48 Hormaeche, C. and Khan, C. (1996) Recombinant bacteria as vaccine carriers of heterologous antigens. In *Concepts in Vaccine Development*. (Kaufmann, S.H.E., ed.) pp. 327–349, Walter de Gruyter
49 Rüssmann, H. *et al.* (1998) Delivery of epitopes by the *Salmonella* type III secretion system for vaccine development. *Science* 281, 565–568
50 Konieczny, M.P.J. *et al.* (2000) Cell surface presentation of recombinant (poly)peptides including functional T-cell epitopes by the AIDA autotransporter system. *FEMS Immunol. Med. Microbiol.* 27, 321–332
51 Spreng, S. *et al.* (1999) The *Escherichia coli* haemolysin secretion apparatus: a potential universal antigen delivery system in Gram-negative bacterial vaccine carriers. *Mol. Microbiol.* 31, 1596–1598
52 Hess, J. *et al.* (1996) Superior efficacy of secreted over somatic antigen display in recombinant *Salmonella* vaccine induced protection against listeriosis. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1458–1463
53 Ryan, E.T. *et al.* (1997) Protective immunity against *Clostridium difficile* toxin A induced by oral immunization with a live, attenuated *Vibrio cholerae* vector strain. *Infect. Immun.* 65, 2941–2949
54 Hess, J. *et al.* (1997) Protection against murine listeriosis by an attenuated recombinant *Salmonella typhimurium* vaccine strain that secretes the naturally somatic antigen superoxide dismutase. *Infect. Immun.* 65, 1286–1292
55 Donner, P. *et al.* (1998) Use of a secretion vector for fertility control by oral vaccination. European Patent 1015023A1.
56 Hahn, H.P. *et al.* (1998) *Salmonella typhimurium* strain genetically engineered to secrete effectively a bioactive human interleukin (hIL)-6 via the *Escherichia coli* hemolysin secretion apparatus. *FEMS Immunol. Med. Microbiol.* 20, 111–119
57 Umelo-Njaka, E. *et al.* (2001) Expression and testing of *Pseudomonas aeruginosa* vaccine candidate proteins prepared with the *Caulobacter crescentus* S-layer protein expression system. *Vaccine* 19, 1406–1415
58 Galen, J.E. and Levine, M.M. (2001) Can a 'flawless' live vector vaccine strain be engineered? *Trends Microbiol.* 9, 372–376
59 Gomez-Duarte, O.G. *et al.* (2001) Expression, extracellular secretion, and immunogenicity of the *Plasmodium falciparum* sporozoite surface protein 2 in *Salmonella* vaccine strains. *Infect. Immun.* 69, 1192–1198
60 Su, G.F. *et al.* (1992) Extracellular export of Shiga toxin B-subunit/haemolysin A (C-terminus) fusion protein expressed in *Salmonella typhimurium aroA*-mutant and stimulation of B-subunit specific antibody responses in mice. *Microb. Pathog.* 13, 465–476
61 Hess, J. *et al.* (2000) Protection against murine tuberculosis by an attenuated recombinant *Salmonella typhimurium* vaccine strain that secreted the 30 kDa-antigen of *Mycobacterium bovis* BCG. *FEMS Immunol. Med. Microbiol.* 27, 283–289
62 Mollenkopf, H. *et al.*, (2001). Protective efficacy against tuberculosis of ESAT-6 secreted by *Salmonella typhimurium* vaccine carrier strain and expressed as naked DNA. *Vaccine* 19, 4028–4035
63 Schlör, S. *et al.* (1997) *In vivo* and *in vitro* studies on interactions between the components of the hemolysin (HlyA) secretion machinery of *Escherichia coli*. *Mol. Gen. Genet.* 256, 306–320
64 Gentschev, I. *et al.* (1998) Delivery of the p67 sporozoite antigen of *Theileria parva* using recombinant *Salmonella*: secretion of the product enhances specific antibody responses in cattle. *Infect. Immun.* 66, 2060–2064
65 Spreng, S. *et al.* (2000) *Salmonella* vaccines secreting measles virus epitopes induce protective immune responses against measles virus encephalitis. *Microbes Infect.* 2, 1687–1689

Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains

Pablo J. Bifani, Barun Mathema, Natalia E. Kurepina and Barry N. Kreiswirth

A large, genetically related group of *Mycobacterium tuberculosis* strains, variously called W or Beijing, is distinguished by specific molecular markers and referred to as the W-Beijing family strains. Molecular epidemiological studies suggest that these strains are highly prevalent throughout Asia and the countries of the former Soviet Union and they have also been reported in several other geographical regions, including North America. Although the spread of W-Beijing family strains in diverse populations is well documented, the underlying host–pathogen factors accounting for their continued dissemination and burden of disease have yet to be determined.

Advances in our understanding of the molecular biology of *Mycobacterium tuberculosis* have proven invaluable in unraveling the epidemiology of tuberculosis (TB) and in constructing the phylogenetic structure of the species. Genomic data have revealed remarkable DNA sequence conservation between chromosomes, and the paucity of synonymous mutations led to the hypothesis that *M. tuberculosis* is a recent human pathogen dating back approximately 15 000 years [1]. Although the *M. tuberculosis* genome is highly conserved in relation to other bacterial species, there are polymorphic regions that are usually associated with insertion sequences and/or repetitive elements and it is these variable regions that form the basis of modern TB genotyping.

Several polymorphic or hypervariable genetic markers have been characterized that, together, can discriminate or sub-speciate clinical isolates of *M. tuberculosis.* The most widespread and robust genotyping tool is the insertion sequence IS*6110*, a member of the IS*3* family of transposable elements [2] that is specific for strains belonging to the *M. tuberculosis* complex [3]. Although noted 'hot spots' have been identified, IS*6110* is more or less randomly distributed around the chromosome and its copy number ranges from the rare clones that lack an insertion to strains that have 26 copies [4,5]. The standardization of a method for IS*6110* Southern blot hybridization has created 'DNA fingerprints' that can be compared between laboratories with the aid of pattern-matching software [6]. As a result, >50 000 *M. tuberculosis* isolates worldwide have been

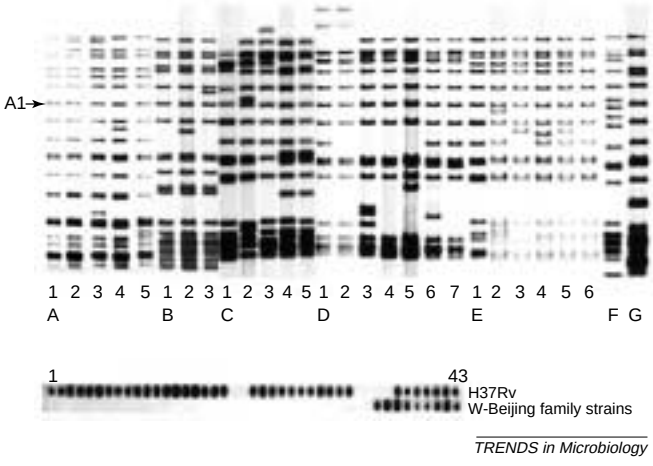
 0966-842X/01/$ – see front matter PII: S0966-842X(01)02277-6

Box 1. Key molecular characteristics that define the W-Beijing family strains

- Principal genetic group 1 [*katG* codon 463 CTG (Leu) and *gyrA* codon 95 ACC (Thr)] [a,b].
- Empirical IS*6110* banding pattern similarity to strain W (copy number range from 15–26).
- Spoligotyping denoted S00034, presence of nine spacers from 35 to 43 [c].
- IS*6110* insertion A1 in the origin of replication (corresponding to 3.36 kb band on a Southern hybridization blot probed with *dnaA-dnaN*) [b].
- One or two IS*6110* insertions (two for strain W) in NTF region, demonstrated by multiplex PCR [b,d].

All W-Beijing family strains share the above molecular profile including W4, W14, W82, strain 210, W148, Beijing members and others (Fig. I). Additional and distinguishing markers specific for NYC strain W and its progenies:

- Second IS*6110* insertion in NTF region (head-to-tail arrangement) [d].
- Rare dinucleotide change in codon 315 of *katG* (AGC→ACA) [e].



References

a Sreevatsan, S. *et al.* (1997) Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc. Natl. Acad. Sci. U. S. A.* 94, 9869–9874

b Kurepina, N.E. *et al.* (1998) Characterization of the phylogenetic distribution and chromosomal insertion sites of five IS*6110* elements in *Mycobacterium tuberculosis*: non-random integration in the *dnaA-dnaN* region. *Tubercle Lung Dis.* 79, 31–42

c Bifani, P.J. *et al.* (1999) Identification of a W variant outbreak of *Mycobacterium tuberculosis* via population-based molecular epidemiology. *J. Am. Med. Assoc.* 282, 2321–2327

d Plikaytis, B.B. *et al.* (1994) Multiplex PCR assay specific for the multidrug-resistant strain W of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* 32, 1542–1546

e Bifani, P.J. *et al.* (1996) Origin and interstate spread of a New York City multidrug-resistant *Mycobacterium tuberculosis* clone family. *J. Am. Med. Assoc.* 275, 452–457

Fig. I. Top panel: IS*6110* Southern blot hybridization DNA fingerprint patterns of W-Beijing family *Mycobacterium tuberculosis* strains isolated from tuberculosis (TB) patients in different studies [Public Health Research Institute (PHRI) TB Center collection]. A, lanes 1–5: representative members of the W4 strain group from a population-based study, New Jersey [15]; B, lanes 1–3: representative members of the W14 strain group from a community cluster in New York City (NYC) [46]; C, lanes 1–5: representative members of the W-Beijing family isolated in China [5,28]; D, lanes 1–7: representative members of the W family strains isolated in the former Soviet Union [41,42]; E, 1–6: Multidrug resistant strain W and descendants isolated from the NYC outbreak [27]; F: W family strain W82 identified in nosocomial transmission study, Tennessee [57]; G, W family strain 210 found in California/Colorado/Texas outbreak [45]. The A1 arrow indicates the IS*6110* insertion (common to all W-Beijing family strains) in the *dnaA-dnaN* region of the *M. tuberculosis* chromosome. Bottom panel: spacer oligonucleotide typing (spoligotyping) patterns of H37Rv (control) and W-Beijing family strains (pattern S00034).

genotyped using IS*6110*. Although highly reliable, IS*6110* fingerprinting does have some limitations, such as poor discrimination between isolates with a low number (≤5 copies) of insertions and in determining genetic relatedness among similar patterns [7]. This method alone cannot distinguish among strains belonging to the *M. tuberculosis* complex. Finally, IS*6110* fingerprinting is also limited by the slow growth of *M. tuberculosis,* as a viable culture is required to produce an accurate genotype.

Consequently, several other genotyping techniques have been developed, including spacer oligonucleotide typing (spoligotyping) [8], variable-number tandem repeat typing (VNTR) [9], polymorphic GC-rich repetitive sequence typing (PGRS) [10] and, more recently, mycobacterial interspersed repetitive units typing (MIRU) [11]. These techniques are not discussed in this review; however, a comprehensive evaluation of the methods of genotyping *M. tuberculosis* has been published elsewhere [12].

Incorporating genotyping into traditional TB epidemiology has been instrumental in confirming suspected outbreaks [13,14], in identifying cases of unsuspected transmission [15], in tracking cases of laboratory cross-contamination [16], in discriminating exogenous versus endogenous disease [17,18] and in demonstrating the occurrence of exogenous superinfection in both immunocompetent and immunocompromised patients [19–21]. Genotyping tools have also been used to evaluate transmission dynamics in specific populations and in defined geographical settings [22,23]. Overall, these epidemiological investigations and studies have used molecular techniques to determine *M. tuberculosis* transmission, based on whether two clinical isolates are identical or different.

Most outbreak- and population-based investigations have identified a prevalent strain that has spread from person to person (clonal spread). This ultimately gives rise to strain variants that are detected on the basis of their subtle genetic changes. Together, a collection of closely related strains derived from a common ancestor is classified as a group and, over time, progenies derived from a group can subsequently disseminate in new populations (clonal expansion), giving rise to a second group of related strains. The identification of genetic

Pablo J. Bifani
U447–Mécanismes Moléculaires de la Pathogénie Bactérienne, Institut Pasteur de Lille – IBL. 1, rue du Professeur Calmette, BP245 – 59019 Lille cedex, France.

Barun Mathema
Natalia E. Kurepina
Barry N. Kreiswirth*
PHRI TB Center, Public Health Research Institute, 455 First Ave, New York, NY 10016, USA.
*e-mail: barry@phri.org

Table 1. Selected studies concerning the global distribution of *M. tuberculosis* W-Beijing family strains[a]

Geographical region	Study design	Description	Ref.
Hong Kong	Prospective study	266 isolates consisting of pretreatment and last sample obtained during chemotherapy. Acquisition of additional IS*6110* in one patient.	[37]
French Polynesia	Cross-sectional	Analysis of 72 isolates from 64 patients identified one from 11 clusters that belong to the W-Beijing family.	[56]
East Asia: China, Mongolia, S. Korea, Thailand	Convenience sample	IS*6110*-based analysis: 'Beijing family' was found predominant among strains tested (China 86%, Mongolia 50%, S. Korea 43%, Thailand 37%). Postulated that these strains are endemic in the region.	[28]
New York City, USA	Outbreak investigation	253 of the 357 MDR strains shared identical 18-band or similar IS*6110* pattern, strain W. Has since been considered the 'index' W strain at PHRI TB Center.	[27]
Taiwan	Cross-sectional	90 isolates representative of 25% of all cases diagnosed at Mackay hospital, Taiwan.	[38]
South Carolina, USA	Nosocomial investigation	IS*6110*-based analysis: nosocomial transmission of MDR strain W1 via contaminated bronchoscope.	[30]
Thailand	Cross-sectional	80/211 isolates from three referral hospitals belong to the W-Beijing family.	[36]
Havana, Cuba	Cross-sectional/prevalence	23/160 TB specimens isolated over a one-year period belong to the W-Beijing family.	[54]
Peru	Convenience sample	IS*6110* and principal genetic group data support a subgroup of isolates, isolated from Peruvian patients, belonging to the W-Beijing family.	[66]
Tennessee, USA	Nosocomial investigation	Nosocomial transmission of a W-Beijing family strain to a healthcare worker and a patient exposed to active TB.	[57]
New York City, USA	Nosocomial investigation	Transmission in a hospital nursery. The source case is believed to be MDR strain W1, a direct descendant of the index MDR strain W.	[58]
California; Colorado; Texas, USA	Population-based study	57 W-Beijing family strains (labeled as strain 210) isolated in three US states. Intrastate patient links were documented with possible interstate transmission	[44]
St Petersburg, Russia	Convenience sample	IS*6110*-based analysis: 15/27 isolates empirically are related to W-Beijing family strains.	[40]
Barcelona, Spain	Routine surveillance	IS*6110*-based analysis: MDR-TB caused by W-Beijing family strains in three immunocompetent patient immigrants from Peru.	[59]
China	Retrospective study on preserved lung tissue	Analysis of 85 paraffin-embedded lung biopsies dated from 1956–1990. 45/49 samples belong to the W-Beijing family.	[60]
New Jersey, USA	Population-based study	43 closely related W-Beijing family clones were found prevalent in one county. Patients shared demographic profiles, however have no known patient links, suggesting both historic and recent transmission.	[15]
Malaysia	Convenience sample	IS*6110*-based analysis: W-Beijing family strains represent 10–19% of all TB isolates analyzed.	[47]
Cape Town, South Africa	Outbreak investigation	Community outbreak of MDR TB identified 17/21 isolates having identical IS*6110* pattern resembling strain W from New York City. None of the patients were HIV-seropositive or institutionalized.	[20]
Western Siberia, Russia	Cross-sectional (operational study)	IS*6110*-based analysis: pattern identical in another Russian study [42] and to the PHRI TB Center samples (D in Fig. I).	[41]
USA	Surveillance study	Surveillance identified strain W in nine states and Puerto Rico.	[24]
Guadeloupe	Population-based study	W-Beijing family strains were identified among 95 TB samples analyzed based on spoligotyping.	[61]
Texas, USA	Population-based study	326 from a total of 1283 isolates were found to belong to the W-Beijing family.	[62]
Vietnam	Cross-sectional/population-based study	IS*6110*-based analysis, spoligotyping: W-Beijing family strains were identified in 301 (54%) of the 563 patients. Cases were associated with younger age, STR^R and INH^R and BCG vaccination.	[35]
Buenaventura, Colombia	Retrospective study	11/111 isolates identified strains related to strain W.	[63]
Iran	Convenience sample	IS*6110*-based analysis, spoligotyping: 6/62 TB patients showed similarity with W-Beijing family strains.	[64]
Israel	Population-based study	IS*6110*-based analysis: 2/13 isolates from TB patients empirically are related to W-Beijing family.	[65]
Thailand	Prospective study	IS*6110*-based analysis, spoligotyping: W-Beijing family strains was found in 90 (44.1%) out of 204 isolates.	[48]
New York City, USA	Convenience sample	IS*6110*-based analysis, spoligotyping, additional molecular typing: 26/13 000 isolates grouped into W14 cluster. All 26 were STR^R. Demographic homogeneity of the patients demonstrated, with no known epidemiological links.	[46]
Jakarta, Indonesia	Prospective study	IS*6110*-based analysis, spoligotyping: 31 (34%) of the 92 patients demonstrated W-Beijing family strains. No relation was found between genotype and BCG vaccination status.	[39]

[a]Abbreviations: BCG, *Mycobacterium bovis* Bacille–Calmette Guérin; INH^R, isoniazid resistant; MDR, multidrug resistant; PHRI, Public Health Research Institute; spoligotyping, spacer oligonucleotide typing; STR^R, streptomycin resistant; TB, tuberculosis.

markers common to members of multiple groups provides a sound framework to classify these strains more broadly into a family structure. One group of TB isolates collectively known as the W or Beijing family has been extensively reported in the TB literature. Here, we review the biological, clinical and

epidemiological data surrounding the clonal expansion and apparent success of the *M. tuberculosis* isolates that are collectively referred to in this review as the W-Beijing family strains.

Epidemiology of the W-Beijing family strains

Members of the W-Beijing family have been responsible for considerable morbidity and mortality worldwide [24–26]. Their epidemiological settings are diverse in that they have been reported to cause outbreaks in institutional and nosocomial environments and have also been implicated in ongoing community transmission, contributing to their endemic nature. Both pan-susceptible and resistant members have been successful in disease transmission. Although strain W was characterized in New York City (NYC) [27], a similar group of predominant isolates was identified in the Far East [28] and named accordingly as the Beijing isolates.

Current molecular data, as we will discuss, indicate that isolates typed to either the W or the Beijing strains are descendants of a common ancestral strain and form a large branch in the *M. tuberculosis* phylogenetic lineage [5]. The genotyping data also defined the chromosomal markers common to the W and Beijing strains and those markers that distinguish the various branches within the W-Beijing family that have evolved in different geographical regions. Here, we profile the genetic markers that define members of the W-Beijing family and highlight several studies that have documented the spread of selected branches within this phylogenetic lineage.

Nosocomial outbreak involving a multidrug resistant strain, strain W

In NYC, during the 43 months from January 1990 to August 1993, a highly multidrug resistant (MDR) strain that was invariably untreatable with streptomycin, isoniazid, rifampin, ethambutol and, in most instances, kanamycin, spread rapidly in an HIV-seropositive population and 357 cases were reported in prisons and hospitals [26,27,29]. DNA fingerprinting of 253 of these patient isolates revealed the clonal spread of a strain with an 18-band IS*6110* pattern that was arbitrarily assigned the name strain W.

Since 1991, >500 cases of TB in NYC have been linked to the clonal spread of strain W (B.N. Kreiswirth *et al.*, unpublished). Analysis of 420 isolates from these cases showed first-line resistance to anti-mycobacterial drugs, and each had the genotype described in Box 1. Progenies of this clone have spread well beyond NYC, with organisms recovered from patients in other regions of New York, Atlanta, Denver, Las Vegas, Miami and Paris, France [27]. Most of the patients could be traced back to contacts in New York hospitals and prisons. In one case, nosocomial transmission documented in South Carolina was the result of a contaminated bronchoscope [30]. Strains related to strain W have been reported in more than ten other states [24].

The genetic analysis showed that over the past ten years the IS*6110* pattern that defines the outbreak strain W has evolved (or diverged) by acquiring or losing an IS*6110* element (panel E, Fig. I), as well as developing additional drug resistance. The clonality of strain W and its progenies was confirmed using a combination of molecular markers, including sequencing of drug target genes [27]. Notably, strain W has a rare dinucleotide change in codon 315 of *katG*, the catalase-peroxidase gene, which is linked to isoniazid resistance. Mutations in codon 315 are the most common genetic alteration associated with isoniazid resistance among clinical isolates, but the dinucleotide substitution (AGC→ACA; Ser→Thr) is unique to NYC strain W and its recent descendants [31].

All strain W isolates and progenies have two copies of IS*6110* in a head-to-tail orientation in the NTF chromosomal region, which are used as the target of a multiplex assay to genotype strain W [5,32], and all share an identical array of mutations in five drug resistance target genes associated with first-line anti-mycobacterial resistance [27]. The *katG* 315 dinucleotide mutation and two IS*6110* insertions in the NTF locus constitute two independent identifiers for NYC strain W and its descendants, but not for other members of the W-Beijing family. Together, these results supported the hypothesis that the MDR strain W outbreak was the result of a primary resistant clone that developed total first-line resistance before its spread in prisons and hospitals. It is noteworthy that fluoroquinolone resistance was acquired repeatedly; this resistance developed independently after the original NYC nosocomial outbreak and a total of five different mutations in *gyrA* have been identified [33].

A decade since the initial outbreak, a total of 11 descendants of strain W have been identified that have subtle differences in IS*6110* fingerprint patterns and that reflect the microevolution of this clone over time (panel E in Fig. I shows six representative patterns). In each case, these MDR strains have the complete complement of genetic markers that define the strain W outbreak as we have described.

Strain W and the W-Beijing family strains

The molecular markers used to type strain W can also be used to group a large collection of related *M. tuberculosis* strains that have spread successfully throughout East Asia and the United States. Collectively, these isolates have distinct molecular features, which, in combination, differentiate them from all other *M. tuberculosis* samples analyzed to date. Although these isolates have a large array of IS*6110* patterns with a high copy number (15–26), they are genetically related to each other on the basis of several independent genetic markers: they each belong to principal genetic group 1 [1], they have an IS*6110* insertion in the origin of replication (A1 insertion) [5] and they have the same spoligotype pattern (S00034) [15].

All W-Beijing family isolates have a single IS*6110* insertion in the NTF region; this genotype also distinguishes these strains from the multidrug resistant NYC strain W and its descendants, which have two copies of IS*6110* in this region (Box 1). The Public Health Research Institute (PHRI) TB Center database (>14 000 *M. tuberculosis* isolates) contains >2000 patient isolates represented by at least 450 different yet similar IS*6110* DNA fingerprint profiles that have been typed to the W-Beijing family. Currently, all members of the W-Beijing family can be segregated from all other clinical isolates on the basis of the genetic markers we have described here. Although other *M. tuberculosis* strains have been grouped to a 'strain family', none has been analyzed to the extent of the W-Beijing family or has been so successful in causing extensive outbreaks.

The strain families that have been cited in the literature, such as the Haarlem family strains [7] and the common laboratory strain H37 and its derivatives [34], have been initially catalogued on the basis of their similar IS*6110* DNA-fingerprint patterns. In these examples, the Haarlem family and H37 are broadly distinguished from the W-Beijing family in being members of principal genetic groups 2 and 3, respectively, and having different secondary-typing patterns.

Like the W-Beijing family, the Haarlem family strains appear to be widespread and have been isolated from patients in Europe, Asia and the Americas, and both spoligotyping and VNTR analysis have provided secondary-typing data to support their genetic relatedness. However, unlike the W-Beijing family strains, which have been further subtyped to groups, no unique chromosomal markers have yet been defined to further distinguish the Haarlem family.

Geographical distribution of the W-Beijing family strains

The common use of IS*6110* genotyping and secondary-typing methods, especially spoligotyping, has provided researchers with the basic tools to differentiate the W-Beijing family strains from all other *M. tuberculosis* complex strains. Consistent with the first reports that coined the name the Beijing family and that suggested these strains were common in Mongolia and China, the ever-growing reports from across Asia have confirmed and extended this hypothesis. Although many of these studies take a 'snapshot' of the strain population in a given geographical region, as shown in Table 1, the prevalence of the W-Beijing family is remarkably high in Vietnam, Hong Kong, Indonesia, Korea, Thailand and Taiwan [28,35–39]. Recent genetic data from isolates from throughout the former Soviet Union also showed the widespread prevalence of the W-Beijing family strains [40,41].

Among a convenience sampling of 2100 *M. tuberculosis* isolates from regions across the Russian Federation, we identified the W-Beijing family strains in 45% of patients. The most prevalent strain, W148, has a distinct 17-band IS*6110* pattern (panel D1, Fig. I), which has been identified throughout the former Soviet Union. In a TB prison in Western Siberia, between 1998 and 2000, 190 prisoners were identified with multidrug resistant W148 isolates, and drug-resistant genotypes indicated this strain spread as a primary resistant clone [42]. As in the NYC outbreak in the early 1990s, the spread of W148 in the Siberian prison was partly owing to the dismantling of public health infrastructure, an ineffective system for testing drug susceptibility and treating highly resistant cases of TB [43]. The recent PHRI TB Center findings that MDR W148 strains have been isolated in NYC from former Soviet Union immigrants raises concerns about their global spread (B.N. Kreiswirth *et al*., unpublished).

The distribution of the W-Beijing family strains in distinct geographical regions and their ability to predominate and spread in large clonal clusters suggests that members of this phylogenetic lineage are better adapted to infect and cause disease in humans. The prevalence of the W-Beijing family strains in Europe, South Africa, and South and North America is probably the result of the large amount of human migration in the 20th century and the reactivation of TB over time. As described below, the clustering and spread of the W-Beijing family strains common in the Asian continent has now been observed among US-born patients in different regions of the United States. These studies further the hypothesis that the W-Beijing family strains have a genetic advantage to cause disease in the human population.

Spread of the W-Beijing family strains in the United States

The low incidence of TB in the United States, juxtaposed with high immigration rates from high-incidence countries, provides a unique situation to study the transmission dynamics of the W-Beijing family strains. As an example, we evaluated all culture-positive TB cases (76% capture) in the state of New Jersey between January 1996 and September 1998 and, using the criteria established in Box 1, we identified 68 patient isolates out of 1207, or 6%, that met the profile of a member of the W-Beijing family over a 45-month period [15]. The strains were further divided into two groups (A and B) of 43 and 25, based on subtle differences in IS*6110* patterns, and distinct PGRS and VNTR patterns. Although all isolates could be shown to be descendants of a common ancestral strain, isolates from group B were found to be distantly related to one another (22 different IS*6110* patterns from 25 patients) as well as to members of group A. By contrast, group A strains (termed W4 – panel A, Fig. I) were very closely related (seven patterns from 43 cases) and hence belong to a distinct branch of the evolutionary tree of the W-Beijing family. Patients in group A were pan-susceptible, primarily US-born, non-Hispanic blacks and clustered within one county in New Jersey. By contrast, most patients infected with group B strains were non-US-born (76% from East Asian countries), and scattered throughout

New Jersey. In addition, group A patients tended to be younger with risk factors for TB such as HIV infection, whereas group B patients were older with no major risk factors. Therefore, group A strains are believed to be part of an ongoing outbreak that has successfully contributed to the endemic rate in these New Jersey communities, whereas group B are most likely cases of reactive TB that have been imported from Asia.

A study carried out in the western United States by Yang and colleagues [44] identified a prevalent strain designated 210 (panel G, Fig. I). This pan-susceptible strain was recovered from 57 patients out of a total of 1324 isolates in a population-based survey and from three geographically separate regions. The 210 IS*6110* banding pattern resembled that of the W-Beijing family strains. Spoligotyping of this clone revealed the signature S00034 pattern. Further analysis confirmed the presence of the A1 IS*6110* insertion in the *dnaA-dnaN* locus and single insertion in NTF [5,45]. The 57 strains were retrieved from California and Texas where the majority of cases were identified (two cases were from Colorado). Epidemiological data demonstrated widespread intrastate transmission of TB among high-risk groups of predominantly US origin. Although no patient links could be established between states, the molecular data suggest that interstate transmission occurred but historically, with secondary local dissemination. In an earlier study by Barnes *et al.* [22], the largest cluster identified in central Los Angeles – 43 of 96 cases – genotyped as strain 210. The 43 cases in this cluster suggest extensive transmission, some of which has been documented within mainly US-born patients in three homeless shelters.

A recent study demonstrates yet another example of dissemination of a W-Beijing family strain (W14) in a US-born population [46]. A convenience sample of 44% of all cases reported in NYC between 1992 and 1999 identified a cluster of 26 isolates belonging to a distinct group, comprised of three variants, of the W-Beijing family strains (panel B, Fig. I). Besides having the global molecular markers inherent in all W-Beijing family strains (Box 1), the W14 cluster bore subtle but distinctive molecular properties in the spoligopattern and VNTR profiles. In addition, all strains belonging to the W14 group have a mutation in codon 43 of the *rpsL* gene conferring high-level streptomycin resistance. In this cluster, 20/26 (77%) of the isolates were resistant only to streptomycin (mono-streptomycin). Acquisition of secondary drug resistance in progenies of the predominant mono-streptomycin-resistant W14 cluster was determined by drug susceptibility testing and confirmed by DNA sequence analysis of target genes known to be associated with resistance to anti-tubercular antibiotics [31]. Molecular data in the W14 study suggests that streptomycin resistance was acquired before its dissemination. Although no patient links were identified, the demographics were very similar in that they were all US-born and had risk factors for TB. As it was a predominantly young, HIV-seropositive population with no demonstrable patient-to-patient links, this suggests both recent and historic transmission.

The spread of the W-Beijing family strains in East Asia

The Beijing type strains were initially described by van Soolingen and co-workers, who analyzed 69 *M. tuberculosis* isolates from TB cases in Beijing, China and showed that 86% of these isolates had similar IS*6110* patterns (15–20 insertions) and identical spoligotype [28]. Subsequent genotyping analysis showed that these pan-susceptible isolates (panel C, Fig. I) had the A1 IS*6110* insertion in the *dnaA-dnaN* locus and a single IS*6110* copy in the NTF region. The predominance of the Beijing strain type in this small sampling collection indicated possible endemic distribution of these strains in Eastern Asia. This notion has been further supported by several subsequent studies that illustrate the variable frequency of the Beijing type (86% in China; 34% in Indonesia; 25% in Malaysia; 50% in Mongolia; 43% in South Korea; 44% in Thailand; and 53% in Vietnam [28,35,39,47,48]). It is important to note that these data are based on a few representative investigations, which often involve convenience sampling, and are not systematic population-based studies (Table 1). Nonetheless, these investigations are representative of the molecular epidemiological studies undertaken in the Far East and clearly indicate that the W-Beijing family strains are widespread in Asia, leading investigators to hypothesize on possible mechanisms that could account for this success. For example, in Vietnam this group of isolates was associated with active transmission in a primarily *Mycobacterium bovis* Bacille Calmette–Guérin (BCG)-immunized young population, whereas in Indonesia they were correlated with increased risk of febrile response [35,39].

The success of the W-Beijing family strains

The factors that contribute to the success of the W-Beijing family strains have yet to be unraveled. The success could stem from increased transmissibility, stability and/or altered gene expression of as-yet-unidentified virulence factors. At least one *in vivo* study has shown that a W-Beijing family member (strain 210) was able to multiply more rapidly in macrophages when compared with strain CDC 1551 and another clinical isolate [49].

There are several hypotheses to explain the large expansion of the W-Beijing family strains throughout the Asian continent. One hypothesis is that these strains spread as a result of their resistance to the BCG vaccine [35]. However, the notion that the vaccine provided selective pressure that could account for the success of the W-Beijing family strains is inconsistent with their ability to cause outbreaks and clustering in non-vaccinated US-born individuals in various regions of the United States. It should also be noted that the remarkable IS*6110* genetic diversity

Questions for future research

- Are there genetic correlates that give rise to certain phenotypic properties to explain the global prevalence of the W-Beijing family strains?
- Will the recent development of gene-chip and proteomic technologies help identify the genetic factors associated with the success of the W-Beijing family strains?
- How do these phenotypic properties facilitate the dissemination of these strains; are they more infectious, more stable or more virulent?
- Are there unique host–bacteria interactions that elicit immune responses specific to this strain family?
- Could we identify ancestral and/or more distant members of the W-Beijing family, and from this could we unravel the phylogenetic structure of this lineage?

among a largely under-representative collection of the W-Beijing family strains in Russia is good evidence that these strains were successful in this region of the world and spread over generations before the administration of BCG in the 1920s. Nonetheless, it is also conceivable that BCG vaccination might have accelerated the spread of the W-Beijing family in these regions, a hypothesis that remains to be confirmed experimentally and/or epidemiologically.

Alternatively, it is conceivable that expansion of the W-Beijing family in the Asian continent resulted from the introduction of a new pathogen into a naive population. Although there are extensive archeological data demonstrating the presence of TB in ancient Egypt and in the Americas [50–52], there are no scientific reports describing the presence of TB in ancient China.

It has also been hypothesized that the predominance of the W-Beijing family strains results from their reduced susceptibility to anti-TB drugs. There are several examples where resistance has been associated with W-Beijing family strains, such as in Vietnam [35], Germany [53] and Cuba [54]; and where multidrug resistant clones have been shown to cause outbreaks, such as in NYC [27], Estonia [40] and in the Russian prisons [41,42]. However, it should also be noted that many of the outbreak clones, such as strain W4 and 210, and the strains endemic to China and Mongolia, are pan-susceptible.

To date, there are no specific reports that indicate the W-Beijing family strains are hypermutable or that they have unique efflux mechanisms that account for a selective advantage against anti-tubercular therapy. In fact, drug resistance among W-Beijing family strains maps to the same genetic targets as in other resistant *M. tuberculosis* strains, and comparative sequence analysis that included several W-Beijing family strains showed no genetic variation among 50 different structural genes [1,55]. It is therefore likely that the simple explanation surrounding the observed association of drug resistance among the W-Beijing family strains is the consequence of their prevalence in any given population and that resistance, *per se*, does not provide the genetic advantage.

Conclusions

The advent of molecular genotyping has clearly shown that members of the W–Beijing family can rapidly establish a niche and expand in the community. In this regard, these strains present a potential public health threat and measures to identify and monitor these strains are therefore essential. The characteristics listed in Box 1 provide a guide to distinguishing the W-Beijing family strains from the rest of the species and to identify the strain W outbreak clone.

Our understanding of the genetic differences present in the W-Beijing family strains could be furthered once the sequencing of strain 210 is completed. The partial genome is available through the TIGR database (http://www.tigr.org). Currently, the mechanisms underlying the success of the spread of the W-Beijing family strains in the human population are not well understood; however, recent molecular epidemiological findings provide a sound basis on which to include these strains in future studies of pathogenicity and virulence factors.

Acknowledgements
We thank B. Salim, W. Eisner and A. Ravikovitch for their help in editing and manuscript preparation.

References

1 Sreevatsan, S. *et al.* (1997) Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc. Natl. Acad. Sci. U. S. A.* 94, 9869–9874
2 McAdam, R.A. *et al.* (1990) Characterization of a *Mycobacterium tuberculosis* insertion sequence belonging to the IS*3* family. *Mol. Microbiol.* 4, 1607–1613
3 Hermans, P.W. *et al.* (1990) Insertion element IS*986* from *Mycobacterium tuberculosis*: a useful tool for diagnosis and epidemiology of tuberculosis. *J. Clin. Microbiol.* 28, 2051–2058
4 McHugh, T.D. and Gillespie, S.H. (1998) Nonrandom association of IS*6110* and *Mycobacterium tuberculosis*: implications for molecular epidemiological studies. *J. Clin. Microbiol.* 36, 1410–1413
5 Kurepina, N.E. *et al.* (1998) Characterization of the phylogenetic distribution and chromosomal insertion sites of five IS*6110* elements in *Mycobacterium tuberculosis*: non-random integration in the *dnaA-dnaN* region. *Tubercle Lung Dis.* 79, 31–42
6 van Embden, J.D. *et al.* (1993) Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J. Clin. Microbiol.* 31, 406–409
7 Kremer, K. *et al.* (1999) Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. *J. Clin. Microbiol.* 37, 2607–2618
8 Groenen, P.M. *et al.* (1993) Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol. Microbiol.* 10, 1057–1065
9 Frothingham, R. and Meeker-O'Connell, W.A. (1998) Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. *Microbiology* 144, 1189–1196
10 Chaves, F. *et al.* (1996) Usefulness of the secondary probe pTBN12 in DNA fingerprinting of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* 34, 1118–1123
11 Supply, P. *et al.* (2000) Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. *Mol. Microbiol.* 36, 762–771
12 Van Soolingen, D. (2001) Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements. *J. Intern. Med.* 249, 1–26
13 Daley, C.L. *et al.* (1992) An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus. An analysis using restriction-fragment-length polymorphisms. *New Engl. J. Med.* 326, 231–235
14 Edlin, B.R. *et al.* (1992) An outbreak of multidrug-resistant tuberculosis among hospitalized patients with the acquired immunodeficiency syndrome. *New Engl. J. Med.* 326, 1514–1521

15 Bifani, P.J. *et al.* (1999) Identification of a W variant outbreak of *Mycobacterium tuberculosis* via population-based molecular epidemiology. *J. Am. Med. Assoc.* 282, 2321–2327
16 Small, P.M. *et al.* (1993) Molecular strain typing of *Mycobacterium tuberculosis* to confirm cross-contamination in the mycobacteriology laboratory and modification of procedures to minimize occurrence of false-positive cultures. *J. Clin. Microbiol.* 31, 1677–1682
17 Pfyffer, G.E. *et al.* (1998) Transmission of tuberculosis in the metropolitan area of Zurich: a 3 year survey based on DNA fingerprinting. *Eur. Respir. J.* 11, 804–808
18 Chaves, F. *et al.* (1999) Evidence of exogenous reinfection and mixed infection with more than one strain of *Mycobacterium tuberculosis* among Spanish HIV-infected inmates. *AIDS* 13, 615–620
19 Small, P.M. *et al.* (1993) Exogenous reinfection with multidrug-resistant *Mycobacterium tuberculosis* in patients with advanced HIV infection. *New Engl. J. Med.* 328, 1137–1144
20 van Rie, A. *et al.* (1999) Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. *New Engl. J. Med.* 341, 1174–1179
21 Caminero, J.A. *et al.* (2001) Exogenous reinfection with tuberculosis on a European island with a moderate incidence of disease. *Am. J. Respir. Crit. Care. Med.* 163, 717–720
22 Barnes, P.F. *et al.* (1997) Patterns of tuberculosis transmission in Central Los Angeles. *J. Am. Med. Assoc.* 278, 1159–1163
23 Yaganehdoost, A. *et al.* (1999) Complex transmission dynamics of clonally related virulent *Mycobacterium tuberculosis* among predominantly HIV-positive gay males associated with 'barhopping' supports lcoation-based control strategies. *J. Infect. Dis.* 180, 1245–1251
24 Agerton, T.B. *et al.* (1999) Spread of strain W, a highly drug-resistant strain of *Mycobacterium tuberculosis*, across the United States. *Clin. Infect. Dis.* 29, 85–92
25 van Rie, A. *et al.* (1999) Transmission of a multidrug-resistant *Mycobacterium tuberculosis* strain resembling 'strain W' among noninstitutionalized, human immunodeficiency virus-seronegative patients. *J. Infect. Dis.* 180, 1608–1615
26 Frieden, T.R. *et al.* (1996) A multi-institutional outbreak of highly drug-resistant tuberculosis: epidemiology and clinical outcomes. *J. Am. Med. Assoc.* 276, 1229–1235
27 Bifani, P.J. *et al.* (1996) Origin and interstate spread of a New York City multidrug-resistant *Mycobacterium tuberculosis* clone family. *J. Am. Med. Assoc.* 275, 452–457
28 van Soolingen, D. *et al.* (1995) Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia. *J. Clin. Microbiol.* 33, 3234–3238
29 Valway, S.E. *et al.* (1994) Multidrug-resistant tuberculosis in the New York State prison system, 1990–1991. *J. Infect. Dis.* 170, 151–156
30 Agerton, T. *et al.* (1997) Transmission of a highly drug-resistant strain (strain W1) of *Mycobacterium tuberculosis*. Community outbreak and nosocomial transmission via a contaminated bronchoscope. *J. Am. Med. Assoc.* 278, 1073–1077
31 Ramaswamy, S. and Musser, J.M. (1998) Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tubercle Lung Dis.* 79, 3–29
32 Plikaytis, B.B. *et al.* (1994) Multiplex PCR assay specific for the multidrug-resistant strain W of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* 32, 1542–1546
33 Xu, C. *et al.* (1996) Fluoroquinolone resistance associated with specific gyrase mutations in clinical isolates of multidrug-resistant *Mycobacterium tuberculosis*. *J. Infect. Dis.* 174, 1127–1130
34 Bifani, P. *et al.* (2000) Molecular characterization of *Mycobacterium tuberculosis* H37Rv/Ra variants: distinguishing the mycobacterial laboratory strain. *J. Clin. Microbiol.* 38, 3200–3204
35 Anh, D.D. *et al.* (2000) *Mycobacterium tuberculosis* Beijing genotype emerging in Vietnam. *Emerg. Infect. Dis.* 6, 302–305
36 Palittapongarnpim, P. *et al.* (1997) Restriction fragment length polymorphism study of *Mycobacterium tuberculosis* in Thailand using IS*6110* as probe. *Int. J. Tuberc. Lung Dis.* 1, 370–376
37 Das, S. *et al.* (1993) Application of DNA fingerprinting with IS*986* to sequential mycobacterial isolates obtained from pulmonary tuberculosis patients in Hong Kong before, during and after short-course chemotherapy. *Tubercle Lung Dis.* 74, 47–51
38 Lin, R. *et al.* (1996) Transmission patterns of tuberculosis in Taiwan: analysis by restriction fragment length polymorphism. *Int. J. Infect. Dis.* 1, 18–21
39 van Crevel, R. *et al.* (2001) *Mycobacterium tuberculosis* Beijing genotype strains associated with febrile response to treatment. *Emerg. Infect. Dis.* 7, 880–883
40 Marttila, H.J. *et al.* (1998) A Ser315Thr substitution in *katG* is predominant in genetically heterogeneous multidrug-resistant *Mycobacterium tuberculosis* isolates originating from the St. Petersburg area in Russia. *Antimicrob. Agents Chemother.* 42, 2443–2445
41 Portaels, F. *et al.* (1999) Addressing multidrug-resistant tuberculosis in penitentiary hospitals and in the general population of the former Soviet Union. *Int. J. Tuberc. Lung Dis.* 3, 582–588
42 Kurepina, N. *et al.* (2001) The sequence analysis of the *pncA* gene determining the PZA-resistance in the predominant *M. tuberculosis* strains isolated in the Tomsk penitentiary system, Western Siberia, Russia. *Int. J. Tuberc. Lung Dis.* 5, S41
43 Kimerling, M.E. (2000) The Russian equation: an evolving paradigm in tuberculosis control. *Int. J. Tuberc. Lung Dis.* 4, S160–S167
44 Yang, Z. *et al.* (1998) Diversity of DNA fingerprints of *Mycobacterium tuberculosis* isolates in the United States. *J. Clin. Microbiol.* 36, 1003–1007
45 Beggs, M.L. *et al.* (2000) Mapping of IS*6110* insertion sites in two epidemic strains of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* 38, 2923–2928
46 Bifani, P. *et al.* (2001) Molecular identification of streptomycin monoresistant *Mycobacterium tuberculosis* related to multidrug-resistant W strain. *Emerg. Infect. Dis.* 7, 842–848
47 Dale, J.W. *et al.* (1999) Molecular epidemiology of tuberculosis in Malaysia. *J. Clin. Microbiol.* 37, 1265–1268
48 Prodinger, W.M. *et al.* (2001) *Mycobacterium tuberculosis* isolates of Beijing genotype in Thailand. *Emerg. Infect. Dis.* 7, 483–484
49 Zhang, M. *et al.* (1998) Enhanced capacity of a widespread strain of *Mycobacterium tuberculosis* to grow in human macrophages. *J. Infect. Dis.* 179, 1213–1217
50 Rothschild, B.M. *et al.* (2001) *Mycobacterium tuberculosis* complex DNA from an extinct bison dated 17,000 years before the present. *Clin. Infect. Dis.* 33, 305–311
51 Nerlich, A.G. *et al.* (2000) Ancient Egyptian prosthesis of the big toe. *Lancet* 356, 2176–2179
52 Salo, W.L. *et al.* (1994) Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. *Proc. Natl. Acad. Sci. U. S. A.* 91, 2091–2094
53 Niemann, S. *et al.* (2000) Double infection with a resistant and a multidrug-resistant strain of *Mycobacterium tuberculosis*. *Emerg. Infect. Dis.* 6, 548–551
54 Diaz, R. *et al.* (1998) Molecular epidemiology of tuberculosis in Cuba outside of Havana, July 1994–June 1995: utility of spoligotyping versus IS*6110* restriction fragment length polymorphism. *Int. J. Tuberc. Lung Dis.* 2, 743–750
55 Musser, J.M. *et al.* (2000) Negligible genetic diversity of *Mycobacterium tuberculosis* host immune system protein targets: evidence of limited selective pressure. *Genetics* 155, 7–16
56 Torrea, G. *et al.* (1995) Chromosomal DNA fingerprinting analysis using the insertion sequence IS*6110* and the repetitive element DR as strain-specific markers for epidemiological study of tuberculosis in French Polynesia. *J. Clin. Microbiol.* 33, 1899–1904
57 Haas, D.W. *et al.* (1998) Nosocomial transmission of a drug-sensitive W-variant *Mycobacterium tuberculosis* strain among patients with acquired immunodeficiency syndrome in Tennessee. *Infect. Control Hosp. Epidemiol.* 19, 635–639
58 Nivin, B. *et al.* (1998) Cross-contamination with *Mycobacterium tuberculosis*: an epidemiological and laboratory investigation. *Infect. Control Hosp. Epidemiol.* 19, 500–503
59 Codina, G. *et al.* (1999) Multidrug-resistant tuberculosis caused by 'W'-related strains in three immunocompetent foreign-born patients. *Int. J. Tuberc. Lung Dis.* 3, 82–84
60 Qian, L. *et al.* (1999) Retrospective analysis of the Beijing family of *Mycobacterium tuberculosis* in preserved lung tissues. *J. Clin. Microbiol.* 37, 471–474
61 Sola, C. *et al.* (1999) Tuberculosis in the Caribbean: using spacer oligonucleotide typing to understand strain origin and transmission. *Emerg. Infect. Dis.* 5, 404–414
62 Soini, H. *et al.* (2000) Characterization of *Mycobacterium tuberculosis* isolates from patients in Houston, Texas, by spoligotyping. *J. Clin. Microbiol.* 38, 669–676
63 Laserson, K.F. *et al.* (2000) Clinical and programmatic mismanagement rather than community outbreak as the cause of chronic, drug-resistant tuberculosis in Buenaventura, Colombia, 1998. *Int. J. Tuberc. Lung Dis.* 4, 673–683
64 Doroudchi, M. *et al.* (2000) IS*6110*-RFLP and spoligotyping of *Mycobacterium tuberculosis* isolates in Iran. *Scand. J. Infect. Dis.* 32, 663–668
65 Ravins, M. *et al.* (2001) Molecular epidemiology of *Mycobacterium tuberculosis* infection in Israel. *J. Clin. Microbiol.* 39, 1175–1177
66 Escalante, P. *et al.* (1998) Genotypic characterization of drug-resistant *Mycobacterium tuberculosis* isolates from Peru. *Tubercle Lung Dis.* 79, 111–118