JCM
Journals.ASM.org

First Worldwide Proficiency Study on Variable-Number Tandem-Repeat Typing of *Mycobacterium tuberculosis* Complex Strains

Jessica L. de Beer,[a] Kristin Kremer,[b] Csaba Ködmön,[c] Philip Supply,[d,e] Dick van Soolingen,[a,f] and the Global Network for the Molecular Surveillance of Tuberculosis 2009

National Tuberculosis Reference Laboratory, Laboratory for Infectious Diseases and Perinatal Screening, Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, Netherlands[a]; Tuberculosis and M/XDR-TB Programme, Division of Communicable Diseases, Health Security & Environment, World Health Organization, Regional Office for Europe, Copenhagen, Denmark[b]; Surveillance Unit, Tuberculosis Programme, European Centre for Disease Prevention and Control, Stockholm, Sweden[c]; Center for Infection and Immunity of Lille, INSERM U1019-CNRS UMR 8204, Université Lille Nord de France and Institut Pasteur de Lille, Lille, France[d]; GenoScreen, Lille, France[e]; and Radboud University Nijmegen Medical Centre, Dekkerswald, and Medical Microbiology, Nijmegen, Netherlands[f]

Although variable-number tandem-repeat (VNTR) typing has gained recognition as the new standard for the DNA fingerprinting of *Mycobacterium tuberculosis* complex (MTBC) isolates, external quality control programs have not yet been developed. Therefore, we organized the first multicenter proficiency study on 24-locus VNTR typing. Sets of 30 DNAs of MTBC strains, including 10 duplicate DNA samples, were distributed among 37 participating laboratories in 30 different countries worldwide. Twenty-four laboratories used an in-house-adapted method with fragment sizing by gel electrophoresis or an automated DNA analyzer, nine laboratories used a commercially available kit, and four laboratories used other methods. The intra- and interlaboratory reproducibilities of VNTR typing varied from 0% to 100%, with averages of 72% and 60%, respectively. Twenty of the 37 laboratories failed to amplify particular VNTR loci; if these missing results were ignored, the number of laboratories with 100% interlaboratory reproducibility increased from 1 to 5. The average interlaboratory reproducibility of VNTR typing using a commercial kit was better (88%) than that of in-house-adapted methods using a DNA analyzer (70%) or gel electrophoresis (50%). Eleven laboratories using in-house-adapted manual typing or automated typing scored inter- and intralaboratory reproducibilities of 80% or higher, which suggests that these approaches can be used in a reliable way. In conclusion, this first multicenter study has documented the worldwide quality of VNTR typing of MTBC strains and highlights the importance of international quality control to improve genotyping in the future.

Since the early 1990s, the molecular typing of *Mycobacterium tuberculosis* complex (MTBC) isolates has revealed important novel insights into the epidemiology of tuberculosis (TB). In a previous study on the intralaboratory reproducibility and discriminatory power of typing methods for MTBC performed in 1999, the initial and widely applied standardized IS*6110* restriction fragment length polymorphism (RFLP) typing technique (29) appeared to be highly reproducible and more discriminatory than other typing techniques existing at that time (17). Therefore, this technique was selected as the standard for the determination of the molecular epidemiology of tuberculosis (17, 29, 30) and, hence, for the study of the international transmission of multidrug-resistant TB (MDR-TB) and extensively drug-resistant TB (XDR-TB) in Europe (6, 7). However, IS*6110* RFLP analysis is technically demanding and time-consuming, and interlaboratory comparisons of the generated banding patterns remained a challenge.

In recent years, variable-number tandem-repeat (VNTR) typing has become the new standard for the typing of MTBC strains (26), because of its short turnaround time, simplified comparisons of digital results, and applicability to small amounts of DNA. VNTR typing of MTBC isolates is a multiple-locus variant-repeat analysis (MLVA) (20) based on the variation in the numbers of DNA tandem repeats at multiple genomic loci. In its simplest form, each target region is individually amplified by using primers annealing to the flanks of the repeat-containing region. The amplicon sizes are determined by electrophoresis on agarose gels in comparison to molecular size markers to deduce the number of tandem-repeat units present at each locus (11, 19, 21, 22, 25, 28). In an automated, high-throughput version, the target loci are amplified in multiplex PCRs by using fluorescently labeled primers, and the amplicon sizes are determined by using a capillary electrophoresis (CE)-based DNA analyzer (1, 4, 27). Although this approach uses more expensive equipment and reagents, it has been found to be more efficient and more accurate for the sizing of amplicons (15) than the manual method. Another, less frequently used method for amplicon size determinations is based on nondenaturing high-performance liquid chromatography (nondHPLC) (8). In all instances, the final result of VNTR typing is a numerical code that corresponds to the numbers of repeats in the target locus and serves as a fingerprint of the respective MTBC isolate. This simple format facilitates the efficient and reliable exchange and comparison of genotyping results both locally and worldwide.

Although MLVA methods have been developed for many mi-

Received 25 March 2011 Returned for modification 5 June 2011
Accepted 11 November 2011
Published ahead of print 14 December 2011
Address correspondence to Dick van Soolingen, Dick.van.Soolingen@rivm.nl.

doi:10.1128/JCM.00607-11

croorganisms in recent years (20), studies regarding the reproducibility of these methods are limited. To our knowledge, the only multilaboratory validation study performed was for MLVA of Shiga toxin-producing *Escherichia coli* O157 (14). For MTBC isolates, two small-scale initial quality control studies were previously performed at the Institut Pasteur, Lille, France (IPL), and at the U.S. Centers for Disease Control and Prevention (CDC), Atlanta, GA, demonstrating 100% intra- and interlaboratory reproducibilities for both agarose gel and automated VNTR typing (24, 27). However, no international proficiency study has been organized since the standardization of the 24-locus VNTR typing method for MTBC isolates in 2006 (26).

In 2009, the European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden, initiated a project, currently outsourced to the Tuberculosis Reference Laboratory at the National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands, on the molecular surveillance of the international transmission of MDR-TB and XDR-TB in the European Union. The participants of the project agreed to use 24-locus VNTR typing as the standard for the typing of MTBC isolates. To test the international intra- and interlaboratory reproducibilities of VNTR typing, panels of 30 MTBC DNAs, including 10 blinded duplicate DNAs, were sent to 18 of the European laboratories participating in the MDR-TB molecular surveillance project and to 19 other laboratories around the world. This report describes the results of this proficiency study on the performance of VNTR typing and provides recommendations to improve the reproducibility and comparability of the method.

MATERIALS AND METHODS

Study design. A test panel consisting of *M. tuberculosis* complex DNAs was prepared at the RIVM and distributed among the 37 participating laboratories. Laboratories were asked to perform VNTR typing as they were accustomed to and to report the VNTR typing results to the RIVM in a numerical format. The RIVM compared the results from each laboratory to those of the reference and reported the individual results to each of the laboratories. In a questionnaire, laboratories were asked to supply details on the VNTR typing method used.

Participants. Thirty-seven laboratories in 30 countries participated in this quality control study. Twenty-three laboratories were located in Europe, including 18 participating within the framework of the ECDC-funded Molecular Surveillance of M/XDR-TB in the European Union project, outsourced to the RIVM. The remaining laboratories were located in Asia ($n = 4$), Oceania ($n = 1$), Africa ($n = 1$), and the Americas ($n = 8$) (including North America [$n = 5$]).

Test panel. The quality control panel was comprised of 30 DNA samples of 20 MTBC strains: 14 *M. tuberculosis* strains (including H37Rv) and 1 strain each of *Mycobacterium bovis*, *M. bovis* BCG, *Mycobacterium microti*, *Mycobacterium canettii*, *Mycobacterium africanum*, and *Mycobacterium pinnipedii*. The selected strains have been extensively typed in previous studies on mycobacterial interspersed repetitive-unit (MIRU)-VNTR typing (27) and on comparisons of DNA typing methods (16, 17) and represent the phylogenetic variety of *M. tuberculosis* complex isolates and allelic ranges of VNTR markers (27). Ten blinded duplicate DNA samples of MTBC strains were included: seven *M. tuberculosis* strains (including H37Rv), one *M. microti* strain, one *M. africanum* strain, and one *M. bovis* BCG strain.

DNA was extracted as previously described (17). The DNA concentration was measured with the Nanodrop-1000 system (Thermo Scientific, Wilmington, DE). For the preparation of all 37 test panels, the DNA samples were diluted to a final concentration of 100 ng/μl and divided over 37 batches. The test panels were stored at −20°C and shipped at room temperature by courier service. Laboratories were asked to dilute the samples by adding molecular-grade water to achieve a final concentration of 10 ng/μl.

Reference typing results. The reference results were obtained at the RIVM by applying both an in-house VNTR typing method and a commercial kit for VNTR typing (GenoScreen, Lille, France). The in-house VNTR method is based on the protocol of the MIRU-VNTR typing manual (publicly available from the MIRU-VNTR*plus* website [www.miru-vntrplus.org]) as originally described by Supply et al. (26), with the following modifications: the amount of DNA polymerase was increased to 0.75 units per multiplex PCR, and the initial concentration of labeled primers for locus 2165 and locus 2163b was increased to 8 μM. For both methods, the amplicon sizes were determined by using the automated ABI 3730 DNA analyzer (Applied Biosystems, CA). The results of both methods were in complete agreement and were also in complete agreement with results previously obtained for these samples (16, 17, 27). Finally, one of the strains included in the panel was *M. tuberculosis* control strain H37Rv, the whole genome of which was sequenced previously (3), and the genome-derived numbers of repeats matched exactly with those found by VNTR typing at the RIVM.

Data analysis. The results of the participating laboratories were compared to those of the reference laboratory, the RIVM, to determine the inter- and intralaboratory reproducibilities. To determine the error rate per VNTR locus, systematic errors were excluded because they would have a high impact on this error rate. The assumption for the cutoff value for systematic errors was >20 errors for an individual VNTR locus with identical differences compared to the reference results.

The conditional chi-square test was used to test the homogeneity of the distribution of errors within a group of participants using the same VNTR method. Assuming that the number of errors in a laboratory has a Poisson distribution and that this distribution is similar for all laboratories, the conditional distribution of the errors among laboratories given the total number of errors is multinomial, with the same probability for each participating laboratory. Thus, the null hypothesis can be tested by means of the conditional, exact chi-square test for goodness of fit. A mixed-effects Poisson model was used to extricate the potential differences between the groups of methods from the evident effect of the laboratory on the number of errors.

RESULTS

The DNA test panel was subjected to VNTR typing by 37 laboratories in 30 countries: 34 laboratories analyzed 24 loci (26), 2 laboratories analyzed a subset of 15 loci (26), and 1 laboratory typed 14 loci. The laboratories could be divided into four groups on the basis of the VNTR typing method used: (i) 15 laboratories used an in-house PCR method and determined amplicon sizes by using agarose gel electrophoresis, (ii) 9 laboratories used an in-house PCR method and determined amplicon sizes by using a CE DNA analyzer, (iii) 9 laboratories used a commercial kit and a CE DNA analyzer, and (iv) 4 laboratories used other methods (Table 1). The alternative methods used by group 4 included the use of PCR reagents from Tiangen Biotech (Beijing, China), the performance of VNTR typing based on methods described previously by Murase et al. (23), adaptations of the VNTR typing method based on multiple previously reported methods (2, 12, 18), and the use of the Transenomic Wave high-performance liquid chromatography system (Transenomic Inc., Omaha, NE) (8).

The number of missing results as a consequence of an amplification failure gives an indication of the typeability of the different methods applied. The percentages of complete VNTR patterns reported by the laboratories did not differ significantly between the various methods used; these percentages were 97% for laboratories using an agarose gel electrophoresis-based in-house method, 92% for laboratories using a DNA analyzer-based in-

TABLE 1 Methods used to perform 24-locus VNTR typing of *M. tuberculosis* complex and average inter- and intralaboratory reproducibilities[a]

Method for VNTR typing	No. of participating laboratories	Avg % reproducibility (range)		Typeability (%)
		Interlaboratory	Intralaboratory	
In-house PCR[b]				
Sizing by gel electrophoresis	15	50 (0–77)	74 (0–100)	97
Sizing by DNA analyzer	9	70 (0–100)	73 (10–100)	92
GenoScreen commercial kit				
Sizing by DNA analyzer	9	88 (50–100)	82 (10–100)	96
Other	4	5 (0–20)	38 (0–60)	79
Total for all methods	37	60	72	93

[a] Data include missing results as mistakes. The typeability of each method is the average percentage of obtained complete VNTR pattern results for the 30 samples of the panel.
[b] Based on or adapted from methods described previously by Supply et al. (26).

house method, and 96% for laboratories using the commercial kit and a DNA analyzer (Table 1). In contrast, the percentage of complete VNTR patterns was lower for those laboratories using other methods, 79%.

If the typeability is calculated at the locus level, i.e., the number of results obtained as a percentage of the total number of loci typed in the VNTR typing of the complete set of samples, the locus typeability score is obtained. The typeability score for the laboratories that used an in-house PCR method and detected the amplicon sizes by gel electrophoresis was 99.8%. For the laboratories in which amplicon sizes were analyzed on a DNA analyzer, the in-house approach and the commercial approach yielded nearly the same results, 99.6% and 99.5%, respectively.

Interlaboratory reproducibility. To determine the interlaboratory reproducibility of VNTR typing, the percentages of the 30 DNA samples for which the numbers of repeats were correctly scored for all loci typed in comparison to the reference laboratory results were calculated for each of the participating laboratories (Fig. 1). Apart from the reference laboratory, only one laboratory yielded a complete and fully correct 24-locus VNTR typing result for all 30 samples and thus had inter- and intralaboratory reproducibilities of 100%. Eleven laboratories had an interlaboratory reproducibility score of 90% or higher, six had a score between 80% and 90%, five had a score between 70% and 80%, and four had a score between 20% and 60%. Nine of the 37 laboratories reported incorrect results for all 30 samples.

The level of interlaboratory reproducibility was in part decreased by the occurrence of incomplete VNTR patterns, i.e., missing results for one or more loci in at least one sample.

Twenty laboratories reported such incomplete VNTR patterns, with missing results for 1 to 9 loci of the set of 24 for one or more samples of the panel. If missing results were not interpreted as incorrect results, the individual interlaboratory reproducibility of 11 laboratories increased up to 40% (mean, 12%; range, 3% to 40%), and the total number of laboratories with a 100% interlaboratory reproducibility increased from 1 to 5 (Fig. 1).

Intralaboratory reproducibility. The intralaboratory reproducibility of VNTR typing was determined for each laboratory by comparing the patterns reported for the 10 duplicate DNA samples (Fig. 2). Incomplete patterns were included in the calculation; i.e., missing results for a particular locus for both samples of a duplicate contributed positively to the intralaboratory reproducibility percentage, while a missing result for only one sample of a duplicate contributed negatively to the intralaboratory reproducibility score. Eleven laboratories (reference excluded) had identical typing results for all 10 duplicated samples and thus had an intralaboratory reproducibility of 100%. Twelve laboratories had an intralaboratory reproducibility of 80% to 90%, and four had an intralaboratory reproducibility of 60% to 70%. Seven laboratories reported identical typing results for less than 5 (<50%) of the

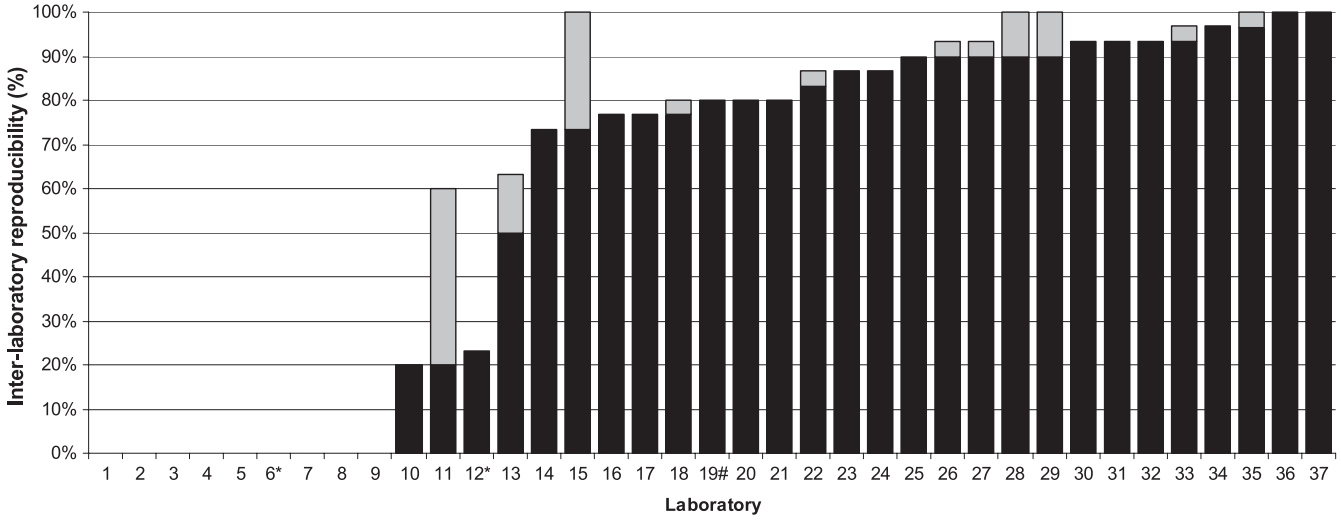

FIG 1 Interlaboratory reproducibility of VNTR typing of 30 *M. tuberculosis* complex DNA samples scored by 36 participating laboratories and the reference laboratory. The black bars represent the reproducibility on the basis of VNTR patterns that were completely identical to those obtained by the reference laboratory (laboratory 37). The gray bars represent the cumulative reproducibility when VNTR patterns that contained missing results were included but which were identical to those obtained by the reference laboratory when not considering the loci with missing results. Most laboratories used 24-locus VNTR typing, and three laboratories used either 14 loci (#) or 15 loci (*) for typing.

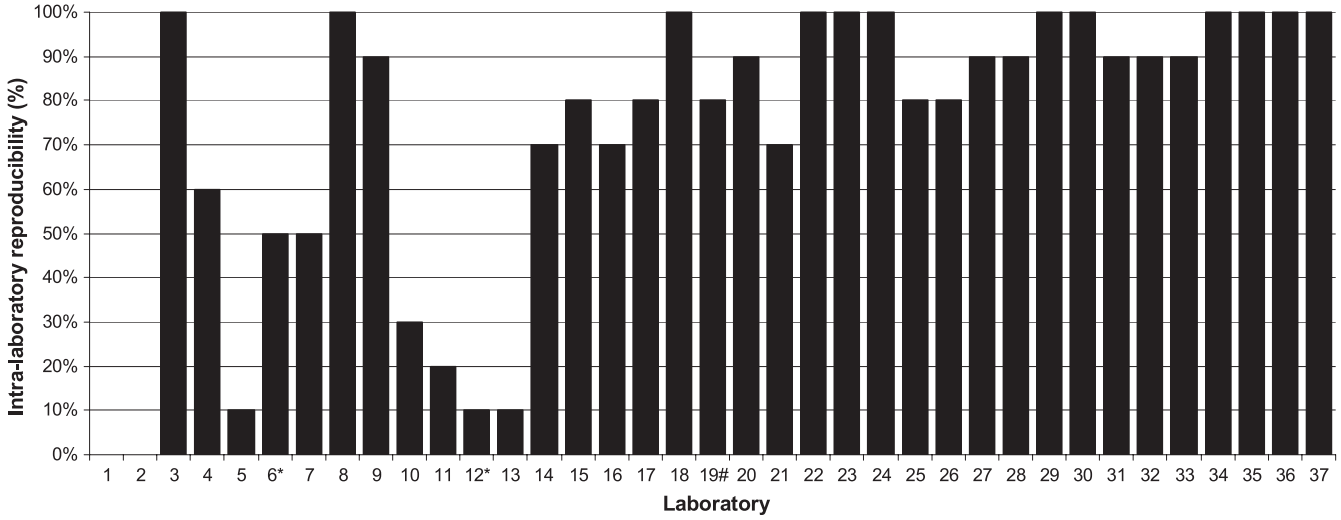

FIG 2 Intralaboratory reproducibility of VNTR typing of 10 duplicated DNA samples of *M. tuberculosis* scored by 36 participating laboratories and the reference laboratory (laboratory 37). The laboratory numbers correspond to those shown in Fig. 1. Most laboratories used 24-locus VNTR typing, and three laboratories used either 14 loci (#) or 15 loci (*) for typing.

duplicated samples, including two laboratories that had different VNTR patterns for all 10 duplicated samples.

Analysis of incorrect results. Deviations from the standard allele-calling system, i.e., the translation of PCR product sizes into repeat numbers, were the main cause of systematic errors. Systematic errors were identified for 17 of the 24 VNTR loci, reported by 1 to 5 laboratories for each of these loci. Among the nine laboratories that had an incorrect analysis of all 30 samples, and thus had 0% interlaboratory reproducibility, six had incorrect analyses caused by systematic errors. The other three laboratories reported random errors.

In two laboratories, incidents of erroneous sample exchange were detected. One laboratory exchanged two samples, and another laboratory tested one sample twice, leaving another sample untested. The errors due to sample exchange were excluded from the error analysis by VNTR locus, as were the complete numbers of errors of two laboratories that produced an extremely high number of errors (responsible for 24% and 28% of the total not-typed or incorrectly typed samples, after the exclusion of the systematic errors). The highest number of errors was identified for the locus at genomic position 4052, which was typed incorrectly 54 times (15% of all the errors, after the exclusion of the errors mentioned above). The error rate for the other VNTR loci ranged from 0% ($n = 0$) for the locus at position 154 to 11.5% ($n = 42$) for the locus at position 4156 (Fig. 3).

Reproducibility by VNTR locus. When the intra- and interlaboratory reproducibilities were analyzed at the level of individual loci instead of the complete 24-locus VNTR patterns, the average scores for inter- and intralaboratory reproducibility increased from 60% to 92% and from 72% to 96%, respectively.

To determine the extent of typing errors in the 24-locus VNTR patterns of the incorrectly typed samples, the results of the erroneous sample substitutions in two laboratories and the missing loci in the results were excluded from the calculation; this left 31 laboratories for analysis. The average number of incorrect loci per strain was calculated for the samples with incorrect VNTR patterns. Thirteen of the participating laboratories had an average of one incorrect locus, which means that for the incorrectly typed samples, there was on average one locus that was different compared to the reference. Six laboratories had on average between 1 and 2 incorrectly analyzed loci, seven laboratories had between 2 and 4 incorrectly analyzed loci, and five laboratories had on average more than 4 incorrectly analyzed loci. The latter laboratories, which were identified as the laboratories with the highest numbers of errors per strain, all made systematic errors.

Reproducibility and typeability by method. The different methodologies applied in the laboratories impacted the typeability and the intra- and interlaboratory reproducibilities (Table 1). The 15 participants that applied the in-house method with amplicon sizing by gel electrophoresis had average inter- and intralaboratory reproducibilities of 50% and 74%, respectively, and the 9 laboratories that used an in-house PCR method and determined amplicon sizes with a DNA analyzer had scores of 70% and 73%, respectively. The best results were obtained by the 9 participants who used a commercial kit and a DNA analyzer, 88% and 82%, respectively. However, the conditional chi-square test showed that none of the groups was homogeneous (P values of <0.001), and certain laboratories had many more errors than other laboratories within the same group, which raises the question of whether these differences in reproducibility are significant or even meaningful. The interlaboratory reproducibility of the methodologies was influenced by systematic errors, which were reported by five laboratories that used the in-house PCR method with amplicon sizing by gel electrophoresis, by a single laboratory that used the in-house method and determination of amplicon sizes with a DNA analyzer, and by three laboratories that used other methods. In contrast, systematic errors were not detected for the laboratories that used the commercial 24-locus VNTR typing kit.

The typeability score represents the ability of a method to deliver a typing result, without considering the quality of this result. The typeability score for the complete VNTR patterns was higher for laboratories that applied the in-house method with amplicon sizing by gel electrophoresis and the laboratories that used the commercial VNTR kit; with these approaches, on average, 97%

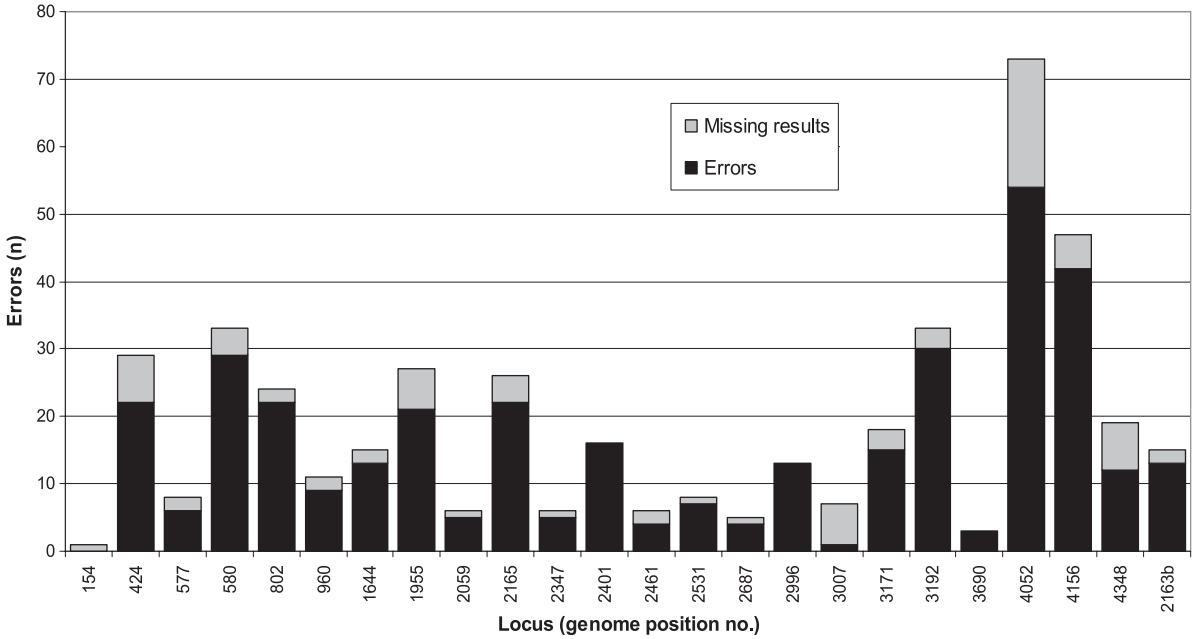

FIG 3 Distribution of detected errors by locus in the 24-locus VNTR panel calculated on the basis of the results of typing of 30 *M. tuberculosis* complex DNA samples by 35 laboratories (excluding laboratories 2 and 5, the systematic errors, and the two incidents of sample exchange described in detail in the text). Among the errors, wrong results and missing results were distinguished.

and 96% of all samples, respectively, yielded a complete typing pattern. Laboratories that used an in-house PCR method and determined amplicon sizes with a DNA analyzer scored 92%, and laboratories that used other methods scored only 79%.

DISCUSSION

This is the first multicenter proficiency study of 24-locus VNTR typing of *M. tuberculosis* complex strains after its standardization in 2006. Initial tests conducted early in the development of this technique, with bilaterally exchanged blinded reference samples, demonstrated that VNTR typing was highly reproducible for a limited number of experienced laboratories at both the intra- and interlaboratory levels (5, 16, 17, 24, 27). However, our study showed that many laboratories were not able to perform 24-locus VNTR typing in agreement with the reference. In fact, our study demonstrates that the reproducibility of 24-locus VNTR typing as it is currently applied is limited, with an average interlaboratory reproducibility of 60% (range, 0% to 100%) and an average intralaboratory reproducibility of 72% (range, 0% to 100%). Errors were caused by systematic deviations from the standard allele-calling system (i.e., rules for translating PCR product sizes into repeat numbers), the challenging amplification of some VNTR loci, and nontechnical issues such as sample exchange and incorrect software use. Appropriate analysis, corrections of the problems encountered, good laboratory practice, and further standardization are essential for the reliable use of this technique to study the molecular epidemiology of TB at local and international levels.

Nontechnical errors. Two incidents of sample exchanges reduced the inter- and intralaboratory reproducibility scores for the two respective laboratories substantially. Mistakes also occurred while managing the results with the frequently used Excel software program (Microsoft Co., Redmond, WA). Systematic labeling and handling of samples during all stages of the VNTR typing technique as well as schemes for data management should be incorporated into protocols to minimize these nontechnical errors.

Allele-calling errors. Systematic deviations from the standard allele-calling system were the cause of 51% of all errors (including the missing results as systematic errors on the locus level) detected in this study, and this has significantly reduced the interlaboratory reproducibility measured in this proficiency study. The use of alternative PCR primer sequences by some groups (e.g., as described previously by Le Fleche et al. [19]) can also affect allele calling based on amplicon sizes. The use of nonstandard allele-calling systems is a general issue in VNTR typing. This problem could be solved by the standardization of the allele calling tables between user laboratories. The use of standard allele calling tables and primers (e.g., publicly available from the MIRU-VNTR*plus* website [www.miru-vntrplus.org]) establishes an unequivocal correspondence between an amplicon size calculated by electrophoresis on agarose gels and a conventional repeat unit number for these loci. The calibration of CE-based DNA analyzers with allelic ladders included in the commercial kit or the use of reference samples enables the integration of the standard allele calling information for subsequent semiautomated allele-calling specialized software. In addition, both calibration methods correct CE-specific effects, i.e., determine the relative migration between the size standard and the PCR products, depending on the VNTR loci and alleles, and differ by polymers used for capillary electrophoresis and between instruments (1, 26). To ensure compliance with the standard allele-calling system, the inclusion of reference strains in each run enables first-line quality control of amplicon size interpretations.

VNTR typing method. The highest reproducibility on average was obtained by the participants using the CE-based commercial kit, followed by laboratories using in-house-adapted CE-based

typing, those using gel-based analysis, and, lastly, those using other methods. Analysis of amplicon sizes by gel electrophoresis is inexpensive and simple but requires a high degree of organization and good laboratory practices, as each sample requires the individual amplification of 24 loci, separate analysis on gels, and, finally, manual interpretation and database management. Another factor that presumably negatively influenced the performances of the gel-based VNTR typing method is the lower level of accuracy of allele sizing on gels, especially for larger PCR fragments. This factor can partly account for the larger proportion of (single- or more-repeat) errors observed for the locus at position 4052, having the largest repeat unit size (111 bp [26]) and therefore generating some of the largest allele sizes among the 24 loci (approximately 700 to 1,000 bp in this panel). Six laboratories were able to achieve an interlaboratory reproducibility of 80% or higher by using manual typing, suggesting that the challenges are not inherent to the VNTR typing method and can be addressed by use of appropriate quality control and quality assurance measures.

However, the conditional chi-square test showed that none of the groups of methods was homogeneous (P values of <0.001). Since each laboratory belonged to a single group of methods, the observed differences in reproducibility could very well be due to the laboratories. As an attempt to extricate the potential effects of methods from the effects of the laboratory, we have fitted a mixed-effects Poisson model to the data, in which the number of errors is modeled in terms of method and locus and each laboratory contributes a random effect. The model's estimates yielded no significant effect of the method (P values of 0.179 and 0.362 for the effects of both in-house methods over the reference commercial method) and suggest that the differences between the three methods are due mainly to the laboratories using them. In particular, the estimates of the random effects reflect the excessive numbers of errors observed for certain laboratories very well. Note, however, that this lack of evidence for differences between the choices of method may, as usual, be due to the actual differences being relatively small and, hence, should not be taken as a positive statement of "no difference."

The typeability, i.e., the ability to report a complete VNTR pattern, was slightly better in laboratories that used gel electrophoresis for the detection of amplicon sizes than in those that used a CE DNA analyzer. This may be explained by differences in the PCR protocols used to amplify the VNTR loci. The in-house methods with gel electrophoresis detection used monoplex PCRs, which have an advantage over multiplex PCRs used in combination with a CE DNA analyzer. For multiplex PCR systems, some loci may not be amplified as efficiently as others because of amplification competition; the amplification of relatively large targets in combination with relatively small targets can be particularly challenging and can be the cause of unamplified loci. The use of monoplex PCRs with the VNTR typing method will improve the amplification efficiency. Therefore, we recommend that laboratories that use CE DNA analyzers to estimate fragment sizes should repeat the amplification of a particular locus with a monoplex PCR in cases of missing loci to either determine the size of the amplicon or confirm the missing results.

The typeability of VNTR typing, as measured in this study, was negatively influenced by the incorporation of an *M. canettii* isolate in the DNA panel. For instance, 15 of 37 participating laboratories were not able to amplify the VNTR locus at position 4052 of the respective DNA sample. This can be explained by the presence of single nucleotide polymorphisms (SNPs) in the primer regions of *M. canettii* isolates (P. Supply et al., unpublished data). *M. canettii* isolates display exceptionally large and therefore more challenging alleles in other loci (e.g., VNTR loci at positions 802 and 2165 are also associated with typeability and/or reproducibility problems) (9, 10, 13). As *M. canettii* strains are found very rarely outside East Africa (9, 10, 31), the problems regarding VNTR typing are not likely to have an impact on the monitoring of the transmission of TB in other areas of the world.

If considered at the level of individual loci instead of complete 24-locus genotypes, the average inter- and intralaboratory reproducibilities increased from 60% to 92% and from 72% to 96%, respectively. However, any mistake in a single locus will finally generate an incorrect 24-locus VNTR typing pattern. Translated into the main objective of typing, namely, the accurate DNA fingerprinting of *M. tuberculosis* isolates for the detection of TB transmission and bacterial population diversity in a certain area, such individual-locus mistakes would lead to the incorrect confirmation or exclusion of an epidemiological link. Therefore, from a public health perspective, reproducibility must thus be considered at the full 24-locus genotype level.

The results of this study cast doubt on the validity of some of the data for VNTR typing reported previously. Therefore, we recommend that each laboratory that obtained unsatisfactory results in this proficiency study should retype part of their previously typed isolates with an improved procedure to evaluate the reliability of earlier studies. Moreover, this study underlines the importance of first-, second-, and third-line quality control of DNA fingerprint methods in general and of MLVA methods in particular. The very low reproducibility scores for some of the laboratories in our study of the reproducibility of VNTR typing of *M. tuberculosis* complex isolates should prompt the development of proficiency schemes for MLVA of other microorganisms (20).

In summary, the results of this proficiency study suggest that the use of the commercial kit will provide optimal levels of VNTR typing reproducibility and typeability. This evaluation proved useful in detecting systematic and incidental errors that reduced the reproducibility of in-house-adapted typing. In fact, 11 laboratories that used in-house-adapted manual typing or automated typing scored inter- and intralaboratory reproducibilities of 80% or higher, which suggests that these approaches can be used in a reliable way, provided that minimal standard conditions are followed.

Based on the information collected in this study, the main observation is the importance of incorporating appropriate quality control and quality assurance measures into protocols for genotyping methods. Continuous improvements of protocols, based on experiences shared by the laboratories in the typing network, will be made available at the RIVM Tuberculosis Reference Laboratory website (www.tuberculosis.rivm.nl). After the introduction of the suggested improvement in the methodology, it is expected that VNTR typing will become more reliable in most laboratories. To investigate this, a second round of proficiency studies will be arranged shortly.

ACKNOWLEDGMENTS

Data from the European Union and European Economic Area Member States included in this study are part of the ECDC MDR-TB molecular surveillance project. The project is coordinated by the ECDC, and the implementation was outsourced to the RIVM under ECDC service con-

tract ECDC/08/019 (February 2009 to February 2012). We acknowledge José Ferreira for statistical analysis of the data. We acknowledge Delaina P. Hooks from the Division of TB Elimination, CDC, Atlanta, GA, for participation and her colleague Lauren Cowan for helpful comments on the manuscript.

All the members of the global network have contributed equally to the results of this first worldwide proficiency study on variable-number tandem-repeat typing of *Mycobacterium tuberculosis* complex strains. Members of the Global Network for the Molecular Surveillance of Tuberculosis 2009 include David Alexander, Public Health Laboratories, Ontario Agency for Health Protection and Promotion, Toronto, Canada; Ewa Augustynowicz-Kopeć, National Tuberculosis Reference Laboratory, National Research Institute of Tuberculosis and Lung Diseases, Warsaw, Poland; Urska Bidovec-Stojkovic, Laboratory for Mycobacteria, University Clinic of Respiratory and Allergic Diseases, Golnik, Slovenia; Martin J. Boeree, Department of Pulmonary Diseases, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; Timothy Brown, United Kingdom Health Protection Agency, London, United Kingdom; Daniela M. Cirillo, Reference Laboratory for Tuberculosis, Emerging Bacterial Pathogens Unit, Milan, Italy; Laura Cruz, California Department of Health Services, Richmond, CA; Susana David, Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisbon, Portugal; Raul Diaz, National Reference Laboratory of Tuberculosis and Mycobacteria, Pedro Kouri Institute of Tropical Medicine, Havana, Cuba; Horng-Yunn Dou, Division of Infectious Diseases, National Health Research Institutes, Miaoli Country, Taiwan, Republic of China; Jason T. Evans, Regional Centre for Mycobacteriology, United Kingdom Health Protection Agency West Midlands Public Health Laboratory, Birmingham, United Kingdom; Maryse Fauville-Dufaux, Tuberculosis and Mycobacteria, Scientific Institute of Public Health, Brussels, Belgium; Margaret M. Fitzgibbon, Irish Mycobacteria Reference Laboratory, Central Pathology Laboratory, Dublin, Ireland; Darío García de Viedma, Clinical Microbiology and Infectious Diseases Department, Gregorio Marañón Hospital, Madrid, Spain, and CIBER Enfermedades Respiratorias, Spain; Maria Globan, Mycobacterium Reference Laboratory, Victorian Infectious Diseases Reference Laboratory, Victoria, Australia; Ramona Groenheit, Swedish Institute for Infectious Disease Control, Solna, Sweden; Marjo Haanperä-Heikkinen, Mycobacterial Reference Laboratory, National Institute for Health and Welfare, Turku, Finland; Alexander Indra, Austrian Agency for Health and Food Safety, Institute for Medical Microbiology and Hygiene, Vienna, Austria; Kai Man Kam, Public Health Laboratory Centre, Kowloon, Hong Kong; Rebecca Kramer, Michigan Department of Community Health, Lansing, MI; Shang Mei, National Tuberculosis Reference Laboratory, Tuberculosis and Thoracic Tumor Institute, Beijing, China; Stefan Niemann, National Reference Center for Mycobacteria, Forschungszentrum Borstel, Borstel, Germany; Mihaela Obrovac, National Mycobacteria Reference Laboratory, Croatian National Institute of Public Health, Zagreb, Croatia; Erik M. Rasmussen, International Reference Laboratory of Mycobacteriology, Copenhagen, Denmark; Guislaine Refrégier, Infection Genetics of Emerging Pathogens and Epidemiology, Orsay, France; Jaime A. Robledo, Centro Colombiano de Investigación en Tuberculosis Corporación para Investigaciones Biológicas-Universidad Pontificia Bolivariana Medellín, Medellín, Colombia; Sofia Samper, Servicio de Microbiología, Hospital Universitario Miguel Servet, IIS Aragon, Zaragoza, and CIBER Enfermedades Respiratorias, Spain; Meenu K. Sharma, National Reference Centre for Mycobacteriology, National Microbiology Laboratory, Public Health Agency of Canada, and Department of Medical Microbiology, University of Manitoba, Winnipeg, Canada; Wladimir Sougakoff, National Reference Centre for Mycobacteria (CNR-MyRMA), Laboratory of Bacteriology, AP_HP, Paris, France; Petras Stakenas, Institute of Biotechnology, Vilnius University, Vilnius, Lithuania; Ruth Stavrum, Section of Microbiology and Immunology, The Gade Institute, University of Bergen, Bergen, Norway; Philip N. Suffys, Laboratory of Molecular Biology Applied to Mycobacteria, Rio de Janeiro, Brazil; Juraj Trenkler, Slovak National Reference laboratory for Mycobacteria, Nitra, Slovakia; Takayuki Wada, Osaka City Institute of Public Health and Environmental Sciences, Osaka, Japan; and Robin M. Warren, Medical Research Council Centre for Molecular and Cellular Biology, Department of Medical Biochemistry, and Centre for Tuberculosis Research and Education, Department of Child Health and Paediatrics, Faculty of Health Sciences, Stellenbosch University, Tygerberg, South Africa.

REFERENCES

1. **Allix C, Supply P, Fauville-Dufaux M.** 2004. Utility of fast mycobacterial interspersed repetitive unit-variable number tandem repeat genotyping in clinical mycobacteriological analysis. Clin. Infect. Dis. **39**:783–789.
2. **Brown TJ, Nikolayevskyy VN, Drobniewski FA.** 2009. Typing Mycobacterium tuberculosis using variable number tandem repeat analysis. Methods Mol. Biol. **465**:371–394.
3. **Cole ST, et al.** 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature **393**:537–544.
4. **Cowan LS, et al.** 2005. Evaluation of a two-step approach for large-scale, prospective genotyping of Mycobacterium tuberculosis isolates in the United States. J. Clin. Microbiol. **43**:688–695.
5. **Cowan LS, Mosher L, Diem L, Massey JP, Crawford JT.** 2002. Variable-number tandem repeat typing of Mycobacterium tuberculosis isolates with low copy numbers of IS6110 by using mycobacterial interspersed repetitive units. J. Clin. Microbiol. **40**:1592–1602.
6. **Devaux I, Kremer K, Heersma H, Van Soolingen D.** 2009. Clusters of multidrug-resistant Mycobacterium tuberculosis cases, Europe. Emerg. Infect. Dis. **15**:1052–1060.
7. **Devaux I, Manissero D, Fernandez de la Hoz K, Kremer K, van Soolingen D.** 2010. Surveillance of extensively drug-resistant tuberculosis in Europe, 2003–2007. Euro Surveill. **15**(11):pii=19518. http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19518.
8. **Evans JT, et al.** 2004. Automated high-throughput mycobacterial interspersed repetitive unit typing of Mycobacterium tuberculosis strains by a combination of PCR and nondenaturing high-performance liquid chromatography. J. Clin. Microbiol. **42**:4175–4180.
9. **Fabre M, et al.** 2010. Molecular characteristics of "Mycobacterium canettii" the smooth Mycobacterium tuberculosis bacilli. Infect. Genet. Evol. **10**:1165–1173.
10. **Fabre M, et al.** 2004. High genetic diversity revealed by variable-number tandem repeat genotyping and analysis of hsp65 gene polymorphism in a large collection of "Mycobacterium canettii" strains indicates that the M. tuberculosis complex is a recently emerged clone of "M. canettii.". J. Clin. Microbiol. **42**:3248–3255.
11. **Frothingham R, Meeker-O'Connell WA.** 1998. Genetic diversity in the Mycobacterium tuberculosis complex based on variable numbers of tandem DNA repeats. Microbiology **144**(Pt 5):1189–1196.
12. **Gopaul KK, Brown TJ, Gibson AL, Yates MD, Drobniewski FA.** 2006. Progression toward an improved DNA amplification-based typing technique in the study of Mycobacterium tuberculosis epidemiology. J. Clin. Microbiol. **44**:2492–2498.
13. **Gutierrez MC, et al.** 2005. Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis. PLoS Pathog. **1**:e5.
14. **Hyytia-Trees E, Lafon P, Vauterin P, Ribot EM.** 2010. Multilaboratory validation study of standardized multiple-locus variable-number tandem repeat analysis protocol for Shiga toxin-producing Escherichia coli O157: a novel approach to normalize fragment size data between capillary electrophoresis platforms. Foodborne Pathog. Dis. **7**:129–136.
15. **Kremer K, et al.** 2005. Discriminatory power and reproducibility of novel DNA typing methods for Mycobacterium tuberculosis complex strains. J. Clin. Microbiol. **43**:5628–5638.
16. **Kremer K, et al.** 2005. Use of variable-number tandem-repeat typing to differentiate Mycobacterium tuberculosis Beijing family isolates from Hong Kong and comparison with IS6110 restriction fragment length polymorphism typing and spoligotyping. J. Clin. Microbiol. **43**:314–320.
17. **Kremer K, et al.** 1999. Comparison of methods based on different molecular epidemiological markers for typing of Mycobacterium tuberculosis complex strains: interlaboratory study of discriminatory power and reproducibility. J. Clin. Microbiol. **37**:2607–2618.
18. **Kwara A, et al.** 2003. Evaluation of the epidemiologic utility of secondary typing methods for differentiation of Mycobacterium tuberculosis isolates. J. Clin. Microbiol. **41**:2683–2685.
19. **Le Fleche P, Fabre M, Denoeud F, Koeck JL, Vergnaud G.** 2002. High

resolution, on-line identification of strains from the Mycobacterium tuberculosis complex based on tandem repeat typing. BMC Microbiol. **2**:37.
20. **Lindstedt BA.** 2005. Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. Electrophoresis **26**:2567–2582.
21. **Magdalena J, Vachee A, Supply P, Locht C.** 1998. Identification of a new DNA region specific for members of Mycobacterium tuberculosis complex. J. Clin. Microbiol. **36**:937–943.
22. **Mazars E, et al.** 2001. High-resolution minisatellite-based typing as a portable approach to global analysis of Mycobacterium tuberculosis molecular epidemiology. Proc. Natl. Acad. Sci. U. S. A. **98**:1901–1906.
23. **Murase Y, Mitarai S, Sugawara I, Kato S, Maeda S.** 2008. Promising loci of variable numbers of tandem repeats for typing Beijing family Mycobacterium tuberculosis. J. Med. Microbiol. **57**:873–880.
24. **Savine E, et al.** 2002. Stability of variable-number tandem repeats of mycobacterial interspersed repetitive units from 12 loci in serial isolates of Mycobacterium tuberculosis. J. Clin. Microbiol. **40**:4561–4566.
25. **Skuce RA, et al.** 2002. Discrimination of Mycobacterium tuberculosis complex bacteria using novel VNTR-PCR targets. Microbiology **148**:519–528.
26. **Supply P, et al.** 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. J. Clin. Microbiol. **44**:4498–4510.
27. **Supply P, et al.** 2001. Automated high-throughput genotyping for study of global epidemiology of Mycobacterium tuberculosis based on mycobacterial interspersed repetitive units. J. Clin. Microbiol. **39**:3563–3571.
28. **Supply P, et al.** 2000. Variable human minisatellite-like regions in the Mycobacterium tuberculosis genome. Mol. Microbiol. **36**:762–771.
29. **van Embden JD, et al.** 1993. Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. **31**:406–409.
30. **van Soolingen D, Arbeit RD.** 2001. Dealing with variation in molecular typing of Mycobacterium tuberculosis: low-intensity bands and other challenges. J. Med. Microbiol. **50**:749–751.
31. **van Soolingen D, et al.** 1997. A novel pathogenic taxon of the Mycobacterium tuberculosis complex, Canetti: characterization of an exceptional isolate from Africa. Int. J. Syst. Bacteriol. **47**:1236–1245.