoresis. Automated sequencing was made, with 0·5 μg product DNA, using the PRISM Taq dye termination system (Perkin Elmer) with the original primers. The cycling profile for sequencing was 25 cycles of 96°C for 15 seconds followed by 60°C for five minutes. Sequencing of both strands of product was made and aligned for base verification. Sequences with ambiguities between strands were checked by repeat sequencing from the same DNA preparation.

#### RESTRICTION DIGESTION AND CAPILLARY ELECTROPHORESIS

The PCR reaction mix (8 μl) was used for restriction digestion by mixing with 1 μl (1 U/μl) HaeIII enzyme (Promega) and 1 μl One-PhorAll buffer (Pharmacia). Digestions were made at 37°C for 1·5 hours. Samples for capillary electrophoresis were prepared by mixing 1 μl sample digest, 18 μl water, and 1 μl (38 μg/ml) standard pBR322 HaeIII digest (Sigma). Samples were electrokinetically loaded for 80 seconds at −5 kV and then electrophoresed for 15 minutes at −13 kV. Capillary electrophoresis profiles were obtained using a DNA fragment analysis kit (Applied Biosystems) according to the manufacturers' instructions on an Applied Biosystems 270HT interfaced with Turbochrom integration software (Perkin Elmer). An estimate of restriction fragment sizes was then calculated by calibrating retention times of the internal standard against sample fragments.

#### PHYLOGENETIC ANALYSIS

Phylogenetic analyses were made using the PHYLIP package available on SEQNET using the FITCH algorithm and Jukes and Cantor DNADIST method[12] for estimations of phylogenetic diversity. Human SOD sequence[13] was aligned to mycobacterial sequences and used as an outgroup to root the tree. Values are given as means (SD).

## Results

#### BIOCHEMICAL IDENTIFICATION AND 16SrRNA PROFILES

A 16SrRNA sequence profile was obtained from all 104 isolates. In the majority of isolates the profile corresponded to the biochemical identification and previously reported 16SrRNA profiles[3] after sequencing one strand only. Isolates giving irregular sequences were verified from both strands and repeated if

```
              193
human         CACCACAGCAAGCACCACGCGGCCTACGTGAACAACCTGAACGTCACCGAGGAGAAGTACCAGGAGGCGTTGGCCAAGGGAGATGTTACAGCCCAGACAG
avium type1   CACCACACCAAGCACCACGCCACCTACGTCAAAGGCGTGAACGACGCTCTTGCCAAGCTCGAAGAGGCCCGCGCCAACGAGGACCACGCTGCG-ATCTTC
avium type2   CACCACACCAAGCACCACGCCACCTACGTCAAAGGCGTGAACGACGCTCTTGCCAAGCTCGAAGAGGCCCGCGCCAACGAGGACCACGCTGCG-ATCTTC
avium type3   CACCACACCAAGCACCACGCCACCTACGTCAAAGGCGTGAACGACGCTCTTGCCAAGCTCGAGGAGGCCCGCGCCAACGAGGACCACGCTGCG-ATCTTC
avium type4   CACCACACCAAGCACCACGCCACCTACGTCAAAGGCGTGAACGACGCTCTTGCCAAGCTCGAAGAGGCCCGCGCCAACGAGGACCACGCTGCG-ATCTTC
avium type5   CACCACACCAAGCACCACGCCACCTACGTCAAAGGCGTGAACGACGCTCTTGCCAAGCTCGAAGAGGCCCGCGCCAACGAGGACCACGCTGCG-ATCTTC
celatum.t3    CACCACAGCAAACACCATGCGACCTACGTGAAGGGCGCCAACGACGCGCTCGAAAAACTCGAGGAAGCACGCGCCAAGGACGACCAGTCGACC-ATCCTG
celatum.t2    CACCACAGCAAACACCATGCGACCTACGTGAAGGGCGCGAACGACGCGCTTGAAAAACTCGAGGAAGCACGCGCCAAGGACGACCAGTCGACC-GTCCTG
chelonei      CAACACAGTAAGCACCACGCGGCCTACGTCGCGGGTGTCAATTCCGCTGTCGCCAAGTTGGAAGAGGCTCGCGAGAAGGGCGACCACGCCGCG-ATCTTC
fortuitum     CACCACAGCAAGCACCACGCGGCGTACGTCAAGGGCGTCAACGACGCCGTGGCCAAGCTCGATGAGGCGCGGGCCAACGGTGACCACGCGGCG-ATCTTC
gordonae      TCACACAGTAAGCACCACGCCACCTACGTCAAGGGCGTCAACGACGCGGTCGCCAAGCTGGAAGAGGCGCGCGCCAAGGCGACCACTCGGCC-ATCTTT
intra type1   CACCACAGTAAGCACCACGCCACCTACGTCAAAGGCGTGAACGACGCTCTGTCCAAGCTCGAAGAGGCCCGTGCCAACGAAGATCACGCTGCG-ATCTTC
intra type2   CACCACAGTAAGCACCACGCCACCTACGTCAAAGGCGTGAACGACGCTCTGTCCAAGCTCGAAGAGGCCCGTGCCAACGAAGATCACGCTGCG-ATCTTC
intra type3   TCACACAGTAAGCACCACGCCACCTACGTCAAAGGCGTGAACGACGCTCTTGCCAAGCTCGAAGAGGCCCGTGCCAACGAAGATCACGCTGCG-ATCTTC
intra type4   TACCACAGTAAGCACCACGCCACATACGTCAAAGGCGTGAATGACGCTCTGTCCAAGCTCGAAGAGGCCCGCGCCAACGAGGACCACGCTGCG-ATCTTC
intra type5   CACCACACTAAGCACCACGCCACGTACGTCAAAGGCGTGAACGACGCTCTGTCCAAGCTCGAAGAGGCCCGCGCCAACGAGGACCACGCTGCG-ATCTTC
leprae        CACCACACCAAGCACCACGCCGCATATGTCAAAGGTGTCAATGACGCGCTTGCCAAACTTGACGAGGCACGCGCCAAAGACGACCACTCCGCG-ATTTTT
malmoense     CACCACAGCAAGCACCACGCCGCCTACGTCAAGGGCGTGAACGACGCCGTCGCCAAGCTTGAAGAGGCGCGGGCCAAGGACGACCACTCGGCG-ATCTTC
marinum       TCACACAGTAAGCACCACGCCACCTACGTCAAGGGTGCCAATGACGCCGTCACCAAACTCGAGGAAGCGCGCGCTAAGGAAGACCACTCGACG-ATCCTG
kansasii t1   CACCACAGTAAGCACCACGCCACCTACGTCAAGGGCGCCAACGATGCGGTCGCCAAACTCGAAGAGGCGCGCGCCAAGGAAGACCACTCGGCG-ATCTTG
kansasii t2   GCACACAGTAAGCACCACGCCACCTACGTCAAGGGCGCCAACGATGCGGTCGCCAAACTCGAAGAGGCGCGCGCCAAGGAAGACCACTCGGCG-ATCTTG
paratb        CACCACACCAAGCACCACGCCACCTACGTCAAAGGCGTGAACGACGCTCTTGCCAAGCTCGAAGAGGCCCGCGCCAACGAGGACCACGCTGCG-ATCTTC
phlei         TCACACAGCAAGCACCACGCCACCTACGTCAAGGGCGTCAACGACGCGATCGCCAAGCTTGAAGAAGCGCGCGCCAACGGCGATCACGGCGCG-ATCTTC
M.sp.40407    TCACACAGCAAACACCATGCGACCTACGTCAAGGGTGCCAACGACGCGCTCGCCAAGCTCGAGGAGGCACGCGCCAACGATGACCAGTCGACG-ATCCTG
scrof t1      CATCACACCAAGCACCACGCCACGTACGTCAAGGGCGTGAACGACGCCGTCGCCAAACTCCAAGAGGCACGCGCCAATGACGACCACGCCGCG-ATCTTC
scrof t2      CATCACACCAAGCACCACGCCACCTACGTCAAGGGCGTGAACGACGCGGTCGCCAAACTCCAAGAGGCGCGCGCCAACGACGACCACGCCGCG-ATCTTC
simiae        CACCACAGCAAGCACCATGCGACGTACGTCAAGGGTTTGAACGACGCCATTGCCAAGCTTGAAGAGGCACGGGCCAACGACGACCACGCCGCG-ATCTTC
smegmatis     CACCACAGCAAGCACCACGCGACCTACGTCAAGGGTGTGAACGACGCGATTGCCAAGCTCGAGGAGGCACGGGCCAACGGTGACCACGCGGCC-ATCTTC
szulgai       TCACACAGCAAGCACCACGCCACCTACGTCAAGGGCGCCAATGACGCTGTCGCCAAACTCGAGGAGGCGCGCGCTCAGGAGGACTTTTCGTCG-ATCTTG
tb            CACCACAGCAAGCACCACGCCACCTACGTAAAGGGCGCCAATGACGCCGTCGCCAAACTCGAAGAGGCGCGCGCCAAGGAAGATCACTCAGCG-ATCTTG
xenopi        CACCACAGCAAGCACCACGCGACGTACGTCAAAGGCGCCAACGACGCGCTCGCCAAGCTGGAGGAGGCGCGCGCCAAAGACGATCATTCCGCG-ATCGTC
              293
human         CTCTTCAGCCTGCAC-TGAAGTTCAATGGTGGTGGTCATATCAATCATAGCATTTTCTGGACAAACCTCAGCCCTAACGGTGGTGGAGAACCCAAAGGGG
avium type1   CTGAACGAAAAGAACCTCGCCTTCCACCTGGGCGGCCACGTCAACCACTCGATCTGGTGGAAGAACCTGTCGCCGGACGGCGGTGACAAGCCCACCGGTG
avium type2   CTGAACGAAAAGAACCTCGCCTTCCACCTGGGCGGCCACGTCAACCACTCGATCTGGTGGAAGAACCTGTCGCCGGACGGCGGTGACAAGCCCACCGGTG
avium type3   CTGAACGAAAAGAACCTCGCCTTCCAGCTGGGCGGCCACGTCAACCACTCGATCTGGTGGAAGAACCTGTCGCCGGACGGCGGTGACAAGCCCACCGGTG
avium type4   CTGAACGAAAAGAACCTCGCCTTCCACCTGGGCGGCCACGTCAACCACTCGATCTGGTGGAAGAACCTGTCGCCGGACGGCGGTGACAAGCCCACCGGCG
avium type5   CTGAACGAAAAGAACCTCGCCTTCCACCTGGGCGGCCACGTCAACCACTCGATCTGGTGGAAGAACCTGTCGCCGGACGGCGGTGACAAGCCCACCGGCG
celatum t3    TTGAACGAGAAGAACCTGGCGTTCAACCTGGCCGGCCACGTGAACCACACCATCTGGTGGAAGAACCTGTCGCCTAACGGCGGTGACAAGCCGACCGGTG
celatum t2    TTGAACGAGAAGAACTTGGCGTTCAACCTGGCCGGCCACGTCAACCACACCATCTGGTGGAAGAACCTGTCGCCCAACGGCGGTGACAAGCCGACCGGGG
chelonei      CTCAACGAGAAGAACCTGGCCTTCCACCTCGGTGGCCACGTGAACCACTCCATCTGGTGGAAGAACCTGTCGCCCAACGGCGGCGACAAGCCCACCGGCG
fortuitum     CTCAACGAGAAGAACCTGGCGTTCCATCTCGGCGGCCACGTGAACCACTCGATCTGGTGGAAGAACCTGTCCCCCAACGGTGGCGACAAGCCGACGGGCG
gordonae      TTGAACGAGAAGAACCTGGCCTTCCACCTGGGCGGTCACGTCAACCACTCCATCTGGTGGAAGAACCTGTCGCCGGACGGCGGCGACAAGCCGACCGGTG
intra type1   CTGAACGAAAAGAACCTGGCCTTTCACCTGGGCGGCCACGTCAACCACTCCATCTGGTGGAAGAACCTGTCGCCGGACGGCGGCGACAAGCCGACCGGCG
intra type2   CTGAACGAAAAGAACCTGGCCTTTCACTTGGGCGGCCACGTCAACCACTCCATCTGGTGGAAGAACCTGTCGCCGGACGGCGGCGACAAGCCGACCGGCG
intra type3   CTGAACGAAAAGAACCTGGCCTTTCACCTGGGCGGCCACGTCAACCACTCCATCTGGTGGAAGAACCTGTCGCCGGACGGCGGCGACAAGCCGACCGGCG
intra type4   CTGAACGAAAAGAACCTCGCATTCCACCTCGGTGGCCACGTCAACCACTCGATCTGGTGGAAGAACCTGTCGCCGGACGGCGGCGACAAGCCGACCGGCG
intra type5   CTGAACGAAAAGAACCTCGCCTTCCACTTGGGCGGCCACGTCAACCACTCCATCTGGTGGAAGAACCTGTCGCCGGACGGCGGCGACAAGCCGACCGGCG
leprae        CTGAACGAGAAGAACCTGGCCTTTCATTTGGGCGGCCACGTCAACCACTCCATCTGGTGGAAGAACCTCTCTCCGAACGGTGGCGATAAGCCGACCGGCG
malmoense     CTGAACGAAAAGAACCTGGCCTTCCACCTGGGCGGGCACGTCAACCACTCGATCTGGTGGAAGAACCTGTCACCGAATGGCGGCGACAAGCCCACCGGCG
marinum       CTGAACGAGAAGAACCTCGCTTTCAACCTCGCCGGCCACGTCAACCACACCATCTGGTGGAAGAACCTCTCCCCCAACGGTGGCGACAAGCCCACCGGTG
kansasii t1   CTGAACGAGAAGAACTTGGCCTTCAACCTCGCCGGCCACGTCAACCACACGATCTGGTGGAAGAACCTTTCTCCCAACGGAGGCGACAAGCCGACCGGCG
kansasii t2   CTGAACGAGAAGAACTTGGCCTTCAACCTCGCCGGCCACGTCAACCACACGATCTGGTGGAAGAACCTTTCTCCCAACGGAGGCGACAAGCCGACCGGCG
paratb        CTGAACGAAAAGAACCTCGCCTTCCACCTGGGCGGCCACGTCAACCACTCGATCTGGTGGAAGAACCTGTCGCCGGACGGCGGTGACAAGCCCACCGGCG
phlei         CTGCACGAGAAGAACCTTGCGTTCCACCTCGGCGGGCACGTCAACCACACGATCTGGTGGAAGAACCTGTCCCCGCACGGTGGCGACAAGCCGACCGGGG
M.sp. 40407   CTGAACGAAAAGAACTTGGCGTTCAACCTGGCGGGCCACGTCAACCACAGCCTGTGGTGGAAGAACCTGTCGCCCGACGGTGGCGACAAGCCGACCGGCG
scrof t1      CTGAACGAAAAGAACCTGGCGTTCCACCTCGGCGGCCACGTGAACCACTCGATCTGGTGGAAGAACCTCTCGCCGGACGGCGGCGACAAGCCGACCGGAG
scrof t2      CTGAACGAGAAGAACCTGGCGTTCCACCTCGGCGGCCACGTGAACCATTCCATCTGGTGGAAGAACCTGTCGCCGGACGGCGGCGACAAGCCGACCGGAG
simiae        TTGAACGAGAAGAACCTGGCATTCCACCTCGGTGGCCACGTCAACCACTCCATCTGGTGGAAAAACCTGTCCCCGAACGGCGGAGACAAGCCGACCGGAG
smegmatis     CTGAACGAGAAGAACCTCGCGTTCCATCTCGGTGGCCACATCAACCACTCGATCTGGTGGAAGAACCTCTCCCCCAACGGTGGCGACAAGCCCACCGGCG
szulgai       CTGAGCGAAAAGAACCTGGCGTTCAACTTGGCCGGCCACGTCAATCACACCATCTGGTGGAAGAATCTGTCTCCCAACGGTGGTGACAAGCCAACGGGCG
tb            CTGAACGAAAAGAATCTAGCTTTCAACCTCGCCGGCCACGTCAATCACACCATCTGGTGGAAGAACCTGTCGCCTAACGGTGGTGACAAGCCCACCGGCG
xenopi        GGGCATGAGAAGGCCCTCGCGTTCAACCTGGCCGGCCATGTCAATCACTGCCTGTGGTGGAAGAACCTGTCCCCCAACGGCGGTGACAAGCCGACCGGCG
              393
human         AGTTGCTGGAAGCCATCAAACGTGACTTTGGTTCCTTTGACAAGTTTAAGGAGAAGCTGACGGCTGCATCTGTTGGTGTCCAAGGCTCAGGTTGGGGTTG
avium type1   AGCTGGCCGCCGCGATCGACGACGCGTTCGGGTCCTTCGACAAGTTCCGAGCGCAATTCAGCGCCGCCGCCAACGGCCTGCAGGGCTCCGGCTGGGCGGT
avium type2   AGCTGGCCGCCGCGATCGACGACGCGTTCGGGTCCTTCGACAAGTTCCGAGCGCAATTCAGCGCCGCCGCCAACGGCCTGCAGGGCTCCGGCTGGGCGGT
avium type3   AGCTGGCCGCCGCGATCGACGACGCGTTCGGGTCCTTCGACAAGTTCCGAGCGCAATTCAGCGCCGCCGCCAACGGCCTGCAGGGCTCCGGCTGGGCGGT
avium type4   AGCTGGCCGCCGCGATCGACGACGCGTTCGGGTCCTTCGACAAGTTCCGGGCGCAATTCAGCGCCGCCGCCAACGGCCTGCAGGGCTCCGGCTGGGCGGT
avium type5   AGCTGGCCGCCGCGATCGACGACGCGTTCGGCTCCTTCGACAAGTTCCGTGCGCAATTCAGCGCCGCCGCCAACGGACTGCAGGGCTCCGGCTGGGCGGT
celatum t3    AGCTGGCCGCGGCCATCGACGACGCGTTCGGCTCGTTCGACAAATTCCGCGCTCAGTTCCACGCCGCCGCCACCACGGTGCAGGGGTCGGGATGGGCAGC
celatum t2    AGCTGGCCGCGGCCATCGACGACGCGTTCGGCTCGTTCGACAAATTCCGCGCTCAGTTCCATGCGGCCGCCACCACCGTGCAGGGGTCGGGTTGGGCGGC
chelonei      ATCTGGCCGCCGCCATCGACGACCAGTTCGGCTCGTTCGACAAGTTCCAGGCACAGTTCACCGCCGCGGCCAACGGCCTGCAGGGCTCGGGCTGGGCCGT
fortuitum     ACCTGGCCGCGGCCATCGACGACCAGTTCGGCTCGTTCGACAAGTTCCAGGCGCAGTTCACCGCCGCCGCCAACGGGCTGCAAGGCTCCGGCTGGGCCGT
gordonae      ATCTGGCCGCCGCGATCGACGACCAGTTCGGCTCGTTCGACAAGTTCCAGGCTCAGTTCAGCGCCGCCGCAAACGGCCTACAGGGCTCGGGCTGGGCGGT
intra type1   AATTGGCCGCCGCGATCGACGACGCCTTCGGATCCTTCGACCGGTTCCGCGCGCAGTTCAGCGCGGCCGCCAACGGCCTGCAGGGGTCGGGCTGGGCGGT
intra type2   AATTGGCCGCCACGATCGATGACGCCTTCGGATCCTTCGACCGGTTCCGCGCGCAGTTCAACGCGGCCGCCAACGGCCTGCAGGGGTCGGGCTGGGCGGT
intra type3   AATTGGCCGCCACGATCGACGACGCCTTCGGATCCTTCGACCGGTTCCGCGCGCAGTTCAGCGCGGCCGCCAACGGCCTGCAGGGGTCGGGCTGGGCGGT
intra type4   AACTGGCTGCTGCGATCGACGACGCGTTCGGGTCCTTCGACCGATTCCGCGCGCAATTCAGCGCGGCCGCCAACGGGCTGCAGGGCTCGGGCTGGGCCGT
intra type5   AACTGGCCGCCGCGATCGATGACGCCTTCGGCTCCTTCGACCGCTTCCGCGCTCAGTTCAGCGCGGCCGCCAACGGACTGCAGGGTTCGGGCTGGGCGGT
leprae        GGCTAGCCACTGACATTGATGAAACGTTTGGGTCGTTCGACAAATTCCGCGCTCAATTCAGCGCGGCCGCCAATGGCCTGCAAGGGTCAGGTTGGGCAGT
malmoense     ACCTCGCGTCCGCGATCGACGACGCGTTCGGATCGTTCGACAAATTCCGCGCCCAGTTCAGCGCGGCCGCCAACGGCTTGCAGGGATCGGGCTGGGCGGT
marinum       AACTCGCCGCGGCCATCGATGAGGCGTTCGGGTCGTTCGACAAGTTTCGGGCCCAGTTCCACGCCGCCGCGACCACCGTGCAGGGGTCGGGTTGGGCAGC
kansasii t1   AACTCGCCGCGGCCATCGACGAGGCGTTCGGGTCCTTCGACAAGTTTCGTGTCCAATTCCACGCCGCCGCCACCACGGTGCAGGGGTCGGTCTGGGCGGC
kansasii t2   AACTCGCCGCGGCCATCGACGAGGCGTTCGGGTCCTTCGACAAGTTTCGTGTCCAGTTCCACGCCGCCGCCACCACGGTGCAGGGGTCGGGCTGGGCGGC
paratb        AGCTGGCCGCCGCGATCGACGACGCGTTCGGCTCCTTCGACAAGTTCCGGGCGCAATTCAGCGCCGCCGCCAACGGCCTGCAGGGCTCCGGCTGGGCGGT
phlei         ATCTCGCGGCTGCGATCGACGACCAGTTCGGATCGTTCGACAACTTTCGCGCGCAGTTCACCGCGGCCGCCAACGGTTTGCAGGGGTCCGGCTGGGCGGT
M.sp.40407    AGCTCGCCGCCGCCATCGACGACGCGTTCGGGTCGTTCGACAAGTTCCGCGCCCAGTTCACGGCGGCGGCCACCACCGTGCAGGGGTCGGGGTGGGCGGC
scrof t1      AACTGGCCGCCGCGATCGATGACGCGTTCGGATCGTTCGACAAATTCCGCGCCCAGTTCAGTGCGGCCGCCAACGGCCTGCAGGGTTCGGGCTGGGCGGT
scrof t2      AGCTGGCCGCCGCGATCGACGACGCGTTCGGATCGTTCGACAAGTTCCGCGCGCAGTTCAGTGCGGCCGCCAACGGCCTGCAGGGTTCGGGCTGGGCGGT
simiae        ATCTCGCCGCCGCCATCGACGACGCCTTCGGTTCGTTCGACAAGTTCCGCGCACAGTTCAGCGCCGCCGCCAACGGCTTGCAGGGCTCGGGCTGGGCGGT
smegmatis     AGCTCGCGGCCGCGATCGACGACCAGTTCGGCTCGTTCGACAAGTTCCAGGCACAGTTCACCGCTGCGGCCAACGGCCTGCAGGGTTCGGGCTGGGCAGT
szulgai       AGCTGGCCGCGGCCATCGATGACGCGTTCGGTTCATTCGACACGTTCAGTGCGCAGTTCCACGCCGCCTCCACCACCGTGCAGGGTTCGGGCTGGGCGGC
tb            AACTCGCCGCAGCCATCGCCGACGCGTTCGGTTCGTTCGACAAGTTCCGTGCGCAGTTCCACGCGGCCGCTACCACCGTGCAGGGGTCGGGCTGGGCGGC
xenopi        AATTGGCCGCCGCCATCGACGACGCGTTCGGCTCGTTCGACAAGTTCCGCGCCCAGTTCACCGCGGCCGCCACGACCGTGCAGGGGTCGGGCTGGGCGGC
```

necessary. Ten isolates identified originally as MAC gave an identical unique profile which we have reported elsewhere[14] as *M celatum* type 3. Isolates of *M scrofulaceum* and *M kansasii* both had two 16SrRNA profiles, identical across the species specific region, but differing

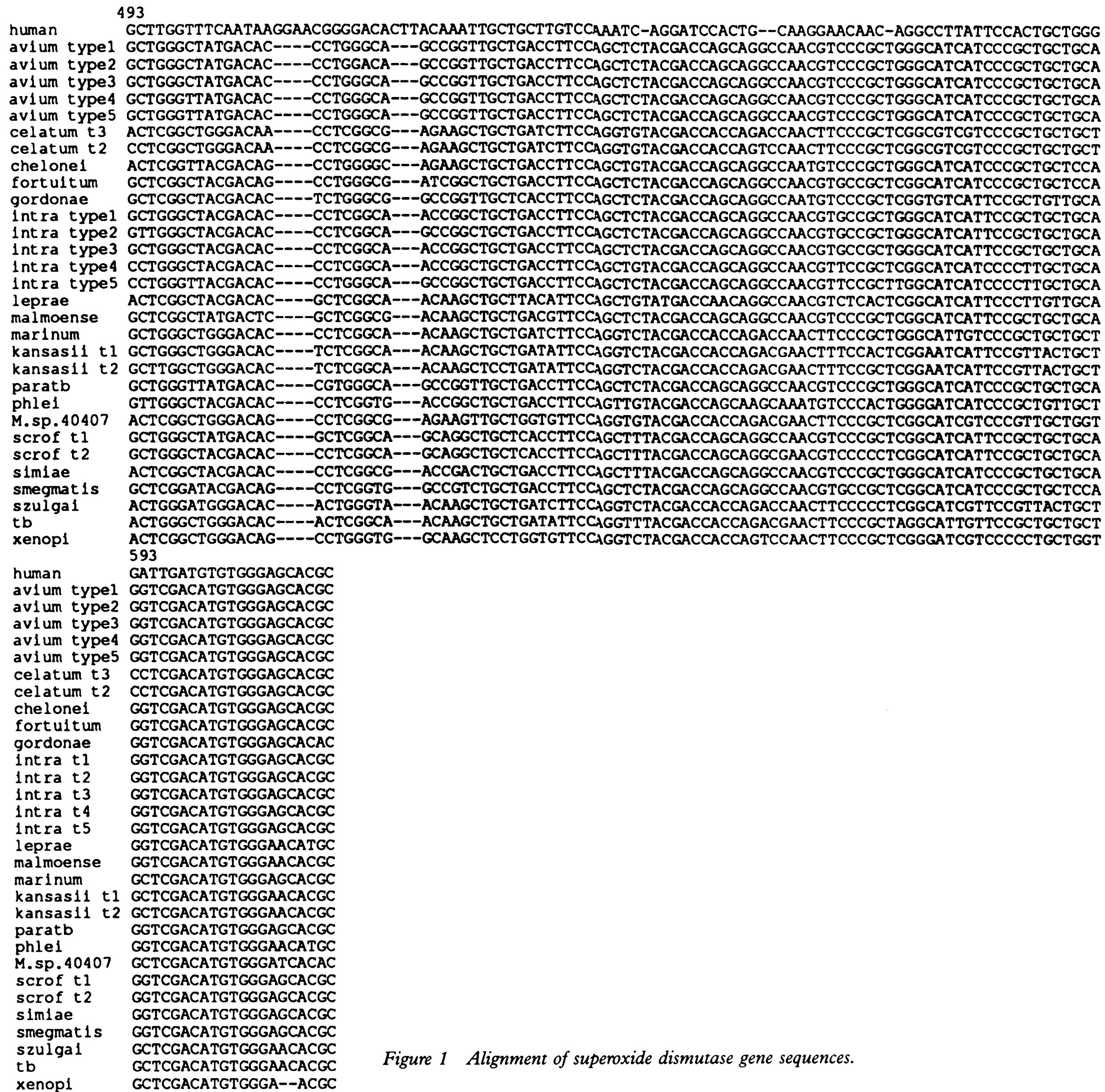

*Figure 1 Alignment of superoxide dismutase gene sequences.*

in up to five bases towards the 5′ end. One isolate showed a profile very similar to *M schimoidei* but was identified biochemically as MAC and is reported here as *M sp.* 40407/94. Four identical isolates identified biochemically as MAC gave a previously unreported unique profile, similar to *M intracellulare*, referred to here as *M intracellulare* type 4. These minor changes in 16SrRNA sequences will be reported elsewhere.

### SOD SEQUENCES

All 104 isolates gave a product with the primers SF1 and SR1 which were sequenced in both directions at least once (fig 1). Five sequence types of *M avium* were observed which showed only small variations (fig 2) in base sequence (0·9(SD 0·4)%). Thirty eight isolates referred to here as *M avium* type 1 included 33 AIDS, five non-AIDS, and the reference strain NCTC 8559. Four further isolates each had unique sequence variations and are referred to here as *M avium* types 2 and 3 (AIDS isolates), and *M avium* types 4 and 5 (non-AIDS isolates). AIDS isolates of *M avium* could not be distinguished from non-AIDS isolates on these sequence data. Isolates identified by 16SrRNA sequencing as similar to *M intracellulare* could be differentiated by minor base variations into five types, referred to as follows: three isolates designated *M intracellulare* type 1 (one AIDS isolate, one non-AIDS isolate, and reference strain NCTC 10682); one isolate designated *M intracellulare* type 2 (AIDS isolate); three isolates designated *M intracellulare* type 3 (AIDS isolates concurrent with *M intracellulare* serotype 18 16SrRNA profile); three isolates designated *M intracellulare* type 4 (AIDS isolates with previously unreported 16SrRNA profile); and one isolate designated *M intracellulare* type 5 (reference strain *M intracellulare* serotype 7; ATCC 35847). Similarly *M scrofulaceum* had one isolate designated *M scrofulaceum* type 1 (reference strain NCTC 10803) and one isolate designated *M scrofulaceum* type 2 (AIDS isolate). Overall the MAC isolates varied by 7·1 (3·7)% and were distinctly clustered in the calculated phylogenetic tree (fig 3) with a 19·5 (4·1)% diversity of this group from all other

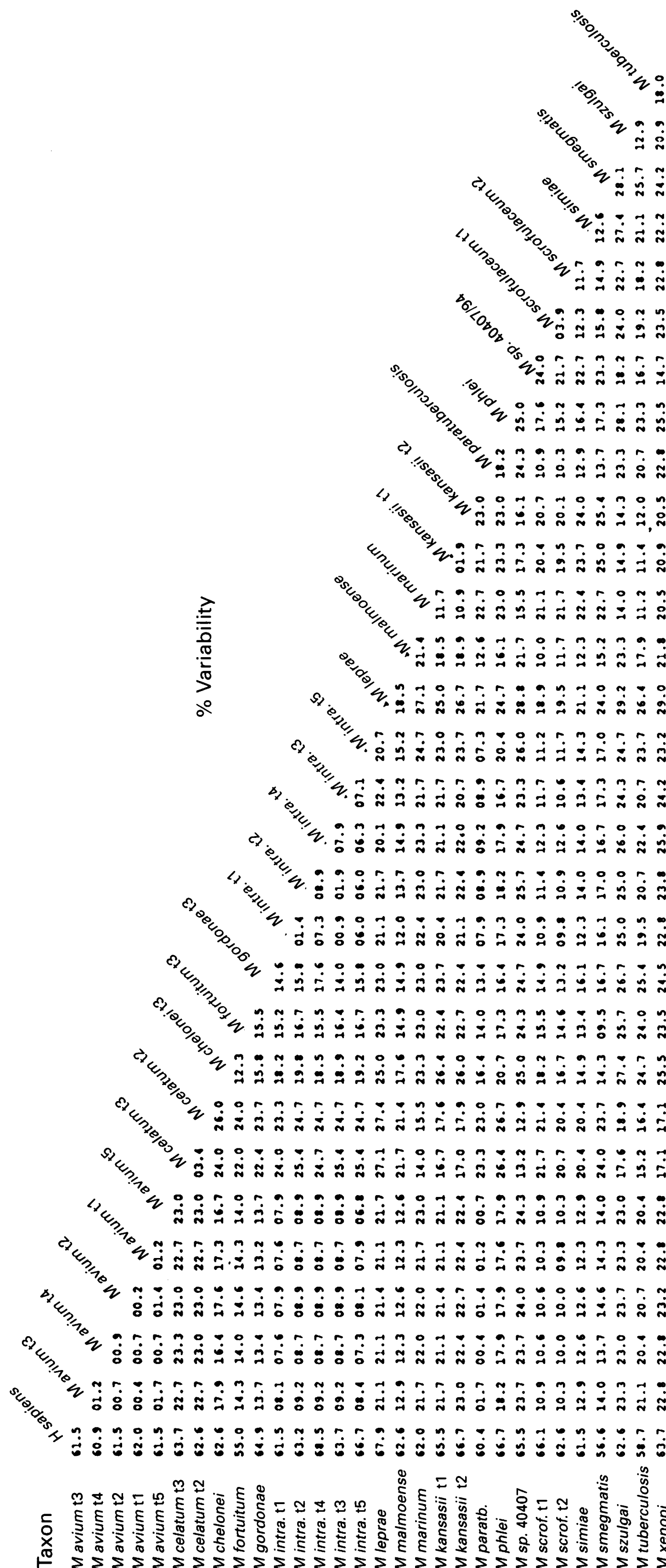


*Figure 2 Superoxide dismutase gene variability matrix calculated with Jukes and Cantor distance method.*[12]

species. *M kansasii* had 12 isolates designated *M kansasii* type 1—including nine AIDS isolates, two non-AIDS isolates, and the reference strain NCTC 10268—and one isolate designated *M kansasii* type 2 (AIDS isolate with minor 16SrRNA sequence variation). *M celatum* type 1 and 10 isolates of *M celatum* type 3 showed identical sequences but had a 3·5% variation from *M celatum* type 2 and 21·5 (3·7)% variation from all other species. All 11 isolates of *M tuberculosis* had identical sequences. Significant variation (range: 10–29) was observed between all the remaining reference strains, with all the rapid growing species also forming a distinct cluster which had 18·6 (4·0)% diversity from all slow growing species. Restriction digest maps produced from each of the 29 sequence variations revealed that HaeIII would be a suitable enzyme for the generation of unique restriction fragment profiles.

## CAPILLARY ELECTROPHORESIS PROFILES

Capillary electrophoresis of samples gave chromatograms for the fragment size range 50–264 bp. The inclusion of a standard digest into each capillary electrophoresis sample allowed an estimation of sample fragment sizes (fig 4). Capillary electrophoresis profiles, in the range 70–264 bp, were plotted for at least one representative of all isolates with unique SOD sequences. Twenty unique profiles were determined (table) which could differentiate all the recognised species studied with the exceptions of *M avium* from *M paratuberculosis*, and *M celatum* types 1 and 3 from *M szulgai*. All internal fragment sizes corresponded to those predicted from the sequence data. The majority of terminal fragments, however, gave estimated sizes of 3–7 bases larger than expected. These increases were constant for each fragment length and were the same for all species and repeat samples tested. In some samples, extra peaks were observed at 57 bp and 60 bp (fig 4). These peaks could not be explained from the sequence data and may have occurred because of non-specific amplification or primer concatenation. Their appearance also varied from sample to sample and did not occur in all repeat samples. These peaks, however, were below the chosen capillary electrophoresis profile range and therefore did not interfere with the differentiation of isolates.

## NUCLEOTIDE SEQUENCE ACCESSION NUMBERS

The following sequences have been submitted to EMBL (European Molecular Bank Library) under the accession numbers: Z48208 (*M marinum*), Z48209 (*M celatum* type 3), Z48210 (*M malmoense*), Z48211 (*M phlei*), Z48212 (*M paratuberculosis*), Z48213 (*M szulgai*), Z48214 (*M smegmatis*), Z48215 (*M celatum* type 3), Z48216 (*M chelonei*), Z48217 (*M intracellulare* serotype 18), and Z48218 (*M intracellulare* serotype 7). Other sequences in this work were found to be identical to the following existing sequences: X52861 (*M tuberculosis*), X16453 (*M leprae*), X81384 (*M avium*), X81387 (*M intracellulare*), X81385 (*M fortuitum*), X81386 (*M gordonae*), X81388 (*M kansasii*), and X81390 (*M simiae*).

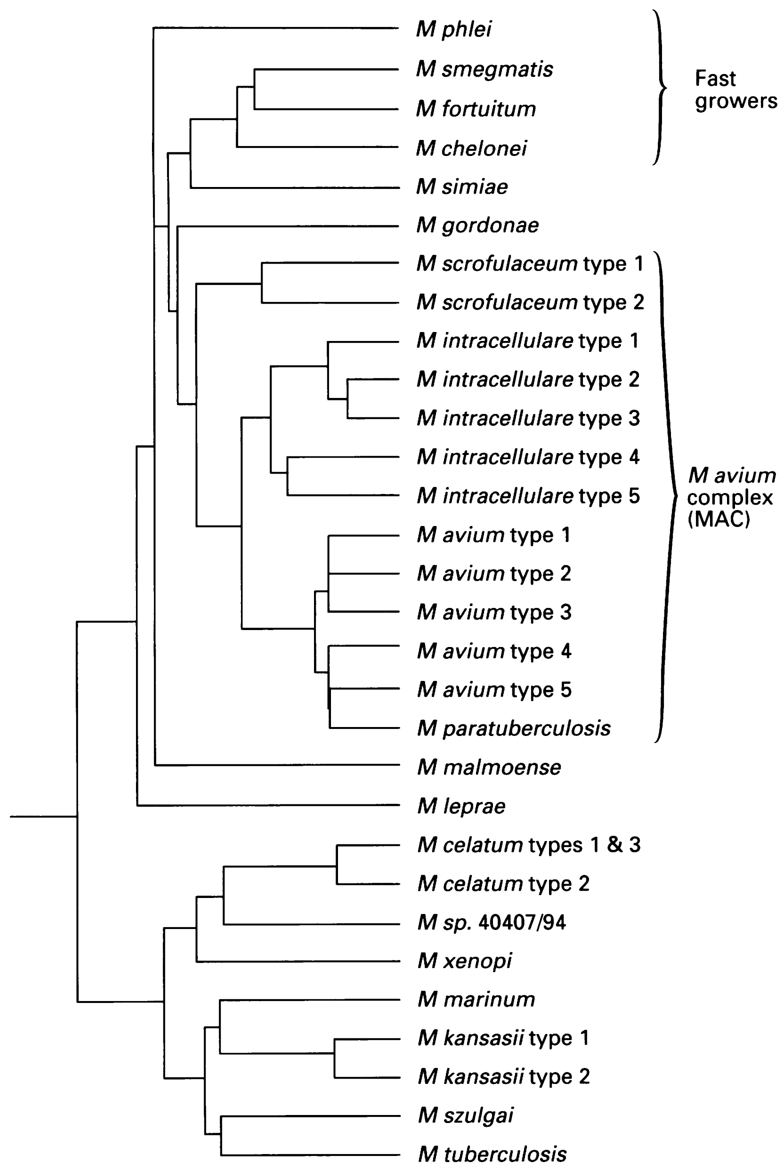


*Figure 3 Phylogenetic tree of mycobacteria derived from superoxide dismutase (SOD) gene by* FITCH *algorithm using* H sapiens *SOD*[13] *as outgroup.*

## Discussion

The diversity shown by the sequences presented in this study indicates that the SOD gene is suitable as an alternative to genes such as 16SrRNA for the genomic identification of the *Mycobacteriaceae*. AIDS isolates of *M avium* could not be distinguished from non-AIDS isolates; however, a phenogram derived from SOD sequences (fig 3) clearly delineates fast and slow growing species and MAC. *M avium*, *M intracellulare*, and *M scrofulaceum* are seen to form a cluster representative of MAC which does not include *M celatum* as suggested by biochemical identifications alone.[15] The phylogenetic diversity of the *M celatum* group from MAC (23·3 (1·3)%) which had been previously suggested using 16SrRNA data is thus confirmed, with *M celatum* being more related to *M xenopi* than MAC. All recognised species used in this study were differentiated by 8–29% base variation with the exception of *M paratuberculosis*, which differed from the observed five *M avium* types by up to four bases (1·1 (0·5)%). This difference could be significant considering the close genetic homology already recognised in studies of other genes from these two species, and that the variation at base No 513 (fig 1) predicts a change in amino acid sequence (valine to leucine). These data concur with 16SrRNA studies in that *M paratuberculosis* should be regarded as a subtype of *M avium*. We believe this report is the first to be able to differentiate *M paratuberculosis* from *M avium* by gene sequence data alone. Minor intraspecies sequence differences were also observed between clinical isolates and reference strains within the species of *M scrofulaceum* (3·9%), *M kansasii* (1·9%), and *M celatum* (3·5%). In the *M intracellulare* samples a larger variation of 5·4 (2·7)% was observed, illustrating the high diversity within this loosely defined species.

We show here that sequence diversity in SOD can be used to predict unique capillary electrophoresis profiles and have rapidly differentiated 16 of 18 species tested with this method. We have only used a single restriction enzyme digest in this study; however, in combination with further restriction digests it would be possible to differentiate all the species described. The advantage of using a single low cut restriction digest can be seen with the results from *M avium* and *M kansasii*, where intraspecies base variation did not occur across restriction sites and thus only one unique profile was seen for all the species types (table).

Terminal restriction digest fragment sizes observed by capillary electrophoresis in the majority of species were constantly larger than predicted from the sequence data. Similar migration retardation anomalies have been reported,[16] possibly due to DNA fragment secondary structure and the concentration and type of polymer used in the capillary. The reproducibility of the technique described here suggests, however, that the observed larger fragment sizes will not present a problem in the identification of species in unknown samples, providing adequate control profiles are used.

A recently described method[17] also uses the SOD gene to detect mycobacterial species-specific sequences. This method presupposes that intraspecies diversity is low enough to allow relatively high stringency hybridisation of short species-specific DNA probes homologous to amplified SOD product. This study shows that four of 42 (10%) clinical isolates of *M avium* contained sequence variations. If variation occurs within the probe sites it could result in failure to hybridise, thereby reducing the specificity of the test. Our approach using identification by low cut restriction analysis circumvents this potential problem by producing capillary electrophoresis profiles for all isolates, regardless of sequence variation. The use of a low cut enzyme guarantees a low probability of variation across restriction sites which would produce profiles dissimilar to the control profile bank. Even in this event, profiles can be extrapolated to give the most likely identification from the nearest fit. This study shows that the high diversity between SOD sequences in mycobacteria ensures that profiles of one species will not be transformed by minor

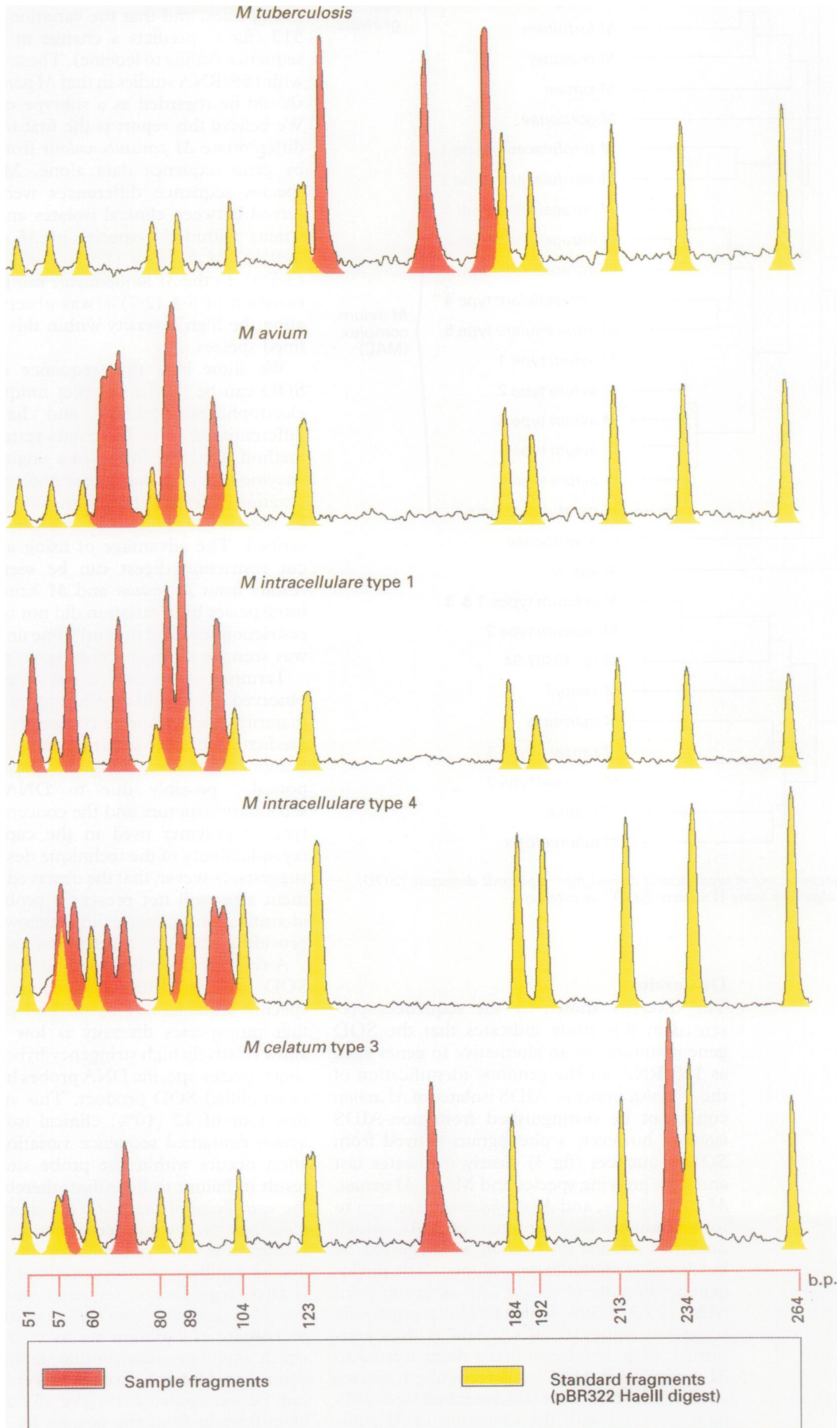


*Figure 4 Capillary electrophoresis profiles of amplified superoxide dismutase gene HaeIII restriction digests from mycobacterium species.*

*Restriction fragment sizes of HaeIII digests of superoxide dismutase from mycobacteria measured by capillary electrophoresis*

| Taxon | Restriction fragment size (bp) | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 70 | 71 | 74 | 77 | 83 | 84 | 87 | 90 | 93 | 98 | 104 | 131 | 134 | 138 | 144 | 145 | 147 | 156 | 160 | 167 | 175 | 179 | 191 | 224 | 227 |
| *M avium* type 1–5 & *M paratuberculosis* | ■ | ■ | | | | | ■ | | | ■ | | | | | | | | | | | | | | | |
| *M celatum* type 1 & 3 *M szulgai* | | ■ | | | | | | | | | | | | | | | | ■ | | | | | | | ■ |
| *M celatum* type 2 | | ■ | | | | | | | | | | | | | | ■ | | ■ | | | | | | | |
| *M tuberculosis* | | | | | | | | | | | | ■ | | | | | | | ■ | | | ■ | | | |
| *M intracellulare* type 1,2 & 3 | | ■ | | | ■ | | ■ | | | ■ | | | | | | | | | | | | | | | |
| *M intracellulare* type 4 | | | | | | | ■ | | | ■ | | ■ | | | | | | | | | | | | | |
| *M intracellulare* type 5 | | ■ | | | | | ■ | | | ■ | | | | | | | | | | | | | | | |
| *M scrofulaceum* type 1 | | ■ | | | ■ | | ■ | | | | | | | | | | | | ■ | | | | | | |
| *M scrofulaceum* type 2 | | ■ | | | | | | | | | | | | | | | | | ■ | ■ | | | | | |
| *M gordonae* | | | | | | | | | | | | | | | | | | | | | | | | | |
| *M chelonae* | | ■ | | | ■ | | ■ | | | | | | | | | | | | | | | | | | |
| *M malmoense* | | | | | | | ■ | | ■ | | ■ | | | | | | ■ | | | | | | | | |
| *M marinum* | | | | ■ | | | | | | | | | | | | | | | ■ | | | | ■ | | |
| *M kansasii* type 1 & 2 | | | | ■ | | | | | | | | | | | ■ | | | | | | | | | | ■ |
| *M phlei* | | | | | | | | | | | | | | | | | | | | | | ■ | | | |
| *M sp.* 40407/94 | | | | | | | | | | | | | ■ | | | | | | ■ | | ■ | | | | |
| *M simiae* | | | | | | | ■ | | | ■ | | | | | | | | | | | | | | ■ | |
| *M fortuitum* | | ■ | | | | ■ | | ■ | | | | | | | | | | | | | | | | | |
| *M smegmatis* | | | ■ | | | | ■ | | | | ■ | | | | | | | | | | | | | | |
| *M xenopi* | | ■ | | | | | | | | | | | | ■ | | | | | | | | ■ | | | |

sequence variation into those corresponding to another species.

The need for rapid detection of small numbers of bacteria present in clinical samples requires an efficient DNA extraction technique to maintain sensitivity. The nature of the mycobacterial cell wall in species such as *M intracellulare* and *M avium* makes it difficult to achieve lysis efficiently without the use of long and technically demanding procedures.[18] Rapid detection techniques have therefore tended to be concentrated on the more easily lysed *M tuberculosis*. We have developed a novel, simple, and rapid method for DNA extraction which, in conjunction with capillary electrophoresis profile analysis of the amplified SOD gene, enables identifications to be carried out on all mycobacterial species within 24 hours. Preliminary work on clinical samples using extraction techniques and PCR conditions described here indicate that our SOD primers are genus specific. We therefore believe that this method will be useful for rapid differential diagnosis of mycobacteria when applied directly to clinical samples.

The authors would like to thank Mr M D Yates of the Regional Centre for Tuberculosis Bacteriology, Dulwich Hospital, London, for his help in providing and identifying biochemically some of the clinical isolates used in this study. This work also benefited from the use of the SEQNET facility (Network No 148.79.64.2 1).

1 Wayne LG, Krichevsky MI, Portyrata D, Jackson CK. Diagnostic probability matrix for identification of slow growing mycobacteria in clinical laboratories. *J Clin Microbiol* 1984;**20**:722–9.
2 Crawford JT. Development of rapid techniques for identification of *M. avium* infections. *Res Microbiol* 1994;**145**: 177–80.
3 Van de Vliet GME, Schukkink RAF, Van Gemen B, Schepers P, Klatser PR. Nucleic acid sequence-based amplification (NASBA) for the identification of mycobacteria. *J Gen Microbiol* 1993;**139**:2423–9.
4 White TJ, Madej R, Pershing DH. The polymerase chain reaction: clinical applications. *Adv Clin Chem* 1992;**29**: 161–96.
5 Bull TJ, Shanson DC. Rapid misdiagnosis by *Mycobacterium avium-intracellulare* masquerading as tuberculosis in PCR/DNA probe tests. *Lancet* 1992;**340**:1360.
6 Masur H. US Public health service task force on prophylaxis and therapy for *Mycobacterium avium* complex. Recommendations on prophylaxis and therapy for disseminated *Mycobacterium avium* complex disease in patients with AIDS. *N Engl J Med* 1993;**329**:896–904.
7 Jacobson MA, Yajko D, Northfelt D, Charlebois E, Gary D, Brosgart C, *et al.* Randomised, placebo-controlled trial of rifampin, ethambutol and ciprofloxacin for AIDS patients with disseminated *Mycobacterium avium* infection. *J Infect Dis* 1993;**168**:112–9.
8 Zhang Y, Lathigra R, Garbe T, Catty D, Young D. Genetic analysis of superoxide dismutase, the 23 kilodalton antigen of *Mycobacterium tuberculosis. Mol Microbiol* 1991;**39**:381–91.
9 Rogall T, Flohr T, Böttger EC. Differentiation of *Mycobacterium* species by direct sequencing for amplified DNA. *J Gen Microbiol* 1990;**136**:1915–20.
10 Thangaraj HS, Lamb I, Davis EO, Jenner PJ, Jeyakumar LH, Colston MJ. Identification, sequencing and expression of *Mycobacterium leprae* superoxide dismutase, a major antigen. *Infect Immun* 1990;**58**:1937–42.
11 Thangaraj HS, Lamb FI, Davis EO, Colston MJ. Nucleotide and deduced amino acid sequence of *Mycobacterium leprae* manganese superoxide dismutase. *Nucleic Acids Res* 1989; **17**:8378.
12 Jukes TH, Cantor CR. Evolution of protein molecules. In: Munro HN, ed. *Mammalian protein metabolism*, vol 3. New York: Academic Press, 1969;21–132.
13 Ho YS, Crapo JD. Isolation and characterization of complementary DNAs encoding human manganese-containing superoxide dismutase. *FEBS Lett* 1988;**229**: 256–60.
14 Bull TJ, Shanson DC, Yates MD, Hamid ME, Minnikin DE. A new type of *Mycobacterium celatum* isolated from AIDS patients in the London area. *Int J Syst Bacteriol* (in press).
15 Butler WR, O'Connor SP, Yakrus MA, Smithwick RW, Plikaytis BB, Moss CW, *et al. Mycobacterium celatum* sp. nov. *Int J Syst Bacteriol* 1993;**43**:539–48.
16 Wenz HM. Capillary electrophoresis as a technique to analyse sequence-induced anomalously migrating DNA fragments. *Nucleic Acids Res* 1994;**22**:4002–8.
17 Zolg JW, Philippi-Schulz S. The superoxide dismutase gene, a target for detection and identification of mycobacteria by PCR. *J Clin Microbiol* 1994;**32**:2801–12.
18 Rastogi N, Barrow WW. Cell envelope constituents and the multifaceted nature of *Mycobacterium avium* pathogenicity and drug resistance. *Res Microbiol* 1994;**145**:243–52.