

Whole genome sequencing–based analysis of tuberculosis (TB) in migrants: rapid tools for cross-border surveillance and to distinguish between recent transmission in the host country and new importations

Estefanía Abascal^{1,2,3}, Laura Pérez-Lago^{1,2,3}, Miguel Martínez-Lirola⁴, Álvaro Chiner-Oms⁵, Marta Herranz^{1,2,6}, Imane Chaoui⁷, Iñaki Comas^{8,9}, My Driss El Messaoudi¹⁰, José Antonio Garrido Cárdenas¹¹, Sheila Santantón^{1,2}, Emilio Bouza^{1,2,6,12}, Darío García-Viedma^{1,2,6}

1. Servicio Microbiología Clínica y Enfermedades Infecciosas, Hospital General Universitario Gregorio Marañón, Madrid, Spain
2. Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain
3. These authors have contributed equally
4. Complejo Hospitalario Torrecárdenas, Almería, Spain
5. Unidad Mixta Genómica y Salud, Centro Superior de Investigación en Salud Pública (FISABIO)-Universitat de València, Valencia, Spain
6. CIBER Enfermedades respiratorias (CIBERES), Spain
7. Unité de Biologie et Recherches Médicales, Division des Sciences du Vivant, Centre National de l'Energie, des Sciences et des Techniques Nucléaires (CNESTEN), Rabat, Morocco
8. Instituto de Biomedicina de Valencia (IBV) Consejo Superior de Investigaciones Científicas (CSIC), Valencia, Spain
9. CIBER Epidemiología y Salud Pública (CIBERESP), Spain
10. Institut Pasteur of Morocco, Casablanca, Morocco
11. Servicio de Secuenciación, Universidad de Almería, Almería, Spain
12. Departamento de Medicina, Facultad de Medicina, Universidad Complutense de Madrid, Madrid, Spain

Correspondence: Darío García de Viedma (dgviedma2@gmail.com)

Citation style for this article:

Abascal Estefanía, Pérez-Lago Laura, Martínez-Lirola Miguel, Chiner-Oms Álvaro, Herranz Marta, Chaoui Imane, Comas Iñaki, El Messaoudi My Driss, Cárdenas José Antonio Garrido, Santantón Sheila, Bouza Emilio, García-de-Viedma Darío. Whole genome sequencing–based analysis of tuberculosis (TB) in migrants: rapid tools for cross-border surveillance and to distinguish between recent transmission in the host country and new importations. *Euro Surveill.* 2019;24(4):pii=1800005. <https://doi.org/10.2807/1560-7917.ES.2019.24.4.1800005>

Article submitted on 26 Dec 2017 / accepted on 01 Jun 2018 / published on 24 Jan 2019

Background: The analysis of transmission of tuberculosis (TB) is challenging in areas with a large migrant population. Standard genotyping may fail to differentiate transmission within the host country from new importations, which is key from an epidemiological perspective. **Aim:** To propose a new strategy to simplify and optimise cross-border surveillance of tuberculosis and to distinguish between recent transmission in the host country and new importations. **Methods:** We selected 10 clusters, defined by 24-locus mycobacterial interspersed repetitive unit-variable number tandem repeat (MIRU-VNTR), from a population in Spain rich in migrants from eastern Europe, north Africa and west Africa and reanalysed 66 isolates by whole-genome sequencing (WGS). A multiplex-allele-specific PCR was designed to target strain-specific marker single nucleotide polymorphisms (SNPs), identified from WGS data, to optimise the surveillance of the most complex cluster. **Results:** In five of 10 clusters not all isolates showed the short genetic distances expected for recent transmission and revealed a higher number of SNPs, thus suggesting independent importations of prevalent strains in the country of origin. In the most complex cluster, rich in Moroccan cases, a multiplex allele-specific

oligonucleotide-PCR (ASO-PCR) targeting the marker SNPs for the transmission subcluster enabled us to prospectively identify new secondary cases. The ASO-PCR-based strategy was transferred and applied in Morocco, demonstrating that the strain was prevalent in the country. **Conclusion:** We provide a new model for optimising the analysis of cross-border surveillance of TB transmission in the scenario of global migration.

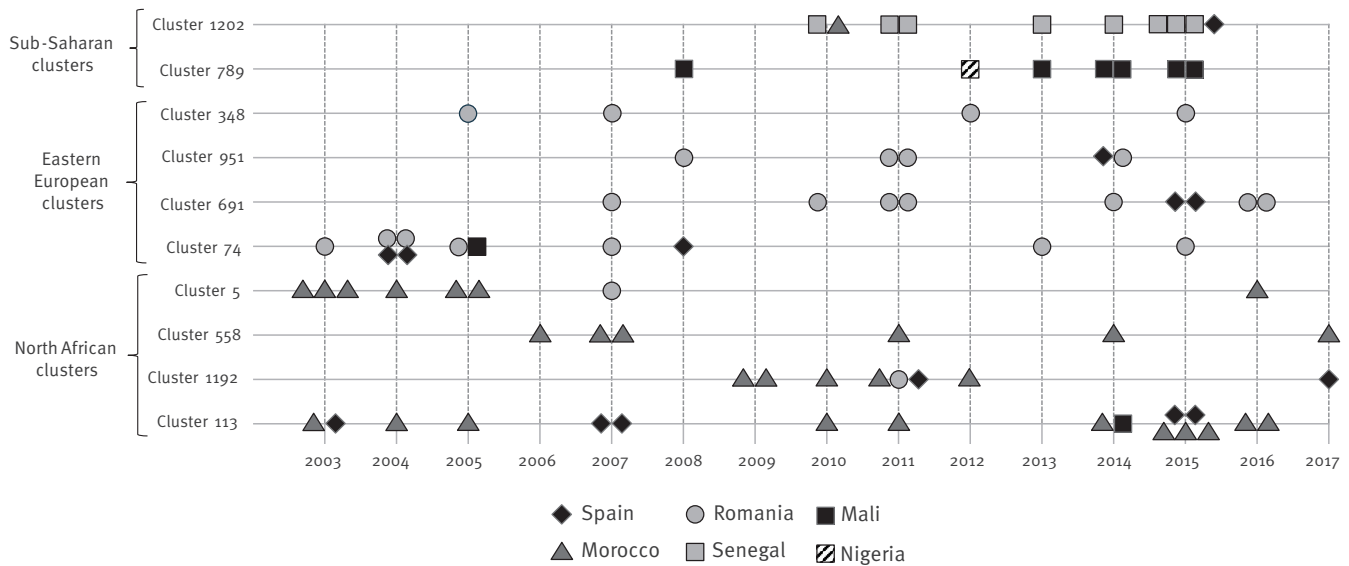
Background

International migration has modified the epidemiology of tuberculosis (TB) in most high-income countries and today, migrants account for up to 40–60% of cases in large cities [1-4]. Some cases are reactivations of infections acquired in the country of origin, with the remainder resulting from recent transmission after arrival in the host country.

Molecular epidemiology provides more accurate data on the transmission dynamics of TB in settings with a complex composition of cases due to migration [5-7]. Several studies have shown variable composition in the nationalities comprising transmission clusters. This variety ranges from settings with marked transmission permeability leading to multinational clusters,

FIGURE 1

Chart summarising the general data of the clusters analysed, rich in cases from sub-Saharan Africa, eastern Europe and north Africa, 2003–2017 (n = 10 clusters)



Clusters are grouped according to their geographic origin (sub-Saharan, eastern European and north African). Each horizontal line corresponds to a cluster and each symbol corresponds to a patient. The patients involved in each cluster are distributed along the timeline (years at the bottom of the chart) with different symbols according to the nationalities shown in the legend.

to other socio-epidemiological contexts where a more homogeneous composition of nationalities is found, with clusters only involving single nationalities [6,8]. Autochthonous clusters and those comprising several nationalities more likely reflect recent transmission events. However, clusters rich in cases from one country of origin are especially difficult to interpret. This is because they can be the result of either of two circumstances: (i) a strain is imported from the country of origin and subsequently transmitted to migrants of the same nationality in the host country; or (ii) genetically closely related strains, which are prevalent in the country of origin, are independently imported by individuals who were exposed in the country of origin but are not epidemiologically related in the host country. Thus, differentiation between these alternatives, i.e. recent transmission in the host country vs importation, is challenging, yet highly relevant in epidemiological terms.

Application of whole-genome sequencing (WGS) for analysis of transmission of TB has given birth to the field of genomic epidemiology, which has markedly increased specificity in the definition of transmission clusters [9–12]. Determination of the number of single nucleotide polymorphisms (SNPs) [12] between the sequences of different isolates allows to split clusters that had been previously defined by standard molecular tools into smaller subclusters that are much more consistent with the geographic distribution of the cases and with the epidemiological links between them [11].

Our aim was to apply WGS in a more in-depth analysis of migrant TB cases involved in clusters in Spain that had been defined by standard genotyping. We attempted to determine whether the clusters corresponded to recent transmission in the host country (because *Mycobacterium tuberculosis* (MTB) isolates show no or a very short genetic distance) or to undetected independent importations of strains that are prevalent in the country of origin and have acquired higher SNP-based diversity as a result of prolonged periods of circulation. In addition, we took advantage of the SNPs identified for either the recently transmitted or imported isolates, to tailor simple PCR tools to simplify and optimise the precise assignation of recent transmission or importation in the new clusters arising. Further, we used these same tools in a new extended and cross-border analysis, for an in-depth surveillance of the MTB strains analysed in unrelated Spanish populations, as well as in the country of origin.

Methods

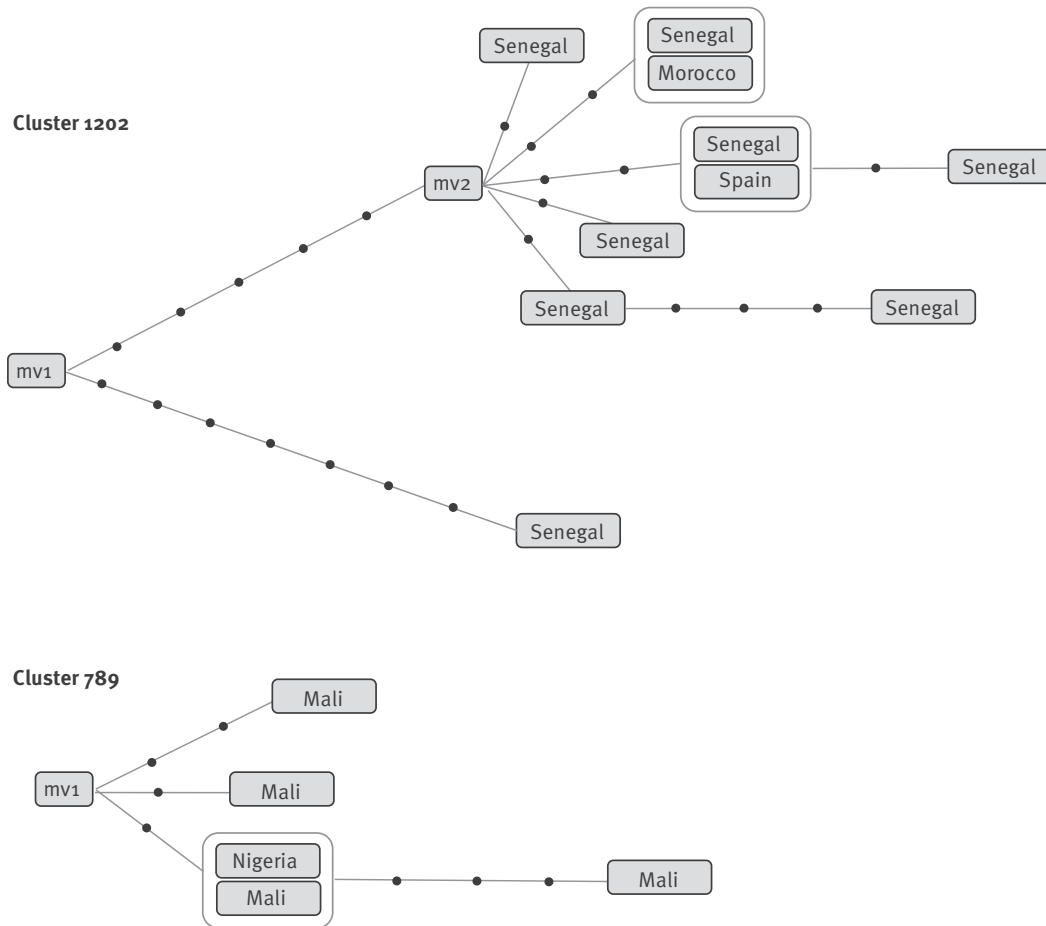
Clusters and strains selected

We retrospectively selected all clusters from the ongoing molecular epidemiology universal genotyping programme in Almería, south-east Spain [7,13] fulfilling the following selection criteria: The clusters analysed were 24 locus mycobacterial interspersed repetitive unit-variable number tandem repeat (MIRU-VNTR)-defined clusters [14] including four or more cases, covering at least 5 years and rich (>60% of the clustered cases) in migrants from a single country from one of

FIGURE 2

Networks of relationships obtained from the whole genome sequencing analysis for clusters rich in cases from sub-Saharan Africa

Sub-Saharan Africa



mv: median vectors; SNP: single nucleotide polymorphisms.

Each black dot corresponds to a SNP. Each box corresponds to a patient and the nationality is indicated within the box. When two or more cases share identical sequences (zero SNPs between them) they are surrounded by a line.

three geographic areas (eastern Europe, north Africa and sub-Saharan Africa). The lineage of the strains involved in the selected clusters was assigned based on the determination of lineage-specific SNP markers [15] by multiplex allele-specific oligonucleotide-PCR (ASO-PCR) [16].

Convenience samples from Valencia (all isolates with available WGS data in IBV, for the period 2004-2017) and Madrid (all isolates with genotypic data available in Hospital Gregorio Marañón, Spain) for the period 2004-10, were also included in the study. A retrospective convenience sample of part of the isolates from northern Morocco (Tangier, Tetouan and Larache) obtained during the same period also were included; no previous genotypic information was available for

these isolates. Finally, a pool of 20 randomly selected TB migrant cases from Morocco (among all those diagnosed in Almería) that were infected with strains other than those analysed in this study were selected as controls.

Genomic analysis

DNA purification

DNA for WGS of the MIRU-VNTR-defined clusters from Almería was purified from subcultures on Mycobacteria Growth Indicator Tube (MGIT) (using Qiagen kit; QIAamp DNA Mini Kit, Qiagen, Courtaboeuf, France) or Lowenstein Jensen medium (CTAB (cetyl trimethylammonium bromide)-based standard purification).

WGS of the strains from the collection in Morocco was performed by purifying (Qiagen kit) the DNA from the remnants of bacterial lysates that had been stored.

WGS of the strains from the collection in Madrid was performed by purifying DNA (Qiagen kit) from freshly inactivated suspensions from the stored frozen isolates.

Whole genome sequencing and single nucleotide polymorphism analysis

WGS was performed as detailed elsewhere [17]. Briefly, DNA libraries were generated following the Nextera XT Illumina protocol (Nextera XT Library Prep kit (FC-131–1024), Illumina, San Diego, United States (US)). Library quality and size distribution were checked on a 2200 TapeStation Bioanalyzer (Agilent Technologies, Santa Clara, US). Libraries were run in a Miseq device (Illumina), which generated 35–151-bp paired-end reads and an average per base coverage of 70 x. Sequences were deposited in www.ebi.ac.uk (PRJEB23664 and PRJEB25814).

We mapped the reads for each strain using the Burrows-Wheeler Aligner and the ancestral MTB genome, which was identical to H37Rv in terms of structure, but which included the maximum likelihood–inferred ancestral nt positions from a virtual ancestor [18]. SNP calls were made with SAMtools and VarScan (coverage of at least 20 x, mean SNP mapping quality of 20). From all the variants detected, we kept only the homozygous calls (those present in at least 90% of the reads in a specific position). Moreover, to filter out potential false positive SNPs due to mapping errors we omitted the variants detected in repetitive regions, phages and PE/PPE regions. Also, SNPs close to indels and those present in areas with an anomalous accumulation of variants (three or more SNPs in 10 bp) were omitted. Alignments and SNP variants (called with a \times coverage in at least one of the isolates in a cluster) were visualised and checked for the remaining isolates in the Integrative Genomics Viewer IGV (version 2.3.59) programme. Multiple comparisons between the SNPs from different isolates were made using an in-house script written in R [19]. We used the reference values (in the number of SNPs) of Walker et al. [12] to determine whether the isolates in a MIRU-VNTR cluster were related. In three isolates we detected an unexpectedly high number of SNPs (> 200) with respect to the other members in the cluster; they were considered to be clustered as the result of homoplasy in the MIRU-VNTR pattern and therefore were eliminated from the study.

The median-joining networks were constructed from the SNP matrix generated for each case using the programme NETWORK 5.0.0.1. Median vectors (mv) were defined when the distribution of SNPs of the isolates analysed indicated the existence of a node that was not represented by the sampled isolates sequenced for each cluster. These median vectors therefore corresponded to non-sampled isolates in the cluster. The

chronology of acquisition of SNPs is represented from left to right in the networks.

Cluster-specific single nucleotide polymorphisms and design of ASO-PCRs

To identify SNPs which were specific for cluster 113, we created a database of variants using sequences from isolates which were representative of the global MTB complex (MTBC) diversity. We downloaded all the accessible raw data from different publications [20–22]. All the fastq files published in these studies were downloaded and aligned against the ancestral MTB genome using the BWA tool. We kept the alignments that had a mean coverage higher than 20. Using this criterion, we kept 7,977 samples representative from the seven lineages. We extracted all the variants present in these samples as described above. The 7,977 samples were filtered to remove transmission clusters so we kept one representative strain of each transmission cluster detected. Once the transmission clusters were filtered, we kept 4,762 sequences. The 207,188 variants present in these samples were used to construct a reference database to evaluate the specificity of the SNPs selected for the ASO-PCRs to be applied in cluster 113.

Two different ASO-PCRs were designed to analyse strain 113. The first ASO-PCR aimed to differentiate new secondary transmitted cases in Almería from independently imported cases. We designed a four-plex single-tube format. Two of the four SNPs targeted were strain 113-marker-SNPs (one targeted the 113 allele and the other the non-113 allele). The remaining two SNPs targeted were only shared by the 113-strain isolates involved in the recent transmission cluster (Supplementary Table S1). The design pursued to obtain three different amplification patterns depending on whether a new case corresponded to recent transmission by strain 113, importation of strain 113 or infection with a strain other than 113.

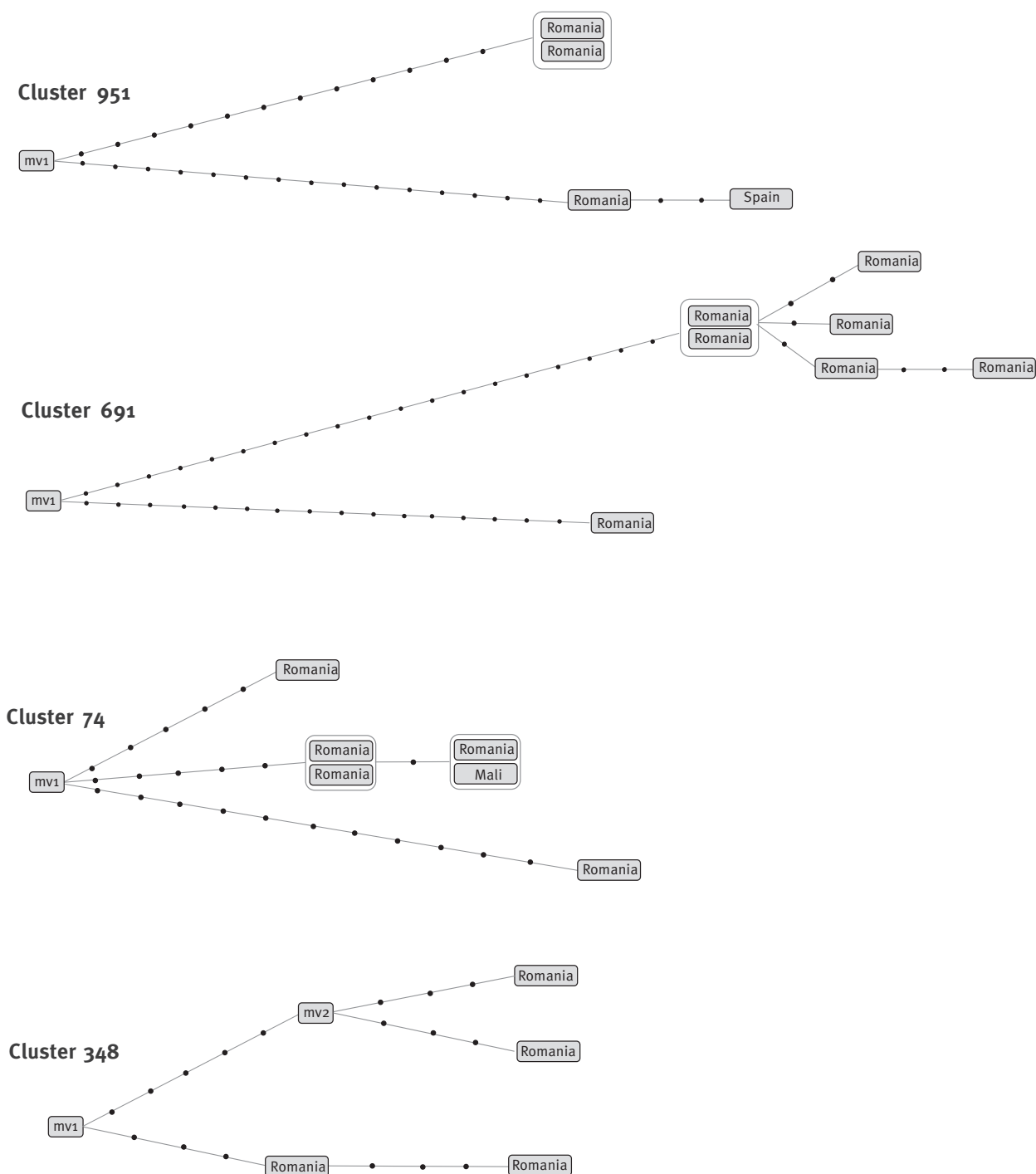
The reaction conditions were as follows: 1.5 mM MgCl₂, 0.2 μ M of each primer (Supplementary Table S1), 200 μ M deoxynucleotides (dNTPs) (Roche, Mannheim, Germany), 1% Dimethyl sulfoxide (DMSO) and 1.5 μ L Taq DNA Polymerase (Roche, Mannheim, Germany). The PCR conditions were 95 °C for 5 min followed by 25–40 cycles (95 °C for 1 min, 61 °C for 1 min, and 72 °C for 1 min) and 72 °C for 10 min. The number of cycles was 25 when using as a template DNA purified from primary positive cultures and 40 when it was purified from sputa.

The second ASO-PCR was applied to assess whether an MTB isolate corresponded to strain 113 or to any other strain. We prepared another version of a four-plex single-tube ASO-PCR to target four SNPs (two alleles specific for isolates 113 and the other two alleles expected for non-113 strains) (Supplementary Table S2). Two different amplification patterns indicated whether a strain corresponded to the 113 strain or to any strain

FIGURE 3

Networks of relationships obtained from the whole genome sequencing analysis for clusters rich in cases from eastern Europe

Eastern Europe



mv: median vectors; SNP: single nucleotide polymorphisms.

Each black dot corresponds to a SNP. Each box corresponds to a patient and the nationality is indicated within the box. When two or more cases share identical sequences (zero SNPs between them) they are surrounded by a line.

other than 113. The reaction conditions were as follows: 1.5 mM MgCl₂, 0.2 µM of each primer (Supplementary Table S2), 200 µM dNTPs (Roche, Mannheim, Germany) and 1.5 µL Taq DNA Polymerase (Roche, Mannheim, Germany). The PCR conditions were 95 °C for 5 min followed by 30 cycles (95 °C for 1 min, 64 °C for 1 min, and 72 °C for 1 min) and 72 °C for 10 min. The ASO-PCR was applied on purified DNA purified or directly on bacterial lysates obtained from boiling stored frozen isolates.

The amplification patterns were analysed by sizing the amplification products using agarose gel electrophoresis.

Results

We selected 10 MIRU-VNTR-defined clusters (Figure 1) from the universal molecular epidemiology survey that has been running in Almería since 2003. The clusters were rich in cases from countries representative of three wide geographic areas, namely, sub-Saharan Africa (two clusters, in which most cases were from Senegal and Mali), north Africa (four clusters in which most cases were from Morocco) and eastern Europe (four clusters in which most cases were from Romania). All the involved strains were pansusceptible and corresponded to lineage four.

Sub-Saharan clusters

In cluster 1202, the analysis of SNPs from the 10 cases indicated the coexistence of a group of nine cases with a genetic distance of 0–7 SNPs between cases (Figure 2). The group included seven cases from Senegal, one from Morocco and one from Spain. Both observations strongly suggested that these nine cases were in fact part of a recent transmission event in Spain. Despite sharing an identical MIRU-VNTR pattern, the remaining case from Senegal showed a higher genetic distance i.e. 12 SNPs, with seven specific for this isolate and not sharing the five SNPs shared by all the isolates in the recent transmission group (Figure 2). These observations made it more likely, that this case corresponded to an unrelated importation from Senegal.

In cluster 789 (Figure 2), we sequenced five of the cases (four from Mali and the only case from Nigeria). The genetic distances between cases were 0–6 SNPs. No cases showed a distribution of SNPs that differed markedly within the group, suggesting the absence of independent importations from the country of origin.

Eastern European clusters

In three of the four clusters that were rich in cases from Romania (Figure 3), we detected the coincidence of cases due to either recent transmission or to independent importations.

In cluster 951, of the five cases, clustered by MIRU-VNTR, (Figure 3) WGS analysis of the four available isolates suggested that the theoretical cluster was hiding two independent subclusters. Two Romanian cases from the year 2011 differed in 27 SNPs and therefore

corresponded to independent importations. Each case caused a secondary case in 2014 due to recent transmission in the host country. The isolates from the secondary cases had two SNPs (Spanish case) and zero SNPs (Romanian case) with respect to the corresponding index case.

A similar situation was observed for cluster 691 (Figure 3). WGS revealed that the MIRU-VNTR-defined cluster included two cases that brought together a high number of SNPs between them (35 SNPs), likely corresponding to two independent importations. A true recent transmission cluster had developed from one of these cases, with another five secondary cases occurring with genetic distances between cases of 0–5 SNPs. The other imported case corresponded to a dead-end branch i.e. it resulted in no secondary cases.

For cluster 74, we identified two different patterns (Figure 3). First, there were four highly related isolates, with 0–1 SNPs between cases, clearly indicative of recent transmission. Second, there were two branches, possibly corresponding to two independently imported cases with five and eleven specific SNPs, respectively, and did not share the five SNPs found in the four isolates belonging to the transmission subgroup. The transmission event (years 2003–2008) was caused by one of these likely imported cases, whereas the remaining two were representative of dead-end branches (years of isolation: 2013 and 2015).

Finally, in cluster 348 (Figure 3), two cases had a genetic distance of three SNPs, suggesting recent transmission between them. However, a definitive interpretation could not be found for the remaining two cases. The cases showed a genetic distance of six SNPs between them, but a non-sampled node (mv2) was inferred to be located between them in the network. It is, therefore, unclear whether these two cases are part of a recent transmission chain involving a non-sampled case in Spain or if they corresponded to two imported cases that were epidemiologically related with a non-sampled case at the host country.

North African clusters

In three of the four clusters, predominately comprising of cases from Morocco, short genetic distances were recorded between all clustered cases (cluster 558: 0–5 SNPs, cluster 1192: 0–3 SNPs and cluster 5: 0–2 SNPs between cases), highly indicative of recent transmission in the host country, Spain (Figure 4).

However, for the remaining cluster, cluster 113, which included 17 cases, WGS of the 14 available isolates revealed a much more complex network of relationships (Figure 4).

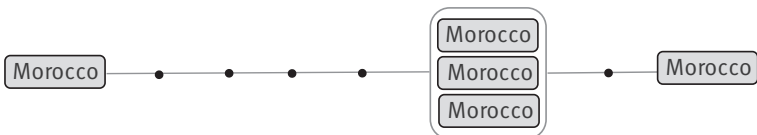
Three median vectors (mv) corresponding to non-sampled cases had to be defined. Seven independent branches were observed (Figure 4), with four, four, seven, eight, nine, 10 and 13 specific SNPs for each of

FIGURE 4

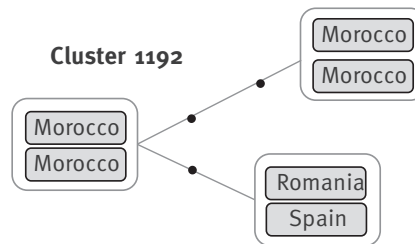
Networks of relationships obtained from the whole genome sequencing analysis for clusters rich in cases from North Africa

North Africa

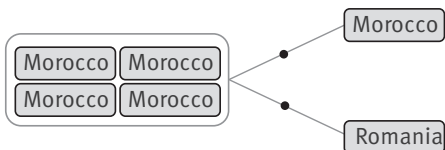
Cluster 558



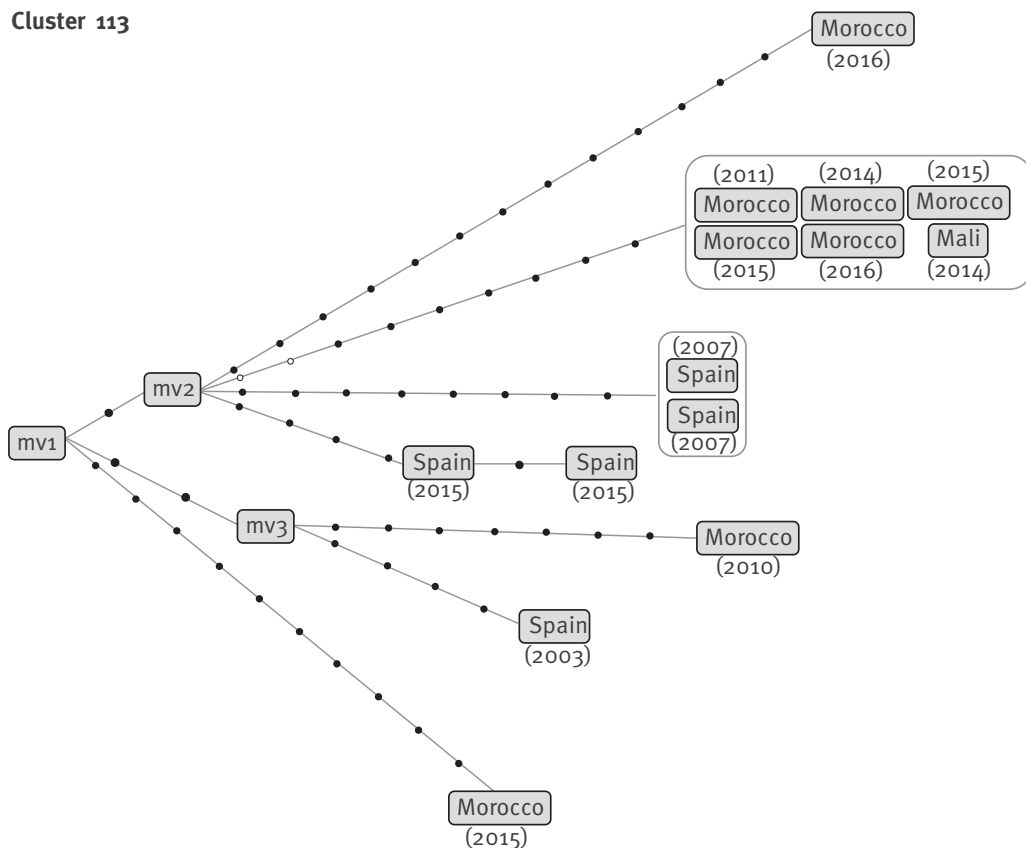
Cluster 1192



Cluster 005



Cluster 113

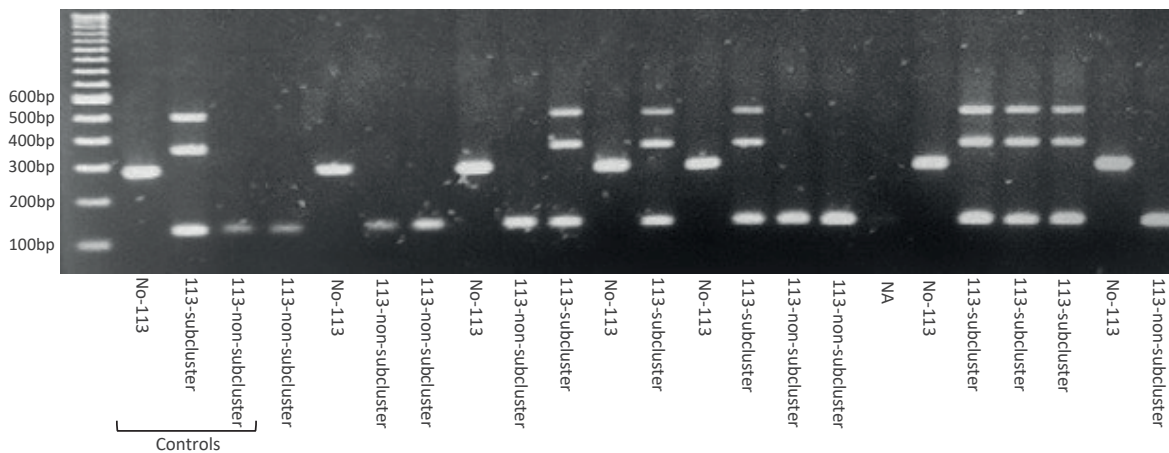


mv: median vectors. SNP: single nucleotide polymorphisms.

Each box corresponds to a patient and the nationality is indicated within the box. When two or more cases share identical sequences (0 SNPs between them) they are surrounded by a line. Each black dot corresponds to a SNP. White dots detailed in cluster 113 correspond to non-fixed alleles, found in heterozygosis in one of the cases, but in homozygosis in the remaining cases. The years of diagnosis are indicated in brackets.

FIGURE 5

Results for the multiplex ASO-PCR designed to precisely assign new incident cases infected by the strain 113 in Almeria and labelling them as due to recent transmission or importation.



Amplification patterns obtained from a selection of isolates representative of the 113 transmission subcluster (113 subcluster), isolates 113 not included in the recent transmission subcluster (113 non-subcluster) and isolates other than 113 (No 113). The different amplification patterns for each group can be observed.

the branches and each more likely corresponding to unrelated cases (distances between each two branches were in the range of 11–24 SNPs). Therefore, these cases were likely due to unrelated importations from Morocco. Of the seven branches, four corresponded to dead-ends, including a single case each (years 2003, 2010, 2015, and 2016); three were from Morocco and were diagnosed 10, 6, and 2 years after arrival. As there were no additional related secondary cases, the findings seem consistent with likely reactivations.

Two of the remaining three branches showed one additional case that was closely related to the imported index case in each branch (zero and one SNPs), which was diagnosed the same year as the index case (year 2007 and 2015, respectively), possibly due to self-limited recent transmission events in Spain.

The remaining branch was the only one with a higher number of cases i.e. six, among which no SNPs were found. Of note, two alleles were in heterozygosis in one of these cases (year 2011) and were fixed as homozygotes in the remaining five cases. Based on this observation, we can infer that the case with heterozygosis was the index case and the remaining five cases were secondary cases and likely due to recent transmissions in Spain.

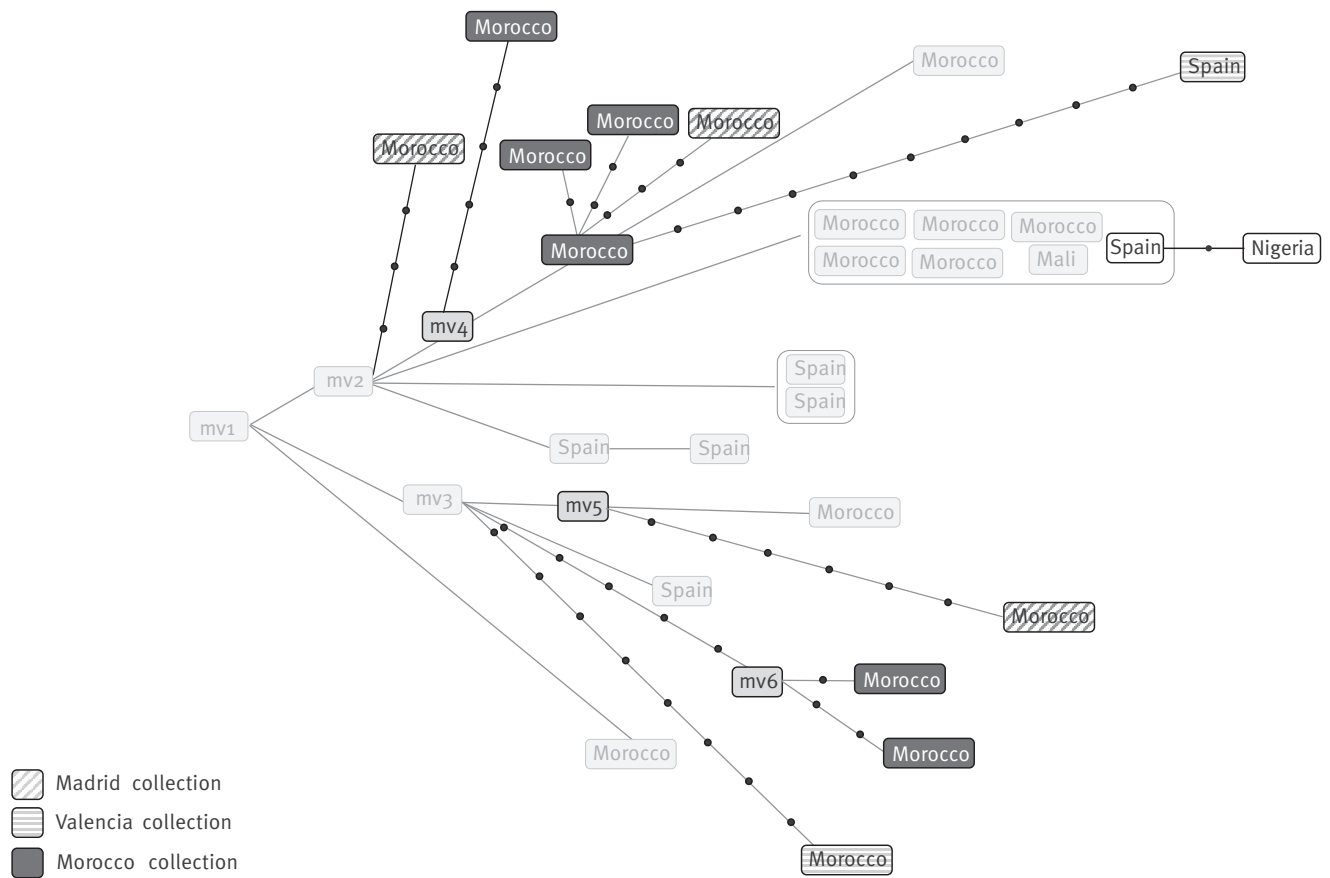
New strategy based on whole genome sequencing data to precisely identify recent transmission

In our context, MIRU-VNTR was proved useless, because it could not discriminate between the three events observed for strain 113 e.g. dead end-imported hosts, self-limited transmission chains and ongoing active transmission events. Among the 17 cases theoretically linked by MIRU-VNTR, only six were really involved in an active recent transmission chain whereas the remaining 11 cases had been misclassified and their epidemiological follow-up was not well oriented. Using standardised interviews with the cases it was possible to establish epidemiological links between the cases in the six-case subcluster, revealing that three cases were customers of the same bar and another case shared a flat with them.

In order to be able to precisely identify the true secondary cases in an active transmission chain, we defined a new approach. We first identified the 71 common SNPs shared by all members in MIRU-VNTR-defined cluster 113 and those SNPs which were specific for the different branches in the network. We designed an allele-specific multiplex PCR (ASO-PCR) including four PCRs, which targeted the following (Supplementary Table S1): (i) two SNPs specific for all the strain 113 isolates in the network, which were selected as a general marker for this strain (one PCR targeting the 113 allele in one

FIGURE 6

Extended network of relationships obtained from the whole genome sequencing analysis for cluster 113 including Almería, Madrid, Valencia and Morocco isolates



mv: median vectors; SNP: single nucleotide polymorphisms.

Each box corresponds to a patient and the nationality is indicated within the box. When two or more cases share identical sequences (0 SNPs between them) they are surrounded by a line. Each black dot corresponds to a SNP. Isolates in Almería detailed in Figure 4 are now fainted. The two new cases identified in Almería by applying the ASO-PCR are shown in white boxes.

of these SNPs and the other PCR the non-113 allele from the other SNP) and (ii) two SNPs among the nine SNPs that were only shared by the branch including the active transmission subcluster (targeting the alleles for the active transmission subcluster).

The ASO-PCR was designed following a four-plex format to target the four SNPs simultaneously in the same reaction tube. This lead to three different amplification patterns depending on whether a new case corresponded to the recently transmitted subcluster 113, to a 113 isolate not involved in this active subcluster (therefore corresponding to a new unrelated importation) or to a strain other than 113 (Figure 5). The specificity of the multiplex ASO-PCR was checked by testing all the 14 isolates with the 113 VNTR pattern and a selection of 20 randomly selected strains for Moroccan migrants among those diagnosed in Almería. The expected

pattern for the three possible profiles was obtained in all cases.

The PCR was transferred to Torrecardenas Hospital in Almería to be prospectively applied on all newly diagnosed TB cases of Moroccan origin or living in the same area as the cases involved in the MIRU-VNTR-defined cluster 113. We first checked that the PCR was sensitive enough to be applied directly on respiratory specimens and were able to obtain an interpretable profile when decontaminated sputa with high or medium bacillary load were used as templates.

An interpretable result was obtained for all the eight stain-positive cases in which the multiplex ASO-PCR was prospectively applied (during a 3-month period) directly on sputa. For the prospective cases with paucibacillary sputa it was necessary to wait until culture was available. In 15 cases, the pattern corresponded

to a non-113 strain; however, in two cases (one from Spain in 2016 and the other from Nigeria in 2017), we obtained the pattern expected for active subcluster 113. Both isolates shared the expected 113 MIRU-VNTR pattern. Subsequent WGS analysis indicated that they showed zero and one SNPs with the six isolates previously included in the active subcluster (Figure 6).

Expanded analysis of strain 113 in unrelated populations

Once the demand for identification of new cases due to recent transmission of the active transmission node was resolved, we focused on the other issue affecting this cluster i.e. the independent importations of closely related (genetically) strains from the country of origin, those likely prevalent in Morocco that have acquired diversity by circulating over extended periods of time. We tried to identify other examples of independent importations for this strain in other unrelated populations.

For this purpose, we selected two Spanish populations: one from Valencia (eastern Spain), a representative of a node with WGS data available from a population-based genomic epidemiology programme and another one from Madrid (central Spain), for which no population-based WGS data were available.

The approach in Valencia was direct and limited to querying on the presence of the 71 SNPs that are specific for the isolates in cluster 113; we identified two cases sharing all the 71 SNPs. When these were integrated into the Almería network, they consistently corresponded to two new subbranches in two of the previously described importation branches (Figure 6).

The approach in Madrid was indirect, involving application of a multiplex ASO-PCR directly on stored isolates from Moroccan migrant TB cases. We prepared a new version of a four-plex ASO-PCR to target four SNPs. Two of the PCRs targeted the alleles that were specific for isolates 113 and the other two targeted the alleles expected for non-113 strains (Supplementary Table S2); the two amplification patterns identified indicated whether a strain corresponded to the 113 MIRU-VNTR cluster or to any strain other than 113 (Figure 7). We applied it to 134 available Moroccan isolates from our retrospective convenience sample and detected the 113 pattern in five cases (Figure 7a). WGS of three of these isolates confirmed them to be 113 (they included all 71 SNPs) and their integration in the network revealed three new branches (Figure 6).

Expanded analysis of strain 113 in the country of origin

We completed the general analysis of strain 113, with a cross-border analysis, by tracking its circulation in the country of origin. The epidemiological information collected from cases by interview aided in determining that most migrant cases were from cities in the north of Morocco.

Molecular epidemiology studies in northern Morocco were checked in which MIRU-VNTR genotypes corresponding to strain 113 could be found. Chaoui et al. [23] reported a cluster involving four cases in Tangier infected by a LAM3 SIT33 strain that could correspond to strain 113. However, only data for the 12-loci version of MIRU-VNTR were available.

To confirm whether strain 113 was circulating in the area, as suggested by the published data, the same multiplex ASO-PCR that had been designed to track strain 113 in Madrid was transferred and locally applied in Morocco. Interrogation of 11 SIT33 isolates revealed seven with the pattern corresponding to strain 113. In addition, testing of 45 additional retrospective isolates from northern Morocco (Tangier, Tetouan and Larache), for which no previous genotypic information was available, revealed a 113 pattern in seven isolates (Figure 7b). WGS was performed in six of the 14 isolates that were positive for 113 and enabled us to integrate them into the network of relationships (Figure 6). Three of the isolates were positioned in two new sub-branches and the other three were located in one of a previously defined importation branch. Furthermore, two probable recent transmission events in Morocco, involving two and three cases respectively, were identified indirectly (with three SNPs between cases in both of them).

Discussion

Molecular epidemiology based on universal genotyping of TB cases in a population allows us to identify clustered cases that are infected by *M. tuberculosis* isolates with identical fingerprints. From the analysis of clustered cases, we can obtain valuable data on transmission dynamics in different epidemiological scenarios.

The increased complexity resulting from changing socio-epidemiological features due to migration demands special attention. The clusters may be autochthonous, mixed multinational, and foreign-born clusters rich in cases from a specific country.

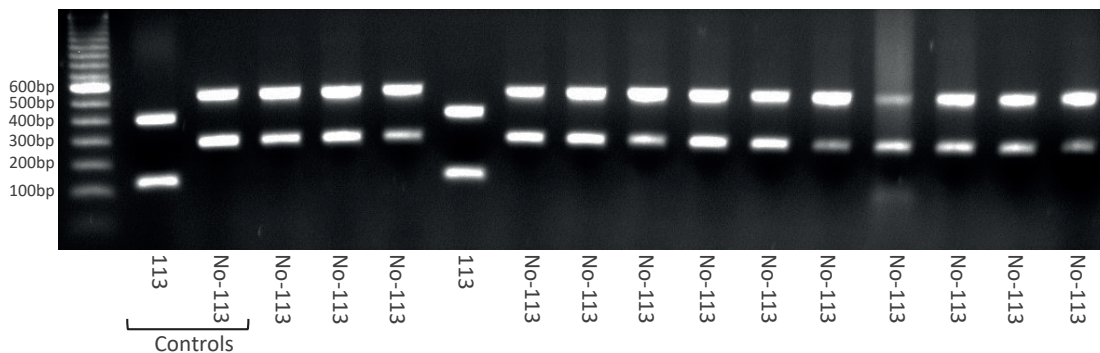
Some of the complex molecular clusters identified in populations with a higher percentage of migrants are not always accompanied by clear epidemiological links between the cases involved [7,24,25]. Here, we tried to analyse whether the lack of epidemiological support could mean that some of the clusters involving migrants were not robust and were misleadingly alerting us to recent transmissions.

We hypothesised that some of the cases in these clusters could correspond to independent importations of strains that might be prevalent in the country of origin. Genetic diversity would be expected to accumulate for a prevalent strain circulating in a high-incidence country over extended periods of time. However, the diversity accumulated is probably insufficient to lead to a change in the MIRU-VNTR pattern, thus explaining why unrelated cases independently importing these strains

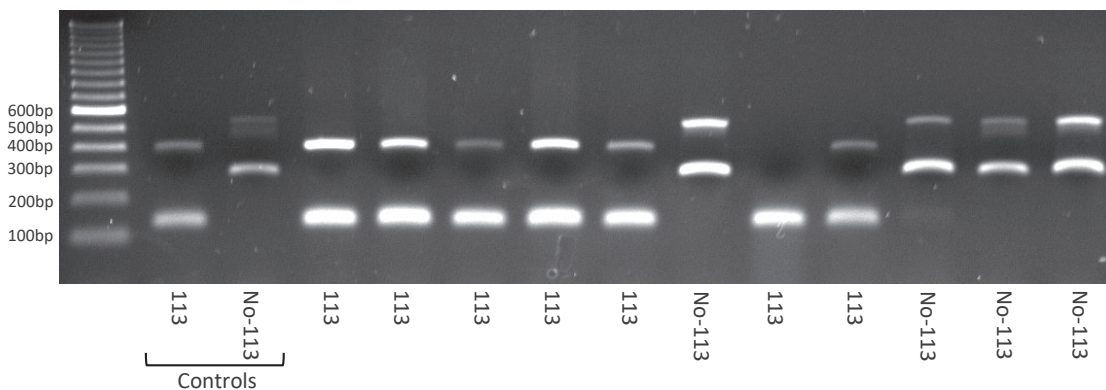
FIGURE 7

Results for the multiplex ASO-PCR designed to (A) retrospectively track strain 113 in Madrid and (B) retrospectively track strain 113 in Morocco

A. Multiplex ASO-PCR results, retrospectively track strain 113, Madrid



B. Multiplex ASO-PCR results, retrospectively track strain 113, Morocco



The different amplification patterns for the strains belonging to the 113 group and those that do not are observed.

appear clustered. MIRU-VNTR types are conserved for highly prevalent strains, as reported in Denmark for a highly prevalent strain responsible for 35% cases over 15 years [26]. However, the application of more discriminative methods e.g. WGS, could help us to reveal some degree of diversity between these prevalent strains and differentiate between true recent transmission in the host country (when no or very limited genetic diversity is found between the corresponding isolates) and independent, unrelated importations of prevalent strains in the country of origin (if we detect greater genetic distance).

Application of this strategy, following the consensus thresholds of diversity to assign or rule out recent transmission with WGS [12], revealed that unrelated importations were hidden within some MIRU-VNTR-defined clusters and had been misinterpreted as recent transmissions in the host country. Due to the size of certain clusters in the analysis we only revealed a minority

(one case in several clusters) that had been misclassified as recent transmission when it was really due to importation. However, in some of the bigger clusters, the magnitude of misclassified cases revealed was higher (eight of 14).

In a 2016 publication, Stucki et al. [27] reported importations within MIRU-VNTR clusters in a nationwide analysis in Switzerland (90 patients in 35 clusters during 2000–08). Only 25% of the MIRU-VNTR-defined clusters including migrants (in this case, mostly from east Africa) were refuted using WGS. The clustering proportion fell from 16.7% to 6.5% for migrant clusters; when only Swiss-born clusters were considered, the decrease was smaller (19.3% to 14.3%). In addition, descriptions of misassignment of recent transmission in MIRU-VNTR-proven migrant clusters revealed by WGS have recently been reported in Canada [28] and the Netherlands [29].

Although our findings are limited to the low number of clusters selected, both these data and ours suggest that the involvement of genetically closely related strains imported independently from high-incidence regions is a widespread phenomenon. We cannot extend the findings from the migrant clusters in our study to all clusters including migrants because in our setting some migrant nationalities were not represented. Nevertheless, our results showed that this phenomenon was not anecdotal or restricted to specific geographic areas and that it was found in clusters with migrants that were representative of different areas e.g. eastern Europe, north Africa and sub-Saharan Africa.

In our study, the identification of imported cases within clusters defined by standard genotyping was mainly supported by the analysis of the total number of SNPs between the clustered cases. However, the analysis of the chronology of diagnosis of the TB cases can also be useful to identify importations. This is because the order of emergence of SNPs is sequential and once acquired they do not reverse [30]. In cluster 558, the last case diagnosed (year 2014) did not present the four SNPs identified in the remaining clustered cases, diagnosed 3–8 years earlier. The most likely explanation is that the 2014 case was imported from a more ancestral branch than the one involved in the recent transmission event in Spain.

The demonstration that both importations and recent transmissions could co-exist in a cluster defined by standard genotyping raised an alert: once one of these genetically closely related strains is imported into a host country, standard molecular epidemiology–surveillance approaches are of very limited value. Based on standard MIRU-VNTR, it would be impossible to discriminate between secondary cases that originated in the host country and unrelated independent importations: all cases would be equally considered clustered.

It is important to differentiate between a new imported case and a recently transmitted secondary case, because each represents a completely different epidemiological situation that has to be managed separately. Consequently, other authors have recommended WGS as the only way to ensure more accurate identification of recent transmission, particularly among migrants from high-incidence areas [27,31]. An alternative to the analysis based on WGS and SNPs calling based on pipelines is the technique of core genome MLST typing, which takes advantage to the discriminatory power of the next generation sequencing (NGS) technique and makes easier the SNP calling by standardised processing and allows a more direct comparative analysis across different laboratories [32]. However, global implementation of WGS is expensive and WGS has been successfully implemented at population level in few settings only [33–35]. With the aim to overcome these limitations and to find a solution that can be implemented in settings where nationwide

WGS application is not a reality, we adapted a strategy previously developed by our group to survey high-risk strains. This strategy is based on tailored ASO-PCRs targeting strain-specific SNPs identified from WGS data of representative isolates for the strains to be surveyed [36]. We implemented it in previous studies to be able to provide a fast response to challenges, such as optimising surveillance of transmission of actively transmitted strains [36], rapid tracking of the presence of specific outbreak strains in a population [37] and confirming the presence of secondary cases due to imported XDR strains from Russia directly on respiratory specimens in the hospital setting [38]. In the current study we adapted the strategy to tailor PCRs targeting the SNPs that were specific for isolates actively involved in recent transmission in the host country and to differentiate these isolates from other independently imported isolates which lacked those SNPs.

To pilot this strategy, we selected the most complex cluster in our study, namely cluster 113, which was rich in cases from Morocco (six different importation branches together with an active transmission cluster). The strategy prospectively identified new secondary cases directly from respiratory specimens. Our proposal not only resolved the epidemiological challenge at the local level, but also enabled us to expand the boundaries of our analysis to other unrelated populations in Spain. If this strain corresponded to a prevalent strain in the country of origin, we would be able to find it in unrelated populations receiving migrants from Morocco. We identified the strain in the two unrelated populations surveyed and proved that importations of the same strain occurred in other settings, thus showing that they were not the result of recent transmissions. For some of the remaining studied strains from migrants from Morocco we also found data indicating they are circulating also in Morocco [23,39] and similar efforts could be done to fully characterise their global distribution.

Conclusion

Tracking transmission of TB through cross-border surveillance is a crucial element in the current epidemiological surveillance of TB, and data from both the country of origin and host countries must be integrated as recently exemplified in a study revealing a cross-border MDR-TB cluster involving several European countries [40]. Our findings revealed standard MIRU-VNTR-based epidemiology was not a suitable approach for cross-border surveillance as it was unable to discriminate between importations and recent transmissions. WGS-based analysis was able to differentiate these two overlapping events, however, genomic analysis is not accessible for many countries involved in cross-border TB transmission. Here, we propose a new strategy, adapted to settings with no or limited access to WGS, based on designing simple PCR tools tailored to be adapted to identify either recent transmission in the host country or independent importations from the

country of origin. Adapted versions of the same PCRs were also designed to be transferred and applied to track the strain circulating in the country of origin.

Our next step will be to extend the approaches used in this study to develop a network of nodes surveying prevalent strains from countries with a high TB incidence that are being exported to countries with low-TB burden. Such a network could contribute to the establishment of a new global cross-border surveillance system, fitted to the challenges associated with international migration.

Acknowledgements

We thank Thomas O'Boyle for proofreading the manuscript. We also thank Remedios Guna (Hospital General de Valencia), Lina Gimeno (Hospital General de Alicante) and Isabel Escribano (Hospital Virgen de los Lirios, Alcoy) for providing the clinical isolates from the Comunidad Valenciana analyzed in this study. This project was funded by ISCIII: ERANET-LAC (TRANS-TB-TRANS REF AC16/00057; ELAC2015/To8-0664), FIS (13/01207; 15/01554) and cofunded by ERDF Funds from the European Commission: "A way of making Europe". Miguel Servet grant (CP15/00075) for LPL. Ministerio de Economía y Competitividad (grant SAF2016-77346-R), ERC (638553-TB-ACCELERATE) to IC. FPU13/00913 (Ministerio de Educación y Ciencia) to ACO.

Conflict of interest

None declared.

Authors' contributions

Design of the study, analysis of results, preparation of the MS: DGV, LPL, EA Experimental tasks: EA, LPL, MML, MH, IC, ICH, JAGC, SS. Analysis of data: DGV, EA, LPL, MML, ACO, MH, IC, MDEM. General revision: EB.

References

1. de Vries G, Aldridge RW, Cayla JA, Haas WH, Sandgren A, van Hest NA, et al. Epidemiology of tuberculosis in big cities of the European Union and European Economic Area countries. *Euro Surveill*. 2014;19(9):20726. <https://doi.org/10.2807/1560-7917.ES2014.19.9.20726> PMID: 24626208
2. Diel R, Rüsche-Gerdes S, Niemann S. Molecular epidemiology of tuberculosis among immigrants in Hamburg, Germany. *J Clin Microbiol*. 2004;42(7):2952-60. <https://doi.org/10.1128/JCM.42.7.2952-2960.2004> PMID: 15243044
3. Iñigo J, García de Viedma D, Arce A, Palenque E, Herranz M, Rodríguez E, et al. Differential findings regarding molecular epidemiology of tuberculosis between two consecutive periods in the context of steady increase of immigration. *Clin Microbiol Infect*. 2013;19(3):292-7. <https://doi.org/10.1111/j.1469-0691.2012.03794.x> PMID: 22404140
4. Ospina JE, Orcau À, Millet JP, Ros M, Gil S, Caylà JAB. Tuberculosis Immigration Working Group. Epidemiology of Tuberculosis in Immigrants in a Large City with Large-Scale Immigration (1991-2013). *PLoS One*. 2016;11(10):e0164736. <https://doi.org/10.1371/journal.pone.0164736> PMID: 27749904
5. De Beer JL, Kodmon C, van der Werf MJ, van Ingen J, van Soolingen D, Collective the ECDC MDR-TB molecular surveillance project participants. Molecular surveillance of multi- and extensively drug-resistant tuberculosis transmission in the European Union from 2003 to 2011. *Euro Surveill*. 2014;19(11):20742. <https://doi.org/10.2807/1560-7917.ES2014.19.11.20742> PMID: 24679719
6. Lillebaek T, Andersen AB, Bauer J, Dirksen A, Glismann S, de Haas P, et al. Risk of Mycobacterium tuberculosis transmission in a low-incidence country due to immigration from high-incidence areas. *J Clin Microbiol*. 2001;39(3):855-61. <https://doi.org/10.1128/JCM.39.3.855-861.2001> PMID: 11230395
7. Martínez-Lirola M, Alonso-Rodríguez N, Sánchez ML, Herranz M, Andrés S, Peñafiel T, et al. Advanced survey of tuberculosis transmission in a complex socioepidemiologic scenario with a high proportion of cases in immigrants. *Clin Infect Dis*. 2008;47(1):8-14. <https://doi.org/10.1086/588785> PMID: 18484876
8. Alonso Rodríguez N, Chaves F, Iñigo J, Bouza E, García de Viedma D, Andrés S, et al. TB Molecular Epidemiology Study Group of Madrid. Transmission permeability of tuberculosis involving immigrants, revealed by a multicentre analysis of clusters. *Clin Microbiol Infect*. 2009;15(5):435-42. <https://doi.org/10.1111/j.1469-0691.2008.02670.x> PMID: 19416291
9. Bryant JM, Schürch AC, van Deutekom H, Harris SR, de Beer JL, de Jager V, et al. Inferring patient to patient transmission of Mycobacterium tuberculosis from whole genome sequencing data. *BMC Infect Dis*. 2013;13(1):110. <https://doi.org/10.1186/1471-2334-13-110> PMID: 23446317
10. Gardy JL, Johnston JC, Ho Sui SJ, Cook VJ, Shah L, Brodtkin E, et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med*. 2011;364(8):730-9. <https://doi.org/10.1056/NEJMoa1003176> PMID: 21345102
11. Roetzer A, Diel R, Kohl TA, Rückert C, Nübel U, Blom J, et al. Whole genome sequencing versus traditional genotyping for investigation of a Mycobacterium tuberculosis outbreak: a longitudinal molecular epidemiological study. *PLoS Med*. 2013;10(2):e1001387. <https://doi.org/10.1371/journal.pmed.1001387> PMID: 23424287
12. Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, et al. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. *Lancet Infect Dis*. 2013;13(2):137-46. [https://doi.org/10.1016/S1473-3099\(12\)70277-3](https://doi.org/10.1016/S1473-3099(12)70277-3) PMID: 23158499
13. Alonso-Rodríguez N, Martínez-Lirola M, Sánchez ML, Herranz M, Peñafiel T, Bonillo MC, et al. Prospective universal application of mycobacterial interspersed repetitive-unit-variable-number tandem-repeat genotyping to characterize Mycobacterium tuberculosis isolates for fast identification of clustered and orphan cases. *J Clin Microbiol*. 2009;47(7):2026-32. <https://doi.org/10.1128/JCM.02308-08> PMID: 19458183
14. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. *J Clin Microbiol*. 2006;44(12):4498-510. <https://doi.org/10.1128/JCM.01392-06> PMID: 17005759
15. Stucki D, Malla B, Hostettler S, Huna T, Feldmann J, Yeboah-Manu D, et al. Two new rapid SNP-typing methods for classifying Mycobacterium tuberculosis complex into the main phylogenetic lineages. *PLoS One*. 2012;7(7):e41253. <https://doi.org/10.1371/journal.pone.0041253> PMID: 22911768
16. Carcelén M, Abascal E, Herranz M, Santantón S, Zenteno R, Ruiz Serrano MJ, et al. Optimizing and accelerating the assignment of lineages in Mycobacterium tuberculosis using novel alternative single-tube assays. *PLoS One*. 2017;12(11):e0186956. <https://doi.org/10.1371/journal.pone.0186956> PMID: 29091913
17. Pérez-Lago L, Comas I, Navarro Y, González-Candelas F, Herranz M, Bouza E, et al. Whole genome sequencing analysis of intrapatient microevolution in Mycobacterium tuberculosis: potential impact on the inference of tuberculosis transmission. *J Infect Dis*. 2014;209(1):98-108. <https://doi.org/10.1093/infdis/jit439> PMID: 23945373
18. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nat Genet*. 2013;45(10):1176-82. <https://doi.org/10.1038/ng.2744> PMID: 23995134
19. The R project for Statistical computing. Available from: <https://www.r-project.org>
20. Coll F, McNeerney R, Guerra-Assunção JA, Glynn JR, Perdigão J, Viveiros M, et al. A robust SNP barcode for typing Mycobacterium tuberculosis complex strains. *Nat Commun*. 2014;5(1):4812. <https://doi.org/10.1038/ncomms5812> PMID: 25176035
21. Comas I, Hailu E, Kiro T, Bekele S, Mekonnen W, Gumi B, et al. Population Genomics of Mycobacterium tuberculosis in Ethiopia Contradicts the Virgin Soil Hypothesis for Human Tuberculosis in Sub-Saharan Africa. *Curr Biol*. 2015;25(24):3260-6. <https://doi.org/10.1016/j.cub.2015.10.061> PMID: 26687624

22. Guerra-Assunção JA, Crampin AC, Houben RM, Mzembe T, Mallard K, Coll F, et al. Large-scale whole genome sequencing of *M. tuberculosis* provides insights into transmission in a high prevalence area. *eLife*. 2015;4:e05166. <https://doi.org/10.7554/eLife.05166> PMID: 25732036
23. Chaoui I, Zozio T, Lahlou O, Sabouni R, Abid M, El Aouad R, et al. Contribution of spoligotyping and MIRU-VNTRs to characterize prevalent *Mycobacterium tuberculosis* genotypes infecting tuberculosis patients in Morocco. *Infect Genet Evol*. 2014;21:463-71. <https://doi.org/10.1016/j.meegid.2013.05.023> PMID: 23732366
24. Anderson LF, Tamne S, Brown T, Watson JP, Mullarkey C, Zenner D, et al. Transmission of multidrug-resistant tuberculosis in the UK: a cross-sectional molecular and epidemiological study of clustering and contact tracing. *Lancet Infect Dis*. 2014;14(5):406-15. [https://doi.org/10.1016/S1473-3099\(14\)70022-2](https://doi.org/10.1016/S1473-3099(14)70022-2) PMID: 24602842
25. Pedersen MK, Andersen AB, Andersen PH, Svensson E, Jensen SG, Lillebaek T. Occupational Tuberculosis in Denmark through 21 Years Analysed by Nationwide Genotyping. *PLoS One*. 2016;11(4):e0153668. <https://doi.org/10.1371/journal.pone.0153668> PMID: 27082745
26. Kamper-Jørgensen Z, Andersen AB, Kok-Jensen A, Bygbjerg IC, Andersen PH, Thomsen VO, et al. Clustered tuberculosis in a low-burden country: nationwide genotyping through 15 years. *J Clin Microbiol*. 2012;50(8):2660-7. <https://doi.org/10.1128/JCM.06358-11> PMID: 22675129
27. Stucki D, Ballif M, Egger M, Furrer H, Altpeter E, Battegay M, et al. Standard Genotyping Overestimates Transmission of *Mycobacterium tuberculosis* among Immigrants in a Low-Incidence Country. *J Clin Microbiol*. 2016;54(7):1862-70. <https://doi.org/10.1128/JCM.00126-16> PMID: 27194683
28. Guthrie J, Kong C, Roth D, Rodrigues M, Hoang L, Walker T, et al. Findings from whole genome sequencing of tuberculosis in a geographically large Canadian province with a diverse population. 25-28 Jun 2017. 38th Congress of the European Society of mycobacteriology.; Sibenik, Croatia. OP 29. Available from: http://vorschau.agentur-konsens.de/ESM/downloads/2017/ESM_Abstract_Book2017_final.pdf
29. Jajou R dNH, Mulder A, Kamst M, van Hunen R, de Vries G, Anthony R, et al. VNTR typing is less reliable in predicting epidemiological links than expected, as indicated by whole genome sequencing (WGS) analysis. *J Clin Microbiol*. Forthcoming 2017.
30. Walker TM, Monk P, Smith EG, Peto TE. Contact investigations for outbreaks of *Mycobacterium tuberculosis*: advances through whole genome sequencing. *Clin Microbiol Infect*. 2013;19(9):796-802. <https://doi.org/10.1111/1469-0691.12183> PMID: 23432709
31. Folkvardsen DB, Norman A, Andersen AB, Michael Rasmussen E, Jelsbak L, Lillebaek T. Genomic Epidemiology of a Major *Mycobacterium tuberculosis* Outbreak: Retrospective Cohort Study in a Low-Incidence Setting Using Sparse Time-Series Sampling. *J Infect Dis*. 2017;216(3):366-74. <https://doi.org/10.1093/infdis/jix298> PMID: 28666374
32. Kohl TA, Diel R, Harmsen D, Rothgänger J, Walter KM, Merker M, et al. Whole-genome-based *Mycobacterium tuberculosis* surveillance: a standardized, portable, and expandable approach. *J Clin Microbiol*. 2014;52(7):2479-86. <https://doi.org/10.1128/JCM.00567-14> PMID: 24789177
33. Cirillo DM, Cabibbe AM, De Filippo MR, Trovato A, Simonetti T, Rossolini GM, et al. Use of WGS in *Mycobacterium tuberculosis* routine diagnosis. *Int J Mycobacteriol*. 2016;5(Suppl 1):S252-3. <https://doi.org/10.1016/j.ijmyco.2016.09.053> PMID: 28043590
34. Walker TM, Lalor MK, Broda A, Ortega LS, Morgan M, Parker L, et al. Assessment of *Mycobacterium tuberculosis* transmission in Oxfordshire, UK, 2007-12, with whole pathogen genome sequences: an observational study. *Lancet Respir Med*. 2014;2(4):285-92. [https://doi.org/10.1016/S2213-2600\(14\)70027-X](https://doi.org/10.1016/S2213-2600(14)70027-X) PMID: 24717625
35. Walker TM, Merker M, Kohl TA, Crook DW, Niemann S, Peto TE. Whole genome sequencing for M/XDR tuberculosis surveillance and for resistance testing. *Clin Microbiol Infect*. 2017;23(3):161-6. <https://doi.org/10.1016/j.cmi.2016.10.014> PMID: 27789378
36. Perez-Lago L, Martinez Lirola M, Herranz M, Comas I, Bouza E, Garcia-de-Viedma D. Fast and low-cost decentralized surveillance of transmission of tuberculosis based on strain-specific PCRs tailored from whole genome sequencing data: a pilot study. *Clin Microbiol Infect*. 2015;21(3):249 e1-9.
37. Pérez-Lago L, Herranz M, Comas I, Ruiz-Serrano MJ, López Roa P, Bouza E, et al. Ultrafast Assessment of the Presence of a High-Risk *Mycobacterium tuberculosis* Strain in a Population. *J Clin Microbiol*. 2016;54(3):779-81. <https://doi.org/10.1128/JCM.02851-15> PMID: 26719445
38. Pérez-Lago L, Martínez-Lirola M, García S, Herranz M, Mokrousov I, Comas I, et al. Urgent Implementation in a Hospital Setting of a Strategy To Rule Out Secondary Cases Caused by Imported Extensively Drug-Resistant *Mycobacterium tuberculosis* Strains at Diagnosis. *J Clin Microbiol*. 2016;54(12):2969-74. <https://doi.org/10.1128/JCM.01718-16> PMID: 27682128
39. Lahlou O, Millet J, Chaoui I, Sabouni R, Filali-Maltouf A, Akrim M, et al. The genotypic population structure of *Mycobacterium tuberculosis* complex from Moroccan patients reveals a predominance of Euro-American lineages. *PLoS One*. 2012;7(10):e47113. <https://doi.org/10.1371/journal.pone.0047113> PMID: 23077552
40. Fiebig L, Kohl TA, Popovici O, Mühlenfeld M, Indra A, Homorodean D, et al. A joint cross-border investigation of a cluster of multidrug-resistant tuberculosis in Austria, Romania and Germany in 2014 using classic, genotyping and whole genome sequencing methods: lessons learnt. *Euro Surveill*. 2017;22(2):30439. <https://doi.org/10.2807/1560-7917.ES.2017.22.2.30439> PMID: 28106529

License and copyright

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence, and indicate if changes were made.

This article is copyright of the authors or their affiliated institutions, 2019.