AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology

Standard Genotyping Overestimates Transmission of *Mycobacterium tuberculosis* among Immigrants in a Low-Incidence Country

David Stucki,[a,b,c] Marie Ballif,[a,b,c] Matthias Egger,[c] Hansjakob Furrer,[d] Ekkehardt Altpeter,[e] Manuel Battegay,[f] Sara Droz,[g] Thomas Bruderer,[h] Mireia Coscolla,[a,b] Sonia Borrell,[a,b] Kathrin Zürcher,[c] Jean-Paul Janssens,[i] Alexandra Calmy,[j] Jesica Mazza Stalder,[k] Katia Jaton,[l] Hans L. Rieder,[m] Gaby E. Pfyffer,[n] Hans H. Siegrist,[o] Matthias Hoffmann,[p] Jan Fehr,[q] Marisa Dolina,[r] Reno Frei,[s] Jacques Schrenzel,[t] Erik C. Böttger,[u] Sebastien Gagneux,[a,b] Lukas Fenner,[a,b,c] for the Swiss HIV Cohort and Molecular Epidemiology of Tuberculosis Study Groups

Swiss Tropical and Public Health Institute, Basel, Switzerland[a]; University of Basel, Basel, Switzerland[b]; Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland[c]; Department of Infectious Diseases, Bern University Hospital, University of Bern, Bern, Switzerland[d]; Division of Communicable Diseases, Federal Office of Public Health, Bern, Switzerland[e]; Division of Infectious Diseases and Hospital Epidemiology, University Hospital Basel, Basel, Switzerland[f]; Mycobacteriology Unit, Institute for Infectious Diseases, University of Bern, Bern, Switzerland[g]; Center for Laboratory Medicine, St. Gallen, Switzerland[h]; Division of Pneumology, University Hospital Geneva, Geneva, Switzerland[i]; Division of Infectious Diseases, University Hospital Geneva, Geneva, Switzerland[j]; Division of Pneumology, University Hospital Lausanne, Lausanne, Switzerland[k]; Institute of Microbiology, University Hospital of Lausanne, Lausanne, Switzerland[l]; Epidemiology, Biostatistics, and Prevention Institute, University of Zürich, Zürich, Switzerland[m]; Department of Medical Microbiology, Luzerner Kantonsspital, Lucerne, Switzerland[n]; ADMed Microbiology, La Chaux-de-Fonds, Switzerland[o]; Division of Infectious Diseases, Kantonsspital St. Gallen, St. Gallen, Switzerland[p]; University Hospital Zürich and University of Zürich, Zürich, Switzerland[q]; Servizio di Microbiologia EOLAB, Bellinzona, Switzerland[r]; Division of Clinical Microbiology, University Hospital Basel, Basel, Switzerland[s]; Laboratory of Bacteriology, University Hospitals of Geneva, Geneva, Switzerland[t]; Institute of Medical Microbiology, National Center for Mycobacteria, University of Zürich, Zürich, Switzerland[u]

Immigrants from regions with a high incidence of tuberculosis (TB) are a risk group for TB in low-incidence countries such as Switzerland. In a previous analysis of a nationwide collection of 520 *Mycobacterium tuberculosis* isolates from 2000 to 2008, we identified 35 clusters comprising 90 patients based on standard genotyping (24-locus mycobacterial interspersed repetitive-unit–variable-number tandem-repeat [MIRU-VNTR] typing and spoligotyping). Here, we used whole-genome sequencing (WGS) to revisit these transmission clusters. Genome-based transmission clusters were defined as isolate pairs separated by ≤12 single nucleotide polymorphisms (SNPs). WGS confirmed 17/35 (49%) MIRU-VNTR typing clusters; the other 18 clusters contained pairs separated by >12 SNPs. Most transmission clusters (3/4) of Swiss-born patients were confirmed by WGS, as opposed to 25% (4/16) of the clusters involving only foreign-born patients. The overall clustering proportion was 17% (90 patients; 95% confidence interval [CI], 14 to 21%) by standard genotyping but only 8% (43 patients; 95% CI, 6 to 11%) by WGS. The clustering proportion was 17% (67/401; 95% CI, 13 to 21%) by standard genotyping and 7% (26/401; 95% CI, 4 to 9%) by WGS among foreign-born patients and 19% (23/119; 95% CI, 13 to 28%) and 14% (17/119; 95% CI, 9 to 22%), respectively, among Swiss-born patients. Using weighted logistic regression, we found weak evidence of an association between birth origin and transmission (adjusted odds ratio of 2.2 and 95% CI of 0.9 to 5.5 comparing Swiss-born patients to others). In conclusion, standard genotyping overestimated recent TB transmission in Switzerland compared to WGS, particularly among immigrants from regions with a high TB incidence, where genetically closely related strains often predominate. We recommend the use of WGS to identify transmission clusters in settings with a low incidence of TB.

Tuberculosis (TB) remains an important public health concern in European countries (1–3). Immigrants from countries with a high TB incidence and HIV-infected populations are common risk groups in Switzerland, as in other European countries (4–8). We and others have previously shown that transmission of *Mycobacterium tuberculosis* occurs but is not more common among immigrants than in the native population (9–12).

Mycobacterial interspersed repetitive-unit–variable-number tandem-repeat (MIRU-VNTR) typing, combined with spoligotyping (13, 14), remains the most commonly used genotyping method in the molecular epidemiology of TB (15). However, MIRU-VNTR typing may not distinguish between genetically closely related strains despite the absence of close epidemiological links between patients (16–18). MIRU-VNTR typing may be suboptimal for studying transmission among immigrants from countries with a high TB incidence, where genetically closely related strains circulate over extended periods of time (19). Hence, recent transmission among immigrants is potentially overestimated, but the extent of this phenomenon is largely unknown. In contrast to standard genotyping methods, whole-genome sequencing (WGS) provides better resolution and has been used to study *M. tuberculosis* transmission (20–24). In this study, we reanalyzed transmis-

Received 8 February 2016 **Returned for modification** 10 March 2016
Accepted 29 April 2016

Accepted manuscript posted online 18 May 2016

Citation Stucki D, Ballif M, Egger M, Furrer H, Altpeter E, Battegay M, Droz S, Bruderer T, Coscolla M, Borrell S, Zürcher K, Janssens J-P, Calmy A, Mazza Stalder J, Jaton K, Rieder HL, Pfyffer GE, Siegrist HH, Hoffmann M, Fehr J, Dolina M, Frei R, Schrenzel J, Böttger EC, Gagneux S, Fenner L, for the Swiss HIV Cohort and Molecular Epidemiology of Tuberculosis Study Groups. 2016. Standard genotyping overestimates transmission of *Mycobacterium tuberculosis* among immigrants in a low-incidence country. J Clin Microbiol 54:1862–1870. doi:10.1128/JCM.00126-16.

Editor: K. C. Carroll, The Johns Hopkins University School of Medicine

Address correspondence to Lukas Fenner, lukas.fenner@ispm.unibe.ch.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JCM.00126-16.

sion clusters previously defined by MIRU-VNTR typing by using WGS to assess the transmission of *M. tuberculosis* among Swiss- and foreign-born TB patients (9).

MATERIALS AND METHODS

Study setting and study population. In 2012, we conducted a nationwide study of the molecular epidemiology of TB in Switzerland as a collaborative project involving the Swiss HIV Cohort Study (SHCS), the National Center for Mycobacteria, diagnostic microbiology laboratories, departments of respiratory medicine and public health, and the Federal Office of Public Health (www.tb-network.ch) (9, 25–28). The study setting was previously described in detail (9). Briefly, all of the patients in the SHCS diagnosed with TB between 2000 and 2008 were enrolled ($n = 93$). In addition, we included a random sample of 288 TB patients from the 4,221 patients with culture-confirmed TB reported to the National TB Surveillance Registry, and all of the drug-resistant TB patients reported in Switzerland ($n = 167$) during the same period (categories not mutually exclusive). Twenty-four-locus MIRU-VNTR typing and spoligotyping were used for the molecular detection of transmission clusters (9). In this follow-up study, we performed WGS of the 90 *M. tuberculosis* isolates belonging to 1 of the 35 MIRU-VNTR typing/spoligotyping clusters.

Clinical data collection and definitions. The clinical data collection was previously described in detail (9). The clustering proportion was determined by the n method and expressed as the number of patients in clusters divided by the total number of individuals (29). MIRU-VNTR typing clusters were defined as a group of isolates with identical MIRU-VNTR typing and spoligotyping patterns (9). In addition, we used IS*6110* restriction fragment length polymorphism (RFLP) patterns when they were available from the National Center for Mycobacteria (30). RFLP has a higher resolution than MIRU-VNTR typing, particularly for strains of the Beijing genotype. Isolates with identical MIRU-VNTR typing and spoligotype patterns but different IS*6110* patterns were considered nonclustered. Mixed molecular clusters were defined as clusters with Swiss- and foreign-born individuals or foreign-born individuals from different continents.

WGS and phylogenetic analyses. We generated whole-genome sequences for all of the 90 patient isolates identified as part of MIRU-VNTR typing clusters (9). We used Illumina Nextera XT or TruSeq library preparation kits and Illumina HiSeq, MiSeq, or NextSeq devices (Illumina, San Diego, CA) for WGS according to the manufacturer's instructions. Isolates were resequenced when the mean read depth was <20×. FastQ files from multiple sequencing runs of the same isolate were merged. We used KvarQ for an initial quality check and determination of *M. tuberculosis* phylogenetic lineages and *in silico* spoligotyping patterns as previously described (31). We then mapped short sequencing reads to a hypothetical *M. tuberculosis* ancestral genome (identical to H37Rv in structure but with maximum-likelihood-inferred ancestral bases [32]) with BWA 0.6.2 (33). Samtools 0.1.19 was used to call variants (SNPs). We only retained positions with read depths of ≥10% and ≤200% of the average read depth for the whole genome and a phred-scaled quality score of ≥30. We excluded positions in known repetitive regions (23), as well as SNPs in genes in which we have previously identified 50-bp sequences with homologous sequences elsewhere in the genome (see Table S1 in the supplemental material). We also excluded positions associated with drug resistance (31). For the analyses of read/allele counts at particular genomic loci, we extracted the number of high-quality bases from variant call format files with SNPeff/SNPsift (34).

Transmission networks based on SNP distances. We generated an alignment of all of the variable positions across the 90 isolates. We then calculated the raw genetic distances (number of SNPs) for each isolate pair in each MIRU-VNTR typing cluster with the compute pairwise distances function (with the pairwise deletion option) in MEGA 5.2.2 (35). We defined a MIRU-VNTR typing cluster as a true transmission cluster if all of the isolate pairs in the cluster were separated by ≤12 SNPs. In a sensitivity analysis, we opted for a stricter definition whereby a MIRU-VNTR typing transmission cluster was considered confirmed if at least one of its isolate pairs was separated by ≤5 SNPs. These thresholds of 12 and 5 SNPs were previously established by Walker et al. (17). We imported an alignment of the variable genomic positions into popart (http://popart.otago.ac.nz) to generate median-joining networks. Networks were generated for all 35 transmission clusters identified by standard genotyping (MIRU-VNTR typing and spoligotyping) (9).

Statistical analysis. We reanalyzed risk factors for transmission by using the WGS-based (true) cluster definition. We used weighted logistic regression models to obtain age- and sex-adjusted odds ratios (aORs) for the probability of belonging to a true molecular cluster. We used the Kruskal-Wallis rank sum test to assess differences between mean genetic distances of Swiss-born, foreign-born, and mixed clusters. As our study sample, by design, included more HIV-infected patients and patients with drug-resistant TB (9), we calculated weights to take sampling proportions into account. As a sensitivity analysis, we restricted the analysis of clustering proportion to the patients in the random sample from the TB registry ($n = 288$) (9). All statistical analyses were performed in Stata version 14 (Stata Corporation, College Station, TX) and R 3.1.2 (36).

In addition, we plotted the mean genetic distances (in SNPs) versus the mean geographic distances (in kilometers) of all of the patient pairs in a molecular cluster (distance between the birth countries' capital cities). Plots were generated with the ggplot2 library in R (37).

Ethics approval. This study was approved by the Ethics Committee of the Canton of Bern, Switzerland (9). Informed consent was obtained from all of the patients enrolled in the SHCS. For patients outside the SHCS, informed consent was obtained by the treating physicians. In some cases, informed consent could not be obtained from the patient because he or she could not be located or was known to have died. In these cases, we obtained permission from the Federal Expert Commission on Confidentiality in Medical Research to use the data provided by the treating physicians.

Nucleotide sequence accession number. Raw sequencing data were submitted to the European Nucleotide Archive and are accessible under BioProject number PRJEB12179.

RESULTS

Study population. The study population consisted of 520 TB patients from the nationwide study in Switzerland (9, 25–28). The patient characteristics are described in Table 1. A total of 119 (22.9%) patients were born in Switzerland, and 401 (77.1%) were born abroad. The median age was 36.5 (interquartile range, 28 to 51) years. Overall, 113 patients (21.7%) were HIV positive. Pulmonary TB accounted for 387 (74.4%) of the cases, and extrapulmonary TB accounted for 133 (25.6%) (Table 1).

Transmission clusters. (i) WGS. Isolates were sequenced with a median sequencing depth of 130× (range, 22× to 274×). For quality assurance, we compared laboratory assay-based phylogenetic lineage classifications and spoligotyping patterns with the results generated from the WGS data with KvarQ (9). We found 100% agreement between the two methods of lineage identification and up to two discordant spacers in the spoligotyping patterns.

(ii) Identification of molecular clusters based on WGS. In the 35 previously defined MIRU-VNTR typing clusters, we found pairwise genetic distances of 0 to 224 (median, 21.5) SNPs (Fig. 1; see Fig. S1 in the supplemental material). In the largest cluster (eight isolates), genetic distances were ≥54 SNPs. Seventeen (48.6%) of 35 MIRU-VNTR typing clusters consisted of pairs separated by ≤12 SNPs, i.e., were confirmed as true transmission clusters, corresponding to 43 (47.8%) of 90 patients (Table 2). The remaining 18 clusters harbored isolate pairs separated by >12 SNPs (47 patients).

The overall clustering proportion decreased from 17.3% (95%

TABLE 1 Characteristics of patients with TB diagnosed in Switzerland between 2000 and 2008 overall and comparing clustered and unclustered TB patients (unweighted), as well as risk factors for transmission (weighted analysis) as defined by genome-based molecular clustering

Characteristic	No. (%) of patients by unweighted analysis (n = 520)			Value for weighted analysis (n = 4,221)	
	All	Clustered (n = 43)	Unclustered (n = 477)	aOR (95% CI)[a]	P
Age (yr) at TB diagnosis					0.46
16–29	154 (29.6)	14 (32.6)	140 (29.4)	1	
30–49	226 (43.5)	16 (37.2)	210 (44.0)	0.56 (0.22–1.41)	
≥50	140 (26.9)	13 (30.2)	127 (26.6)	0.71 (0.29–1.73)	
Sex					0.018
Male	254 (48.8)	26 (60.5)	228 (47.8)	1	
Female	266 (51.2)	17 (39.5)	249 (52.2)	0.4 (0.18–0.85)	
Country of birth					0.026
Switzerland	119 (22.9)	17 (39.5)	102 (21.4)	1	
Europe but not Switzerland	106 (20.4)	2 (4.7)	104 (21.8)	0.16 (0.03–0.83)	
Asia	120 (23.1)	3 (7.0)	117 (24.5)	0.24 (0.05–1.08)	
Sub-Saharan Africa	132 (25.4)	16 (37.2)	116 (24.3)	1.11 (0.36–3.45)	
Central and South America	20 (3.8)	3 (7.0)	17 (3.6)	1.99 (0.40–9.85)	
Other regions/Unknown	23 (4.4)	2 (4.7)	21 (4.4)	0.80 (0.13–5.02)	
Born in Switzerland					0.091
No	401 (77.1)	26 (60.5)	375 (78.6)	1	
Yes	119 (22.9)	17 (39.5)	102 (21.4)	2.21 (0.88–5.52)	
Residence status					0.41
Swiss national	130 (25.0)	14 (32.6)	116 (24.3)	1	
Foreigner/resident	199 (38.3)	11 (25.6)	188 (39.4)	0.45 (0.17–1.23)	
Asylum seeker/refugee	101 (19.4)	8 (18.6)	93 (19.5)	0.42 (0.11–1.56)	
Unknown	90 (17.3)	10 (23.3)	80 (16.8)	0.80 (0.22–2.81)	
HIV status					0.033
Negative	407 (78.3)	34 (79.1)	373 (78.2)	1	
Positive	113 (21.7)	9 (20.9)	104 (21.8)	0.39 (0.16–0.93)	
Site of TB					0.054
Extrapulmonary	133 (25.6)	5 (11.6)	128 (26.8)	1	
Pulmonary	387 (74.4)	38 (88.4)	349 (73.2)	3.14 (0.98–10.07)	
Cavitary disease					0.039
No	407 (78.3)	28 (65.1)	379 (79.5)	1	
Yes	113 (21.7)	15 (34.9)	98 (20.5)	2.31 (1.05–5.10)	
Positive smear microscopy					0.049
No	352 (67.7)	21 (48.8)	331 (69.4)	1	
Yes	168 (32.3)	22 (51.2)	146 (30.6)	2.18 (1.00–4.73)	
TB in family or social surroundings in last 2 yr					0.068
No	482 (92.7)	38 (88.4)	444 (93.1)	1	
Yes	38 (7.3)	5 (11.6)	33 (6.9)	2.95 (0.92–9.41)	

[a] Adjusted for age and sex and weighted for sampling proportions.

confidence interval [95% CI], 14.2 to 20.8%) based on standard genotyping (spoligotyping and MIRU-VNTR typing) to 8.3% (95% CI, 6.0 to 11.0%) based on WGS. When restricting the analysis to the 288 randomly selected patients, we found 27 clustered patients in 11 genome-based clusters, resulting in a clustering proportion of 9.4% (95% CI, 6.3 to 13.3%). When using a more stringent WGS definition of transmission clusters (at least one isolate pair at a cluster distance of ≤5 SNPs), 13 of 35 (37.1%; CI, 21.5 to 55.1%) MIRU-VNTR typing clusters were confirmed. These 13 transmission clusters included 35/520 patients, corresponding to a clustering proportion of 6.7% (95% CI, 4.7 to 9.2%).

(iii) Infection with multiple *M. tuberculosis* strains. In five isolates that were part of transmission clusters defined by MIRU-VNTR typing, we detected multiple alleles at several MIRU-VNTR typing loci, potentially indicating infection with multiple *M. tuberculosis* strains. We therefore conducted an allele frequency analysis based on sequencing reads for each SNP call. De-

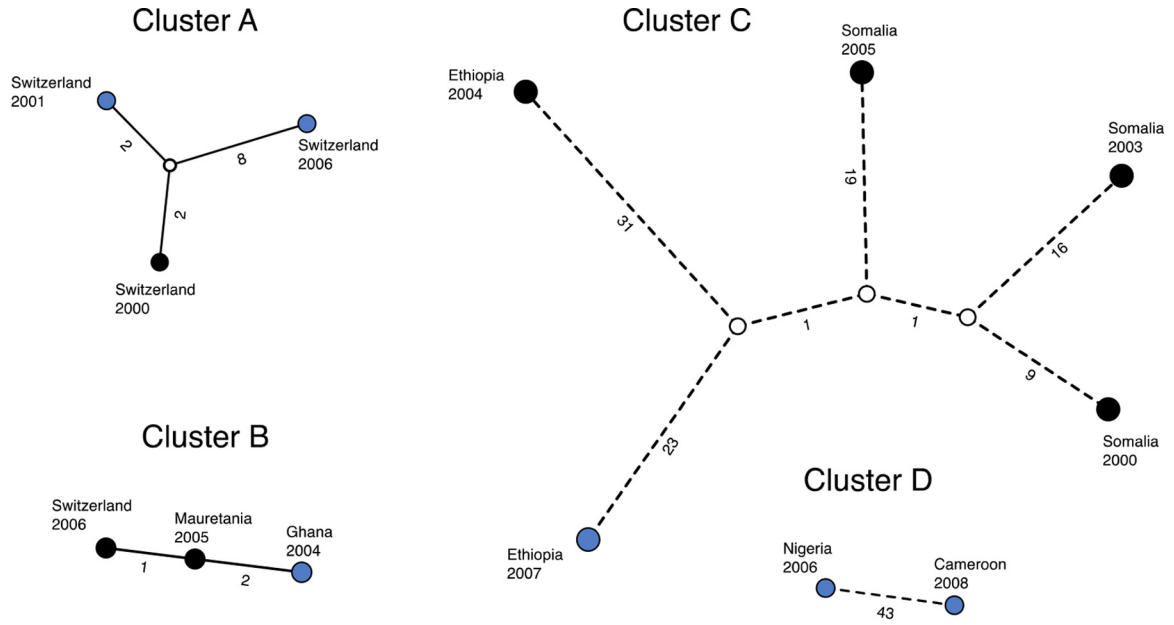

FIG 1 *M. tuberculosis* transmission networks with SNP data from WGS. Representative examples of different types of transmission clusters that were previously identified by MIRU-VNTR typing are shown. Cluster A (all patients from Switzerland) was confirmed as a true transmission cluster by WGS, with distances of ≤12 SNPs between all isolates. Cluster B is a mixed cluster (one patient from Switzerland and two patients from West Africa) that was confirmed to be a true transmission cluster with one and two SNPs between patient isolates. Clusters C and D were identified by MIRU-VNTR typing as transmission clusters, but WGS did not confirm these clusters as true clusters (genetic distances of >12 SNPs). Filled circles represent patient isolates; white circles represent median vectors, i.e., hypothetical isolates inferred from the sequencing data; and blue circles indicate HIV-positive patients. The values next to lines are SNP distances. Countries of birth and years of TB diagnosis are indicated next to circles. Clusters with solid lines are true clusters, i.e., clusters confirmed by WGS (≤12 SNPs), whereas clusters with dashed lines are clusters not confirmed by WGS (>12 SNPs). All other transmission clusters are shown in Fig. S1 in the supplemental material.

spite the presence of multiallelic variant calls in all of the isolates (potential microevolutionary events), none of the five isolates with multiple MIRU-VNTR typing bands showed evidence of lineage- or sublineage-specific markers with multiple alleles in the sequencing reads.

TABLE 2 Numbers of MIRU-VNTR typing clusters and numbers of patients in MIRU-VNTR typing clusters confirmed by WGS[a]

Category	No. (%) of MIRU-VNTR typing clusters or patients		
	Confirmed by WGS	Not confirmed by WGS	Total no.
Molecular clusters[b]			
Swiss-born only	3 (75.0)	1 (25.0)	4
Mixed	10 (66.7)	5 (33.3)	15
Foreign-born only	4 (25.0)	12 (75.0)	16
Total	17 (48.6)	18 (51.4)	35
Patients in molecular clusters[c]			
Swiss-born only	8 (80.0)	2 (20.0)	10
Mixed	27 (62.8)	16 (37.2)	43
Foreign-born only	8 (21.6)	29 (78.4)	37
Total	43 (47.8)	47 (52.2)	90

[a] MIRU-VNTR typing clusters not confirmed by WGS (false-positive clusters) and clusters confirmed by WGS (true clusters) are presented according to the countries of birth of the patients involved as follows: clusters involving Swiss-born TB patients only, foreign-born patients only, or mixed clusters (e.g., involving both Swiss- and foreign-born TB patients or foreign-born patients from different continents).
[b] Fisher's exact test *P* value of 0.031.
[c] Fisher's exact test *P* value of <0.0001.

Molecular clustering in Swiss-born, foreign-born, and HIV-positive patients. Four MIRU-VNTR typing clusters involved Swiss-born patients only, 16 clusters involved foreign-born patients only, and 15 clusters were of mixed birth group origin. Three (75.0%) of four clusters involving only Swiss-born patients were confirmed by WGS as true clusters (8/10 [80.0%] clustered Swiss-born patients). On the other hand, only 4/16 (25.0%) immigrant clusters (born on the same continent) were true clusters (8/37 [21.6%] patients). Of the 15 mixed clusters, 10 (66.7%) were true clusters (27/43 [62.8%] clustered patients) (Table 2). We also assessed whether foreign-born patients were overrepresented in MIRU-VNTR typing clusters not confirmed by WGS. Among the 90 patients in the MIRU-VNTR typing clusters, foreign-born patients were more likely to be in clusters not confirmed by WGS than in true clusters (aOR, 4.5; CI, 1.5 to 13.6; $P = 0.008$).

We then calculated the true (genome-based) clustering proportions of both Swiss- and foreign-born patients. The clustering proportion among Swiss-born patients decreased only slightly, from 19.3% (23/119; 95% CI, 12.7 to 27.6%) on the basis of MIRU-VNTR typing to 14.3% (17/119; 95% CI, 8.5 to 21.9) on the basis of WGS data. In contrast, the clustering proportion among immigrants was more than halved, from 16.7% (67/401; 95% CI, 13.2 to 20.7) to 6.5% (26/401; 95% CI, 4.3 to 9.4). Figure 2 summarizes the possible factors leading to an overestimation of *M. tuberculosis* transmission on the basis of standard genotyping among foreign-born and native TB patients in settings with a low incidence of TB.

The median genetic distance differed significantly among the three groups: 9 SNPs (range, 8 to 15) in clusters with Swiss-born

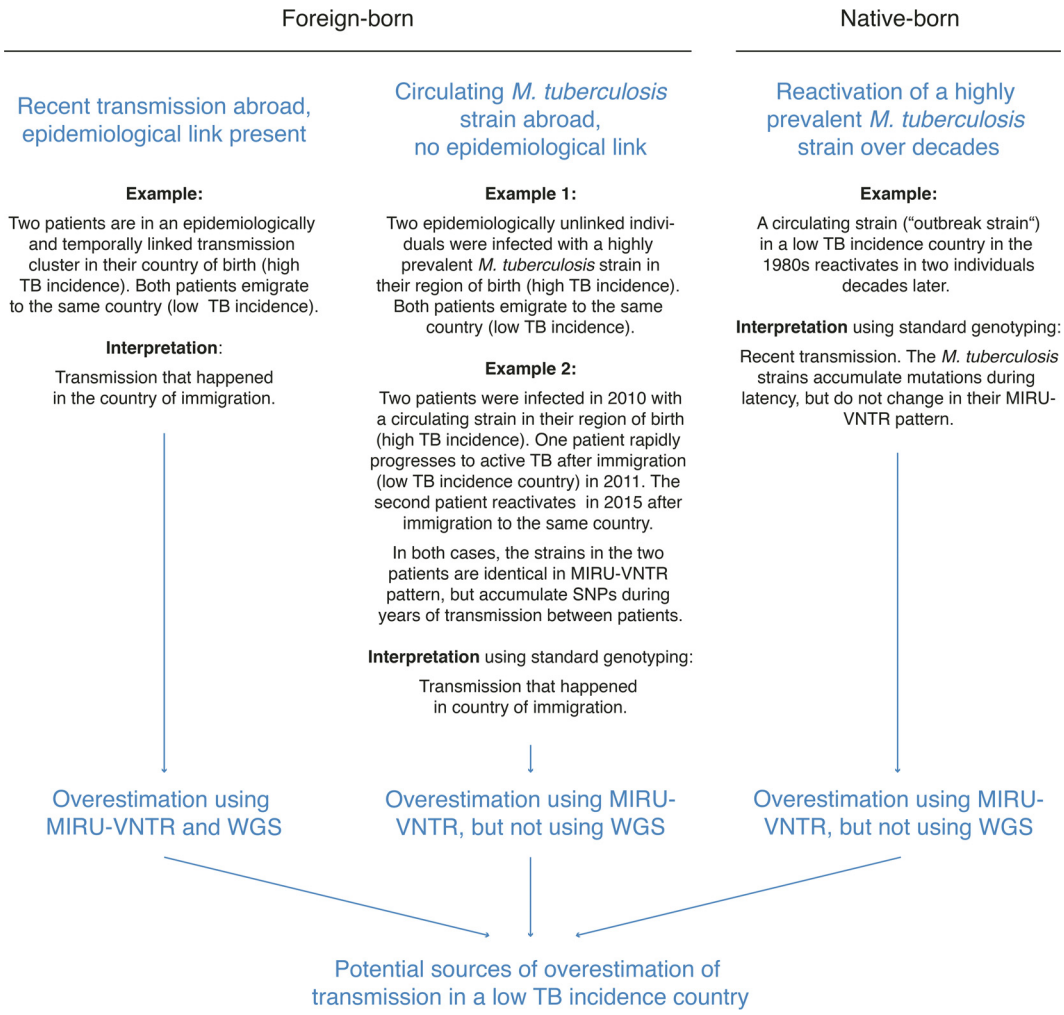

FIG 2 Possible situations leading to overestimation of *M. tuberculosis* transmission in a country with a low incidence of TB.

individuals only, 2 SNPs (range, 0 to 162) in mixed clusters, and 24 SNPs (range, 2 to 224) in clusters with foreign-born individuals only ($P = 0.030$) (Fig. 3).

Geographic and genetic distances within molecular clusters. Plotting the mean genetic distance (in SNPs) versus the mean geographic distance between patient origins (capital cities) for each molecular cluster further supported the different patterns of clustering by the birth origins of the patients (see Fig. S2 in the supplemental material). Among the foreign-born patients, a majority of MIRU-VNTR typing clusters (11/16, 68.8%) consisted of patients born within a geographic distance of 3,500 km but harboring genetic distances of >12 SNPs. This indicates that genetically closely related strains circulating in a geographically restricted area in the region of origin were imported into Switzerland. Among the mixed MIRU-VNTR typing clusters, the transmission of which is expected to have happened in Switzerland, a majority (9/15, 60%) had mean geographic distances of >3,500 km but genetic distances of ≤12 SNPs, indicating recent transmission in Switzerland (see Fig. S2 in the supplemental material).

Risk factors for transmission. Female (aOR, 0.39; CI, 0.18 to 0.85) and HIV-positive (aOR, 0.39; CI, 0.16 to 0.93) patients were significantly less likely to be involved in true transmission clusters (Table 1). In contrast, patients with cavitary disease were more likely to be associated with transmission (aOR, 2.31; CI, 1.05 to 5.10). There was weak evidence of an association between being born in Switzerland and being involved in a true transmission cluster (an aOR of 2.21 and a 95% CI of 0.88 to 5.52 for Swiss-born versus foreign-born patients). Overall, the risk factors for transmission remained similar to those previously reported on the basis of standard genotyping methods (9).

DISCUSSION

In this nationwide study of 520 TB patients, 18 of 35 transmission clusters identified by standard molecular genotyping (spoligotyping and MIRU-VNTR typing) were refuted by WGS. This suggests that transmission of *M. tuberculosis* is generally overestimated in countries with a low incidence of TB such as Switzerland. Furthermore, we found a striking difference between transmission clusters involving Swiss-born patients and clusters involving foreign-born patients. WGS confirmed three quarters of the clusters involving Swiss-born individuals only but only one quarter of the clusters involving foreign-born patients only, hence indicating

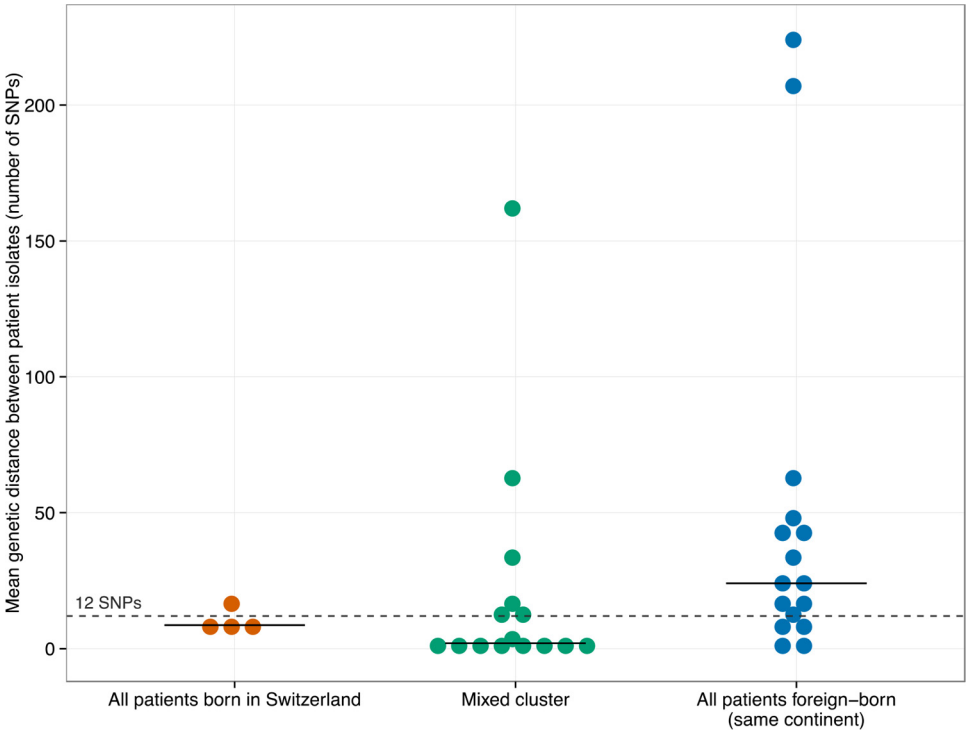

FIG 3 Median genetic distances in MIRU-VNTR typing-defined transmission clusters. Each data point shows the mean genetic distance (as the number of SNPs) in 1 of the 35 MIRU-VNTR typing-defined transmission clusters. Solid black lines indicate median values of mean pairwise distances within molecular clusters. The distribution of clusters of patients born in Switzerland, mixed clusters, and clusters of immigrant patients born on the same continent (except Switzerland) was significantly different (Kruskal-Wallis nonparametric test, $P = 0.030$). The dashed line represents the cutoff for the definition of genome-based molecular clusters (≥12 SNPs).

that transmission was especially overestimated in the immigrant population.

M. tuberculosis strains from immigrants, which were defined as clustered by MIRU-VNTR typing but not by WGS, are likely genetically closely related genotypes imported independently from a high-incidence region where they are highly prevalent (19). Such strains accumulate genetic mutations over time (often leading to pairwise SNP distances of >12 SNPs), but the MIRU-VNTR typing pattern may not change. In such a situation, identical MIRU-VNTR typing will be wrongly interpreted as recent transmission in the country of immigration (9). Similar observations were made in the United Kingdom, where immigrant TB patients were identified in transmission clusters on the basis of standard MIRU-VNTR genotyping, although no epidemiological link could be found during contact investigations (16).

The clustering proportion (indicating recent transmission) among Swiss-born individuals was similar when standard genotyping and WGS were used but more than 2-fold lower among foreign-born individuals when WGS was used. In reality, the clustering proportion among foreign-born individuals might even be lower, as we cannot exclude the possibility that WGS-confirmed clusters (≤12 SNPs) involving immigrants might partly represent transmission that happened in the country of origin and not in Switzerland. Only social contact tracing could provide further insights into transmission dynamics, but such investigations are notoriously difficult to perform, particularly among immigrants (23, 38). The low proportion of true transmission clusters among immigrants in our study was further supported by the weighted analysis of predictors of transmission, which showed that foreign-born TB patients tended to be less likely than Swiss-born patients to be involved in true clusters. Of note, the clustering proportion among immigrants is remarkably similar to previous observations among immigrant patients with multidrug resistance diagnosed in Switzerland, which showed a clustering proportion of 8% (compared to 7% in our study) on the basis of standard genotyping and contact tracing (30).

The majority of mixed molecular clusters as defined by MIRU-VNTR typing (i.e., involving Swiss- and foreign-born individuals) showed small SNP distances (≤12 SNPs), confirming the intuitive explanation that transmission between Swiss-born and foreign-born patients likely occurred in Switzerland. This was further supported by the analysis of geographic distances between patient birth countries, which indicated that most of the isolate pairs in the mixed clusters were from patients born far away from each other, despite small genetic distances. The five mixed clusters harboring larger SNP distances may reflect TB cases due to infections by circulating global or European *M. tuberculosis* genotypes, such as the recently described large cluster in Eastern Europe (39).

We found no evidence of infection with multiple strains among clustered TB cases in the WGS data, despite the presence of double alleles in the MIRU-VNTR typing patterns of five clustered isolates. Infections with multiple strains could potentially also influence the identification of molecular clusters, as individual strains in an infection with multiple strains cannot be resolved by MIRU-VNTR typing. The prevalence and relevance of such multiple infections need to be studied further (40–42).

A potential limitation of our study is the definition of transmission clusters by WGS. The threshold of 12 SNPs that we used to exclude transmission has been established by Walker et al. (17) and is in the range of what other studies have used (21, 43, 44). However, an adequate cluster definition may be adapted according to the setting (low versus high TB incidence), the study population, and the technical specifications of the WGS analysis pipeline (i.e., whether particular genomic regions such as the PE/polymorphic GC-repetitive sequence genes are excluded from the analysis, which was the case here). For comparison, we repeated the analyses with a stricter definition of transmission clusters also proposed by Walker et al. (MIRU-VNTR typing clusters in which at least one pairwise distance was ≤5 SNPs [17]), which further reduced the number of true clusters to 13 but did not change the overall clustering proportion significantly. A further limitation may be our sample size: we included 12.3% of the TB cases diagnosed between 2000 and 2008 in Switzerland, which potentially underestimates the overall clustering proportion (29). Indeed, SNP distances could, in fact, become shorter upon the inclusion of additional patient isolates with intermediate genotypes, hence increasing the proportion of true clusters.

In conclusion, only one quarter of the foreign-born transmission clusters previously identified by MIRU-VNTR typing were confirmed as true transmission clusters by WGS. We therefore recommend the use of WGS for more accurate identification of recent transmission of *M. tuberculosis* among immigrants in countries with a low incidence of TB but also in countries with a high TB incidence, where genetically closely related strains circulate. Although WGS analysis remains resource intensive, the strategy adopted in the United Kingdom documents that implementing WGS in the routine public laboratory surveillance system is feasible (21, 45) and allows the prompt identification of transmission clusters, as well as information about the drug resistance genotype (45, 46). Our results also indicate that the native population in Switzerland may also play a role in spreading TB, particularly individuals belonging to high-risk populations (22, 23). Additional prospective studies using WGS are needed, possibly complemented with social network analyses (20), to evaluate the usefulness of real-time analyses of TB transmission dynamics in countries with a low incidence of TB.

ACKNOWLEDGMENTS

We thank all of the TB patients who participated in this study, the treating physicians who provided clinical information, and the microbiology laboratories providing strains. We are grateful to the National TB Surveillance Registry at the Federal Office of Public Health.

The members of the Molecular Epidemiology of Tuberculosis Study Group are as follows. The central coordinating team consists of Lukas Fenner and Matthias Egger, Institute of Social and Preventive Medicine, Bern; Sebastien Gagneux and Marcel Tanner, Swiss Tropical and Public Health Institute, Basel; and Hansjakob Furrer, Inselspital Bern. The National Center for Mycobacteria is represented by Erik C. Böttger, Institute of Medical Microbiology, University of Zürich, Zürich, Switzerland. The microbiology laboratory members are Reno Frei, Clinical Microbiology, University Hospital Basel; Thomas Bodmer and Sara Droz, Institute for Infectious Diseases, University of Bern; Jacques Schrenzel, Laboratory of Bacteriology, University Hospitals of Geneva; Katia Jaton, Institute of Microbiology, University Hospital of Lausanne; Hans Siegrist, ADMed Microbiology, La Chaux-de-Fonds; Gaby E. Pfyffer, Department of Medical Microbiology, Luzerner Kantonsspital, Lucerne; Thomas Bruderer and Detlev Schultze, Center for Laboratory Medicine, St. Gallen; Marisa Dolina, Clinical Microbiology, EOLAB, Bellinzona, Switzerland; and Olivier Dubuis, Viollier AG Switzerland, Allschwil. The SHCS is represented by Manuel Battegay, University Hospital Basel; Enos Bernasconi and Andrea Parini, Lugano; Matthias Hoffmann, St. Gallen; Hansjakob Furrer, Inselspital Bern; Matthias Cavassini, University Hospital of Lausanne; Bernard Hirschel and Alexandra Calmy, University Hospital of Geneva; and Jan Fehr, University Hospital of Zürich. The respiratory clinic members are Jean-Paul Janssens, University Hospital of Geneva, and Jesica Mazza Stalder, University Hospital of Lausanne. The Federal Office of Public Health is represented by Peter Helbling and Ekkehardt Altpeter, Division of Communicable Diseases. Also included is Hans L. Rieder, Epidemiology, Biostatistics, and Prevention Institute, University of Zürich, Zürich, Switzerland.

The members of the SHCS are J. Barth, M. Battegay, E. Bernasconi, J. Böni, H. C. Bucher, C. Burton-Jeangros, A. Calmy, M. Cavassini, C. Cellerai, M. Egger, L. Elzi, J. Fehr, J. Fellay, M. Flepp, P. Francioli (president of the SHCS), H. Furrer (chairman of the Clinical and Laboratory Committee), C. A. Fux, M. Gorgievski, H. Günthard (chairman of the Scientific Board), D. Haerry (deputy of the Positive Council), B. Hasse, H. H. Hirsch, B. Hirschel, I. Hösli, C. Kahlert, L. Kaiser, O. Keiser, C. Kind, T. Klimkait, H. Kovari, B. Ledergerber, G. Martinetti, B. Martinez de Tejada, K. Metzner, N. Müller, D. Nadal, G. Pantaleo, A. Rauch, S. Regenass, M. Rickenbach (head of the Data Center), C. Rudin (chairman of the Mother & Child Substudy), P. Schmid, D. Schultze, F. Schöni-Affolter, J. Schüpbach, R. Speck, P. Taffé, P. Tarr, A. Telenti, A. Trkola, P. Vernazza, R. Weber, and S. Yerly.

We have no conflicts of interest to declare.

FUNDING INFORMATION

This work was supported by the Swiss National Science Foundation (grants 324730-12544 and PP00P3_150750), the SHCS (grants 588 and 740), the Federal Office of Public Health (grant 09.007368), the National Center for Mycobacteria, and the European Research Council (grant 309540-EVODRTB). The SHCS is supported by the Swiss National Science Foundation (grant 33CS30-134277). Sequencing analyses were performed at the sciCORE (http://scicore.unibas.ch/) scientific computing core facility at the University of Basel. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. **World Health Organization.** 2014. Global tuberculosis report 2014. World Health Organization, Geneva, Switzerland. http://apps.who.int/iris/bitstream/10665/137094/1/9789241564809_eng.pdf.
2. **Falzon D, Kudjawu Y, Desenclos JC, Fernandez de la Hoz K, Dadu A, Zaleskis R.** 2008. Stopping TB in Europe: some progress but still not there. Euro Surveill **13:**27–30. http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=8073.
3. **Zellweger JP.** 2013. Current issues in the management of tuberculosis in Europe. Panminerva Med **55:**145–155.
4. **Swiss Federal Office of Public Health.** 2011. Tuberkulose in der Schweiz 2005–2009. Bull BAG **10:**205–213. (In German.) http://www.bag.admin.ch/themen/medizin/00682/00684/01108/index.html?lang=de.
5. **Abgrall S, Del Giudice P, Melica G, Costagliola D, FHDH-ANRS CO4.** 2010. HIV-associated tuberculosis and immigration in a high-income country: incidence trends and risk factors in recent years. AIDS **24:**763–771. http://dx.doi.org/10.1097/QAD.0b013e3283366747.
6. **Dahle UR, Eldholm V, Winje BA, Mannsåker T, Heldal E.** 2007. Impact of immigration on the molecular epidemiology of *Mycobacterium tuberculosis* in a low-incidence country. Am J Respir Crit Care Med **176:**930–935. http://dx.doi.org/10.1164/rccm.200702-187OC.
7. **Wolff H, Janssens JP, Bodenmann P, Meynard A, Delhumeau C, Rochat T, Sudre P, Costanza MC, Gaspoz JM, Morabia A.** 2010. Undocumented migrants in Switzerland: geographical origin versus legal status as risk factor for tuberculosis. J Immigr Minor Health **12:**18–23. http://dx.doi.org/10.1007/s10903-009-9271-6.
8. **Lillebaek T, Andersen AB, Bauer J, Dirksen A, Glismann S, de Haas P, Kok-Jensen A.** 2001. Risk of *Mycobacterium tuberculosis* transmission in a

low-incidence country due to immigration from high-incidence areas. J Clin Microbiol **39:**855–861. http://dx.doi.org/10.1128/JCM.39.3.855-861.2001.

9. **Fenner L, Gagneux S, Helbling P, Battegay M, Rieder HL, Pfyffer GE, Zwahlen M, Furrer H, Siegrist HH, Fehr J, Dolina M, Calmy A, Stucki D, Jaton K, Janssens J-P, Stalder JM, Bodmer T, Ninet B, Böttger EC, Egger M.** 2012. *Mycobacterium tuberculosis* transmission in a country with low tuberculosis incidence: role of immigration and HIV infection. J Clin Microbiol **50:**388–395. http://dx.doi.org/10.1128/JCM.05392-11.
10. **Barniol J, Niemann S, Louis VR, Brodhun B, Dreweck C, Richter E, Becher H, Haas W, Junghanss T.** 2009. Transmission dynamics of pulmonary tuberculosis between autochthonous and immigrant subpopulations. BMC Infect Dis **9:**197. http://dx.doi.org/10.1186/1471-2334-9-197.
11. **Kamper-Jørgensen Z, Andersen AB, Kok-Jensen A, Bygbjerg IC, Andersen PH, Thomsen VO, Kamper-Jørgensen M, Lillebaek T.** 2012. Clustered tuberculosis in a low-burden country: nationwide genotyping through 15 years. J Clin Microbiol **50:**2660–2667. http://dx.doi.org/10.1128/JCM.06358-11.
12. **Guzman Herrador BR, Rønning K, Borgen K, Mannsåker T, Dahle UR.** 2015. Description of the largest cluster of tuberculosis notified in Norway 1997–2011: is the Norwegian tuberculosis control programme serving its purpose for high risk groups? BMC Public Health **15:**367. http://dx.doi.org/10.1186/s12889-015-1701-x.
13. **Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B, Locht C.** 2000. Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. Mol Microbiol **36:**762–771.
14. **Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, van Embden J.** 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol **35:**907–914.
15. **Kato-Maeda M, Metcalfe JZ, Flores L.** 2011. Genotyping of *Mycobacterium tuberculosis*: application in epidemiologic studies. Future Microbiol **6:**203–216. http://dx.doi.org/10.2217/fmb.10.165.
16. **Munang ML, Browne C, Khanom S, Evans JT, Smith EG, Hawkey PM, Kunst H, Welch SB, Dedicoat MJ.** 2015. Tuberculosis microepidemics among dispersed migrants, Birmingham, UK, 2004–2013. Emerg Infect Dis **21:**524–527. http://dx.doi.org/10.3201/eid2103.140209.
17. **Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, Eyre DW, Wilson DJ, Hawkey PM, Crook DW, Parkhill J, Harris D, Walker AS, Bowden R, Monk P, Smith EG, Peto TE.** 2013. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. Lancet Infect Dis **13:**137–146. http://dx.doi.org/10.1016/S1473-3099(12)70277-3.
18. **Glynn JR, Bauer J, de Boer AS, Borgdorff MW, Fine PE, Godfrey-Faussett P, Vynnycky E.** 1999. Interpreting DNA fingerprint clusters of *Mycobacterium tuberculosis*. European concerted action on molecular epidemiology and control of tuberculosis. Int J Tuberc Lung Dis **3:**1055–1060.
19. **Wampande EM, Mupere E, Debanne SM, Asiimwe BB, Nsereko M, Mayanja H, Eisenach K, Kaplan G, Boom HW, Sebastien G, Joloba ML.** 2013. Long-term dominance of *Mycobacterium tuberculosis* Uganda family in peri-urban Kampala-Uganda is not associated with cavitary disease. BMC Infect Dis **13:**484. http://dx.doi.org/10.1186/1471-2334-13-484.
20. **Gardy JL, Johnston JC, Ho Sui SJ, Cook VJ, Shah L, Brodkin E, Rempel S, Moore R, Zhao Y, Holt R, Varhol R, Birol I, Lem M, Sharma MK, Elwood K, Jones SJ, Brinkman FS, Brunham RC, Tang P.** 2011. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. N Engl J Med **364:**730–739. http://dx.doi.org/10.1056/NEJMoa1003176.
21. **Walker TM, Lalor MK, Broda A, Ortega LS, Morgan M, Parker L, Churchill S, Bennett K, Golubchik T, Giess AP, Del Ojo Elias C, Jeffery KJ, Bowler ICJW, Laurenson IF, Barrett A, Drobniewski F, McCarthy ND, Anderson LF, Abubakar I, Thomas HL, Monk P, Smith EG, Walker AS, Crook DW, Peto TE, Conlon CP.** 2014. Assessment of *Mycobacterium tuberculosis* transmission in Oxfordshire, UK, 2007-12, with whole pathogen genome sequences: an observational study. Lancet Respir Med **2:**285–292. http://dx.doi.org/10.1016/S2213-2600(14)70027-X.
22. **Roetzer A, Diel R, Kohl TA, Rückert C, Nübel U, Blom J, Wirth T, Jaenicke S, Schuback S, Rüsch-Gerdes S, Supply P, Kalinowski J, Niemann S.** 2013. Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study. PLoS Med **10:**e1001387. http://dx.doi.org/10.1371/journal.pmed.1001387.
23. **Stucki D, Ballif M, Bodmer T, Coscolla M, Maurer A-M, Droz S, Butz C, Borrell S, Längle C, Feldmann J, Furrer H, Mordasini C, Helbling P, Rieder HL, Egger M, Gagneux S, Fenner L.** 2015. Tracking a tuberculosis outbreak over 21 years: strain-specific single-nucleotide polymorphism typing combined with targeted whole-genome sequencing. J Infect Dis **211:**1306–1316. http://dx.doi.org/10.1093/infdis/jiu601.
24. **Coscolla M, Barry PM, Oeltmann JE, Koshinsky H, Shaw T, Cilnis M, Posey J, Rose J, Weber T, Fofanov VY, Gagneux S, Kato-Maeda M, Metcalfe JZ.** 2015. Genomic epidemiology of multidrug-resistant *Mycobacterium tuberculosis* during transcontinental spread. J Infect Dis **212:**302–310. http://dx.doi.org/10.1093/infdis/jiv025.
25. **Fenner L, Egger M, Bodmer T, Altpeter E, Zwahlen M, Jaton K, Pfyffer GE, Borrell S, Dubuis O, Bruderer T, Siegrist HH, Furrer H, Calmy A, Fehr J, Stalder JM, Ninet B, Böttger EC, Gagneux S.** 2012. Effect of mutation and genetic background on drug resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother **56:**3047–3053. http://dx.doi.org/10.1128/AAC.06460-11.
26. **Fenner L, Gagneux S, Janssens J-P, Fehr J, Cavassini M, Hoffmann M, Bernasconi E, Schrenzel J, Bodmer T, Böttger EC, Helbling P, Egger M, Swiss HIV Cohort and Molecular Epidemiology of Tuberculosis Study Groups.** 2012. Tuberculosis in HIV-negative and HIV-infected patients in a low-incidence country: clinical characteristics and treatment outcomes. PLoS One **7:**e34186. http://dx.doi.org/10.1371/journal.pone.0034186.
27. **Fenner L, Malla B, Ninet B, Dubuis O, Stucki D, Borrell S, Huna T, Bodmer T, Egger M, Gagneux S.** 2011. "Pseudo-Beijing": evidence for convergent evolution in the direct repeat region of *Mycobacterium tuberculosis*. PLoS One **6:**e24737. http://dx.doi.org/10.1371/journal.pone.0024737.
28. **Fenner L, Egger M, Bodmer T, Furrer H, Ballif M, Battegay M, Helbling P, Fehr J, Gsponer T, Rieder HL, Zwahlen M, Hoffmann M, Bernasconi E, Cavassini M, Calmy A, Dolina M, Frei R, Janssens J-P, Borrell S, Stucki D, Schrenzel J, Böttger EC, Gagneux S, Swiss HIV Cohort and Molecular Epidemiology of Tuberculosis Study Groups.** 2013. HIV infection disrupts the sympatric host-pathogen relationship in human tuberculosis. PLoS Genet **9:**e1003318. http://dx.doi.org/10.1371/journal.pgen.1003318.
29. **Glynn JR, Vynnycky E, Fine PE.** 1999. Influence of sampling on estimates of clustering and recent transmission of *Mycobacterium tuberculosis* derived from DNA fingerprinting techniques. Am J Epidemiol **149:**366–371. http://dx.doi.org/10.1093/oxfordjournals.aje.a009822.
30. **Somoskovi A, Helbling P, Deggim V, Hömke R, Ritter C, Böttger EC.** 2014. Transmission of multidrug-resistant tuberculosis in a low-incidence setting, Switzerland, 2006 to 2012. Euro Surveill **19:**24–28. http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20736.
31. **Steiner A, Stucki D, Coscolla M, Borrell S, Gagneux S.** 2014. KvarQ: targeted and direct variant calling from fastq reads of bacterial genomes. BMC Genomics **15:**881. http://dx.doi.org/10.1186/1471-2164-15-881.
32. **Comas I, Chakravartti J, Small PM, Galagan J, Niemann S, Kremer K, Ernst JD, Gagneux S.** 2010. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. Nat Genet **42:**498–503. http://dx.doi.org/10.1038/ng.590.
33. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25:**1754–1760. http://dx.doi.org/10.1093/bioinformatics/btp324.
34. **Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, Lu X.** 2012. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. Front Genet **3:**35. http://dx.doi.org/10.3389/fgene.2012.00035.
35. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28:**2731–2739. http://dx.doi.org/10.1093/molbev/msr121.
36. **R Core Team.** 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
37. **Wickham H.** 2009. ggplot2: elegant graphics for data analysis. Springer Verlag New York, NY.
38. **Cook VJ, Shah L, Gardy J, Bourgeois A-C.** 2012. Recommendations on modern contact investigation methods for enhancing tuberculosis control. Int J Tuberc Lung Dis **16:**297–305. http://dx.doi.org/10.5588/ijtld.11.0350.

39. **De Beer JL, Kodmon C, van der Werf MJ, van Ingen J, van Soolingen D, ECDC MDR-TB Molecular Surveillance Project Participants.** 2014. Molecular surveillance of multi- and extensively drug-resistant tuberculosis transmission in the European Union from 2003 to 2011. Euro Surveill **19:**5–13. http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20742.
40. **Huyen MN, Kremer K, Lan NT, Cobelens FG, Buu TN, Dung NH, Caws M, Tiemersma EW, van Soolingen D.** 2012. Mixed tuberculosis infections in rural South Vietnam. J Clin Microbiol **50:**1586–1592. http://dx.doi.org/10.1128/JCM.00434-12.
41. **Cohen T, Wilson D, Wallengren K, Samuel EY, Murray M.** 2011. Mixed-strain *Mycobacterium tuberculosis* infections among patients dying in a hospital in KwaZulu-Natal, South Africa. J Clin Microbiol **49:**385–388. http://dx.doi.org/10.1128/JCM.01378-10.
42. **Mallard K, McNerney R, Crampin AC, Houben R, Ndlovu R, Munthali L, Warren RM, French N, Glynn JR.** 2010. Molecular detection of mixed infections of *Mycobacterium tuberculosis* strains in sputum samples from patients in Karonga District, Malawi. J Clin Microbiol **48:**4512–4518. http://dx.doi.org/10.1128/JCM.01683-10.
43. **Cohen KA, Abeel T, Manson McGuire A, Desjardins CA, Munsamy V, Shea TP, Walker BJ, Bantubani N, Almeida DV, Alvarado L, Chapman SB, Mvelase NR, Duffy EY, Fitzgerald MG, Govender P, Gujja S, Hamilton S, Howarth C, Larimer JD, Maharaj K, Pearson MD, Priest ME, Zeng Q, Padayatchi N, Grosset J, Young SK, Wortman J, Mlisana KP, O'Donnell MR, Birren BW, Bishai WR, Pym AS, Earl AM.** 2015. Evolution of extensively drug-resistant tuberculosis over four decades: whole-genome sequencing and dating analysis of *Mycobacterium tuberculosis* isolates from KwaZulu-Natal. PLoS Med **12:**e1001880. http://dx.doi.org/10.1371/journal.pmed.1001880.
44. **Bryant JM, Harris SR, Parkhill J, Dawson R, Diacon AH, van Helden P, Pym A, Mahayiddin AA, Chuchottaworn C, Sanne IM, Louw C, Boeree MJ, Hoelscher M, McHugh TD, Bateson AL, Hunt RD, Mwaigwisya S, Wright L, Gillespie SH, Bentley SD.** 2013. Whole-genome sequencing to establish relapse or re-infection with *Mycobacterium tuberculosis*: a retrospective observational study. Lancet Respir Med **1:**786–792. http://dx.doi.org/10.1016/S2213-2600(13)70231-5.
45. **Pankhurst LJ, del Ojo Elias C, Votintseva AA, Walker TM, Cole K, Davies J, Fermont JM, Gascoyne-Binzi DM, Kohl TA, Kong C, Lemaitre N, Niemann S, Paul J, Rogers TR, Roycroft E, Smith EG, Supply P, Tang P, Wilcox MH, Wordsworth S, Wyllie D, Xu L, Crook DW.** 2016. Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. Lancet Respir Med **4:**49–58. http://dx.doi.org/10.1016/S2213-2600(15)00466-X.
46. **Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P, Iqbal Z, Feuerriegel S, Niehaus KE, Wilson DJ, Clifton DA, Kapatai G, Ip CL, Bowden R, Drobniewski FA, Allix-Béguec C, Gaudin C, Parkhill J, Diel R, Supply P, Crook DW, Smith EG, Walker AS, Ismail N, Niemann S, Peto TE.** 2015. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. Lancet Infect Dis **15:**1193–1202. http://dx.doi.org/10.1016/S1473-3099(15)00062-6.