

Characterization of the Genetic Diversity of Extensively-Drug Resistant Mycobacterium tuberculosis Clinical Isolates from Pulmonary Tuberculosis Patients in Peru

Omar Cáceres, Nalin Rastogi, Carlos Bartra, David Couvin, Marco Galarza, Luis Asencios, Alberto Mendoza-Ticona

▶ **To cite this version:**

Omar Cáceres, Nalin Rastogi, Carlos Bartra, David Couvin, Marco Galarza, et al.. Characterization of the Genetic Diversity of Extensively-Drug Resistant Mycobacterium tuberculosis Clinical Isolates from Pulmonary Tuberculosis Patients in Peru. PLoS ONE, 2014, 9 (12), pp.e112789. 10.1371/journal.pone.0112789 . pasteur-01153013

HAL Id: pasteur-01153013
https://riip.hal.science/pasteur-01153013

Submitted on 19 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PLOS ONE

RESEARCH ARTICLE

Characterization of the Genetic Diversity of Extensively-Drug Resistant *Mycobacterium tuberculosis* Clinical Isolates from Pulmonary Tuberculosis Patients in Peru

Omar Cáceres[1]*, Nalin Rastogi[2], Carlos Bartra[3], David Couvin[2], Marco Galarza[1], Luis Asencios[3], Alberto Mendoza-Ticona[4]

1. Biotechnology and Molecular Biology Laboratory, Instituto Nacional de Salud, Lima, Peru, **2.** WHO Supranational TB Reference Laboratory, TB and Mycobacteria Unit, Institut Pasteur de la Guadeloupe, Guadeloupe, France, **3.** Mycobacterias National Reference Laboratory, Instituto Nacional de Salud, Lima, Peru, **4.** Ministerio de Salud (MINSA), Lima, Peru

*caceresreyomar@gmail.com

CrossMark
click for updates

OPEN ACCESS

Citation: Cáceres O, Rastogi N, Bartra C, Couvin D, Galarza M, et al. (2014) Characterization of the Genetic Diversity of Extensively-Drug Resistant *Mycobacterium tuberculosis* Clinical Isolates from Pulmonary Tuberculosis Patients in Peru. PLoS ONE 9(12): e112789. doi:10.1371/journal.pone.0112789

Editor: Riccardo Manganelli, University of Padova, Medical School, Italy

Received: April 28, 2014

Accepted: October 20, 2014

Published: December 9, 2014

Copyright: © 2014 Cβceres et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: AM and OC were supported by International Association of National Public Health Institutes (IANPHI) Research Seed Grant (www.ianphi.org) and Instituto Nacional de Salud, Peru (www.ins.gob.pe). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Background: Peru holds the fourth highest burden of tuberculosis in the Americas. Despite an apparently well-functioning DOTS control program, the prevalence of multidrug resistant tuberculosis (MDR-TB) continues to increase. To worsen this situation, cases of extensively drug resistance tuberculosis (XDR-TB) have been detected. Little information exists about the genetic diversity of drug-susceptible vs. MDR-TB and XDR-TB.

Methods: Cryopreserved samples of XDR strains from 2007 to 2009 (second semester), were identified and collected. Starting from 227 frozen samples, a total of 142 XDR-TB strains of *Mycobacterium tuberculosis* complex (MTBC; 1 isolate per patient) were retained for this study. Each strain DNA was analyzed by spoligotyping and the 15-loci Mycobacterial Interspersed Repetitive Unit (MIRU-15).

Results: Among the 142 isolates analyzed, only 2 samples (1.41%) could not be matched to any lineage. The most prevalent sublineage was Haarlem (43.66%), followed by T (27.46%), LAM (16.2%), Beijing (9.15%), and X clade (1.41%). Spoligotype analysis identified clustering for 128/142 (90.1%) isolates vs. 49/142 (34.5%) with MIRUs. Of the samples, 90.85% belonged to retreated patients. The drug resistant profile demonstrated that 62.67% showed resistance to injectable drugs capreomycin (CAP) and kanamycin (KAN) vs. 15.5% to CAP alone and 21.8% to KAN alone. The SIT219/T1 and SIT50/H3 were the most prevalent

patterns in our study. The spoligoforest analysis showed that SIT53/T1 was at the origin of many of the T lineage strains as well as a big proportion of Haarlem lineage strains (SIT50/H3, followed by SIT47/H1, SIT49/H3, and SIT2375/H1), as opposed to the SIT1/Beijing strains that did not appear to evolve into minor Beijing sublineages among the XDR-TB strains.

Conclusion: In contrast with other Latin-American countries where LAM sublineage is the most predominant, we found the Haarlem to be the most common followed by T sublineage among the XDR-TB strains.

Introduction

With almost 9 million new cases in 2011 and 1.4 million deaths, tuberculosis (TB) caused by *Mycobacterium tuberculosis* ranks as the second leading cause of death from an infectious disease in the world [1]. The emergence of multidrug-resistant (MDR) strains showing combined resistance to two major first-line drugs isoniazid (INH) and rifampicin (RIF) and the increased HIV/TB coinfection not only contribute to the spread and re-emergence of this disease, but also constitute a threat of developing added resistance to second-line drugs.

From 2007 to 2010 the proportion of new TB cases reported as MDR-TB ranged from 0%–28.9% while the proportion of previously treated MDR-TB ranged from 0% to 65.1% [2]. MDR-TB complicates management of patients due to increased pressure on public health systems and cost of the treatment. It further aggravates the emergence of extensively drug-resistant TB (XDR-TB), defined as MDR-TB plus resistance to a fluoroquinolone and at least one of the three second-line injectable drugs (Amikacin, Kanamycin or Capreomycin). The fact that XDR-TB requires longer, more expensive and more toxic treatment regimens, that at the same time are less likely to cure the disease [2], further worsens the situation. Hence, tackling XDR-TB represents a formidable challenge to public health programs, particularly in low-resource settings.

Following Haiti, Bolivia and Guyana, Peru holds the fourth highest burden of tuberculosis in the Americas. In 2012, 29,760 cases were reported with an incidence of 95 cases per 100,000 inhabitants [1]. MDR-TB in Peru is increasing; in 2012, 1225 new cases of pulmonary MDR-TB were reported [1]. Peru has 41.3% of all MDR-TB cases in the region of the Americas.

The first XDR-TB cases were detected by Instituto Nacional de Salud (INS) in 2007 [3]. Since then, the number of new XDR-TB cases has been increasing, e.g., 50 new cases were detected in 2010 vs. 92 in 2013. From the total cases of TB, the highest prevalence of MDR-TB and XDR-TB cases occurred in Lima (the capital of Peru) with 80% and 92% respectively [4].

To better comprehend the molecular epidemiology of MTBC, techniques based on two-step typing strategies combining rapid and high resolution PCR-based methods such as spoligotyping [5] and MIRU-VNTRs [6, 7] have been

successfully used. Among the latter, 15-loci MIRU-VNTRs were shown to possess enough discriminatory power for epidemiological studies permitting assignment of clusters with epidemiological data efficiently [7,8]. We therefore decided to characterize the genetic diversity of the XDR-TB strains isolated from pulmonary TB patients in Peru, and to analyze their population structure using spoligotyping and 15-loci MIRU-VNTRs in conjunction with available demographic, clinical and epidemiological data.

Materials and Methods

Bacteria and strain information

The Mycobacteria Laboratory at the INS is the National Reference Laboratory for the diagnosis and detection of MTBC drug resistance in Peru. Under its routine activity, INS isolated, identified, and cryo-preserved all MDR and XDR strains in a MTBC strain bank, which were duly confirmed for their drug-resistance using drug-susceptibility testing (DST) for first and second line antituberculosis drugs by the agar proportion method [9]. Starting from 227 frozen samples obtained between 2007 to the second semester of 2009, a total of 142 XDR-TB strains of *Mycobacterium tuberculosis* complex (MTBC; 1 isolate per patient) were successfully subcultured, and retained for this retrospective genotyping study. All cryo-preserved strains were thawed and reactivated in 2 ml of 7H9 liquid culture media for 10 days. After the strains were confirmed to be MTBC by rapid chromatographic immunoassay (BD MGI TBc Identification Test), 1.5 ml of culture was centrifuged at 10000 rpm for 5 minutes. The pellet was resuspended in 1 ml of TE buffer, 500 uL of the resuspension was heated at 100℃ for 30 minutes, and the DNA was extracted using the CTAB-NaCl method [10]. The remaining 0.5 ml was subcultured in Löwenstein-Jensen (L-J) medium to be cryopreserved thereafter.

For strain information, an Excel database was generated with demographic data (age, sex, geographic area of isolation, year of isolation) and drug resistant profiles (RIF: rifampicin, INH: isoniazid, EMB: ethambutol, PZ: pyrazinamide, SM: streptomycin, CFX: ciprofloxacin, KAN: kanamycin, CAP: capreomycin, ETH: ethionamide, PAS: p-amino salicylic acid, CS: cycloserine). All of this data was taken from the lab registry for cryo-preserved strains. Each sample evaluated was a single strain from a unique patient. Although serial strains are available in the strain bank, we were careful to only examine one strain per patient for this analysis. The strains covered all Peruvian departments where XDR strains were primarily isolated.

Genotyping methods

Spoligotyping was carried out using a commercial kit (Isogen Bioscience, BV Maarsen, The Netherlands) according to the protocol previously described by Kamerbeek et al. [5]. Briefly, the DR region of the TB genome was amplified using

primers DRa and DRb, and the amplified biotinylated products were hybridized to a set of 43 oligonucleotides covalently bound to a membrane. The hybridized PCR products were then incubated with a streptavidin-peroxidase conjugate and the membrane then exposed to chemiluminescence reaction (Amersham ECL Direct nucleic acid labeling and detection system, GE Healthcare Limited, UK). The membrane was exposed on a gel documentation system (Chemidoc XRS, Biorad, USA). DNA extracts of *M. tuberculosis* H37Rv and *M. bovis* BCG were used as controls. Spoligotypes in binary format were analyzed and compared with the SITVIT2 proprietary database of the Pasteur Institute of Guadeloupe, which is an updated ''in-house'' version of the SITVITWEB database [11]. A cluster was defined as two or more strains sharing identical spoligotyping patterns, and assigned a Spoligotype International Type (SIT) number in the database.

15-loci MIRU typing was performed as described elsewhere [8]. Briefly, each locus was amplified individually using 2 μL of mycobacterial DNA (20 ng) in 23 μL of a reaction mixture containing 0.4 μM of loci-respective primers and PCR Master Mix (Invitrogen, California, USA) according to the manufacturer's instructions. The PCR conditions for each set of primers were carried out as described [8] with a minor modification; we used Betaine 1M instead of DMSO. PCR products were subjected to electrophoresis in a 2% weight/volume agarose gel (Invitrogen Life Technologies, SP, Brazil). 100-bp DNA Ladders (Fermentas, Vilnius, Lithuania) were used as molecular markers. The gels were stained with ethidium bromide and visualized under ultraviolet light, then photodocumented with Chemidoc XRS System. PCR fragment size was determined by Quantity One software (BioRad, CA, USA) with the molecular markers as reference and the MIRU allele scoring was determined according to Supply et al. [6]. For data entry in the SITVIT2 database, the results from each of the 15 loci were combined to create a 15-digit allelic profile in the following order: MIRU-4, MIRU-10, MIRU-16, MIRU-26, MIRU-31, MIRU-40, ETR-A, ETR-C, QUB-11b, QUB-26, QUB-4156, Mtub04, Mtub21, Mtub30, and Mtub39. A cluster was defined as two or more strains sharing identical 15-loci MIRU patterns, and assigned a MIRU International Type (15-MIT) number in the database.

Evolutionary relationship analysis

A minimum spanning tree (MST) illustrating evolutionary relationships between spoligotypes and MIRU patterns was drawn using BioNumerics version 6.6 (Applied Maths NV, Sint-Maartens-Latem, Belgium). MST connects each spoligotype based on the degree of changes required to go from one allele to another. The MST structure is represented by branches (continuous vs. dashed and dotted lines) and circles representing each individual pattern. The length of the branches represents the distance between patterns while the complexity of the lines (continuous, gray dashed and gray dotted) denotes the number of allele/spacer changes between two patterns: solid lines, 1 or 2 or 3 changes (thicker ones indicate a single change, while the thinner ones indicate 2 or 3 changes); gray

dashed lines represent 4 changes; and gray dotted lines represent 5 or more changes. The size of the circle is proportional to the total number of isolates.

We drew MSTs to visualize the relationships between spoligotype patterns and the sites of isolation of the strains, the treatment history of patients, and resistance to injectable drugs, namely CAP and KAN. Peruvian XDR strains belonging to Beijing lineage were further compared by drawing a MST with Beijing isolates from other countries in the SITVIT2 database (n=863); for this analysis we retained countries with a significant number of Beijing strains with 15-loci MIRU typing data (35 or more) being available, i.e., Japan (n=603), China (n=209) and France (n=35).

In addition, the comparative diversity of these strains was also evaluated by WebLogo graphical representation (available at: http://weblogo.berkeley.edu) which was previously used to represent spoligotyping motifs based on the presence or absence of specific spacer sequences [12]. This application was initially designed to generate a graphical representation of amino acid or nucleic acid sequence logo analysis [13, 14]. We adapted this application to create sequence codes for 15-loci MIRUs as follows; WebLogo Label/number of copies for a loci: A/1, B/2, C/3, D/4, E/5, F/6, G/7, H/8, I/9, J/10, K/11, L/12, M/13, N/14, O/15, P/16, Q/0, U/Unknown.

Relationships among spoligotypes were estimated using the spoligoforest program in the SpolTools webpage (http://www.emi.unsw.edu.au/spoltools/; [15, 16]) for all SITs observed. The method makes use of a model that considers mutations in spoligotypes as irreversible deletions of spacers, and assigns probabilities to the lengths of these deletions. The size of each node is an increasing function of the number of isolates (i.e., the cluster size); edges between nodes reflect evolutionary relationships between spoligotypes with arrowheads pointing to descendants. The spoligoforest tree was colored using the GraphViz software (http://www.graphviz.org).

Genetic diversity analysis

The Hunter–Gaston discriminatory index (HGDI) [17] was used to estimate the discriminatory power of genotyping methods. Cluster analyses of Spoligotyping and MIRU profiles were also recorded as character data and analyzed using MIRU-VNTRplus program [18]. Dendrograms were generated by using the Jaccard's distance option and the unweighted pair group method of averages (UPGMA) clustering method.

Statistical analysis

Statistical analysis was performed with Epi Info software 3.51 (Centers for Disease Control and Prevention, Atlanta, GA, USA), by using $\chi 2$ test or Fisher exact test for the comparison of proportions. Median age and interquartile (IQ) ranges were calculated using MegaStat software (http://highered.mheducation.com/sites/0077425995/information_center_view0/index.html). A p value$<$0.04 was

considered significant, and a p value between 0.04 and 0.06 was considered ''marginally significant''.

Ethical consideration

The study was approved at the Institutional Review Board and Ethical Committee at Instituto Nacional de Salud in Peru. All information related to the patients was completely anonymized prior to analysis.

Results

Characteristics of the population studied

Starting from 227 frozen samples in the INS strain-bank, a total of 142 XDR-TB strains of *Mycobacterium tuberculosis* complex (MTBC; 1 isolate per patient) representing 62.5% of the sample were retained for this study; 85 strains were not included in the study for a diversity of reasons (strains not reactivated, duplicated strains from the same patient and/or epidemiological data missing). All subsequent analyses are based on the 142 MTBC strains genotyped under this study. The age of the patients ranged between 15 and 72 years, with an average of 34 years and a median of 31 years (S1 Table). The majority of the subjects (66.2%) were aged between 25 and 54 years. From this sample, 90 (63.4%) were male and 52 (36.6%) were female. In regards to age groups the difference between male and female was marginally significant ($p=0.0505$) (Table 1).

All samples were pulmonary XDR-TB and HIV negative, 13 samples (9.15%) were new XDR-TB cases and 129 (90.85%) were relapsed. DST confirmed that all samples were XDR-TB, but showed a variable resistance to injectable drugs: 89 samples (62.67%) showed resistance to CAP and KAN, 22 (15.5%) were resistant to CAP alone and 31 samples (21.8%) were resistant to KAN alone. (S2 Table)

In regards to the origin of the strains, 119 (83.8%) were isolated in Lima and 23 samples (16.2%) in one of the following departments: Ica, Tacna, Arequipa Madre de Dios, Junin and Callao (constitutional province). The cases detected in Callao represented 60.8% of all cases detected outside of Lima.

Analysis by spoligotyping

We performed spoligotyping to determine the population structure of the 142 XDR strains (S2 Table). The most dominant spoligotype family in the XDR cases was the Haarlem (H) sublineage (43.6%, n=62), followed by the T (28.2%, n=40), Latin-American and Mediterranean (LAM, 16.2%, n=23), Beijing (9.2%, n=13) and X3 (1.4%, n=2) sublineages. Two isolates (1.4%) displayed unknown patterns with no matches to any of the major clades present in the database. Moreover, we found nine (6.3%) new SITs (Table 2) and 5 orphan strains (3.5%) that were not present in the SITVIT2 database (see S2 Table for detailed genotyping and drug-resistance data and demographic information).

Table 1. Age distribution between males and females with XDR-TB.

Age Groups	Frequency in Females	Percent	Frequency in Males	Percent
1<21	4	7.7	9	10.0
21<41	40	76.9	49	54.4
41<61	7	13.5	27	30.0
61<81	1	1.9	5	5.6
Total	52	100.0	90	100.0

p-value=0.050579770369946.
The difference between Females and Males is marginally significant (regarding the Age groups).

doi:10.1371/journal.pone.0112789.t001

One hundred and twenty eight isolates, collected in Lima and different departments, were categorized into 26 shared-types (Table 2). Nine strains exhibited unique SIT patterns. The remaining isolates formed 17 different clusters. SIT50 (H3 clade) and SIT219 (T1 clade) were the predominant patterns - each pattern was presented in 21 different isolates, each one accounting for 14.79% of all isolates in the study (Table 2). Thirteen predominant SITs representing 120 strains were identified and their worldwide distribution was determined (Table 3). As mentioned previously, SIT50-H3 and SIT219-T1 were the predominant type (each with 14.79% of isolates in our study), followed by SIT47-H1 (11.27%), SIT1-Beijing (9.15%), SIT3001-H3 (7.75%), SIT53-T1 and SIT1355-LAM represented 4.93% each. Finally, SIT469-LAM1 and SIT3778-H3; SIT42-LAM9 and SIT49-H3 and SIT52-T2 and SIT93-LAM5 accounted for 3.52%, 2.82% and 2.11% respectively (Table 2). The discriminative power of the spoligotyping method, measured by the Hunter-Gaston index, was 0.924

Analysis by MIRU-VNTRs

The results of the MIRU analysis (15-MIT) showed that the 142 isolates were classified into clustered (n=49 or 34.5% grouped in 11 clusters) and unclustered (n=93 or 65.5%) patterns. The clustered strains corresponded to following lineages by spoligotyping: Beijing (n=4, 8.2%); Haarlem (n=22, 44.9%); T (n=14, 28.6%); LAM (n=8, 16.3%) and an unknown lineage (n=1, 2%) (S2 Table). The 93 unclustered isolates showed unique MIRU patterns, all except one (which corresponded to SIT49/MIT369 in the SITVIT2 database), the remaining patterns were not yet reported and corresponded to orphan patterns. A dendrogram was constructed based on both spoligotyping and MIRU results (S1 Figure), and showed that the isolates could be divided into three groups based on their phylogenetic clustering and genotypic characteristics. Groups I, II, and III contained 13, 47 and 82 isolates respectively. Group I presented one cluster (3 isolates), group II presented 4 clusters (12 isolates) and group III presented 6 clusters (19 isolates). We observed that fourteen isolates were not grouped in the dendrogram because all of them had a different spoligotype (S1 Figure) in regards to the cluster generated by MIRU-15 and the spoligotyping profile. The HGDI for MIRU-15 was 0.993

Table 2. Description of 26 shared-types (SITs; n = 137 isolates) and corresponding spoligotyping defined lineages/sublineages starting from a total of 227 cryopreserved *M. tuberculosis* strains isolated from adults with pulmonary tuberculosis in Lima, Peru.

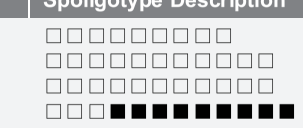
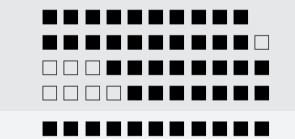
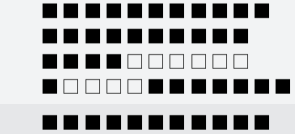
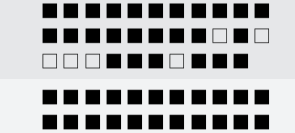
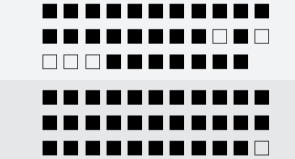
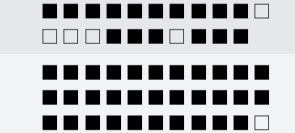
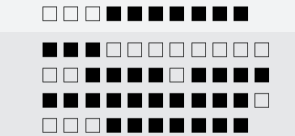
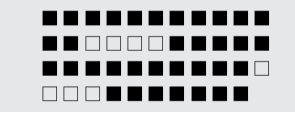
SIT	Spoligotype Description	Octal Number	Nb in this study	% in study	% in this study vs. data-base	Line-age**	Match with a strain from (for newly created SITs)***	Unique vs. Clus-tered¥
1		000000000 003771	13	9.15	0.13	Beijing		Clustere-d
42		7777776077 60771	4	2.82	0.12	LAM9		Clustere-d
47		777777774 020771	16	11.27	1.07	H1		Clustere-d
49		7777777777 20731	4	2.82	2.26	H3		Clustere-d
50		777777777 720771	21	14.79	0.63	H3		Clustere-d
52		777777777 760731	3	2.11	0.33	T2		Clustere-d
53		777777777 760771	7	4.93	0.12	T1		Clustere-d
91		700036777 760771	1	0.7	0.4	X3		Unique
93		777737607 760771	3	2.11	0.84	LAM5		Clustere-d
189		777741777 760771	1	0.7	9.09	T1		Unique

Table 2. Cont.

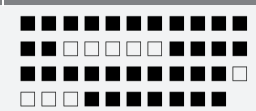
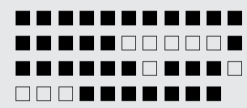
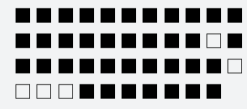
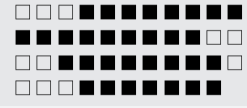
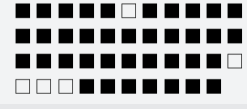
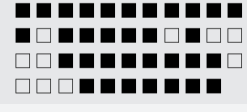
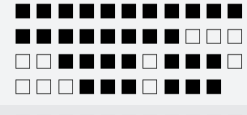
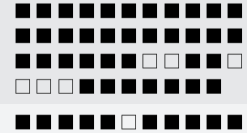
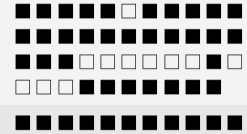
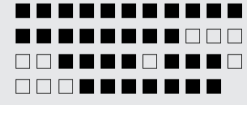
SIT	Spoligotype Description	Octal Number	Nb in this study	% in study	% in this study vs. data-base	Line-age**	Match with a strain from (for newly created SITs)***	Unique vs. Clus-tered¥
219		777740777 760771	21	14.79	35.59	T1		Clustered
222		777774077 560771	1	0.7	1.89	Un-known		Unique
291		777777677 760771	1	0.7	1.15	T1		Unique
469		07777760 7760771	5	3.52	14.71	LAM1		Clustered
1122		76777777 7760771	1	0.7	1.69	T1		Unique
1160		77773720 7760771	1	0.7	25	LAM5		Unique
1355		7777774 07560731	7	4.93	12.73	LAM		Clustered
1905		7777777 77460771	2	1.41	9.52	T1		Clustered
2375		767777 774020771	2	1.41	15.38	H1		Clustered
2502		777777 407560771	2	1.41	40	LAM6		Clustered

Table 2. Cont.

SIT	Spoligotype Description	Octal Number	Nb in this study	% in study	% in this study vs. database	Lineage**	Match with a strain from (for newly created SITs)***	Unique vs. Clustered¥
2940	■■■■■■■■■■ ■■■■■■■■■□□ □□■■■■■■■■□ □□□■□□■■■■	777777 607760471	1	0.7	14.29	LAM9		Unique
3001	■■■■■■■■■■ ■■■■■□□□□□■ ■■■■■□□□□■□ □□□■■■■■■	7777740 77020771	11	7.75	55	H3		Clustered
3777	■■■■■■■■■■ ■■■■■□■□□□■ ■■■■■■□■■■□ □□□■■■■■■	77777507 7560771	1	0.7	50	T1	This study (n=1), ESP (n=1)	Unique
3778	■■■■■■■■■■ ■■■■■□■□□□■ ■■■■■□□□□■□ □□□■■■■■■	77777507 7020771	5	3.52	100	H3	This study (n=5)	Clustered
3779	■■■■■□■■■■ ■■■■■□□□□□■ ■■■■■□□□□■□ □□□■■■■■■	76777407 7020771	2	1.41	100	H3	This study (n=2)	Clustered
3780	■■■□□□□□□□ □□■■■□■■■■ ■■■■■■■■□□ □□□■■■■■■	7000367 77740771	1	0.7	50	X3	This study (n=1), PER (n=1)	Unique

*A total of 22/26 SITs (n=128) matched a preexisting shared-type in the database, whereas 4/26 SITs (n=9 isolates) were newly created either within the present study or after a match with an orphan in the database.

A total of 17 SITs containing 128 isolates were clustered within this study (2 to 21 isolates per cluster), while 9 SITs contained a unique strain within this study. Note that SITs in bold indicates "newly created shared-type" (n=4 containing 9 isolates) due to 2 or more strains belonging to an identical new pattern within this study or after a match with an orphan in the database.

**Lineage designations according to SITVIT2; "Unknown" designates patterns with signatures that do not belong to any of the major clades described in the database.

*** SIT designations are indicated in the table.

ESP=Spain, PER=Peru.

¥Clustered strains correspond to a similar spoligotype pattern shared by 2 or more strains "within this study"; as opposed to unique strains harboring a spoligotype pattern that does not match with another strain from this study.

doi:10.1371/journal.pone.0112789.t002

Table 3. Description of clusters composed of predominant shared types (defined as SITs representing>2% strains, n=13) in our study and their worldwide distribution in the SITVIT2 database.

SIT	Spoligotype Description	Octal Number	Number (%) in study	% in study vs. SITVIT2	Lineage	Distribution in Regions with>=3% of a given SIT*	Distribution in countries with>=3% of a given SIT**
1	□□□□

Table 3. Cont.

SIT	Spoligotype Description	Octal Number	Number (%) in study	% in study vs. SITVIT2	Lineage	Distribution in Regions with>=3% of a given SIT*	Distribution in countries with>=3% of a given SIT**
53	■■■■■■■■■■■■■■■■■■■■■■■■■ ■■■■■■■□□□□■■■■■■■	77777777776077-1	7 (4.93)	0.12	T1	AMER-N 16.92, EURO-W 16.21, AMER-S 12.47, EURO-S 9.75, ASIA-W 7.13, EURO-N 5.53, AFRI-S 5.15, AFRI-E 4.67, ASIA-E 4.39, AFRI-N 3.65, EURO-E 3.38	USA 13.67, FXX 8.16, ITA 5.53, BRA 5.31, ZAF 5.03, TUR 3.6, AUT 3.55, CHN 3.2
93	■■■■■■■■■■■■□■■■■■■■□□□□■ ■■■■■■■□□□□■■■■■■■	77773760776077-1	3 (2.11)	0.84	LAM5	AMER-S 48.6, AMER-N 21.35, CARI 11.8, EURO-S 8.43, EURO-W 5.06	VEN 26.4, USA 21.07, BRA 12.08, ITA 5.34, GLP 5.34, PER 4.78, HTI 4.21, FXX 3.09, ESP 3.09
219	■■■■■■■■■■■■□□□□□■■■■■■ ■■■■■■■□□□□■■■■■■■	77774077776077-1	21 (14.79)	35.59	T1	AMER-N 52.54, AMER-S 47.46	USA 52.54, PER 47.46
469	□□□■■■■■■■■■■■■■■■□□□ □■■■■■■■■□□□□■■■■■■	07777760776077-1	5 (3.52)	14.71	LAM1	AMER-N 38.24, AMER-S 32.35, EURO-S 5.88, ASIA-SE 5.88, AFRI-N 5.88	USA 38.24, PER 26.47, ITA 5.88, EGY 5.88
1355	■■■■■■■■■■■■■■■■■□□□□□■ ■■■□■■■□□□□■■■□■■■	77777740756073-1	7 (4.93)	12.73	LAM	AMER-S 61.82, EURO-S 23.64, AMER-N 12.73	PER 60.0, ITA 18.18, USA 12.73, ESP 5.46
3001	■■■■■■■■■■■■■■■□□□□□■■■■ ■■□□□□■□□□□■■■■■■■	77777407702077-1	11 (7.75)	55	H3	AMER-S 95.0, AMER-N 5.0	PER 95.0, USA 5.0
3778	■■■■■■■■■■■■■■■□■□□□■■■■ ■■□□□□■□□□□■■■■■■■	77777507702077-1	5 (3.52)	100	H3	AMER-S 100.0	PER 100.0

* Worldwide distribution is reported for regions with more than 3% of a given SITs as compared to their total number in the SITVIT2 database.
The definition of macro-geographical regions and sub-regions (http://unstats.un.org/unsd/methods/m49/m49regin.htm) is according to the United Nations; Regions: AFRI (Africa), AMER (Americas), ASIA (Asia), EURO (Europe), and OCE (Oceania), subdivided in: E (Eastern), M (Middle), C (Central), N (Northern), S (Southern), SE (South-Eastern), and W (Western). Furthermore, CARIB (Caribbean) belongs to Americas, while Oceania is subdivided in 4 sub-regions, AUST (Australasia), MEL (Melanesia), MIC (Micronesia), and POLY (Polynesia).
Note that in our classification scheme, Russia has been attributed a new sub-region by itself (Northern Asia) instead of including it among rest of the Eastern Europe. It reflects its geographical localization as well as due to the similarity of specific TB genotypes circulating in Russia (a majority of Beijing genotypes) with those prevalent in Central, Eastern and South-Eastern Asia.
** The 3 letter country codes are according to http://en.wikipedia.org/wiki/ISO_3166-1_alpha-3; countrywide distribution is only shown for SITs with ≥3% of a given SITs as compared to their total number in the SITVIT2 database.

doi:10.1371/journal.pone.0112789.t003

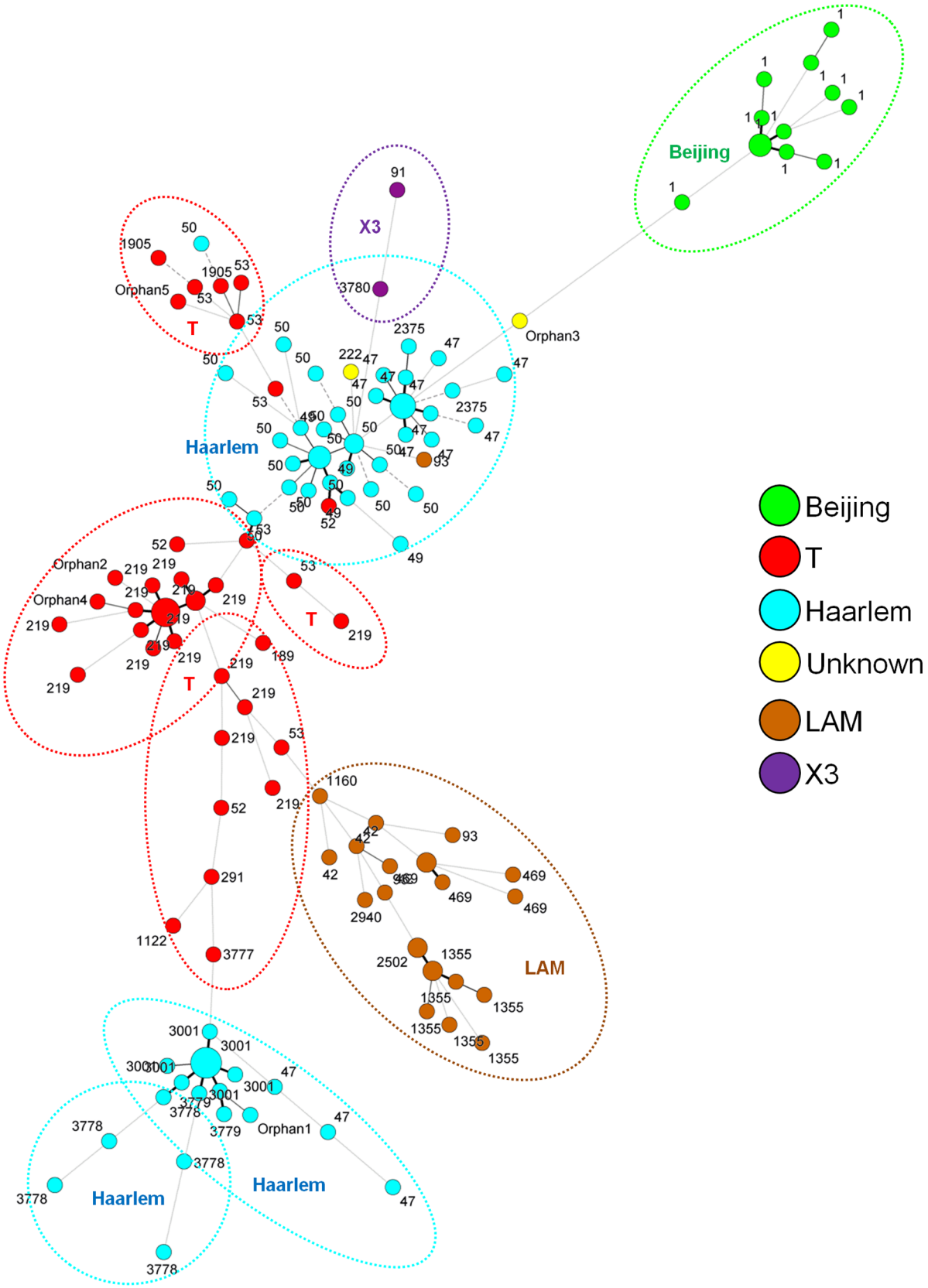

Beijing
X3
T
Haarlem
LAM
Unknown
Orphan1
Orphan2
Orphan3
Orphan4
Orphan5

Fig. 1. MST of spoligotyping in conjunction with MIRU-15 typing. Distinction of the genotypic lineages is shown by circles of different colors. Patterns colored in yellow indicate a strain with an unknown signature (unclassified in the SITVIT2 database). The MST allows a finer discrimination of the main spoligotype central nodes.

doi:10.1371/journal.pone.0112789.g001

Relationship between lineage of MTB and Peruvian XDR strains

A minimum spanning tree based spoligotyping data was constructed to visualize the patterns connected with cities of isolation of XDR strains in Peru (S2 Figure). All the strains isolated in the departments of Peru were present in Lima except the SIT1122, orphan 2 and orphan 5 that were present exclusively in Callao, SIT1160 in Ica and SIT3779 in Callao and Junin. A composite MST based on both spoligotyping and MIRU results was drawn for a better discrimination of

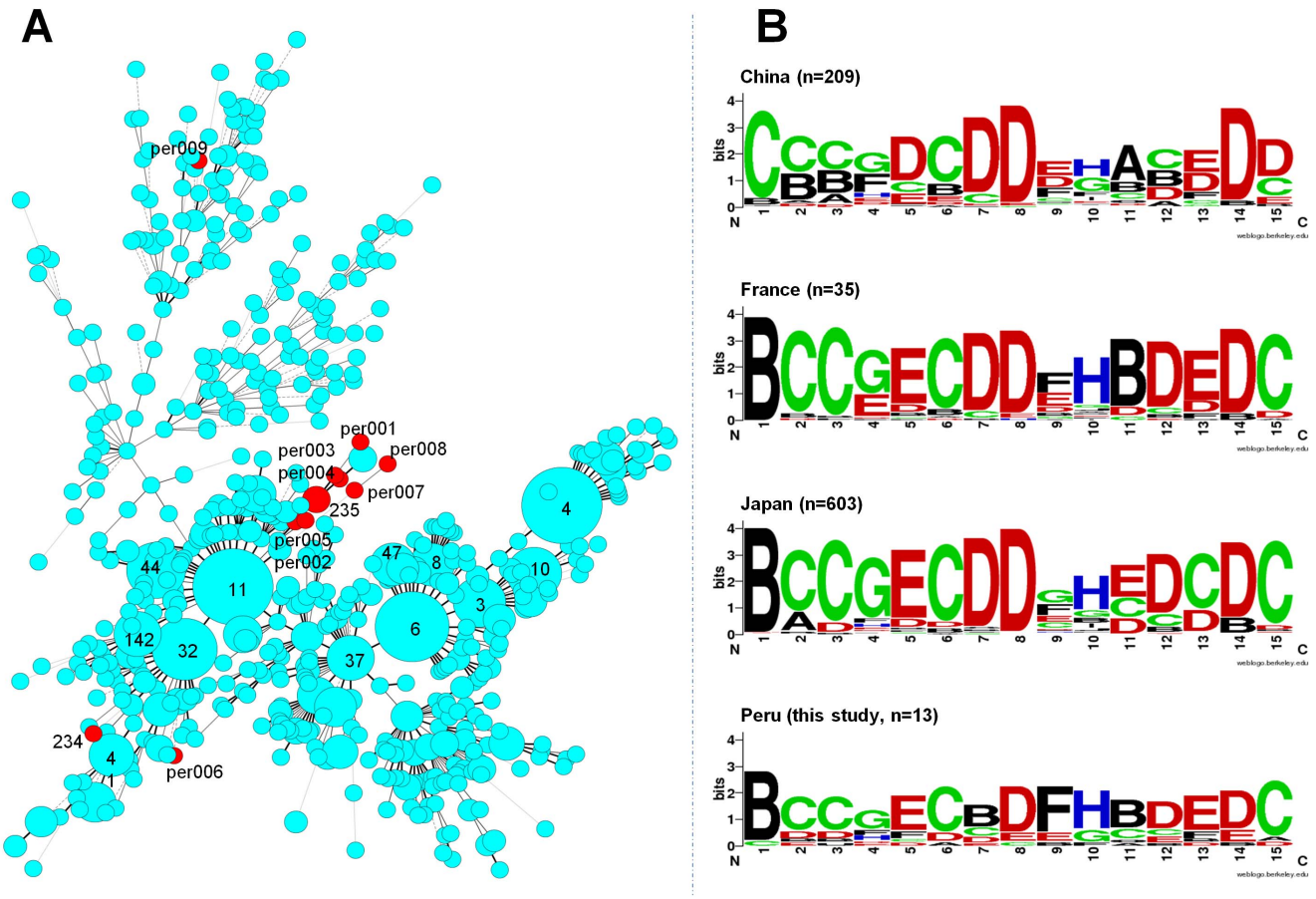


Fig. 2. Evolutionary relationships between Beijing lineages isolated from Peruvian XDR-TB strains (n=13) vs. other countries in the SITVIT2 database (n=863). (A). A minimum spanning tree based on 15-loci MIRU-VNTR of the Peruvian Beijing isolates (highlighted in red) versus isolates from other countries (shown in cyan-blue color). (B). Comparative diversity of Peruvian Beijing vs. other countries using the WebLogo application. Each logo consists of stacks of symbols, one stack for each MIRU loci. The overall height of the stack indicates the conservation of a given MIRU loci with a fixed number of copies at that position (i.e., if 100% of the strains conserve the same number of copies for a given MIRU loci, it corresponds to 4 bits), while the height of individual symbols within the stack indicates the relative frequency of number of copies of a given MIRU loci at that position. WebLogo stack label/ number of copies for MIRU loci: A/1, B/2, C/3, D/4, E/5, F/6, G/7, H/8, I/9, J/10, K/11, L/12, M/13, N/14, O/15, P/16, Q/0, U/Unknown.

doi:10.1371/journal.pone.0112789.g002

circulating lineages among XDR strains (Fig. 1), and allowed separation of strains in six well-defined groups around main spoligotype central nodes, comprised of Beijing, T, Haarlem, LAM, X3 and strains with unknown signatures. Among these identified lineages, all were well distributed; nonetheless one may notice the predominance of Haarlem group followed by T sublineages (Fig. 1).

We also drew a spoligoforest tree as a hierarchical layout, where the continuity of the lines indicates the weight of the hypothetical evolutionary relationship between spoligotypes (S3 Figure). In this illustration, each pattern from the study is represented by a node with area size being proportional to the total number of isolates. Changes (loss of spacers) are represented by directed edges between nodes, with the arrowheads pointing to descendant spoligotypes. Using this model, solid black lines link patterns that are very similar, i.e., loss of one spacer only as opposed to dashed and dotted lines that represent respectively 2 or more spacer changes. The spoligoforest obtained showed four subtrees with connected components and two unconnected nodes. One may notice that SIT50/H3 and SIT219/T1 are the biggest nodes (n=21 each), followed by SIT47/H1 (n=16), SIT1/Beijing (n=13), SIT3001/H3 (n=11), SIT53/T1 and SIT1355/LAM (n=7 each); followed by smaller nodes of 5 strains and less. Eighteen spoligotypes descended from SIT53, six of which are in small clusters (range 1–4), and two in larger clusters (21 isolates) - one from a lineage distinct to SIT53. The hypothetical evolutionary relationship between spoligotypes SIT53 and SIT50 and SIT53 and SIT219, the spoligotypes with the largest clusters in the data, were strong and weak respectively. Five other spoligotypes (SIT52, SIT189, SIT291, SIT1905 and SIT1122) also showed a strong relationship with SIT53, while two other subtrees (rooted by spoligotype SIT3777 and SIT2502) lead to Haarlem and LAM lineage strains. Globally, this analysis suggested that SIT53/T1 was at the origin of many of the T lineage strains as well as a big proportion of Haarlem lineage strains (SIT50/H3, followed by SIT47/H1, SIT49/H3, and SIT2375/H1), as opposed to the SIT1/Beijing strains that did not appear to evolve into minor Beijing sublineages among the XDR-TB strains.

Association between of Peruvian XDR strains and resistance to injectable drugs

We drew a spoligotyping based MST to visualize a possible link between lineages and treatment history of the patients (S4 Figure). It shows that a majority of XDR-TB cases concerned relapsed patients (n=129, 90.85%) with only rare new cases (n=13, 9.15%); the latter concerned SIT3001/H3, SIT3778/H3, SIT50/H3, SIT 47/H1, SIT219/T1, SIT 1160/LAM5, and SIT1/Beijing. Regarding drug-resistance to injectable drugs CAP and KAN (explained earlier, see also S2 Table), the following distribution patterns were noticed: CAP alone, n=22/142 (15.5%); KAN alone, n=31/142 (21.8%); and both CAP+KAN, n=89/142 (62.7%). Interestingly, the MST showed that only 5 shared-types (SIT50, 53, 219, 3001 and 3778) contained all the 3 patterns of drug resistance observed. Among the remaining cases, 4 shared-types (SIT1, 52, 93, and 49) contained strains with 2

drug resistance patterns (KAN-R and CAP+KAN both), while 2 shared-types (SIT47, 1355) strains with CAP-R and CAP+KAN.

Evolutionary relationships between Beijing lineages isolated from Peruvian XDR strains

A MIRU based MST (Fig. 2A) and WebLogo (Fig. 2B) were drawn to compare the Beijing lineage *M. tuberculosis* strains encountered in our study (n=13) vs. other countries in the SITVIT2 database (n=863) for which a significant number of strains with 15-loci typing data (35 or more) were available: Japan n=603, China n=209 and France n=35. The results obtained showed that a majority of the Peruvian strains have a specific phylogenetic position on the MST, close to one of the predominant MIRU-15 International Type-11 (15-MIT11), found in Japan (Fig. 2A). However, 3 Beijing strains from Peru are well isolated from others: 2 orphans per006 and per009 and a shared-type strain MIT234. As illustrated in Fig. 2A, per006 and 15-MIT234 are phylogenetically close to Japanese Beijing strains, while per009 at the top of the MST is close to Chinese patterns. The WebLogo representation (Fig. 2B) of each stack of symbols corresponding to each MIRU loci indicates that each region (Japan, China, France or Peru) has some specificities, but despite this fact, some correlation and similarity can be seen, notably between the Japanese, French and Peruvian (this study) strains. Despite these similarities, we can distinguish the Peruvian MIRU-15 isolates by its particular specific variation of number of copies on the 7^{th}, 9^{th} and 11^{th} loci positions. However, the statistical analysis based on the WebLogo data (data not shown) did not yield statistically significant variations by Fisher's Exact Test, essentially because of the small sample size of the Peruvian Beijing strains.

Discussion

TB is a prevalent disease in Peru. Since 2009 the government declared the disease as sanitary emergency due to the constant increase in the number of TB cases, mainly MDR-TB cases and the emergence of XDR-TB [19]. The distribution of TB cases is not homogeneous in Peru; the central coast (mainly in Lima) is the setting which presents the 90% of all TB cases and 96.7% of MDR-TB cases [19]. Since the first cases of XDR-TB were detected in Peru in 2007 [3], their number has been constantly increasing. Subsequent emergence of primary XDR-TB cases in children [4] has further worsened the situation, and controlling XDR-TB presents today a formidable challenge for the public health system. However, there is little information concerning the molecular epidemiology and genotypic diversity of circulating XDR-TB clones in Peru. There were only very few Peruvian strains reported in the previous versions of the databases including in SITVITWEB [11]. However, thanks to recent studies [20, 21, 22], and the fact that the number of Peruvian strains was considerably increased in the updated

SITVIT2 database (to almost 900), we thought it desirable to investigate the genetic diversity of XDR-TB isolates from Peru.

In the present study we analyzed 142 XDR-TB isolates which were grouped into 17 clusters. With the exception of 2 strains (1.41%) that did not match lineages reported so far, the remaining strains were distributed among following lineages: Haarlem (43.66%), T (27.46%), LAM (16.2%), Beijing (9.15%), and X clade (1.41%). Spoligotype led to a clustering for 90.1% strains vs. 34.5% by 15-loci MIRUs. Furthermore, 9 isolates (6.3%) were recorded as new SITs and 5 isolates (3.5%) corresponded to orphan patterns.

A recent study reported that the predominant genotypes in susceptible and resistant MTB isolates from Peru were LAM (23.8%), T (23.8%), Haarlem (22.3%), and Beijing (9.3%). Forty-three isolates were not reported previously (13.3%). The author concluded that the relatively high number of clusters suggests that recent transmission may be one major cause of the high incidence of TB in Peru [22]. Other studies carried out in Venezuela [23, 24], Paraguay [25], Honduras [26], and Brazil [27] showed similar results with LAM being the most prevalent lineage. For example, the predominant MTBC lineages in Brazil in decreasing order were: LAM (46%); the ill-defined T (18.6%); the Haarlem (12.2%), the X (4.7%), the S (1.9%), and the East African Indian (EAI) (0.85%) families [27]. Interestingly, this descriptive information of MTBC lineages/sublineages differs from our results on XDR isolates in Peru, where Haarlem was the most prevalent lineage, followed by T, LAM, and Beijing. One may notice that both SIT219/T1 and SIT50/H3 are the two most predominant SITs in our study: 14.8% each (**Table 2**), are also predominant in USA (**Table 3**). However, as opposed to SIT50/H3 which is equally present in North, Central and South America (between 4 to 6%), the SIT219/T1 is almost exclusively found in Peru ($p<0.001$), and to a lesser extent in North America (**S3 Table**). This observation may indirectly suggest that the strain was brought to the US by Peruvian migrants.

A study by Dalla Costa et.al [28] showed that Haarlem sublineage, mainly SIT50/H3, had a high frequency of *katG* S315T mutation in INH resistant MTB strains. Furthermore, it has been reported [29] that this sublineage presents mutations in certain genes allowing greater adaptability to hostile environments, such as those present following challenge by anti-TB drugs or engulfment by macrophages. These characteristics may partially explain the successful spread of Haarlem lineage strains, often associated with drug-resistance outbreaks in South America and elsewhere [28]. Though this explanation alone might not be sufficient to clarify the high prevalence of Haarlem sublineage in our study, mainly SIT50/H3 lineage strains, it could be partially responsible for the high prevalence of MDR-TB and XDR-TB cases observed. Future investigations should ideally focus on these aspects in Peru by comparing the relationship between Haarlem and other sublineages versus drug resistance mutations.

There is scarce information about genotyping studies in XDR isolates; in 2008 the first description of XDR genotypes concerned cases observed in South Africa [30]. Out of 41 isolates genotyped by spoligotyping, thirty-one isolates matched a previously described spoligotype; among these, the Beijing lineage was the largest

group (34%) followed by LAM, EAI, T, Haarlem, S and X3 sublineages. Another study done in Colombia characterized 10 XDR isolates [31] which were identified as SIT190/Beijing, SIT62/H1, SIT881/unknown, SIT545/LAM2 and SIT3010/S. Surprisingly, our results showed that the Beijing lineage is not predominant among Peruvian XDR isolates since the proportion of SIT1/Beijing strains in our study (9.15%) actually matches with the proportion of Beijing strains in other recently published studies from Peru [32, 33]. Ritacco et al. [33] speculated that the Beijing family strains were first introduced into Peru, and eventually into other South American countries, when Peru received a significant number of Chinese immigrants in the mid-19th century. In this context, our results using MST and WebLogo analysis shows that Peruvian Beijing XDR strains are more related to Japanese strains than Chinese strains (**Fig. 2**). This finding is possible because Peru also received Japanese immigrants at the end of the 19th century. A previous study [32] together with our results suggests the co-circulation of Beijing family with Japanese and Chinese ancestors in Peru.

In regards to the SIT distribution and the gender of patients, our study found that the difference between males and females was significant (p-value=0.039). We noticed that the proportion of SIT47/H1 was particularly important among females. The proportions of SIT1355/LAM, SIT3001/H3, and SIT3778/H3 were notably more important among male patients. Although the exact reasons are not well understood, these results might underline a preference for certain sublineages between male and female patients; nonetheless these differences should be verified in future analyses.

We also found that the XDR-TB affected slightly more females than males in our study (**Table 1**). This appears unusual since males are usually more frequently affected by TB than females, probably because of a higher exposition to various well-known risk factors. However, as high as 90% of the XDR-TB cases in our study concerned relapse cases, there is a possibility that the higher frequency of XDR-TB in female patients was due to a greater likelihood of treatment abandonment.

In Peru, the HIV cases are restricted to risk groups at difference to other countries. Pulmonary tuberculosis cases (susceptible, MDR and XDR tuberculosis) appear in general population. This is the reason for which HIV associated to tuberculosis appears in very low proportion (<3% of total TB cases in Peru, source: Ministry of Health). In our study, none of the 142 patients were HIV-positive. We may also mention that of our 227 cryopreserved samples, 2 samples were isolated from HIV-positive patients that unfortunately could not be reactivated upon subculturing.

In regards to the treatment history of patients, the majority were relapsed cases with a long history of treatment, initially for susceptible and then MDR-TB. The new cases almost always concerned predominant SITs (with the only exception of SIT1160), suggesting that MTBC isolates with acquired drug-resistance from retreated patients might be actively being transmitted to newly infected patients (primary XDR-TB cases).

We must nonetheless acknowledge limitations of the data presented here. Most importantly, our sampling strategy was opportunistic making use of a strain bank, and the period of sampling was relatively short (2.5 years). However, since XDR-TB cases represent 6% of MDR-TB cases, we considered that these cumulative cases of XDR strains over a period of 2.5 years might be considered as being representative of Peruvian XDR strains. There are advantages in having a strain bank which can be used for genotyping studies when epidemiological data are delinked from patient identifiers, but it also leads to obvious drawbacks, e.g., only limited clinical data are available and returning to clinical notes for further detail is not possible.

In conclusion, our study report for the first time in Peru, the genetic characterization and evolutionary relationships of XDR-TB strains, and highlights a significant proportion of Haarlem sublineage, followed by T sublineages – which are not among the usually predominant lineages in Peru. Furthermore, against all odds – we did not find Beijing lineage strains as the major cause of prevailing XDR-TB cases in Peru. Further studies are necessary to corroborate these results and to investigate whether these lineages continue to be a major cause of XDR-TB in Peru.

Supporting Information

S1 Figure. Dendrogram of Peruvian XDR-TB strains generated by MIRU-VNTRplus software (www.miru-vntrplus.org). The dendrogram shows three groups (I, II, III) containing 11clusters (n=34 strains; see text for details).
doi:10.1371/journal.pone.0112789.s001 (PDF)

S2 Figure. A minimum spanning tree illustrating the relationships between spoligotype patterns and the cities of isolation of the strains.
doi:10.1371/journal.pone.0112789.s002 (TIF)

S3 Figure. A spoligoforest tree drawn as Hierarchical Layout showing the parent to descendant relationships of the *M. tuberculosis* spoligotypes of Peruvian XDR isolates. The heuristic used selects a single inbound edge with a maximum weight using a Zipf model; solid black lines link patterns that are very similar, i.e., loss of one spacer only (maximum weigh being 1.0), while dashed lines represent links of weight comprised between 0.5 and 1, and dotted lines a weight less than 0.5. Note that orphan isolates (double circled), either appear at terminal positions on the tree, or as isolated strain without interconnections with the other.
doi:10.1371/journal.pone.0112789.s003 (TIF)

S4 Figure. A minimum spanning tree illustrating the relationships between spoligotype patterns and the treatment history of patients.
doi:10.1371/journal.pone.0112789.s004 (TIF)

S1 Table. Descriptive statistics on age of patients.
doi:10.1371/journal.pone.0112789.s005 (PDF)

S2 Table. Detailed genotyping and drug-resistance data and demographic information on *M. tuberculosis* XDR strains (n=142) isolated from adults with pulmonary tuberculosis in Peru.
doi:10.1371/journal.pone.0112789.s006 (PDF)

S3 Table. A comparison of the proportion of all SITs found in this study as compared to the other strains isolated in Peru and neighboring regions (Northern America, Southern America, Central America and Caribbean), recorded in the SITVIT2 database.
doi:10.1371/journal.pone.0112789.s007 (PDF)

Acknowledgments

We thank all professional staff of the Peruvian National Reference Laboratory of Mycobacteria, for the original isolations of the XDR-TB strains and drug resistance identification. This research was supported by the Peruvian National Institute of Health and International Association of National Public Health Institutes (IANPHI) Research Seed Grant (www.ianphi.org).

Author Contributions

Conceived and designed the experiments: OC AM LA. Performed the experiments: OC CB MG. Analyzed the data: OC DC NR. Contributed reagents/materials/analysis tools: OC LA AM. Wrote the paper: OC NR DC AM.

References

1. **World Health Organization** (2013) Global tuberculosis report. WHO, Geneva.
2. **Zignol M, van Gemert W, Falzon D, Sismanidis C, Glaziou P, et al.** (2012) Surveillance of anti-tuberculosis drug resistance in the world: an updated analysis, 2007–2010. Bull World Health Organ 90:111–119.
3. **Mendoza-Ticona A, Asencios-Solís L, Quispe-Torres N, Leo-Hurtado E** (2007). Evidencia de tuberculosis con resistencia extendida a drogas de segunda linea (TB-XDR) en el Perú. Rev Peru Med Exp Salud Publica 24(3):313–314.
4. **Del Castillo H, Mendoza-Ticona A, Saravia JC, Somocurcio JG** (2009) Epidemia de tuberculosis multidrogo resistente y Extensivamente resistente a drogas (TB MDR/XDR) en el Perú: situación y propuestas para su control. Rev Peru Med Exp Salud Publica 26(3): 380–386.
5. **Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, et al.** (1997) Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol 35(4):907–914.
6. **Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D, et al.** (2001) Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. J. Clin. Microbiol 39:3563–3571.
7. **Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, et al.** (2006) Proposal for Standardization of Optimized Mycobacterial Interspersed Repetitive Unit–Variable-Number Tandem Repeat Typing of *Mycobacterium tuberculosis*. J Clin Microbiol 44(12): 4498–4510.
8. **Alonso-Rodríguez N, Martínez-Lirola M, Herránz M, Sanchez-Benitez M, Barroso P, et al.** (2008) Evaluation of the new advanced 15-loci MIRU-VNTR genotyping tool in *Mycobacterium tuberculosis* molecular epidemiology studies. *BMC Microbiology* 8:34.

9. Kent PT, Kubica GP, (ed). Public health mycobacteriology. A guide for the level III laboratory. Atlanta, GA: U.S. Department of Health, Education, and Welfare, Centers for Disease Control and Prevention; 1995.

10. **van Soolingen D, Hermans PW, de Haas PEW, Soll DR, van Embden JD** (1991) Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. J Clin Microbiol 29:2578–2586.

11. **Demay C, Liens B, Burguière T, Hill V, Couvin D, et al.** (2012) SITVITWEB–a publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. Infect Genet Evol 12(4):755–66.

12. **Driscoll JR, Bifani PJ, Mathema B, McGarry MA, Zickas GM, et al.** (2002). Spoligologos: a bioinformatic approach to displaying and analyzing *Mycobacterium tuberculosis* data. Emerg Infect Dis 8:1306–1309.

13. **Crooks GE, Hon G, Chandonia JM, Brenner SE** (2004) WebLogo: A sequence logo generator. Genome Research 14:1188–1190.

14. **Schneider TD, Stephens RM** (1990) Sequence Logos: A New Way to Display Consensus Sequences. Nucleic Acids Res 18:6097–6100.

15. **Tang C, Reyes JF, Luciani F, Francis AR, Tanaka MM** (2008) SpolTools: online utilities for analyzing spoligotypes of the *Mycobacterium tuberculosis* complex. Bioinformatics 24(20):2414–5.

16. **Reyes JF, Francis AR, Tanaka MM** (2008) Models of deletion for visualizing bacterial variation: an application to tuberculosis spoligotypes. BMC Bioinformatics 9:496.

17. **Hunter PR, Gaston MA** (1988) Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J Clin Microbiol 26(11):2465–2466.

18. **Allix-Beguec C, Harmsen D, Weniger T, Supply P, Niemann S** (2008) Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. J Clin Microbiol 46(8):2692–2699.

19. **Ministerio de Salud.** Dirección de Salud V Lima Ciudad: Análisis de la situación de salud 2010 de la Dirección de Salud V Lima Ciudad. Lima, Peru: Dirección de Salud V Lima Ciudad. Available: 2011 Jan 20 http://www.disavlc.gob.pe/epi/

20. **Barletta F, Otero L, Collantes J, Asto B, de Jong BC, et al.** (2013) Genetic variability of *Mycobacterium tuberculosis* complex in patients with no known risk factors for MDR-TB in the North-eastern part of Lima, Peru. BMC Infectious Diseases 13:397.

21. **Sheen P, Couvin D, Grandjean L, Zimic M, Dominguez M, et al.** (2013) Genetic diversity of *Mycobacterium tuberculosis* in Peru and exploration of phylogenetic associations with drug resistance. PLoS ONE 8(6): e65873.

22. **Taype CA, Agapito JC, Accinelli RA, Espinoza JR, Godreuil S, et al.** (2012). Genetic diversity, population structure and drug resistance of *Mycobacterium tuberculosis* in Peru. Infect Genet Evol 12: 577–585.

23. **Aristimuno L, Armengol R, Cebollada A, Espana M, Guilarte A, et al.** (2006) Molecular characterisation of *Mycobacterium tuberculosis* isolates in the First National Survey of Anti-tuberculosis Drug Resistance from Venezuela. BMC Microbiol 6:90.

24. **Abadía E, Sequera M, Ortega D, Méndez MV, Escalona A, et al.** (2009) *Mycobacterium tuberculosis* ecology in Venezuela: epidemiologic correlates of common spoligotypes and a large clonal cluster defined by MIRU-VNTR-24. BMC Infect Dis 6(9):122.

25. **Candia N, Lopez B, Zozio T, Carrivale M, Diaz C, et al.** (2007) First insight into *Mycobacterium tuberculosis* genetic diversity in Paraguay. BMC Microbiol 8(7):75.

26. **Rosales S, Pineda-García L, Ghebremichael S, Rastogi N, Hoffner SE** (2010) Molecular diversity of *Mycobacterium tuberculosis* isolates from patients with tuberculosis in Honduras. BMC Microbiology 10:208.

27. **Gomes HM, Elias AR, Oelemann MA, Pereira MA, Montes FF, et al.** (2012) Spoligotypes of *Mycobacterium tuberculosis* complex isolates from patient's residents of 11 states of Brazil. Infect Genet Evol 12(4):649–56.

28. **Dalla Costa ER, Ribeiro MO, Silva MS, Arnold LS, Rostirolla DC, et al.** (2009) Correlations of mutations in katG, oxyR-ahpC and inhA genes and in vitro susceptibility in *Mycobacterium tuberculosis* clinical strains segregated by spoligotype families from tuberculosis prevalent countries in South America. BMC Microbiol 9:39.

29. **Olano J, López B, Reyes A, Del Pilar Lemos M, Correa N, et al.** (2007) Mutations in DNA repair genes are associated with the Haarlem lineage of *Mycobacterium tuberculosis* independently of their antibiotic resistance. Tuberculosis (Edinb) 87(6):502–8.

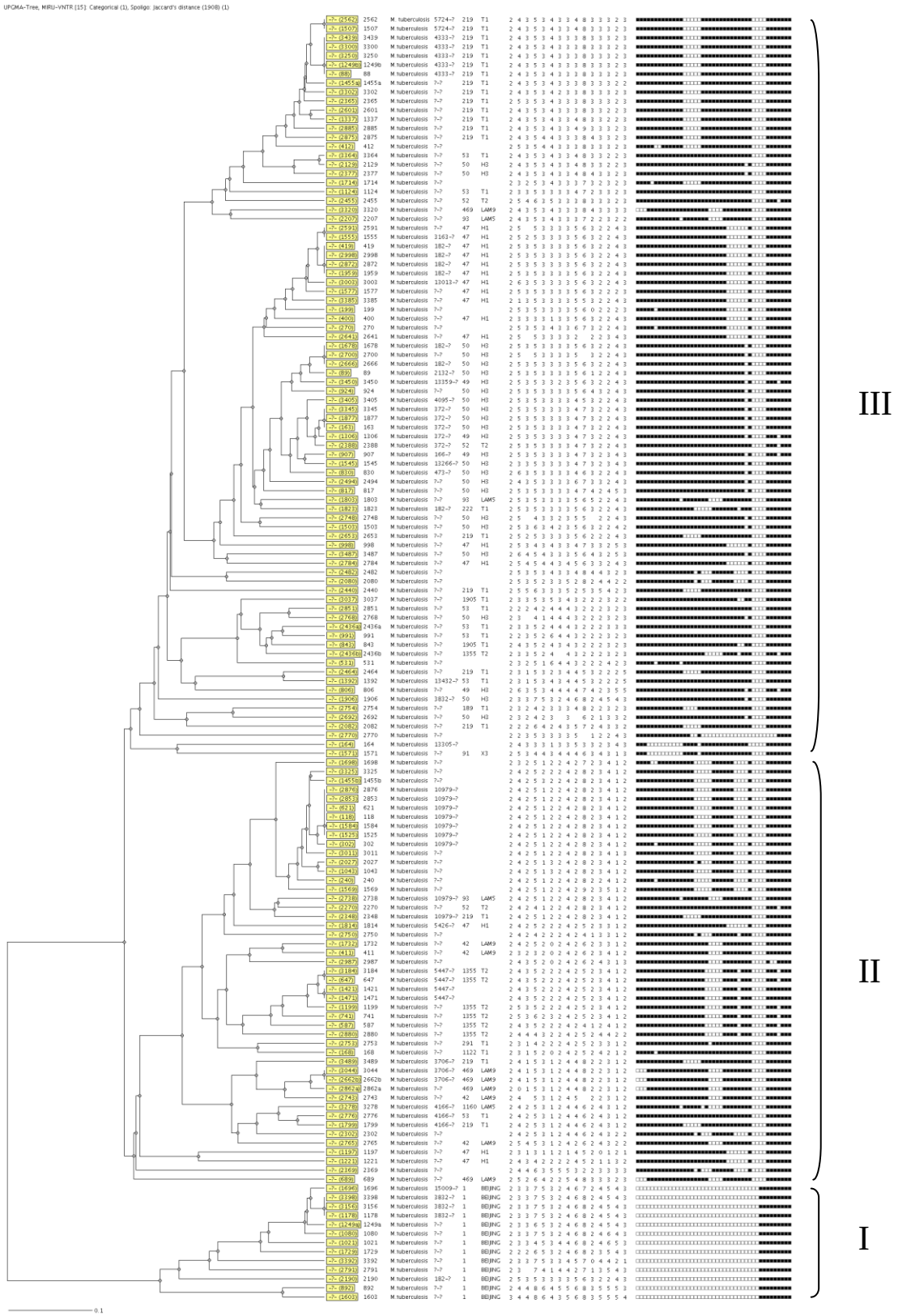
30. **Mlambo CK, Warren RM, Poswa X, Victor TC, Duse AG, et al.** (2008) Genotypic diversity of extensively drug-resistant tuberculosis (XDR-TB) in South Africa. Int J Tuberc Lung Dis 12(1):99–104

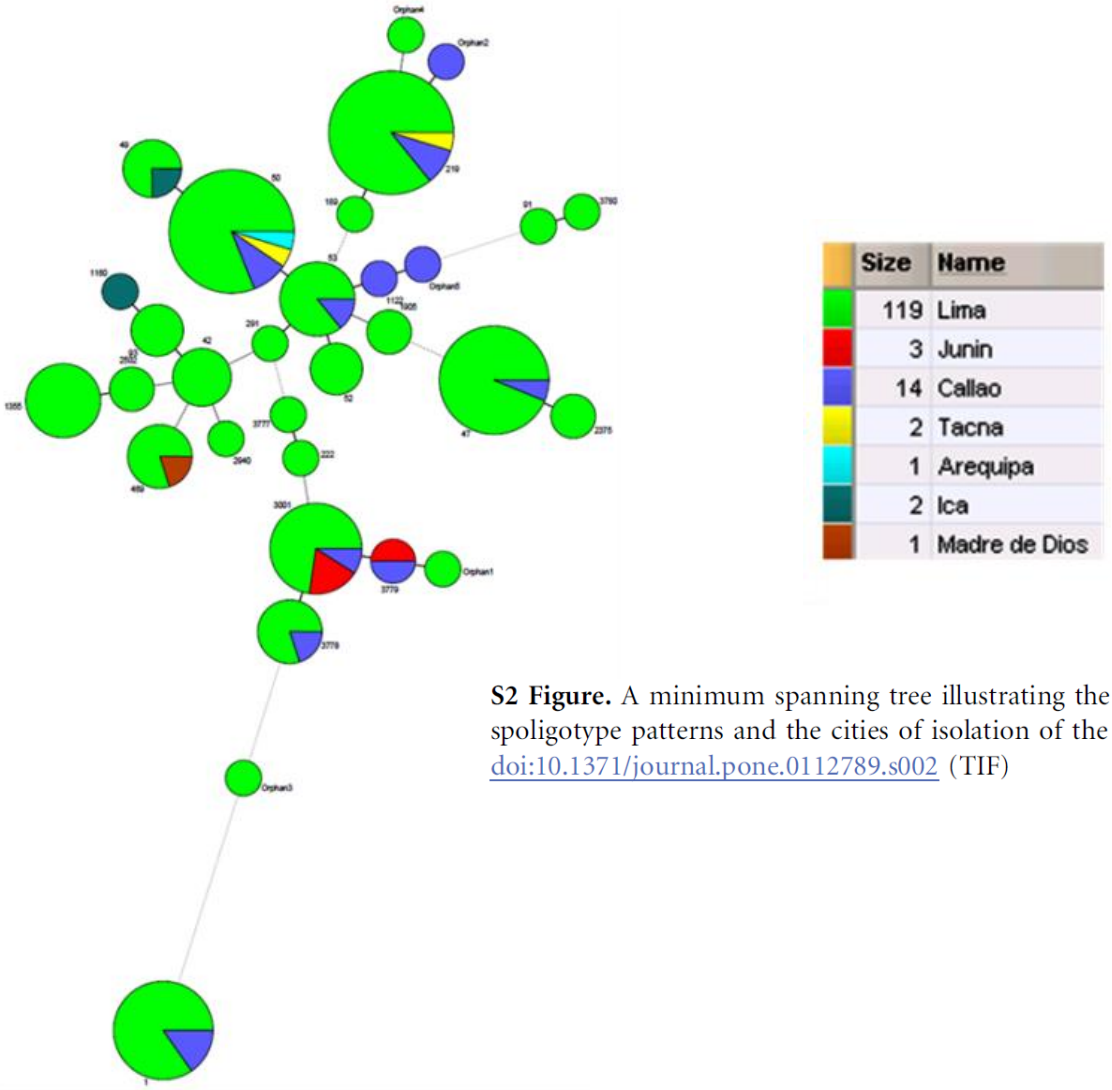
31. **Nieto LM, Ferro BE, Villegas SL, Mehaffy C, Forero L, et al.** (2012) Characterization of Extensively Drug-Resistant Tuberculosis Cases from Valle del Cauca, Colombia J. Clin. Microbiol 50(12):4185–4187.

32. **Iwamoto T, Grandjean L, Arikawa K, Nakanishi N, Caviedes L, et al.** (2012) Genetic Diversity and Transmission Characteristics of Beijing Family Strains of *Mycobacterium tuberculosis* in Peru. PLoS ONE 7(11): e49651.

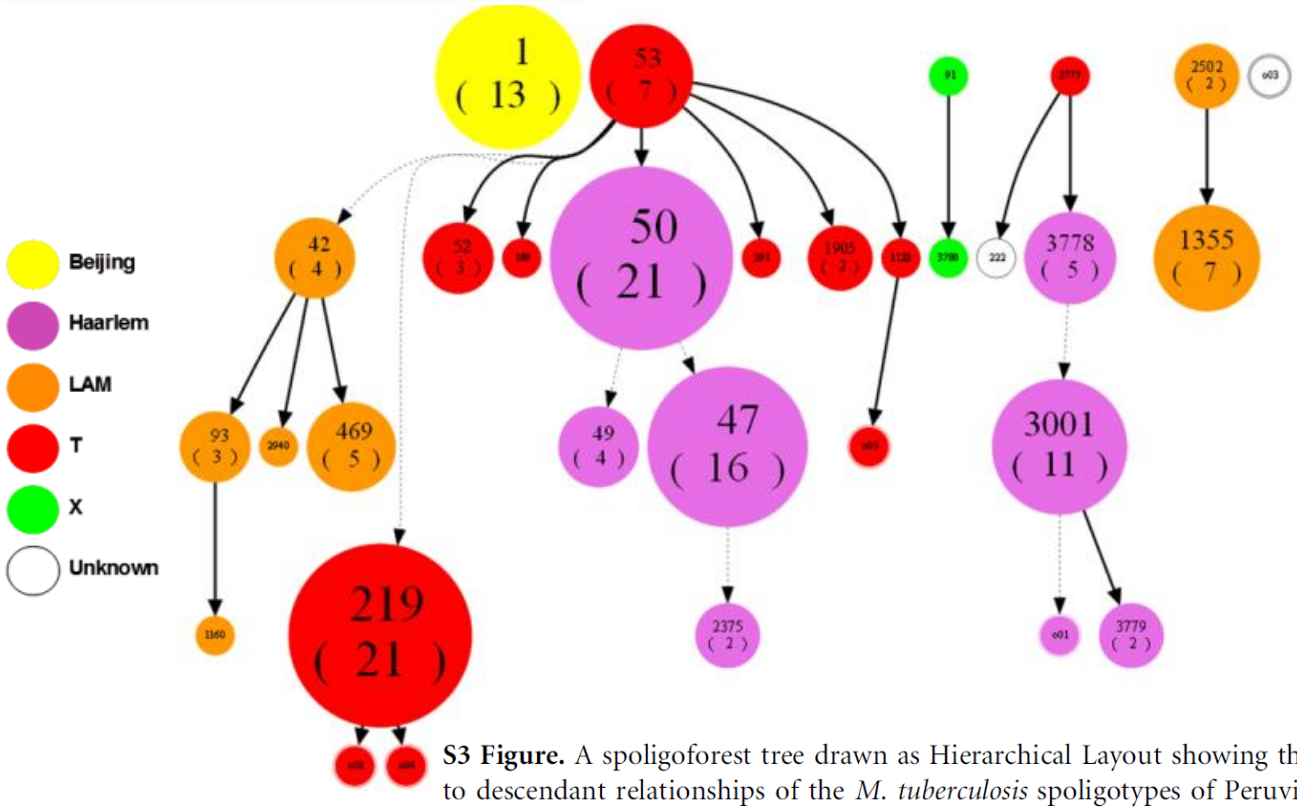
33. **Ritacco V, Lopez B, Cafrune PI, Ferrazoli L, Suffys PN, et al.** (2008) *Mycobacterium tuberculosis* strains of the Beijing genotype are rarely observed in tuberculosis patients in South America. Mem Inst Oswaldo Cruz 103: 489–492.

Figure S1. Dendrogram of Peruvian M. tuberculosis XDR strains generated by MIRU-VNTRplus software (www.miru-vntrplus.org). The dendrogram shows three groups (I, II, III) which grouped 11 clusters containing 34 strains (See text)

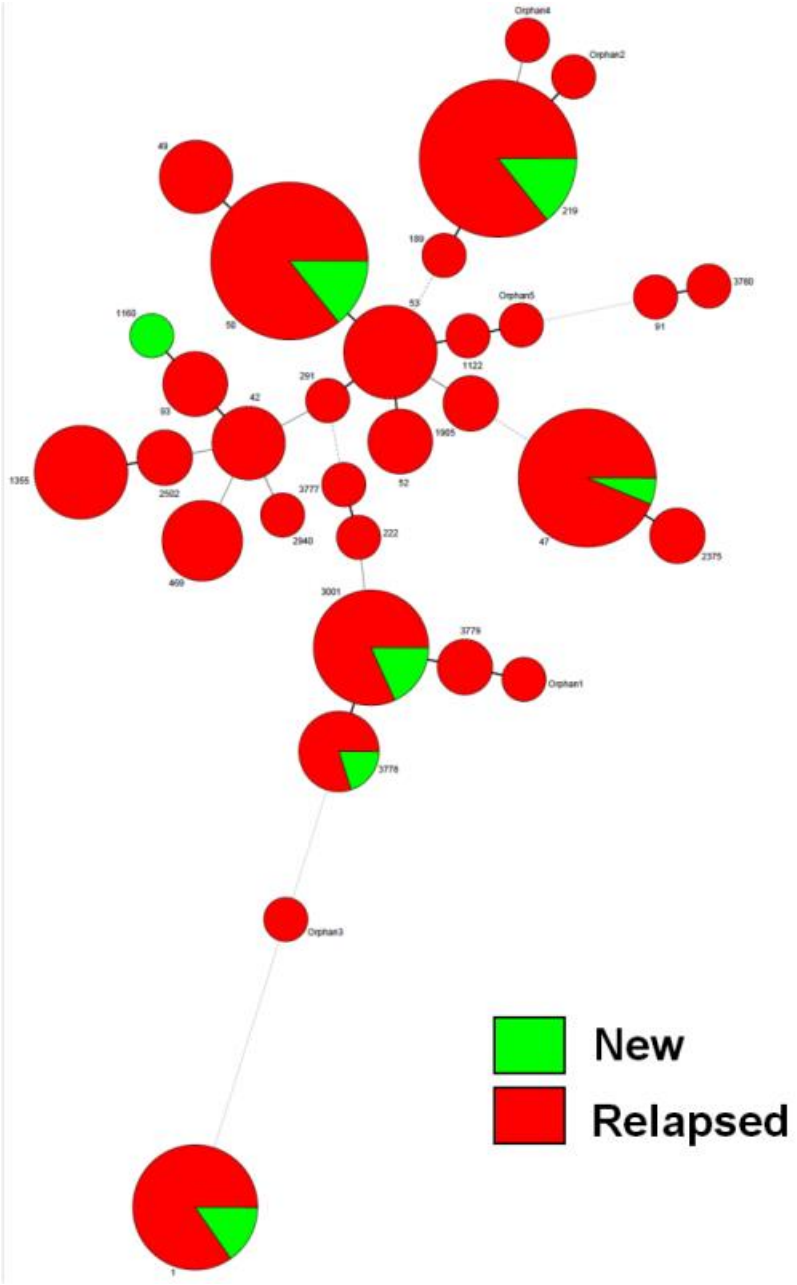




S2 Figure. A minimum spanning tree illustrating the relationships between spoligotype patterns and the cities of isolation of the strains.
doi:10.1371/journal.pone.0112789.s002 (TIF)



S3 Figure. A spoligoforest tree drawn as Hierarchical Layout showing the parent to descendant relationships of the *M. tuberculosis* spoligotypes of Peruvian XDR isolates. The heuristic used selects a single inbound edge with a maximum weight using a Zipf model; solid black lines link patterns that are very similar, i.e., loss of one spacer only (maximum weigh being 1.0), while dashed lines represent links of weight comprised between 0.5 and 1, and dotted lines a weight less than 0.5. Note that orphan isolates (double circled), either appear at terminal positions on the tree, or as isolated strain without interconnections with the other.



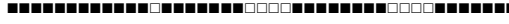
S4 Figure. A minimum spanning tree illustrating the relationships between spoligotype patterns and the treatment history of patients.
doi:10.1371/journal.pone.0112789.s004 (TIF)

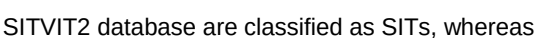
Supplemental Table S1: Descriptive statistics on age of patients

	Male patients	*Female patients*	***Total***
Count	90	52	**142**
Mean	36.33	30.06	**34.04**
Sample variance	180.49	108.88	**162.52**
Sample standard deviation	13.43	10.43	**12.75**
Minimum	15	17	**15**
Maximum	72	70	**72**
Range	57	53	**57**
1st quartile	26.00	23.00	**24.00**
Median	34.00	26.00	**31.00**
3rd quartile	46.50	34.00	**42.00**
Interquartile range	20.50	11.00	**18.00**
Mode	40.00	24.00	**24.00**

Supplemental Table S2: Detailed genotyping and drug-resistance data and demographic information on cryopreserved M. tuberculosis XDR strains (n=142) isolated from adults with pulmonary tuberculosis in Peru.

IsoNumber	Year	Strain	Sex/age	Isolation city	Spoligotype Description	Octal code	Clade	SIT	MIRU15	15-MIT	Drug-R to CAP or KAN or both	Patient status
PER0620094892	2009	892	F/41	Callao	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	244864556835553	Orph007	CAP S+KAN R	Relapsed
PER06200941021	2009	1021	M/20	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	233453446824653	Orph002	CAP S+KAN R	Relapsed
PER06200941080	2009	1080	M/26	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	233753246824643	Orph005	CAP S+KAN R	Relapsed
PER06200941178	2009	1178	F/25	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	233753246824543	235	CAP S+KAN R	Relapsed
PER06200941603	2009	1603	F/20	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	344864356835554	Orph008	CAP S+KAN R	Relapsed
PER06200741696	2007	1696	M/25	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	233753246724543	Orph004	CAP R+KAN R	Relapsed
PER06200741729	2007	1729	M/40	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	222653246823543	Orph001	CAP S+KAN R	Relapsed
PER06200842190	2008	2190	M/50	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	253533335632243	234	CAP R+KAN R	New
PER06200842791	2008	2791	M/19	Callao	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	23-741442713543	Orph009	CAP S+KAN R	Relapsed
PER06200843156	2008	3156	M/40	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	233753246824543	235	CAP S+KAN R	Relapsed
PER06200843392	2008	3392	F/24	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	233753345704421	Orph006	CAP S+KAN R	Relapsed
PER06200843398	2008	3398	M/28	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	233753246824543	235	CAP R+KAN R	Relapsed
PER06200741249a	2007	1249a	M/50	Lima	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	000000000003771	Beijing	1	233653246824543	Orph003	CAP R+KAN R	New
PER0620094411	2009	411	M/40	Lima	■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■■	777777607760771	LAM9	42	232320242623412	Orph010	CAP R+KAN R	Relapsed
PER06200741732	2007	1732	F/44	Lima	■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■■	777777607760771	LAM9	42	242520242623312	Orph011	CAP R+KAN R	Relapsed
PER06200842743	2008	2743	F/34	Lima	■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■■	777777607760771	LAM9	42	24-531245-22312	Orph013	CAP R+KAN R	Relapsed
PER06200842765	2008	2765	F/56	Lima	■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■■	777777607760771	LAM9	42	254531242624322	Orph012	CAP R+KAN R	Relapsed
PER0620074400	2007	400	F/22	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	233331335632243	Orph016	CAP R+KAN S	Relapsed
PER0620074419	2007	419	F/26	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	253533335632243	234	CAP R+KAN R	Relapsed
PER0620094998	2009	998	F/40	Callao	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	253434334733253	Orph020	CAP R+KAN R	Relapsed
PER06200741197	2007	1197	F/24	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	231311214520121	Orph015	CAP R+KAN S	Relapsed
PER06200741221	2007	1221	F/22	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	243422224521132	Orph018	CAP R+KAN S	Relapsed
PER06200741555	2007	1555	F/26	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	252533335632243	Orph019	CAP R+KAN R	Relapsed
PER06200941577	2009	1577	M/53	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	253533335632223	Orph021	CAP R+KAN R	Relapsed
PER06200741814	2007	1814	M/41	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	242522242523312	Orph017	CAP R+KAN R	Relapsed
PER06200841959	2008	1959	F/25	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	253533335632243	234	CAP R+KAN S	Relapsed
PER06200842591	2008	2591	F/24	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	25-533335632243	Orph025	CAP R+KAN S	Relapsed
PER06200842641	2008	2641	M/43	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	25-533332-22343	Orph024	CAP R+KAN S	Relapsed
PER06200842784	2008	2784	F/36	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	254544345633243	Orph022	CAP R+KAN S	New
PER06200842872	2008	2872	F/26	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	253533335632243	234	CAP R+KAN R	Relapsed
PER06200842998	2008	2998	M/24	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	253533335632243	234	CAP R+KAN S	Relapsed
PER06200843003	2008	3003	F/22	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	263533335632243	Orph023	CAP R+KAN S	Relapsed
PER06200843385	2008	3385	F/25	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	47	213533335532243	Orph014	CAP R+KAN S	Relapsed
PER0620094806	2009	806	M/47	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■□■■■	777777777720731	H3	49	263534444742355	Orph027	CAP S+KAN R	Relapsed
PER0620094907	2009	907	M/15	Ica	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■□■■■	777777777720731	H3	49	253533334732343	369	CAP S+KAN R	Relapsed
PER06200941306	2009	1306	M/19	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■□■■■	777777777720731	H3	49	253533334732243	240	CAP S+KAN R	Relapsed
PER06200843450	2008	3450	F/24	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■□■■■	777777777720731	H3	49	253533235632243	Orph026	CAP R+KAN R	Relapsed
PER062008489	2008	89	F/39	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	253533335612243	Orph033	CAP S+KAN R	Relapsed
PER0620084163	2008	163	M/25	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	253533334732243	240	CAP R+KAN R	Relapsed
PER0620094817	2009	817	M/72	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	253533334742453	Orph032	CAP R+KAN R	Relapsed
PER0620094830	2009	830	M/32	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	263533334632243	Orph036	CAP S+KAN R	Relapsed
PER0620094924	2009	924	M/18	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	253533335643243	Orph034	CAP R+KAN R	New
PER06200941503	2009	1503	F/43	Callao	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	253634235632242	Orph035	CAP R+KAN R	New
PER06200941545	2009	1545	F/24	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	233533334732343	Orph028	CAP R+KAN R	Relapsed
PER06200741678	2007	1678	M/26	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	253533335632243	234	CAP R+KAN R	Relapsed
PER06200841877	2008	1877	M/20	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	253533334732243	240	CAP R+KAN R	New
PER06200841906	2008	1906	M/25	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	233753246824543	235	CAP R+KAN R	Relapsed
PER06200842129	2008	2129	M/29	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	243534334833223	237	CAP S+KAN R	Relapsed
PER06200842377	2008	2377	M/54	Callao	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	243534334843323	Orph030	CAP R+KAN S	Relapsed
PER06200842494	2008	2494	F/34	Tacna	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	243533336733243	Orph029	CAP R+KAN S	Relapsed
PER06200842666	2008	2666	F/26	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	253533335632243	234	CAP R+KAN R	Relapsed
PER06200842692	2008	2692	M/29	Lima	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	50	232423-3-621332	Orph038	CAP R+KAN R	Relapsed

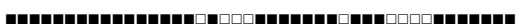
PER06200842700	2008	2700	M/63	Arequipa		777777777720771	H3	50	25-533335-32243	Orph041	CAP R+KAN R	Relapsed
PER06200842748	2008	2748	F/27	Lima		777777777720771	H3	50	25-4332355-2243	Orph040	CAP R+KAN R	Relapsed
PER06200842768	2008	2768	F/45	Lima		777777777720771	H3	50	23-414443222323	Orph039	CAP R+KAN R	Relapsed
PER06200843345	2008	3345	M/72	Lima		777777777720771	H3	50	253533334732243	240	CAP R+KAN R	Relapsed
PER06200843405	2008	3405	M/43	Lima		777777777720771	H3	50	253533334532243	Orph031	CAP R+KAN R	Relapsed
PER06200843487	2008	3487	M/26	Lima		777777777720771	H3	50	264543335643253	Orph037	CAP R+KAN R	Relapsed
PER06200842270	2008	2270	F/31	Lima		777777777760731	T2	52	242412242823412	Orph042	CAP S+KAN R	Relapsed
PER06200842388	2008	2388	M/72	Lima		777777777760731	T2	52	253533334732243	240	CAP S+KAN R	Relapsed
PER06200842455	2008	2455	M/42	Lima		777777777760731	T2	52	254635333833323	Orph043	CAP R+KAN R	Relapsed
PER0620094991	2009	991	M/28	Callao		777777777760771	T1	53	223526443222323	Orph045	CAP R+KAN R	Relapsed
PER06200941124	2009	1124	M/24	Lima		777777777760771	T1	53	233533334723323	Orph048	CAP R+KAN R	Relapsed
PER06200941392	2009	1392	F/21	Lima		777777777760771	T1	53	231534344532225	Orph046	CAP S+KAN R	Relapsed
PER06200842776	2008	2776	F/32	Lima		777777777760771	T1	53	242531244624312	239	CAP R+KAN R	Relapsed
PER06200842851	2008	2851	F/41	Lima		777777777760771	T1	53	222424443222323	Orph044	CAP R+KAN R	Relapsed
PER06200843364	2008	3364	F/29	Lima		777777777760771	T1	53	243534334833223	237	CAP R+KAN S	Relapsed
PER062008424 36a	2008	2436a	M/38	Lima		777777777760771	T1	53	233524443222333	Orph047	CAP R+KAN R	Relapsed
PER06200941571	2009	1571	M/30	Lima		700036777760771	X3	91	253443444634313	Orph049	CAP R+KAN R	Relapsed
PER06200741803	2007	1803	F/26	Lima		777737607760771	LAM5	93	253533335652243	Orph051	CAP R+KAN R	Relapsed
PER06200842207	2008	2207	F/23	Lima		777737607760771	LAM5	93	243534333722322	Orph050	CAP S+KAN R	Relapsed
PER06200842738	2008	2738	M/17	Lima		777737607760771	LAM5	93	242512242823412	242	CAP R+KAN R	Relapsed
PER06200842754	2008	2754	F/23	Lima		777741777760771	T1	189	232423334822323	Orph052	CAP R+KAN R	Relapsed
PER062008488	2008	88	F/23	Lima		777740777760771	T1	219	243534333833323	236	CAP R+KAN R	Relapsed
PER06200941337	2009	1337	M/24	Lima		777740777760771	T1	219	243534334833223	237	CAP S+KAN R	New
PER06200741507	2007	1507	F/23	Lima		777740777760771	T1	219	243534334833323	233	CAP S+KAN R	Relapsed
PER06200741799	2007	1799	M/43	Lima		777740777760771	T1	219	242531244624312	239	CAP R+KAN R	Relapsed
PER06200842082	2008	2082	F/24	Lima		777740777760771	T1	219	222642435724332	Orph053	CAP S+KAN R	Relapsed
PER06200842348	2008	2348	F/31	Lima		777740777760771	T1	219	242512242823412	242	CAP R+KAN R	Relapsed
PER06200842365	2008	2365	M/33	Lima		777740777760771	T1	219	253534333833323	Orph061	CAP R+KAN R	Relapsed
PER06200842440	2008	2440	M/22	Lima		777740777760771	T1	219	255633352535423	Orph062	CAP R+KAN S	Relapsed
PER06200842464	2008	2464	F/34	Tacna		777740777760771	T1	219	231532344532225	Orph054	CAP R+KAN S	Relapsed
PER06200842562	2008	2562	F/21	Lima		777740777760771	T1	219	243534334833323	233	CAP R+KAN R	Relapsed
PER06200842601	2008	2601	M/27	Lima		777740777760771	T1	219	243534333833223	Orph056	CAP R+KAN S	Relapsed
PER06200842653	2008	2653	M/34	Lima		777740777760771	T1	219	252533335622243	Orph060	CAP R+KAN R	Relapsed
PER06200842875	2008	2875	M/31	Lima		777740777760771	T1	219	243544333843323	Orph059	CAP R+KAN R	Relapsed
PER06200842885	2008	2885	M/54	Callao		777740777760771	T1	219	243534334933323	Orph058	CAP R+KAN S	Relapsed
PER06200843250	2008	3250	F/23	Lima		777740777760771	T1	219	243534333833323	236	CAP R+KAN R	Relapsed
PER06200843300	2008	3300	F/19	Lima		777740777760771	T1	219	243534333833323	236	CAP R+KAN R	Relapsed
PER06200843302	2008	3302	M/16	Callao		777740777760771	T1	219	243534233833323	Orph055	CAP R+KAN R	Relapsed
PER06200843439	2008	3439	M/21	Lima		777740777760771	T1	219	243534333833323	236	CAP R+KAN S	Relapsed
PER06200843489	2008	3489	M/24	Lima		777740777760771	T1	219	241531244822312	238	CAP S+KAN R	New
PER062007412 49b	2007	1249b	M/50	Lima		777740777760771	T1	219	243534333833323	236	CAP R+KAN R	New
PER062007414 55a	2007	1455a	M/27	Lima		777740777760771	T1	219	243534333833322	Orph057	CAP R+KAN S	Relapsed
PER06200741823	2007	1823	M/47	Lima		777774077560771	Unknown	222	253533335632243	234	CAP R+KAN R	Relapsed
PER06200842753	2008	2753	M/32	Lima		777777677760771	T1	291	231422242523312	Orph063	CAP R+KAN R	Relapsed
PER0620094689	2009	689	F/20	Lima		077777607760771	LAM1	469	252642254833323	Orph066	CAP R+KAN R	Relapsed
PER06200843044	2008	3044	M/35	Lima		077777607760771	LAM1	469	241531244822312	238	CAP R+KAN R	Relapsed
PER06200843320	2008	3320	M/25	Madre de Dios		077777607760771	LAM1	469	243534333843333	Orph065	CAP R+KAN R	Relapsed
PER062008428 62a	2008	2862a	M/32	Lima		077777607760771	LAM1	469	201531244822312	Orph064	CAP R+KAN R	Relapsed
PER062008428 62b	2008	2862b	M/32	Lima		077777607760771	LAM1	469	241531244822312	238	CAP R+KAN R	Relapsed
PER0620084168	2008	168	M/49	Callao		767777777760771	T1	1122	231520242524212	Orph067	CAP S+KAN R	Relapsed
PER06200843298	2008	3298	F/17	Ica		777737207760771	LAM5	1160	242531244624312	239	CAP R+KAN R	New
PER0620094587	2009	587	M/38	Lima		777777407560731	LAM	1355	243522242412412	Orph068	CAP R+KAN R	Relapsed
PER0620094647	2009	647	M/51	Lima		777777407560731	LAM	1355	243522242523412	241	CAP R+KAN R	Relapsed
PER0620094741	2009	741	M/55	Lima		777777407560731	LAM	1355	253623242523412	Orph071	CAP R+KAN R	Relapsed
PER06200941199	2009	1199	M/55	Lima		777777407560731	LAM	1355	253522242523412	Orph070	CAP R+KAN R	Relapsed
PER06200842880	2008	2880	F/23	Lima		777777407560731	LAM	1355	244432242524422	Orph069	CAP R+KAN R	Relapsed
PER06200843184	2008	3184	M/49	Lima		777777407560731	LAM	1355	243522242523412	241	CAP R+KAN S	Relapsed
PER062008424 36b	2008	2436b	M/38	Lima		777777407560731	LAM	1355	233524-43222323	Orph072	CAP R+KAN R	Relapsed
PER0620094843	2009	843	M/48	Lima		777777777460771	T1	1905	243524343222323	Orph074	CAP S+KAN R	Relapsed
PER06200843037	2008	3037	M/48	Lima		777777777460771	T1	1905	233535343222322	Orph073	CAP S+KAN R	Relapsed
PER0620084199	2008	199	M/53	Lima		767777774020771	H1	2375	253533335602223	Orph075	CAP R+KAN R	Relapsed

PER0620084270	2008	270	M/34	Lima		767777774020771	H1	2375	253534336732243	Orph076	CAP R+KAN R	Relapsed
PER06200941421	2009	1421	M/51	Lima		777777407560771	LAM6	2502	243522242523412	241	CAP R+KAN R	Relapsed
PER06200941471	2009	1471	F/34	Lima		777777407560771	LAM6	2502	243522242523412	241	CAP R+KAN R	Relapsed
PER06200842987	2008	2987	M/34	Lima		777777607760471	LAM9	2940	243520242624313	Orph077	CAP S+KAN R	Relapsed
PER0620084118	2008	118	F/70	Junin		777774077020771	H3	3001	242512242823412	242	CAP R+KAN R	Relapsed
PER0620094621	2009	621	M/35	Lima		777774077020771	H3	3001	242512242823412	242	CAP R+KAN R	Relapsed
PER06200941043	2009	1043	M/52	Lima		777774077020771	H3	3001	242513242823412	243	CAP R+KAN R	Relapsed
PER06200941525	2009	1525	M/22	Lima		777774077020771	H3	3001	242512242823412	242	CAP R+KAN R	New
PER06200941569	2009	1569	M/61	Lima		777774077020771	H3	3001	242512242923512	Orph079	CAP S+KAN R	Relapsed
PER06200941584	2009	1584	M/42	Junin		777774077020771	H3	3001	242512242823412	242	CAP S+KAN R	Relapsed
PER06200842853	2008	2853	M/42	Callao		777774077020771	H3	3001	242512242823412	242	CAP R+KAN R	Relapsed
PER06200842876	2008	2876	F/40	Lima		777774077020771	H3	3001	242512242823412	242	CAP R+KAN R	Relapsed
PER06200843011	2008	3011	M/42	Lima		777774077020771	H3	3001	242512242823413	Orph078	CAP R+KAN R	Relapsed
PER06200843325	2008	3325	M/22	Lima		777774077020771	H3	3001	242522242823412	Orph080	CAP R+KAN R	New
PER06200741455b	2007	1455b	M/27	Lima		777774077020771	H3	3001	242532242823412	Orph081	CAP R+KAN S	Relapsed
PER06200842750	2008	2750	M/26	Lima		777775077560771	T1	3777*	242422242413312	Orph082	CAP R+KAN R	Relapsed
PER06200842027	2008	2027	M/40	Callao		777775077020771	H3	3778*	242513242823412	243	CAP R+KAN R	Relapsed
PER06200842080	2008	2080	F/27	Lima		777775077020771	H3	3778*	253523352824422	Orph085	CAP R+KAN R	Relapsed
PER06200842302	2008	2302	M/16	Lima		777775077020771	H3	3778*	242531244624322	Orph083	CAP R+KAN S	Relapsed
PER06200842369	2008	2369	M/38	Lima		777775077020771	H3	3778*	244635553223333	Orph084	CAP R+KAN R	Relapsed
PER06200842482	2008	2482	M/24	Lima		777775077020771	H3	3778*	253534334844323	Orph086	CAP S+KAN R	New
PER0620084240	2008	240	M/45	Junin		767774077020771	H3	3779*	242512242822412	Orph087	CAP R+KAN R	Relapsed
PER0620094302	2009	302	M/40	Callao		767774077020771	H3	3779*	242512242823412	242	CAP R+KAN R	Relapsed
PER0620084164	2008	164	F/53	Lima		700036777740771	X3	3780*	243331335332343	Orph088	CAP R+KAN R	Relapsed
PER06200741698	2007	1698	F/31	Lima		747774077020771	H3	Orphan1	232512242723412	Orph089	CAP R+KAN R	Relapsed
PER06200741714	2007	1714	M/31	Callao		757740777760771	T1	Orphan2	232534333732323	Orph091	CAP R+KAN R	Relapsed
PER06200842770	2008	2770	M/27	Lima		777771000000071	Unknown	Orphan3	223533335-12243	Orph093	CAP S+KAN R	Relapsed
PER0620094412	2009	412	M/33	Lima		763740777760771	T1	Orphan4	253544333833323	Orph092	CAP R+KAN R	Relapsed
PER0620094531	2009	531	M/28	Callao		767737777760771	T3	Orphan5	232516443222423	Orph090	CAP R+KAN R	Relapsed

Unique strains matching a preexisting pattern in the SITVIT2 database are classified as SITs, whereas in case of no match, they are designated as ‘‘orphan", and highlighted in blue.

*SITs followed by an asterisk and highlighted in yellow indicate "newly created shared-type" after match with another orphan in the database, or due to 2 or more strains belonging to a new pattern within this study.

Supplemental Table S3: A comparison of the proportion of all SITs found in this study as compared to the other strains isolated in Peru and neighbouring regions (Northern America, Southern America, Central America and Caribbean), recorded in the SITVIT2 database.

SIT	Spoligotype Description	Octal value	This study n/t (%)	Distribution In Neighboring Regions			
				Peru (Without Study) n/t (%)	AMER-N n/t (%)	AMER-C & CARI n/t (%)	AMER-S (without study) n/t (%)
1		000000000003771	13/142 (9,15)	50/886 (5.64)	1994/14557 (13.70)	65/3197 (2.03)	91/9716 (0.94)
42		777777607760771	4/142 (2,82)	69/886 (7.79)	397/14557 (2.73)	206/3197 (6.44)	973/9716 (10.01)
47		777777774020771	16/142 (11,27)	32/886 (3.61)	243/14557 (1.67)	58/3197 (1.81)	177/9716 (1.82)
49		777777777720731	4/142 (2,82)	9/886 (1.02)	26/14557 (0.18)	9/3197 (0.28)	19/9716 (0.20)
50		777777777720771	21/142 (14,79)	139/886 (15.69)	595/14557 (4.09)	160/3197 (5.00)	579/9716 (5.96)
52		777777777760731	3/142 (2,11)	1/886 (0.11)	111/14557 (0.76)	41/3197 (1.28)	18/9716 (0.19)
53		777777777760771	7/142 (4,93)	111/886 (12.53)	847/14557 (5.82)	359/3197 (11.23)	749/9716 (7.71)
91		700036777760771	1/142 (0,7)	24/886 (2.71)	126/14557 (0.87)	37/3197 (1.16)	55/9716 (0.57)
93		777737607760771	3/142 (2,11)	14/886 (1.58)	75/14557 (0.52)	43/3197 (1.35)	170/9716 (1.75)
189		777741777760771	1/142 (0,7)	1/886 (0.11)	3/14557 (0.02)	2/3197 (0.06)	1/9716 (0.01)
219		777740777760771	21/142 (14,79)	7/886 (0.79)	31/14557 (0.21)	0/3197 (0.00)	7/9716 (0.07)
222		777774077560771	1/142 (0,7)	26/886 (2.93)	17/14557 (0.12)	1/3197 (0.03)	27/9716 (0.28)
291		777777677760771	1/142 (0,7)	1/886 (0.11)	1/14557 (0.01)	5/3197 (0.16)	23/9716 (0.24)
469		077777607760771	5/142 (3,52)	4/886 (0.45)	13/14557 (0.09)	2/3197 (0.06)	6/9716 (0.06)
1122		767777777760771	1/142 (0,7)	1/886 (0.11)	4/14557 (0.03)	0/3197 (0.00)	2/9716 (0.02)
1160		777737207760771	1/142 (0,7)	0/886 (0.00)	2/14557 (0.01)	0/3197 (0.00)	1/9716 (0.01)
1355		777777407560731	7/142 (4,93)	26/886 (2.93)	7/14557 (0.05)	0/3197 (0.00)	27/9716 (0.28)
1905		777777777460771	2/142 (1,41)	0/886 (0.00)	0/14557 (0.00)	0/3197 (0.00)	18/9716 (0.19)
2375		767777774020771	2/142 (1,41)	0/886 (0.00)	0/14557 (0.00)	0/3197 (0.00)	0/9716 (0.00)
2502		777777407560771	2/142 (1,41)	0/886 (0.00)	0/14557 (0.00)	0/3197 (0.00)	3/9716 (0.03)
2940		777777607760471	1/142 (0,7)	0/886 (0.00)	0/14557 (0.00)	0/3197 (0.00)	5/9716 (0.05)
3001		777774077020771	11/142 (7,75)	8/886 (0.90)	1/14557 (0.01)	0/3197 (0.00)	8/9716 (0.08)
3777		777775077560771	1/142 (0,7)	0/886 (0.00)	0/14557 (0.00)	0/3197 (0.00)	0/9716 (0.00)
3778		777775077020771	5/142 (3,52)	0/886 (0.00)	0/14557 (0.00)	0/3197 (0.00)	0/9716 (0.00)
3779		767774077020771	2/142 (1,41)	0/886 (0.00)	0/14557 (0.00)	0/3197 (0.00)	0/9716 (0.00)
3780		700036777740771	1/142 (0,7)	1/886 (0.11)	0/14557 (0.00)	0/3197 (0.00)	1/9716 (0.01)
	TOTAL		**137/142 (96.48)**	**524/886 (56.14)**	**4493/14557 (30.86)**	**988/3197 (30.90)**	**2960/9716 (30.47)**