

## RESEARCH ARTICLE

## Developing customized stepwise MIRU-VNTR typing for tuberculosis surveillance in Georgia

Nino Maghradze<sup>1,2,3</sup>, Levan Jugheli<sup>1,2,3</sup>, Sonia Borrell<sup>1,2</sup>, Nestani Tukvadze<sup>1,2,3</sup>, Russell R. Kempker<sup>4</sup>, Henry M. Blumberg<sup>4,5</sup>, Sebastien Gagneux<sup>1,2\*</sup>

**1** Swiss Tropical and Public Health Institute, Basel, Switzerland, **2** University of Basel, Basel, Switzerland, **3** National Center for Tuberculosis and Lung Diseases (NCTLD), Tbilisi, Georgia, **4** Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, GA, United States of America, **5** Departments of Epidemiology and Global Health, Rollins School of Public Health of Emory University, Atlanta, GA, United States of America

\* [sebastien.gagneux@unibas.ch](mailto:sebastien.gagneux@unibas.ch)



## Abstract

## Introduction

Mycobacterial Interspersed Repetitive Units–Variable Tandem Repeats (MIRU-VNTR) typing has been widely used for molecular epidemiological studies of tuberculosis (TB). However, genotyping tools for *Mycobacterium tuberculosis* (Mtb) may be limiting in some settings due to high cost and workload. In this study developed a customized stepwise MIRU-VNTR typing that prioritizes high discriminatory loci and validated this method using penitentiary system cohort in the country of Georgia.

## Methods

We used a previously generated MIRU-VNTR dataset from recurrent TB cases (32 cases) in Georgia and a new dataset of TB cases from the penitentiary system (102 cases) recruited from 2014 to 2015. A Hunter-Gaston Discriminatory Index (HGDI) was calculated utilizing a 24 standard loci panel, to select high discriminatory power loci, subsequently defined as the customized Georgia-specific set of loci for initial typing. The remaining loci were scored and hierarchically grouped for second and third step typing of the cohort. We then compared the processing time and costs of the customized stepwise method to the standard 24-loci method.

## Results

For the customized Georgia-specific set that was used for initial typing, 10 loci were selected with a minimum value of 0.32 to the highest HGDI score locus. Customized 10 loci (step 1) typing of 102 Mtb patient isolates revealed 35.7% clustered cases. This proportion was reduced to 19.5% after hierarchical application of 2<sup>nd</sup> and 3<sup>rd</sup> step typing with the corresponding groups of loci. Our customized stepwise MIRU-VNTR genotyping approach reduced the quantity of samples to be typed and therefore overall processing time and costs by 42.6% each.

## OPEN ACCESS

**Citation:** Maghradze N, Jugheli L, Borrell S, Tukvadze N, Kempker RR, Blumberg HM, et al. (2022) Developing customized stepwise MIRU-VNTR typing for tuberculosis surveillance in Georgia. PLoS ONE 17(3): e0264472. <https://doi.org/10.1371/journal.pone.0264472>

**Editor:** Pradeep Kumar, North Eastern Regional Institute of Science and Technology, INDIA

**Received:** October 20, 2021

**Accepted:** February 10, 2022

**Published:** March 1, 2022

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** This work was supported by the Swiss National Science Foundation (grant numbers CRSII5\_177163 and 310030\_188888) – Sebastien Gagneux, the European Research Council (ECOEVODRTB\_883582) – Sebastien Gagneux, and from the U.S. National Institute of Health (NIH) Fogarty International Center (D43TW007124) – Russell R Kempker, Henry M. Blumberg. The funders had no role in study design, data collection

and analysis, decision to publish, or preparation of the manuscript.

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

## Conclusion

Our study shows that our customized stepwise MIRU-VNTR typing approach is a valid alternative of standard MIRU-VNTR typing panels for molecular epidemiological investigation in Georgia that saves time, workload and costs. Similar approaches could be developed for other settings.

## Introduction

Genotyping plays an important role in the surveillance and molecular epidemiology of tuberculosis (TB). However, the implementation of TB genotyping methods has been limited in many settings, especially in low- and middle-income countries (LMIC), given that they can be computationally demanding, time-consuming and costly [1, 2]. Mycobacterial Interspersed Repetitive Unit-Variable Number of Tandem Repeats (MIRU-VNTR) typing is widely used to evaluate the population structure, strain genetic diversity, and transmission of the *Mycobacterium tuberculosis* complex (MTBC) [3, 4]. Globally, the human-adapted MTBC can be classified into nine phylogenetic lineages that differ in their geographic distribution [5, 6]. This phylo-geographic genetic diversity also impacts the discriminatory power of MIRU-VNTR loci in a given population, which has led to the development of customized MIRU-VNTR loci sets for application in different geographical settings [2, 7, 8]. Compared to the MIRU-VNTR typing using standard 12, 15 or 24 loci panels proposed by Supply *et al* [9], customized approaches have focused on the most sensitive and highly discriminatory loci [10, 11], allowing for reduced workload, labor and consumable-dependent cost and sample processing time.

However, customization and reduction of the MIRU-VNTR loci set, might be inefficient for reaching higher level of discrimination, especially where the variety of MTBC lineages is limited and the bacterial population is homogeneous. In such instances, various combinations of genotyping tools for increasing resolution have been proposed [11–13], but no systematic approach has yet been established. Considering the limited resources in LMIC, combining several distinct typing methods increases the complexity and costs, and the absence of standardization results in poor comparability between the laboratories. In the present study, we implemented customized stepwise MIRU-VNTR typing specifically for the epidemiological setting of Georgia [14]. This approach is based on selecting and developing groups of MIRU-VNTR loci, to reduce the number of samples to be typed after each step.

Recently, standard 24-loci MIRU-VNTR typing has been established at the National Reference Laboratory (NRL) of National Center of Tuberculosis and Lung Diseases (NCTLD) in Tbilisi, Georgia, which has been used to differentiate between relapse and reinfection in recurrent TB cases [15]. In this study, we modified the standard 24-loci MIRU-VNTR panel into a customized stepwise typing approach for the Georgian Mtb population and applied it to a patient population from the penitentiary system.

While important progress in TB control and a substantial decrease in TB incidence in Georgia from 2002 (228 per 100,000) to 2019 (80 per 100,000) has occurred, many challenges to TB elimination remain, including high rates of drug-resistant TB cases. In particular, the penitentiary system has been identified as a hotspot of TB disease transmission including multidrug-resistant (MDR)-TB [16–18]. Over the past six years, with the advent of new programmatic initiatives, cases of TB diagnosed in the penitentiary system have decreased, including MDR and extensively drug-resistant (XDR) TB, from 33% of all MDR cases in the country in 2011 to the 4.1% in 2020 (National Surveillance Program, unpublished data). However, the

penitentiary system remains a persistent site for transmission of drug-resistant TB [17–19], bearing a risk of infection spillover into the community. Hence, rapid detection of TB transmission in prisons, along with infection surveillance in the country, is crucial, and resource-adapted molecular methods for surveillance are therefore needed.

## Materials and methods

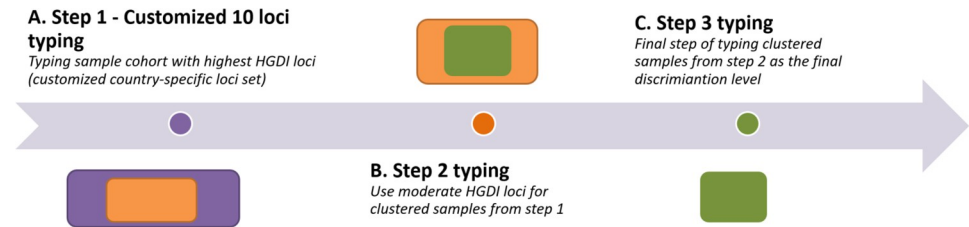
### Ethics

Ethical approval for the study was obtained from the relevant Institutional Review Boards (IRB) of the NCTLD in Georgia and in Switzerland (Ethikkommission Nordwest- und Zentralschweiz). Due to observational nature of the study which do not involve any physical intervention and was solely using the routinely collected samples from TB patient in Georgia, the need of informed consent was waived by the National Council on Bioethics of Ministry of Labor, Health and Social Affairs of Georgia, and therefore by local IRB. The data from archived samples were fully anonymized to guarantee the confidentiality.

### Customization and validation of country-specific minimal set of MIRU-VNTR loci

**Genotyping dataset and parameters for customization.** In order to develop a customized loci set for the Mtb strain population in Georgia, we first re-analyzed a dataset from a previous study of recurrent TB cases, with a retrospective cohort of 32 patients (64 samples) from 2014–2016, typed with the standard 24 loci panel [15]. Inclusion criteria of the samples for the relapsed cases was availability of the samples in the National Reference Laboratory biobank. From this dataset, we used allelic data for 37 samples; 27 single Mtb isolates from relapsed patients and paired isolates from 5 reinfection cases. We defined genotypic clusters with a single locus variants (SLV), and calculated discriminatory indexes for each locus of the 24-loci panel using Hunter and Gaston discriminatory index (HGDI) [20]. Based on the Simpson's index of diversity, HGD index was calculated using the online tool [http://insilico.ehu.es/mini\\_tools/discriminatory\\_power](http://insilico.ehu.es/mini_tools/discriminatory_power). Following the HGDI calculations and acknowledged cut-offs, the loci were classified into high ( $HGDI \geq 0.6$ ), moderate ( $0.3 \leq HGDI < 0.6$ ) and low ( $HGDI < 0.3$ ) discriminant categories [18]. We ranked the 24 loci based on their HGDI score and selected the 10 loci with highest discriminatory indexes.

**Genotyping dataset for customized set validation.** For the purpose of validating our customized set of loci, we focused on patients with pulmonary TB residing and diagnosed in the penitentiary system in 2014–2015. Cases were not overlapping the recurrent TB case samples and were identified through the national TB surveillance database and were linked to the National Reference Laboratory (NRL) sample ID numbers to identify the NRL strain biobank and phenotypic drug susceptibility testing (pDST) results. We were able to link 136 TB cases in prisons from 2014–2015, with bio banked Mtb isolates for further use in the study, however, total of 102 Mtb isolates were recovered from the laboratory biobank and were typed using the customized 10 loci set. Specifically, the 10 customized MIRU-VNTR loci were multiplexed in four separate PCRs– 1) MIRU16, MIRU31 and ETRB; 2) MIRU40, MIRU26 and Mtub30; 3) Mtub 21, QUB11b and QUB26; 4) MIRU39. Negative (H<sub>2</sub>O) and positive (H37Rv DNA) controls were included in each reaction. PCR amplification was performed under following conditions: the initial denaturation at 95°C for 15 minutes, followed by 40 cycles of 94°C for 1 minute, 59°C for 1 minute and 72°C for 1.5 minutes with final extension at 72°C for 10 minutes. We used low-resolution electrophoresis for amplification check, followed by higher resolution electrophoresis with 1.8% agarose gel, using 1kb ladder as size marker. A numerical



**Fig 1. Schematic representation of the MIRU-VNTR stepwise approach.** Schematic representation of the stepwise approach, with reduced amount of the samples to be typed after application of each step of locus group, resulting reduced time, cost and workload.

<https://doi.org/10.1371/journal.pone.0264472.g001>

code was obtained by comparison of the allelic table published by Supply *et al* [16]. Genotypic clusters were defined based on SLV.

### Re-evaluation of the customized set and HGD indexes and stepwise approach development

For the stepwise typing approach development, we used the remaining 14 loci to type 35 Mtb isolates from the penitentiary system that were clustered after typing with the customized 10 loci set. Additionally, we re-assessed the discriminatory power of each locus based on the combined datasets of i) recurrent TB cases (37 samples) and ii) prison cohort samples for which full 24 loci results were available (35 samples). The HDGI was re-calculated and the loci were ranked following the cut-off scores described above. The criteria of locus hierarchy and inclusion into the two-step typing panel was defined by i) HGDI, calculated within the recurrent TB cases and penitentiary system cohort; and ii) ability of the locus to discriminate and reduce proportion of the samples to be typed by the stepwise addition of each locus to the predefined 10 loci customized set (Fig 1).

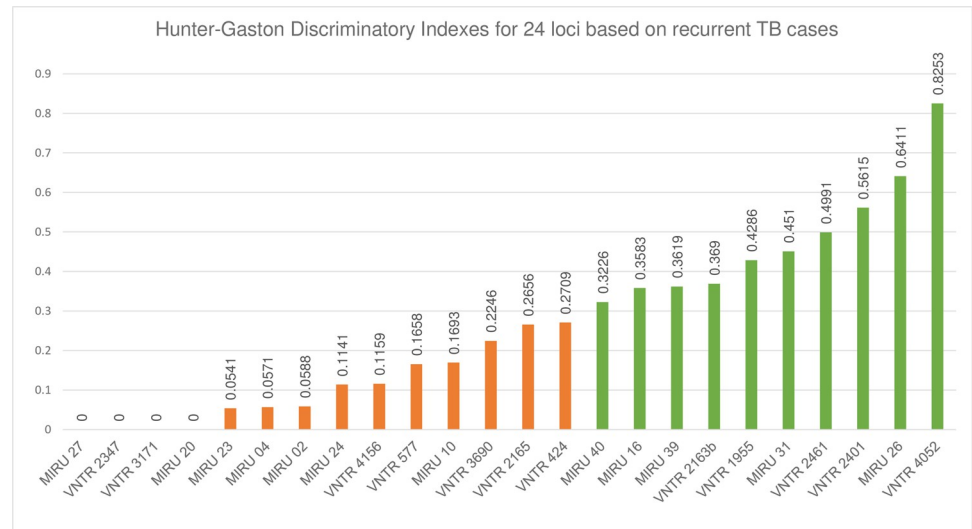
### Cost and time assessment

Along with method modification, we assessed the processing time and approximate direct cost based on the technician per-hour labor cost, consumables and reagents pricing in Georgia for a batch of 54 samples processed with the customized stepwise MIRU-VNTR typing (excluding DNA extraction) and standard 24-locus panel, with 2 additional controls. The number of samples in a batch was defined based on the number of wells available on our electrophoresis tray.

## Results

### Customization of the Georgia-specific MIRU-VNTR loci set using typing results from recurrent TB cases

Utilizing the standard 24-loci panel, MIRU-VNTR typing of 37 samples from our previously published cohort of recurrent TB cases [15] revealed two highly discriminatory loci, namely—VNTR4052 (QUB-26) and MIRU26, with an HGDI of 0.83 and 0.64, respectively (Fig 2). Eight moderately discriminant loci were identified: VNTR2401/Mtub30 (HGDI = 0.56), VNTR2461/ETR B (HGDI = 0.50), MIRU31/ETR E (HGDI = 0.45), VNTR1955/Mtub21 (HGDI = 0.43), VNTR2163b/QUB-11b (HGDI = 0.34), MIRU39 (HGDI = 0.36), MIRU16 (HGDI = 0.36), MIRU40 (HGDI = 0.32). Another ten loci showed low discriminatory indices: VNTR424/Mtub04 (HGDI = 0.27), VNTR2165/ETR A (HGDI = 0.27), VNTR3690/Mtub39 (HGDI = 0.22), MIRU10 (HGDI = 0.17), VNTR577/ETR C (HGDI = 0.17), VNTR4156/QUB-4156c (HGDI = 0.12), MIRU24 (HGDI = 0.11), MIRU02 (HGDI = 0.06), MIRU04/ETR D



**Fig 2. HGDI scores of 24-loci panel based on recurrent TB cases.** Graph represents individual discriminatory index for 24-loci panel generated from recurrent TB cohort typing results. Green bars indicate “the Georgian customized set” with high and moderate index loci included in the customized set.

<https://doi.org/10.1371/journal.pone.0264472.g002>

(HGDI = 0.06) and MIRU23 (0.05). Finally, four loci, MIRU20, VNTR2347/Mtub 29, MIRU27/QUB-5 and VNTR3171/Mtub34 showed no discriminatory power with an HGDI score of zero (Fig 2). A Georgia-specific panel of ten loci was defined, which included the two highly and the eight moderately discriminatory loci (Fig 2). The loci in the customized Georgia-specific set were compared to the standard 15- and 12-loci panels [2]. Our customized panel shared four loci with both the 12- and 15-loci standard panels, while four and one loci were exclusively shared with the 15-loci set and the 12-loci set, respectively. VNTR2461 with a moderate discriminatory power in the Georgia-specific panel has so far not been included in earlier proposed typing sets.

## MIRU-VNTR typing of the Mtb samples from the penitentiary system

**Genotyping with the customized 10-loci set.** From total of unique 102 Mtb isolates from prison TB cases, full allelic data was obtained for 98 (96.1%) samples. MIRU-VNTR genotyping of 98 Mtb isolates with our customized 10-loci set revealed seven clusters by SLV comprising 35 clustered isolates and 63 singletons (non-clustered isolates). The 35 clustered isolates were distributed in the seven clusters as follows: cluster 1–12 isolates (34.4%), cluster 2–11 (31.4%), cluster 4–4 (11.4%), cluster 4 to 7 were represented with two (5.7%) isolates per cluster (Fig 3).

For the stepwise typing approach development of the clustered isolates, we used the remaining 14 loci from the standard 24-loci panel not included in our customized 10-loci set. These resulted in an additional 16 singletons, while the number of clusters remained the same, with a sample distribution as follows: cluster 1 –five isolates (26.3%), cluster 2, 3, 6 and 7 each with two isolates (10.5%), while cluster 4 and 5 had three (15.8%) isolates each, resulting in a total of 19/98 (19.5%) clustered isolates (Fig 3).

**Defining loci sets for the stepwise approach.** In order to combine loci for the second typing step, we re-evaluated and compared rankings of HGDI scores (Fig 4A, 4B) by assessing the number of discriminated samples by each locus added separately to the customized 10-loci set within the combined dataset. VNTR424/Mtub04 showed an increased HGDI (within the

Step 1 - Customized 10 loci typing																																		
MIRU16, MIRU26, MIRU31, MIRU40, VNTR1955, VNTR2163b, VNTR2401, VNTR4052, VNTR2461, MIRU39																																		
C1						C2						C3			C4	C5	C6	C7																
10	15	16	18	19	14	12	17	28	21	25	30	37	38	39	51	53	55	56	52	60	65	70	74	75	80	81	5	20	3	4	13	24	46	47
Step 2 typing																																		
VNTR424, MIRU10, VNTR577, VNTR3690, VNTR2165, VNTR3171																																		
C1						C2			C3			C4			C5			C6			C7			C8										
10	15	16	18	19	14	12	17	21	25	37	38	39	53	55	56	74	75	80	81	3	4													
Step 3 typing																																		
MIRU02, MIRU04, MIRU20, MIRU23, MIRU24, VNTR2347, VNTR4156																																		
C1						C2			C3			C4			C5			C6			C7													
10	15	16	18	19	14	12	17	21	25	37	38	39	53	55	56	74	75	80	81															

**Fig 3. Stepwise application of customized and additional 14 loci and distribution of the clustered samples.** Selected loci for the stepwise approach and distribution of the clustered samples after typing with customized 10-loci set, step two typing with 6 loci and step 3 typing with remaining 8 loci.

<https://doi.org/10.1371/journal.pone.0264472.g003>



**Fig 4. Individual locus HGDI of 10 customized and additional 14 loci.** (A) HGDI calculated for the customized 10 loci and (B) the additional 14 loci, based on typing of recurrent TB cases (civil cohort) in blue bars along with same indexes calculated within penitentiary system samples in orange. Grey bars indicate indexes calculated from combined datasets with 24 loci typing. Category definition:  $HGDI \geq 0.6$  high discriminatory power,  $0.3 \leq HGDI < 0.6$  moderate discriminatory power,  $HGDI < 0.3$  low discriminatory power.

<https://doi.org/10.1371/journal.pone.0264472.g004>

additional 14 loci set) in the penitentiary system compared to recurrent TB typing data (Fig 4B). Four loci—MIRU10, VNTR577/ETR C, VNTR2165/ETR A and VNTR3690/Mtub39 added most to the increase in HGDI. VNTR3171/Mtub34 showed an increased capacity to discriminate between the samples when added to these four loci, while the HGDI of VNTR3171/Mtub34 on its own was 0.05 and 0 in the penitentiary and recurrent TB cohort was, respectively. Application of the 2<sup>nd</sup> step loci VNTR424/Mtub04, MIRU10, VNTR577/ETR C, VNTR2165/ETR A, VNTR3690/Mtub39, and VNTR3171/Mtub34 reduced the number of clustered isolates from 35 to 22, leading to the reduction of the samples to be genotyped by 37%.

The remaining eight loci (MIRU02, MIRU04/ETR D, MIRU20, MIRU23, MIRU24, MIRU27/QUB-5, VNTR2347/Mtub 29, and VNTR4156/QUB-4156c) were included in a third and last step of genotyping. The HGDI of MIRU27/QUB-5 remained zero in both cohorts as it did not identify any discrete sample from the population. Application of the third step loci on 22 clustered samples identified 2 singletons, defining the final outcome of the customized stepwise MIRU-VNTR typing of penitentiary system cohort— 19 (19.4%) clustered cases from the total of 98 samples.

### Processing time for the customized stepwise MIRU-VNTR typing

Considering a conventional method with two gel trays with 32-well gel comb, it is possible to run 54 samples with 2 controls and 8 molecular size markers. To compare the total amount of time required for our customized stepwise genotyping approach and for the standard 24-loci method, we calculated the processing time for 56 samples (including controls) for both methods. Our calculation included the following steps: master mix preparation (20 min.), adding sample DNA (20 min.), PCR (165 min.), and loading samples for gel electrophoresis (40 min.), running the gel (240 min), dyeing (20 min.) and imaging (8 min).

Applying the full 24-loci panel on the 102 patient isolates resulted in a total of 2,448 amplicons to run on gel electrophoresis, with a total processing time of 387.6 person-hours. By contrast, our customized stepwise typing approach resulted in a reduced number of amplicons to be processed from 2,448 to 1,020 (41.7%) after 1<sup>st</sup> step customized 10 loci typing. This was followed by 210 (8.6%) amplicons after applying the 2<sup>nd</sup> step typing with 6 loci, and to 176 (7.2%) amplicons to be typed for the 3<sup>rd</sup> step with the remaining 8 loci. Hence, our customized stepwise approach resulted in a total of 1,406 amplicons to run, with 41.7% and 8.6% reduction of typing with 1<sup>st</sup> step customized loci and 2<sup>nd</sup> step typing with an additional 6 loci, respectively. Compared to the 387.6 person-hours with 24-loci panel typing, the stepwise approach required 222.6 person-hours, which corresponds to a reduction in total processing time of 42.6%.

### Cost comparison of the customized stepwise genotyping approach

The price of the 24-loci MIRU-VNTR per sample based on consumable prices in Georgia in addition to laboratory technician salary per-hour was approximately 1,124 Euros, while the customized stepwise approach using the 10-6-8 loci sequentially resulted in a decreased cost of approximately 648 Euros. This corresponds to a cost reduction of 42.6% compared to the standard method, indicating a major dependence of the total cost on the sample quantity and processing time.

### Discussion

We adapted the standard 24-loci MIRU-VNTR panel for the Georgian TB population by i) reducing the number of loci to a customized set of 10 loci for an initial step of genotyping; ii) developing a stepwise typing approach to retain adequate discriminatory power; iii) applying

this approach on a retrospective TB patient cohort in a penitentiary system in Georgia. Our findings demonstrate that the customized stepwise MIRU-VNTR typing approach reduces the processing time, cost and workload, while achieving equal discriminatory power to 24 standard MIRU-VNTR loci panel. The modified MIRU-VNTR typing can be a benefit for the TB disease transmission control in the most effective way, especially for LMIC.

For the purpose of the MIRU-VNTR customization, we modified the 24-loci panel based on the discriminatory power of each loci using HGDI. Initially, we chose the 10 highest HGDI score loci and compared them to the standardized 12- and 15-loci panel sets (Table 1). While four loci included in our customized set were shared with both the standard sets, one locus was shared only with the 12-loci panel, four loci only with the 15-loci panel, and one locus—VNTR2461 (ETRB) was not present either of the two standard sets. The bias of 15-loci and 24-loci panels towards MTBC lineage 4 has been well-documented [9]. In Georgia, most MTBC belongs to lineage 4 and lineage 2/Beijing. Deviation from the standard loci panels is mostly influenced by HGDI sensitivity to local strain variation. Therefore, our results from assessing a customized set, based on the MTBC lineages circulating in Georgia, are consistent with previous studies showing the influence of strain diversity on the discriminatory power of genotyping methods [8, 10, 11]. This variation thus should be considered when choosing appropriate methodologies for molecular epidemiological studies, whereby customizing based on local strain variation is one possible way, particularly in resource-limited settings [10, 11, 21].

**Table 1. MIRU-VNTR 24 loci set with individual HGDI and inclusion in standardized and customized genotyping panel.**

Locus (24 loci panel)	12 Loci Panel	15 Loci Panel	Customized 10-loci set
MIRU 02	X		
MIRU 04	X	X	
MIRU 40	X	X	X
MIRU 10	X	X	
MIRU 16	X	X	X
MIRU 20	X		
MIRU 23	X		
MIRU 24	X		
MIRU 26	X	X	X
MIRU 27	X		
MIRU 31	X	X	X
MIRU 39	X		X
VNTR 424		X	
VNTR 577		X	
VNTR 1955		X	X
VNTR 2163b		X	X
VNTR 2165		X	
VNTR 2347			
VNTR 2401		X	X
VNTR 2461			X
VNTR 3171			
VNTR 3690		X	
VNTR 4052		X	X
VNTR 4156		X	

<https://doi.org/10.1371/journal.pone.0264472.t001>



For increased discrimination, we additionally typed samples that were clustered based on the initial typing with the remaining loci from the 24-loci panel and re-evaluated the pre-defined customized set. The initial customized panel revealed 35.7% of clustered cases, which was subsequently reduced to 19.4% after applying an additional 14 loci. Re-evaluation of the HGDI indexes, along with assessment of the ability of each locus to identify the majority of the discrete samples during the step-by-step addition to the customized 10-loci set, allowed us to arrange the remaining 14 loci into the stepwise typing approach. While the 10-loci customized set might be sufficient as a standalone panel for particular epidemiological studies, it can be also considered as a primary typing set—step 1, for the hierarchical application of additional MIRU-VNTR loci. However, if a higher discriminatory power is required, our customized stepwise approach showed the capacity to reduce the number of samples to be typed by 41.7% and 8.6% using customized/step 1 and step two typing, respectively, leading to a reduced workload, processing time and cost.

We did observe the influence of the particular study population on the discriminatory power of specific loci, for example for locus VNTR424 and VNTR3171. Therefore, evaluation of the discriminatory power by HGDI needs to consider the homogeneity of the population to be typed.

Our customized stepwise approach advances classical 24 MIRU-VNTR panel by reduced workload, processing time and direct costs (technician per-hour labor cost, consumables and reagents) by 42.6%. Given that our findings are based on local pricing of the consumables and laboratory technician average salary, the result should thus be treated with caution as they will vary based on the specific context with regards to costs of technician salaries and consumables. Variation in cost reduction (absolute number) is attributable to local pricing. In Georgia, a middle income country, the consumable prices are much higher compared to Europe and United states due to procurement and shipping complexities, whereas estimated labor costs are lower. Therefore, the direct cost savings discussed in the paper are specific for the Georgian setting but this approach is likely to result in cost savings in other settings as well.

Based on our findings, we propose “The Georgian customized 10 MIRU-VNTR loci set”, to be used for molecular epidemiological studies in the country of Georgia as a primary typing tool for identifying major clusters or defining relapse/reinfection. Moreover, the additional 14 loci could be used for further discrimination in clusters defined based on the initial typing. Such a stepwise approach using two typing sets will reduce time, cost and workload, while maintaining high discriminatory power and reproducibility in the most practical way.

Our study has several limitations. Firstly, our sample size was moderate. However, the study covered all the available samples from the penitentiary system and the combination with previous dataset from the general population had increased the variation within the sample set. Secondly, even though MIRU-VNTR typing can produce sufficient data for specific questions, if higher resolution is needed, this method can be limiting, especially in settings with a large proportion of lineage2/Beijing strains [7]. Ideally, whole genome sequencing (WGS) would be the method of choice, however it requires the high-level facility with sequencing instruments which currently are not available in Georgia. Even more importantly, the bioinformatics expertise is not widely available in Georgia, which increases complexity of using the WGS methodology. Tentatively, compared to WGS pricing in Europe and United States, the cost of the sequencing in Georgia would be much higher than MIRU-VNTR. Our study results combined with previous data showed that usage of the combination of a customized set of MIRU-VNTR loci and a stepwise approach gives useful information in the most feasible and cost-effective way.

Although conventional MIRU-VNTR is still a widely used tool, this method may still overestimate Mtb recent transmission events as confirmed by several studies [22–24]. The next

decade is likely to witness transition to sequencing technologies due to considerable decrease in sequencing cost [25]. Additionally, decline of the TB burden within the country would give a significant advantage to these technologies to be used for not only epidemiological investigations, but for diagnosis and drug resistance determination as well. However, until this transition occurs in developing countries, MIRU-VNTR typing based on a customized stepwise approach could be valuable addition to the surveillance tools used in Georgia to enhance the control and prevention of TB.

## Author Contributions

**Conceptualization:** Nino Maghradze, Levan Jugheli, Sonia Borrell, Russell R. Kempker, Henry M. Blumberg, Sebastien Gagneux.

**Data curation:** Nino Maghradze, Levan Jugheli, Sonia Borrell, Nestani Tukvadze.

**Formal analysis:** Nino Maghradze, Levan Jugheli, Sonia Borrell, Sebastien Gagneux.

**Funding acquisition:** Nestani Tukvadze, Henry M. Blumberg, Sebastien Gagneux.

**Methodology:** Nino Maghradze, Levan Jugheli, Sonia Borrell, Sebastien Gagneux.

**Writing – original draft:** Nino Maghradze, Levan Jugheli, Sonia Borrell, Sebastien Gagneux.

**Writing – review & editing:** Nino Maghradze, Levan Jugheli, Sonia Borrell, Nestani Tukvadze, Henry M. Blumberg, Sebastien Gagneux.

## References

1. Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN. Molecular Epidemiology of Tuberculosis: Current Insights. *Clinical Microbiology Reviews*. 2006; 19(4):658–85. <https://doi.org/10.1128/CMR.00061-05> PMID: 17041139
2. Comas I, Homolka S, Niemann S, Gagneux S. Genotyping of Genetically Monomorphic Bacteria: DNA Sequencing in *Mycobacterium tuberculosis* Highlights the Limitations of Current Methodologies. *PLOS ONE*. 2009; 4(11):e7815. <https://doi.org/10.1371/journal.pone.0007815> PMID: 19915672
3. Garcia De Viedma D, Perez-Lago L. The Evolution of Genotyping Strategies To Detect, Analyze, and Control Transmission of Tuberculosis. *Microbiol Spectr*. 2018; 6(5).
4. Gagneux S. Strain Variation in the *Mycobacterium tuberculosis* Complex: Its Role in Biology, Epidemiology and Control: Springer International Publishing; 2017. 314 p.
5. Coscolla M, Gagneux S. Consequences of genomic diversity in *Mycobacterium tuberculosis*. *Semin Immunol*. 2014; 26(6):431–44. <https://doi.org/10.1016/j.smim.2014.09.012> PMID: 25453224
6. Ngabonziza JCS, Loiseau C, Marceau M, Jouet A, Menardo F, Tzafadia O, et al. A sister lineage of the *Mycobacterium tuberculosis* complex discovered in the African Great Lakes region. *Nat Commun*. 2020; 11(1):2917. <https://doi.org/10.1038/s41467-020-16626-6> PMID: 32518235
7. Mokrousov I, Narvskaya O, Limeschenko E, Vyazovaya A, Otten T, Vyshnevskiy B. Analysis of the Allelic Diversity of the Mycobacterial Interspersed Repetitive Units in *Mycobacterium tuberculosis* Strains of the Beijing Family: Practical Implications and Evolutionary Considerations. *Journal of Clinical Microbiology*. 2004; 42(6):2438–44. <https://doi.org/10.1128/JCM.42.6.2438-2444.2004> PMID: 15184416
8. Weerasekera D, Pathirane H, Madegedara D, Dissanayake N, Thevanesam V, Magana-Arachchi DN. Evaluation of the 15 and 24-loci MIRU-VNTR genotyping tools with spoligotyping in the identification of *Mycobacterium tuberculosis* strains and their genetic diversity in molecular epidemiology studies. *Infect Dis (Lond)*. 2019; 51(3):206–15. <https://doi.org/10.1080/23744235.2018.1551619> PMID: 30689510
9. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsche-Gerdes S, Willery E, et al. Proposal for Standardization of Optimized Mycobacterial Interspersed Repetitive Unit-Variable-Number Tandem Repeat Typing of *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology*. 2006; 44(12):4498–510. <https://doi.org/10.1128/JCM.01392-06> PMID: 17005759
10. Asante-Poku A, Nyaho MS, Borrell S, Comas I, Gagneux S, Yeboah-Manu D. Evaluation of Customised Lineage-Specific Sets of MIRU-VNTR Loci for Genotyping *Mycobacterium tuberculosis* Complex Isolates in Ghana. *PLOS ONE*. 2014; 9(3):e92675. <https://doi.org/10.1371/journal.pone.0092675> PMID: 24667333

11. Rasoahanimitalisoa R, Rakotosamimanana N, Stucki D, Sola C, Gagneux S, Rasolofo Razanamparany V. Evaluation of spoligotyping, SNPs and customised MIRU-VNTR combination for genotyping *Mycobacterium tuberculosis* clinical isolates in Madagascar. *PloS One*. 2017; 12(10):e0186088. <https://doi.org/10.1371/journal.pone.0186088> PMID: 29053711
12. San LL, Aye KS, Oo NAT, Shwe MM, Fukushima Y, Gordon SV, et al. Insight into multidrug-resistant Beijing genotype *Mycobacterium tuberculosis* isolates in Myanmar. *Int J Infect Dis*. 2018; 76:109–19. <https://doi.org/10.1016/j.ijid.2018.06.009> PMID: 29936318
13. Devi KR, Bhutia R, Bhowmick S, Mukherjee K, Mahanta J, Narain K. Genetic Diversity of *Mycobacterium tuberculosis* Isolates from Assam, India: Dominance of Beijing Family and Discovery of Two New Clades Related to CAS1\_Delhi and EAI Family Based on Spoligotyping and MIRU-VNTR Typing. *PLoS One*. 2015; 10(12):e0145860. <https://doi.org/10.1371/journal.pone.0145860> PMID: 26701129
14. Esen N, Ozkütük A, Coban H, Atlas E. [Three step MIRU-VNTR for routine mycobacteriology laboratory practice]. *Mikrobiyoloji bulteni*. 2010; 44(3):375–83. PMID: 21063987
15. Maghradze N, Jugheli L, Borrell S, Tukvadze N, Aspindzelashvili R, Avaliani Z, et al. Classifying recurrent *Mycobacterium tuberculosis* cases in Georgia using MIRU-VNTR typing. *PLoS One*. 2019; 14(10):e0223610. <https://doi.org/10.1371/journal.pone.0223610> PMID: 31626647
16. Gegia M, Kalandadze I, Madzgharashvili M, Furin J. DEVELOPING A HUMAN RIGHTS-BASED PROGRAM FOR TUBERCULOSIS CONTROL IN GEORGIAN PRISONS. *Health Hum Rights*. 2011; 13(2):E73–E81. PMID: 22773034
17. Droznin M, Johnson A, Johnson AM. Multidrug resistant tuberculosis in prisons located in former Soviet countries: A systematic review. *PloS One*. 2017; 12(3):e0174373. <https://doi.org/10.1371/journal.pone.0174373> PMID: 28334036
18. Gygli SM, Loiseau C, Jugheli L, Adamia N, Trauner A, Reinhard M, et al. Prisons as ecological drivers of fitness-compensated multidrug-resistant *Mycobacterium tuberculosis*. *Nat Med*. 2021.
19. Shamputa IC, Jugheli L, Sadradze N, Willery E, Portaels F, Supply P, et al. Mixed infection and clonal representativeness of a single sputum sample in tuberculosis patients from a penitentiary hospital in Georgia. *Respir Res*. 2006; 7(1):99. <https://doi.org/10.1186/1465-9921-7-99> PMID: 16846493
20. Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *Journal of Clinical Microbiology*. 1988; 26(11):2465–6. <https://doi.org/10.1128/jcm.26.11.2465-2466.1988> PMID: 3069867
21. Hanekom M, van der Spuy GD, Gey van Pittius NC, McEvoy CR, Hoek KG, Ndabambi SL, et al. Discordance between mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing and IS6110 restriction fragment length polymorphism genotyping for analysis of *Mycobacterium tuberculosis* Beijing strains in a setting of high incidence of tuberculosis. *J Clin Microbiol*. 2008; 46(10):3338–45. <https://doi.org/10.1128/JCM.00770-08> PMID: 18716230
22. Wyllie DH, Davidson JA, Grace Smith E, Rathod P, Crook DW, Peto TEA, et al. A Quantitative Evaluation of MIRU-VNTR Typing Against Whole-Genome Sequencing for Identifying *Mycobacterium tuberculosis* Transmission: A Prospective Observational Cohort Study. *EBioMedicine*. 2018; 34:122–30. <https://doi.org/10.1016/j.ebiom.2018.07.019> PMID: 30077721
23. Meehan CJ, Moris P, Kohl TA, Pecerska J, Akter S, Merker M, et al. The relationship between transmission time and clustering methods in *Mycobacterium tuberculosis* epidemiology. *EBioMedicine*. 2018; 37:410–6. <https://doi.org/10.1016/j.ebiom.2018.10.013> PMID: 30341041
24. Stucki D, Ballif M, Egger M, Furrer H, Altpeter E, Battegay M, et al. Standard Genotyping Overestimates Transmission of *Mycobacterium tuberculosis* among Immigrants in a Low-Incidence Country. *Journal of Clinical Microbiology*. 2016; 54(7):1862–70. <https://doi.org/10.1128/JCM.00126-16> PMID: 27194683
25. Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, et al. Whole genome sequencing of *Mycobacterium tuberculosis*: current standards and open issues. *Nat Rev Microbiol*. 2019; 17(9):533–45. <https://doi.org/10.1038/s41579-019-0214-5> PMID: 31209399