

Forensic application of Y-chromosomal STR analysis in Lithuanian population

Giedrė Ruzgaitė¹*,

Marija Čaplinskienė²,

Rima Baranovienė²,

Jūratė Jankauskienė²,

Jolanta Kukienė²,

Kristina Savanevskytė²,

Daiva Bunokienė²

¹ *Vytautas Magnus University,
K. Donelaičio St. 58,
LT-44248 Kaunas, Lithuania*

² *State Forensic Medicine Service
under the Ministry of Justice
of the Republic of Lithuania,
Didlaukio St. 86E,
LT-08303 Vilnius, Lithuania*

This paper presents a comprehensive Y-chromosomal STR haplotype analysis in the Lithuanian population in order to evaluate Lithuanians' Y chromosome diversity, to infer genetic relations between Lithuanian and other European neighbouring populations and to introduce population reference data for generation of reliable Y-STR haplotype frequency estimates to be used in the quantitative assessment of Y-STR haplotype match in the forensic casework. Data were collected from the peripheral blood samples of 194 unrelated males throughout various regions of Lithuania. The amplification of 17 Y-STRs was carried out in one multiplex PCR using an AmpF/STR[®] Yfiler[™] PCR Amplification Kit according to the supplier's protocol. The results indicated that the Y-chromosomal haplotype diversity in the Lithuanian population rises as the number of the analyzed Y-STRs is increased. However, all additional Y-STR loci are not hypervariable and only their whole makes a large diversity of Y-STR haplotypes in Lithuanian males. The analysis of molecular variance revealed low but significant interpopulation differences except the pair of Lithuanian and Latvian populations. The phylogenetic analysis showed that the clustered Y chromosome gene pool of Lithuanians and Latvians has a closer phylogenetic relation to Russian and Estonian populations and is less genetically related to other neighbouring populations of Belarus and Poland. Yet Y-STRs alleles and haplotypes differentiate effectively inside the Lithuanian population and between Lithuanians and its geographical neighbours excluding the Latvian population. Comparison of the Y-STR data suggests that Lithuanian and Latvian populations are closely related not only by geography and language but also by the Y chromosome gene pool represented by forensic Y-STR markers. Consequently, more forensic Y-STR markers should be included in the Y-STR haplotype in order to achieve a resolution between the Y chromosomes of Lithuanian and Latvian males. Lithuanian Y-STR haplotype data were submitted to the 34th release of the Y-STR Haplotype Reference Database 3.0 for match probability calculations in the forensic casework.

Key words: Lithuanian population, genetic diversity, Y chromosome, STR, forensic genetics

* Corresponding author. E-mail: giedre.ruzgaite@yahoo.com

INTRODUCTION

Lithuania is a tiny dot on the map of the world but the name of Lithuania was first mentioned in historical sources as early as 1009. It is situated in Northeastern Europe, the southernmost of the three Baltic states. The western coast of Lithuania is washed by the Baltic Sea. In the north, east and west it borders with the neighbouring countries of Latvia, Belarus, Poland and the Kaliningrad region of the Russian Federation. The geographical center of Europe is in Lithuania, in the capital Vilnius.

The first people settled in the territory of Lithuania after the last glacial period in the 10th millennium BC (Česnys, 1991). Over a millennium, the Proto-Indo-Europeans, who arrived in the 3rd–2nd millennium BC, mixed with the local population and formed various Baltic tribes (Česnys, 1991), which in the 1230s were united by Mindaugas, who was crowned as King of Lithuania in 1253. The Grand Duchy of Lithuania expanded rapidly overtaking the former Slavic principalities of Kievan Rus, and by the end of the 14th century, Lithuania was the largest country in Europe (present-day Belarus, Ukraine, Estonia, Latvia, parts of Poland and Russia were the territories of the multicultural Grand Duchy of Lithuania). With the Lublin Union of 1569, Poland and Lithuania formed a new state, the Polish–Lithuanian Commonwealth. The Commonwealth lasted more than two centuries until neighbouring countries systematically dismantled it from 1772 to 1795 and the Russian Empire annexed most of Lithuania's territory. Nevertheless, the Lithuanian National Revival laid the foundations of the modern Lithuanian nation and independent Lithuania.

These historical events contributed greatly to the ethnic composition of Lithuania. The majority of the population is ethnic Lithuanians, who speak Lithuanian that belongs to the Baltic group of Indo-European languages (Lithuanian and Latvian are the only languages of the Baltic group that are still spoken today). Sizeable minorities are the Poles (the largest

minority), the Russians (the second largest minority) and Belarusians. Also the unification of Baltic tribes resulted in linguistic differentiation among various regions of Lithuania. Six ethno-linguistic groups can be distinguished (west, south, east Aukštaičiai and north, west, south Žemaičiai), but without apparent genetic differences among them (Kasperavičiūtė, Kučinskas, 2004; Kasperavičiūtė et al., 2004). Consequently, a gene pool of the present day Lithuanian population is relatively homogeneous and is comprised of a complex mixture of the former Baltic tribes with the varying influence from Finno-Ugric and Slavic cultures within the history of the Lithuanian population.

Genetic studies of the Lithuanian population were performed using a range of different genetic markers, starting from blood groups and serum proteins through phenylketonuria and cystic fibrosis disease mutations to RFLPs (restriction fragment length polymorphism), Alu insertions, SNPs (single nucleotide polymorphism), STRs (short tandem repeat) and CNVs (copy number variation) located in mitochondrial, Y chromosome or autosomal DNA sequences (Kasperavičiūtė, Kučinskas, 2004; Kasperavičiūtė et al., 2004; Kučinskas, 2001; Zerjal et al., 2001; Lessig et al., 2001; Laitinen et al., 2002; Roewer et al., 2005). However, the genomic origin of the Lithuanian population and its genetic relationships with neighbouring populations are still under question as results from different population genetic studies using various genetic markers are inconsistent. In our Lithuanian population study we chose Y chromosome STR markers since due to unique Y chromosome genetic characteristics such as male-specificity, paternal inheritance and absence of genetic recombination, Y chromosome STRs are proved to be an important tool in answering specific population and forensic genetics questions. For example, in population genetics, Y chromosome STR testing has become useful for making inferences on human migration and evolutionary studies because the lack of recombination enables comparison of male

individuals separated by large periods of time. In the field of forensic genetics, these STR markers are especially helpful in deficient paternities, missing persons and mass disaster victim identification where the presence of male relatives having the same Y chromosome expands the number of possible reference samples or rape cases where male specific Y-STR testing permits isolation of the male component from the female DNA.

During the forensic DNA analysis in the event of a match between the Y-STR haplotypes for a case DNA sample and a reference DNA sample, it is a must to have an estimate of the probability that the Y-STR profile match would occur by chance in a population (Roewer et al., 2000). That is, in order to make statistically accurate Y-STR test results in the forensic casework, it is required to calculate the frequency of a given Y-STR haplotype in a population using large and diverse human male population reference databases since the significance of a match between genetically typed DNA samples depends on the frequency at which the observed Y-STR haplotype occurs in a population (Roewer et al., 2000). The significance of the observed Y-STR haplotype match in forensic genetics decreases as the frequency of a particular Y-STR haplotype in a population increases. It should be noted that observation of a match with Y-STRs does not have the same power of discrimination and weight in court as a match with autosomal STRs able to individualize DNA profiles since the genetic information the Y chromosome contains is consequently shared by all male individuals related through the paternal line as it is transmitted unchanged from father to son unless a mutation occurs. Although Y-STRs have a relatively high mutation rate that correlates with an extensive polymorphism and provides high-resolution Y-STR haplotyping, there are shared Y-STR haplotypes in populations because either males share identity by descent or because a particular set of Y-STRs does not distinguish closely related, but different, paternal lineages due to the lack of sufficient polymorphism in a population

(Kayser et al., 2001). This led to pay close attention to Y-STR diversity assessment within and between human male populations.

Although studies of Y-STR haplotype distribution have been performed in a variety of populations worldwide and the data on Y-STR markers have been submitted to the Y-STR population reference databases, there is a lack of information on Lithuanian Y chromosome STR haplotypes. Thus, in this study we will evaluate Lithuanians' Y-chromosomal diversity with respect to the Y-STR allelic and haplotype variation, address the question how Lithuanians are genetically related to other European, in particular, neighbouring populations and introduce Lithuanian Y chromosome STR haplotype reference data for forensic purposes.

This paper presents a Y-chromosomal STR haplotype analysis in the Lithuanian population according to 17 forensic Y-STR loci: DYS456, DYS389I, DYS390, DYS389II, DYS458, DYS19, DYS385a/b, DYS393, DYS391, DYS439, DYS635, DYS392, Y GATA H4, DYS437, DYS438, DYS448. All these markers are highly sensitive, easy to use in forensic practice and are explored in many other populations of the world (Kayser et al., 2001; Roewer et al., 2000; Kayser et al., 1997; Pascali et al., 1998).

MATERIALS AND METHODS

DNA samples. For the Lithuanian population genetic analysis, 194 peripheral blood samples were collected from unrelated males involved in paternity cases during the years 2008–2009. Previous studies (Chakraborty et al., 1992; Evett et al., 1991) suggested that 100–150 tested samples per population might provide sufficient population and forensic genetics calculations based on allele frequencies for DNA loci. Obviously, collecting genetic information from more samples per population usually adds to the precision of the obtained results (Butler, 2005). In this study, samples were selected from all ten regions of Lithuania, i. e. Vilnius, Kaunas, Klaipėda, Šiauliai, Panevėžys,

Utena, Alytus, Marijampolė, Tauragė and Telšiai, in order to represent six ethno-linguistic groups of Lithuania. Related males or males of non-local origin were not included. Males with obvious non-Lithuanian surnames were omitted as well as data with the same surname were deleted. The geographical origin of each sample was based on the place of birth of the index person. Therefore, samples represent an unbiased and presentable part of male population in Lithuania. Genomic DNA was extracted from peripheral blood samples using the Proteinase K/Chelex 100 method (Butler, 2005).

Y chromosome genotyping. For each DNA sample, the amplification of Y chromosomal STR loci was performed in one multiplex polymerase chain reaction by the use of the AmpF/STR® Yfiler™ PCR Amplification Kit (Applied Biosystems) according to the supplier's protocol (Mulero et al., 2006). This kit coamplifies 17 Y-STRs, including the markers defined as the "European minimal haplotype" (DYS19, DYS385 a/b, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393) (Gill et al., 2001), plus two loci (DYS438 and DYS439), added to this panel by the Scientific Working Group on DNA Analysis Methods (SWGDM) (Lee et al., 2004), and additional highly polymorphic loci: DYS437, DYS448, DYS456, DYS458, DYS635 and Y GATA H4 (Mulero et al., 2006). In the final 25 µl polymerase chain reaction volume, 10 µl (0.1 ng/µl) of genomic DNA was mixed with 9.2 µl of AmpF/STR® Yfiler™ PCR Reaction Mix, 0.8 µl (5 U/µl) of AmpliTaq Gold® DNA Polymerase and 5.0 µl of AmpF/STR® Yfiler™ Primer Set. Amplification conditions were as follows: 1 ng DNA, 4 U AmpliTaq Gold® DNA Polymerase, forward and reverse primer concentration ranged from 0.12–1.5 µM, 800 µM dNTPs, 1.60 mM MgCl₂, 50 mM KCl, 0.05% sodium azide, 160 µg/ml bovine serum albumin, 10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0. The forward primers were fluorescently labelled by the five-dye set system consisting of 6-FAM™, VIC®, NED™, PET® and LIZ®. The samples were amplified in the GeneAmp®

PCR System 9700 (Applied Biosystems). The thermal cycling conditions comprised enzyme activation at 95 °C for 11 min, followed by 30 cycles of denaturation at 94 °C for 1 min, annealing at 61 °C for 1 min, and extension at 72 °C for 1 min. The final extension was performed at 60 °C for 80 min with 4 °C temperature hold until the samples were removed from the thermal cycler.

The PCR products were separated and detected by capillary electrophoresis in the Performance Optimized Polymer (3130 POP-4™) on the ABI PRISM® 3130 Genetic Analyzer (Applied Biosystems) using the HID Fragment Analysis 36_POP4 module. Samples for electrophoresis were prepared as follows: 1 µl of the amplification product or AmpF/STR® Yfiler™ Allelic Ladder was added to the mixture of 0.3 µl of Gene Scan™ – 500 LIZ® size standard and 8.7 µl of deionized Hi-Di™ Formamide (Applied Biosystems), denatured at 95 °C for 3 min, and then chilled on ice for 3 min. Capillary electrophoresis results were analyzed using the Gene Mapper® ID software v.3.2. Y-STR allele assignment was made by comparison to the AmpF/STR® Yfiler™ Allelic Ladder. Alleles were named according to the contained number of repeat units under the recommendations of the DNA Commission of the International Society for Forensic Genetics (Gill et al., 2001; Gusmão et al., 2006).

Statistical analyses. Y-STR haplotype data were prepared for the analysis using the MS Excel™ with Microsatellite Toolkit add-in (Park, 2001). The basic parameters of molecular diversity were calculated using the Arlequin software ver. 3.1 (Excoffier et al., 2005). Allele and haplotype frequencies were estimated by a simple gene counting method. The allele frequency of the multicopy locus DYS385 a/b was analyzed as a combination of two alleles. Gene diversity (GD) of each locus and haplotype diversity (HD) was computed using the formula $D = (n/n-1)(1-\sum p_i^2)$, where n represents the sample size, i. e. The total number of alleles or haplotypes, and p_i is the frequency of the i th allele or haplotype (Nei, 1987). For each of

these values the standard deviation (SD) was counted. Population genetic characteristics, i. e. The number of different, unique and repeating haplotypes, were calculated. Haplotype discrimination capacity was determined by dividing the number of different haplotypes seen in a given population by the total number of the population sample (Kayser et al., 1997). Analysis of molecular variance (AMOVA) (Roewer et al., 1996) for 9 minimal haplotype Y-STR loci was performed to measure the genetic distance between the Lithuanian population sample and reference data from neighbouring populations submitted to the Y-STR Haplotype Reference Database 3.0 (YHRD) (Willuweit, Roewer, 2007). Pairwise distances based on Φ_{ST} values were computed from R_{ST} by using an online AMOVA tool (Willuweit, Roewer, 2007). The results of AMOVA were tested for statistical significance with 10,000 permutations and visualized by a multidimensional scaling (MDS) plot provided at the YHRD website (Willuweit, Roewer, 2007). To illustrate the comparative analysis of Y-STR haplotypes among populations, a neighbour-joining tree was designed using the PHYLIP package ver. 3.69 (Felsenstein, 2005) and was viewed with the TreeView software (Page, 1996). Y-STR haplotype data ($n = 8029$) for comparison represent five European neighbouring countries that relate with the history of Lithuanian population: Estonia, 1 population sample with 133 haplotypes; Latvia, 1 population sample with 145 haplotypes; Belarus, a population set of 12 samples with 489 haplotypes, approached as one metapopulation of the Belarusian ethnic group; Poland, a population set of 21 samples with 4 730 haplotypes, approached as one metapopulation with the majority of the Polish ethnic group ($n = 4195$); Russian Federation, a population set of 49 samples with 2 532 haplotypes, approached as one metapopulation with the majority of the Russian ethnic group ($n = 1466$) (Willuweit, Roewer, 2007).

Quality control. The State Forensic Medicine Service participates in the quality control of the GEDNAP, i. e. The German DNA

Profiling Group (Rand et al., 2002; 2004) as well as uses internal control standards and kit controls. The presented Y-STR haplotype data were submitted to the 34th release of the Y-STR Haplotype Reference Database 3.0 (YHRD) (Willuweit, Roewer, 2007) and received the accession number (YA003661) after evaluation of the Lithuanian population Y-STR data set.

RESULTS

Genetic diversity of Y chromosome STR markers in Lithuanian population. The main measure of population genetic diversity revealed by genetic markers in the population genetics is the allele frequency. The allele range from 4 to 10 alleles was detected for 16 Y-STR loci. The frequencies of single alleles and genotypes (for duplicated locus DYS385 a/b) segregating at each Y-STR polymorphism are presented in Table 1. In comparison with other populations (information obtained from YHRD (Willuweit, Roewer, 2007)), most of the Y-STR loci in the Lithuanian population have a low number of allelic variants, i. e. from many possible Y-STR alleles observed in populations around the world Lithuanian males have only a few. Table 1 also lists the Y chromosome gene diversity values. For isolated microsatellite loci gene diversity ranged between 0.3137 (DYS437) and 0.8012 (DYS385 a/b, when approached as genotype).

Genetic variation at different Y-STR loci is combined to a large variety of Y chromosome STR haplotypes. When Y-STR haplotypes were comprised from the most variable Y-STR loci to the least variable Y-STR loci, it revealed that over the 9 most informative Y-STR loci, little resolution was made by adding other Y-STR loci (Fig. 1). In order to evaluate the additional discrimination power within the Lithuanian population, we calculated the Y-STR haplotype diversity for different Y-STR marker combinations (Table 2). The analyzed Y-STR marker combinations consisted of frequently used sets including the “European minimal haplotype” (Gill et al., 2001), the “SWGDM

Table 1. Observed alleles, allele frequencies and gene diversity (GD) of the 16 Y-STR loci in the Lithuanian population

Locus	DYS 19	DYS 389I	DYS 389II	DYS 390	DYS 391	DYS 392	DYS 393	DYS 437	DYS 438	DYS 439	DYS 448	DYS 456	DYS 458	DYS 635	Y_GATA H4	Locus Allele	DYS 385 a/b
6																10/13	0.0052
7																10/14	0.0155
8																11/11	0.0103
9					0.0052				0.0361	0.0103						11/12	0.0052
10				0.4278	0.0103			0.4124	0.5670						0.0103	11/13	0.2629
11	0.0052			0.5464	0.5722	0.0155		0.4381	0.2784						0.2010	11/14	0.3299
12	0.0722			0.0206	0.0361	0.0361		0.1082	0.1237			0.0258			0.5309	11/15	0.1443
13	0.0464	0.5825		0.0619	0.0103	0.0103	0.0052	0.0155	0.0052	0.2010	0.0052	0.2010	0.0052		0.2577	11/16	0.0052
14	0.1392	0.3247		0.2629	0.3402	0.8196				0.1237	0.0258					11/17	0.0052
15	0.4691	0.0155		0.0515	0.0103	0.1134				0.2990	0.3402					12/13	0.0052
16	0.2938			0.0052		0.0567				0.2423	0.2062					12/14	0.0258
17	0.0515									0.1031	0.2474					12/15	0.0206
18									0.0103	0.0052	0.1340					13/13	0.0052
19									0.3918		0.0361					13/14	0.0206
20				0.0052					0.5412		0.0052	0.0206				13/15	0.0103
21				-					0.0412							13/16	0.0155
22				0.0567					0.0155							13/20	0.0052
23				0.3402											0.4124	14/14	0.0155
24				0.2423											0.1753	14/15	0.0309
25				0.3247											0.0361	14/16	0.0052
26				0.0309												14/17	0.0052
27				0.0052												15/15	0.0103
28				0.0722												15/17	0.0103
29				0.2216												15/18	0.0103
30				0.4845												16/17	0.0103
31				0.1804												16/18	0.0103
32				0.0361													
GD	0.6729	0.5526	0.6806	0.7196	0.5207	0.5986	0.5279	0.3137	0.6282	0.5884	0.5544	0.7889	0.7645	0.7380	0.6144		0.8012
±SD	0.0216	0.0249	0.0236	0.0127	0.0141	0.0297	0.0241	0.0404	0.0194	0.0266	0.0194	0.0113	0.0135	0.0185	0.0223		0.0183

* Values of the most frequent alleles are indicated in bold.

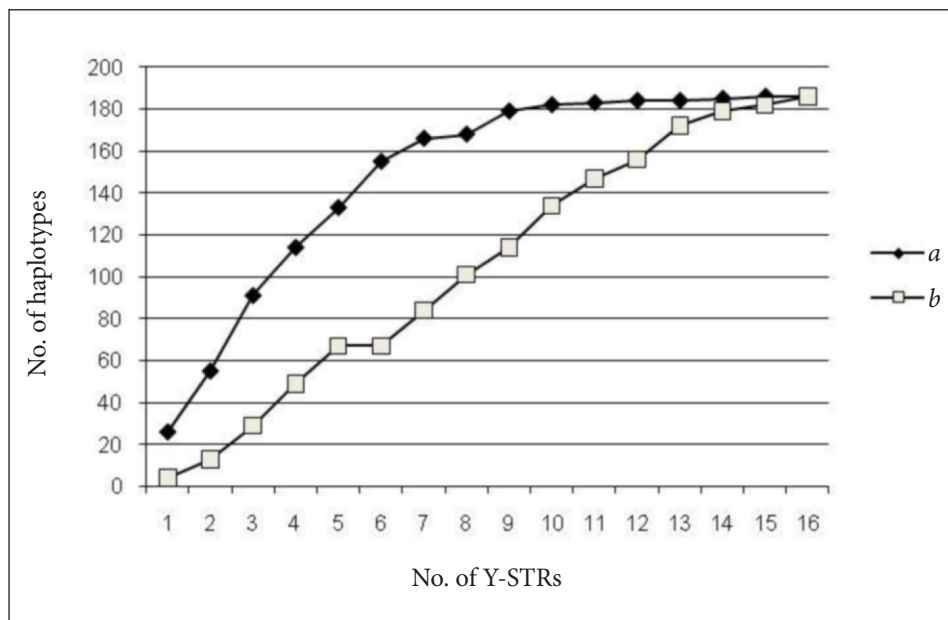


Fig. 1. Number of different Y-chromosomal haplotypes composed by combining the most variable loci to the least variable (*a*) and vice versa (*b*)

extended haplotype” (Lee et al., 2004) and the “Yfiler™ kit 17 Y-STR marker set” (Mulero et al., 2006). Also the discrimination capacity and the number of unique Y-STR haplotypes for various Y-STR marker combinations were determined (Table 2). The results indicated that the Y-STR haplotype diversity value rises as the number of analyzed Y-STRs is increased. However, all additional Y-STR loci are not hypervariable and only their whole makes large diversity of Y-STR haplotypes in Lithuanian males. As anticipated, the discrimination capacity and the number of unique Y-STR haplotypes correlate with the number of Y-STRs used in the analysis. The comparative

analysis of Y chromosomal STR haplotypes demonstrated that the “Yfiler™ kit 17 Y-STR marker set” (Mulero et al., 2006) in the Lithuanian population provides higher Y-STR haplotype diversity (0.99952 ± 0.0006) and discrimination capacity (95.9%) than the “SWGDAM extended haplotype” (Lee et al., 2004) or the “European minimal haplotype” (Gill et al., 2001). The total number of observed repeating Y chromosomal haplotypes is presented in Table 3. According to the “17 Y-STR marker set” (Mulero et al., 2006), only the 15-14-30-23-11-14-14-11,13-14-10-10-19-13-17-21-12 Y-STR haplotype occurred in three copies. The most common “SWGDAM extended haplotype”

Table 2. Y-chromosomal haplotype diversity and discrimination capacity in the Lithuanian population

	Minimal haplotype	SWGDAM extended haplotype	17 Y-STRs haplotype
Sample size	194	194	194
Different haplotypes	145	164	186
Unique haplotypes	123	148	179
Repeating haplotypes	71	46	15
HD \pm SD	0.99375 ± 0.0018	0.99664 ± 0.0013	0.99952 ± 0.0006
Discrimination capacity	74.7%	84.5%	95.9%

HD: haplotype diversity; SD: standard deviation.

Table 3. Analysis of molecular variance (AMOVA) for minimal haplotype Y-STR loci between Lithuania (n = 194) and five European neighbouring populations (n = 8029)

	Belarus	Estonia	Latvia	Poland	Russia	Lithuania
Belarus	–	0.0000	0.0000	0.0001	0.0000	0.0000
Estonia	0.0996	–	0.0028	0.0000	0.0004	0.0076
Latvia	0.0393	0.0261	–	0.0000	0.0006	0.3788*
Poland	0.0066	0.0910	0.0261	–	0.0000	0.0000
Russia	0.0412	0.0173	0.0186	0.0503	–	0.0016
Lithuania	0.0615	0.0176	0.0001	0.0479	0.0120	–

Pairwise distances based on Φ_{ST} values were calculated using RST.

p values based on 10,000 permutations are shown above the diagonal and Φ_{ST} values are below the diagonal.

* Non-significant Φ_{ST} values are bold, $p > 0.05$.

(Lee et al., 2004) 15-14-30-23-10-14-14-11,13-10-10 and the “European minimal haplotype” (Gill et al., 2001) 15-14-30-23-10-14-14-11,13 were found in 8 and 9 copies, respectively, of this Lithuanian population sample. To infer the specificity of these Y-chromosomal STR haplotypes to the Lithuanian population, we searched the YHRD database release 34 (Willuweit, Roewer, 2007). Under the given “European minimal haplotype” (Gill et al., 2001) the database (n = 87440) provided 17 copies (2 in Latvia, 5 in Poland, 7 in Germany, 1 in Finland and 2 in Russia). As for the above-mentioned “17 Y-STR marker set haplotype” (Mulero et al., 2006) and “SWGDM extended haplotype” (Lee et al., 2004), only 1 (Poland) and 5 (1 in Russia, 2 in Germany, 1 in Poland, 1 in Finland) matches among the database sample (n = 27531) and (n = 59735), respectively, were detected. Lithuanians’ Y chromosome STR parameters of forensic importance are summarized in Tables 1 and 2.

Comparative analysis of Y-STR haplotypes among Lithuanian and neighbouring populations. Analysis of molecular variance (AMOVA) was performed to investigate the degree and significance of differentiation between Lithuanian and neighbouring populations. The matrix of pairwise Φ_{ST} values is shown in Table 3. The pairwise population comparisons revealed significant interpopulation differences except the pair of Lithuanian

and Latvian populations. The lowest interpopulation distance (0.0066) was observed among closely related populations residing in Belarus and Poland territories. In contrast, the highest Φ_{ST} value was assigned between the clearly distinct populations, Belarusians and Estonians (0.0996). Between other pairs of Lithuania’s neighbouring populations interpopulation variances ranged from 0.0120 (Lithuania vs Russia) to 0.0910 (Poland vs Estonia). Pairwise distances among the populations are visualized in the MDS plot (Fig. 2). By a multidimensional scaling analysis two main groups of populations with a low interpopulation distance were distinguished: one containing the Balts, i. e. Lithuanians and Latvians, and the other consisting of Belarus and Poland populations, which represent Slavic cultures. Further distant were multicultural Russian metapopulation and Finno-Ugric speaking Estonians. However, small Φ_{ST} values ($\Phi_{ST} < 0.0996$) suggest low interpopulation variances and genetic influences from all neighbouring cultures. Focusing on the genetic relationships between all neighbouring populations, the neighbour-joining tree was structured from the Φ_{ST} distance matrix (Fig. 3). It shows that clustered Lithuanians and Latvians have a closer phylogenetic relation to Russian and Estonian populations and are less genetically related to other neighbouring Slavic populations in Belarus and Poland.

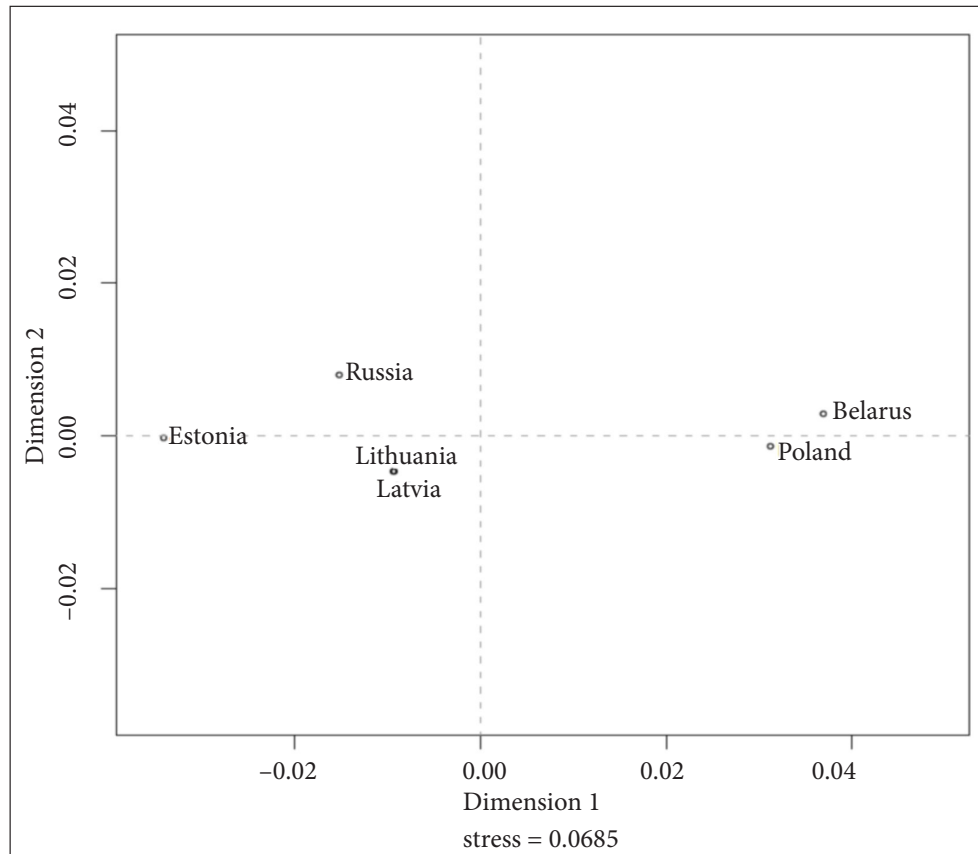


Fig. 2. Multidimensional scaling (MDS) plot based on the pairwise Φ_{ST} values from the Y-STR haplotypes of Lithuanian and five European neighbouring populations

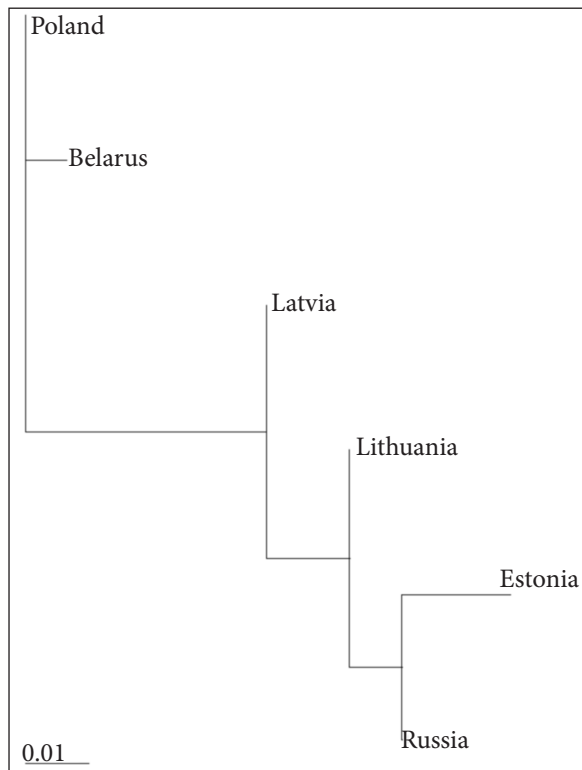


Fig. 3. Neighbour-joining tree based on the pairwise genetic distance Φ_{ST} values from the Y-STR haplotypes of Lithuanian and five European neighbouring populations

DISCUSSION

One of the reasons that Y-STRs make advantageous for population genetics studies and forensic purposes is their discrimination power and marked genetic variation, which enable to compose highly informative Y chromosome STR haplotypes. Due to the greater sensitivity of nonrecombining Y-chromosomal markers to founder effects and genetic drift, Y-STRs are very powerful in the detection of genetic differences between populations, especially among the European populations (Kayser et al., 2001). As our results indicate, a high resolution Y-chromosomal DNA profile can be achieved in the Lithuanian population (Table 2). However, the Y-STR haplotype sharing with the European populations of the close geographical neighbourhood that could be influenced by recurrent mutations and/or common ancestry is observed.

As noticed in the study (Kayser et al., 2001), with the Y-STRs, it is possible to analyse simultaneously closely and distantly related populations by AMOVA and phylogenetic analysis. To study the interpopulation genetic links according to forensic Y-STR markers we chose Lithuania's neighbouring countries, which include reference population sets representing different regions and different linguistic groups of Europe. The results revealed the genetic division into Slavic speaking Belarusians and Poles, Finno-Ugric speaking Estonians, the Balts, who speak Baltic languages, and multicultural Russian metapopulation (Fig. 2). Yet exhibited genetic division is not strict as small Φ_{ST} values suggest (Table 3).

It should be emphasized that when Russia is approached as one metapopulation of the entirely Russian ethnic group from the European part of Russia, in view of Y-STR markers it shows a close genetic relation to the core group of Slavic speaking populations of Belarus and Poland with a clear genetic distance to linguistically different groups from Estonia, Latvia and Lithuania (Roewer et al., 2008). On the other hand, we considered Russia as one multicultural metapopulation consisting of not only Slavic

speaking Russians but also of Uralic and Altaic speaking populations residing in the Siberian and Caucasian part of the Russian territory, therefore in our Y-STR analysis the Russian metapopulation is more phylogenetically related to Estonians, Latvians and Lithuanians than Belarusian and Polish populations. This genetic relationship became apparent due to genetic contribution of Altaic and Uralic languages speaking populations, which represent male founder population movements from Eastern Europe or Northern Asia to the northeast Baltic Sea region (Rosser et al., 2000), in particular where these three Baltic countries locate. Despite a close geographical vicinity, our results indicate significant population differentiation between Estonians and the Balts, i. e. Lithuanians and Latvians, that is explicable by a linguistic barrier and suggests a distinct history of these Baltic male populations. This is in agreement with the previously reported studies (Kasperavičiūtė et al., 2004; Zerjal et al., 2001; Lessig et al., 2001; Roewer et al., 2005). The similarity of the Y-chromosomal STR gene pool among Lithuanians and Latvians could be a result of the geographical proximity, linguistic coherence and common history of the populations. Also a close genetic relation between the populations of Poland and Belarus is evident as Polish and Belarusian cultures have been tied together for many centuries. Overall, genetic differences and similarities between populations revealed in the Y-STR analysis may depend on the geographical, linguistic and/or historical influences, which could come through population bottlenecks, founder effects, genetic drift, gene flow and/or non-random mating. These forces of population evolution should be considered before drawing any decisive conclusions about the genomic origin of the Lithuanian population and its genetic relationships with neighbouring populations.

From a forensic genetics perspective, this study provides an informative analysis of the Y-chromosomal STR diversity in the Lithuanian population and emphasizes the discrimination capacity of high-resolution Y-STR typing that

can be achieved in the Lithuanian population for the forensic DNA casework. Our data indicates that Y-STRs alleles and haplotypes differentiate effectively inside the Lithuanian population, and between Lithuanians and its geographical neighbours, excluding the Latvian population. This suggests that Lithuanian and Latvian populations are closely related not only by geography and language but also by the Y chromosome gene pool represented by forensic Y-STR markers. Consequently, more forensic Y-STR markers should be included in the Y-STR haplotype in order to achieve a resolution between the Y chromosomes of Lithuanian and Latvian males. Lithuanian Y-STR haplotype data were submitted to the 34th release of the Y-STR Haplotype Reference Database 3.0 for Y-STR haplotype match probability calculations in the forensic casework.

ACKNOWLEDGEMENTS

Authors of the article thank the administration of the State Forensic Medicine Service for collaboration during the research.

Received 30 June 2015

Accepted 15 August 2015

References

1. AmpFlSTR® Yfiler™ PCR Amplification Kit. User's Manual. Foster City, CA: Applied Biosystems; 2004.
2. Butler JM. Forensic DNA typing: biology, technology and genetics of STR markers. 2nd ed. London: Elsevier Academic Press; 2005.
3. Butler JM. Forensic DNA typing. London: Elsevier Academic Press; 2005.
4. Chakraborty R. Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. *Hum Biol.* 1992; 64: 141–59.
5. Česnys G. Anthropological roots of the Lithuanians. *Science, Arts and Lithuania.* 1991; 1: 4–10.
6. Evett IW, Gill P. A discussion of the robustness of methods for assessing the evidential value of DNA single locus profiles in crime investigations. *Electrophoresis.* 1991; 12: 226–30.
7. Excoffier L, Laval G, Schneider S. Arlequin (version 3.1): an integrated software package for population genetics data analysis. *Evol Bioinform.* 2005; 47–50.
8. Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.69. Distributed by the Author. Seattle: Department of Genome Sciences, University of Washington; 2005.
9. Gill P, Brenner C, Brinkmann B, Budowle B, Carracedo A, Jobling MA, et al. DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y chromosome STRs. *Forensic Sci Int.* 2001; 124: 5–10.
10. Gusmão L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR, et al. DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic Sci Int.* 2006; 157: 187–97.
11. Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al. Evaluation of Y chromosomal STRs: a multicenter study. *Int J Legal Med.* 1997; 110: 125–33.
12. Kayser M, Krawczak M, Excoffier L, Dieltjes P, Corach D, Pascali V, et al. An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. *Am J Hum Genet.* 2001; 68(4): 990–1018.
13. Kasperavičiūtė D, Kučinskas V. Mitochondrial DNA sequence analysis in the Lithuanian population. *Acta Med. Litu.* 2004; 11(1): 1–6.
14. Kasperavičiūtė D, Kučinskas V, Stoneking M. Y chromosome and mitochondrial DNA variation in Lithuanians. *Ann Hum Genet.* 2004; 68: 443–52.
15. Kučinskas V. Population genetics of Lithuanians. *Ann Hum Biol.* 2001; 28(1): 1–14.
16. Kučinskas V. Gene and gene geography of Lithuanians. *Laboratorinė medicina.* 2001; 3(11): 11–20.

17. Kučinskas V. Genomo įvairovė: lietuviai Europoje. Vilnius: Spalvų šalis; 2004.
18. Laitinen V, Lahermo P, Sistonen P, Savontaus ML. Y-chromosomal diversity suggests that Baltic males share common Finno-Ugric-speaking forefathers. *Hum Hered.* 2002; 53: 68–78.
19. Lee D, Sims G, Newman J, Laad C, Smerick S, Barna C, et al. Report on the current activities of the Scientific Working Group on DNA Analysis Methods YSTR Subcommittee. *Forensic Sci. Commun.* 2004; 6(3).
20. Lessig R, Edelmann J, Krawczak M. Population genetics of Y-chromosomal microsatellites in Baltic males. *Forensic Sci Int.* 2001; 118: 153–7.
21. Mulero JJ, Chang CW, Calandro LM, Green RL, Li Y, Johnson CL, Hennessy LK. Development and validation of the AmpFL-STR® Yfiler™ PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system. *J Forensic Sci.* 2006; 51(1): 64–75.
22. Nei M. *Molecular evolutionary genetics.* New York: Columbia University Press; 1987.
23. Page RDM. TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci.* 1996; 12: 357–8.
24. Park SDE. Trypanotolerance in West African cattle and the population genetic effects of selection [PhD Thesis]. University of Dublin; 2001. Available from: <http://animalgenomics.ucd.ie/sdepark/ms-toolkit/>
25. Pascali VL, Dobosz M, Brinkmann B. Coordinating Y-chromosomal STR-research for the courts. *Int J Legal Med.* 1998; 112(1): 1.
26. Rand S, Schurenkamp M, Brinkmann B. The GEDNAP (German DNA profiling group) blind trial concept. *Int J Legal Med.* 2002; 116: 199–206.
27. Rand S, Schurenkamp M, Hohoff C, Brinkmann B. The GEDNAP blind trial concept. Part II. Trends and developments. *Int J Legal Med.* 2004; 118: 83–9.
28. Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawczak M, de Knijff P. Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related human populations. *Hum Mol Genet.* 1996; 5(7): 1029–33.
29. Roewer L, Kayser M, de Knijff P, Anslinger K, Betz A, Caglià A, et al. A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. *Forensic Sci Int.* 2000; 114: 31–43.
30. Roewer L, Croucher P, Willuweit S, Lu TT, Kayser M, Lessig R, et al. Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. *Hum Genet.* 2005; 116: 279–91.
31. Roewer L, Willuweit S, Krüger C, Nagi M, Rychkov S, Morozowa I, et al. Analysis of Y chromosome STR haplotypes in the European part of Russia reveals high diversities but non-significant genetic distances between populations. *Int J Legal Med.* 2008; 122(3): 219–23.
32. Rosser Z, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, et al. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet.* 2000; 67: 1526–43.
33. Zerjal T, Pandya A, Santos FR, Adhikari R, Tarazona E, Kayser M, et al. The use of Y-chromosomal DNA variation to investigate population history: recent male spread in Asia and Europe. In: Papiha SS, Deka R, Chakraborty R, editors. *Genomic diversity: Applications in human population genetics.* New York: Kluwer Academic/Plenum Publishers; 1999. p. 91–101.
34. Zerjal T, Beckman L, Beckman G, Mikesaar AV, Krumina A, Kucinskas V, Hurles ME, Tyler-Smith C. Geographical, linguistic, and cultural influences on genetic diversity: Y chromosomal distribution in Northern European populations. *Mol Biol Evol.* 2001; 18: 1077–87.
35. Willuweit S, Roewer L on behalf of the International Forensic Y Chromosome User Group. Y Chromosome Haplotype Reference Database (YHRD): update. *Forensic Sci Int Genet.* 2007; 1: 83–7. Available from: <http://www.yhrd.org>

**Giedrė Ruzgaitė, Marija Čaplinskienė,
Rima Baranovienė, Jūratė Jankauskienė,
Jolanta Kukienė, Kristina Savanevskytė,
Daiva Bunokienė**

Y CHROMOSOMOS STR ŽYMENŲ ANALIZĖ LIETUVOS POPULIACIJOJE TEISMO GENETIKOS ASPEKTAIS

Santrauka

Darbe pateikta pirmoji plataus masto JAV ir Europos teismo genetinės ekspertizės jurisdikcijose „auksiniu standartu“ laikomų Y chromosomos STR žymenų analizė Lietuvos populiacijoje. Siekiama įvertinti Lietuvos populiacijos Y chromosomos genofondo įvairovę; nustatyti genetines sąsajas tarp Lietuvos ir kaimyninių populiacijų; sudaryti Lietuvos populiacijos genetinę struktūrą ir įvairovę reprezentuojančią Y-STR haplotipų duomenų bazę, kuri yra būtina atliekant teismo genetinę ekspertizę DNR profilių lyginamojo tyrimo išvadų statistiniam patikimumui nustatyti. Tyrimo objektą sudaro 194 giminystės ryšiais tarpusavyje nesusijusių Lietuvos populiacijos vyrų genominės DNR profiliai, surinkti iš visų Lietuvos Respublikos apskričių. Tiriamųjų Y chromosomos STR genetinės sritys pagausintos naudojant „AmpFI STR® Yfiler™ PCR Amplification Kit“ komercinį rinkinį pagal gamintojo protokolą. Tyrimo rezultatai parodė, kad Y chromosomos haplotipų įvairovė Lietuvos populiacijoje gausėja didėjant analizuojamųjų Y-STR žymenų skaičiui, tačiau tiriamieji Y-STR žymenys nėra itin varia-

bilūs ir tik tiriamųjų Y-STR žymenų bendra visu-
ma sukuria didelę Y chromosomos STR haplotipų
įvairovę Lietuvos vyrų populiacijoje. Molekulinės
genetinės įvairovės dispersinė analizė pagal Y-STR
žymenis atskleidė mažą, bet statistiškai reikšmin-
gą genetinį atstumą tarp visų tiriamųjų populiaci-
jų porų, išskyrus Lietuvos ir Latvijos populiacijų
porą. Filogenetinės analizės duomenimis, Lietuvos
ir Latvijos Y chromosomos genofondas pagal ti-
riamuosius Y-STR žymenis yra filogenetiškai arti-
mesnis Rusijos ir Estijos nei Baltarusijos ir Lenkijos
populiacijų Y chromosomos genofondui. Nustatyta,
kad tiriamųjų Y-STR žymenų aleliai ir haplotipai
gali reikšmingai skirtis tiek Lietuvos populiacijoje,
tiek tarp Lietuvos ir kaimyninių populiacijų, išsky-
rus Latvijos populiaciją. Tai rodo, kad Lietuvos ir
Latvijos populiacijos yra artimos ne tik geografiškai
ir kalbiškai, bet ir genetiškai pagal Y chromosomos
genofondą, reprezentuojamą JAV ir Europos teis-
mo genetinės ekspertizės jurisdikcijose pripažinta
Y-STR žymenų sistema. Taigi, daugiau Y-STR žyme-
nų turėtų būti analizuojama siekiant diferencijuoti
Lietuvos ir Latvijos Y chromosomos haplotipus.
Lietuvos populiacijos genetinę struktūrą ir įvairovę
reprezentuojantys Y chromosomos STR haplotipai
įtraukti į tarptautinę Y chromosomos STR popu-
liacinių palyginamųjų duomenų bazę, naudojamą
teismo genetinėje ekspertizėje „Y-STR Haplotype
Reference Database 3.0“.

Raktažodžiai: Lietuvos populiacija, genetinė
įvairovė, Y chromosoma, STR, teismo genetika