

Genes and gene geography of Lithuanians

Vaidutis Kučinskas

Abstract. The primary objective of this article was to overview the present-day knowledge on genetic features of the Lithuanian population. Genetic differentiation within the Lithuanian population and the relationship between Lithuanians and other European populations was analysed by means of blood groups, serum protein polymorphisms and DNA markers including mtDNA and Y chromosome markers. The results of the research have shown small differences between present-day Lithuanian ethnolinguistic groups, which probably go back to the pre-historic Baltic tribal structure. The Baltic peoples show a mixture of eastern and western genetic traits, e. g. a high frequency of the blood group B combined with a very high frequency of the Rh-negative blood group. Studies of the Baltic "tribal gene" LWb indicate the presence of a considerable Baltic admixture in the neighbouring Finno-Ugric and Slavic populations.

Keywords: Lithuanians, gene geography, blood groups, protein polymorphism, DNA markers.

INTRODUCTION

Migrations and admixture tend to obliterate the unique and peculiar genetic features of original historical populations. The Lithuanian population is an interesting object for genetic studies, but Lithuania has long been a 'white spot' on the population genetic map of Europe. Although initial studies of blood groups had been performed in Lithuania in the first decades of the 20th century [1, 2, 3], contemporary genetic investigation of the Lithuanian population started just 15 years ago and revealed some internal genetic variation [4]. This article is a contribution to the population genetics of Lithuania.

HISTORICAL BACKGROUND

The original populations of the Baltic Sea region have widely different cultural and linguistic origins (figure 1, A). Thus those on the western (Scandinavian) side speak Germanic languages, whereas those on the eastern

(Baltic) side speak Baltic or Finno-Ugric languages.

Balts who gave rise to the Lithuanians were in ancient times surrounded by other Baltic tribes, who spoke closely related languages and did rarely come into contact with non-Baltic tribes. Having lived for a long time far from foreign influences, the future Lithuanians preserved the ancient social structure and also the archaic nature of their language [5].

The Lithuanian language is a descendant from the Baltic branch of the Indo-European family. Baltic languages originated from the Proto-Indo-European language region from which also the Slavic and Germanic languages descended. Only two languages of this branch have survived to the present day: Lithuanian and Latvian.

In pre-historic times, the Baltic languages were spoken in a large area from the Vistula in the west and to the Dnepr River Basin and the upper reaches the Volga, the Oka and the Moscow rivers in the east (figure 2). In the period between 4000 YBP and 3000 YBP the Baltic tribes are be-

Human Genetics Centre, Vilnius
University, Vilnius, Lithuania

Correspondence to:

Prof. V. Kučinskas, Human Genetics
Centre of Vilnius University,
Santariškių street 2, LT-2021 Vilnius,
Lithuania;
phone: +370-2-365195,
fax: +370-2-365199,
e-mail: vaidutis.kucinskas@mf.vu.lt

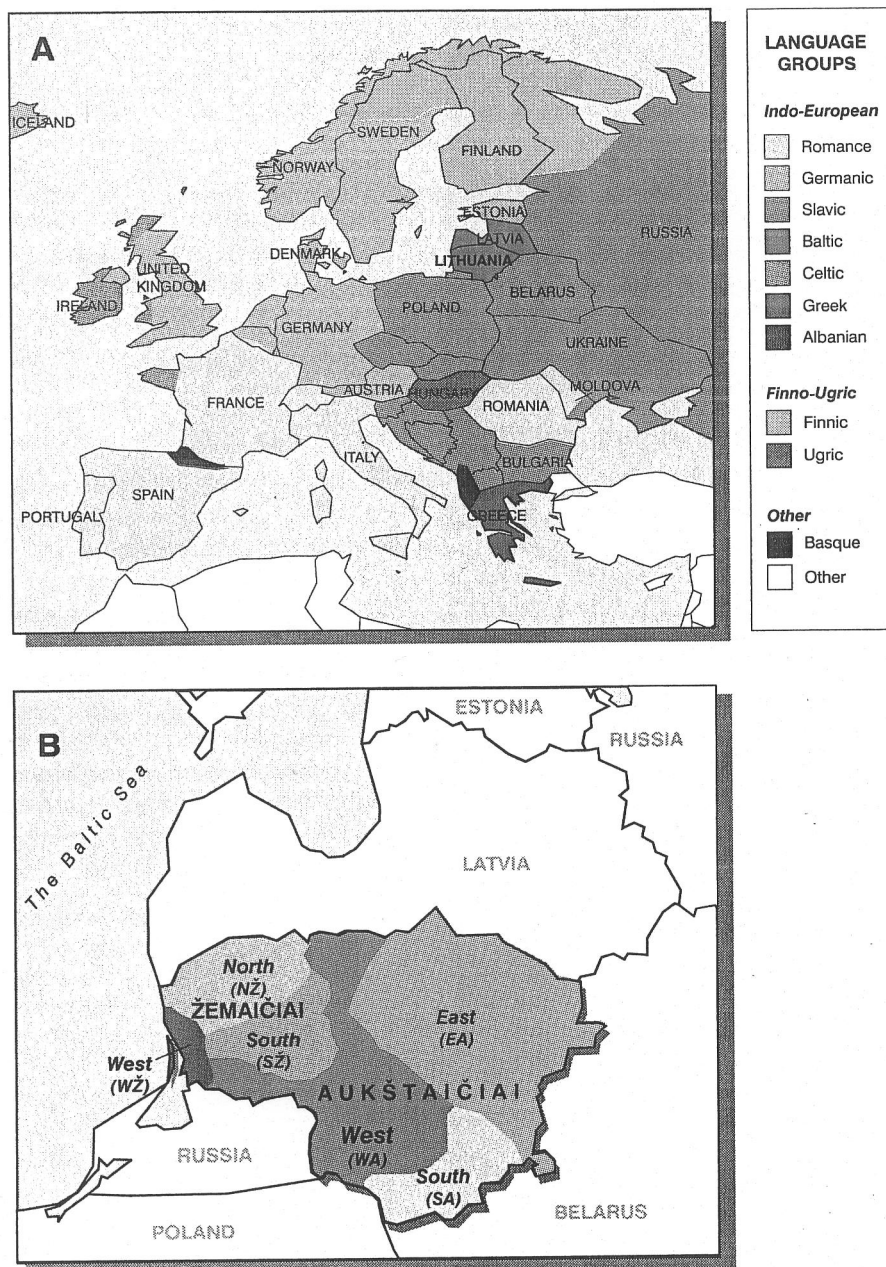


Figure 1. A, European language groups. B, Lithuanian ethnolinguistic subgroups

lieved to have inhabited a larger territory than the Slavic and the Germanic peoples [5]. This large Baltic territory was at that time covered by impenetrable forests far from the major migration and trade routes. Thus the Balts appear to have lived for a long time in relative isolation. This facilitated the preservation of the original Indo-European way of life, the ancient cultural stratification and the extremely archaic language that had hardly changed from the Proto-Indo-European.

The West and East Balts began to diverge from one another around 3000 YBP and were finally separated around 2400–2300 YBP. No West Baltic tribes or languages have survived to the present. After their separation

from the West Balts, the East Balts continued to exist for about a thousand years as a tribal group with some internal differentiation.

In the middle of the first millennium A.D. or somewhat earlier Slavic tribes began to force their way into the territory inhabited by the Balts. They came suddenly, moved swiftly, perhaps in several waves along the Dnepr, Sozh, Desna and other rivers in a north and north-easterly direction into the very heart of the Baltic territory. They reached the present-day Pskov–Novgorod region and in the 5th–6th centuries came into direct contact with the northern neighbours of the Balts, the Finno-Ugric tribes. In this manner a deep Slavic wedge was driven into the ancient Baltic territory

which was split geographically into the western portion – the historic Prussian, Latvian and Lithuanian lands – and the remaining Baltic tribes in the east which were scattered in a large sparsely inhabited area. The Balts who remained were eventually assimilated by the Slavs.

In the northern region of the East Balts there was a considerable admixture with Finno-Ugric tribes. According to available data, around 2500 YBP the Finno-Ugric tribes still inhabited all of the northern Curonian lands up to the Abava River, the Riga shore and further to the east encompassing all of Latvia to the north of the Daugava River. In the northern area the Curonian, Semigalian, Selonian and Lettigalian tribes and languages began to merge. The fusion of these tribes gave rise to the Latvian nation. In the south the East Baltic tribes remained and formed the Lithuanian nation.

Lithuanians were first mentioned in historical documents at the beginning of the 11th century. At that time the name Lithuania was most likely not used for the entire present-day Lithuanian region, but only for a portion of it, most likely the territory between the Neris, Nemunas and Merkys Rivers. The initial unification of Lithuanian lands into one state probably took place somewhere in this vicinity. To the south Lithuania bordered to the old Yotvingian lands and to the north it reached the lands of Žemaičiai. To the southwest the early Lithuanians bordered to the West Balts – the Prussian tribes. At that time a very intensive consolidation of the related Baltic tribes (figure 2) took place. This contributed greatly to the formation of the present-day Lithuanian dialects. The Curonian (in the west of Lithuania) substratum had a particularly distinct influence on the formation of the current Žemaitish dialect. During the German invasion (in the middle of the 13th century) many Curonians were killed and many migrated from their lands or were forcibly moved to other areas. In almost two hundred years of fierce battles between the Lithuanians and the Teutonic Order, many of the older inhabitants from Nadruva and Scalovia died or retreated further from the battle zone. The Yotvingian lands conquered by the Teutonic knights were brutally devastated and the Yotvingian territory became a wasteland, which was resettled once the wars had subsided. Into this territory came Lithuanians and East Slavs (the Byelorussian an-



Figure 2. Baltic tribes (A.D. 1200)

cestors) who had lived in the Grand Duchy of Lithuania.

During later centuries the Lithuanian State (Grand Duchy of Lithuania) was gradually losing its territory and finally its independence. Lithuanian tribes consolidated in the area of the present-day Republic of Lithuania (gained independence in 1918 and regained it in 1990). Thus the contemporary population of Lithuania, which is an object of the present study, is comprised of a complex mixture of former

Baltic tribes and states. The main periods in the formation of the Lithuanian population are shown on figure 3.

Since the Neolithic period the native inhabitants of the Lithuanian territory have not been replaced by any other ethnic group. In other words, the roots of the present-day Lithuanian population are deep and the probability that the inhabitants of the present-day Lithuania have preserved the ancient genetic composition is high [6].

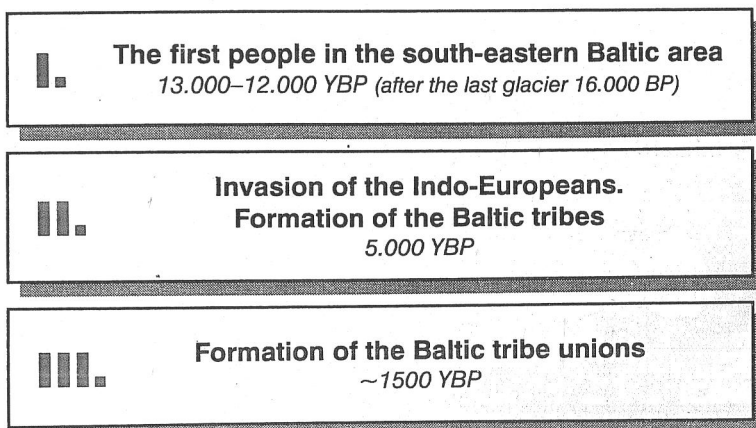


Figure 3. The main periods in the formation of the Lithuanian population. YBP, years before present

DATA USED IN THIS REVIEW

Blood samples were collected from unrelated individuals from six ethnolinguistic subgroups: North Žemaičiai, West Žemaičiai, South Žemaičiai, East Aukštaičiai, West Aukštaičiai, and South Aukštaičiai [5]. These six subgroups later were grouped into two main regions of Lithuania: Žemaičiai and Aukštaičiai (figure 1, B) according to Z. Zinkevičius [5].

In this review the following genetic systems are discussed:

1) Blood groups: A1A2B0, Rh (C, c, D, E, e), MNS, P, Lutheran (Lu), Kell, Lewis, Secretor (FUT2), Duffy (Fy), Kidd (Jk), Landsteiner Wiener (LW), Colton (Co);

2) Serum proteins: TF, Gc, α 1-AT;

3) mt-DNA RFLPs (*Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II, *Hinc*II) and hypervariable region I (HVI), RSA1 RFLPs;

4) Frequency distribution of:

a) CFTR gene mutation Δ DF508,

b) PAH gene mutation R408W/RFLP haplotype/VNTR/STR minihaplotypes,

c) peripherin/RDS gene mutations Glu304Gln, Lys310Arg and Gly338Asp,

e) CMKBR5 gene mutation Δ Dccr5 [7],

f) recent Alu insertions: ACE, TPA25, APO, FXIIIIB, A25, PV92, D1, B65, HS2.43, HS4.14, HS3.23, HS4.32, HS4.59, HS4.65, HS4.69, HS4.75;

5) Y chromosome haplotype variation according to 11 biallelic markers: 92R7, SRY-1532, DYS257, YAP, SRY-2627, M9, sY81, Tat, SRY-8299, 12f2, LLY22g.

GENETIC VARIATION BETWEEN DIFFERENT LITHUANIAN ETHNOLINGUISTIC SUBGROUPS

Data on some allelic frequencies in the Lithuanian ethnolinguistic subgroups are summarized in tables 1 and 2 [8–15]. Table 1 shows some selected allele and haplotype frequencies of different blood group systems.

The A2 gene in the A1A2B0 system was most frequent in the South Aukštaičiai, while in the West and North Žemaičiai it was not significantly less frequent ($p > 0.05$) [12, 16]. There was a possible gradient (northwest to southeast) of the A2 frequency. The absolute values of the A2 frequencies were low compared to Finns and Swedes.

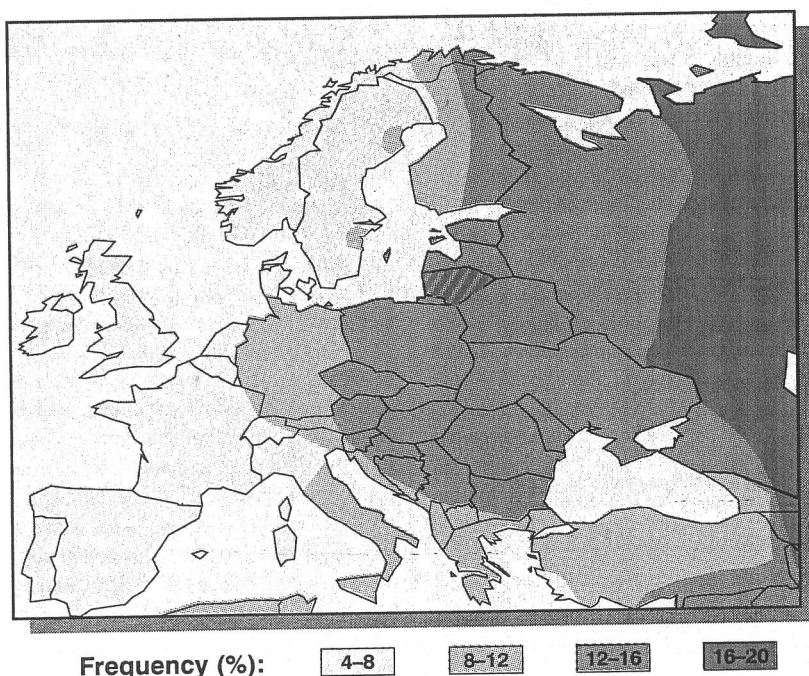
Table 1. Blood group markers in Lithuanian ethnolinguistic subgroups [8, 11, 12]

| Ethno-linguistic sub-group | | Genetic markers | | | | | | | | | | | |
|----------------------------|-------|--------------------------------------|--------------------|--------------------|--------------------|--------------------|------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|-------------------------|
| | | A2 (A1A2B0) | cde (Rh) | Ms (MNSs) | P1 (P) | Lua (LU) | K Kpb U1 (A) (Kell) | Cob (Co) | Le (Lewis) | Se (FUT2) | Fya (FY) | Jka (JK) | LW ^b (LW) |
| | | F ^a (SD ^b) | F (SD) | F (SD) | F (SD) | F (SD) | F (SD) | F (SD) | F (SD) | F (SD) | F (SD) | F (SD) | F (SD) |
| Aukštaičiai | East | 0.0293 (0.0103) | 0.4504 (0.0281) | 0.3806 (0.0234) | 0.4284 (0.0310) | 0.0057 (0.004) | 0.0429 (0.0108) | 0.0289 (0.0090) | 0.3703 (0.0206) | 0.6825 (0.0386) | 0.4400 (0.0265) | 0.5200 (0.0267) | 0.0657 (0.0132) |
| | West | 0.0447 (0.0147) | 0.4032 (0.0336) | 0.2726 (0.0249) | 0.5064 (0.3421) | 0.0195 (0.009) | 0.0430 (0.0127) | 0.0520 (0.0141) | 0.3644 (0.0233) | 0.6298 (0.0440) | 0.4766 (0.0312) | 0.5195 (0.0312) | 0.0508 (0.0137) |
| | South | 0.0506 (0.0156) | 0.4773 (0.0307) | 0.3177 (0.0270) | 0.3421 (0.0327) | 0.0303 (0.011) | 0.0417 (0.0123) | 0.0346 (0.0113) | 0.3371 (0.0232) | 0.6768 (0.0437) | 0.3902 (0.0300) | 0.5144 (0.0308) | 0.0265 (0.0099) |
| Žemaičiai | North | 0.0282 (0.0106) | 0.4867 (0.0293) | 0.2985 (0.0287) | 0.4749 (0.0344) | 0.0131 (0.0065) | 0.0327 (0.0018) | 0.0468 (0.0122) | 0.3792 (0.0226) | 0.7205 (0.0419) | 0.4608 (0.0285) | 0.5131 (0.0286) | 0.0752 (0.0151) |
| | South | 0.0316 (0.0118) | 0.4145 (0.0321) | 0.2683 (0.0288) | 0.3746 (0.0328) | 0.0177 (0.0079) | 0.0355 (0.0110) | 0.0735 (0.0158) | 0.3151 (0.0231) | 0.6654 (0.0418) | 0.4929 (0.0298) | 0.5674 (0.0295) | 0.0603 (0.0142) |
| | West | 0.0243 (0.0121) | 0.4629 (0.0381) | 0.2426 (0.0256) | 0.4883 (0.0429) | 0.0250 (0.011) | 0.0200 (0.0099) | 0.0887 (0.0206) | 0.3742 (0.0273) | 0.6081 (0.0495) | 0.4400 (0.0351) | 0.5550 (0.0351) | 0.0550 (0.0161) |

^a Frequency; ^b Standard deviation

Table 2. Some serum protein and DNA markers in Lithuanian ethnolinguistic subgroups [11, 8, 9, 14, 13]

| Ethno-linguistic subgroup | | Genetic markers | | | | | | | | | | | | | | | |
|---------------------------|-------|-----------------|--------|--------|--------|--------|--------|--------|--------|--------|----------------|----------------|-------------------------------------|----------------|----------------|----------------|----------------|
| | | TF | | | GC | | PI | | | | Peripherin/RDS | | Alu insertions | | | | |
| | | TF*C1 | TF*C2 | TF*C3 | Gc*1F | Gc*1S | PI*M2 | PI*M3 | PI*S | PI*Z | 2 ^a | 3 ^b | ACE | TPA25 | APO | FXIIIB | A25 |
| | | | | | | | | | | | | | if ^c het ^d | if het | if het | if het | if het |
| Aukštaičiai | East | 0.8415 | 0.0732 | 0.0732 | 0.1098 | 0.6585 | 0.0854 | 0.1341 | 0.0366 | 0.0122 | 0.2300 | 0.1100 | 0.400 0.300 | 0.550 0.400 | 0.925 0.050 | 0.075 0.050 | 0.050 0.100 |
| | West | 0.7727 | 0.1273 | 0.0773 | 0.0654 | 0.6168 | 0.1500 | 0.1000 | 0.0045 | 0.0000 | 0.3400 | 0.1400 | 0.300 0.400 | 0.550 0.300 | 0.975 0.050 | 0.300 0.100 | 0.125 0.250 |
| | South | 0.8144 | 0.1082 | 0.0773 | 0.1082 | 0.6546 | 0.1649 | 0.0722 | 0.0103 | 0.0155 | 0.3300 | 0.0900 | 0.425 0.450 | 0.375 0.750 | 0.950 0.100 | 0.325 0.150 | 0.050 0.100 |
| Žemaičiai | North | 0.7600 | 0.1400 | 0.0950 | 0.1000 | 0.6000 | 0.2200 | 0.0550 | 0.0200 | 0.0050 | 0.2500 | 0.0600 | 0.375 0.150 | 0.625 0.750 | 0.925 0.150 | 0.175 0.150 | 0.025 0.050 |
| | South | 0.7431 | 0.1600 | 0.0963 | 0.0596 | 0.6835 | 0.1651 | 0.0688 | 0.0229 | 0.0183 | 0.3200 | 0.1200 | 0.525 0.350 | 0.550 0.400 | 0.975 0.050 | 0.150 0.100 | 0.100 0.200 |
| | West | N ^e | N | N | N | N | N | N | N | N | 0.3300 | 0.1100 | 0.305 0.300 | 0.475 0.550 | 0.875 0.150 | 0.200 0.200 | 0.075 0.150 |

^a Glu304Gln+Gly338Asp; ^b Glu304Gln+Lys310Arg+Gly338Asp; ^c Insertion frequency; ^d Heterozygosity; ^e Not investigatedFigure 4. Prevalence rate of the I^B allele (A1A2B0 blood group system) in Europe

The frequency of blood group B (IB allele), which is a marker of eastern (Asiatic) influence, is relatively high in the Balts (figure 4).

The frequency of the Rh haplotype cde was very high in the Lithuanian ethnolinguistic subgroups with a maximum value in the North Žemaičiai (figure 5) [12, 16].

The frequency of the Ms haplotype of MNSs blood group system varied around 30% in the Lithuanian subgroups. In East Aukštaičiai subgroup it was higher although not significantly ($p > 0.05$) than in the other subgroups. This subgroup is geographically close to East Latvians, who have a relatively high Ms haplotype frequency in comparison to Lithuanians. Thus the high frequency of Ms in the East Aukštaičiai may be due to admixture with Latvians.

The frequency of P1*1 allele (P blood group system) also varied in

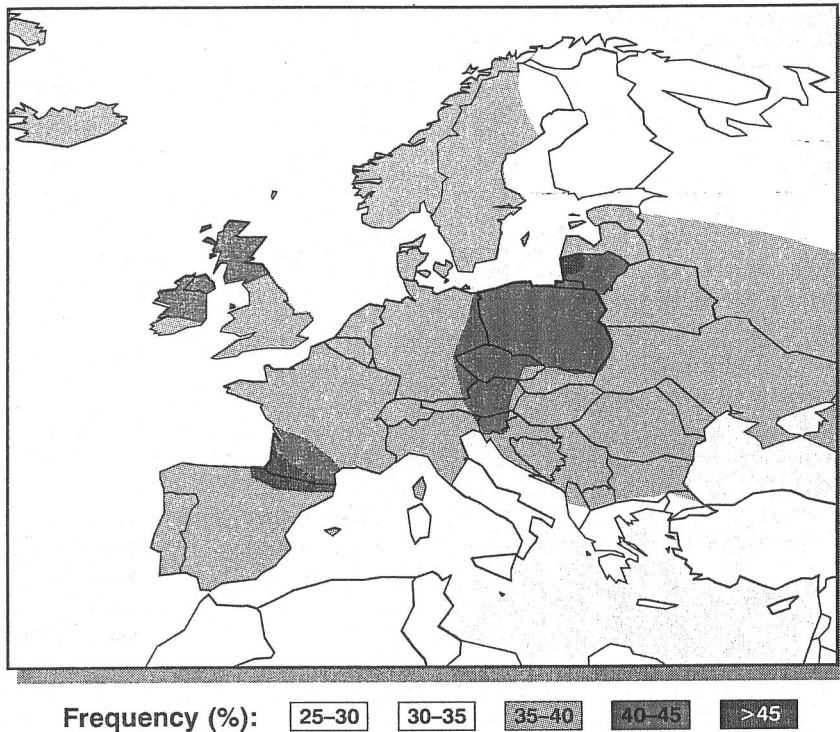


Figure 5. Prevalence rate of the cde (Rh-) haplotype (Rh blood group system) in Europe

Lithuania in a relatively wide range. The lowest its value was found in the South Aukštaičiai subgroup (34.21%) and it was statistically significant ($p < 0.01$) in comparison to other ethnolinguistic subgroups. This frequency is more than 10% lower than in other populations of northern Europe [17]. The frequency of Lu_a gene (Lutheran blood group system) varied in the range of 1%–2% in most European populations. The value was found in East Aukštaičiai subgroup was particularly low ($p > 0.05$).

The frequency of KKpb UI(A) gene (Kell blood group system) showed very little variation in the Žemaičiai and Aukštaičiai (around 3% and 4%). The frequency of Cob gene (Colton blood group system) varied between 2% and 8% in the Lithuanian subgroups, and was lower in the Aukštaičiai although statistically not significant. The frequency of the le allele (Lewis blood group system) appeared to be rather uniform in the Lithuanian population. The Aukštaičiai group was found to be homogeneous with respect to the frequency of the

FUT2 gene (Secretor blood group system) while Žemaičiai were more variable. The frequency of the Fya allele (Duffy blood group system) was similar to European figures with exception for a lower (although statistically not significantly) value in the South Aukštaičiai. The frequency of the Jka gene (Kidd blood group system) showed a very small variation

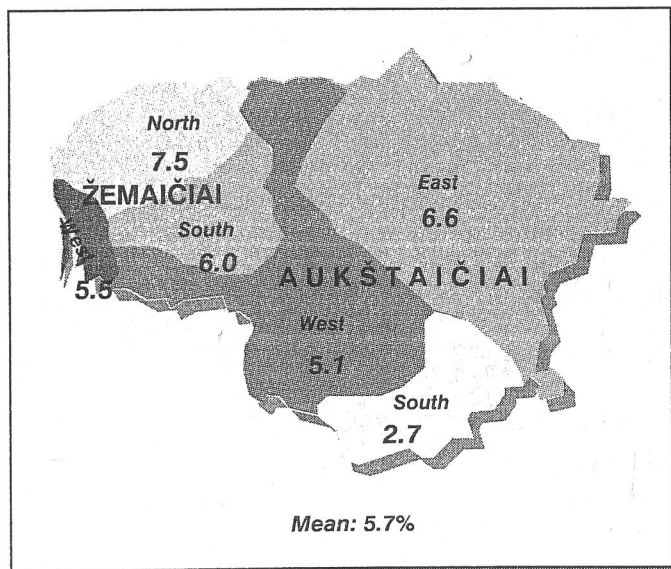


Figure 6. Frequencies (%) of the LW^b allele (Landsteiner Wiener blood group system) in Lithuanian ethnolinguistic groups

between Lithuanian ethnolinguistic groups.

The results for the Landsteiner Wiener blood group system are very interesting and deserve special attention. The frequencies of LW phenotypes and the LW^b allele in the six Lithuanian ethnolinguistic subgroups are shown in figure 6 [11, 15]. In all population samples there was a good agreement between observed and expected phenotype frequencies assuming a Hardy–Weinberg equilibrium. The LW^b allele frequency varied between 7.5% in the North Žemaičiai and 2.7% in the South Aukštaičiai (figure 6). In the latter subgroup it was significantly lower than in the rest of Lithuania ($p = 0.02$). LW^b is assumed a Baltic 'tribal' gene [18]. Its peculiar geographic distribution, relatively high frequencies in the Balts and zero or almost zero levels in Asiatic, African and West European populations makes this blood group very informative marker for a Baltic genetic influence, i. e. its presence in other populations in any appreciable frequency is likely to be due to a Baltic admixture. The finding of the highest LW^b allele frequency in the North Žemaičiai and the lowest one in the South Aukštaičiai is in agreement with the assumption that the North Žemaičiai subgroup is an autochthonous population with Curonian influence in the North, whereas the South Aukštaičiai subgroup in the south may involve some influence from the Yotvingian gene pool.

Data on some serum protein and DNA markers are shown in table 2. Comparison of the variation in the frequencies of the allelic genes TF (transferrin), GC (group-specific protein) and PI (α_1 -antitrypsin) (table 2) shows that these frequencies only slightly differ between the Žemaičiai and Aukštaičiai. In two adjacent subgroups, the West Aukštaičiai and South Žemaičiai low frequencies of the $GC*1F$ allele were found.

Direct DNA analysis also revealed some variation in nucleotide sequence variation in different Lithuanian subgroups. The peripherin/RDS gene polymorphisms $Glu304Gln + Gly338Asp$ were detected in approximately 30% in the Lithuanian population investigated [14]. Similar results were obtained from the analysis of the frequency of the polymorphisms $Glu304Gln + Lys310Arg + Gly338Asp$ of this gene: East Aukštaičiai and North Žemaičiai showed lower values.

The Lithuanian population was also examined with respect to a num-

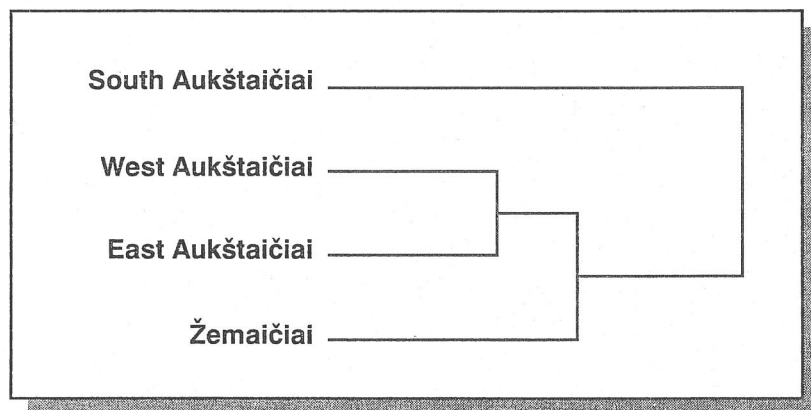


Figure 7. Results of clustering of the Lithuanian ethnolinguistic subgroups on the basis of dermatoglyphic features

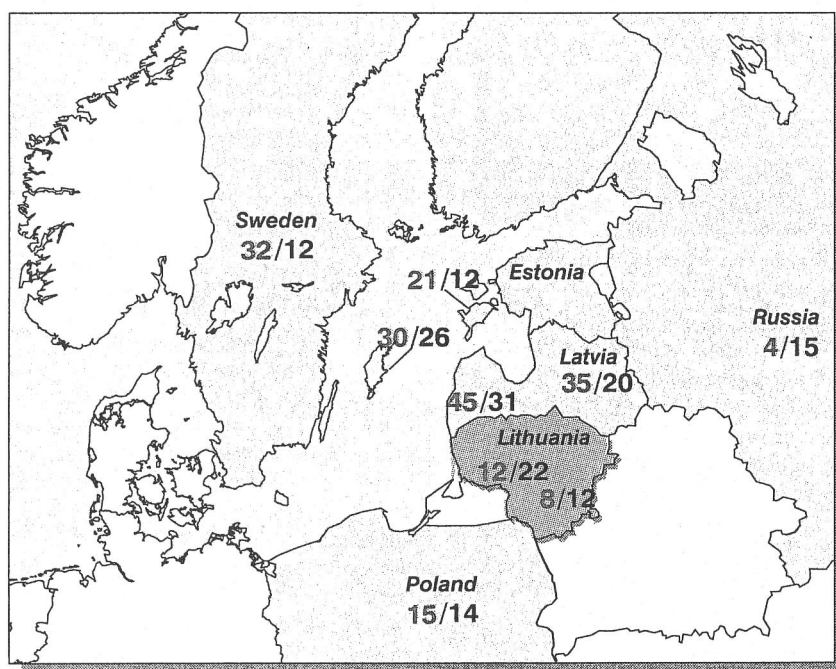


Figure 8. Frequencies (per 10,000 population) of the α 1-antitrypsin (PI) gene alleles Z (red numbers) and S (black numbers) in the Baltic Sea region

Table 3. Frequencies of mtDNA haplotypes^a in the Lithuanian population [20]

| mtDNA haplotype | Lithuanians | | Aukštaičiai | Žemaičiai |
|-----------------|-----------------------|-----------|-------------|-----------|
| | <i>n</i> ^b | Frequency | Frequency | Frequency |
| 2-1-1-1-1-2 | 125 | 0.8117 | 0.791 | 0.889 |
| 2-1-1-1-8-2 | 2 | 0.0130 | 0.018 | 0 |
| 2-1-1-1-2-2 | 7 | 0.0454 | 0.064 | 0 |
| 2-1-2-1-1-2 | 5 | 0.0325 | 0.046 | 0 |
| 2-1-1-1-3-2 | 1 | 0.0065 | 0.009 | 0 |
| 2-3-3-1-2-2 | 3 | 0.0194 | 0.027 | 0 |
| 2-3-3-1-1-2 | 1 | 0.0065 | 0.009 | 0 |
| 2-1-6-1-1-2 | 1 | 0.0065 | 0.009 | 0 |
| 2-1-1-1-15-2 | 1 | 0.0065 | 0.009 | 0 |
| 2-1-12-1-2-2 | 1 | 0.0065 | 0.009 | 0 |
| 5-1-1-1-1-2 | 1 | 0.0065 | 0.009 | 0 |
| 2-1-4-1-1-2 | 1 | 0.0065 | | 0.028 |
| 2-3-1-1-3-2 | 2 | 0.0130 | | 0.055 |
| 2-3-12-1-1-2 | 1 | 0.0065 | | 0.028 |

^a Morphs are presented in the following order: *HpaI*-*Bam*HI-*Hae*II-*Msp*I-*Ava*II-*Hinc*II

^b Number of individuals

ber of Alu insertions [13, 19]. The frequencies of insertions in the ethnolinguistic subgroups varied independently without any distinguishable pattern. Due to the small numbers in each group ($n=20$) the power to detect significant differences was low. However, in agreement with some other markers TPA25 Alu insertion showed a significant difference between the South Aukštaičiai and North Žemaičiai ($p=0.025$).

Thus, analysis of genetic markers in the six ethnolinguistic subgroups of the Lithuanians shows that most of the variation is insignificant and only the South Aukštaičiai and North Žemaičiai subgroups can be distinguished with some degree of certainty. These two subgroups appeared to be statistically distinguishable in some cases. Moreover, the South Aukštaičiai formed a definite cluster calculated from genetic distances according to a number of dermatoglyphic features and using different methods for the estimation of genetic distance and clustering (figure 7) [4].

In this context it is important to recall the background distribution of Baltic tribes. It is likely that the observed frequency variations of investigated genes in different ethnolinguistic subgroups reflect gene pools of these original Baltic tribes. Thus, an influence of the gene pool of Curonians formerly inhabiting the northwestern part of the present-day territory of Lithuania might be present in the contemporary North Žemaičiai. Also a genetic influence of the Yotvingians – former inhabitants of southern part of contemporary Lithuania and the neighbouring territory of Poland – may have caused the deviation of South Aukštaičiai from the rest of Lithuanians.

Other minor genetic differences observed between the two main ethnolinguistic groups of Lithuanians, Aukštaičiai and Žemaičiai. These two groups have been developing for a long time as two independent Baltic tribes Lietuviai (Aukštaičiai) and Žemaičiai and it is likely that their original gene pools are reflected in the present-day differences between these two ethnolinguistic groups. Such differences are particularly evident in the analysis of mt-DNA RFLPs using restriction enzymes *Hpa*I-*Bam*HI-*Hae*II-*Msp*I-*Ava*II-*Hinc*II [20]. Only haplotype 1 appeared to be common for the Aukštaičiai and Žemaičiai groups (table 3) while other haplotypes were found just in either Aukštaičiai or in Žemaičiai.

COMPARISON BETWEEN LITHUANIANS AND OTHER EUROPEAN POPULATIONS

Lithuanians and Latvians generally show similar gene frequencies, but there are also some differences e. g. Latvians (especially Curonians) have significantly higher frequencies of the α 1-antitrypsin (PI) deficiency alleles S and Z, two typically Western European mutations (figure 8) [8]. The Baltic peoples had previously a much wider geographic distribution when they came into contact with populations of Finno-Ugric, Germanic and Slavic origins, which may be one reason for the finding of both western and eastern genetic influences in the Balts. The frequency of the blood group B, which is a marker of eastern (Asiatic) influence, is relatively high in the Balts (figure 4), but the frequency of the Rh-negative gene (homozygosity for the cde haplotype) is very frequent and similar to that in the Basques (figure 5) [17].

Some genetic markers, e. g. blood group B, are unspecific markers of eastern influence unable to discriminate between Finno-Ugric and Baltic peoples. There are, however, also specific (tribal) markers for Finno-Ugric and Baltic influence. Some genetic transferrin variants, particularly TF*DCHI, are Finno-Ugric markers. TF*DCHI is e. g. found in Finns and Estonians, but not in Latvians and Lithuanians (figure 9, A). The previously mentioned blood group LWb is a Baltic tribal marker, but interestingly its frequency is also relatively high in Estonians, Finns, Russians and Poles suggesting a Baltic genetic influence in the neighbouring Finno-Ugric and Slavic populations (figure 9, A) [15]. The occurrence in Finns is in agreement with the frequent occurrence of Baltic loanwords in the Finnish language [18] and the import to Finland of the Battle-Axe culture from the Baltic area.

The analysis of mitochondrial DNA (mtDNA) has proved to be a useful tool in our understanding of the origin and migration patterns of Lithuanians. Direct sequencing of mtDNA hypervariability region I (HV1) showed that 73% of the obtained nucleotide sequences can be classified to clusters defined by M. Richards and colleagues [21]. The remaining sequences presented nucleotide substitutions, which did not allow associating them with certainty to any cluster. Sequences

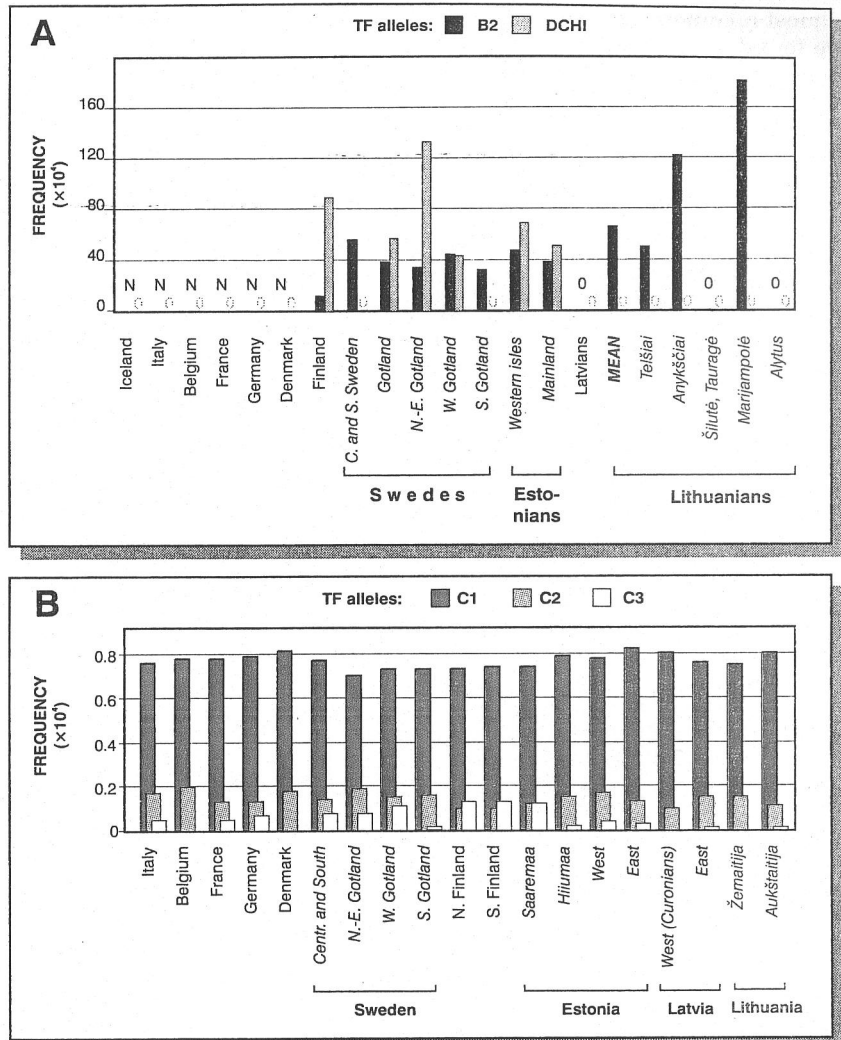


Figure 9. Frequencies (per 10,000 population) of the transferrin (TF) alleles DCHI, B2 (panel A) and C1, C2, C3 (panel B) in some European populations. N, no data

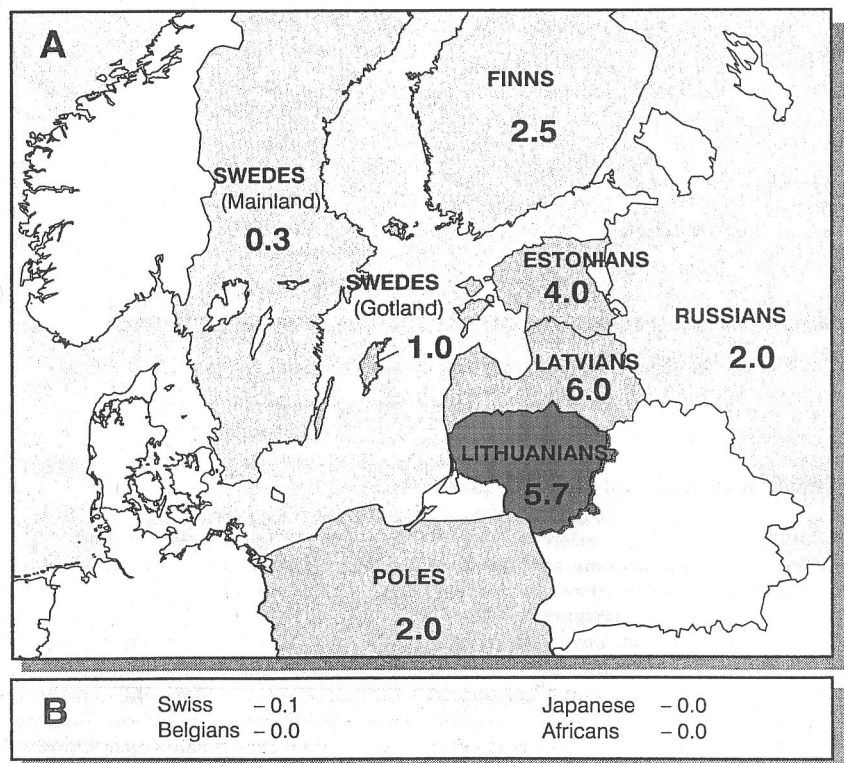


Figure 10. Frequencies (%) of the LW^b allele (Lansteiner Wiener blood group system) in the Baltic Sea region (A) and some other populations (B)

of almost all major European clusters were found in Lithuanians [22]. The frequency of the most common cluster H seems to be lower in the popula-

tion of Lithuania in comparison with other European populations, while sequences of more ancestral cluster U are more frequent [22].

Polymorphic biallelic and micro-satellite DNA markers specific to the Y chromosome have been already recognised to be highly variable in human evolutionary studies and population genetics. The distribution of haplogroups frequencies showed very small variation between Lithuanian ethnolinguistic groups [23]. The highest frequency was found for the haplogroup 3 (46%), haplogroups 16, 2 and 1 were less frequent (36%, 13% and 5%, respectively) [23]. Comparison of Lithuanian population with other European populations was made on the basis of principal component (PC) analysis. Such analysis allows to display graphically, in a few dimensions, the maximum amount of variance within a multivariate data set, with minimal loss of information. The results of PC analysis of the data on human Y-chromosome haplogroups are shown in figure 11, in which population are labelled according to linguistic affiliation. PC1–PC3 summarise 71.4% of the variance [24]. Y-chromosomal diversity within the populations surrounding the Baltic Sea shows good agreement between geographical and genetical data (figure 12) [25].

mtDNA and Y-chromosomal diversity shows that Lithuanian population is an old homogeneous group residing on its territory for a long period of time.

Variations in the prevalence rates of inherited diseases are also informative in studies of population genetics and ethnic heterogeneity. Investigation of PAH gene mutation R408W in phenylketonuria patients has shown it to be the most frequent one not only in Lithuania but also in the populations of East, Central and Southwest Europe [26]. Analysis of the possible origin of the mutation R408W on the basis of RFLP, VNTR and STR haplotypes [27, 28] has led to the conclusion that this mutation is most likely of the Baltic origin (table 4), bearing in mind the wider distribution of the Baltic tribes in Europe several thousand years ago when initial mutation arose and haplotypes were formed [28].

Another interesting case is the prevalence rate of the CFTR gene mutation Δ DF508 in the Lithuanian cystic fibrosis (CF) patients. CF is a rather common inherited disease and Δ DF508 is the most common mutation causing this disease. Investigation of the carrier status for this mutation in a population screening revealed that its frequency was 1:120 in Lithuania

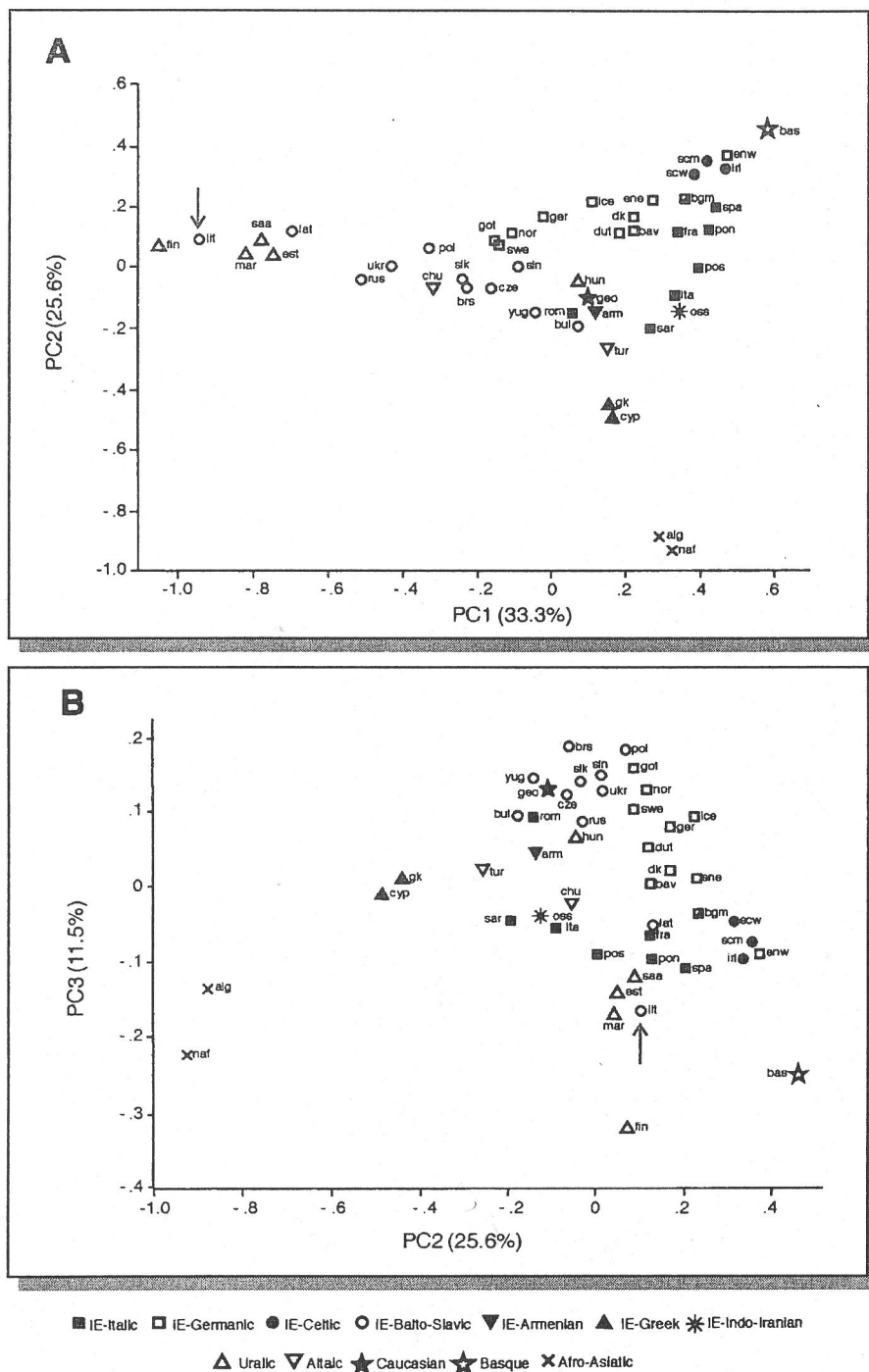


Figure 11. PC analysis of Y-chromosomal haplogroups diversity [24]. A, PC2 plotted against PC1. B, PC3 plotted against PC2

The percentage of variance explained by each component is given on the axes. Linguistic affiliation for each population is indicated symbolically; the Belgian sample is part Dutch/part French-speaking and has a hybrid symbol. Abbreviated population names: alg = Algerian; arm = Armenian; bas = Basque; bav = Bavarian; bgm = Belgian; brs = Belarussian; bul = Bulgarian; chu = Chuvash; cyp = Cypriot; cze = Czech; dk = Danish; ene = east Anglian; enw = Cornish; est = Estonian; fin = Finnish; fra = French; geo = Georgian; ger = German; gk = Greek; got = Gotlander; hun = Hungarian; ice = Icelandic; irl = Irish; ita = Italian; lat = Latvian; lit = Lithuanian; mar = Mari; naf = northern African; nor = Norwegian; oss = Ossetian; pol = Polish; pon = northern Portuguese; pos = southern Portuguese; rom = Romanian; rus = Russian; saa = Saami; sar = Sardinian; scm = Scottish; scw = western Scottish; slk = Slovakian; sln = Slovenian; spa = Spanish; swe = northern Swedish; tur = Turkish; ukr = Ukrainian; yug = Yugoslavian. Red arrows point to the data from Lithuanian population

Table 4. Association between VNTR/STR minihaplotypes and PAH mutations in chromosomes of Baltic (B), Slavic (S), Tatar (T) and unknown (U) origin in the Lithuanian population

| VNTR/STR | R408W | R158Q | R261Q | G272X | IVS10nt-11g→a | IVS12nt1g→a | Unknown | |
|----------|------------------------------------|-----------|-------|-------|---------------|-------------|-----------|-----------|
| 3/236 | 4(B) (4%) 1(U) | | | | | | 1(B) | |
| 3/240 | 65(B) (68%) 19(S) (68%) 3(U) | | | | | | | |
| 3/244 | | | | | | | 4(B) | |
| 3/248 | 1(B) | 3(B) (3%) | | | | | 3(B) | |
| 3/252 | | 1(B) | | | | | | B: 96% |
| 7/236 | | 1(S) 1(B) | | | | | | |
| 7/240 | 1(U) | | | | | | 1(B) | |
| 7/244 | | | 1(B) | | | | 1(T) | |
| 7/248 | | | | | | 1(B) | 1(B) | |
| 8/228 | | | 1(S) | 1(B) | | | 1(B) | |
| 8/232 | | | | 1(B) | | | | |
| 8/236 | | | | 1(B) | | | | |
| 8/240 | | | | | 1(B) | | 1(S) | B: 4% |
| 8/244 | | | | | | | 1(B) 1(S) | |
| 8/248 | | | | | | | 1(B) | |
| 9/236 | | | | | | | 1(S) | S: 25% |
| 9/248 | | 1(B) | | | | | 3(S) | |
| 9/252 | | | | | | | 1(S) | |

[29, 30] while it is much higher in other European populations, e.g. 1:38 in Denmark [31]. The relative frequency of this mutation in CF patients is 39% in Lithuania, 45% in Sweden and Finland, 55% in Poland, 71% in Germany, 87% in Denmark [32], thus the lowest was found in Lithuania in comparison to other European populations investigated.

CONCLUDING REMARKS

In summary it can be pointed out that the haplotype and allele frequencies are not randomly distributed among the Lithuanian ethnolinguistic subgroups showing more or less clear patterns of distribution, which are associated with the differentiation of these subgroups in present-day Lithuania. A quite distinct pattern was found for the South Aukštaičiai and North Žemaičiai, which can be explained by their historical background together with a marked and long standing genetic isolation from other populations.

The Žemaičiai ethnolinguistic group is more homogeneous according to the genetic markers data both within the North, West and South subgroups and between these. Among the Žemaitish and Aukštaitish parts of Lithuania there is a significant heterogeneity. The distribution of

groups of the Lithuanians. In both groups type 1 was the most common one. In the Aukštaičiai there were 10 additional types in altogether 23 individuals. In the Žemaičiai group none of these types was found. On the other hand, there were three mt-DNA types in the latter group not found in the former one.

According to the data on rare genes causing inherited diseases, mutations that are the most frequent in the Lithuanian population often appear to be also the most frequent in Europe (e. g. PAH gene mutation R408W), the least frequent in Europe (e. g. CFTR gene mutation ΔDF508) or their frequency values take an intermediate position in the established gene frequency gradient [7].

ACKNOWLEDGEMENTS

The author is grateful to Prof. Lars Beckman (University of Umeå, Umeå, Sweden) for most helpful discussion, to Prof. M. Stoneking (Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany) for valuable suggestions in the investigation of Alu insertions and to Dr. D. Steponavičiūtė (Human Genetics Centre of Vilnius University, Vilnius, Lithuania) for assistance in preparing the manuscript. ♦

Gauta:
Priimta spaudai:

2001 05 11
2001 08 06

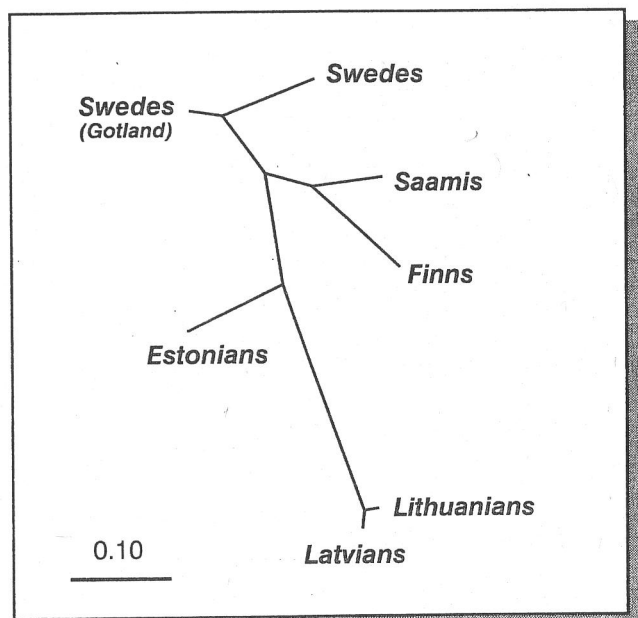


Figure 12. Phylogenetic tree according to Y-chromosome haplogroup 16 of north, east and south populations of the Baltic Sea region

mt-DNA RFLP types differed between these two main

Santrauka

LIETUVIŲ GENAI IR GENŲ GEOGRAFIJA

Vaidutis Kučinskas

Pagrindinis šio straipsnio tikslas – apžvelgti šiuolaikines žinias apie lietuvių populiacijos genetinius ypatumus. Lietuvių populiacijos genetinė diferenciacija ir lietuvių bei kitų Europos populiacijų ryšiai buvo analizuojami kraujo grupių, serumo baltymų polimorfizmo ir DNR žyme-

nų (tarp jų mtDNR ir Y chromosomos žymenų) pagrindu. Tyrimų rezultatai atskleidė nedidelius dabartinių Lietuvos etnolingvistinių grupių skirtumus, kurie tikriausiai kyla iš priešistorinės baltų genčių struktūros. Baltų tautose susimaišę rytų ir vakarų požymiai, pavyzdžiui, didelis B kraujo grupės dažnis kartu su labai dideliu Rh(-) kraujo grupės dažniu. Baltiškojo „gentinio“ geno LW^b tyrimų duomenys rodo didelę baltų įtaką kaimyninėms finų-ugrų ir slavų populiacijoms.

Prasminiai žodžiai: lietuviai, genų geografijs, kraujo grupės, baltymų polimorfizmas, DNR žymenys.

REFERENCES

1. Lašas V. Our and our neighbours' blood groups // *Medicina*. 1929, no. 10, p. 541–547 (in Lithuanian).
2. Natkevičaitė M. Human blood groups and their distribution among the inhabitants of Lithuania // *Kosmos*. 1929, no. 8, p. 237–289 (in Lithuanian).
3. Žilinskas J., Masalskis R. Blood group subordination (parents–children blood peculiarities) and relations between adult and infant blood groups // *VDU MF*. 1936, no 3, p. 186–229 (in Lithuanian).
4. Kučinskas V. The genetic process in present-day Lithuanian population // *Science, Arts and Lithuania*. 1990, no. 1, p. 48–53.
5. Zinkevičius Z. The History of the Lithuanian Language. 1st edition. Vilnius: Mokslo ir enciklopedijų leidykla, 1998, p. 333.
6. Česnys G. Anthropological roots of Lithuanians // *Science, Arts and Lithuania*. 1991, no. 1, p. 4–10.
7. Libert F., Cochaux P., Beckman G., et al. The Δ ccr5 mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe // *Hum. Mol. Genet.* 1998, vol. 7, p. 399–406.
8. Beckman L., Sikström C., Mikelsaar A., et al. Alpha1-Antitrypsin (PI) alleles as markers of West-European influence in the Baltic Sea region // *Hum. Hered.* 1999, vol. 49, p. 52–55.
9. Beckman L., Sikström C., Mikelsaar A.-V., et al. Transferrin variants as markers of migrations and admixture between populations in the Baltic Sea region // *Hum. Hered.* 1998, vol. 48, p. 185–191.
10. Van Landeghem G. F., Beckman L. E., Sikström C., et al. New DNA polymorphisms define ethnically distinct haplotypes in the human transferrin receptor gene // *Hum. Hered.* 1998, vol. 48, p. 245–250.
11. Kučinskas V., Sistonen P., Beckman L. The LW (Landsteiner Wiener) blood group system phenotypes and LW^b allele frequencies in six Lithuanian ethnolinguistic subgroups // *Acta medica Lituanica*. 1998, no. 4, p. 281–283.
12. Kučinskas V., Ambrasienė D., Sistonen P. A1A2B0 (AB0) and C, c, D, E, e (Rh) blood groups polymorphisms and geography in six ethnolinguistic regions of Lithuania // *Biologija* (Vilnius). 1998, no. 4, p. 3–6.
13. Kučinskas V., Kasnauskienė J. Genetic variants of recent Alu insertions // *Biologija* (Vilnius). 1998, no. 4, p. 7–9.
14. Kučinskas V., Payne A. M., Ambrasienė D., et al. Mutations in the human rhodopsin gene and polymorphisms in peripherin/RDS gene in Lithuanian autosomal dominant retinitis pigmentosa patients // *J. Appl. Genet.* 1999, vol. 40, p. 53–61.
15. Sistonen P., Virtaranta-Knowles K., Denisova R., et al. The LW^b blood group as a marker of prehistoric Baltic migrations and admixture // *Hum. Hered.* 1999, vol. 49, p. 154–158.
16. Kučinskas V., Radikas J., Rasmuson M. Genetic diversity in the Lithuanian rural population as illustrated by variation in the AB0 and Rh(D) blood groups // *Hum. Hered.* 1994, vol. 44, p. 344–349.
17. Cavalli-Sforza L.L., Menozzi P., Piazza A. The History and Geography of Human Genes. 1st edition. Princeton, New Jersey: Princeton University Press, 1994. 518 p.
18. Sistonen P. The LW (Landsteiner Wiener) blood group system. Elucidation of the genetics of the LW blood group based on the finding of a 'new' blood group system. PhD thesis, University of Helsinki, 1984.
19. Kučinskas V., Kasnauskienė J., Cimbališienė L. Fenilketonurijos molekuliniai tyrimai Lietuvoje: 92% mutacijų, nustatytų denatūruojančio gradientinio gelio elektroforezės metodu // *Medicina*. 2001, Nr. 3, p.
20. Kučinskas V. Human mitochondrial DNA variation in Lithuania // *Anthropologischer Anzeiger*. 1994, vol. 52, p. 289–295.
21. Richards M., Corte-Real H., Forster P., et al. Paleolithic and neolithic lineages in the European mitochondrial gene pool // *Am. J. Hum. Genet.* 1996, vol. 59, p. 185–203.
22. Kasperavičiūtė D., Kučinskas V., this issue.
23. Ambrasienė D., Kučinskas V. Lietuvos vyrų Y chromosomos genetinės įvairovės analizė pagal bialelinius genetinius žymenis // *Laboratorinė medicina*. 2001, Nr. 4 (in press).
24. Rosser Z. H., Zerjal T., Hurler M. E., et al. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language // *Am. J. Hum. Genet.* 2000, vol. 67, no. 6, p. 1526–1543.
25. Zerjal T., Beckman L., Beckman G., et al. Geographical, linguistic, and cultural influences on genetic diversity: Y-chromosomal distribution in northern European populations // *Mol. Biol. Evol.* 2001, vol. 18, No. 6, p. 1077–1087.
26. Kučinskas V., Jurgelevičius V., Cimbališienė L., Holmgren, G. Distributions of phenylalanine hydroxylase mutations and haplotypes in Lithuanian phenylketonuria patients // *Hum. Hered.* 1994, vol. 44, p. 110–113.
27. Kalaydjieva L., Dworniczak B., Kučinskas V., et al. Geographical distribution gradients of the major PKU mutations and the linked haplotypes // *Hum. Genet.* 1991, vol. 86, p. 411–413.
28. Giannattasio S., Jurgelevičius V., Lattanzio P., et al. Phenylketonuria mutations and the linked haplotypes in the Lithuanian population: origin of the most common R408W mutation // *Hum. Hered.* 1997, vol. 47, p. 155–160.
29. Wennberg C., Kučinskas V. Low frequency of the Δ F508 mutation in Finno-Ugrian and Baltic populations // *Hum. Hered.* 1994, vol. 44, p. 169–171.
30. Ambrasienė D., Kučinskas V. Heterozygote screening for major cystic fibrosis mutation Δ F508 in the Lithuanian population // *Acta medica Lituanica*. 1997, Suppl. 1, p. 3–5.
31. Schwartz M., Brandt N.J., Skovby F. Screening for carriers of cystic fibrosis among pregnant women: a pilot study // *Eur. J. Hum. Genet.* 1993, vol. 1, p. 239–244.
32. The Cystic Fibrosis Genetic Analysis Consortium. Population variation of common cystic fibrosis mutations // *Hum. Mutat.* 1994, vol. 4, p. 167–177.