

Appendix S1.**MATERIALS AND METHODS (19–32)***Ethical considerations*

Written informed consent was obtained from all subjects, and the study was reviewed and approved by the institutional ethics review board of the Institute of Dermatology, Chinese Academy of Medical Sciences and Peking Union Medical College, Nanjing, Jiangsu, China (No. KYZR2009-016).

Study population, sample collection and strain isolation

From July 2002 to March 2013, 30 patients were diagnosed with CTB after seeking clinical care for dermatological conditions at the Institute of Dermatology, Chinese Academy of Medical Sciences and Peking Union Medical College, Nanjing, China. These patients originated from Jiangsu ($n=15$) and Anhui ($n=15$) provinces. The diagnosis of CTB was based on a combination of clinical, histopathological, and laboratory features, as described previously (6). Clinical isolates had been cultured from skin specimens and banked in a CTB repository. All microbiological culture procedures were performed one specimen at a time, in biological safety cabinets equipped with a negative-pressure system.

Fifteen of the 30 strains were from patients diagnosed during the period 2002 to 2007, as described previously (6). Fifteen strains from subsequently diagnosed CTB patients were added. An ongoing epidemiology database, consisting of patients' age, sex, clinical features, and findings from histopathology, bacteriology, and routine blood and urine tests, has been created. Results from PPD screens, HIV antibody analysis, liver and renal function analysis, and chest radiography are included.

Skin tissue homogenates from each patient suspected of TB infection were inoculated on Löwenstein–Jensen (L–J) medium. Two growth temperatures were tested in parallel by incubating the inoculated media at 32°C and 37°C for at least 1 month and up to 2 months. Ziehl–Neelsen staining was used to confirm the cultured organisms as acid-fast bacilli (AFB).

Genomic DNA extraction and drug resistance gene testing

Mycobacterial genomic DNA was extracted from a loop of colonies growing on L–J medium. The bacterial cells were resuspended in 200 µl TE buffer (10 mM Tris–HCl, 1 mM EDTA). The suspensions were frozen in liquid nitrogen and transferred to boiling water and held for 1 min. The freeze–thaw steps were repeated 4 more times to release genomic DNA. After centrifugation of the suspension, the supernatant fluid containing genomic DNA was obtained and stored at –20°C until further genotyping analysis. Standard TB strain H37Rv, obtained from the World Health Organization (WHO) was used as a control for all microbiological and genetic procedures. PCR amplification of *M. tuberculosis* specific *hsp65* (19) and *16S rRNA* (20, 21) coding sequences were performed to confirm that the isolated pathogen was *M. tuberculosis*.

From a bank of DNAs prepared from clinical PTB strains collected during the period 2002 to 2013, 58 were randomly selected and kindly provided by Shanghai Hospital. These 58 DNAs

served as laboratory controls for PCR amplification and baseline interpretation of CTB molecular strain typing and drug resistance results. The drug resistance determining regions (DRDR) within *katG*, *inhA*, *rpoB*, *embB*, *pncA*, *rpsL*, and *rrs* genes were amplified (22–26). Fifteen of the 58 PTB patients originated from Jiangsu and the remainder from 4 other provinces.

Identification of lineage

RD105 is a marker of Lineage 2 of *M. tuberculosis*. This lineage is composed mainly of Beijing family strains. The presence or absence of RD105 was analysed using PCR. PCR products were resolved on 2% agarose gel and sized using a 100-bp DNA ladder (Thermo Scientific, Loughborough, UK). An amplified fragment length of 300 bp is a diagnostic feature for Beijing family, as described previously (16).

Spoligotyping of the isolates was performed as described previously (9). In brief, genomic DNA of *M. tuberculosis* isolates was amplified using PCR with the primers Dra (biotin-labelled) and Drb. The PCR products were subsequently hybridized to a set of 43 oligonucleotide probes corresponding to each spacer. Spoligotypes (STs) in binary format were compared against those in the SpolDB4 database.

To identify suitable MIRU–VNTR loci for genotyping *M.tb* in this area, 1 panel consisting of 24 loci was assembled. The panel included all but Mtub 29 and Mtub 34 of the 24 locus VNTR panel (<http://www.miru-vnrplus.org/MIRU/index.faces>), i.e. 10 MIRU–VNTR loci (MIRU-2, -10, -16, -20, -23, -24, -26, -27, -39, and -40), 5 Mtub loci (Mtub-04, -21, -30, -38, and -39), 6 exact tandem repeats (ETRs; ETR-A, -B, -C, -D, -E, and -F), and 3 Queen's University of Belfast (QUB) loci (QUB-11b, -26, and -4156c). The primers for amplification of each locus have been described in previous studies (27–30). PCR products were resolved on 2% agarose gel and sized against a 100-bp DNA ladder. The images of ethidium bromide-stained gels were captured, and the copy number of each locus was calculated using Image Lab software (Bio-Rad, Richmond, CA, USA).

The discriminatory power of each of the VNTR loci was determined using Hunter–Gaston Discriminatory Index (HGDI) (31). The BioNumerics 7.0 (Applied Maths, St-Martens-Latem, Belgium) software package was used to generate a dendrogram based on 24 VNTR loci and ST of the CTBs and PTBs using the unweighted-pair group method using average linkages (UPGMA), while clustering and visualizing were possible with metric multi-dimensional scaling (MDS). An MDS plot was generated with the isoMDS function in the MASS library (32) including both MIRUs and spoligotype. The MIRU–VNTR plus online web application was used to find strain type relatedness of the CTB among the 58 PTB and a reference global panel of MTBC isolates by the Neighbor-Joining method.

Statistical analyses

SPSS 19.0 (SPSS Science, Chicago, IL, USA) was used for statistical analysis. Upon comparison of the demographic, clinical and genetic characteristics of the 2 groups (CTB vs. PTB), significant differences were determined by unpaired *t*-test and χ^2 test. Statistical significance was defined as a *p*-value < 0.05.