



Prevalence and Characterization of Heterogeneous Variable-Number Tandem-Repeat Clusters Comprising Drug-Susceptible and/or Variable Resistant *Mycobacterium tuberculosis* Complex Isolates in the Netherlands from 2004 to 2016

Inge Roof,^a Rana Jajou,^a Miranda Kamst,^a Arnout Mulder,^a Albert de Neeling,^a Rianne van Hunen,^{a,b} Wim van der Hoek,^a Dick van Soolingen^a

^aNational Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands

^bKNCV Tuberculosis Foundation, The Hague, the Netherlands

ABSTRACT The variable-number tandem-repeat (VNTR) typing method is used to study tuberculosis (TB) transmission. Clustering of *Mycobacterium tuberculosis* isolates with identical VNTR patterns is assumed to reflect recent transmission. Hence, clusters are thought to be homogeneous regarding antibiotic resistance. In practice, however, heterogeneous clusters are also identified. This study investigates the prevalence and characteristics of heterogeneous VNTR clusters and assesses whether isolates in these clusters remain clustered when subjected to whole-genome sequencing (WGS). In the period from 2004 to 2016, 9,072 isolates were included. Demographic and epidemiological linkage data were obtained from the Netherlands Tuberculosis Register. VNTR clusters were defined as homogeneous when isolates shared identical resistance profiles or as heterogeneous if both susceptible and (variable) resistant isolates were found. Multivariate logistic regression analysis was performed to identify factors associated with heterogeneous clustering. Isolates from 2016 were subjected to WGS, and a genetic distance of 12 single nucleotide polymorphisms (SNPs) was used as the cutoff for WGS clustering. In total, 4,661/9,072 (51%) isolates were clustered into 985 different VNTR clusters, of which 217 (22%) were heterogeneous. Patient characteristics associated with heterogeneous clustering were non-Dutch ethnicity (odds ratio [OR], 1.46 [95% confidence interval {CI}, 1.22 to 1.75]), asylum seeker (OR, 1.51 [95% CI, 1.24 to 1.85]), extrapulmonary TB (OR, 1.26 [95% CI, 1.09 to 1.46]), previous TB diagnosis (OR, 1.38 [95% CI, 1.04 to 1.82]), and not being a contact of a TB patient (OR, 1.35 [95% CI, 1.08 to 1.69]). With WGS, 34% of heterogeneous and 78% of homogeneous isolates from 2016 remained clustered. Heterogeneous VNTR clusters are common but seem to be explained by a substantial degree of false clustering by VNTR typing compared to WGS.

KEYWORDS tuberculosis, VNTR, WGS, antibiotic resistance

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, remains a major public health threat since this disease is among the top 10 causes of death worldwide (1). The Netherlands has the status of a low-TB-incidence country, with 5.2 cases per 100,000 inhabitants in 2016 (2). In 2016, 889 TB patients were reported in the Netherlands, which was a 3% increase compared to the incidence in 2015, due mainly to an increased influx of asylum seekers from high-TB-incidence countries (2). In the National Tuberculosis Control Plan, the Netherlands has set a target of reducing TB incidence and transmission by 25% in 5 years. Early diagnosis, adequate treatment, and comprehensive contact investigation are important key elements in the strategy to eliminate TB in low-incidence countries like the Netherlands.

Received 5 June 2018 Returned for modification 9 July 2018 Accepted 28 August 2018

Accepted manuscript posted online 29 August 2018

Citation Roof I, Jajou R, Kamst M, Mulder A, de Neeling A, van Hunen R, van der Hoek W, van Soolingen D. 2018. Prevalence and characterization of heterogeneous variable-number tandem-repeat clusters comprising drug-susceptible and/or variable resistant *Mycobacterium tuberculosis* complex isolates in the Netherlands from 2004 to 2016. *J Clin Microbiol* 56:e00887-18. <https://doi.org/10.1128/JCM.00887-18>.

Editor Betty A. Forbes, Virginia Commonwealth University Medical Center

Copyright © 2018 American Society for Microbiology. All Rights Reserved.

Address correspondence to Inge Roof, inge.roof@rivm.nl.

In this era of elimination, it is important to identify the origin of TB infection, i.e., index case, in all cases. DNA fingerprinting of *M. tuberculosis* complex isolates is a useful tool to assist conventional contact tracing of TB patients and newly infected individuals and the evaluation of TB control interventions (3, 4). In the Netherlands, variable-number tandem-repeat (VNTR) typing has been the standard DNA typing method for *M. tuberculosis* complex isolates since 2009, after retyping of all isolates from the period from 2004 to 2008. VNTR typing is based on the PCR amplification of tandem repeated DNA sequences in 24 defined VNTR loci of the *M. tuberculosis* genome (5), and isolates sharing identical VNTR fingerprint patterns are assigned to the same VNTR cluster. The municipal health services perform cluster investigation to identify epidemiological links between clustered TB cases, and an epidemiological link is confirmed when patients within the same VNTR cluster can be linked in time and place.

Generally, it is assumed that VNTR clustering of isolates reflects recent TB transmission, although some observations refute this (6, 7). Assuming that isolates within a VNTR cluster share the same transmission chain, they are thought to be homogeneous, i.e., sharing an identical antibiotic resistance profile (8, 9). However, in daily practice, heterogeneous VNTR clusters are also observed, comprising both susceptible and resistant isolates or comprising resistant isolates with different resistance profiles. A previous study showed that within a cluster that initially consisted of only susceptible isolates, more than 30% of the cases developed isoniazid resistance over time (10), while in the majority of cases, there was no indication of acquired resistance. This may lead to confusion for the municipal health services during cluster investigations.

The prevalence of heterogeneous TB clusters has to date not been quantified. It is also unclear what causes this phenomenon or whether heterogeneity is due merely to an inaccuracy of VNTR clustering. In 2016, a whole-genome sequencing (WGS) project was initiated in the Netherlands, in which WGS is conducted in parallel with VNTR typing of all *M. tuberculosis* complex isolates, making it possible to investigate whether isolates belonging to heterogeneous VNTR clusters remain clustered according to WGS.

This study aims to investigate (i) the prevalence of heterogeneous VNTR clusters in the Netherlands in a large database, (ii) patient characteristics related to the occurrence of heterogeneous clusters, (iii) the composition of drug resistance profiles within heterogeneous clusters, and (iv) to what extent resistant strains are transmitted to a subsequent patient. Finally, this study determines whether isolates belonging to heterogeneous VNTR clusters remain clustered when WGS is applied.

MATERIALS AND METHODS

Study population. All *M. tuberculosis* complex isolates from the Netherlands in the period from January 2004 to December 2016 were included. *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) isolates, laboratory cross-contaminations, isolates with double alleles in >1 VNTR locus, and isolates from overseas areas were excluded.

Definition of homogeneous and heterogeneous clusters. Isolates sharing an identical 24-locus VNTR pattern were assigned to the same VNTR cluster. Clusters were classified as homogeneous or heterogeneous based on the results of drug susceptibility testing (DST) for the first-line drugs isoniazid (INH), rifampin (RIF), ethambutol (EMB), pyrazinamide (PZA), and streptomycin (STR) (11), as determined with the Mycobacteria Growth Indicator Tube (MGIT) assay (12). Clusters were defined as homogeneous when all isolates within the cluster were susceptible to all first-line antibiotics or when all isolates revealed the same resistance pattern. Heterogeneous clusters comprised susceptible and resistant isolates or resistant isolates with differences in resistance profiles.

DNA fingerprinting. All *M. tuberculosis* complex isolates were subjected to VNTR typing as described previously by Supply et al. (5), and since 2015, an optimized version regarding the costs and turnaround time of this typing method has been applied (13). WGS was performed simultaneously with VNTR typing for all isolates cultured between 1 January 2016 and 31 December 2016, using an Illumina HiSeq sequencer. Fastq.gz files were mapped against version 3 of the H37Rv reference genome (GenBank accession number [AL123456.3](https://www.ncbi.nlm.nih.gov/nuccore/AL123456.3)), using Breseq software 0.28.1, and single nucleotide polymorphisms (SNPs) were detected at a minimum allele frequency of 80%.

To compare the degrees of clustering between VNTR and WGS, homogeneous and heterogeneous clusters with at least two isolates from 2016 were selected. Isolates from 2016 that clustered with isolates from previous years were excluded from this analysis, as no WGS data were available for isolates cultured before 2016. Heterogeneous clusters with at least two isolates from 2016 and known epidemiological information were included in the minimum spanning tree. Isolates were clustered on the basis of WGS when the genetic distance between isolates was a maximum of 12 SNPs (14). All SNPs within a 12-bp

distance; SNPs annotated as PE, PGRS, *esx*, repeat, polyketide, *pks*, or transposase; and isolates with a genome coverage depth of $<20\times$ were excluded.

Data collection. Data on VNTR typing were available in BioNumerics software version 7.6. Patient characteristics and information on epidemiological linkage between patients were obtained from the Netherlands Tuberculosis Register (NTR). This included the characteristics gender, age, ethnicity, residence, pulmonary tuberculosis (PTB)/extrapulmonary tuberculosis (ETB), previous TB episode, and TB risk groups (i.e., contact of a TB patient, immigrant and asylum seeker, homeless, drug or alcohol abuser, prisoner, health care worker, traveler to areas where TB is endemic, HIV status, and being otherwise immunocompromised). Cluster investigation based on VNTR has been performed since 2009 in the Netherlands, and since 2014, detailed data on the source case and the secondary patients have been recorded in the NTR.

Statistical analysis. Demographic characteristics and the proportions of epidemiological links were compared between homogeneous and heterogeneous clusters by using a chi-square test. The Mann-Whitney U test was used to analyze differences in cluster size between homogeneous and heterogeneous clusters. All variables with *P* values of <0.2 in univariate analysis were included as independent variables in the multivariate logistic regression analysis. A *P* value of <0.05 was considered statistically significant. Univariate and multivariate logistic regression analyses were performed by using SPSS version 24.0.0.1. R Statistics version 3.2.2 was used for WGS analysis. A minimum spanning tree was built with Ridom SeqSphere⁺ software version 4.1.9 (Ridom GmbH, Münster, Germany).

RESULTS

Description of heterogeneous VNTR clusters. For 106 clusters comprising 232 isolates, DST data were missing or available for only one isolate within a VNTR cluster and were therefore excluded from further analysis. In total, 9,072 *M. tuberculosis* complex isolates from the period from 2004 to 2016 were analyzed, of which 4,661 isolates (51.4%) clustered into 985 distinct VNTR clusters; 1,695/4,661 (36.4%) isolates belonged to heterogeneous clusters (Fig. 1). The number of heterogeneous clusters was 217 (22%), versus 768 (78%) homogeneous clusters. The large majority ($n = 202$) of heterogeneous clusters included both susceptible and resistant isolates, while the remaining 15 clusters consisted of resistant isolates with different resistance profiles. The numbers of isolates varied from 2 to 111 within homogeneous clusters and from 2 to 135 within heterogeneous clusters, with the majority of homogeneous and heterogeneous clusters containing 2 isolates. The cluster size was significantly larger among the heterogeneous clusters than among the homogeneous clusters ($P < 0.001$).

Of 1,695 isolates in heterogeneous clusters, 1,124 (66.3%) were susceptible to all first-line antibiotics, 450 (26.6%) were resistant to at least one antibiotic, and 121 (7.1%) isolates had no DST data available. In total, 22 different resistance profiles were identified, varying from monoresistance to resistance against all five included antibiotics. The most prevalent resistance profiles in heterogeneous clusters were STR monoresistance (25.8%), INH monoresistance (22.4%), STR and INH resistance (19.6%), and PZA monoresistance (9.1%) (Fig. 2). Of the 41 isolates in heterogeneous clusters with PZA monoresistance, 16 (39%) were *M. bovis*. Almost 90% of heterogeneous clusters contained at least one isolate harboring one of the four most prevalent resistance profiles.

Characteristics associated with heterogeneous clustering. Compared to homogeneous clusters, the heterogeneous clusters comprised a significantly high proportion of patients who were non-Dutch (84.0% versus 76.4%), asylum seekers (17.2% versus 10.5%), diagnosed with ETB (34.7% versus 28.8%), and previously diagnosed with TB (6.5% versus 4.9%). The percentage of patients who were a contact of a TB patient was significantly lower in the heterogeneous clusters (7.8% versus 11.4%) (Table 1). The variables gender, age, homeless, prisoner inmate, and health care worker had a *P* value of <0.2 in univariate analysis and were therefore included in multivariate analysis; the variable "previous TB treatment" was not included due to the relatively small number of patients for whom these data were available ($n = 141$). Multivariate analysis showed that patients belonging to heterogeneous clusters were more likely to be non-Dutch (odds ratio [OR], 1.46 [95% confidence interval [CI], 1.22 to 1.75]), to be asylum seekers (OR, 1.51 [95% CI, 1.24 to 1.85]), to not be a contact of another TB patient (OR, 1.35 [95% CI, 1.08 to 1.69]), to have ETB (OR, 1.26 [95% CI, 1.09 to 1.46]), and to be previously diagnosed with TB (OR, 1.38 [95% CI, 1.04 to 1.82]) (Table 2).

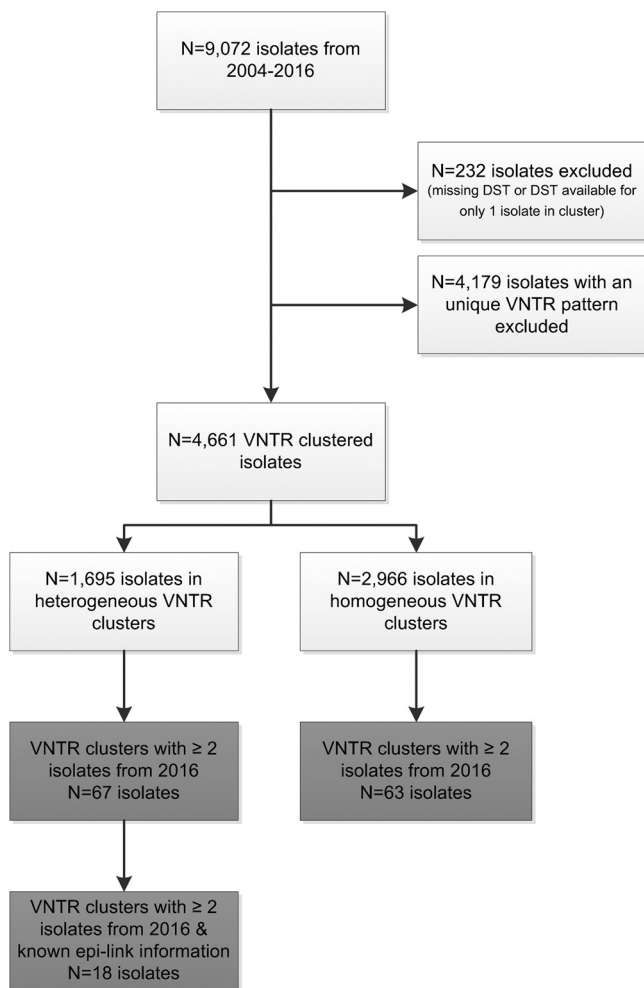


FIG 1 Flowchart of the inclusion and exclusion of isolates from January 2004 to December 2016 for VNTR and WGS analyses. Light gray, inclusion of isolates for VNTR typing; dark gray, inclusion of isolates for WGS.

Cluster investigation of heterogeneous clusters. Data on epidemiological links were available in the NTR for 1,955 (75.6%) of all 2,585 VNTR-clustered isolates between 2009 and 2016. An epidemiological link was confirmed for 16.5% of the isolates within heterogeneous clusters, which was significantly lower than the proportion of confirmed epidemiological links in homogeneous clusters (24.6%; $P < 0.001$). To investigate to what extent resistant strains were transmitted, 24 epidemiological links between patients in heterogeneous clusters with known source case and secondary patient were studied. For 14/24 epidemiological links, there was no transmission of resistant strains, as strains from both the source case and the secondary patient were susceptible to all first-line antibiotics. Among another epidemiologically linked pair, the strain from the source case was STR resistant, and that from the secondary patient was susceptible, meaning that here, there was also no indication of transmission of resistant strains. Eight epidemiological links were confirmed between isolates from a susceptible source case and from a secondary patient that was monoresistant to STR ($n = 7$) or resistant to both INH and RIF ($n = 1$). In the remaining epidemiologically linked pair, an INH- and STR-resistant strain was transmitted from the source case to the subsequent patient.

Clustering on the basis of WGS. VNTR clusters with at least 2 isolates from 2016 were selected for WGS; WGS data were available for 67 isolates belonging to 13 different heterogeneous VNTR clusters and for 63 isolates from 27 different homogeneous VNTR clusters (Fig. 1). The average genome coverage depth of these 130 isolates

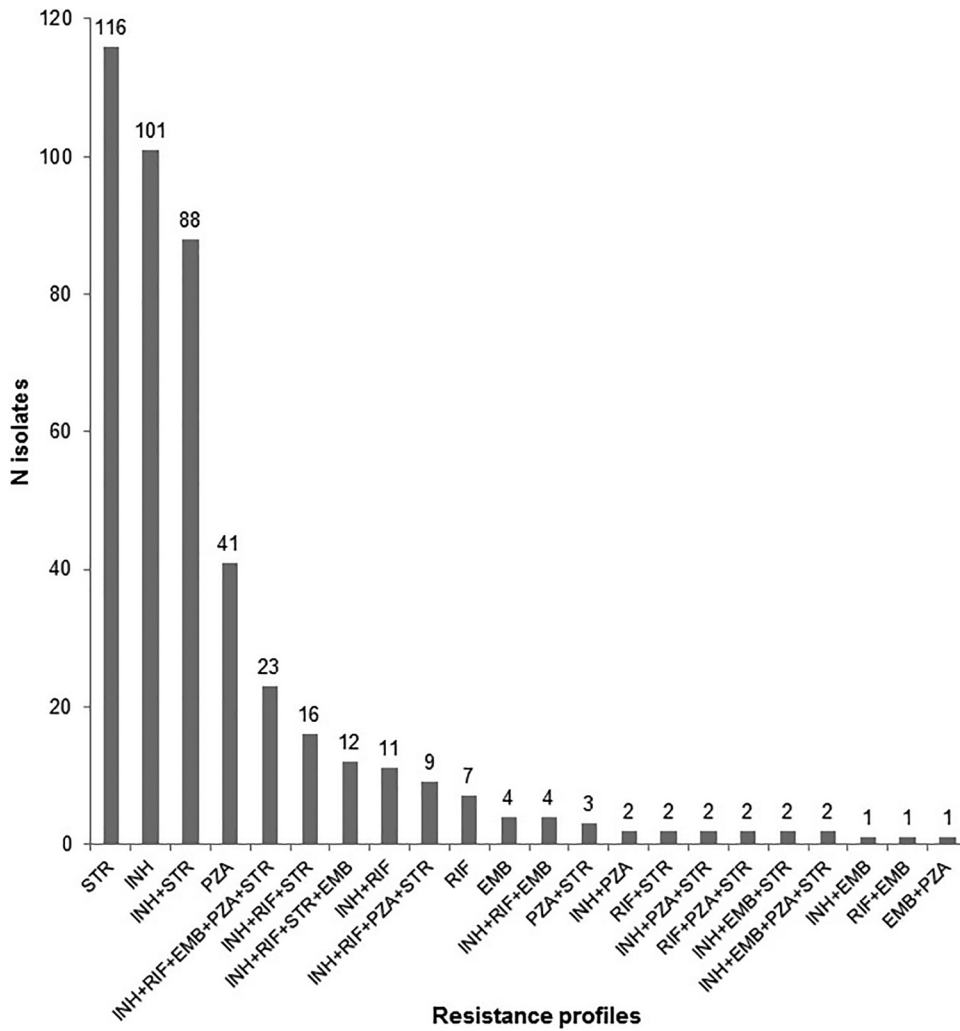


FIG 2 Overview of the variation in resistance profiles of the 450 resistant isolates in the heterogeneous clusters. INH, isoniazid; RIF, rifampin; STR, streptomycin; EMB, ethambutol; PZA, pyrazinamide.

was 113× (range, 35× to 252×). WGS clustered 34% (23/67) of the isolates in heterogeneous clusters, compared to 78% (49/63) of the isolates in homogeneous clusters. A minimum spanning tree was built from five heterogeneous VNTR clusters from 2016 (comprising in total 18 isolates) with available WGS data and epidemiological data. Epidemiological links were confirmed by the Municipal Health Services only for cluster A, comprising two susceptible and two STR-resistant isolates, and this was the only cluster that remained clustered by WGS and, thus, the only one out of five heterogeneous clusters that actually resulted from recent transmission. The remaining isolates from patients belonging to clusters B to E were not clustered by WGS, and this was in line with the epidemiological investigations by municipal health services, as no epidemiological links could be confirmed for these patients (Fig. 3).

DISCUSSION

In this study, the prevalence and characteristics of heterogeneous clusters were investigated in a large study population from 2004 to 2016. Based on VNTR typing, heterogeneity of antibiotic resistance in clusters was found to be common (22%). However, when WGS was applied, a substantial degree of heterogeneous clustering appeared to be falsely clustered by VNTR typing.

The relatively lower degree of clustering by WGS is comparable to data from a previous study in which the clustering of isolates of TB patients by WGS was halved

TABLE 1 Patient characteristics of *M. tuberculosis* complex isolates clustered by VNTR typing in the period from 2004 to 2016^a

Patient characteristic	No. of heterogeneous isolates (n = 1,646) (%)	No. of homogeneous isolates (n = 2,856) (%)	P value
Sex			0.143
Male	1,028 (62.5)	1,847 (64.7)	
Female	617 (37.5)	1,009 (35.3)	
Age group (yr)			0.087
0–24	412 (25.1)	642 (22.5)	
25–44	713 (43.0)	1,278 (44.7)	
45–64	362 (22.1)	610 (21.4)	
65+	159 (9.8)	326 (11.4)	
Ethnicity			<0.001
Dutch	260 (16.0)	657 (23.6)	
Non-Dutch	1,364 (84.0)	2,125 (76.4)	
Residence			0.636
Urban	601 (36.5)	1,063 (37.2)	
Rural	1,045 (63.5)	1,793 (62.8)	
Localization of TB			<0.001
PTB	831 (50.5)	1,646 (57.6)	
ETB	571 (34.7)	822 (28.8)	
PTB + ETB	244 (14.8)	388 (13.6)	
Previous TB diagnosis			0.027
Yes	99 (6.5)	129 (4.9)	
No	1,413 (93.5)	2,492 (95.1)	
Result of previous treatment			0.088
Discontinued	22 (36.7)	19 (23.5)	
Completed	38 (63.3)	62 (76.5)	
Risk group			
Contact of TB patient	129 (7.8)	327 (11.4)	<0.001
Asylum seeker	283 (17.2)	300 (10.5)	<0.001
Homeless	53 (3.2)	73 (2.6)	0.194
Alcohol abuser	33 (2.0)	67 (2.3)	0.455
Drug abuser	60 (3.6)	98 (3.4)	0.707
Prisoner	42 (2.6)	93 (3.3)	0.183
Health care worker	9 (0.5)	29 (1.0)	0.103
Traveler to area of endemicity	29 (1.8)	61 (2.1)	0.389
Immunocompromised			
HIV positive	79 (10.8)	103 (9.2)	0.257
Diabetes	60 (3.6)	120 (4.2)	0.359
Malignancy	38 (2.3)	62 (2.2)	0.763
Renal disease	19 (1.2)	35 (1.2)	0.833
Immunosuppressive use	31 (1.8)	67 (2.3)	0.307

^aTB, tuberculosis; PTB, pulmonary tuberculosis; ETB, extrapulmonary tuberculosis; HIV, human immunodeficiency virus.

compared to that by VNTR typing (15). In our study, only one of the five heterogeneous VNTR clusters remained clustered by WGS. This finding is in line with data from a previous study that concluded that a drug-susceptible isolate and a multidrug-resistant isolate with identical VNTR DNA fingerprints harbored a considerable degree of genomic diversity, with a genetic distance of 130 SNPs and one large deletion between the two isolates (16). However, in that study, isolates were clustered when they shared 23 out of 24 identical loci in the VNTR pattern, instead of the applied identical 24-locus pattern in our study. The only heterogeneous cluster in our study that remained clustered by WGS was also epidemiologically confirmed by the municipal health services. The absence of confirmed epidemiological links and the large SNP distances in the remaining heterogeneous clusters support the hypothesis of false clustering by

TABLE 2 Multivariate analysis of factors associated with heterogeneous clustering^a

Parameter	Multivariate OR (95% CI)	P value
Ethnicity		
Dutch	1	
Non-Dutch	1.46 (1.22–1.75)	<0.001
Localization of TB		
PTB	1	
ETB	1.26 (1.09–1.46)	0.002
PTB + ETB	1.17 (0.96–1.42)	0.100
Previous TB diagnosis		
Yes	1.38 (1.04–1.82)	0.024
No	1	
Contact of TB patient		
Yes	1	
No	1.35 (1.08–1.69)	0.009
Asylum seeker		
Yes	1.51 (1.24–1.85)	<0.001
No	1	

^aTB, tuberculosis; PTB, pulmonary tuberculosis; ETB, extrapulmonary tuberculosis; OR, odds ratio; CI, confidence interval.

VNTR. A similar outcome was previously described by Jajou et al., who concluded that WGS is a more discriminative tool to predict epidemiological links than VNTR typing, thereby improving the efficiency of cluster investigations (15).

Some variables in this study appeared to be significantly associated with cluster heterogeneity. First, cases in heterogeneous clusters were more likely to be of non-Dutch ethnicity, which can be explained by the fact that most of the drug resistance (about 80%) in the Netherlands is found among foreign-born patients (17, 18). The variables non-TB contact, asylum seeker, and ETB might be indicative of the import of resistant strains by foreign-born patients from high-TB-incidence countries rather than recent transmission in the Netherlands. This suggestion is supported by the low proportion (16.5%) of confirmed epidemiological links found between isolates in heterogeneous clusters. In addition, a previous study showed that TB among immigrants in Hamburg, Germany, was often not due to recent transmission of TB (19). In general, clustering of isolates from asylum seekers in terms of recent transmission should be interpreted with caution, because the isolates from foreign-born patients can belong to predominating strains with a high degree of genetic homogeneity in their home countries (20–22) and might thereby be incorrectly clustered by VNTR typing. This hypothesis is supported by a recent study in which a large VNTR cluster of mostly asylum seekers was divided into several subclusters when analyzed by WGS (23). Furthermore, the association of a previous TB diagnosis with heterogeneity may indicate endogenous reactivation of TB in foreign-born patients and asylum seekers.

Other explanations for the occurrence of heterogeneous clusters can be discussed, for example, the existence of heteroresistance within *M. tuberculosis* isolates, i.e., the presence of drug-susceptible and drug-resistant bacteria in the same isolate. Recently, a study investigating the prevalence of heteroresistance by WGS showed that for around 50% of the isolates tested, heteroresistance was present in at least one resistance-associated genomic locus (24). Moreover, the existence of resistant minority subpopulations in *M. tuberculosis* was described previously (25). Using a targeted deep sequencing technique, Metcalfe et al. described the presence of microheteroresistance (<5% drug-resistant subpopulation of the total population) in phenotypically resistant isolates with no resistance-associated mutations determined by Sanger sequencing (25). In our study, one of the heterogeneous clusters remained clustered by WGS despite differences in STR resistance in the respective isolates, which can be explained by the existence of minority resistant subpopulations at lower allele frequencies than our current standard of 80%.

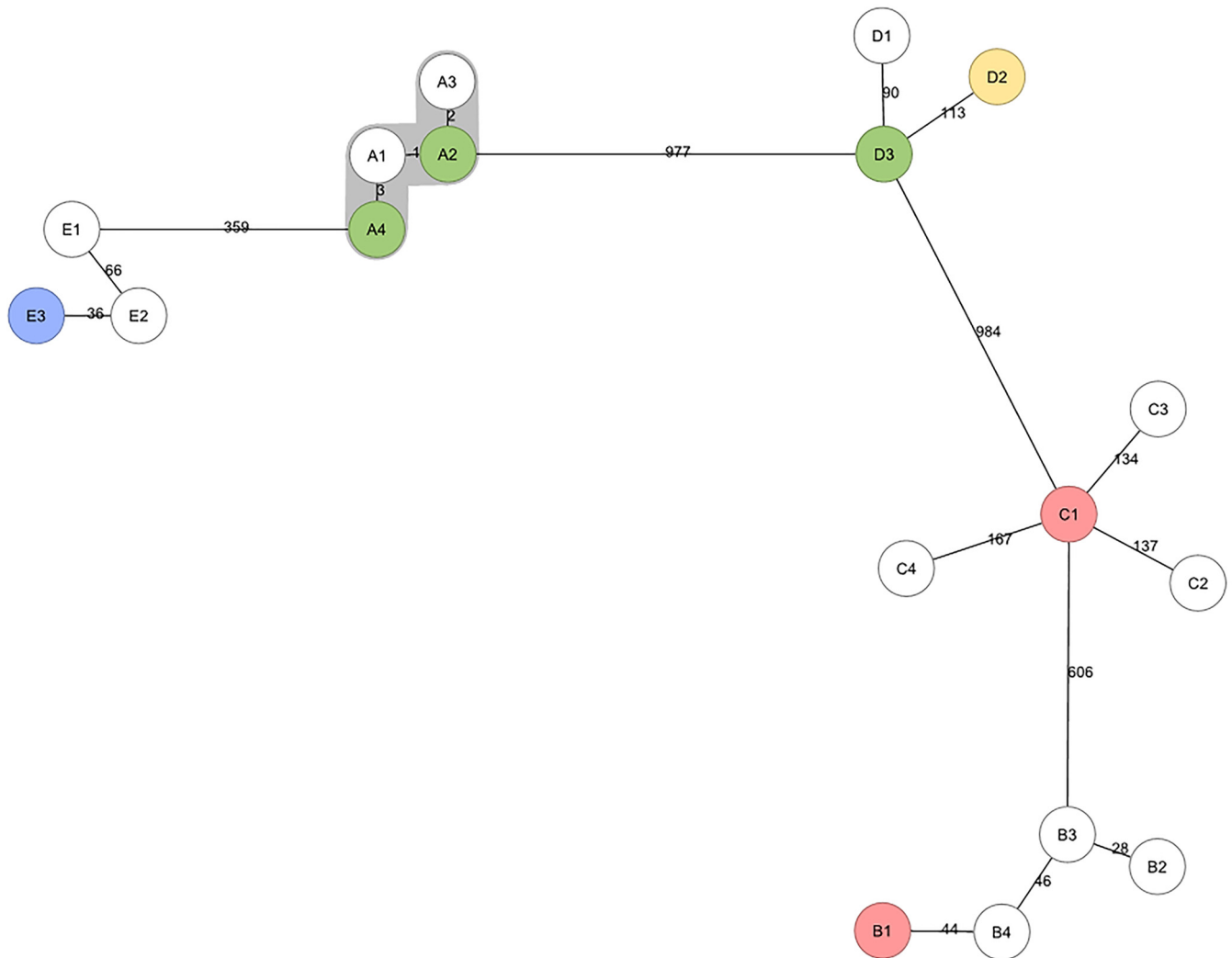


FIG 3 Minimum spanning tree of five heterogeneous VNTR clusters (clusters A to E) analyzed by WGS. Clustering by WGS is represented in gray, and the genetic distances in SNPs to the nearest isolate are shown on the branches. White, susceptible samples; green, streptomycin-resistant samples; red, isoniazid-resistant samples; blue, isoniazid- and streptomycin-resistant samples; yellow, rifampin-, isoniazid-, ethambutol-, and streptomycin-resistant samples.

Finally, the bacterial diversity present in the lung of individual TB patients might explain the heterogeneity of resistance profiles within VNTR clusters. Results from previous studies suggest that sputum samples may not represent the complete strain diversity present in a patient's lung and that DST on isolates from patient sputa may not represent the susceptibility of all bacterial populations within the lung (26–28). Moreover, if sputum samples do not contain the real diversity of the bacterial population in the respiratory tract, it might be that only a subpopulation is transmitted from patient to patient.

The most prevalent resistance profiles observed in heterogeneous clusters were INH and STR monoresistance. This is in accordance with the overall TB drug resistance pattern previously described in the Netherlands using data from 1993 to 2011 (17). In the same study, a low overall occurrence of PZA resistance was reported for the Netherlands, which was the fourth most prevalent resistance profile among isolates in heterogeneous clusters in our study. The relatively high frequency of PZA monoresistance is partially because 39% of these isolates were *M. bovis* strains, which are known to be intrinsically resistant to PZA (29).

The rate of transmission of resistant strains to a secondary patient was found to be relatively low in our study. Among the majority of epidemiologically linked pairs, the

strains from the source case and secondary patient were susceptible to first-line antibiotics. In a few instances, an epidemiological link between a susceptible source patient and a resistant secondary patient was confirmed, which can indicate transmission of resistant strains; however, we could not completely rule out that resistance was acquired over time.

This study is based on a comprehensive data collection process with genotyping of all *M. tuberculosis* complex isolates at the Dutch national tuberculosis reference laboratory, generating comparable and generalizable results. Despite this, the study has some limitations. First, a minority of clusters could not be classified as homo- or heterogeneous due to missing DST data and was excluded from the analysis. As this was only a small proportion of the total number of isolates included in the study, we believe that this did not have an impact on the results. Second, epidemiological cluster investigation has been performed based on VNTR typing since 2009; however, detailed information on the source cases and secondary patients has been recorded in the NTR only since 2014, meaning that only a subset was available for analysis and might therefore not be representative. Furthermore, WGS data were available for only a subset of isolates, in which most of the isolates were susceptible or resistant to STR only, resulting in a low variability of included isolates compared to the complete study population.

In conclusion, 22% of the VNTR clusters were heterogeneous; however, the low level of agreement of heterogeneous clustering between VNTR typing and WGS suggests the inaccuracy of VNTR typing to correctly cluster isolates belonging to the same transmission chain, and this was also represented in the low extent of transmission of resistant strains between epidemiologically linked cases. With the increasing availability and application of WGS, differentiation of *M. tuberculosis* complex isolates can be performed with a higher degree of discriminatory power, and false clustering by VNTR typing can be prevented.

REFERENCES

- World Health Organization. 2018. Tuberculosis key facts. World Health Organization, Geneva, Switzerland. <http://www.who.int/news-room/fact-sheets/detail/tuberculosis>. Accessed 5 August 2018.
- Slump E, Bregman IM, Erkens CGM, van Hunen R, Schimmel HJ, van Soolingen D, de Vries G. 2017. Tuberculose in Nederland 2016: surveillancerapport inclusief rapportage monitoring van interventies. Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven, the Netherlands. <https://doi.org/10.21945/rivm-2017-0160>.
- Barnes PF, Cave MD. 2003. Molecular epidemiology of tuberculosis. *N Engl J Med* 349:1149–1156. <https://doi.org/10.1056/NEJMra021964>.
- de Vries G, van Hest RA, Burdo CC, van Soolingen D, Richardus JH. 2009. A Mycobacterium tuberculosis cluster demonstrating the use of genotyping in urban tuberculosis control. *BMC Infect Dis* 9:151. <https://doi.org/10.1186/1471-2334-9-151>.
- Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S, Willery E, Savine E, de Haas P, van Deutekom H, Roring S, Bifani P, Kurepina N, Kreiswirth B, Sola C, Rastogi N, Vatin V, Gutierrez MC, Fauville M, Niemann S, Skuce R, Kremer K, Locht C, van Soolingen D. 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. *J Clin Microbiol* 44:4498–4510. <https://doi.org/10.1128/JCM.01392-06>.
- de Vries G, Baars HW, Sebek MM, van Hest NA, Richardus JH. 2008. Transmission classification model to determine place and time of infection of tuberculosis cases in an urban area. *J Clin Microbiol* 46:3924–3930. <https://doi.org/10.1128/JCM.00793-08>.
- Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC, Schecter GF, Daley CL, Schoolnik GK. 1994. The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. *N Engl J Med* 330:1703–1709. <https://doi.org/10.1056/NEJM199406163302402>.
- Dolzani L, Rosato M, Sartori B, Banfi E, Lagatolla C, Predominato M, Fabris C, Tonin E, Gombac F, Monti-Bragadin C. 2004. Mycobacterium tuberculosis isolates belonging to katG gyrA group 2 are associated with clustered cases of tuberculosis in Italian patients. *J Med Microbiol* 53:155–159. <https://doi.org/10.1099/jmm.0.05471-0>.
- Chaidir L, Sengstake S, de Beer J, Krismawati H, Lestari FD, Ayawaila S, van Soolingen D, Anthony R, van Crevel R, Alisjahbana B. 2015. Mycobacterium tuberculosis genotypic drug resistance patterns and clustering in Jayapura, Papua, Indonesia. *Int J Tuberc Lung Dis* 19:428–433. <https://doi.org/10.5588/ijtld.14.0350>.
- Perri BR, Proops D, Moonan PK, Munsiff SS, Kreiswirth BN, Goranson C, Ahuja SD. 2011. Mycobacterium tuberculosis cluster with developing drug resistance, New York, New York, USA, 2003–2009. *Emerg Infect Dis* 17:372–378. <https://doi.org/10.3201/eid1703.101002>.
- World Health Organization. 2010. Treatment of tuberculosis: guidelines, 4th ed. World Health Organization, Geneva, Switzerland.
- Palomino JC, Traore H, Fissette K, Portaels F. 1999. Evaluation of Mycobacteria Growth Indicator Tube (MGIT) for drug susceptibility testing of Mycobacterium tuberculosis. *Int J Tuberc Lung Dis* 3:344–348.
- de Beer JL, Akkerman OW, Schürch AC, Mulder A, van der Werf TS, van der Zanden AGM, van Ingen J, van Soolingen D. 2014. Optimization of standard in-house 24-locus variable-number tandem-repeat typing for Mycobacterium tuberculosis and its direct application to clinical material. *J Clin Microbiol* 52:1338–1342. <https://doi.org/10.1128/JCM.03436-13>.
- Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dediccoat 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. [https://doi.org/10.1016/S1473-3099\(12\)70277-3](https://doi.org/10.1016/S1473-3099(12)70277-3).
- Jajou R, de Neeling A, van Hunen R, de Vries G, Schimmel H, Mulder A, Anthony R, van der Hoek W, van Soolingen D. 2018. Epidemiological links between tuberculosis cases identified twice as efficiently by whole genome sequencing than conventional molecular typing: a population-based study. *PLoS One* 13:e0195413. <https://doi.org/10.1371/journal.pone.0195413>.

16. Niemann S, Koser CU, Gagneux S, Plinke C, Homolka S, Bignell H, Carter RJ, Cheetham RK, Cox A, Gormley NA, Kokko-Gonzales P, Murray LJ, Rigatti R, Smith VP, Arends FP, Cox HS, Smith G, Archer JA. 2009. Genomic diversity among drug sensitive and multidrug resistant isolates of *Mycobacterium tuberculosis* with identical DNA fingerprints. *PLoS One* 4:e7407. <https://doi.org/10.1371/journal.pone.0007407>.
17. Ruesen C, van Gageldonk-Lafeber AB, de Vries G, Erkens CG, van Rest J, Korthals Altes H, de Neeling H, Kamst M, van Soolingen D. 2014. Extent and origin of resistance to antituberculosis drugs in the Netherlands, 1993 to 2011. *Euro Surveill* 19;pii=20738. <https://doi.org/10.2807/1560-7917.ES2014.19.11.20738>.
18. Lambregts-van Weezenbeek CS, Jansen HM, Veen J, Nagelkerke NJ, Sebek MM, van Soolingen D. 1998. Origin and management of primary and acquired drug-resistant tuberculosis in The Netherlands: the truth behind the rates. *Int J Tuberc Lung Dis* 2:296–302.
19. Diel R, Rusch-Gerdes S, Niemann S. 2004. Molecular epidemiology of tuberculosis among immigrants in Hamburg, Germany. *J Clin Microbiol* 42:2952–2960. <https://doi.org/10.1128/JCM.42.7.2952-2960.2004>.
20. Cardoso Oelemann M, Gomes HM, Willery E, Possuelo L, Batista Lima KV, Allix-Beguec C, Locht C, Goguet de la Salmoniere YO, Gutierrez MC, Suffys P, Supply P. 2011. The forest behind the tree: phylogenetic exploration of a dominant *Mycobacterium tuberculosis* strain lineage from a high tuberculosis burden country. *PLoS One* 6:e18256. <https://doi.org/10.1371/journal.pone.0018256>.
21. Chihota V, Apers L, Mungofa S, Kasongo W, Nyoni IM, Tembwe R, Mbulo G, Tembo M, Streicher EM, van der Spuy GD, Victor TC, van Helden P, Warren RM. 2007. Predominance of a single genotype of *Mycobacterium tuberculosis* in regions of Southern Africa. *Int J Tuberc Lung Dis* 11: 311–318.
22. van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, Portaels F, Qing HZ, Enkhsaikan D, Nymadawa P, van Embden JD. 1995. Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. *J Clin Microbiol* 33:3234–3238.
23. Jajou R, de Neeling A, Rasmussen EM, Norman A, Mulder A, van Hunen R, de Vries G, Haddad W, Anthony R, Lillebaek T, van der Hoek W, van Soolingen D. 2018. A predominant variable-number tandem-repeat cluster of *Mycobacterium tuberculosis* isolates among asylum seekers in the Netherlands and Denmark, deciphered by whole-genome sequencing. *J Clin Microbiol* 56:e01100-17. <https://doi.org/10.1128/JCM.01100-17>.
24. Operario DJ, Koepfel AF, Turner SD, Bao Y, Pholwat S, Banu S, Foongladda S, Mpagama S, Gratz J, Ogarkov O, Zhadova S, Heysell SK, Hout ER. 2017. Prevalence and extent of heteroresistance by next generation sequencing of multidrug-resistant tuberculosis. *PLoS One* 12:e0176522. <https://doi.org/10.1371/journal.pone.0176522>.
25. Metcalfe JZ, Streicher E, Theron G, Colman RE, Allender C, Lemmer D, Warren R, Engelthaler DM. 14 June 2017. Cryptic micro-heteroresistance explains *M. tuberculosis* phenotypic resistance. *Am J Respir Crit Care Med* <https://doi.org/10.1164/rccm.201703-0556OC>.
26. Ford C, Yusim K, Ioerger T, Feng S, Chase M, Greene M, Korber B, Fortune S. 2012. *Mycobacterium tuberculosis*—heterogeneity revealed through whole genome sequencing. *Tuberculosis (Edinb)* 92:194–201. <https://doi.org/10.1016/j.tube.2011.11.003>.
27. Kaplan G, Post FA, Moreira AL, Wainwright H, Kreiswirth BN, Tanverdi M, Mathema B, Ramaswamy SV, Walther G, Steyn LM, Barry CE, III, Bekker L-G. 2003. *Mycobacterium tuberculosis* growth at the cavity surface: a microenvironment with failed immunity. *Infect Immun* 71:7099–7108. <https://doi.org/10.1128/IAI.71.12.7099-7108.2003>.
28. Shamputa IC, Jugheli L, Sadradze N, Willery E, Portaels F, Supply P, Rigouts L. 2006. Mixed infection and clonal representativeness of a single sputum sample in tuberculosis patients from a penitentiary hospital in Georgia. *Respir Res* 7:99. <https://doi.org/10.1186/1465-9921-7-99>.
29. Scorpio A, Zhang Y. 1996. Mutations in *pncA*, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in *tubercle bacillus*. *Nat Med* 2:662–667. <https://doi.org/10.1038/nm0696-662>.