



HAL
open science

Analysis of Mycobacterium tuberculosis genetic lineages circulating in Riga and Riga region, Latvia, isolated between 2008 and 2012

Ilva Pole, Julija Trofimova, Inga Norvaisa, Philip Supply, Girts Skenders, Anda Nodieva, Iveta Ozere, Vija Riekstina, Viktorija Igunnova, Jelena Storozenko, et al.

► To cite this version:

Ilva Pole, Julija Trofimova, Inga Norvaisa, Philip Supply, Girts Skenders, et al.. Analysis of Mycobacterium tuberculosis genetic lineages circulating in Riga and Riga region, Latvia, isolated between 2008 and 2012. *Infection, Genetics and Evolution*, 2020, 78, pp.104126. 10.1016/j.meegid.2019.104126 . hal-03054014

HAL Id: hal-03054014

<https://hal.science/hal-03054014v1>

Submitted on 11 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Manuscript Number: MEEGID-D-19-00518R2

Title: Analysis of Mycobacterium tuberculosis genetic lineages circulating in Riga and Riga region, Latvia, isolated between 2008 and 2012.

Article Type: VSI: St. Petersburg Symposium

Keywords: Molecular epidemiology; genotyping; tuberculosis; drug resistance; Latvia

Corresponding Author: Dr. Renate Ranka, Ph.D.

Corresponding Author's Institution: Latvian Biomedical Research and Study Centre

First Author: Ilva Pole

Order of Authors: Ilva Pole; Jūlija Trofimova; Inga Norvaisa; Philip Supply; Girts Skenders; Anda Nodieva; Iveta Ozere; Vija Riekstina; Viktorija Igumnova; Jelena Storozenko; Inta Jansone; Ludmila Viksna; Renate Ranka, Ph.D.

Abstract: Although the number of new tuberculosis (TB) cases registered per year has decreased by 3-fold between 2001 and 2017 in Latvia, the TB incidence and rates of multidrug resistant TB in this Baltic country remain substantially higher than in most other European countries. Molecular typing methods of Mycobacterium tuberculosis (MTB) play an important role both in clinical studies of the disease and the epidemiological investigations, allowing to describe and characterize the pathogen's population structure and spread of particular genotypes. Aim of this study was to examine the prevalence of MTB lineages in Riga and Riga region of Latvia within a five-year period (2008 - 2012), and to evaluate the discriminatory power (DP) of spoligotyping, standard 24-locus MIRU-VNTR and IS6110-RFLP methods in this setting. The results showed that the main MTB spoligotype families were Beijing (25.3%) and LAM (24.3%), followed by T (22.1%), Ural (11.2%), Haarlem (6.6%) and X superfamily (3.4%). This distribution remained stable over the five consecutive years. 67.6% of MTB isolates were pan-susceptible, and 32.4% were resistant to any drug; multi-drug resistance was found in 5.8% of MTB strains, and 7.6% of MTB isolates were extensively drug-resistant. Drug resistance was associated with SIT1, SIT283 and SIT42 genotypes, while SIT1 and SIT42 were overrepresented among multi drug-resistant MTB strains. Overall, DP of spoligotyping method alone was 0.8953, while DP of both 24-locus MIRU-VNTR analysis and IS6110 RFLP was higher (DP=0.9846 and 0.9927, respectively), mainly due to the improvement of the resolution for the Beijing strains. In conclusion, this work represents the first comprehensive molecular epidemiological description of TB in Latvia, highlighting the high genetic diversity of MTB strains circulating in Riga and Riga region. In combination with detailed epidemiological data this approach was helpful for the in-depth understanding of epidemiological processes in settings where the Next-Gen sequencing is not available as a routine method.

June 21, 2019

Editor-in-Chief

Michel Tibayrenc

Journal of Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases
(MEEGID)

Enclosed, please find our manuscript entitled „*Analysis of Mycobacterium tuberculosis genetic lineages circulating in Riga and Riga region, Latvia, isolated between 2008 and 2012*”.

This article is submitted to the Special issue on St. Petersburg Symposium on TB & Mycobacteria (Edited by Igor Mokrousov).

It was preliminarily discussed with Dr Igor Mokrousov who encouraged this submission and thus may serve as its handling editor.

Tuberculosis is one of the major infectious diseases in Latvia, a country in the Baltic region of the northeastern Europe. Molecular genotyping methods of *Mycobacterium tuberculosis* play an important role both in clinical studies of the disease and in the epidemiological investigations, allowing to describe and characterize pathogen`s population structure. The aim of this study was to examine the prevalence of *M. tuberculosis* lineages in Riga and Riga region of Latvia within the five year period (2008 – 2012), and to evaluate the discrimination power of spoligotyping, MIRU-VNTR and IS6110-RFLP methods for genotyping of *M. tuberculosis* clinical isolates in this setting.

This report for the first time describes the molecular epidemiology of tuberculosis in Latvia.

The corresponding author of the manuscript, Renate Ranka, is a senior researcher in Latvian Biomedical Research and Study Centre.

I certify that this manuscript, or any part of it, has not been published and will not be submitted elsewhere for publication while being considered by the MEEGID.

No competing interests are present.

Sincerely yours,

Corresponding author:

Renate Ranka

Latvian Biomedical Research and Study Centre,

Ratsupites street 1, Riga, Latvia, LV1067.

e-mail: renate_r@biomed.lu.lv

Response to Reviewers.

We thank the Editor and the reviewers for constructive comments that we feel substantially improved the manuscript. We have done our best to change the manuscript according to the recommendations.

Please note, that the numbering of tables and supplementary tables was changed as new data were added to the manuscript.

Reviewers' comments:

Reviewer #1: Very good and sound paper. There will be one of the significant epidemiological works, with an excellent sampling.

My comments are below.

1. It is a pity that there is no information on drug sensitivity, although this would undoubtedly strengthen the article and shed light on the epidemiological significance of the identified genotypes (spoligotypes, clusters, RFLP types).

Moreover, the introduction gives such a characteristic: "Latvia is among the multidrug-resistant tuberculosis (MDR-TB) middle burden countries in the world: there were 36 (7.5%) notified. MDR TB cases in 2017 (data from The Center for Disease Prevention and Control of Latvia)."

and the discussion mentions the association of the Beijing genotype with resistance, including reference to earlier studies in Latvia ("In 281 Latvia, too, Beijing genotype was prevalent among MDR-TB patients (Tracevska et al., 2003)."

Thus a comparison would allow to find out what proportion in this study are MDR strains, of which genotypes, clusters, and dynamics can be traced from 2008 to 2012.

Our response.

The information on drug resistance for the studied isolates was added (Table 4).

Sections of the manuscript (Abstract, Materials and Methods, Results, Discussion) were changed accordingly.

2. The article does not discuss clusters 94-32 and 100-32 at all. It is interesting if there are any changes in their proportions over the years.

Our response.

The information in Discussion section was added.

However, it was not possible to track the changes of the proportions of these MIRU-VNTR clusters over the years, as this analysis was not performed for all MTB isolates (please see also the answer below).

3. It would also be interesting to know about SIT262 (that "was associated with imprisonment"): how similar were VNTR profiles and DST profiles? Also, take note of recent paper of Sinkov et al. BMC Genomics 2018 about different association with DR of the Ural spoligotypes.

Our response.

Again, unfortunately, MIRU-VNTR was not available for all samples. The obtained data showed that MIRU/VNTR clustering rate for Ural family was not high (17 out of 46 isolates in 5 clusters); only two clusters (4 and 3 isolates) for SIT262 (Suppl. Table 3).

The drug susceptibility testing revealed that 22 of 23 SIT262 isolates were pan-susceptible (Table 4).

The information in the Discussion section was added.

4. The above questions also concern SIT254 LAM, in particular in terms of drug resistance.

Our response.

The information was added (Table 4). The majority of samples were drug-susceptible.

5. I noticed that the studied isolates were also obtained from children (43 aged below 20 years, in table 1). What % of all TB-diagnosed children during this period. Any idea about family contacts? Indeed this may be a subject of a separate article.

Our response:

The information was added in the Results section.

In total, within the youngest age group 16 MTB isolates were from children; no infectious source case was identified for any of these cases. All other culture-positive pediatric cases for whom epidemiological links were identified were included in a separate study (this study is in progress).

6. It would be interesting to track the dynamics of the number of clusters (VNTR) in different genotypes. Did it change or remained stable? Or at least make a table by years of main spoligotypes, to see if the structure changed.

Our response:

The authors agree that the dynamics of VNTR clusters could be interesting to explore. Unfortunately, in this study, the VNTR analysis was not available for all isolates. Only 411 samples, which were obtained mainly in years 2009-2010, were studied to explore the DP of genotyping methods in our settings. As the distribution of these samples within the years is not uniform, the analysis of the dynamics will not be accurate.

The prevalence of main spoligotypes (overall prevalence >1%) by years is presented in new Supplementary Table S2.

Highlights.

- Molecular epidemiology of tuberculosis in Latvia was described for the first time
- High diversity of *M. tuberculosis* strains was found
- Drug resistance of *M. tuberculosis* isolates was analyzed.
- Association between spoligotypes and TB patients` age was found
- Combination of genotyping methods increased the resolution of the genotypes

1 Virtual Special issue on

2 St. Petersburg Symposium on TB & Mycobacteria.

3

4 **Analysis of Mycobacterium tuberculosis genetic lineages circulating in Riga and Riga**
5 **region, Latvia, isolated between 2008 and 2012.**

6

7 **The running title: Molecular epidemiology of tuberculosis in Latvia.**

8

9 Ilva Pole ^{a, b}, Julija Trofimova ^a, Inga Norvaisa ^a, Philip Supply ^c, Girts Skenders ^a, Anda
10 Nodieva^d, Iveta Ozere ^{a, d}, Vija Riekstina ^a, Viktorija Igumnova ^{b, d}, Jelena Storozenko ^{d, e}, Inta
11 Jansone ^b, Ludmila Viksna ^{d, e}, Renate Ranka ^{b, d}

12

13 a - Riga East University Hospital, Centre of Tuberculosis and Lung Diseases;

14 b - Latvian Biomedical Research and Study Centre;

15 c – Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 8204 -

16 CIIL - Center for Infection and Immunity of Lille, F-59000 Lille, France;

17 d - Riga Stradiņš University;

18 e - Riga East University Hospital, Latvian Centre of Infectious Diseases.

19

20 **Corresponding author:**

21 Renate Ranka, PhD, Latvian Biomedical Research and Study Centre, Ratsupites Str. 1, Riga,

22 Latvia, LV-1067. Phone: 371 67808219, Fax: 371 67442407, e-mail: rena_r@biomed.lu.lv

23

24 **Abstract**

25 Although the number of new tuberculosis (TB) cases registered per year has decreased by 3-fold
26 between 2001 and 2017 in Latvia, the TB incidence and rates of multidrug resistant TB in this
27 Baltic country remain substantially higher than in most other European countries. Molecular
28 typing methods of *Mycobacterium tuberculosis* (MTB) play an important role both in clinical
29 studies of the disease and the epidemiological investigations, allowing to describe and
30 characterize the pathogen's population structure and spread of particular genotypes. Aim of this
31 study was to examine the prevalence of MTB lineages in Riga and Riga region of Latvia within a
32 five-year period (2008 – 2012), and to evaluate the discriminatory power (DP) of spoligotyping,
33 standard 24-locus MIRU-VNTR and IS6110-RFLP methods in this setting.

34 The results showed that the main MTB spoligotype families were Beijing (25.3%) and LAM
35 (24.3%), followed by T (22.1%), Ural (11.2%), Haarlem (6.6%) and X superfamily (3.4%). This
36 distribution remained stable over the five consecutive years. 67.6% of MTB isolates were pan-
37 susceptible, and 32.4% were resistant to any drug; multi-drug resistance was found in 5.8% of
38 MTB strains, and 7.6% of MTB isolates were extensively drug-resistant. Drug resistance was
39 associated with SIT1, SIT283 and SIT42 genotypes, while SIT1 and SIT42 were overrepresented
40 among multi drug-resistant MTB strains. Overall, DP of spoligotyping method alone was 0.8953,
41 while DP of both 24-locus MIRU-VNTR analysis and IS6110 RFLP was higher (DP=0.9846 and
42 0.9927, respectively), mainly due to the improvement of the resolution for the Beijing strains.

43 In conclusion, this work represents the first comprehensive molecular epidemiological
44 description of TB in Latvia, highlighting the high genetic diversity of MTB strains circulating in
45 Riga and Riga region. In combination with detailed epidemiological data this approach was

46 helpful for the in-depth understanding of epidemiological processes in settings where the Next-

47 Gen sequencing is not available as a routine method.

48

49 **1. Introduction**

50 The epidemiological situation of tuberculosis (TB) in Latvia, a country in the Baltic
51 region of the northeastern Europe, during the last decade has improved, with 1729 (73.5 per 100
52 000 population) and 483 (24.8 per 100 000 population) new cases registered in year 2001 and
53 2017, respectively. However, this incidence remains substantially above the average incidence in
54 most other European countries. Furthermore, Latvia is among the multidrug-resistant
55 tuberculosis (MDR-TB) middle burden countries in the world: there were 36 (7.5%) notified
56 MDR TB cases in 2017 (data from The Centre for Disease Prevention and Control of Latvia).

57 Molecular typing methods of *Mycobacterium tuberculosis* (MTB) isolates play an
58 important role both in clinical studies of TB and in epidemiological investigations. The
59 description of predominant MTB lineages, diversity of strains and clustering rates is necessary
60 for monitoring the spread of TB in a particular geographic region or epidemiological settings.
61 While several genotyping methods are available to date, all these methods have their own
62 particular advantages and disadvantages in particular settings and they differ with their
63 sensitivity, reproducibility, resolution, and discriminatory power (DP) (Supply et al., 2000; van
64 Soolingen, 2001; Supply et al., 2001; Gori et al., 2005; Supply et al., 2006; reviewed in Merker
65 et al., 2017). Compared to classical typing methods, whole genome sequencing (WGS) method
66 provides the most precise and comprehensive identification and characterization of MTB strain
67 genotypes. However, several barriers for clinical adoption exists including costs and the lack of
68 bioinformatic expertise among clinical microbiologists (Satta et al., 2018). Among the classical
69 typing methods, restriction fragment length polymorphism analysis of IS6110 (IS6110-RFLP),
70 which have been previously known as the first gold standard for MTB strain typing, is time-
71 consuming and needs good quality, purified and concentrated DNA (van Soolingen, 2001). PCR-

72 based methods such as spoligotyping (Kamerbeek et al., 1997) and 24-locus mycobacterial
73 interspersed repetitive unit variable-number tandem repeat (MIRU-VNTR) typing (Supply et al.,
74 2006) are rapid, easily performed, and can be done with very small quantities of crude DNA
75 extracts. Because of its resolution power, generally superior compared to spoligotyping and in
76 the same range as IS6110 RFLP (Supply et al., 2006; De Beer et al., 2013), molecular clustering
77 based on the latter method has been and is still largely used as a proxy to identify recent
78 transmission of MTB strains, when WGS is not accessible.

79 Riga is the capital of Latvia and is also the largest city in the Baltic States. With
80 approximately one million inhabitants (year 2017), Riga city and Riga region is home to ~52%
81 of Latvia's population (Central Statistical Bureau of Latvia, csb.gov.lv.). The aim of this study
82 was to examine the prevalence of MTB lineages in Riga and Riga region of Latvia within the
83 five year period (2008 – 2012), and to evaluate the DP of spoligotyping, MIRU-VNTR and
84 IS6110 RFLP methods for genotyping of MTB clinical isolates in this specific setting. This is
85 the first comprehensive molecular epidemiological study of TB in Latvia.

86

87

88 **2. Materials and methods**

89 **2.1. Study Setting and Design**

90 The Center of Tuberculosis and Lung Diseases, Riga East University Hospital plays a central
91 role in TB management in the country. The Center handles all MTB isolates for the Riga and
92 Riga region, and oversees routine diagnosis, and provides phenotypic drug sensitivity testing.

93 **2.2. Mycobacterial isolates and epidemiological data**

94 The source of the MTB isolates was mainly sputum or induced sputum, and occasionally other
95 biological samples such as urine, faeces, bronchial aspirates, biopsy samples, pleural effusion
96 and operation material, from TB patients living in Riga and Riga region and admitted to Center
97 of Tuberculosis and Lung Diseases, Riga East University Hospital between January 2008 and
98 December 2012. Patient diagnosis was confirmed both microscopically and by culturing on solid
99 Lowenstein-Jensen (LJ) medium for 4–6 weeks. Only one isolate from each patient at time of
100 diagnosis was included in this study.

101 Epidemiological and clinical data for this study was provided by the National TB registry of
102 Latvia.

103

104 **2.3. DNA extraction and genotyping**

105 Mycobacterial chromosomal DNA was isolated from mycobacterial colonies grown on
106 Lowenstein Jensen (LJ) media using Cetyltrimethyl Ammonium Bromide (CTAB) method
107 (van Soolingen et al., 1991). Spoligotyping was used to identify MTB genotype for all
108 available clinical isolates. Spoligotyping was performed using commercially available kits
109 (Isogen Life Science, Netherlands; later Ocimum Biosolutions, India) following previously
110 described standard protocol (Kamerbeek et al., 1997). Spoligotyping results were analysed
111 and compared with those in SITVIT2 MTBC database was used for discrimination of MTB
112 spoligotype families (SIT numbers and families) (Couvin D et al., 2018).

113 24-locus MIRU-VNTR typing was performed using PCR followed by capillary sequencer
114 analysis, as described in (Supply et al., 2006). The 24-allele profiles were analyzed with the
115 tools implemented in the MIRU-VNTRplus online database (Weniger et al., 2010).

116 DNA fingerprinting with IS6110 as a probe was performed by using standardized protocols as
117 described elsewhere (van Embden et al., 1993). The BioNumerics software Version 5.3
118 (Applied Maths NV, Sint-Martens-Latern, Belgium) was used for data analysis.

119 **2.4. Drug susceptibility testing**

120 MTB drug susceptibility testing (DST) for the 1st line (isoniazid, rifampicin, ethambutol and
121 pyrazinamide) and 2nd line drugs (streptomycin, ofloxacin, kanamycin, amikacin,
122 capreomycin, ethionamide, cycloserine; para-aminosalicylic acid) was carried out for all
123 strains using the method of absolute concentrations on the solid LJ medium and Bactec MGIT
124 960 system (Becton Dickinson, Heidelberg, Germany) according to WHO recommendations
125 (WHO, 2018). The laboratory is externally quality assured according to the requirements of
126 standard LVS EN ISO:15189-2013.

127 **2.5. Statistical analysis**

128 The 2-sided Fisher's exact test was employed to determine a *P* value (GraphPad Software, La
129 Jolla, California, USA). A *P* value equal to or less than 0.05 was considered significant. The
130 distribution of the spoligotypes among different age groups was analyzed with principal
131 components analysis (PCA) in R using FactoMineR (Lê et al., 2008). The discriminatory
132 power of each method was calculated by the Hunter-Gaston discriminatory index (HGDI).

133

134 **3. Results**

135 **3.1. Epidemiological data of TB patients**

136 From January 2008 to December 2012, 4241 new TB cases were recorded in Latvia including
137 1625 TB cases (38.3 %) in Riga and Riga region. Among these cases 7.71% (327/4241) were
138 children (age 0 - 17 years); in Riga and Riga region the proportion of pediatric cases was 9.17%
139 (149/1625). In total, in Riga and Riga region, 1308 cases were culture-positive, and 908 MTB
140 clinical isolates were available for this study.

141 Table 1 presents an overview of the demographic and clinical characteristics of culture-positive
142 TB patients included in this study. The male/female ratio was 581/327, the median age was 42
143 (1-89 years old). The majority of patients were 21-60 years old (775 individuals, 85.3%) and
144 were born in Latvia (813 individuals, 89.5%). Within the youngest age group (i.e. individuals 0-
145 20 years old) 16 MTB isolates were from children. Pulmonary TB was diagnosed in 861 (94.8%)
146 cases, while 13 (1.4%) patients had extra-pulmonary TB. Both types of disease were diagnosed
147 in 33 (3.6%) individuals.

148

149 **3.2. MTB isolate diversity by spoligotyping**

150 Spoligotyping analysis of 908 MTB clinical isolates revealed a total of 138 different patterns
151 (DP: 0.9034) (Supplemental Table S1).

152 Of 908 isolates, 821 isolates were divided into 51 clusters (cluster size 2-224 isolates). In total,
153 846 samples (93.2%) showed shared-types or SITs when compared to the SITVIT2 MTBC
154 database. In contrast, only 11 MTB isolates (1.2%) belonged to known orphan spoligotypes,
155 while 51 isolates (5.6% of all samples) corresponded to a total of 38 new spoligotyping patterns
156 that have not yet been reported (Supplemental Table S1). The spoligotyping results showed that

157 almost equal proportions of MTB strains, Latvia, belonged to the Beijing and LAM lineages in
158 Riga and the Riga region (25.4% and 26.5%, respectively) followed by T lineage (20.3%) (Table
159 2). Less prevalent genotypes in Riga region were Haarlem (8.5%), Ural (7.9%) and X
160 superfamily (2.4%), while Central Asian and East-African Indian strains were absent. Other
161 spoligotypes accounted for 8.9% of all MTB isolates analysed. The distribution of the MTB
162 spoligotype families within the 5 year period was further analyzed, and the results are shown in
163 Table 2. A slight fluctuation in the distribution of MTB families within different years was
164 observed, however, this difference was not statistically significant ($P>0.05$). Similarly,
165 fluctuations in the distribution of most prevalent MTB spoligotypes (prevalence $\geq 1\%$) within this
166 study period were not statistically significant (Supplemental Table S2).

167 The Beijing family was represented by four spoligotypes; SIT1 was the dominant spoligotype
168 as it was detected in 224 (24.7%) of all studied clinical isolates. SIT190 was found in only
169 three samples (0.3%), while both SIT255 and SIT265 spoligotypes were detected in 2 samples
170 each (0.2%). By contrast, higher spoligotype diversity was observed for LAM, T, and Haarlem
171 families. The LAM family comprised 241 clinical isolates and was represented by 19
172 spoligotypes including isolates belonging to LAM-RUS spoligotypes SIT254, 264, 496, 1240
173 and 2246. Further, 184 MTB isolates that belonged to the T family showed 32 different
174 spoligotypes, while the Haarlem family consisted of 77 studied isolates and was represented
175 by 13 spoligotypes (Supplemental Table S1). Overall, spoligotypes with a prevalence greater
176 than 10% among all isolates included SIT1 (Beijing, 24.7%), SIT254 (LAM-RUS, 11.9%) and
177 SIT53 (T family, 11.0%), while spoligotypes with a prevalence greater than 2% included
178 SIT42, SIT283, SIT766, SIT262, SIT40 and SIT119; other spoligotypes with a prevalence
179 lower than 2% were represented by 280 (30.8%) MTB isolates in total (Table 1).

180

181 **3.3. Association between spoligotypes and TB patients` characteristics**

182 Possible associations between spoligotypes and TB patients` characteristics were analyzed by
183 PCA analysis followed by statistical evaluation. Upon PCA analysis, age-related differences
184 were interestingly revealed among spoligotype patterns. In a score plot, the spoligotype SIT1
185 as well as spoligotypes with a prevalence less than 2% in the population (collectively named
186 as others in Fig. 1) were clearly separated; on the other hand, the PCA showed very close
187 distances between SIT254, SIT53, SIT42, SIT283, SIT766, SIT262, SIT40 and SIT119 for the
188 age groups studied (Figure 1). Indeed, the prevalence of SIT1 in the individuals who were
189 older than 60 years of age was decreased, while the prevalence of rare spoligotypes in older
190 individuals was significantly increased in comparison with young (age ≤ 20 years old) patients
191 ($P = 0.0374$) (Table 1).

192 In contrast, the data analysis revealed no differences of the distribution of spoligotypes among
193 male and female patients. Similarly, no significant differences were observed between Latvia-
194 born and foreign-born individuals, and for the type of the disease (Table 1).

195

196 **3.4. Association between spoligotypes and TB risk factors**

197 Further, possible association between spoligotypes and TB risk factors was explored. In this
198 study, seven different TB risk factors were analyzed which were included in the National TB
199 registry of Latvia since year 2000: housing status, alcohol abuse, drug abuse, HIV infection,
200 imprisonment, malnutrition and reported contact with a TB patient (Table 3). The majority of
201 patients had a permanent home (92.2%), while 70 (7.7%) patients were homeless. Of the patients
202 for whom HIV status was known (99.7%), 139 (15.3%) individuals were HIV-positive. Alcohol

203 abuse as a risk factor was detected for 270 (29.7%) patients, while 115 (12.7%) individuals were
204 recorded as being in prison, and 79 (8.7%) - using drugs. Malnutrition was recorded for a half of
205 all patients (51.1%), while less than 30% of individuals have reported the contact with TB
206 patient (27.1%). The presence of any of these seven TB risk factors was detected in the majority
207 of patients (76.5%). Overall, the prevalence of the main spoligotypes among individuals with
208 risk factors was similar to those observed in non-risk patients ($P>0.05$) (Table 3). However,
209 statistically significant difference was detected in the prevalence of both SIT1 and SIT53
210 spoligotypes depending on the housing status of the patients: SIT53 was more often present in
211 the homeless persons, and SIT1 was more often present in patients with permanent home
212 ($P=0.0083$ and $P=0.0424$, respectively). In addition, the spoligotype SIT262 was more often
213 detected in former prisoners (7.0% vs 1.9%, $P=0.005$).

214

215 **3.5. Drug resistance analysis**

216 Analysis of drug resistance profiles showed that 67.6% of MTB isolates were pan-susceptible
217 (n=614) and 32.4% were resistant to any drug (n=294) (Table 4). The distribution of
218 spoligotypes among drug resistant isolates was as follows: 43.2% SIT1 (Beijing, n=127), 9.9%
219 SIT42 (LAM family, n=29), 5.8% SIT53 (T family, n=17), 14.6% SIT283 (Haarlem, n=43); all
220 other spoligotypes accounted for 26.5% (n=78) of drug-resistant MTB strains. Spoligotypes with
221 a drug resistance prevalence greater than 30% among the isolates included SIT1 (Beijing,
222 56.7%), SIT42 (LAM, 56.9%), SIT264 (LAM-RUS, 66.7%), SIT283 (Haarlem, 100%) and
223 SIT1292 (Ural, 41.7%); statistically significant difference was reached for SIT1, SIT42 and
224 SIT283 ($P\leq 0.05$; Table 4).

225 In total, MDR and extensively-drug-resistance (XDR) was found in 5.8% (53/908) and 7.6%
226 (69/908) of MTB strains, respectively.

227 The analysis of the distribution of spoligotypes among MDR/XDR isolates showed a significant
228 overrepresentation of both SIT1 and SIT42 isolates: these spoligotypes accounted for 47.2% and
229 17.0% of MDR MTB, and for 55.1% and 34.0% of XDR strains, respectively (Table 4).

230

231 **3.6. Analysis of MTB isolates by different genotyping approaches**

232 From 908 MTB isolates, 411 samples were available for additional *IS6110* RFLP and MIRU-
233 VNTR analyses and thus were used for more in-depth genotypic characterization of MTB

234 isolates in Riga and Riga region. In total, these samples represented 25.3% of all MTB-positive
235 cultures isolated in the 2008 – 2012 time period. Among them, 78 spoligotypes were identified.

236 The isolates belonged to the following MTB families: Beijing (n=104; 25.3%), T (n=91; 22.1%),
237 LAM (n=100; 24.3%), Haarlem (n=27; 6.6%), Ural (n=46; 11.2%) and X (n=14; 3.4%); the
238 remaining 29 (7.1%) represented other families and/or unpublished SITs (Table 2). This overall

239 representation of spoligotypes for these 411 isolates was similar to that observed for the full
240 MTB dataset, indicating that this selection was quite representative of the total strain set.

241 Spoligotyping alone identified 78 genotypes, 50 of which were unique; 87.8% of isolates (361 of
242 411) formed 27 clusters with a cluster range 2-31 samples.

243 Upon 24-locus MIRU-VNTR typing, 242 different MIRU-VNTR patterns were obtained; 213
244 MTB isolates (51.8%) were grouped into 44 clusters with 2 to 37 isolates per cluster, while the
245 remaining 198 samples showed a unique MIRU-VNTR profile (Table 5; Supplemental Table
246 3). Thus, MIRU-VNTR typing had a substantially greater DP (DP=0.9846) in comparison with
247 spoligotyping (DP=0.8953).

248 When 24 MIRU-VNTR-based cluster analysis was performed for the main MTB families
249 separately, the Beijing family isolates (n=104) were divided into 45 different MIRU-VNTR
250 genotypes, forming ten clusters and 35 unique genotypes. According to the MLVA MtbC15-9
251 nomenclature implemented in the MIRU-VNTR Plus database, the 3 most common allelic
252 profiles were: (1) 94-32 (244233352644425153353823) represented by 37 isolates (35.6%);
253 (2) 100-32 (244233352644425173353723) represented by 11 isolates (10.6%); and (3) 170-32
254 (244243352544425153353823) represented by 5 isolates (4.8%) (Table 5; Supplemental Table
255 3). MTB isolates of the LAM lineage (n=100) were divided into 44 MIRU-VNTR genotypes
256 forming ten clusters with 2-25 members and 34 unique genotypes. The 3 most common
257 genetic profiles were following: (1) 121-52 (132244332224125153322622) represented by 25
258 isolates (25.0%); (2) 843-52 (132254332224125153322622) represented by 15 isolates
259 (15.0%); (3) 15664-52 (132244332224125153321622) represented by 7 isolates (7.0%).

260 In contrast, MIRU-VNTR types within the other spoligotype families were relatively more
261 diverse/less clustered. For instance, MTB isolates of T family (n=91) were divided into 61
262 MIRU-VNTR genotypes, forming 11 clusters and 50 unique genotypes. The 3 most common
263 genotypes were represented by 5 (5.5%; 20574-15, 223114322534425153342512), 6 (6.6%;
264 12677-15, 224243122334225143334322) and 7 isolates (7.7%; 20556-15,
265 224243122434225143335322), respectively. Likewise, the largest cluster observed among the
266 isolates part of the Haarlem spoligotype family (n=27) comprised only 4 isolates (14.8%;
267 3527-15, 223235332434425153324732)), while a second cluster included only 2 isolates
268 (7.4%). The Ural MTB family members (N=46) formed five clusters, of which 2 were of four
269 isolates each (163-15, 235237232244425113323632 and 16668-15,
270 215237232244425113323732); three other clusters had three isolates each. The X family

271 members (n=14) have formed three clusters which had only two isolates each (Supplemental
272 Table 3).

273 Next, we applied *IS6110* RFLP typing to all 411 isolates. In total, 257 different patterns were
274 obtained. 207 isolates (50.4%) were grouped into 53 clusters, and the calculated DP for the
275 RFLP genotyping method in our dataset was as high as 0.9927 (Table 5). Importantly, in the
276 case of Beijing strains, *IS6110* RFLP typing showed superior resolution of the genotypes
277 (DP=0.9766) to both spoligotyping and MIRU-VNTR typing (DP=0.0571 and 0.8615,
278 respectively).

279 When all three genotyping methods were combined, only 119 MTB isolates (29.0%) were
280 grouped into 39 clusters with 2 to 12 isolates per one cluster, while the remaining 292 samples
281 have shown a unique RFLP and/or MIRU-VNTR/spoligotype patterns (Supplemental Table
282 4). The largest cluster was composed of 12 MTB isolates belonged to SIT254 (representative
283 of LAM-RUS family). In total, the LAM lineage isolates formed 9 clusters with 2-12 isolates,
284 the Beijing genotype isolates - 9 clusters (2-8 isolates), T family – 11 clusters (2-4 isolates),
285 Ural – four clusters (2-3 isolates), and X – three clusters with 2, 3 and 4 isolates each. In
286 addition, two clusters with two isolates each were obtained for the new spoligotypes
287 (spoligotype 757777777720761 and 477777663760771) (Supplemental Table 4).

288

289 **4. Discussion**

290 Although MTB genotypes were previously studied in association with drug resistance in
291 Latvia (Tracevska et al., 2003; Nodieva et al., 2010; Bauskenieks et al., 2014), this is the first
292 comprehensive report describing the molecular epidemiology of TB in Riga and Riga region,

293 which is home for more than 50% of Latvia's population. The analysis of the prevalence of the
294 main MTB spoligotypes within the five-year time frame revealed the presence of 138
295 spoligotypes. Among them, 38 spoligotypes representing 5.6 percent of the 908 study isolates
296 had not been previously reported to the global SITVIT database. At the other extremity, the 3
297 most common spoligotypes, namely SIT1 (Beijing family), SIT53 (T family) and SIT254
298 (LAM family) accounted altogether for approximately one-half of all MTB isolates, and
299 represented 24.7, 11.0 and 11.9% of isolates, respectively. These patterns are also among the
300 most common SITs in the world (Demay et al., 2012). In addition to be dominant in East Asia,
301 strains with the canonical Beijing SIT1 spoligotype are present in all seven geographic areas
302 of the world (Merker et al, 2015). In Europe, Beijing isolates were previously reported to be
303 very common in the Latvia's eastern and northern neighboring countries, Russia and Estonia,
304 often in association with drug resistance (Toungousova et al., 2002; Kruuner et al., 2001). In
305 Latvia, too, Beijing genotype was prevalent among MDR-TB patients (Tracevska et al., 2003).
306 More recent data indicated that the XDR *M. tuberculosis* strain population in northwestern
307 Russia is heavily dominated by Beijing genotype isolates (Vyazovaya et al., 2015). In Pskov
308 oblast, which is located in northwestern Russia on the Russian/Latvian border, the Beijing
309 genotype was the most prevalent followed by LAM, T, Haarlem, Ural, and Manu2 families
310 (Mokrousov et al., 2012).

311 In this study, in total, 32.4% of MTB isolates were resistant to any drug, and 13.4% were
312 MDR/XDR strains. Unsurprisingly, Beijing genotype was highly associated with drug
313 resistance, as 43.2% of drug-resistant isolates belonged to the spoligotype SIT1. SIT1 was also
314 the predominant genotype among MDR/XDR isolates: approximately half of all MDR and
315 XDR isolates (i.e. 47.2% and 55.1%, respectively) belonged to this spoligotype. While SIT1

316 was also the most common spoligotype in our study, the predominance of this genotype
317 among drug-resistant isolates is not related only to the simple majority. The analysis of drug
318 resistance profiles showed that resistance to any drug and MDR/XDR phenotype was observed
319 in 56.7% and 28.1% of the SIT1 strains, respectively. Both proportions were statistically larger
320 when compared to the proportion of drug-resistant and MDR/XDR strains among the total
321 number of MTB isolates (i.e. 32.4% and 13.4%, respectively). MIRU-VNTR analysis of
322 Beijing isolates in Latvia showed the dominance of MTB strains belonged to clusters 94-32
323 (35.6% of Beijing isolates) and 100-32 (10.6%) which were reported to be the most common
324 cross-border molecular clusters of MDR TB in Europe (ECDC, 2016). Similarly, Beijing
325 clusters 94–32 and 100–32 were the dominant MTB genotypes in Central Asia, and the clonal
326 spread of resistant TB strains, particularly of the Beijing lineage, was highlighted as a root of
327 the MDR-TB epidemic in Central Asia (Engström et al., 2019).

328 However, the distribution of Beijing strains was not solely responsible for the
329 transmission of drug-resistant TB in Latvia. Less prevalent genotypes among drug-resistant
330 isolates were SIT283 (Haarlem, 14.6%), SIT42 (LAM, 9.9%), SIT53 (T family, 5.8%) and
331 SIT264 (LAM-RUS, 2.7%), followed by other spoligotypes with the prevalence less than 2%
332 each. The proportion of drug-resistant isolates among SIT42, SIT264 and SIT283 samples
333 (56.9%, 66.7% and 100%, respectively), as well as the proportion of MDR/XDR isolates
334 among SIT42 samples (52.9%) was significantly increased. Along with Beijing strains, the
335 combination of spoligotyping/MIRU-VNTR analysis revealed high clustering rate for LAM
336 family in Latvia; the two most common types were 121-52 and 843-52, which accounted for
337 25.0% and 15.0% of LAM isolates, respectively. While SIT254 was not associated with drug
338 resistance, as 95.4% of isolates were pan-susceptible in our study, the distribution of other

339 LAM family genotypes, including SIT42 and SIT264, is of high importance. These findings
340 indicate that the high burden of TB drug resistance in Latvia is at least partially related to the
341 circulation of MTB strains of specific genotypes. Thus, genotyping of MTB strains is highly
342 important to identify significant factors involved in the distribution of resistance strains which,
343 in turn, may affect TB control programs.

344 Overall, the prevalence of different genotype families in Latvia seen in this study
345 resembles those reported for the neighboring Estonia, northwestern Russia and, to a smaller
346 extent, Belarus and Ukraine (where Ural genotypes were not observed according to previous
347 reference reports) (Figure 2) (Chernyaeva et al., 2012; Mokrousov et al., 2012; Toit et al.,
348 2014; Zalutskaya et al., 2013; Nikolaevskaya et al., 2016;). Relatively to these Baltic and
349 Eastern countries, the prevalence of Beijing genotypes is substantially lower in Western
350 Europe and Scandinavia, and the proportion of isolates with spoligotype patterns typical of the
351 Euro-American lineage (alias lineage 4) of MTB is archetypally much higher (Figure 2)
352 (Demay et al., 2012; Smit et al., 2013).

353 Interestingly, among the 51 isolates with a total of 38 novel spoligotypes that were
354 discovered among circulating MTB strains in Riga region, six shared the same spoligotype
355 477777663760771, unreported in any other world place thus far. These isolates might thus
356 represent a genotype specific to Latvia. Further studies would help to explore the frequency
357 and underlying clonal/genome diversity of this spoligotype more in details.

358 Importantly as well, the distribution pattern of the most common genotypes with a
359 dominance of Beijing, LAM and T spoligotypes remained stable over the five-year study
360 period, indicating an overall unchanged prevalence and circulation of different MTB strains in
361 the country and society. Similar result was obtained for the distribution of the most common

362 spoligotypes including drug resistance-related SIT1, SIT42 and SIT283. This stability at least
363 within this time span limits contrasts with the substantial changes in MTB strain population
364 structures observed over time in other European countries like Ireland and Germany,
365 associated with human migration and mobility (Roycroft et al., 2018; Andrés et al., 2017).
366 These results is consistent with the fact that the vast majority of TB patients in this study, i.e.
367 89.5%, were local born.

368 While we did not find any associations of spoligotypes with patients` characteristics
369 such as gender, country of birth or type of the disease, our PCA results showed that the SIT1
370 spoligotype is more frequent in 21–50 year-old adults while the rare spoligotypes were more
371 frequent in elderly patients. In keeping with the notion that the majority of infected individuals
372 remain asymptomatic over most or all of their lifetime (Zumla et al. 2013), and many of the
373 TB cases in the elderly result from reactivation of a latent MTB infection acquired at much
374 younger ages (Negin et al. 2015), the association of these rare spoligotypes with older patients
375 might reflect genotypes that were more prevalent in the past but have become less abundant
376 nowadays.

377 We also found that SIT53 (T family, representing a generic spoligotype signature that
378 can be shared by different branches of the Euro-American lineage of *M. tuberculosis*) isolates
379 were more common in homeless patients, while SIT1 spoligotype (Beijing family) was
380 significantly more frequent in patients with permanent home than in homeless persons, and
381 SIT262 (Ural family) was associated with imprisonment. Prisons are settings in which TB
382 transmission occurs and the incidence of TB in prisons is typically several fold higher
383 compared to the general population (reviewed in Biadglegne et al., 2015). Ural family strains
384 have not previously been associated with increased transmissibility (Mokrousov, 2012).

385 However, additional environmental, social, and host related risk factors (like history of
386 previous imprisonment in other (Eastern) countries, of contacts with prisoners with such a
387 history) could facilitate the spread of particular TB strains in these places (Biadglegne et al.,
388 2015). Similarly, low socioeconomic status and enhanced exposure in crowded places such as
389 homeless shelters may have an impact on the transmission dynamics of particular strains (part
390 of the Euro-American lineage in this case) in specific patient groups, such as homeless people.
391 Also, recent report of Sinkov and colleagues highlighted the recent emergency of Ural Clade C
392 (belongs to SIT262) in Eastern European countries Moldova and Belarus which was
393 significantly associated with XDR or pre-XDR status (Sinkov et al., 2019). In our study, the
394 majority of SIT262 isolates were pan-susceptible (22 of 23), while drug resistant strains
395 including those with MDR phenotype were found among isolates belonging to other Ural
396 family spoligotypes SIT35, 1292 and 3344. These findings demonstrated the circulation of
397 Ural strains with concern for drug resistance and the need for detection and surveillance of
398 these strains in Eurasia to prevent further dissemination.

399 Consistently with multiple previous studies, our results show that, although it is a rapid
400 and convenient genotyping method for broad identification of genotype families in clinical
401 strains (Gori et al., 2005) (Millet et al., 2007), spoligotyping cannot be used alone in order to
402 investigate the possible epidemiological links between patients, especially in settings
403 dominated by the Beijing lineage. Isolates with Beijing spoligotypes were the most common in
404 our study, resulting in a spoligotyping clustering rate of 97.1% (101/104 isolates) and a
405 calculated DP for as low as 0.057 for these isolates. Unsurprisingly as well, the 24-locus
406 MIRU-VNTR analysis clearly increased the overall resolution of the obtained MTB genotypes
407 (calculated DP=0.9846), and especially among isolates with a Beijing spoligotype with a

408 resulting clustering rate of 66.4% (69/104) among them. Likewise, IS6110 RFLP typing had a
409 much higher resolution in comparison with spoligotyping (DP=0.9927), and in the case of
410 Beijing lineage the DP was the highest among the methods applied (DP=0.9766). Given that
411 spoligotyping is both an appropriate routine method for a clinical laboratory in Latvia and a
412 useful screen to identify genotype families, a combination of this method with at least one of
413 the two other molecular genotyping approaches should be used in order to correctly interpret
414 typing results for the needs in epidemiological tracing.

415

416 In conclusions, we report the first comprehensive molecular epidemiological study of TB
417 in Latvia. This study highlighted the high genetic diversity of *M. tuberculosis* strains circulating
418 in Riga and Riga region. The combination of the spoligotyping with the 24-locus MIRU-VNTR
419 and IS6110 RFLP clearly increased the overall resolution of the MTB genotypes. In combination
420 with detailed epidemiological data, this approach was helpful for a first in-depth understanding
421 of epidemiological processes in a settings where the Next-Gen sequencing is not available as a
422 routine method.

423

424

425 **Acknowledgments**

426 This study was supported by the EU 7th Framework TB - PANNET and Latvian National
427 Research program VPP-2010-1/9.3 and VPP “BIOMEDICINE”. P.S. was a consultant for
428 Genoscreen.

429

430

431 **References**

- 432 1. Andrés, M., Göhring-Zwacka, E., Fiebig, L., Priwitzer, M., Richter, E., Rüscher-Gerdes, S.,
433 Haas, W., Niemann, S., Brodhun, B., 2017. Integration of molecular typing results into
434 tuberculosis surveillance in Germany-A pilot study. *PLoS One*. 12(11):e0188356. doi:
435 10.1371/journal.pone.0188356.
- 436 2. Bauskenieks, M., Pole, I., Skenders, G., Jansone, I., Broka, L., Nodieva, A., Ozere, I.,
437 Kalvisa, A., Ranka, R., Baumanis, V., 2015. Genotypic and phenotypic characteristics of
438 aminoglycoside-resistant *Mycobacterium tuberculosis* isolates in Latvia. *Diagn.*
439 *Microbiol. Infect. Dis.* 81(3):177-82. doi: 10.1016/j.diagmicrobio.2014.12.004.
- 440 3. Biadlegne, F., Rodloff, A.C., Sack, U., 2015. Review of the prevalence and drug
441 resistance of tuberculosis in prisons: a hidden epidemic. *Epidemiol. Infect.* 143(5):887-
442 900. doi: 10.1017/S095026881400288X.
- 443 4. Brudey, K., Gordon, M., Moström, P., Svensson, L., Jonsson, B., Sola, C., Ridell, M.,
444 Rastogi, N., 2004. Molecular epidemiology of *Mycobacterium tuberculosis* in western
445 Sweden. *J. Clin. Microbiol.* 42(7):3046-51.
- 446 5. Chernyaeva, E., Dobrynin, P., Pestova, N., Matveeva, N., Zhemkov, V., Kozlov, A.,
447 2012. Molecular genetic analysis of *Mycobacterium tuberculosis* strains spread in
448 different patient groups in St. Petersburg, Russia. *Eur. J. Clin. Microbiol. Infect. Dis.*
449 31(8):1753-7. doi: 10.1007/s10096-011-1497-2.
- 450 6. Couvin, D., David, A., Zozio, T., Rastogi, N., 2018. Macro-geographical specificities of
451 the prevailing tuberculosis epidemic as seen through SITVIT2, an updated version of the
452 *Mycobacterium tuberculosis* genotyping database. *Infect Genet Evol.* S1567-
453 1348(18)30969-9. doi: 10.1016/j.meegid.2018.12.030.

- 454 7. De Beer, J.L., van Ingen, J., de Vries, G., Erkens, C., Sebek, M., Mulder, A., Sloom, R.,
455 van den Brandt, A.M., Enaimi, M., Kremer, K., Supply, P., van Soolingen D., 2013.
456 Comparative study of IS6110 RFLP and VNTR typing of Mycobacterium tuberculosis in
457 the Netherlands, based on a five year nationwide survey. J. Clin. Microbiol. 51(4):1193-
458 8. doi: 10.1128/JCM.03061-12.
- 459 8. Demay, C., Liens, B., Burguière, T., Hill, V., Couvin, D., Millet, J., Mokrousov, I., Sola,
460 C., Zozio, T., Rastogi, N., 2012. SITVITWEB-a publicly available international
461 multimer database for studying Mycobacterium tuberculosis genetic diversity and
462 molecular epidemiology. Infect. Genet. Evol. 12(4):755-66. doi:
463 10.1016/j.meegid.2012.02.004.
- 464 9. Engström, A., Antonenka, U., Kadyrov, A., Kalmambetova, G., Kranzer, K., Merker,
465 M., Kabirov, O., Parpieva, N., Rajabov, A., Sahalchyk, E., Sayfudtinov, Z., Niemann,
466 S., Hoffmann, H., 2019. Population structure of drug-resistant Mycobacterium
467 tuberculosis in Central Asia. BMC Infect Dis. 29;19(1):908.
- 468 10. European Centre for Disease Prevention and Control. Molecular typing for surveillance
469 of multidrug-resistant tuberculosis in the EU/EEA – January 2016. Stockholm: ECDC;
470 2016.
- 471 11. Gori, A., Bandera, A., Marchetti, G., Degli Esposti, A., Catozzi, L., Nardi, G.P., Gazzola,
472 L., Ferrario, G., van Embden, J.D., van Soolingen, D., Moroni, M., Franzetti, F., 2005.
473 Spoligotyping and Mycobacterium tuberculosis. Emerg Infect Dis. 11(8):1242-8.
- 474 12. Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., et al., 1997.
475 Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for
476 diagnosis and epidemiology. J. Clin. Microbiol. 35: 907–914.

- 477 13. Kruuner, A., S. E. Hoffner, H. Sillastu, M. Danilovits, K. Levina, S. B. Svenson, S.
478 Ghebremichael, T. Koivula, and G. Kallenius. 2001. Spread of drug-resistant pulmonary
479 tuberculosis in Estonia. *J. Clin. Microbiol.* 39: 3339–3345.
- 480 14. Lê, S., Josse, J., Husson, F., 2008. FactoMineR: An R Package for Multivariate Analysis.
481 *Journal of Statistical Software.* 25(1):1-18.
- 482 15. Merker, M., Blin, C., Mona, S., Duforet-Frebourg, N., Lecher, S., Willery, E., Blum,
483 M.G., Rüsç-Gerdes, S., Mokrousov, I., Aleksic, E., Allix-Béguet, C., Antierens, A.,
484 Augustynowicz-Kopeć, E., Ballif, M., Barletta, F., Beck, H.P., Barry, C.E., Bonnet, M.,
485 Borroni, E., Campos-Herrero, I., Cirillo, D., Cox, H., Crowe, S., Crudu, V., Diel, R.,
486 Drobniowski, F., Fauville-Dufaux, M., Gagneux, S., Ghebremichael, S., Hanekom, M.,
487 Hoffner, S., Jiao, W.W., Kalon, S., Kohl, T.A., Kontsevaya, I., Lillebæk, T., Maeda, S.,
488 Nikolayevskyy, V., Rasmussen, M., Rastogi, N., Samper, S., Sanchez-Padilla, E., Savic,
489 B., Shamputa, I.C., Shen, A., Sng, L.H., Stakenas, P., Toit, K., Varaine, F., Vukovic, D.,
490 Wahl, C., Warren, R., Supply, P., Niemann, S., Wirth, T., 2015. Evolutionary history and
491 global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nat. Genet.* 47(3):242-
492 9. doi: 10.1038/ng.3195.
- 493 16. Merker, M., Kohl, T.A., Niemann, S., Supply, P., 2017. The Evolution of Strain Typing
494 in the *Mycobacterium tuberculosis* complex. *Adv. Exp. Med. Biol.* 1019:43-78. doi:
495 10.1007/978-3-319-64371-7_3.
- 496 17. Millet, J., Miyagi-Shiohira, C., Yamane, N., Sola, C., Rastogi, N., 2007. Assessment of
497 mycobacterial interspersed repetitive unit-QUB markers to further discriminate the
498 Beijing genotype in a population-based study of the genetic diversity of *Mycobacterium*

- 499 tuberculosis clinical isolates from Okinawa, Ryukyu Islands, Japan. *J. Clin. Microbiol.*
500 45(11):3606-15.
- 501 18. Mokrousov, I., 2012. The quiet and controversial: Ural family of *Mycobacterium*
502 tuberculosis. *Infect. Genet. Evol.* 12(4):619-29. doi: 10.1016/j.meegid.2011.09.026.
- 503 19. Mokrousov, I., Vyazovaya, A., Otten, T., Zhuravlev, V., Pavlova, E., Tarashkevich, L.,
504 Krishevich, V., Vishnevsky, B., Narvskaya, O., 2012. *Mycobacterium tuberculosis*
505 population in northwestern Russia: an update from Russian-EU/Latvian border region.
506 *PLoS One.* 7(7):e41318.
- 507 20. Mokrousov, I., Vyazovaya, A., Solovieva, N., Sunchalina, T., Markelov, Y., Chernyaeva,
508 E., Melnikova, N., Dogonadze, M., Starkova, D., Vasilieva, N., Gerasimova, A.,
509 Kononenko, Y., Zhuravlev, V., Narvskaya, O., 2015. Trends in molecular epidemiology
510 of drug-resistant tuberculosis in Republic of Karelia, Russian Federation. *BMC*
511 *Microbiol.* 15:279. doi: 10.1186/s12866-015-0613-3.
- 512 21. Negin, J., Abimbola, S., Marais, B.J., 2015. Tuberculosis among older adults--time to take
513 notice. *Int. J. Infect. Dis.* 32():135-7.
- 514 22. Nikolaevskaya, E., Pavlovska, O., Molina, B., Brännberg, P., Marynova, I., Krylova, K.,
515 Ivanitsa, T., Filuk, V., Lacom, A., Dominguez, J., Rzhepishevskaya, O., Prat, C., 2016.
516 Spoligotypes of *Mycobacterium tuberculosis* isolated from tuberculosis patients in
517 Odessa, Ukraine. *European Congress of Clinical Microbiology and Infectious Diseases.*
518 Amsterdam, Netherlands, 9-12 April, 2016. P0155.
- 519 23. Nodieva, A., Jansone, I., Broka, L., Pole, I., Skenders, G., Baumanis, V., 2010. Recent
520 nosocomial transmission and genotypes of multidrug-resistant *Mycobacterium*
521 tuberculosis. *Int. J. Tuberc. Lung Dis.* 14(4):427-33.

- 522 24. Roycroft, E., O'Toole, R.F., Fitzgibbon, M.M., Montgomery, L., O'Meara, M., Downes,
523 P., Jackson, S., O'Donnell, J., Laurenson, I.F., McLaughlin, A.M., Keane, J., Rogers,
524 T.R., 2018. Molecular epidemiology of multi- and extensively-drug-resistant
525 *Mycobacterium tuberculosis* in Ireland, 2001-2014. *J. Infect.* 76(1):55-67. doi:
526 10.1016/j.jinf.2017.10.002.
- 527 25. Satta, G., Lipman, M., Smith, G.P., Arnold, C., Kon, O.M., McHugh, T.D., 2018.
528 *Mycobacterium tuberculosis* and whole-genome sequencing: how close are we to
529 unleashing its full potential? *Clin. Microbiol. Infect.* 24(6):604-609. doi:
530 10.1016/j.cmi.2017.10.030.
- 531 26. Sinkov, V., Ogarkov, O., Mokrousov, I., Bukin, Y., Zhdanova, S., Heysell, S.K., 2019.
532 New epidemic cluster of pre-extensively drug resistant isolates of *Mycobacterium*
533 *tuberculosis* Ural family emerging in Eastern Europe. *BMC Genomics.* 19(1):762. doi:
534 10.1186/s12864-018-5162-3.
- 535 27. Smit, P.W., Haanperä, M., Rantala, P., Couvin, D., Lyytikäinen, O., Rastogi, N., Ruutu,
536 P., Soini, H., 2013. Molecular Epidemiology of Tuberculosis in Finland, 2008-2011.
537 *PLoS One.* 8(12): e85027.
- 538 28. Supply, P., C. Allix, S. Lesjean, M. Cardoso-Oelemann, S. Rusch-Gerdes, E. Willery, E.
539 Savine, P. de Haas, H. van Deutekom, S. Roring, P. Bifani, N. Kurepina, B. Kreiswirth,
540 C. Sola, N. Rastogi, V. Vatin, M. C. Gutierrez, M. Fauville, S. Niemann, R. Skuce, K.
541 Kremer, C. Locht, and D. van Soolingen. 2006. Proposal for standardization of optimized
542 mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of
543 *Mycobacterium tuberculosis* . *J. Clin. Microbiol.* 44:4498–4510.

- 544 29. Supply, P., Mazars, E., Lesjean, S., Vincent, V., Gicquel, B., et al., 2000. Variable human
545 minisatellite-like regions in the Mycobacterium tuberculosis genome. *Mol. Microbiol.* 36:
546 762–771.
- 547 30. Supply, P., Lesjean, S., Savine, E., Kremer, K., van Soolingen, D., Locht, C., 2001.
548 Automated high-throughput genotyping for study of global epidemiology of
549 Mycobacterium tuberculosis based on mycobacterial inter- spersed repetitive units. *J.*
550 *Clin. Microbiol.* 39: 3563–3571.
- 551 31. Toit, K., Altraja, A., Acosta, C.D., Viiklepp, P., Kremer, K., Kummik, T., Danilovič, M.,
552 Van den Bergh, R., Harries, A.D., Supply, P., 2014. A four-year nationwide molecular
553 epidemiological study in Estonia: risk factors for tuberculosis transmission. *Public Health*
554 *Action.* 4(Suppl 2):S34-40. doi: 10.5588/pha.14.0045.
- 555 32. Tounghousova, O. S., P. Sandven, A. O. Mariandyshev, N. I. Nizovtseva, G. Bjune, and
556 D. A. Caugant. 2002. Spread of drug-resistant Mycobacterium tuberculosis strains of the
557 Beijing genotype in the Archangel Oblast, Russia. *J. Clin. Microbiol.* 40:1930–1937.
- 558 33. Tracevska, T., Jansone, I., Baumanis, V., Marga, O., Lillebaek, T., 2003. Prevalence of
559 Beijing genotype in Latvian multidrug-resistant Mycobacterium tuberculosis isolates. *Int.*
560 *J. Tuberc. Lung Dis.* 7(11):1097-103.
- 561 34. Vyazovaya, A., Levina, K., Zhuravlev, V., Viiklepp, P., Kütt, M., Mokrousov, I., 2018.
562 Emerging resistant clones of Mycobacterium tuberculosis in a spatiotemporal context. *J*
563 *Antimicrob. Chemother.* 73(2):325-331. doi: 10.1093/jac/dkx372.
- 564 35. Vyazovaya, A., Mokrousov, I., Zhuravlev, V., Solovieva, N., Otten, T., Vishnevsky, B.,
565 Narvskaya, O., 2015. Dominance of the Beijing genotype among XDR Mycobacterium
566 tuberculosis strains in Russia. *Int. J. Mycobacteriol.* 4 (S1):84-85.

- 567 36. Van Embden, J.D., Cave, M.D., Crawford, J.T., Dale, J.W., Eisenach, K.D., Gicquel, B.,
568 Hermans, P., McAdam, R., Shinnick, T.M., Small, P.M., 1993. Strain identification of
569 *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized
570 methodology. J. Clin. Microbiol. 31: 406-409.
- 571 37. Van Soolingen, D., 2001. Molecular epidemiology of tuberculosis and other
572 mycobacterial infections: main methodologies and achievements. J. Intern. Med. 249 (1):
573 p. 1-26.
- 574 38. van Soolingen, D., Hermans, P.W., de Haas, P.E., Soll, D.R., van Embden, J.D., 1991.
575 Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex
576 strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in
577 the epidemiology of tuberculosis. J. Clin. Microbiol. 29(11):2578-86.
- 578 39. Weniger, T., Krawczyk, J., Supply, P., Niemann, S., Harmsen, D., 2010. MIRU-
579 VNTRplus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis*
580 complex bacteria. Nucleic Acids Res. 38(Web Server issue):W326-31. doi:
581 10.1093/nar/gkq351.
- 582 40. World Health Organization (WHO). Technical manual for drug susceptibility testing of
583 medicines used in the treatment of tuberculosis. Geneva: World Health Organization;
584 2018.
- 585 41. Zalutskaya, A., Wijkander, M., Jureen, P., Skrahina, A., Hoffner, S., 2013. Multidrug-
586 resistant *Myobacterium tuberculosis* caused by the Beijing genotype and a specific T1
587 genotype clone (SIT No. 266) is widely transmitted in Minsk. Int. J. Mycobacteriol.
588 2(4):194-8. doi: 10.1016/j.ijmyco.2013.08.001. Epub 2013 Sep 6.

589 42. Zumla, A., Raviglione, M., Hafner ,R., von Reyn, C.F., 2013. Tuberculosis. N. Engl. J.
590 Med. 368(8):745-55. doi: 10.1056/NEJMra1200894.

591

592

593 **Figure captures**

594 **Figure 1.** Principal component analysis (PCA) of the *M. tuberculosis* spoligotypes in age groups.

595 PCA map showed remarkable segregation of SIT1 and spoligotypes with a prevalence less than

596 2% in the population (collectively named as Others).

597

598 **Figure 2.** Distribution of *M. tuberculosis* lineages in Latvia and neighboring countries. Data

599 were obtained: Latvia, this study; Russia, Mokrousov et al., 2015; Estonia, Vyazovaya et al.,

600 2018; Finland, Smit et al., 2013; Sweden, Brudey et al., 2004; Belarus, Zalutskaya et al., 2013;

601 Ukraine, Nikolaevskaya et al., 2016; Poland, SITWIT database; Germany, SITVIT database;

602 Denmark, SITVIT database.

603

Table 2. Prevalence of MTB spoligotype families in Riga and Riga region, 2008 – 2012.

Year	Registered NEW TB cases in Riga and Riga region (MTB-positive cultures)	MTB isolates in this study No. (%)*	Spoligotype family, No. of strains (%)						
			Beijing	LAM	T	Haarlem	Ural	X	Others
2008	354 (288)	181 (62.8)	41 (22.7)	50 (27.6)	38 (21.0)	23 (12.7)	9 (5.0)	2 (1.1)	18 (9.9)
2009	305 (253)	212 (83.8)	55 (25.9)	42 (19.8)	41 (19.3)	14 (6.6)	27 (12.7)	9 (4.3)	24 (11.3)
2010	305 (245)	183 (74.7)	40 (21.9)	52 (28.4)	37 (20.2)	14 (7.7)	17 (9.3)	5 (2.7)	18 (9.8)
2011	315 (250)	180 (72.0)	50 (27.8)	54 (30.0)	37 (20.6)	16 (8.9)	9 (5.0)	4 (2.2)	10 (5.6)
2012	346 (272)	152 (56.0)	45 (29.6)	43 (28.3)	31 (20.4)	10 (6.6)	10 (6.6)	2 (1.3)	11 (7.2)
Total	1625 (1308)	908 (69.4)	231 (25.4)	241 (26.5)	184 (20.3)	77 (8.5)	72 (7.9)	22 (2.4)	81 (8.9)
Set of samples **	-	411 (25.3)	104 (25.3)	100 (24.3)	91 (22.1)	27 (6.6)	46 (11.2)	14 (3.4)	29 (7.1)

* The proportion of the MTB-positive cultures.

** Set of MTB samples which were genotyped by all three genotyping methods (Spoligotyping, 24-loci MIRU-VNTR and IS6110 RFLP (n=411)).

Table 3. Spoligotyping of 908 M. tuberculosis clinical samples and association between main spoligotypes and patients` risk factors.

Risk factor		MTB isolates, No. (%)	Spoligotype, No. of strains (%)									
			SIT1 N=224 (24.7)	SIT254 N=108 (11.9)	SIT53 N=100 (11.0)	SIT42 N=51 (5.6)	SIT283 N=43 (4.7)	SIT766 N=35 (3.9)	SIT262 N=23 (2.5)	SIT40 N=23 (2.5)	SIT119 N=21 (2.3)	Others N=280 (30.8)
Housing status	Permanent home	837 (92.2)	214 (25.6)*	100 (12.0)	85 (10.2)**	48 (5.7)	41 (4.9)	31 (3.7)	20 (2.4)	21 (2.5)	17 (2.0)	260 (31.1)
	Homeless	70 (7.7)	10 (14.3)	8 (11.4)	15 (21.4)	3 (4.3)	2 (2.9)	4 (5.7)	3 (4.3)	1 (1.4)	4 (5.7)	20 (28.6)
	Unknown	1 (0.1)										
Alcohol abuse	Yes	270 (29.7)	67 (24.8)	24 (8.9)	35 (13.0)	13 (4.8)	12 (4.4)	11 (4.1)	10 (3.7)	6 (2.2)	8 (3.0)	84 (31.1)
	No	637 (70.2)	157 (24.7)	84 (13.2)	65 (10.2)	38 (6.0)	31 (4.9)	24 (3.8)	13 (2.0)	17 (2.7)	13 (2.0)	195 (30.6)
	Unknown	1 (0.1)										
Drug abuse	Yes	79 (8.7)	18 (22.8)	15 (19.0)	6 (7.6)	3 (3.8)	3 (3.8)	2 (2.5)	4 (5.1)	2 (2.5)	3 (3.8)	23 (29.1)
	No	828 (91.2)	206 (24.9)	93 (11.2)	94 (11.4)	48 (5.8)	40 (4.8)	33 (4.0)	19 (2.3)	21 (2.5)	18 (2.2)	256 (30.9)
	Unknown	1 (0.1)										
HIV positive	Yes	139 (15.3)	38 (27.3)	23 (16.6)	10 (7.2)	8 (5.8)	8 (5.8)	3 (2.2)	5 (3.6)	2 (1.4)	5 (3.6)	37 (26.6)
	No	768 (84.6)	186 (24.2)	85 (11.1)	90 (11.7)	43 (5.6)	35 (4.6)	32 (4.2)	18 (2.3)	21 (2.7)	16 (2.1)	242 (31.5)
	Unknown	1 (0.1)										
Imprisonment	Yes	115 (12.7)	21 (18.3)	15 (13.0)	13 (11.3)	5 (4.4)	6 (5.2)	4 (3.5)	8 (7.0)***	3 (2.6)	3 (2.6)	37 (32.2)
	No	792 (87.2)	203 (25.6)	93 (11.7)	87 (11.0)	46 (5.8)	37 (4.7)	31 (3.9)	15 (1.9)	20 (2.5)	18 (2.3)	242 (30.6)
	Unknown	1 (0.1)										
Malnutrition	Yes	464 (51.1)	116 (25.0)	50 (10.8)	46 (9.9)	25 (5.4)	20 (4.3)	19 (4.1)	14 (3.0)	12 (2.6)	13 (2.8)	149 (32.1)
	No	442 (48.7)	108 (24.4)	58 (13.1)	54 (12.2)	26 (5.9)	23 (5.2)	16 (3.6)	9 (2.0)	11 (2.5)	8 (1.8)	129 (29.2)
	Unknown	2 (0.2)										
Reported contact with a TB patient	Yes	246 (27.1)	67 (27.2)	30 (12.2)	28 (11.4)	15 (6.1)	12 (4.9)	11 (4.5)	3 (1.2)	7 (2.9)	8 (3.3)	65 (26.4)
	No	661 (72.8)	157 (23.8)	78 (11.8)	72 (10.9)	36 (5.5)	31 (4.7)	24 (3.6)	20 (3.0)	16 (2.4)	13 (2.0)	214 (32.4)
	Unknown	1 (0.1)										
Any of 7 risk factors	Yes	695 (76.5)	175 (25.2)	82 (11.8)	76 (10.9)	41 (5.9)	32 (4.6)	30 (4.3)	17 (2.5)	16 (2.3)	20 (2.9)	206 (29.6)
	No	212 (23.4)	49 (23.1)	26 (12.3)	24 (11.3)	10 (4.7)	11 (5.2)	5 (2.4)	6 (2.8)	7 (3.3)	1 (0.5)	73 (34.4)

* This difference was statistically significant: $P=0.0424$; ** $P=0.0083$; *** $P=0.0050$.

Table 4. Drug susceptibility of MTB isolates of the most common spoligotypes (prevalence greater than 1%) in Riga and Riga region, 2008-2012 (n=908).

SIT**	Spoligotype family	Isolates in study, No. (%)	Pan-susceptible, No (%)	Drug-resistant*, No (%)	MDR, No (%)	XDR, No (%)	Pan-susceptible/drug resistant* proportions within SIT, %	MDR/XDR proportions within SIT, %
1	Beijing	224 (24.7)	97 (15.8)	127 (43.2)	25 (47.2)	38 (55.1)	43.3/56.7	11.2/17.0
20	LAM	10 (1.1)	9 (1.5)	1 (0.3)	1 (1.9)	0	90.0/10.0	10/0
35	Ural	12 (1.3)	9 (1.5)	3 (1.0)	2 (3.8)	0	75.0/25.0	16.7/0
40	T	23 (2.5)	22 (3.6)	1 (0.3)	0	0	95.7/4.3	0/0
42	LAM	51 (5.6)	22 (3.6)	29 (9.9)	9 (17.0)	18 (34.0)	43.1/56.9	17.6/35.3
47	Haarlem	9 (1.0)	8 (1.3)	1 (0.3)	0	0	88.9/11.1	0/0
50	Haarlem	11 (1.2)	8 (1.3)	3 (1.0)	0	0	72.7/27.3	0/0
53	T	100 (11.0)	83 (13.5)	17 (5.8)	4 (7.6)	0	83.0/17.0	4/0
119	X	21 (2.3)	20 (3.3)	1 (0.3)	0	0	95.2/4.8	0/0
254	LAM-RUS	108 (11.9)	103 (16.8)	5 (1.7)	0	1 (1.5)	95.4/4.6	0/0.9
262	Ural	23 (2.5)	22 (3.6)	1 (0.3)	0	0	95.7/4.3	0/0
264	LAM-RUS	12 (1.3)	4 (0.7)	8 (2.7)	1 (1.9)	1 (1.5)	33.3/66.7	8.3/8.3
283	Haarlem	43 (4.7)	0	43 (14.6)	3 (5.7)	0	0/100	7.0/0
766	LAM	35 (3.9)	32 (5.2)	3 (1.0)	0	1 (1.5)	91.4/8.6	0/2.9
1292	Ural	12 (1.3)	7 (1.1)	5 (1.7)	0	0	58.3/41.7	0/0
3344	Ural	9 (1.0)	8 (1.3)	1 (0.3)	0	0	88.9/11.1	0/0
Others		205 (22.6)	160 (26.1)	45 (15.3)	8 (15.1)	10 (14.5)	78.0/22.0	3.9/4.9
Total		908	614	294	53	69	67.6/32.4	5.8/7.6

* Including MDR and XDR isolates

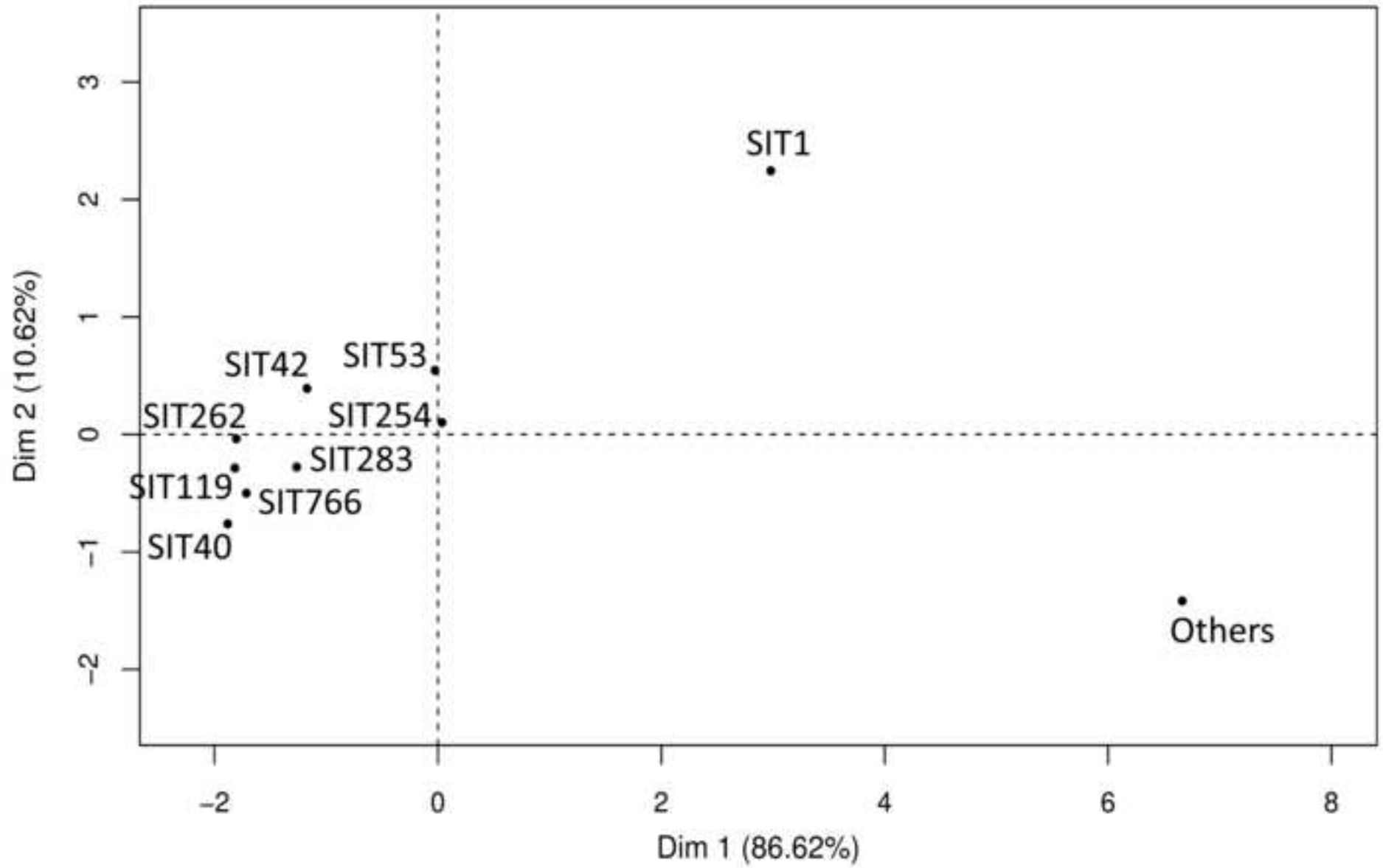
Table 5.

Cluster analysis of *M. tuberculosis* isolates by spoligotyping, 24-locus MIRU-VNTR and IS6110 RFLP genotyping methods.

Genotype (No. of strains)*		Genotyping method		
		Spoligotyping	MIRU-VNTR	IS6110 RFLP
Beijing (104)	No. of clusters	1	10	15
	No. of isolates in clusters	101	69	55
	No. of genotypes	4	45	64
	DP	0.0571	0.8615	0.9776
LAM (100)	No. of clusters	4	10	12
	No. of isolates in clusters	94	66	59
	No. of genotypes	10	44	53
	DP	0.6263	0.9099	0.9459
T family (91)	No. of clusters	8	11	14
	No. of isolates in clusters	80	41	45
	No. of genotypes	19	61	60
	DP	0.6232	0.9829	0.9690
Haarlem (27)	No. of clusters	5	2	2
	No. of isolates in clusters	25	6	10
	No. of genotypes	7	23	19
	DP	0.7692	0.9801	0.9174
Ural (46)	No. of clusters	4	5	6
	No. of isolates in clusters	39	17	22
	No. of genotypes	11	34	30
	DP	0.7913	0.9797	0.9681
X family (14)	No. of clusters	1	3	2
	No. of isolates in clusters	13	6	12
	No. of genotypes	2	11	4
	DP	0.1429	0.967	0.6264
Others (29)	No. of clusters	4	3	2
	No. of isolates in clusters	9	8	4
	No. of genotypes	24	24	27
	DP	0.9852	0.9828	0.9951
TOTAL (411 strains)	No. of clusters	27	44	53
	No. of isolates in clusters	361	213	207
	No. of genotypes	77	242	257
	DP	0.8953	0.9846	0.9927

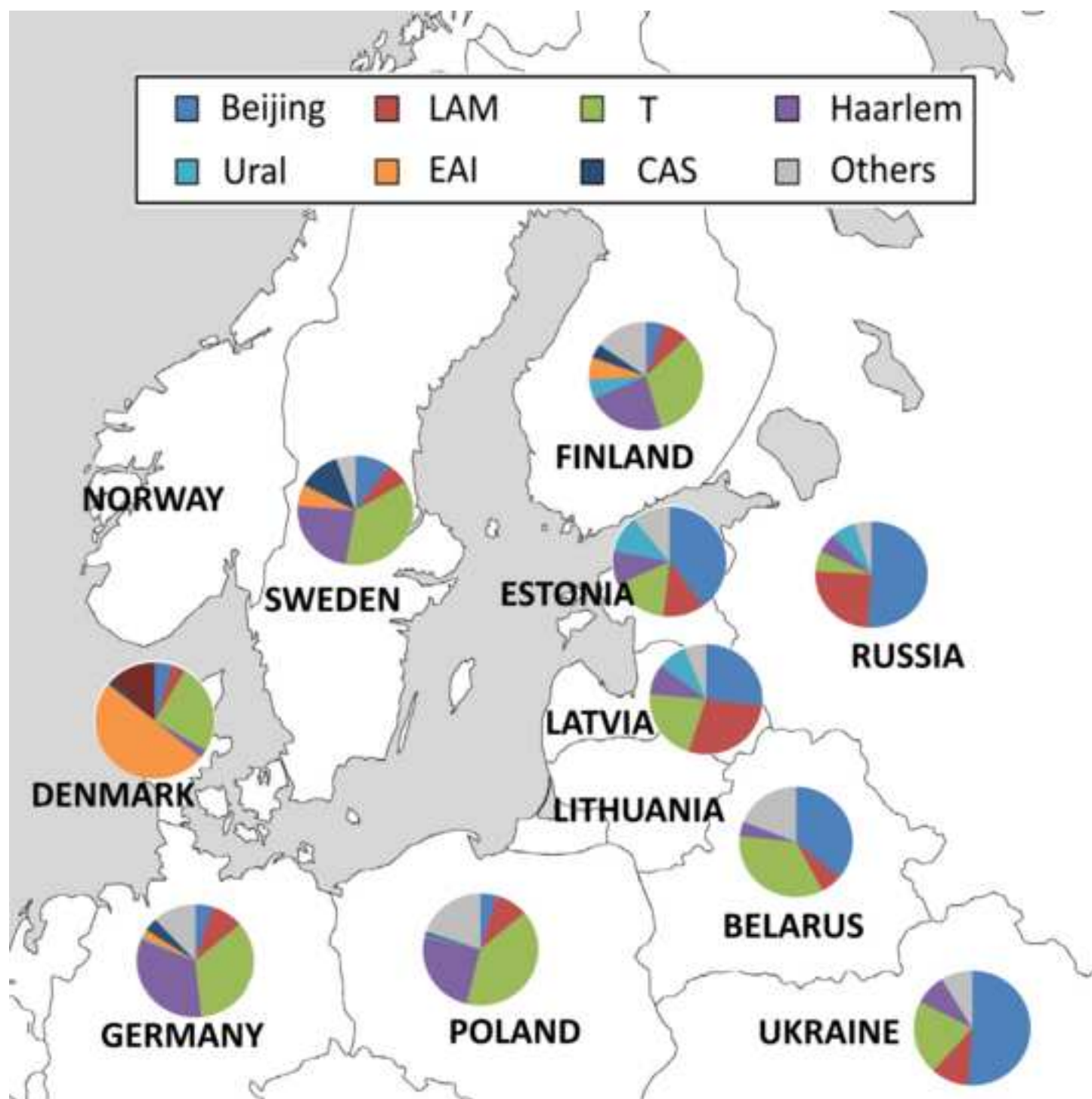
* DP: discrimination power

Factor map (PCA)



Figure(s)

[Click here to download high resolution image](#)



Supplementary Material

[Click here to download Supplementary Material: Supplem_Table_S1_2019____.pdf](#)

Supplementary Material

[Click here to download Supplementary Material: Suppl_Table_S2_new.pdf](#)

Supplementary Material

[Click here to download Supplementary Material: Suppl_Table_S3.pdf](#)

Supplementary Material

[Click here to download Supplementary Material: Suppl_Table_S4.pdf](#)

1 Virtual Special issue on

2 St. Petersburg Symposium on TB & Mycobacteria.

3

4 **Analysis of Mycobacterium tuberculosis genetic lineages circulating in Riga and Riga**
5 **region, Latvia, isolated between 2008 and 2012.**

6

7 **The running title: Molecular epidemiology of tuberculosis in Latvia.**

8

9 Ilva Pole ^{a, b}, Julija Trofimova ^a, Inga Norvaisa ^a, Philip Supply ^c, Girts Skenders ^a, Anda
10 Nodieva^d, Iveta Ozere ^{a, d}, Vija Riekstina ^a, Viktorija Igumnova ^{b, d}, Jelena Storozenko ^{d, e}, Inta
11 Jansone ^b, Ludmila Viksna ^{d, e}, Renate Ranka ^{b, d}

12

13 a - Riga East University Hospital, Centre of Tuberculosis and Lung Diseases;

14 b - Latvian Biomedical Research and Study Centre;

15 c – Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 8204 -

16 CIIL - Center for Infection and Immunity of Lille, F-59000 Lille, France;

17 d - Riga Stradiņš University;

18 e - Riga East University Hospital, Latvian Centre of Infectious Diseases.

19

20 **Corresponding author:**

21 Renate Ranka, PhD, Latvian Biomedical Research and Study Centre, Ratsupites Str. 1, Riga,

22 Latvia, LV-1067. Phone: 371 67808219, Fax: 371 67442407, e-mail: rena_r@biomed.lu.lv

23

24 Abstract

25 Although the number of new tuberculosis (TB) cases registered per year has decreased by 3-fold
26 between 2001 and 2017 in Latvia, the TB incidence and rates of multidrug-resistant TB in this
27 Baltic country remain substantially higher than in most other European countries. Molecular
28 typing methods of *Mycobacterium tuberculosis* play an important role both in clinical studies of
29 the disease and the epidemiological investigations, allowing to describe and characterize the
30 pathogen's population structure and spread of particular genotypes. The aim of this study was to
31 examine the prevalence of *M. tuberculosis* lineages in Riga and Riga region of Latvia within a
32 five-year period (2008–2012), and to evaluate the discriminatory power of spoligotyping,
33 standard 24-locus MIRU-VNTR and IS6110-RFLP methods for genotyping of *M. tuberculosis*
34 clinical isolates in this setting.

35 Analysis of spoligotyping results of 908 isolates showed that the main *M. tuberculosis*
36 spoligotype families were Beijing and LAM (25.3% and 24.3%, respectively) followed by T
37 (22.1%), Ural (11.2%), Haarlem (6.6%) and X superfamily (3.4%). This distribution remained
38 similar over the five consecutive years. Analysis of possible associations between spoligotypes
39 and TB patients' characteristics revealed some age-related differences. Moreover, SIT53 (T) and
40 SIT1 (Beijing) spoligotypes were more often detected among homeless patients and patients with
41 permanent home, respectively ($P=0.0083$ and $P=0.0424$, respectively), and SIT262 was more
42 often identified in former prisoners (7.0% vs 1.9%, $P=0.005$).

43 Spoligotyping, MIRU-VNTR and IS6110-RFLP results were available for a total of 411 *M.*
44 *tuberculosis* clinical isolates. Overall, the discriminatory power of spoligotyping method alone
45 was 0.8953, but was very low in case among Beijing family isolates, as expected. The
46 discriminatory power of both 24-locus MIRU-VNTR analysis and IS6110-RFLP was higher

47 ($DP=0.9846$ and 0.9927 , respectively), mainly due to the improvement of the resolution for the
48 Beijing strains.

49 In conclusion, this work represents the first comprehensive molecular epidemiological
50 description of TB in Latvia, highlighting the high genetic diversity of *M. tuberculosis* strains
51 circulating in Riga and Riga region. In combination with detailed epidemiological data this
52 approach was helpful for the in-depth understanding of epidemiological processes in settings
53 where the Next Gen sequencing is not available as a routine method.

54 Although the number of new tuberculosis (TB) cases registered per year has decreased by 3-fold
55 between 2001 and 2017 in Latvia, the TB incidence and rates of multidrug resistant TB in this
56 Baltic country remain substantially higher than in most other European countries. Molecular
57 typing methods of *Mycobacterium tuberculosis* (MTB) play an important role both in clinical
58 studies of the disease and the epidemiological investigations, allowing to describe and
59 characterize the pathogen's population structure and spread of particular genotypes. Aim of this
60 study was to examine the prevalence of MTB lineages in Riga and Riga region of Latvia within a
61 five-year period (2008 – 2012), and to evaluate the discriminatory power (DP) of spoligotyping,
62 standard 24-locus MIRU-VNTR and IS6110-RFLP methods in this setting.

64 The results showed that the main MTB spoligotype families were Beijing (25.3%) and LAM
65 (24.3%), followed by T (22.1%), Ural (11.2%), Haarlem (6.6%) and X superfamily (3.4%). This
66 distribution remained stable over the five consecutive years. 67.6% of MTB isolates were pan-
67 susceptible, and 32.4% were resistant to any drug; multi-drug resistance was found in 5.8% of
68 MTB strains, and 7.6% of MTB isolates were extensively drug-resistant. Drug resistance was
69 associated with SIT1, SIT283 and SIT42 genotypes, while SIT1 and SIT42 were overrepresented
70 among multi drug-resistant MTB strains. Overall, DP of spoligotyping method alone was 0.8953,

71 while DP of both 24-locus MIRU-VNTR analysis and IS6110 RFLP was higher (DP=0.9846 and
72 0.9927, respectively), mainly due to the improvement of the resolution for the Beijing strains.
73 In conclusion, this work represents the first comprehensive molecular epidemiological
74 description of TB in Latvia, highlighting the high genetic diversity of MTB strains circulating in
75 Riga and Riga region. In combination with detailed epidemiological data this approach was
76 helpful for the in-depth understanding of epidemiological processes in settings where the Next-
77 Gen sequencing is not available as a routine method.

78

79 **1. Introduction**

80 The epidemiological situation of tuberculosis (TB) in Latvia, a country in the Baltic
81 region of the northeastern Europe, during the last decade has improved, with 1729 (73.5 per 100
82 000 population) and 483 (24.8 per 100 000 population) new cases registered in year 2001 and
83 2017, respectively. However, this incidence remains substantially above the average incidence in
84 most other European countries. Furthermore, Latvia is among the multidrug-resistant
85 tuberculosis (MDR-TB) middle burden countries in the world: there were 36 (7.5%) notified
86 MDR TB cases in 2017 (data from The Centre for Disease Prevention and Control of Latvia).

87 Molecular typing methods of *Mycobacterium tuberculosis* (MTB) isolates play an
88 important role both in clinical studies of TB and in epidemiological investigations. The
89 description of predominant MTB lineages, diversity of strains and clustering rates is necessary
90 for monitoring the spread of TB in a particular geographic region or epidemiological settings.
91 While several genotyping methods are available to date, all these methods have their own
92 particular advantages and disadvantages in particular settings and they differ with their
93 sensitivity, reproducibility, resolution, and discriminatory power (DP) (Supply et al., 2000; van
94 Soolingen, 2001; Supply et al., 2001; Gori et al., 2005; Supply et al., 2006; reviewed in Merker
95 et al., 2017). Compared to classical typing methods, whole genome sequencing (WGS) method
96 provides the most precise and comprehensive identification and characterization of MTB strain
97 genotypes. However, several barriers for clinical adoption exists including costs and the lack of
98 bioinformatic expertise among clinical microbiologists (Satta et al., 2018). Among the classical
99 typing methods, restriction fragment length polymorphism analysis of IS6110 (IS6110-RFLP),
100 which have been previously known as the first gold standard for MTB strain typing, is time-
101 consuming and needs good quality, purified and concentrated DNA (van Soolingen, 2001). PCR-

102 based methods such as spoligotyping (Kamerbeek et al., 1997) and 24-locus mycobacterial
103 interspersed repetitive unit variable-number tandem repeat (MIRU-VNTR) typing (Supply et al.,
104 2006) are rapid, easily performed, and can be done with very small quantities of crude DNA
105 extracts. Because of its resolution power, generally superior compared to spoligotyping and in
106 the same range as IS6110 RFLP (Supply et al., 2006; De Beer et al., 2013), molecular clustering
107 based on the latter method has been and is still largely used as a proxy to identify recent
108 transmission of MTB strains, when WGS is not accessible.

109 Riga is the capital of Latvia and is also the largest city in the Baltic States. With
110 approximately one million inhabitants (year 2017), Riga city and Riga region is home to ~52%
111 of Latvia's population (Central Statistical Bureau of Latvia, csb.gov.lv.). The aim of this study
112 was to examine the prevalence of MTB lineages in Riga and Riga region of Latvia within the
113 five year period (2008 – 2012), and to evaluate the DP of spoligotyping, MIRU-VNTR and
114 IS6110 RFLP methods for genotyping of MTB clinical isolates in this specific setting. This is
115 the first comprehensive molecular epidemiological study of TB in Latvia.

116

117

118 **2. Materials and methods**

119 **2.1. Study Setting and Design**

120 The Center of Tuberculosis and Lung Diseases, Riga East University Hospital plays a central
121 role in TB management in the country. The Center handles all MTB isolates for the Riga and
122 Riga region, and oversees routine diagnosis, and provides phenotypic drug sensitivity testing.

123 **2.2. Mycobacterial isolates and epidemiological data**

124 The source of the MTB isolates was mainly sputum or induced sputum, and occasionally other
125 biological samples such as urine, faeces, bronchial aspirates, biopsy samples, pleural effusion
126 and operation material, from TB patients living in Riga and Riga region and admitted to Center
127 of Tuberculosis and Lung Diseases, Riga East University Hospital between January 2008 and
128 December 2012. Patient diagnosis was confirmed both microscopically and by culturing on solid
129 Lowenstein-Jensen (LJ) medium for 4–6 weeks. Only one isolate from each patient at time of
130 diagnosis was included in this study.

131 Epidemiological and clinical data for this study was provided by the National TB registry of
132 Latvia.

133

134 **2.3. DNA extraction and genotyping**

135 Mycobacterial chromosomal DNA was isolated from mycobacterial colonies grown on
136 Lowenstein Jensen (LJ) media using Cetyltrimethyl Ammonium Bromide (CTAB) method
137 (van Soolingen et al., 1991). Spoligotyping was used to identify MTB genotype for all
138 available clinical isolates. Spoligotyping was performed using commercially available kits
139 (Isogen Life Science, Netherlands; later Ocimum Biosolutions, India) following previously
140 described standard protocol (Kamerbeek et al., 1997). Spoligotyping results were analysed
141 and compared with those in SITVIT2 MTBC database was used for discrimination of MTB
142 spoligotype families (SIT numbers and families) (Couvin D et al., 2018).

143 24-locus MIRU-VNTR typing was performed using PCR followed by capillary sequencer
144 analysis, as described in (Supply et al., 2006). The 24-allele profiles were analyzed with the
145 tools implemented in the MIRU-VNTRplus online database (Weniger et al., 2010).

146 DNA fingerprinting with IS6110 as a probe was performed by using standardized protocols as
147 described elsewhere (van Embden et al., 1993). The BioNumerics software Version 5.3
148 (Applied Maths NV, Sint-Martens-Latern, Belgium) was used for data analysis.

149

150 **2.4. Drug susceptibility testing**

151 MTB drug susceptibility testing (DST) for the 1st line (isoniazid, rifampicin, ethambutol and
152 pyrazinamide) and 2nd line drugs (streptomycin, ofloxacin, kanamycin, amikacin,
153 capreomycin, ethionamide, cycloserine; para-aminosalicylic acid) was carried out for all
154 strains using the method of absolute concentrations on the solid LJ medium and Bactec MGIT
155 960 system (Becton Dickinson, Heidelberg, Germany) according to WHO recommendations
156 (WHO, 2018). The laboratory is externally quality assured according to the requirements of
157 standard LVS EN ISO:15189-2013.

158 **2.5.4. Statistical analysis**

159 The 2-sided Fisher's exact test was employed to determine a *P* value (GraphPad Software, La
160 Jolla, California, USA). A *P* value equal to or less than 0.05 was considered significant. The
161 distribution of the spoligotypes among different age groups was analyzed with principal
162 components analysis (PCA) in R using FactoMineR (Lê et al., 2008). The discriminatory
163 power of each method was calculated by the Hunter-Gaston discriminatory index (HGDI).

164

165 **3. Results**

166 **3.1. Epidemiological data of TB patients**

167 From January 2008 to December 2012, 4241 new TB cases were recorded in Latvia including
168 1625 TB cases (38.3 %) in Riga and Riga region. Among these cases 7.71% (327/4241) were
169 children (age 0 - 17 years); in Riga and Riga region the proportion of pediatric cases was 9.17%
170 (149/1625). In total, in Riga and Riga region, 1308 cases were culture-positive, and 908 MTB
171 clinical isolates were available for this study.

172 ~~Among them, 1308 cases were culture positive, and a total of 908 MTB clinical isolates were~~
173 ~~available for this study.~~

174 Table 1 presents an overview of the demographic and clinical characteristics of culture-positive
175 TB patients included in this study. The male/female ratio was 581/327, the median age was 42
176 (1-89 years old). The majority of patients were 21-60 years old (775 individuals, 85.3%) and
177 were born in Latvia (813 individuals, 89.5%). Within the youngest age group (i.e. individuals 0-
178 20 years old) 16 MTB isolates were from children. Pulmonary TB was diagnosed in 861 (94.8%)
179 cases, while 13 (1.4%) patients had extra-pulmonary TB. Both types of disease were diagnosed
180 in 33 (3.6%) individuals.

181

182 **3.2. MTB isolate diversity by spoligotyping**

183 Spoligotyping analysis of 908 MTB clinical isolates revealed a total of 138 different patterns
184 (DP: 0.9034) (Supplemental Table S1).

185 Of 908 isolates, 821 isolates were divided into 51 clusters (cluster size 2-224 isolates). In total,
186 846 samples (93.2%) showed shared-types or SITs when compared to the SITVIT2 MTBC
187 database. In contrast, only 11 MTB isolates (1.2%) belonged to known orphan spoligotypes,

188 while 51 isolates (5.6% of all samples) corresponded to a total of 38 new spoligotyping patterns
189 that have not yet been reported (Supplemental Table S1). The spoligotyping results showed that
190 almost equal proportions of MTB strains, Latvia, belonged to the Beijing and LAM lineages in
191 Riga and the Riga region (25.4% and 26.5%, respectively) followed by T lineage (20.3%) (Table
192 2). Less prevalent genotypes in Riga region were Haarlem (8.5%), Ural (7.9%) and X
193 superfamily (2.4%), while Central Asian and East-African Indian strains were absent. Other
194 spoligotypes accounted for 8.9% of all MTB isolates analysed. The distribution of the MTB
195 spoligotype families within the 5 year period was further analyzed, and the results are shown in
196 Table 2. A slight fluctuation in the distribution of MTB families within different years was
197 observed, however, this difference was not statistically significant ($P>0.05$). Similarly,
198 fluctuations in the distribution of most prevalent MTB spoligotypes (prevalence $\geq 1\%$) within this
199 study period were not statistically significant (Supplemental Table S2).
200 The Beijing family was represented by four spoligotypes; SIT1 was the dominant spoligotype
201 as it was detected in 224 (24.7%) of all studied clinical isolates. SIT190 was found in only
202 three samples (0.3%), while both SIT255 and SIT265 spoligotypes were detected in 2 samples
203 each (0.2%). By contrast, higher spoligotype diversity was observed for LAM, T, and Haarlem
204 families. The LAM family comprised 241 clinical isolates and was represented by 19
205 spoligotypes including isolates belonging to LAM-RUS spoligotypes SIT254, 264, 496, 1240
206 and 2246. Further, 184 MTB isolates that belonged to the T family showed 32 different
207 spoligotypes, while the Haarlem family consisted of 77 studied isolates and was represented
208 by 13 spoligotypes (Supplemental Table S1). Overall, spoligotypes with a prevalence greater
209 than 10% among all isolates included SIT1 (Beijing, 24.7%), SIT254 (LAM-RUS, 11.9%) and
210 SIT53 (T family, 11.0%), while spoligotypes with a prevalence greater than 2% included

211 SIT42, SIT283, SIT766, SIT262, SIT40 and SIT119; other spoligotypes with a prevalence
212 lower than 2% were represented by 280 (30.8%) MTB isolates in total (Table 1).

213

214 **3.3. Association between spoligotypes and TB patients` characteristics**

215 Possible associations between spoligotypes and TB patients` characteristics were analyzed by
216 PCA analysis followed by statistical evaluation. Upon PCA analysis, age-related differences
217 were interestingly revealed among spoligotype patterns. In a score plot, the spoligotype SIT1
218 as well as spoligotypes with a prevalence less than 2% in the population (collectively named
219 as others in Fig. 1) were clearly separated; on the other hand, the PCA showed very close
220 distances between SIT254, SIT53, SIT42, SIT283, SIT766, SIT262, SIT40 and SIT119 for the
221 age groups studied (Figure 1). Indeed, the prevalence of SIT1 in the individuals who were
222 older than 60 years of age was decreased, while the prevalence of rare spoligotypes in older
223 individuals was significantly increased in comparison with young (age ≤ 20 years old) patients
224 ($P = 0.0374$) (Table 1).

225 In contrast, the data analysis revealed no differences of the distribution of spoligotypes among
226 male and female patients. Similarly, no significant differences were observed between Latvia-
227 born and foreign-born individuals, and for the type of the disease (Table 1).

228

229 **3.4. Association between spoligotypes and TB risk factors**

230 Further, possible association between spoligotypes and TB risk factors was explored. In this
231 study, seven different TB risk factors were analyzed which were included in the National TB
232 registry of Latvia since year 2000: housing status, alcohol abuse, drug abuse, HIV infection,
233 imprisonment, malnutrition and reported contact with a TB patient (Table 3). The majority of

234 patients had a permanent home (92.2%), while 70 (7.7%) patients were homeless. Of the patients
235 for whom HIV status was known (99.7%), 139 (15.3%) individuals were HIV-positive. Alcohol
236 abuse as a risk factor was detected for 270 (29.7%) patients, while 115 (12.7%) individuals were
237 recorded as being in prison, and 79 (8.7%) - using drugs. Malnutrition was recorded for a half of
238 all patients (51.1%), while less than 30% of individuals have reported the contact with TB
239 patient (27.1%). The presence of any of these seven TB risk factors was detected in the majority
240 of patients (76.5%). Overall, the prevalence of the main spoligotypes among individuals with
241 risk factors was similar to those observed in non-risk patients ($P>0.05$) (Table 3). However,
242 statistically significant difference was detected in the prevalence of both SIT1 and SIT53
243 spoligotypes depending on the housing status of the patients: SIT53 was more often present in
244 the homeless persons, and SIT1 was more often present in patients with permanent home
245 ($P=0.0083$ and $P=0.0424$, respectively). In addition, the spoligotype SIT262 was more often
246 detected in former prisoners (7.0% vs 1.9%, $P=0.005$).

247

248 **3.5. Drug resistance analysis**

249 Analysis of drug resistance profiles showed that 67.6% of MTB isolates were pan-susceptible
250 (n=614) and 32.4% were resistant to any drug (n=294) (Table 4). The distribution of
251 spoligotypes among drug resistant isolates was as follows: 43.2% SIT1 (Beijing, n=127), 9.9%
252 SIT42 (LAM family, n=29), 5.8% SIT53 (T family, n=17), 14.6% SIT283 (Haarlem, n=43); all
253 other spoligotypes accounted for 26.5% (n=78) of drug-resistant MTB strains. Spoligotypes with
254 a drug resistance prevalence greater than 30% among the isolates included SIT1 (Beijing,
255 56.7%), SIT42 (LAM, 56.9%), SIT264 (LAM-RUS, 66.7%), SIT283 (Haarlem, 100%) and

256 SIT1292 (Ural, 41.7%); statistically significant difference was reached for SIT1, SIT42 and
257 SIT283 (P≤0.05; Table 4).
258 In total, MDR and extensively-drug-resistance (XDR) was found in 5.8% (53/908) and 7.6%
259 (69/908) of MTB strains, respectively.
260 The analysis of the distribution of spoligotypes among MDR/XDR isolates showed a significant
261 overrepresentation of both SIT1 and SIT42 isolates: these spoligotypes accounted for 47.2% and
262 17.0% of MDR MTB, and for 55.1% and 34.0% of XDR strains, respectively (Table 4).

263 **3.65. Analysis of MTB isolates by different genotyping approaches**

264 From 908 MTB isolates, 411 samples were available for additional IS6110 RFLP and MIRU-
265 VNTR analyses and thus were used for more in-depth genotypic characterization of MTB
266 isolates in Riga and Riga region. In total, these samples represented 25.3% of all MTB-positive
267 cultures isolated in the 2008 – 2012 time period. Among them, 78 spoligotypes were identified.
268 The isolates belonged to the following MTB families: Beijing (n=104; 25.3%), T (n=91; 22.1%),
269 LAM (n=100; 24.3%), Haarlem (n=27; 6.6%), Ural (n=46; 11.2%) and X (n=14; 3.4%); the
270 remaining 29 (7.1%) represented other families and/or unpublished SITs (Table 2). This overall
271 representation of spoligotypes for these 411 isolates was similar to that observed for the full
272 MTB dataset, indicating that this selection was quite representative of the total strain set.
273 Spoligotyping alone identified 78 genotypes, 50 of which were unique; 87.8% of isolates (361 of
274 411) formed 27 clusters with a cluster range 2-31 samples.
275 Upon 24-locus MIRU-VNTR typing, 242 different MIRU-VNTR patterns were obtained; 213
276 MTB isolates (51.8%) were grouped into 44 clusters with 2 to 37 isolates per cluster, while the
277 remaining 198 samples showed a unique MIRU-VNTR profile (Table 54; Supplemental Table

278 | [32](#)). Thus, MIRU-VNTR typing had a substantially greater DP (DP=0.9846) in comparison
279 | with spoligotyping (DP=0.8953).

280 | When 24 MIRU-VNTR-based cluster analysis was performed for the main MTB families
281 | separately, the Beijing family isolates (n=104) were divided into 45 different MIRU-VNTR
282 | genotypes, forming ten clusters and 35 unique genotypes. According to the MLVA MtbC15-9
283 | nomenclature implemented in the MIRU-VNTR Plus database, the 3 most common allelic
284 | profiles were: (1) 94-32 (244233352644425153353823) represented by 37 isolates (35.6%);
285 | (2) 100-32 (244233352644425173353723) represented by 11 isolates (10.6%); and (3) 170-32
286 | (244243352544425153353823) represented by 5 isolates (4.8%) (Table [54](#); Supplemental
287 | Table [32](#)). MTB isolates of the LAM lineage (n=100) were divided into 44 MIRU-VNTR
288 | genotypes forming ten clusters with 2-25 members and 34 unique genotypes. The 3 most
289 | common genetic profiles were following: (1) 121-52 (132244332224125153322622)
290 | represented by 25 isolates (24.0%); (2) 843-52 (132254332224125153322622) represented by
291 | 15 isolates (15.0%); (3) 15664-52 (132244332224125153321622) represented by 7 isolates
292 | (7.0%).

293 | In contrast, MIRU-VNTR types within the other spoligotype families were relatively more
294 | diverse/less clustered. For instance, MTB isolates of T family (n=91) were divided into 61
295 | MIRU-VNTR genotypes, forming 11 clusters and 50 unique genotypes. The 3 most common
296 | genotypes were represented by 5 (5.5%; 20574-15, 223114322534425153342512), 6 (6.6%;
297 | 12677-15, 224243122334225143334322) and 7 isolates (7.7%; 20556-15,
298 | 224243122434225143335322), respectively. Likewise, the largest cluster observed among the
299 | isolates part of the Haarlem spoligotype family (n=27) comprised only 4 isolates (14.8%;
300 | 3527-15, 223235332434425153324732)), while a second cluster included only 2 isolates

301 (7.4%). The Ural MTB family members (N=46) formed five clusters, of which 2 were of four
302 isolates each (163-15, 235237232244425113323632 and 16668-15,
303 215237232244425113323732); three other clusters had three isolates each. The X family
304 members (n=14) have formed three clusters which had only two isolates each (Supplemental
305 Table [32](#)).

306 Next, we applied *IS6110* RFLP typing to all 411 isolates. In total, 257 different patterns were
307 obtained. 207 isolates (50.4%) were grouped into 53 clusters, and the calculated DP for the
308 RFLP genotyping method in our dataset was as high as 0.9927 (Table [54](#)). Importantly, in the
309 case of Beijing strains, *IS6110* RFLP typing showed superior resolution of the genotypes
310 (DP=0.9766) to both spoligotyping and MIRU-VNTR typing (DP=0.0571 and 0.8615,
311 respectively).

312 When all three genotyping methods were combined, only 119 MTB isolates (29.0%) were
313 grouped into 39 clusters with 2 to 12 isolates per one cluster, while the remaining 292 samples
314 have shown a unique RFLP and/or MIRU-VNTR/spoligotype patterns (Supplemental Table
315 [43](#)). The largest cluster was composed of 12 MTB isolates belonged to SIT254 (representative
316 of LAM-RUS family). In total, the LAM lineage isolates formed 9 clusters with 2-12 isolates,
317 the Beijing genotype isolates - 9 clusters (2-8 isolates), T family – 11 clusters (2-4 isolates),
318 Ural – four clusters (2-3 isolates), and X – three clusters with 2, 3 and 4 isolates each. In
319 addition, two clusters with two isolates each were obtained for the new spoligotypes
320 (spoligotype 75777777720761 and 477777663760771) (Supplemental Table [43](#)).

321

322 4. Discussion

323 Although MTB genotypes were previously studied in association with drug resistance in
324 Latvia (Tracevska et al., 2003; Nodieva et al., 2010; Bauskenieks et al., 2014), this is the first
325 comprehensive report describing the molecular epidemiology of TB in Riga and Riga region,
326 which is home for more than 50% of Latvia's population. The analysis of the prevalence of the
327 main MTB spoligotypes within the five-year time frame revealed the presence of 138
328 spoligotypes. Among them, 38 spoligotypes representing 5.6 percent of the 908 study isolates
329 had not been previously reported to the global SITVIT database. At the other extremity, the 3
330 most common spoligotypes, namely SIT1 (Beijing family), SIT53 (T family) and SIT254
331 (LAM family) accounted altogether for approximately one-half of all MTB isolates, and
332 represented 24.7, 11.0 and 11.9% of isolates, respectively. These patterns are also among the
333 most common SITs in the world (Demay et al., 2012). In addition to be dominant in East Asia,
334 strains with the canonical Beijing SIT1 spoligotype are present in all seven geographic areas
335 of the world (Merker et al, 2015). In Europe, Beijing isolates were previously reported to be
336 very common in the Latvia's eastern and northern neighboring countries, Russia and Estonia,
337 often in association with drug resistance (Toungousova et al., 2002; Kruuner et al., 2001). In
338 Latvia, too, Beijing genotype was prevalent among MDR-TB patients (Tracevska et al., 2003).
339 More recent data indicated that the XDR *M. tuberculosis* strain population in northwestern
340 Russia is heavily dominated by Beijing genotype isolates (Vyazovaya et al., 2015). In Pskov
341 oblast, which is located in northwestern Russia on the Russian/Latvian border, the Beijing
342 genotype was the most prevalent followed by LAM, T, Haarlem, Ural, and Manu2 families
343 (Mokrousov et al., 2012).

344 In this study, in total, 32.4% of MTB isolates were resistant to any drug, and 13.4% were
345 MDR/XDR strains. Unsurprisingly, Beijing genotype was highly associated with drug

346 resistance, as 43.2% of drug-resistant isolates belonged to the spoligotype SIT1. SIT1 was also
347 the predominant genotype among MDR/XDR isolates: approximately half of all MDR and
348 XDR isolates (i.e. 47.2% and 55.1%, respectively) belonged to this spoligotype. While SIT1
349 was also the most common spoligotype in our study, the predominance of this genotype
350 among drug-resistant isolates is not related only to the simple majority. The analysis of drug
351 resistance profiles showed that resistance to any drug and MDR/XDR phenotype was observed
352 in 56.7% and 28.1% of the SIT1 strains, respectively. Both proportions were statistically larger
353 when compared to the proportion of drug-resistant and MDR/XDR strains among the total
354 number of MTB isolates (i.e. 32.4% and 13.4%, respectively). MIRU-VNTR analysis of
355 Beijing isolates in Latvia showed the dominance of MTB strains belonged to clusters 94-32
356 (35.6% of Beijing isolates) and 100-32 (10.6%) which were reported to be the most common
357 cross-border molecular clusters of MDR TB in Europe (ECDC, 2016). Similarly, Beijing
358 clusters 94–32 and 100–32 were the dominant MTB genotypes in Central Asia, and the clonal
359 spread of resistant TB strains, particularly of the Beijing lineage, was highlighted as a root of
360 the MDR-TB epidemic in Central Asia (Engström et al., 2019).
361 However, the distribution of Beijing strains was not solely responsible for the transmission of
362 drug-resistant TB in Latvia. Less prevalent genotypes among drug-resistant isolates were
363 SIT283 (Haarlem, 14.6%), SIT42 (LAM, 9.9%), SIT53 (T family, 5.8%) and SIT264 (LAM-
364 RUS, 2.7%), followed by other spoligotypes with the prevalence less than 2% each. The
365 proportion of drug-resistant isolates among SIT42, SIT264 and SIT283 samples (56.9%,
366 66.7% and 100%, respectively), as well as the proportion of MDR/XDR isolates among SIT42
367 samples (52.9%) was significantly increased. Along with Beijing strains, the combination of
368 spoligotyping/MIRU-VNTR analysis revealed high clustering rate for LAM family in Latvia;

369 the two most common types were 121-52 and 843-52, which accounted for 25.0% and 15.0%
370 of LAM isolates, respectively. While SIT254 was not associated with drug resistance, as
371 95.4% of isolates were pan-susceptible in our study, the distribution of other LAM family
372 genotypes, including SIT42 and SIT264, is of high importance. These findings indicate that
373 the high burden of TB drug resistance in Latvia is at least partially related to the circulation of
374 MTB strains of specific genotypes. Thus, genotyping of MTB strains is highly important to
375 identify significant factors involved in the distribution of resistance strains which, in turn, may
376 affect TB control programs.

377 Overall, the prevalence of different genotype families in Latvia seen in this study
378 resembles those reported for the neighboring Estonia, northwestern Russia and, to a smaller
379 extent, Belarus and Ukraine (where Ural genotypes were not observed according to previous
380 reference reports) (Figure 2) (Chernyaeva et al., 2012; Mokrousov et al., 2012; Toit et al.,
381 2014; Zalutskaya et al., 2013; Nikolaevskaya et al., 2016;). Relatively to these Baltic and
382 Eastern countries, the prevalence of Beijing genotypes is substantially lower in Western
383 Europe and Scandinavia, and the proportion of isolates with spoligotype patterns typical of the
384 Euro-American lineage (alias lineage 4) of *M. tuberculosis* is archetypally much higher
385 (Figure 2) (Demay et al., 2012; Smit et al., 2013).

386 Interestingly, among the 51 isolates with a total of 38 novel spoligotypes that were
387 discovered among circulating MTB strains in Riga region, six shared the same spoligotype
388 477777663760771, unreported in any other world place thus far. These isolates might thus
389 represent a genotype specific to Latvia. Further studies would help to explore the frequency
390 and underlying clonal/genome diversity of this spoligotype more in details.

391 Importantly as well, the distribution pattern of the most common genotypes with a
392 dominance of Beijing, LAM and T spoligotypes remained stable over the five-year study
393 period, indicating an overall unchanged prevalence and circulation of different MTB strains in
394 the country and society. Similar result was obtained for the distribution of the most common
395 spoligotypes including drug resistance-related SIT1, SIT42 and SIT283. This stability at least
396 within this time span limits contrasts with the substantial changes in MTB strain population
397 structures observed over time in other European countries like Ireland and Germany,
398 associated with human migration and mobility (Roycroft et al., 2018; Andrés et al., 2017).
399 These results is consistent with the fact that the vast majority of TB patients in this study, i.e.
400 89.5%, were local born.

401 While we did not find any associations of spoligotypes with patients` characteristics
402 such as gender, country of birth or type of the disease, our PCA results showed that the SIT1
403 spoligotype is more frequent in 21–50 year-old adults while the rare spoligotypes were more
404 frequent in elderly patients. In keeping with the notion that the majority of infected individuals
405 remain asymptomatic over most or all of their lifetime (Zumla et al. 2013), and many of the
406 TB cases in the elderly result from reactivation of a latent MTB infection acquired at much
407 younger ages (Negin et al. 2015), the association of these rare spoligotypes with older patients
408 might reflect genotypes that were more prevalent in the past but have become less abundant
409 nowadays.

410 We also found that SIT53 (T family, representing a generic spoligotype signature that
411 can be shared by different branches of the Euro-American lineage of *M. tuberculosis*) isolates
412 were more common in homeless patients, while SIT1 spoligotype (Beijing family) was
413 significantly more frequent in patients with permanent home than in homeless persons, and

414 SIT262 (Ural family) was associated with imprisonment. Prisons are settings in which TB
415 transmission occurs and the incidence of TB in prisons is typically several fold higher
416 compared to the general population (reviewed in Biadglegne et al., 2015). Ural family strains
417 have not previously been associated with increased transmissibility (Mokrousov, 2012).
418 However, additional environmental, social, and host related risk factors (like history of
419 previous imprisonment in other (Eastern) countries, of contacts with prisoners with such a
420 history) could facilitate the spread of particular TB strains in these places (Biadglegne et al.,
421 2015). Similarly, low socioeconomic status and enhanced exposure in crowded places such as
422 homeless shelters may have an impact on the transmission dynamics of particular strains (part
423 of the Euro-American lineage in this case) in specific patient groups, such as homeless people.

424 Also, recent report of Sinkov and colleagues highlighted the recent emergency of Ural Clade C
425 (belongs to SIT262) in Eastern European countries Moldova and Belarus which was
426 significantly associated with XDR or pre-XDR status (Sinkov et al., 2019). In our study, the
427 majority of SIT262 isolates were pan-susceptible (22 of 23), while drug resistant strains
428 including those with MDR phenotype were found among isolates belonging to other Ural
429 family spoligotypes SIT35, 1292 and 3344. These findings demonstrated the circulation of
430 Ural strains with concern for drug resistance and the need for detection and surveillance of
431 these strains in Eurasia to prevent further dissemination.

432 Consistently with multiple previous studies, our results show that, although it is a rapid
433 and convenient genotyping method for broad identification of genotype families in clinical
434 strains (Gori et al., 2005) (Millet et al., 2007), spoligotyping cannot be used alone in order to
435 investigate the possible epidemiological links between patients, especially in settings
436 dominated by the Beijing lineage. Isolates with Beijing spoligotypes were the most common in

437 our study, resulting in a spoligotyping clustering rate of 97.1% (101/104 isolates) and a
438 calculated DP for as low as 0.057 for these isolates. Unsurprisingly as well, the 24-locus
439 MIRU-VNTR analysis clearly increased the overall resolution of the obtained MTB genotypes
440 (calculated DP=0.9846), and especially among isolates with a Beijing spoligotype with a
441 resulting clustering rate of 66.4% (69/104) among them. Likewise, IS6110 RFLP typing had a
442 much higher resolution in comparison with spoligotyping (DP=0.9927), and in the case of
443 Beijing lineage the DP was the highest among the methods applied (DP=0.9766). Given that
444 spoligotyping is both an appropriate routine method for a clinical laboratory in Latvia and a
445 useful screen to identify genotype families, a combination of this method with at least one of
446 the two other molecular genotyping approaches should be used in order to correctly interpret
447 typing results for the needs in epidemiological tracing.

448 In conclusions, we report the first comprehensive molecular epidemiological study of TB
449 in Latvia. This study highlighted the high genetic diversity of *M. tuberculosis* strains circulating
450 in Riga and Riga region. The combination of the spoligotyping with the 24-locus MIRU-VNTR
451 and IS6110 RFLP clearly increased the overall resolution of the MTB genotypes. In combination
452 with detailed epidemiological data, this approach was helpful for a first in-depth understanding
453 of epidemiological processes in a settings where the Next-Gen sequencing is not available as a
454 routine method.

455

456

457 **Acknowledgments**

458 This study was supported by the EU 7th Framework TB - PANNET and Latvian National
459 Research program VPP-2010-1/9.3 and VPP “BIOMEDICINE”. P.S. was a consultant for
460 Genoscreen.

461

462

463 **References**

- 464 1. Andrés, M., Göhring-Zwacka, E., Fiebig, L., Priwitzer, M., Richter, E., Rüscher-Gerdes, S.,
465 Haas, W., Niemann, S., Brodhun, B., 2017. Integration of molecular typing results into
466 tuberculosis surveillance in Germany-A pilot study. *PLoS One*. 12(11):e0188356. doi:
467 10.1371/journal.pone.0188356.
- 468 2. Bauskenieks, M., Pole, I., Skenders, G., Jansone, I., Broka, L., Nodieva, A., Ozere, I.,
469 Kalvisa, A., Ranka, R., Baumanis, V., 2015. Genotypic and phenotypic characteristics of
470 aminoglycoside-resistant *Mycobacterium tuberculosis* isolates in Latvia. *Diagn.*
471 *Microbiol. Infect. Dis.* 81(3):177-82. doi: 10.1016/j.diagmicrobio.2014.12.004.
- 472 3. Biadlegne, F., Rodloff, A.C., Sack, U., 2015. Review of the prevalence and drug
473 resistance of tuberculosis in prisons: a hidden epidemic. *Epidemiol. Infect.* 143(5):887-
474 900. doi: 10.1017/S095026881400288X.
- 475 4. Brudey, K., Gordon, M., Moström, P., Svensson, L., Jonsson, B., Sola, C., Ridell, M.,
476 Rastogi, N., 2004. Molecular epidemiology of *Mycobacterium tuberculosis* in western
477 Sweden. *J. Clin. Microbiol.* 42(7):3046-51.
- 478 5. Chernyaeva, E., Dobrynin, P., Pestova, N., Matveeva, N., Zhemkov, V., Kozlov, A.,
479 2012. Molecular genetic analysis of *Mycobacterium tuberculosis* strains spread in
480 different patient groups in St. Petersburg, Russia. *Eur. J. Clin. Microbiol. Infect. Dis.*
481 31(8):1753-7. doi: 10.1007/s10096-011-1497-2.
- 482 6. Couvin, D., David, A., Zozio, T., Rastogi, N., 2018. Macro-geographical specificities of
483 the prevailing tuberculosis epidemic as seen through SITVIT2, an updated version of the
484 *Mycobacterium tuberculosis* genotyping database. *Infect Genet Evol.* S1567-
485 1348(18)30969-9. doi: 10.1016/j.meegid.2018.12.030.

- 486 7. De Beer, J.L., van Ingen, J., de Vries, G., Erkens, C., Sebek, M., Mulder, A., Sloom, R.,
487 van den Brandt, A.M., Enaimi, M., Kremer, K., Supply, P., van Soolingen D., 2013.
488 Comparative study of IS6110 RFLP and VNTR typing of Mycobacterium tuberculosis in
489 the Netherlands, based on a five year nationwide survey. J. Clin. Microbiol. 51(4):1193-
490 8. doi: 10.1128/JCM.03061-12.
- 491 8. Demay, C., Liens, B., Burguière, T., Hill, V., Couvin, D., Millet, J., Mokrousov, I., Sola,
492 C., Zozio, T., Rastogi, N., 2012. SITVITWEB-a publicly available international
493 multimer database for studying Mycobacterium tuberculosis genetic diversity and
494 molecular epidemiology. Infect. Genet. Evol. 12(4):755-66. doi:
495 10.1016/j.meegid.2012.02.004.
- 496 9. Engström, A., Antonenka, U., Kadyrov, A., Kalmambetova, G., Kranzer, K., Merker, M.,
497 Kabirov, O., Parpieva, N., Rajabov, A., Sahalchik, E., Sayfudinov, Z., Niemann, S.,
498 Hoffmann, H., 2019. Population structure of drug-resistant Mycobacterium tuberculosis
499 in Central Asia. BMC Infect Dis. 29:19(1):908.
- 500 8.10. European Centre for Disease Prevention and Control. Molecular typing for
501 surveillance of multidrug-resistant tuberculosis in the EU/EEA – January 2016.
502 Stockholm: ECDC; 2016.
- 503 9.11. Gori, A., Bandera, A., Marchetti, G., Degli Esposti, A., Catozzi, L., Nardi, G.P.,
504 Gazzola, L., Ferrario, G., van Embden, J.D., van Soolingen, D., Moroni, M., Franzetti, F.,
505 2005. Spoligotyping and Mycobacterium tuberculosis. Emerg Infect Dis. 11(8):1242-8.
- 506 10.12. Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., et al.,
507 1997. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for
508 diagnosis and epidemiology. J. Clin. Microbiol. 35: 907–914.

509 | ~~11.~~13. Kruuner, A., S. E. Hoffner, H. Sillastu, M. Danilovits, K. Levina, S. B. Svenson,
510 | S. Ghebremichael, T. Koivula, and G. Kallenius. 2001. Spread of drug-resistant
511 | pulmonary tuberculosis in Estonia. *J. Clin. Microbiol.* 39: 3339–3345.

512 | ~~12.~~14. Lê, S., Josse, J., Husson, F., 2008. FactoMineR: An R Package for Multivariate
513 | Analysis. *Journal of Statistical Software.* 25(1):1-18.

514 | ~~13.~~15. Merker, M., Blin, C., Mona, S., Duforet-Frebourg, N., Lecher, S., Willery, E.,
515 | Blum, M.G., Rüsç-Gerdes, S., Mokrousov, I., Aleksic, E., Allix-Béguet, C., Antierens,
516 | A., Augustynowicz-Kopeć, E., Ballif, M., Barletta, F., Beck, H.P., Barry, C.E., Bonnet,
517 | M., Borroni, E., Campos-Herrero, I., Cirillo, D., Cox, H., Crowe, S., Crudu, V., Diel, R.,
518 | Drobniewski, F., Fauville-Dufaux, M., Gagneux, S., Ghebremichael, S., Hanekom, M.,
519 | Hoffner, S., Jiao, W.W., Kalon, S., Kohl, T.A., Kontsevaya, I., Lillebæk, T., Maeda, S.,
520 | Nikolayevskyy, V., Rasmussen, M., Rastogi, N., Samper, S., Sanchez-Padilla, E., Savic,
521 | B., Shamputa, I.C., Shen, A., Sng, L.H., Stakenas, P., Toit, K., Varaine, F., Vukovic, D.,
522 | Wahl, C., Warren, R., Supply, P., Niemann, S., Wirth, T., 2015. Evolutionary history and
523 | global spread of the Mycobacterium tuberculosis Beijing lineage. *Nat. Genet.* 47(3):242-
524 | 9. doi: 10.1038/ng.3195.

525 | ~~14.~~16. Merker, M., Kohl, T.A., Niemann, S., Supply, P., 2017. The Evolution of Strain
526 | Typing in the Mycobacterium tuberculosis complex. *Adv. Exp. Med. Biol.* 1019:43-78.
527 | doi: 10.1007/978-3-319-64371-7_3.

528 | ~~15.~~17. Millet, J., Miyagi-Shiohira, C., Yamane, N., Sola, C., Rastogi, N., 2007.
529 | Assessment of mycobacterial interspersed repetitive unit-QUB markers to further
530 | discriminate the Beijing genotype in a population-based study of the genetic diversity of

531 Mycobacterium tuberculosis clinical isolates from Okinawa, Ryukyu Islands, Japan. J.
532 Clin. Microbiol. 45(11):3606-15.

533 | ~~16.~~18. Mokrousov, I., 2012. The quiet and controversial: Ural family of Mycobacterium
534 tuberculosis. Infect. Genet. Evol. 12(4):619-29. doi: 10.1016/j.meegid.2011.09.026.

535 | ~~17.~~19. Mokrousov, I., Vyazovaya, A., Otten, T., Zhuravlev, V., Pavlova, E.,
536 Tarashkevich, L., Krishevich, V., Vishnevsky, B., Narvskaya, O., 2012. Mycobacterium
537 tuberculosis population in northwestern Russia: an update from Russian-EU/Latvian
538 border region. PLoS One. 7(7):e41318.

539 | ~~18.~~20. Mokrousov, I., Vyazovaya, A., Solovieva, N., Sunchalina, T., Markelov, Y.,
540 Chernyaeva, E., Melnikova, N., Dogonadze, M., Starkova, D., Vasilieva, N., Gerasimova,
541 A., Kononenko, Y., Zhuravlev, V., Narvskaya, O., 2015. Trends in molecular
542 epidemiology of drug-resistant tuberculosis in Republic of Karelia, Russian Federation.
543 BMC Microbiol. 15:279. doi: 10.1186/s12866-015-0613-3.

544 | ~~19.~~21. Negin, J., Abimbola, S., Marais, B.J., 2015. Tuberculosis among older adults--
545 time to take notice. Int. J. Infect. Dis. 32():135-7.

546 | ~~20.~~22. Nikolaevskaya, E., Pavlovska, O., Molina, B., Brännberg, P., Marynova, I.,
547 Krylova, K., Ivanitsa, T., Filuk, V., Lacom, A., Dominguez, J., Rzhepishevskaya, O., Prat,
548 C., 2016. Spoligotypes of Mycobacterium tuberculosis isolated from tuberculosis patients
549 in Odessa, Ukraine. European Congress of Clinical Microbiology and Infectious
550 Diseases. Amsterdam, Netherlands, 9-12 April, 2016. P0155.

551 | ~~21.~~23. Nodieva, A., Jansone, I., Broka, L., Pole, I., Skenders, G., Baumanis, V., 2010.
552 Recent nosocomial transmission and genotypes of multidrug-resistant Mycobacterium
553 tuberculosis. Int. J. Tuberc. Lung Dis. 14(4):427-33.

554 | ~~22-24.~~____ Roycroft, E., O'Toole, R.F., Fitzgibbon, M.M., Montgomery, L., O'Meara, M.,
555 | Downes, P., Jackson, S., O'Donnell, J., Laurenson, I.F., McLaughlin, A.M., Keane, J.,
556 | Rogers, T.R., 2018. Molecular epidemiology of multi- and extensively-drug-resistant
557 | Mycobacterium tuberculosis in Ireland, 2001-2014. *J. Infect.* 76(1):55-67. doi:
558 | 10.1016/j.jinf.2017.10.002.

559 | 25. Satta, G., Lipman, M., Smith, G.P., Arnold, C., Kon, O.M., McHugh, T.D., 2018.
560 | Mycobacterium tuberculosis and whole-genome sequencing: how close are we to
561 | unleashing its full potential? *Clin. Microbiol. Infect.* 24(6):604-609. doi:
562 | 10.1016/j.cmi.2017.10.030.

563 | ~~23-26.~~ Sinkov, V., Ogarkov, O., Mokrousov, I., Bukin, Y., Zhdanova, S., Heysell, S.K.,
564 | 2019. New epidemic cluster of pre-extensively drug resistant isolates of Mycobacterium
565 | tuberculosis Ural family emerging in Eastern Europe. *BMC Genomics.* 19(1):762. doi:
566 | 10.1186/s12864-018-5162-3.

567 | ~~24-27.~~____ Smit, P.W., Haanperä, M., Rantala, P., Couvin, D., Lyytikäinen, O., Rastogi, N.,
568 | Ruutu, P., Soini, H., 2013. Molecular Epidemiology of Tuberculosis in Finland, 2008-
569 | 2011. *PLoS One.* 8(12): e85027.

570 | ~~25-28.~~____ Supply, P., C. Allix, S. Lesjean, M. Cardoso-Oelemann, S. Rusch-Gerdes, E.
571 | Willery, E. Savine, P. de Haas, H. van Deutekom, S. Roring, P. Bifani, N. Kurepina, B.
572 | Kreiswirth, C. Sola, N. Rastogi, V. Vatin, M. C. Gutierrez, M. Fauville, S. Niemann, R.
573 | Skuce, K. Kremer, C. Locht, and D. van Soolingen. 2006. Proposal for standardization of
574 | optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat
575 | typing of Mycobacterium tuberculosis . *J. Clin. Microbiol.* 44:4498–4510.

576 | ~~26-29.~~ Supply, P., Mazars, E., Lesjean, S., Vincent, V., Gicquel, B., et al., 2000. Variable
577 | human minisatellite-like regions in the Mycobacterium tuberculosis genome. Mol.
578 | Microbiol. 36: 762–771.

579 | ~~27-30.~~ Supply, P., Lesjean, S., Savine, E., Kremer, K., van Soolingen, D., Locht, C.,
580 | 2001. Automated high-throughput genotyping for study of global epidemiology of
581 | Mycobacterium tuberculosis based on mycobacterial inter- spersed repetitive units. J.
582 | Clin. Microbiol. 39: 3563–3571.

583 | ~~28-31.~~ Toit, K., Altraja, A., Acosta, C.D., Viiklepp, P., Kremer, K., Kummik, T.,
584 | Danilovič, M., Van den Bergh, R., Harries, A.D., Supply, P., 2014. A four-year
585 | nationwide molecular epidemiological study in Estonia: risk factors for tuberculosis
586 | transmission. Public Health Action. 4(Suppl 2):S34-40. doi: 10.5588/pha.14.0045.

587 | ~~29-32.~~ Toungousova, O. S., P. Sandven, A. O. Mariandyshev, N. I. Nizovtseva, G.
588 | Bjune, and D. A. Caugant. 2002. Spread of drug-resistant Mycobacterium tuberculosis
589 | strains of the Beijing genotype in the Archangel Oblast, Russia. J. Clin. Microbiol.
590 | 40:1930–1937.

591 | ~~30-33.~~ Tracevska, T., Jansone, I., Baumanis, V., Marga, O., Lillebaek, T., 2003.
592 | Prevalence of Beijing genotype in Latvian multidrug-resistant Mycobacterium
593 | tuberculosis isolates. Int. J. Tuberc. Lung Dis. 7(11):1097-103.

594 | ~~31-34.~~ Vyazovaya, A., Levina, K., Zhuravlev, V., Viiklepp, P., Kütt, M., Mokrousov, I.,
595 | 2018. Emerging resistant clones of Mycobacterium tuberculosis in a spatiotemporal
596 | context. J Antimicrob. Chemother. 73(2):325-331. doi: 10.1093/jac/dkx372.

597 | ~~32.~~35.____ Vyazovaya, A., Mokrousov, I., Zhuravlev, V., Solovieva, N., Otten, T.,
598 | Vishnevsky, B., Narvskaya, O., 2015. Dominance of the Beijing genotype among XDR
599 | *Mycobacterium tuberculosis* strains in Russia. *Int. J. Mycobacteriol.* 4 (S1):84-85.

600 | ~~33.~~36.____ Van Embden, J.D., Cave, M.D., Crawford, J.T., Dale, J.W., Eisenach, K.D.,
601 | Gicquel, B., Hermans, P., McAdam, R., Shinnick, T.M., Small, P.M., 1993. Strain
602 | identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations
603 | for a standardized methodology. *J. Clin. Microbiol.* 31: 406-409.

604 | ~~34.~~37.____ Van Soolingen, D., 2001. Molecular epidemiology of tuberculosis and other
605 | mycobacterial infections: main methodologies and achievements. *J. Intern. Med.* 249 (1):
606 | p. 1-26.

607 | ~~35.~~38.____ van Soolingen, D., Hermans, P.W., de Haas, P.E., Soll, D.R., van Embden, J.D.,
608 | 1991. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis*
609 | complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a
610 | tool in the epidemiology of tuberculosis. *J. Clin. Microbiol.* 29(11):2578-86.

611 | ~~36.~~39.____ Weniger, T., Krawczyk, J., Supply, P., Niemann, S., Harmsen, D., 2010. MIRU-
612 | VNTRplus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis*
613 | complex bacteria. *Nucleic Acids Res.* 38(Web Server issue):W326-31. doi:
614 | 10.1093/nar/gkq351.

615 | ~~37.~~40.____ Zalutskaya, A., Wijkander, M., Jureen, P., Skrahina, A., Hoffner, S., 2013.
616 | Multidrug-resistant *Mycobacterium tuberculosis* caused by the Beijing genotype and a
617 | specific T1 genotype clone (SIT No. 266) is widely transmitted in Minsk. *Int. J.*
618 | *Mycobacteriol.* 2(4):194-8. doi: 10.1016/j.ijmyco.2013.08.001. Epub 2013 Sep 6.

619 | ~~38.41.~~ Zumla, A., Raviglione, M., Hafner ,R., von Reyn, C.F., 2013. Tuberculosis. N.
620 | Engl. J. Med. 368(8):745-55. doi: 10.1056/NEJMra1200894.

621

622

623 **Figure captures**

624 **Figure 1.** Principal component analysis (PCA) of the *M. tuberculosis* spoligotypes in age groups.

625 PCA map showed remarkable segregation of SIT1 and spoligotypes with a prevalence less than

626 2% in the population (collectively named as Others).

627

628 **Figure 2.** Distribution of *M. tuberculosis* lineages in Latvia and neighboring countries. Data

629 were obtained: Latvia, this study; Russia, Mokrousov et al., 2015; Estonia, Vyazovaya et al.,

630 2018; Finland, Smit et al., 2013; Sweden, Brudey et al., 2004; Belarus, Zalutskaya et al., 2013;

631 Ukraine, Nikolaevskaya et al., 2016; Poland, SITWIT database; Germany, SITVIT database;

632 Denmark, SITVIT database.

633

*Author Contributions Section

Author contributions

Ilva Pole: Investigation, Methodology, Writing- Original draft

Julija Trofimova: Investigation

Inga Norvaisa: Investigation

Philip Supply: Funding acquisition, Writing- Reviewing and Editing

Girts Skenders: Investigation

Anda Nodieva: Investigation

Iveta Ozere: Investigation

Vija Riekstina: Data curation

Viktorija Igumnova: Investigation

Jelena Storozenko: Supervision

Inta Jansone: Conceptualization, Supervision, Investigation

Ludmila Viksna: Funding acquisition, Supervision

Renate Ranka: Conceptualization, Funding acquisition, Supervision, Writing- Reviewing and Editing