

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Minisatellite DNA Markers in Population Studies

Svetlana Limborska, Andrey Khrunin and Dmitry Verbenko  
*Institute of Molecular Genetics, Russian Academy of Sciences, Moscow  
Russia*

## 1. Introduction

The discovery of an anonymous multiallelic locus in 1980 demonstrated for the first time that the human DNA contains hypervariable regions (Wyman & White, 1980). Eight alleles of a polymorphic locus that was not associated with any known gene were identified during the blot hybridization of total human genomic DNA treated with the restriction enzyme EcoRI with a human DNA fragment (16 bp in length) isolated from a phage genomic library. The multiallelic nature of the polymorphism at this locus did not stem from a variation in restriction sites; rather, it originated from a variable number of tandem repeats in the short core DNA sequence. Later, other similar polymorphic sites were detected near the 5' end of the insulin gene (Bell et al., 1982), in the Harvey *ras* oncogene (Capon et al., 1983), in the  $\zeta$ -globin pseudogene (Proudfoot et al., 1982), and within the  $\beta$ -globin cluster (Weller et al., 1984). In 1985, Jeffreys et al. published the results of their research, in which they described a fourfold repeat of 33 nucleotides in one of the introns of the human myoglobin gene (Jeffreys et al., 1985). These polymorphic regions consisted of tandem repeats of a short sequence (~11–60 bp) and were termed variable number tandem repeats (VNTRs) (Kendrew & Lawrence, 1994).

The other term for VNTR loci, minisatellites, was attributed based on the similarity of some of their properties with those of highly repetitive satellite DNA sequences. Tandem satellite DNA repeats are combined into a huge and structurally diverse group arranged into continuous clusters with monomeric units positioned in a head-to-tail configuration. The differences between minisatellite loci and satellite DNA are the greater variability in the length of the repeating unit of the latter (varying from 10 to 1000 bp) and chromosomal localization: satellite DNA is located in the regions of near-centromere heterochromatin in metaphase chromosomes and in the chromocenters of interphase nuclei, whereas minisatellite sequences are distributed evenly over most regions of all chromosomes (Miklos & John, 1979).

The classification of Jeffreys et al. discriminates between microsatellites (with an elementary link of 2–6 bp), minisatellites (up to 100 bp), midisatellites (100–400 bp), and macrosatellites (up to several thousand bp) (Jeffreys et al., 1994). Later, a great emphasis was placed on loci that were generally classified as microsatellites with elementary links of less than 10 bp (short tandem repeats, STRs) and minisatellites with links of more than 10 bp in size (VNTRs) (Schlotterer C., 1998; Gemayel et al., 2010). These two types of variable DNA were used as genetic markers in genomic and population studies.

Many minisatellites are characterized by hypervariability and a high degree of polymorphism. Their heterozygosity is 85–99%, whereas the maximum heterozygosity of biallelic loci is 50% (Bois, 2003; Gemayel et al., 2010). Minisatellite loci have been used actively in criminalistics, in the panels of highly informative markers used for the identification of individuals and ascertainment of filiation (Jurka & Gentles, 2006; Hong-Sheng et al., 2009). Data on the population frequencies of hypervariable markers were important for the determination of data reliability in such investigations (Zhivotovskii, 1996).

Minisatellite markers have also been used intensively in genetic studies, including the study of genomic diversity in human populations. However, during the last few years, because of the great technological efforts that are necessary for the typing of minisatellite markers, the main focus of population genetics research has been the investigation of short tandem repeats (STRs) and single nucleotide polymorphisms (SNPs) (Kelkar et al., 2008). Difficulties in the classification and determination of minisatellite allele variants halted the progress of the use of these DNA markers in population research. However, recently, the peculiar properties of minisatellites (i.e., new data regarding their functional significance) have rendered this class of DNA markers actual again (Babushkina & Kucher, 2011, Gemayel et al., 2010). Taking into account the high mutation rate of hypervariable minisatellites, their polymorphic nature allows not only the determination of population divergence over long periods, but also the detection of the specificity of the relatively modern (hundreds of years ago) ethnic history of populations.

## 2. Minisatellite loci in human DNA

The high degree of polymorphism of minisatellite sequences is evidence of the high rate of their evolution; nevertheless, most of them are rather stable. Hypermutability was demonstrated for only a few minisatellite loci (Bois, 2003). These loci are very suitable models for studying the mechanisms that lead to the variability of minisatellite sequences. Minisatellites are classified as hypermutable only when their mean mutation frequency in germ line cells is greater than 0.5%; this rate may be equal (e.g., MS1 minisatellites) or vary (e.g., CEB1 minisatellites) between male and female germ cells. According to rough estimates, less than 10 out of about 300 minisatellites typed in families were hypermutable (Amarger et al., 1998). No structural peculiarities distinguishing hypermutable minisatellites from other tandem repeats were found; the ratio of repeats with near-telomere localization to those distributed evenly over the genome was the same for minisatellites of the human genome as a whole (Vergnaud & Denoeud, 2000).

Several models have been accepted for the interpretation of the mechanisms underlying the mutational processes in minisatellite loci that result in the duplication of a repeating unit. Levinson and Gutman suggested that tandem repeat duplication occurs at random, but can be repeated many times after such an occurrence because of nonspecific mating of DNA chains resulting from replisome slippage (Levinson & Gutman, 1987). At present, this model is considered as acceptable to explain the appearance and increase in the length of microsatellites with an elementary repeat unit smaller than 10 bp (Kasai et al., 1990; Gemayel et al., 2010). The model developed by Buard and Vergnaud envisages the effect of *cis* activators on the stability of minisatellites (Buard & Vergnaud, 1994). According to this model, initiation of the recombination hotspot located outside the minisatellite structure

may stimulate a perfect mosaic of intra- and interallele events. As a whole, the process appears to be an exchange that is analogous to gene conversion, with the involvement of sister chromatids. The increase of the length of minisatellites is usually polar: i.e., it consists of the addition of a new region of tandem repeats to the 3'-end region of the minisatellite. The conservativeness of the flanking regions of the minisatellite cluster is indirect evidence of the correctness of this model, whereas highly intensive reorganizations, such as intra- and interallele exchange of repeating units, occur inside the cluster (Harris, 2002).

The processes that determine the instability of minisatellites are different in somatic cells and germ line cells (Bois, 2003). The events of crossing over and gene conversion occur in the germ line, supposedly during meiosis. The somatic instability of minisatellites is determined mainly by intra-allele duplications. Figure 1 shows the events that lead to the instability of the minisatellite loci of human DNA are as follows. Moreover, the appearance of the double-stranded breaks of DNA outside the minisatellite region can lead to their migration to the minisatellite region (for a detailed review see Bois, 2003). Some models have been proposed to explain the mechanisms of this migration. The transfer of a double-stranded break results in the accumulation of mutations within the minisatellite cluster, but not in other regions of genomic DNA, which explains the hypervariability of minisatellites (Bois & Jeffreys, 1999, Bois, 2003).

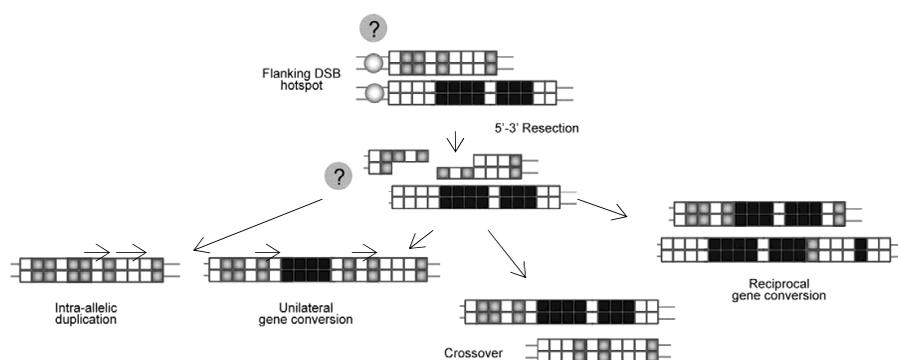


Fig. 1. Simplified model of human minisatellite instability (according to Bois, 2003). The allele destined to be mutated is shown with dark shadows, and the recipient allele is in black. An initial double strand break (DSB) outside or within minisatellite repeat can, after 5'-3' resection, generate simple intra-allelic duplication (no strand invasion), unilateral conversion, crossing over, or reciprocal gene conversion (strand invasion). The two steps not yet fully understood is indicated with a circled question mark. The mechanism leading to the generation of the initial DSB remain elusive. The identification of intermediates of conversion or recombination need to be characterized to dissect the various pathways. Reprinted from *Genomics* Vol.81, No.4, Bois, P.R. Hypermutable minisatellites, a human affair?, pp. 349-355, issn 0888-7543, Copyright 2003, with permission from Elsevier (<http://www.sciencedirect.com/science/article/pii/S0888754303000211>)

The availability of complete genome sequences and increased knowledge of genome biology indicate that minisatellites may occur within coding and regulatory regions, and that some minisatellites are involved in the processes of genome regulation (Gemayel et al., 2010). These sequences not only have specific biological functions, but, via their intrinsic instability, may also lead to faster rates of evolution of genes and their associated

phenotypes. Variable minisatellites, as is the case for the *FLO1* gene of the benign brewer's yeast *S. cerevisiae*, lead to gradual, quantitative functional changes that may allow the rapid adaptation of the organism to changes in the environment (Verstrepen et al., 2005). Correlative observations between minisatellite allele size and gene expression patterns have been made in higher eukaryotes, including humans. One example is the minisatellite located upstream of the promoter of the human insulin gene *INS*. Shorter minisatellite alleles are linked to altered expression of *INS*, both *in vitro* and *in vivo*, and to predisposition to insulin-dependent diabetes mellitus (Bell et al., 1992; Bennett et al., 1995). Another study showed that this minisatellite affects the expression of the nearby *IGF2* gene, which encodes the insulin-like growth factor II (Paquette et al., 1999). The locus encompassing the *INS* and *IGF2* genes is termed *IDDM2*, and the minisatellite is located between the genes. The exact mechanism underlying this upregulation was not determined in the study; however, it was suggested that the tandem repeated structure of minisatellite loci might potentiate Z-DNA formation, which alters gene expression (Gemayel et al., 2010).

Several studies of human genes showed that minisatellites might interact with transcriptional factors. For example, the minisatellite located in intron 2 of the serotonin transporter gene *5-HTT* may influence its transcription levels by binding the transcription factor YB-1 (Klenova et al., 2004). The location of minisatellites involved in the regulation of gene expression and function need not be limited to promoter sequences. These loci have been found in other expression regulatory sequences, including the 5' and 3'UTRs of transcripts and introns. Many of these loci have a function in regulating gene expression, and variation in repeat units often affects their activity (Kawakami et al., 2001; Fuke et al., 2001).

Another mechanism of transcriptional regulation involves the possible interaction of miRNA with minisatellites. The unusual properties of the 27 bp minisatellite situated at exon 4 of endothelial NO synthase have been demonstrated *in vitro* using endothelial cell culture (Song et al., 2003). *eNOS* gene expression and eNOS protein concentration and enzyme activity correlate with the allele size of this VNTR. Moreover, the transcript of this VNTR is a short intronic repeat small RNA (sirRNA) that inhibits *eNOS* expression during transcription via a negative regulation mechanism (Zhang et al., 2005).

### 3. Hypervariable minisatellite DNA markers in human population genetics

In 1985, Jeffreys et al. published the results of their research, in which they described a minisatellite—a fourfold repeat of 33 nucleotides—in one of the introns of the human myoglobin gene (Jeffreys et al., 1985). Using this minisatellite as a probe, the authors isolated and characterized several hypervariable sequences (cores) from a human genomic library; all of these also proved to be minisatellites with 3–29 repeating units. These cores, despite their similarities to each other, exhibited some differences in their nucleotide composition and varied in length from 16 to 64 bp. They were similar in that all repeating links comprised an almost identical 15-nucleotide sequence, referred to by the authors as the core sequence, which may be described as a consensus sequence bearing certain similarities to the  $\chi$  sequence of  $\lambda$  phage DNA (GCTGTGG). The ability of probes based on the core sequence to identify many such minisatellite loci during blot hybridization with total human DNA demonstrated the multiplicity of their localization in the genome. This property, together with the high level of polymorphism determined by the variability of the number

of repeats, underscored the method of multilocus DNA fingerprinting (Jeffreys et al., 1994). Using different "policore" samples for blot hybridization, Jefferies et al. were the first to show that the hybridization pattern detected comprises many loci containing a family of minisatellites (Jeffreys et al., 1988). Blot hybridization of such sites using restricted genomic DNA revealed a picture of multiple hybridization, i.e., the presence of these loci in the genome set. The level of polymorphism of the blot hybridization patterns was extremely high, because it was determined by a combination of a large number of independent hypervariable genome loci.

Later, two independent groups of researchers, led by Ryskov in the USSR and Vassart in Belgium, discovered another family of hypervariable regions with multiple localizations in the genome, which was detected using M13 phage DNA (Ryskov et al., 1988; Dzhincharadze et al., 1987; Vassart et al., 1987). The use of two small regions that are typical of minisatellite sequences within the M13 phage DNA as a natural probe for blot hybridization with restricted DNA yielded multiple patterns of hybridization and a high level of interindividual polymorphism, which was comparable to that observed for the minisatellites described by Jeffreys et al. M13 minisatellites exhibited universal distribution in animate nature; they were found in microorganisms, plants, animals, and humans (Ryskov et al., 1988; 1990). The genomic fingerprinting technique, which analyzes many genomic polymorphic systems in the same experiment, has been used widely in forensic medicine and parentage testing (Semenova et al., 1996; Shabrova et al., 2006).

The DNA multilocus patterns determined using M13 phage hybridization were used for the first time in human population studies by Barysheva et al. (Barysheva et al., 1989, 1991a, 1991b; Semina et al., 1993). These authors determined the characteristics of M13 DNA fingerprint patterns, performed segregation analyses, and estimated the mutation frequency in M13 human minisatellite loci. The result of the cluster analysis of a genetic distance matrix confirmed data on the relationships in the group of local populations under consideration. Kalnin et al. were the first to apply multiple correspondence analysis (MCA) to the analysis of DNA fingerprinting data in human population studies (Kalnin et al., 1995). These studies demonstrated the great potential of this approach for the analysis of complex DNA multilocus blot hybridization patterns, which yield adequate results despite the inevitable errors of fragment identification that arise from the analysis of multiple autoradiographs (Shabrova et al., 2006).

However, the difficulty in analyzing DNA multilocus blot hybridization patterns, combined with the need to use the cumbersome method of blot hybridization, limited the widespread use of this approach in population genetics. Minisatellites detected later, which had unique localizations in the genome, were used widely in subsequent studies, as their detection was based on the simpler and more accessible PCR method. It eventuated that a set of singly localized (monolocus) minisatellites yielded identification patterns with an information content similar to that of the minisatellites of Jeffreys et al. or of the M13 multilocus probes. This fact favors the significant development of studies of the population characteristics of these monolocus minisatellites. The minisatellites used most frequently in population studies are those included in forensic identification panels, the most popular of which are the 3'APOB and D1S80 minisatellites; the properties of these minisatellites will be addressed in detail.

#### 4. The 3'APOB minisatellite polymorphism in human population research

One of the VNTR loci maps to chromosome 2 and is located 75 bp from the second polyadenylation signal at the 3' end of the *APOB* gene (Huang & Breslow, 1987). The APOB protein is one of the major low-density lipoproteins and plays a central role in the metabolism of serum cholesterol. The 3'APOB hypervariable region consists of a tandem-repeat sequence that is rich in A and T. Two basic types of 15-nucleotide-long core repeats have been identified (Buresi et al., 1996). The tandem repeating unit of the 3'APOB minisatellite consists of two tandem sequences of 14 and 16 bp; thus, neighboring allelic variants differ from each other by 30 bp, or by 2 repeats. Allelic variants differ in the number of repeats and contain 25 to 55 repeat units (Ludwig et al., 1989). The literature describes several systems of alleles for the 3'APOB locus. Since the establishment of one such system by Ludwig, allelic variants are denoted as 30, 32, 34, etc., according to the number of repeated units. According to the system of Boerwinkle (Boerwinkle et al., 1989), which counts one structure-segment sequence before the minisatellite cluster, the same allelic variants are designated as 31, 33, 35, etc., respectively. Thus, the allelic variant with number 36 of Ludwig's system corresponds to variant 37 of Boerwinkle's system. In addition to the two types of core segment structure, the 3'APOB alleles of a number of core segments have sequence microvariations (a pure AT sequence is interrupted by a C or a G), usually concentrated at the 3' end of the minisatellite (Chen et al., 1999; Marz et al., 1993; Buresi et al., 1996). In the human population samples analyzed by Buresi et al., a haplotype analysis of such substitutions revealed the presence of five allelic sequences. The authors found the ancestral state of the 3'APOB minisatellite allelic sequence via comparison with another allele sequence variant discovered in primates and showed that different types of allele sequences in humans appeared during 3'APOB minisatellite evolution due to three possible conversions at the minisatellite locus (Buresi et al., 1996).

The 3'APOB polymorphism has been used widely in investigations of the history and diversity of humans, both worldwide and in individual population groups (Buresi et al., 1996; Destro-Bisol et al., 2000; Renges et al., 2002; Kravchenko et al., 1996; Poltl et al., 1996; Zago et al., 1996; Verbenko et al., 2005). It has been considered as a suitable locus for a pilot study of the relationships between the shape of allele-size distributions of minisatellites and the microevolutionary processes leading to their present-day distribution (Destro-Bisol et al., 2000). The allele-size frequencies can be used to calculate interpopulation genetic distances. Higher differences in the level of polymorphism have been found in populations with different origins and ethnicities (Buresi et al., 1996; Destro-Bisol et al., 2000; Renges et al., 2002; Kravchenko et al., 1996; Poltl et al., 1996).

The use of the 3'APOB minisatellite as a marker in the study of evolutionary models was launched in 1992 (Renges et al., 2002). Subsequently, scientists from many countries studied the 3'APOB minisatellite polymorphism in a large number of human populations. Although the allele frequencies vary considerably among different populations, the similarities in their distribution shapes should be noted. Generally, two allelic variants (34 and 36 repeats) are detected most often (their total frequency is 57–77%), one or two allelic variants are less frequent (usually a variant containing 32, 46, or 48 repetitions, with frequencies up to 12%), whereas other alleles occur with a frequency of less than 5% (Boerwinkle et al., 1989a; Ludwig et al., 1989; Deka et al., 1992; Friedl et al., 1990; Renges et al., 1992; Lahermo et al.,

1996; Spitsyn et al., 2000; Destro-Bisol et al., 2000; Khusnutdinova et al., 1999; Khusnutdinova and et al. 2003; Akhmetov et al, 2006; Bermisheva et al., 2007). As an example, we considered the population characteristics of the 3'APOB polymorphism in Eastern European populations.

Eastern Europe is inhabited by a great number of ethnic groups that differ significantly in their characteristics (Kuzeev, 1985; Bunak, 1965). East Slavs are the main population group in Eastern Europe. The formation of East Slavic peoples (Russians, Ukrainians, and Belarusians) is supposed to have occurred because of the long-term migration and expansion of ancestral Slav tribes from Central Europe to the territory of the Russian Plain, which was settled by pre-Finno-Ugric tribes, since the Late Paleolithic (Sedov, 1979; Alekseeva, 1973). Ethnic groups of the southern region and surrounding Ural Mountains, which neighbor the East Slavic peoples, have an even more luxuriant history resulting in the formation of the Turkic language groups of Tatars and Bashkirs, the peoples of the North Caucasus, and Mongoloid populations such as the Kalmyks.

The Eastern Slavonic linguistic group (Indo-European linguistic family) was represented by samples from Russian populations from the European (northwestern) part of Russia (Oschevensk, Belaia Sluda, Kholmogory, Mezen, Kursk, Novgorod, Cossacks, Sychevka, Kostroma, and Smolensk), and six Byelorussian populations (Grodno, Pinsk, Mjadel, Bobruisk, Nesvij, and Khoyniki) from different regions of the Republic of Belarus (for a detailed description of the Byelorussians, see Popova *et al.* (Popova et al., 2001). The Belaia Sluda group is an isolated Russian population living at the border of the Arkhangelsk region of northern Russia with the Republic of Komi. The Oschevensk group is an isolated Russian population living in the Kargopol district of the Arkhangelsk region. From ethnohistorical and anthropological points of view, these Russian groups might carry an admixture of ancient Vepsian (Ageeva, 2000) or Saami lineages (Sedov, 1979; Alekseeva, 1973). The Kholmogory are based in a town near the city of Arkhangelsk, representing Russian north-coast dwellers, and the Mezen group is from the same lineage, which is derived from Russians who migrated from Novgorod to the northeast, starting in the 16th century. The Novgorod group is from the northwestern European part of Russia. The Kursk group is a southwestern Russian population. The Cossacks are a southern Russian population from the Krasnodar region (settled at the Kuban River). The Smolensk group is from the town of Ugra, with a complex history of population movements (southwestern part of Eastern Europe), and the Sychevka group is also from the Smolensk district of Russia (from the central part of the Russian Plain, which borders the Tver district).

Populations from the western Ural region were represented by Finno-Ugric speakers—the Komi-Permyats from the Perm district of Russia, the Komi-Ziryans (Izhemski and Priluszki Komi subpopulations), the Udmurts, and the Meadow Maris—and by Turkic speakers (Altaic linguistic family): the Bashkirs, from the Beloretsky region of the Republic of Bashkiria, and the Tatars, from the town of Almetyevsk (for a detailed description of these groups see Bermisheva *et al.*, 2003). The Komi (Komi-Zyryans) are one of the most numerous peoples of the Finno-Ugric group: they occupy the northeasternmost location among European ethnic groups, which adjacent to the Nentsy. They inhabit the territory of the basins and the tributaries of the Vichegda, Mezen and Pechora rivers. The contemporary Komi people consist of some distinct ethnographic groups, which formed during the 8th to the 19th centuries. Two geographically different ethnographic groups were studied in this

work. One of them, the Izhemski Komi, takes a particular place among the Komi groups. They stand out because of a number of peculiarities of their language and traditional economy. The latter has long been based on the commodity of reindeer breeding. Moreover, the Izhemski Komi exhibit some anthropological traits that differentiate them from other Komi groups. Unlike the Izhemski Komi, the Priluzski Komi have traditionally occupied themselves with farming and cattle breeding. In addition, the Priluzski Komi belong to ethnic groups that, historically, formed before the Izhemski Komi people. Two ethnic groups originate from the Altaic linguistic family, but do not inhabit the Ural region: these are the Kalmyk and Yakut populations. Kalmyks inhabit the steppe region located to the northwest of the Caspian Sea. This ethnic group settled in their current region of inhabitation during a migration from the Dzungaria region of Central Asia in the 16th century (northwestern China). The Yakut population lives in East Siberia and belongs to the Turkic linguistic group (Altaic linguistic family). In classical anthropology, they are classified as the Central Asian type (Cavalli-Sforza et al., 1994). In this study, samples from the Elista region of Kalmykia and from a central group of Yakut people were examined.

The blood samples used in this study were obtained by venipuncture into EDTA-coated Vacutainer tubes after obtaining informed consent from each individual. To fulfill the selection criteria, all individuals had to belong to the native ethnic group of the region studied (descended from at least three generations living in the region), be unrelated to each other, and be healthy. DNA isolation and purification, PCR analysis, gel electrophoresis, and multidimensional statistical analyses were performed as described in Verbenko et al (Verbenko et al., 2003a, 2006). Calculations of population characteristics, pairwise genetic distances, and molecular variances were performed using POPGENE version 1.32 (Yeh et al., 1999) and GDA (Weir, 1996, Lewis & Zaykin, 2001) software. Fisher's exact testing of contingency tables was performed using RxC software (Miller, 1997). The phylogenetic interrelation in the populations of Eastern Europe was studied based on data on the variability of these minisatellite loci, and genetic distances were calculated according to Nei (Nei, 1972). The matrix obtained was used for analysis using the method of multidimensional scaling, which visualizes interrelations between the populations and facilitates the significant interpretation of results, especially regarding the multimodal distribution of the frequencies of allele variants.

The data obtained revealed the presence of 31 allelic variants of the 3'*APOB* minisatellite, ranging in size from 24 to 54 repetitions with varying frequency in the population. Nine (in a population of Meadow Mari) to 17 (in a population of Bobruisk Belarusians) different alleles were found in different populations. A wide range of variability in 3'*APOB* minisatellite alleles indicates a high level of polymorphism in the populations under study. In the Eastern European populations studied here, the most frequent allelic variants of the 3'*APOB* minisatellite had 34 and 36 repeats. The allele with 36 repeats is dominant in European populations of Eastern Slavs (Russian, Belarusian, and Ukrainian), whereas the allele with 34 repeats is most frequent in the Asian populations (Kalmyk and Yakut). This is consistent with the results of the majority of the works addressing the variability of this marker in populations worldwide. Major allelic variants represent the largest contribution to the peculiarity of allele frequency differences in the populations; however, a wide range of information stems from minor allelic variants containing 30, 32, 38, 40, or 42 repeats, which contribute to the identity of particular groups and specific populations. Modality is a peculiarity of the distribution of allele frequencies at the 3'*APOB* minisatellite locus, which

can be treated as the main human-group diagnostic feature. The histogram (Figure 2) displays the allele frequency distributions of the 3' *APOB* minisatellite in populations of Russians, Yakuts, and Africans (Cameroon) (Destro-Bisol et al., 2000). The pronounced differences in the distribution of the major alleles containing 34 and 36 repeats can be observed. In contrast to those observed for European and Asian populations, the profile of the allele frequency distribution in populations of sub-Saharan African origin is unimodal (Deka et al., 1992; Renges et al., 1992; Destro-Bisol et al., 2000).

The allele spectrum of the 3' *APOB* minisatellite in Eastern European populations (Russians, Belarusians, Ukrainians, Adygeis, Circassians, and Abkhazians) is bimodal, with peaks at alleles 34–36 and 48. In contrast, the Asian populations of Kalmyk and Yakut and the population of the Volga-Ural region do not exhibit the second peak; the frequency distribution of the allelic variants observed in these populations is unimodal.

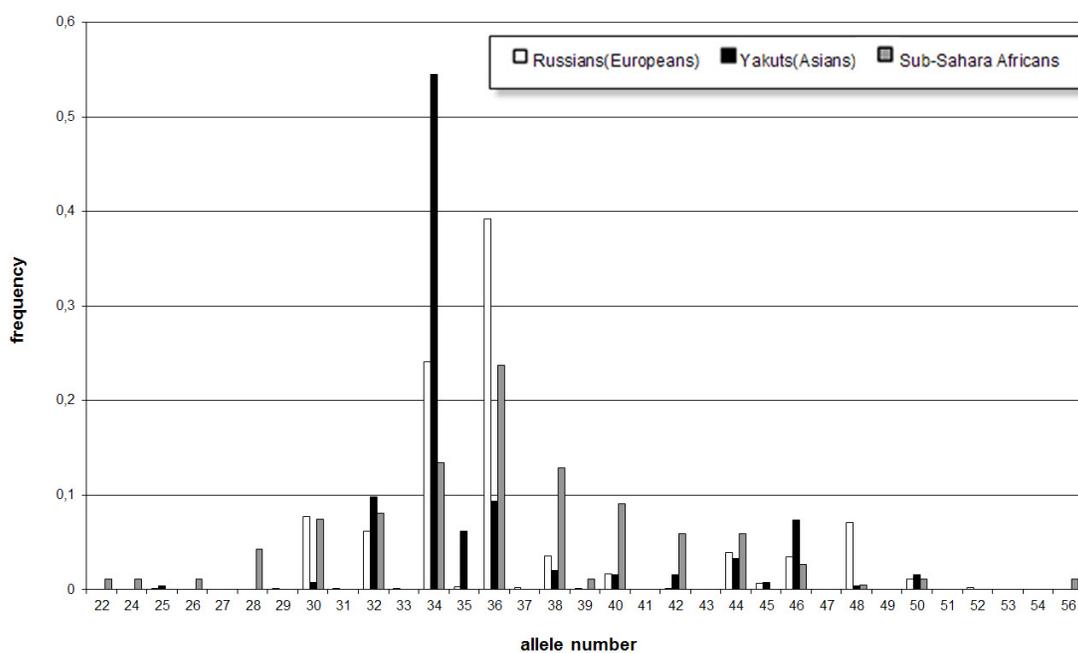


Fig. 2. 3'-end *APOB* minisatellite allele frequency distributions in populations of three main human groups.

A comparative analysis of the data obtained with European and Asian population data from literature revealed the similarity between the Eastern Slavs and the European populations of western and central Europe, and the Yakut and Kalmyk – with Asian populations (Chinese and Japanese) (Verbenko et al., 2003a). The analysis of the data using multidimensional scaling showed two clusters: Asian and European, which has a compact core. At the heart of the European population cluster were Germans, French, Swedish, Russians, Ukrainians, and Belarusians. The proximity of the East Slavic populations to the main Western Europeans supports the view of archaeologists and anthropologists regarding a Central European origin for the Eastern Slavs. According to this type of research, the ancestral home of the Slavs was settled between the Oder and the Vistula (Sedov, 1979 as cited in Verbenko et al., 2003a).

The Figure 3 shows the results of multidimensional scaling of Nei's pairwise genetic distances calculated for Eastern European populations. The resulting graph can be

interpreted as both first and second dimensions, and considering the axes together. Based on the first dimension, it should be noted that a core group of populations inhabiting Eastern Europe is concentrated at the origin of the coordinates, whereas the populations with Asian origin of Kalmyk and Yakut are visibly removed from it. A common alliance of Eastern Slavs (Russian, Ukrainian, and Byelorussian populations), Northern Caucasians (Adygeys – including Adygei-Shapsugs of the Black Sea coast – Abkhazians, and Circassian populations), and populations living close to the Ural Mountains region (Komi, Bashkirs, and Mari) may be distinguished further taking into account the second dimension.

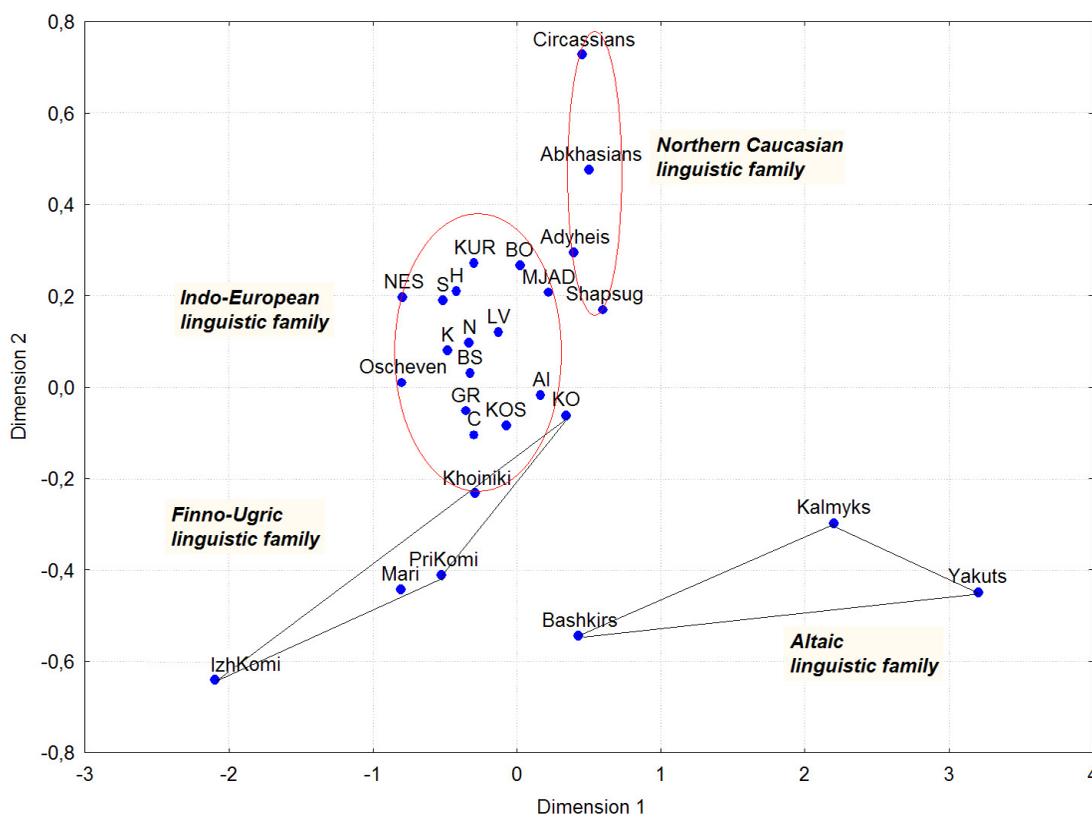


Fig. 3. Multidimensional scaling plot (two dimensions) of Nei's genetic distances among 26 populations of Eastern Europe and one population from Siberia based on 3' APOB minisatellite variability. Linguistic affiliations of populations are designated with geometrical figures. Abbreviations are: Russians: Cossacks (C), Belaya Sluda (BS), Kholmogory (H), Kostroma (KOS), Novgorod (N), Oschevensk (Oscheven), Smolensk (S), Kursk (KUR); Belarussians: Grodno (GR), Khojniki, Nesvij (NES), Mjadel' (MJAD), Bobruisk (BO); Ukrainians: Kiev (K), Lviv (LV), Alchevsk (AL); Other ethnic groups: Circassians, Abkhazians, Adygeis, Shapsugs (North Caucasus geographic region); Bashkirs (Beloretsky region), Komi-Permyats (KO), Izhemski Komi (IzhKomi), Priluzski Komi (PriKomi), Maris (Ural geographic region), Kalmyks, Yakuts.

If the linguistic classification of populations is taken into account, Eastern Slavonic, Northern Caucasian, Altaic, and Finno-Ugric clusters can be assigned. Within the main cluster of Eastern Slavs, discrete Russian, Ukrainian (Kravchenko et al., 1996), and Belarusian populations can be differentiated easily. Despite their wide geographical distribution, the Eastern Slavonic populations (Russians, Ukrainians, and Byelorussians) have a common

historical lineage and are also closely associated according to their 3'*APOB* polymorphisms. For example, although Kuban Cossacks (an ethnic community of Russians from the Krasnodar region) inhabit part of the Northern Caucasus, they are closer to the Eastern Slavonic populations than to other populations of their locality. The closest relationships among the Eastern Slavonic linguistic group are between the Russian and Ukrainian populations. The greatest diversity is found for the Byelorussian populations, which are distributed around other Eastern Slavonic populations on the plot. This diversity was possibly caused by long-term gene flow during numerous migrations through the Belarus region.

The cluster of the Finno-Ugric linguistic family and the Northern Caucasian linguistic family are close to the Eastern Slavonic linguistic group cluster. The Bashkirs belong to the Altaic language family, although their localization in the graph is close to the populations of the Finno-Ugric language family. However, there is some level of proximity of the Bashkir population to other peoples of the Altaic language family (the Kalmyk and Yakut). The position of Komi-Permyats, in this case in the immediate vicinity of the Eastern Slavonic populations, alienates them from the populations living close to the Ural Mountains region and may be due to the peculiarities of the ethnic history of the Komi-Permyats (Bunak, 1965; Kuzeev, 1985). We know that this ethnic group was separated from the Komi people only a few centuries ago, and is characterized recently by very close contact with Eastern Slavonic populations, which apparently left an imprint on the formation of the gene pool of the Komi-Permyats. The special arrangement of the Izhemski Komi, which alienates them from other groups, may also be due to the peculiarities of the ethnic history of this group (Khrunin et al., 2007). We found similar 3'*APOB* diversity among Eastern Slavonic and Northern Caucasus ethnic groups. However, there were significant differences for the Kalmyk and Yakut populations of Asian origin, as well as for Uralic Komis, Mari, and Bashkirs. The differences observed are similar to those obtained based on other DNA polymorphisms (Belyaeva et al., 1999, 2003; Bermisheva et al., 2001; Khar'kov et al., 2004; Kravchenko et al., 2002; Malyarchuk et al., 2001, 2002; Mirabal et al., 2009; Orekhov et al., 1999; Popova et al., 1999, 2001; Shabrova et al., 2004) and are in good agreement with ethnohistorical (Ageeva, 2000) and anthropological data (Alexeeva, 1973, Sedov, 1979). These observations underscore the significance of the 3'*APOB* minisatellite locus for population genetics research.

Thus, the 3'*APOB* minisatellite locus exhibits decreased genetic heterogeneity among 16 broadly distributed Eastern Slavonic populations, in contrast with its significant heterogeneity among Northern Caucasian populations and among Altaic and Finno-Ugric-speaking populations. This peculiarity may reflect the integrity of the Eastern Slavonic gene pool and suggests negligible influences from neighboring ethnic groups during the process of origin and differentiation of Eastern Slavs at the 3'*APOB* genome site. At the same time, the genotype frequency distribution revealed significant differences between Eastern Slavonic groups—both among the ethnic groups and between closely related populations belonging to one ethnic group that reveals high-differentiation properties of this marker.

## 5. D1S80 minisatellite polymorphism in human population research

The hypervariable minisatellite locus D1S80 (pMCT118) is the second most frequently used marker in population studies; it is located in the short arm of chromosome 1 and is a tandemly organized repeating region with an elementary link of 16 bp (Nakamura et al, 1988).

D1S80 is located at a distance of 16.5 kb from the start of the gene that encodes the  $\eta 2$  subunit of phospholipase, which plays an important role in the calcium metabolism of cerebral neurons, but no imbalance due to coupling between D1S80 region and the sequence of this gene has been found (Jeffreys et al, 1985; Sajantila et al., 1992; Tanaka, 2005). The allele variants of D1S80 vary in length because of variable repetition of the elementary link (15 to 41 or more repeats). According to the notation of alleles by Nakamura et al. (Nakamura et al, 1988), different alleles are designated in compliance with the number of repeats. The spectrum and frequencies of D1S80 alleles have been described fairly comprehensively, as this locus has been used intensively in criminalistics and forensic medical examinations (Kasai et al., 1990). Subsequently, Budowle et al. suggested the possibility of using D1S80 to differentiate populations (Budowle et al., 1991, 1995). The first investigation of the variability of this locus on a global scale, namely, in 43 populations from different regions of the world, was published by Duncan et al. in 1996 (Duncan et al., 1996-97). Clear distinctions were noted between populations of different main human groups and high similarity was shown among populations of the same main human group.

The subsequent report of a global analysis of D1S80 variability, performed by Mastana and Papiha (Mastana & Papiha, 2001), described the study of the marker in 84 world populations. The authors presented the spectra of D1S80 allele frequencies and, using the method of factor correspondence analysis, revealed clear-cut distinctions between European, Asian, Afro-American, American Indian, and Indian ethnic groups. Subsequently, the D1S80 polymorphism was analyzed in 33 world populations with a focus on the variability of the marker in sub-Saharan African populations (aboriginals of Africa) and a population of Arabian origin (the population of Egypt) (Herrera et al., 2004). As the differentiation of ethnic groups based on D1S80 data provided a very good description of the peculiarities of the groups, which were demonstrated previously via the analysis of biochemical markers, and conforms with the geographical locations of the populations with the peculiarities of their origin, the authors drew a conclusion allowing the applicability of only one marker, D1S80, to the study of the phylogenetic interrelationships of populations (Herrera et al., 2004).

A multimodal distribution is the distinctive feature of the spectrum of D1S80 allele frequencies. Some D1S80 alleles occur quite frequently, e.g., the total frequency of allele variants with 18 and 24 repeats is as high as 70% (Das & Mastana, 2003, Herrera et al., 2004, Walsh & Eckhoff, 2007). The first major allele, which contains 18 repeat units, occurs in 5.5–9% of sub-Saharan African populations, in 15–21% of Asian populations, and in 13–35% of European populations. The second major allele (24 repeats) has a frequency of 26–45% in Europeans, 6–29% in sub-Saharan Africans, and 17–24% in Asians (Das & Mastana, 2003; Duncan et al., 1996-97; Budowle et al., 1991, 1995; Herrera et al., 2004; Sajantila et al., 1992).

Some of the allele variants may either have a very high or very low frequency in particular populations, or even in the main human groups. Thus, in comparison with other main human groups, sub-Saharan Africans are characterized by very high frequencies of alleles 17 (up to 10%), 21 (up to 16%), 22 (up to 12%), 28 (up to 20%), and 34 (up to 31%), and a low frequency of allele 18. Moreover, the diversity of the aboriginal groups of Africa is so high that each population has its own typical profile of distribution with different numbers of repeats and positions of modes (modal values). This picture is also typical for the aboriginals of northern Australia.

We studied the polymorphism of the D1S80 minisatellite in 32 populations from the Eastern European region (Verbenko et al., 2003b, 2004, 2006, 2007; Khrunin et al., 2007; Limborska et al., 2011a). The study revealed the presence of 27 allele variants of the D1S80 minisatellite, with sizes ranging from 15 to more than 41 repeats and with varying frequencies in these populations. Various populations of Eastern Europe have 11 (in Ukrainians from Lviv (Kravchenko et al., 2001) and Byelorussians from Nesvij) to 20 (in Kalmyks) different allele variants. The broad spectrum of variability observed for the alleles of the D1S80 minisatellite provides evidence of the high level of polymorphism present in the populations under study. Though the allele frequencies vary significantly in different populations, their common features can be traced in their distribution. As a rule, three alleles (with 18, 24, and 31 repeats) occur at maximal frequency (the total frequency of their occurrence is 50–75%), one or two alleles occur more rarely (usually, these are variants containing 22, 25, 28, and 30 repeats, with frequencies of up to 11%), whereas other alleles usually occur with a frequency of less than 5%. Allele 24 is predominant in the European populations of Eastern Slavs (Russians, Belarusians, and Ukrainians) and allele 18 predominates in the populations of Kalmyks and Yakuts, which have an Asian origin. The populations of the Volga-Ural region (Tatars, Udmurts, Bashkirs, Maris, and Komis) and the populations of the Adygei-Abkhazia group (Adygeis, Abkhazians, and Circassians) have approximately the same frequencies of alleles 18 and 24. The frequency of allele 31 is low in European populations, intermediate in Asians, and maximal (up to 17%) in some populations of the Volga-Ural region.

The distribution of D1S80 allele frequencies in the populations studied is multimodal (see examples in Figure 4). The spectrum of alleles in European (Russian), Asian (Yakut), and Uralic (Udmurt) populations has common maxima for alleles 18 and 24. The ratio of the frequencies of alleles 18 and 24 is unequal among populations of the main human groups; the phenomenon of inversion of the frequency of the major alleles 18 and 24 is particularly noticeable between Asian and European populations. The comparison of D1S80 allele frequency distributions between the populations studied and worldwide populations revealed a similarity between Central European populations and Eastern Slavs, and between Yakut and Kalmyk populations and other Asian populations of China and Japan (Verbenko et al., 2006).

The capacity of minisatellite D1S80 to differentiate Eastern European populations was studied using multidimensional scaling of Nei's genetic distance matrix based on D1S80 allele distributions (Fig. 5, mathematical space). The main group including the Eastern Slav and Adygei-Abkhazian populations is concentrated in the cluster to the right of the origin of coordinates; the genetic relationship of these two groups can be interpreted easily based on their common European origin. Thus, European populations form one of the main clusters on the multidimensional scaling plot. Populations with an Asian origin (Kalmyks and Yakuts) are characterized by significant remoteness from Europeans. Populations of the Ural geographic region (Udmurts, Maris, Komis, Bashkirs, and Tatars) are located to the right of the origin of coordinates, in an intermediate position between the European and Asian populations. The second dimension provides the distinct differentiation between the populations that live close to the region of the Ural Mountains and the populations with an Asian origin.

The grouping of populations according to linguistic classification is indicated in the Figure 5. The populations of the Eastern Slavonic linguistic family (Russians, Ukrainians, and

Belarusians) form a single cluster, which also includes the populations of the Northern Caucasian linguistic family, as a subcluster. Russian populations from the Arkhangelsk region (Oschevensk and Belaya Sluda settlements), which are located within the cluster of the Finno-Ugric linguistic group, are the exception. The peculiar positions of these populations may be a reflection of peculiarities in their ethnic history regarding the variability of the D1S80 minisatellite; they were formed in the course of the development of northern lands by Russians under the active assimilation of native Finno-Ugric peoples, starting from the second millennium AD. This fact seems to have left a mark on the formation of their gene pool. A comparison of these findings with the results obtained in the study of mitochondrial DNA and Y chromosome markers revealed the corresponding peculiarities of these populations. Thus, a detailed analysis of the combination of nine polymorphic restriction sites with mutations of the first hypervariable segment of the mitochondrial DNA of the Russian population of Oschevensk in the Arkhangelsk region revealed a high frequency of the U5b1 mitotype, which is typical of the Finno-Ugric population of Lapps (Belyaeva et al, 2003; Balanovsky et al, 2008; Limborska et al., 2011b). An analysis of the haplotypes of Y chromosome microsatellite markers in the populations of Eastern Slavs demonstrated the kinship between these populations; only the population of Oschevensk in the Arkhangelsk region was characterized by a special set of haplotypes. Comparison of these findings with the variability of Y chromosome haplotypes in European populations revealed the kinship between the Arkhangelsk population from Oschevensk and not only Slavic, but also Finno-Ugric populations (Khrunin et al., 2005). The analysis of the data of mitochondrial DNA and Y chromosome polymorphisms, taking into consideration the peculiarities of ethnic history, leads to the conclusion that the gene pools of some Russian populations from the Arkhangelsk region contain an ancestral Finno-Ugric component.

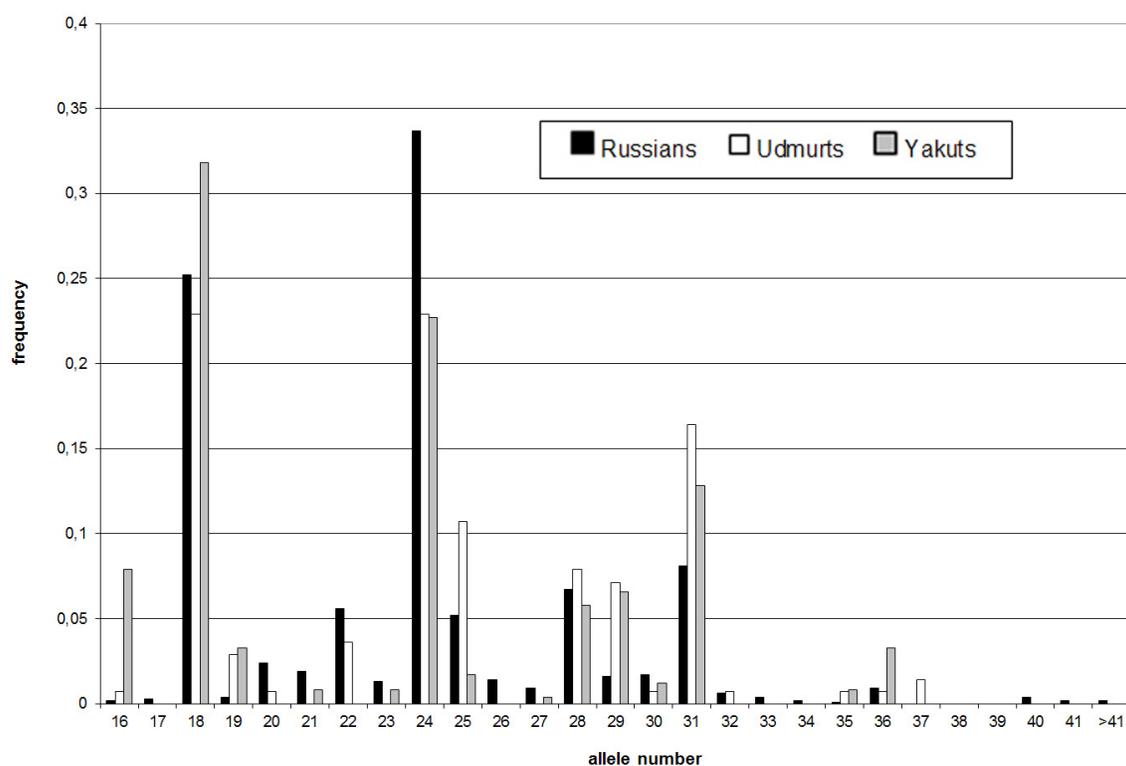


Fig. 4. D1S80 minisatellite allele frequency distributions in populations from Russia.

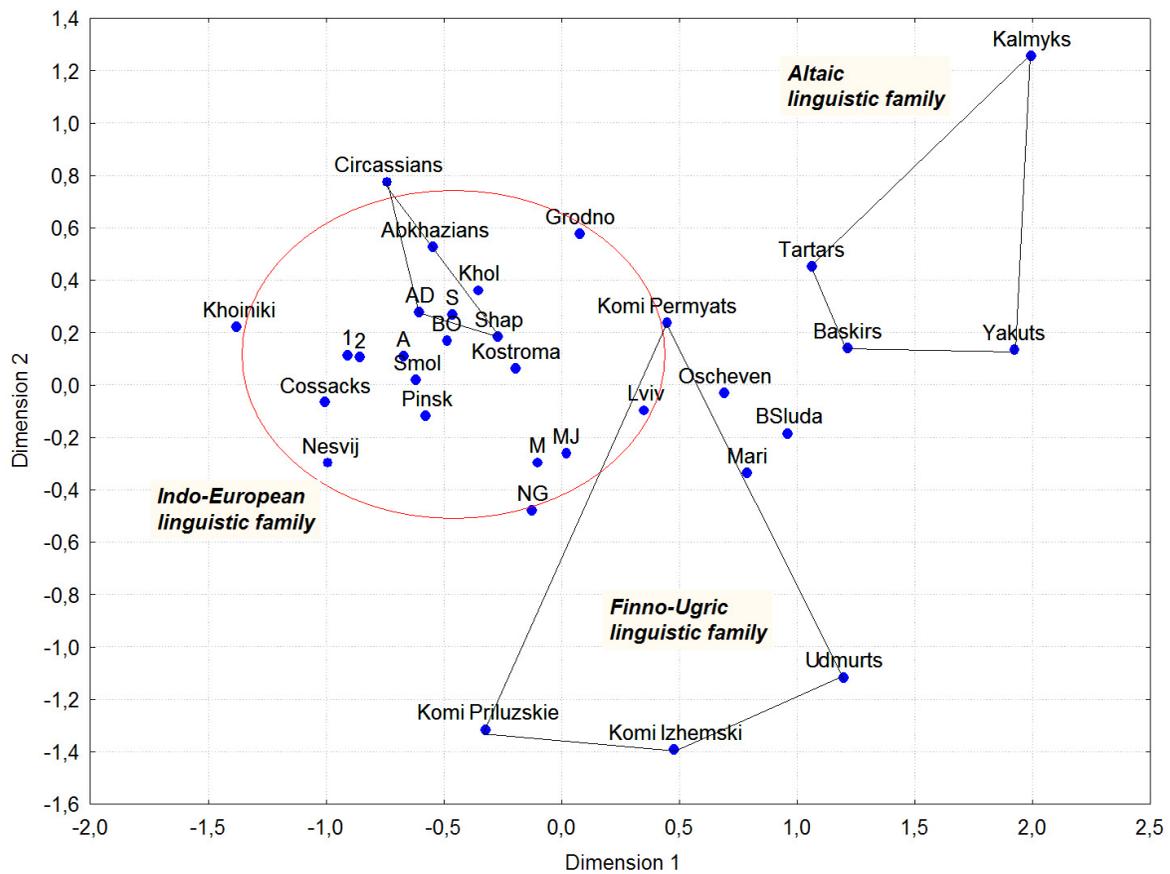


Fig. 5. Multidimensional scaling plot (two dimensions) of Nei's genetic distances among 31 populations of Eastern Europe and one population from Siberia. Linguistic affiliations of populations are designated with geometrical figures. The triangle includes the populations of North Caucasus linguistic family (except for Russians). Abbreviations are: European origin populations: Circassians, Abkhazians, Adygeis (AD), Shapsugs (Shap) (North Caucasus geographic region); Russians: Cossacks, Belaya Sluda (BSluda), Kholmogory (Khol), Kostroma, Novgorod (NG), Oschevensk (Osheven), Mezen (M), Smolensk (Ugra district) (Smol), Smolensk (Sychevka district) (S), Kursk (1); Belarussians: Grodno, Khojiniki, Nesvij, Mjadel' (MJ), Bobruisk (BO), Pinsk; Ukrainians: Kiev (2), Lviv, Alchevsk (A); Circassians, Abkhazians, Adygeis (AD), Shapsugs (Shap) (North Caucasus geographic region); Other ethnic groups: Bashkirs, Tartars, Komi-Permyats, Izhemski Komi, Priluzski Komi, Udmurts, Maris (Ural geographic region), Kalmyks, Yakuts.

Based on the aforementioned data, we can conclude that there is a spectrum of distribution of allele frequencies of specific alleles of D1S80 for the main human groups and for individual populations. It should be noted that the nature of the variability of this minisatellite differs from that of the 3'APOB minisatellite locus; however, both represent markers that can be used to differentiate clearly the major human groups and identify the peculiarities of individual populations. A good example is the case of the Russian populations of European origin that live in the Arkhangelsk district: the Oschevensk and the Belaia Sluda. These two populations from the same ethnic group can be readily distinguished from other Russian populations using multidimensional scaling of the D1S80 minisatellite variability genetic distance matrix, whereas the 3'APOB minisatellite, which is

located in a different chromosome, does not differentiate these populations among other Russians. One can propose that the characteristics of *3'APOB* and D1S80 reflect different aspects of the history of population evolution.

Thus, the minisatellite DNA markers *3'APOB* and D1S80 are sensitive and informative markers that can be applied to the study of the genetic structure of populations and population differences, and can be used to determine the genetic affinities among populations and to reconstruct their evolutionary relationships. The minisatellites D1S80 and *3'APOB* have been used extensively in population studies worldwide. Analysis of these loci in the population of Russia is an important part of population studies, both in terms of describing the variability of the gene pool, and as an annex to the global analysis of the origin and differentiation of human populations.

## 6. Allele spectrum shape subdivision using the SNP–VNTR haplotype at D1S80

To explore the evolutionary scenario underlying the complex allele distribution shape of D1S80 in different populations, we used an innovative technique of simultaneous determination of SNPs and minisatellites (Limborska et al., 2011a). Using fluorescently labeled primers and fragment analysis via capillary electrophoresis, we identified a hypervariable combination of the minisatellite polymorphism at the D1S80 locus and a SNP ( $G>T$ , rs16824398) adjacent to (74 bp) the minisatellite. The approach applied allows the determination of autosomal haplotypes representing the combinations of VNTR alleles with certain repeat numbers and alleles of the flanking SNP (Figure 6). A comparison of the SNPrs1682498-D1S80 haplotype frequencies was performed in populations of European (Russians), Asian (Yakuts), and African origin (from the sub-Saharan region; student volunteers from the Peoples' Friendship University of Russia, Moscow).

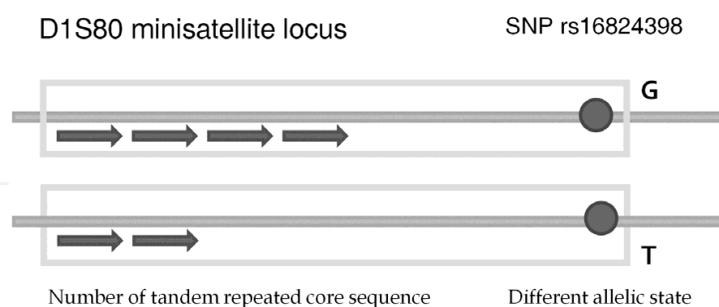


Fig. 6. Schematic of SNP-VNTR system (hypervariable minisatellite polymorphism combination of locus D1S80 and a single-nucleotide polymorphism) depicting double heterozygote autosomal haplotype for a diploid organism. In this example, one homolog has a G allele at the SNP and minisatellite locus of four (as e.g.) repeat units. The other homolog has a T allele at the SNP and minisatellite locus of two (as e.g.) repeat units.

The distributions of D1S80 allele frequencies in the populations studied are shown in Figures 7-9. Twenty-two alleles containing 16–41 repeats were detected among the 820 chromosomes typed. The allele spectra of all populations were multimodal, with the main peaks at alleles 18 and 24 in the Russian and Yakut samples. The frequency of allele 24 was high in the two Russian populations and the frequency of allele 18 was highest among the

Yakuts. The African allele spectrum was expanded and had peaks at alleles 18 and 24, albeit at lower frequency; the other major alleles were alleles 21 and 28.

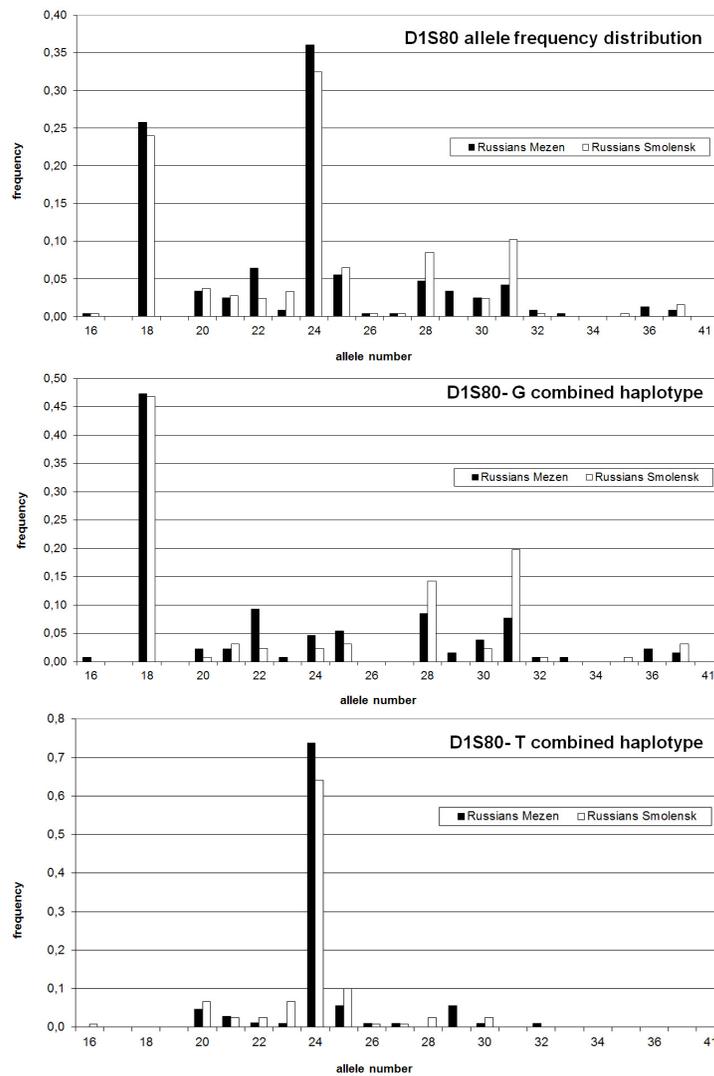


Fig. 7. D1S80 minisatellite allele frequency distributions in two geographically Russian populations (Europeans ).

An analysis of the amplified SNPrs1682498-D1S80 allelic pairs revealed the chromosome-related specificity of D1S80 allele spectra. A likelihood-ratio test was used to assess the significance of the LD between each D1S80 allele and SNP alleles. One of the most frequent variants, allele 24, was significantly associated with the T allele in non-African samples ( $D' = 0.75-0.93$ ;  $P < 1 \times 10^{-4}$ ). This combination was 10 times more frequent than the combination including the G allele. No significant association between allele 24 and the SNP background was observed in the African group, in which both combinations occurred almost equally (20.5% and 15.1%, respectively;  $D' = 0.11$ ;  $P = 0.465$ ). In Africans, the T allele was associated most strongly with allele 21. This combination also occurred 10 times more frequently than the combination including the G allele. Another frequent D1S80 allele containing 18 repeats was in complete LD ( $D' = 1.00$ ) with the G allele and was not detected on the T background in any of the populations. However, it occurred about three times more frequently in non-

African populations (40.4–47.3%) than in African samples. The other common D1S80 alleles (28 and 31) and population-specific alleles (e.g., 16 in Yakuts and 33 in Africans) were also associated mainly with the G allele. One exception was allele 31 among the Yakuts, which had no specific linkage with any of the SNP alleles (G, 12.3%; T, 10.9%;  $D' = 0.09$ ;  $P = 0.785$ ).

Generally, the African samples showed greater genetic diversity for each SNP allele background than did the samples from the other populations. Lower values of allele diversity were estimated on the T background in each population. This could be explained by the observation that most D1S80 alleles were less frequent on the T background, and that the absolute number of alleles was also lower compared with that observed for the G background (including the absence of alleles with more than 32 repeats on the T background). A comparison of the SNP-linked D1S80 allele distributions between pairs of populations revealed a high degree of similarity ( $P > 0.01$ ) among the allele spectra on the T background in non-African samples. In the case of the G-background-related distributions, a similarity was evident only among Russian samples.

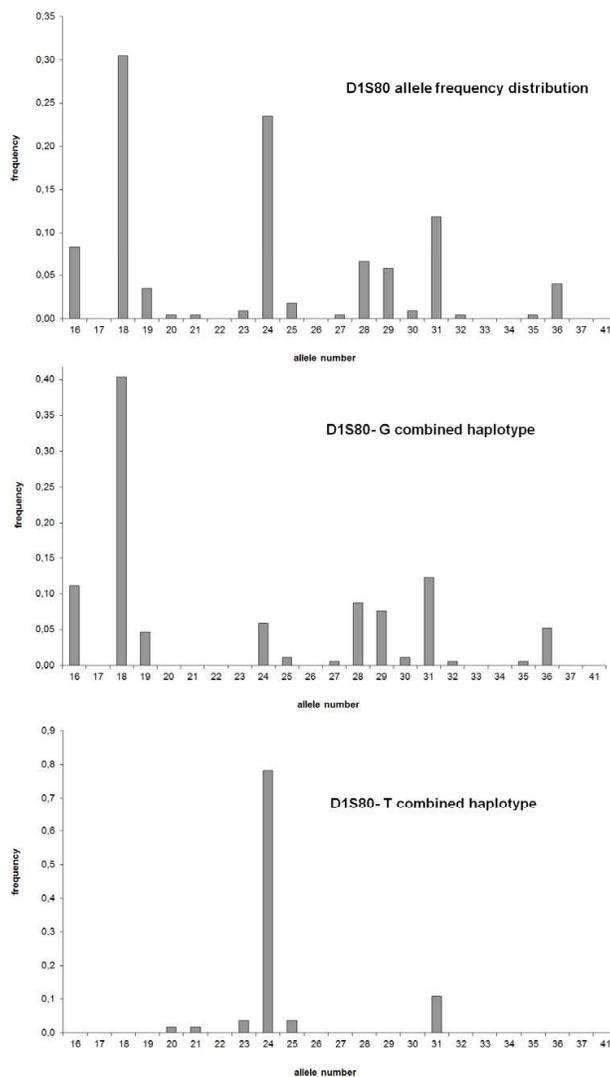


Fig. 8. D1S80 minisatellite allele frequency distributions in Yakut population (Asians).

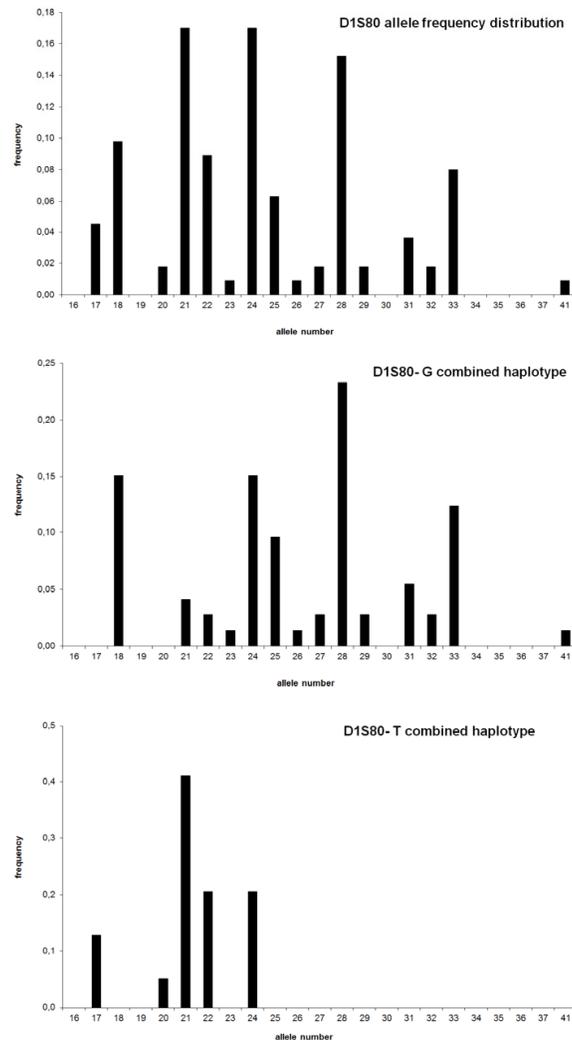


Fig. 9. D1S80 minisatellite allele frequency distributions in Sub-Saharan African population

The allele frequency patterns of the D1S80 locus are multimodal in many different populations (Duncan et al. 1997, Verbenko et al. 2006, Khrunin et al., 2007; Walsh and Eckhoff 2007), and similar patterns were observed in the population samples presented in the current study. However, unlike other studies that did not provide any data on the potential contribution of homoplasy to the allele spectra observed, the use of a system including SNPs allowed us to describe and analyze the distribution patterns of the D1S80 alleles according to their chromosomal location. We started with the analysis of the chromosome-related patterns of D1S80 alleles in samples from Russia (Russians and Yakuts). We paid particular attention to the analysis of D1S80 polymorphisms among northern Russians (individuals from the Mezen district of the Arkhangelsk region), who exhibit differences from other Russians and proximity to the Finno-Ugric- and Baltic-speaking populations (Balanovsky et al. 2008; Khrunin et al. 2009; Limborska et al., 2011b). Yakuts are typical North Asians (Cavalli-Sforza et al. 1996) who are one of the groups that is genetically most distant from Russians (Khrunin et al. 2005, 2007; Verbenko et al., 2005; Flegontova et al., 2009). Subsequently, taking into account the extensive data available on

the D1S80 polymorphism among African populations, we compared the D1S80 allele patterns of Russians and Yakuts with our African samples. Although our African DNAs were collected from students who presumably came from all over sub-Saharan Africa, the final distribution of D1S80 allele frequencies observed did not differ significantly from the spectra described for African populations in general (Herrera et al., 2004).

The empirical haplotype phase determination subdivided the total D1S80 allele set into two haploid allele spectra marked with the corresponding alleles of the rs16824398 SNP. In non-African samples, the major D1S80 alleles (18 and 24) were associated with a different SNP background. In the context of this finding, the differences in D1S80 allele spectra between populations of different ethnic origins described previously may be explained by the different ratio of chromosomes with T and G alleles. In our study, these frequencies were close in Russian samples from Smolensk and Mezen, and were different in the Yakut population.

The comparison of D1S80 allele distributions on each of the SNP backgrounds in the populations studied suggests an African origin for both the European and Asian SNP-VNTR haplotypes. In the non-African samples, the most frequent VNTR allele had 24 repeats on chromosomes carrying the T allele and 18 repeats on chromosomes carrying the G allele; the frequencies of these alleles were lower in the African samples. The most frequent alleles in the African samples were VNTR alleles with 21 repeats (on chromosomes with the T allele) and 28 repeats (on chromosomes with the G allele). Generally, the African samples revealed an expanded spectrum of frequent VNTR alleles on both types of chromosomes. This observation is consistent with the corresponding greater values of genetic diversity estimated for Africans.

Our results suggest that, during their migration out of Africa, modern humans carried only a subset of the VNTR spectrum on each of the SNP backgrounds tested. The further evolutionary history of non-African groups (European and Asian ancestors) was accompanied by founding bottlenecks, which were stronger in Asian populations (Keinan et al. 2007, Auton et al. 2009), resulting in the current different ratio of chromosomes with G and T alleles. The greater genetic diversity estimated on the G background for each population suggests that the rs16824398 locus with G at a polymorphic position is a more ancestral sequence than the T-containing variant. This is also consistent with the results of the analysis of corresponding reference sequences in apes, where only G variants are found (the Ensembl project).

Thus, our study demonstrated the effectiveness of applying a haplotyping approach to the analysis of VNTR polymorphisms. Using this method, we identified the chromosome-related characteristics of D1S80 allele patterns, as well as their population features. Taken in the context of other studies, our main findings also illustrate clearly the potential advantages of this SNPrs1682498-D1S80 system over single D1S80 locus testing in population studies and in cases of complex kinship diagnoses.

The subdivision of the D1S80 allele spectrum shape on the linked SNP background is indicative of populations of the main human groups. Considering the SNP ancestral state, the data obtained conform to the out-of-Africa hypothesis of evolution of human populations and provide some details regarding the migration scenario of the main human groups. The application of this SNP-VNTR approach to different sites of the autosomal genome may provide detailed insights into population microevolution.

## 7. Conclusion

In the past few years, minisatellites have been on the periphery of the attention of researchers. The current focus of interest is the use of SNPs in population studies because of new high-yield methods of testing (DNA microarrays).

Recently, another type of tandemly repeated hypervariable regions of the genome—microsatellites—received renewed interest, as some of them turn out to be the basis of a number of hereditary diseases. This has contributed to the development of new studies of microsatellites, which resulted in certain fundamental conclusions regarding their origin in evolution, the direction of their variability, and their role in the functioning of genomes.

However, existing situation does not negate the significance of minisatellites as very important regions of the genome, especially considering that it has become clear that some of them have a regulatory role regarding the genome. Because of the significant advances in whole-genome sequencing technologies, it can be assumed that, in the near future, the detailed study of minisatellites will be possible as it becomes clear how the features of their appearance and their manner of variability are different from those of microsatellites.

Here, we have tried to show the relevance, importance, and effectiveness of minisatellites based on the example of the use of two minisatellites (*D1S80* and *3'APOB*) in population research. Each of the two minisatellites is located at different chromosomes and can trace the evolutionary trajectory of the population, as it is marked by the corresponding genome segment. It is believed that the genome contains a lot of sites, each of which keeps its own evolutionary history record.

The variability of each of the markers studied (*D1S80* and *3'APOB*) in populations of Eastern Europe agrees quite well with the picture obtained from the analysis of anthropological data, as well as with historical and ethnographic data. Most analyses of other types of DNA markers have yielded similar differentiation patterns for populations of Eastern Europe (for a review see Verbenko & Limborska, 2008; Limborska et al., 2011b). In the case of multidimensional scaling of the matrix of genetic distances for each of the minisatellites studied, Ural populations were located between the European and Asian populations, which is a visual representation of the human origin and diagnostic properties of the markers. The individual features of the *3'APOB* minisatellite marker include a good population segregation of the Eastern Slavs and Northern Caucasian language families. The use of the minisatellite marker *D1S80* together with the main human-group origin/diagnostic features, allows the determination of the specific evolutionary trajectories of individual populations. For example, this marker can be used to distinguish Eastern Slavonic populations from those with stronger expression of the Finno-Ugric component (northern Russian Arkhangelsk region).

The high degree of differentiation of populations based on the variability of the minisatellite markers *D1S80* and *3'APOB* in combination with the method of multidimensional scaling of data processing allows the differentiation of not only distant, but also closely related populations of one ethnic group. Despite the coincidence of main cluster patterns of population differentiation in Eastern Europe detected for each marker, the analysis of individual minisatellite markers can lead to the identification of the specific features of the population. The pattern of variation observed for each minisatellite marker is specific

because of the different positions of the markers in the genome. According to Dobzhansky, the action of evolutionary events (mutations and genetic drift, followed by the action of selection) on populations may result in different effects on some parts of the genome, which leaves a trace on the frequency distribution of alleles in the population (Dobzhansky T, 1970). Thus, we can assume that the frequency distribution of the allelic variants of the D1S80 and 3'APOB minisatellite loci can be used to characterize certain features of the genetic history of populations. Taking into account the significant differences observed between the frequency distributions of the allelic variants of minisatellite markers among populations of different main human groups and ethnic groups, these markers should be classified as highly differentiated and origin-diagnostic tools.

Special possibilities appeared in evolutionary studies with the simultaneous detection of minisatellites and a flanking SNP. The analysis of these combined haplotypes allows the subdivision of the allele spectrum based on differences in the allelic state of SNPs with decreased mutation rate. Considering the ancestral state of the SNP, these data provide some details regarding the migration scenario of the main human groups. The application of this SNP-VNTR approach to different sites of the autosomal genome may provide detailed insights into population microevolution. It can be assumed that the use of a set of such markers would allow the most informative description of the structure of the gene pool of a particular region, and the identification of the features of the microevolution, origin, and genetic history of its populations.

## 8. Acknowledgment

The authors are sincerely grateful to all of their colleagues, including the coauthors of publications that provided the basis of this chapter. The work was partially supported by the Ministry of Science and Education of Russian Federation, the Program of support of the leading scientific schools by the President of Russian Federation, programs of Presidium of Russian Academy of Sciences: "Molecular and Cell Biology" and "Fundamental Sciences for Medicine" (subprogram "Human Polymorphism"), and Russian Foundation for Basic Research. We express sincere appreciation to Dr. Denis V. Khokhrin for preparation of this manuscript.

## 9. References

- Ageeva, R.A. (2000). *Where one's kin from? Peoples of Russia: Names and Fates. Explanatory Dictionary*. Academia, ISBN 5-87444-033-X, Moscow
- Akhmetova, V.L., Khusainova, R.I., Iur'ev, E.B., Tuktarova, I.A., Petrova, N.V., Makarov, S.V., Kravchuk, O.I., Pai, G.V., Balanovskaia, E.V., Ginter, E.K. & Khusnutdinova, E.K. (2006). Analysis of polymorphism at nine nuclear genome DNA loci in maris. *Genetika*, Vol.42, No.2, (Feb), pp. 256-273, issn 0016-6758
- Alekseeva, T.I. (1973). *Ethnogenesis of Eastern Slavs*. (Book in Russian). Moscow State University, Moscow
- Amarger, V., Gauguier, D., Yerle, M., Apiou, F., Pinton, P., Giraudeau, F., Monfouilloux, S., Lathrop, M., Dutrillaux, B., Buard, J. & Vergnaud, G. (1998). Analysis of distribution in the human, pig, and rat genomes points toward a general

- subtelomeric origin of minisatellite structures. *Genomics*, Vol.52, No.1, (Aug 15), pp. 62-71, issn 0888-7543
- Auton, A., Bryc, K., Boyko, A.R., Lohmueller, K.E., Novembre, J., Reynolds, A., Indap, A., Wright, M.H., Degenhardt, J.D., Gutenkunst, R.N., King, K.S., Nelson, M.R. & Bustamante, C.D. (2009). Global distribution of genomic diversity underscores rich complex history of continental human populations. *Genome Res*, Vol.19, No.5, (May), pp. 795-803, issn 1088-9051
- Babushkina, N.P. & Kucher, A.N. Functional role of VNTR polymorphism of human genes. *Genetika*, Vol.47, No.6, (Jun), pp. 725-734, issn 0016-6758
- Balanovsky, O., Rootsi, S., Pshenichnov, A., Kivisild, T., Churnosov, M., Evseeva, I., Pocheshkhova, E., Boldyreva, M., Yankovsky, N., Balanovska, E. & Villems, R. (2008). Two sources of the Russian patrilineal heritage in their Eurasian context. *Am J Hum Genet*, Vol.82, No.1, (Jan), pp. 236-250, issn 1537-6605 (Electronic)
- Barysheva, E.V., Prosnjak, M.I., Vlasov, M.S., Golubtsov, V.I., Revazov, A.A., Limborskaia, S.A. & Ginter, E.K. (1989). The use of DNA from phage M13 for the analysis of interindividual polymorphism of human DNA as demonstrated by a population study in Krasnodar city. *Genetika*, Vol.25, No.11, (Nov), pp. 2079-2082, issn 0016-6758
- Barysheva, E.V., Bukina, A.M., Petrova, N.V., Limborskaia, S.A. & Ginter, E.K. (1991). Use of DNA polymorphism detected by M13 phage DNA in population studies. *Genetika*, Vol.27, No.3, (Mar), pp. 399-403, issn 0016-6758
- Barysheva, E.V., Bukina, A.M., Limborskaia, S.A. & Ginter, E.K. (1991). Analysis of genetic distances between populations using human DNA "fingerprints" detected by a phage M13 DNA probe. *Genetika*, Vol.27, No.9, (Sep), pp. 1493-1498, issn 0016-6758
- Bell, G.I., Selby, M.J. & Rutter, W.J. (1982). The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature*, Vol.295, No.5844, (Jan 7), pp. 31-35, issn 0028-0836
- Belyaeva, O.V., Balanovsky, O.P., Ashworth, L.K., Lebedev, Y.B., Spitsyn, V.A., Guseva, N.A., Erdes, S., Mikulich, A.I., Khusnutdinova, E.K. & Limborska, S.A. (1999). Fine mapping of a polymorphic CA repeat marker on human chromosome 19 and its use in population studies. *Gene*, Vol.230, No.2, (Apr 16), pp. 259-266, issn 0378-1119
- Belyaeva, O., Bermisheva, M., Khrunin, A., Slominsky, P., Bebyakova, N., Khusnutdinova, E., Mikulich, A. & Limborska, S. (2003). Mitochondrial DNA variations in Russian and Belorussian populations. *Hum Biol*, Vol.75, No.5, (Oct), pp. 647-660, issn 0018-7143
- Bennett, S.T., Lucassen, A.M., Gough, S.C., Powell, E.E., Undlien, D.E., Pritchard, L.E., Merriman, M.E., Kawaguchi, Y., Dronsfield, M.J., Pociot, F. & et al. (1995). Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nat Genet*, Vol.9, No.3, (Mar), pp. 284-292, issn 1061-4036
- Bermisheva, M.A., Viktorova, T.V., Beliaeva, O., Limborskaia, S.A. & Khusnutdinova, E.K. (2001). Polymorphism of hypervariable segment I of mitochondrial DNA in three ethnic groups of the Volga-Ural region. *Genetika*, Vol.37, No.8, (Aug), pp. 1118-1124, issn 0016-6758
- Bermisheva, M.A., Viktorova, T.V. & Khusnutdinova, E.K. (2003). Polymorphism of human mitochondrial DNA. *Genetika*, Vol.39, No.8, (Aug), pp. 1013-1025, issn 0016-6758

- Bermisheva, M.A., Petrova, N.V., Zinchenko, R.A., Timkovskaia, E.E., Malyshev, P., Gavrilina, S.G., Ginter, E.K. & Kusnutdinova, E.K. (2007). Population study of the Udmurt population: analysis of ten polymorphic DNA loci of the nuclear genome. *Genetika*, Vol.43, No.5, (May), pp. 688-705, issn 0016-6758
- Boerwinkle, E., Xiong, W.J., Fourest, E. & Chan, L. (1989). Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: application to the apolipoprotein B 3' hypervariable region. *Proc Natl Acad Sci U S A*, Vol.86, No.1, (Jan), pp. 212-216, issn 0027-8424
- Bois, P. & Jeffreys, A.J. (1999). Minisatellite instability and germline mutation. *Cell Mol Life Sci*, Vol.55, No.12, (Sep), pp. 1636-1648, issn 1420-682X
- Bois, P.R. (2003). Hypermutable minisatellites, a human affair? *Genomics*, Vol.81, No.4, (Apr), pp. 349-355, issn 0888-7543
- Buard, J. & Vergnaud, G. (1994). Complex recombination events at the hypermutable minisatellite CEB1 (D2S90). *EMBO J*, Vol.13, No.13, (Jul 1), pp. 3203-3210, issn 0261-4189
- Budowle, B., Chakraborty, R., Giusti, A.M., Eisenberg, A.J. & Allen, R.C. (1991). Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. *Am J Hum Genet*, Vol.48, No.1, (Jan), pp. 137-144, issn 0002-9297
- Budowle, B., Baechtel, F.S., Smerick, J.B., Presley, K.W., Giusti, A.M., Parsons, G., Alevy, M.C. & Chakraborty, R. (1995). D1S80 population data in African Americans, Caucasians, southeastern Hispanics, southwestern Hispanics, and Orientals. *J Forensic Sci*, Vol.40, No.1, (Jan), pp. 38-44, issn 0022-1198
- Bunak, V.V. (1965). *Origin and Ethnic History of the Russian Nation. (Proiskhozhdeniye i etnicheskaya istoriya russkogo naroda)*, Nauka, Moscow.
- Buresi, C., Desmarais, E., Vigneron, S., Lamarti, H., Smaoui, N., Cambien, F. & Roizes, G. (1996). Structural analysis of the minisatellite present at the 3' end of the human apolipoprotein B gene: new definition of the alleles and evolutionary implications. *Hum Mol Genet*, Vol.5, No.1, (Jan), pp. 61-68, issn 0964-6906
- Capon, D.J., Chen, E.Y., Levinson, A.D., Seeburg, P.H. & Goeddel, D.V. (1983). Complete nucleotide sequences of the T24 human bladder carcinoma oncogene and its normal homologue. *Nature*, Vol.302, No.5903, (Mar 3), pp. 33-37, issn 0028-0836
- Cavalli-Sforza, L.L. & Feldman, M.W. (2003). The application of molecular genetic approaches to the study of human evolution. *Nat Genet*, Vol.33 Suppl, (Mar), pp. 266-275, issn 1061-4036
- Chen, B., Guo, Z., He, P., Ye, P., Buresi, C. & Roizes, G. (1999). Structure and function of alleles in the 3' end region of human apoB gene. *Chin Med J (Engl)*, Vol.112, No.3, (Mar), pp. 221-223, issn 0366-6999
- Chistiakov, D.A., Gavrilov, D.K., Ovchinnikov, I.V. & Nosikov, V.V. (1993). Analysis of the distribution of alleles of four hypervariable tandem repeats among unrelated Russian individuals living in Moscow, using the polymerase chain reaction. *Mol Biol (Mosk)*, Vol.27, No.6, (Nov-Dec), pp. 1304-1314, issn 0026-8984
- Choong, M.L., Koay, E.S., Khaw, M.C. & Aw, T.C. (1999). Apolipoprotein B 5'-Ins/Del and 3'-VNTR polymorphisms in Chinese, Malay and Indian Singaporeans. *Hum Hered*, Vol.49, No.1, (Jan), pp. 31-40, issn 0001-5652
- Das, K. & Mastana, S.S. (2003). Genetic variation at three VNTR loci in three tribal populations of Orissa, India. *Ann Hum Biol*, Vol.30, No.3, (May-Jun), pp. 237-249, issn 0301-4460

- Deka, R., Chakraborty, R., DeCroo, S., Rothhammer, F., Barton, S.A. & Ferrell, R.E. (1992). Characteristics of polymorphism at a VNTR locus 3' to the apolipoprotein B gene in five human populations. *Am J Hum Genet*, Vol.51, No.6, (Dec), pp. 1325-1333, issn 0002-9297
- Destro-Bisol, G., Capelli, C. & Belledi, M. (2000). Inferring microevolutionary patterns from allele-size frequency distributions of minisatellite loci: a worldwide study of the APOB 3' hypervariable region polymorphism. *Hum Biol*, Vol.72, No.5, (Oct), pp. 733-751, issn 0018-7143
- Dobzhansky, T. (1970). *Genetics of the Evolutionary Process*, Columbia Univ. Press, ISBN-13: 978-0231083065, New York.
- Duncan, G., Thomas, E., Gallo, J.C., Baird, L.S., Garrison, J. & Herrera, R.J. (1996). Human phylogenetic relationships according to the D1S80 locus. *Genetica*, Vol.98, No.3, pp. 277-287, issn 0016-6707
- Dzhincharadze, A.G., Ivanov, P.L. & Ryskov, A.P. (1987). Genome "dactyloscopy". Characteristics of the human cloned sequence JIN 600 in vector M13 with properties of highly polymorphic DNA marker. *Dokl Akad Nauk SSSR*, Vol.295, No.1, (Jul-Aug), pp. 230-233, issn 0002-3264
- Flegontova, O.V., Khrunin, A.V., Lylova, O.I., Tarskaia, L.A., Spitsyn, V.A., Mikulich, A.I. & Limborska, S.A. (2009). Haplotype frequencies at the DRD2 locus in populations of the East European Plain. *BMC Genet*, Vol.10, pp. 62, issn 1471-2156 (Electronic)
- Fuke, S., Suo, S., Takahashi, N., Koike, H., Sasagawa, N. & Ishiura, S. (2001). The VNTR polymorphism of the human dopamine transporter (DAT1) gene affects gene expression. *Pharmacogenomics J*, Vol.1, No.2, pp. 152-156, issn 1470-269X
- Gemayel, R., Vinces, M.D., Legendre, M. & Verstrepen, K.J. Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu Rev Genet*, Vol.44, pp. 445-477, issn 1545-2948
- Harris, R.F. (2002). Hapmap flap. *Curr Biol*, Vol.12, No.24, (Dec 23), pp. R827, issn 0960-9822
- Herrera, R.J., Adrien, L.R., Ruiz, L.M., Sanabria, N.Y. & Duncan, G. (2004). D1S80 single-locus discrimination among African populations. *Hum Biol*, Vol.76, No.1, (Feb), pp. 87-108, issn 0018-7143
- Hong-Sheng, G., Peng, Z., Cheng-Bo, Y. & Sheng-Bin, L. (2009). HGD-Chn: The Database of Genome Diversity and Variation for Chinese Populations. *Leg Med (Tokyo)*, Vol.11 Suppl 1, (Apr), pp. S201-202, issn 1873-4162 (Electronic)
- Huang, L.S. & Breslow, J.L. (1987). A unique AT-rich hypervariable minisatellite 3' to the ApoB gene defines a high information restriction fragment length polymorphism. *J Biol Chem*, Vol.262, No.19, (Jul 5), pp. 8952-8955, issn 0021-9258
- Jeffreys, A.J., Wilson, V. & Thein, S.L. (1985). Hypervariable 'minisatellite' regions in human DNA. *Nature*, Vol.314, No.6006, (Mar 7-13), pp. 67-73, issn 0028-0836
- Jeffreys, A.J., Wilson, V., Neumann, R. & Keyte, J. (1988). Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. *Nucleic Acids Res*, Vol.16, No.23, (Dec 9), pp. 10953-10971, issn 0305-1048
- Jeffreys, A.J., Tamaki, K., MacLeod, A., Monckton, D.G., Neil, D.L. & Armour, J.A. (1994). Complex gene conversion events in germline mutation at human minisatellites. *Nat Genet*, Vol.6, No.2, (Feb), pp. 136-145, issn 1061-4036
- Jurka, J. & Gentles, A.J. (2006). Origin and diversification of minisatellites derived from human Alu sequences. *Gene*, Vol.365, (Jan 3), pp. 21-26, issn 0378-1119

- Kalnina, V.V., Kalnina, O.V., Prośniak, M.I., Khidiatova, I.M., Khusnutdinova, E.K., Raphicov, K.S., Limborska, S.A. (1995). Use of DNA fingerprinting for human population genetic studies, Vol. 247, No. 4, (May 20), pp. 488-93.
- Kawakami, K., Salonga, D., Park, J.M., Danenberg, K.D., Uetake, H., Brabender, J., Omura, K., Watanabe, G. & Danenberg, P.V. (2001). Different lengths of a polymorphic repeat sequence in the thymidylate synthase gene affect translational efficiency but not its gene expression. *Clin Cancer Res*, Vol.7, No.12, (Dec), pp. 4096-4101, issn 1078-0432
- Keinan, A., Mullikin, J.C., Patterson, N. & Reich, D. (2007). Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans. *Nat Genet*, Vol.39, No.10, (Oct), pp. 1251-1255, issn 1546-1718 (Electronic)
- Kelkar, Y.D., Tyekucheva, S., Chiaromonte, F. & Makova, K.D. (2008). The genome-wide determinants of human and chimpanzee microsatellite evolution. *Genome Res*, Vol.18, No.1, (Jan), pp. 30-38, issn 1088-9051
- Kendrew J.C., Lawrence E.L.(1994) *The encyclopedia of molecular biology*. Blackwell Science, ISBN 0-632-02182-9, Oxford
- Khar'kov, V.N., Stepanov, V.A., Borinskaia, S.A., Kozhekbaeva Zh, M., Gusar, V.A., Grechanina, E., Puzyrev, V.P., Khusnutdinova, E.K. & Iankovskii, N.K. (2004). Structure of the gene pool of eastern Ukrainians from Y-chromosome haplogroups. *Genetika*, Vol.40, No.3, (Mar), pp. 415-421, issn 0016-6758
- Khrunin, A., Mihailov, E., Nikopensus, T., Krjutskov, K., Limborska, S. & Metspalu, A. (2009). Analysis of allele and haplotype diversity across 25 genomic regions in three Eastern European populations. *Hum Hered*, Vol.68, No.1, pp. 35-44, issn 1423-0062 (Electronic)
- Khrunin, A.V., Bebiakova, N.A., Ivanov, V.P., Solodilova, M.A. & Limborskaia, S.A. (2005). Polymorphism of Y-chromosomal microsatellites in Russian populations from the northern and southern Russia as exemplified by the populations of Kursk and Arkhangel'sk Oblast. *Genetika*, Vol.41, No.8, (Aug), pp. 1125-1131, issn 0016-6758
- Khusnutdinova, E.K., Viktorova, T.V., Fatkhislamova, R.I. & Galeeva, A.R. (1999). Evaluation of the relative contribution of Caucasoid and Mongoloid components in the formation of ethnic groups of the Volga-Ural region according to data of DNA polymorphism. *Genetika*, Vol.35, No.8, (Aug), pp. 1132-1137, issn 0016-6758
- Khusnutdinova, E.K., Viktorova, T.V., Akhmetova, V.L., Mustafina, O.E., Fatkhislamova, R.I., Balanovskaia, E.V., Petrova, N.V., Makarov, S.V., Kravchuk, O.I., Pai, G.V. & Ginter, E.K. (2003). Population-genetic structure of Chuvashia (from data on eight DNA loci in the nuclear genome). *Genetika*, Vol.39, No.11, (Nov), pp. 1550-1563, issn 0016-6758
- Klenova, E., Scott, A.C., Roberts, J., Shamsuddin, S., Lovejoy, E.A., Bergmann, S., Bubb, V.J., Royer, H.D. & Quinn, J.P. (2004). YB-1 and CTCF differentially regulate the 5-HTT polymorphic intron 2 enhancer which predisposes to a variety of neurological disorders. *J Neurosci*, Vol.24, No.26, (Jun 30), pp. 5966-5973, issn 1529-2401 (Electronic)
- Kravchenko, S.A., Maliarchuk, O.S. & Livshits, L.A. (1996). A population genetics study of the allelic polymorphism in the hypervariable region of the apolipoprotein B gene in the population of different regions of Ukraine. *Tsitol Genet*, Vol.30, No.5, (Sep-Oct), pp. 35-41, issn 0564-3783

- Kravchenko, S.A., Slominskii, P.A., Bets, L.A., Stepanova, A.V., Mikulich, A.I., Limborskaia, S.A. & Livshits, L.A. (2002). Polymorphism of the STR-locus of Y chromosomes in Eastern Slavs in three populations from Belorussia, Russia and the Ukraine. *Genetika*, Vol.38, No.1, (Jan), pp. 97-104, issn 0016-6758
- Kravchenko, S.A., Malyarchuk, S.G. & Pampukha, V.M. (2001). *Genetika i selektsiia na Ukraine na rubezhe tysyacheletiy (Genetics and Selection in the Ukraine at the Boundary of Millennia)*, Vol. 4, pp. 410-422, Logos, Kiev
- Kuzeev, R.G. (1985). Peoples of Povolzhye and Priuralye. (Narody Povolzhya i Priuralya), Nauka, Moscow.
- Lahermo, P., Sajantila, A., Sistonen, P., Lukka, M., Aula, P., Peltonen, L. & Savontaus, M.L. (1996). The genetic relationship between the Finns and the Finnish Saami (Lapps): analysis of nuclear DNA and mtDNA. *Am J Hum Genet*, Vol.58, No.6, (Jun), pp. 1309-1322, issn 0002-9297
- Latorra, D., Stern, C.M. & Schanfield, M.S. (1994). Characterization of human AFLP systems apolipoprotein B, phenylalanine hydroxylase, and D1S80. *PCR Methods Appl*, Vol.3, No.6, (Jun), pp. 351-358, issn 1054-9803
- Levinson, G. & Gutman, G.A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol*, Vol.4, No.3, (May), pp. 203-221, issn 0737-4038
- Lewis, P.O., Zaykin, D. (2001). *Genetic Data Analysis: Computer program for the analysis of allelic data.* Version 1.0 (d16c). Available from: <http://lewis.eeb.uconn.edu/lewishome/software.html>
- Limborska, S.A., Khrunin, A.V., Flegontova, O.V., Tasitz, V.A. & Verbenko, D.A. (2011a). Specificity of genetic diversity in D1S80 revealed by SNP-VNTR haplotyping. *Ann Hum Biol*, Vol.38, No.5, (Sep), pp. 564-569, issn 1464-5033 (Electronic)
- Limborska, S.A., Verbenko, D.A., Khrunin, A.V., Slominsky, P.A., Bebyakova, N.A. (2011b) Ethnic genomics: analysis of genome polymorphism of Archangelsk region. *Bulletin of Moscow University. Series XXIII. Anthropology*, No.3, pp. 100-119, ISSN 0201-7385
- Ludwig, E.H., Friedl, W. & McCarthy, B.J. (1989). High-resolution analysis of a hypervariable region in the human apolipoprotein B gene. *Am J Hum Genet*, Vol.45, No.3, (Sep), pp. 458-464, issn 0002-9297
- Malyarchuk, B.A. & Derenko, M.V. (2001). Mitochondrial DNA variability in Russians and Ukrainians: implication to the origin of the Eastern Slavs. *Ann Hum Genet*, Vol.65, No.Pt 1, (Jan), pp. 63-78, issn 0003-4800
- Malyarchuk, B.A., Grzybowski, T., Derenko, M.V., Czarny, J., Wozniak, M. & Miscicka-Sliwka, D. (2002). Mitochondrial DNA variability in Poles and Russians. *Ann Hum Genet*, Vol.66, No.Pt 4, (Jul), pp. 261-283, issn 0003-4800
- Marz, W., Ruzicka, V., Fisher, E., Russ, A.P., Schneider, W. & Gross, W. (1993). Typing of the 3' hypervariable region of the apolipoprotein B gene: approaches, pitfalls, and applications. *Electrophoresis*, Vol.14, No.3, (Mar), pp. 169-173, issn 0173-0835
- Mastana, S.S. & Papiha, S.S. (2001). D1S80 distribution in world populations with new data from the UK and the Indian sub-continent. *Ann Hum Biol*, Vol.28, No.3, (May-Jun), pp. 308-318, issn 0301-4460
- Miklos, G.L. & John, B. (1979). Heterochromatin and satellite DNA in man: properties and prospects. *Am J Hum Genet*, Vol.31, No.3, (May), pp. 264-280, issn 0002-9297

- Miller, MP. (1997). *R by C: A program that performs Fisher's Exact test on any sized contingency table through the use of the Metropolis algorithm*, Available from: <http://bioweb.usu.edu/mpmbio/>
- Mirabal, S., Regueiro, M., Cadenas, A.M., Cavalli-Sforza, L.L., Underhill, P.A., Verbenko, D.A., Limborska, S.A. & Herrera, R.J. (2009). Y-chromosome distribution within the geo-linguistic landscape of northwestern Russia. *Eur J Hum Genet*, Vol.17, No.10, (Oct), pp. 1260-1273, issn 1476-5438 (Electronic)
- Nakahara, M., Shimozawa, M., Nakamura, Y., Irino, Y., Morita, M., Kudo, Y. & Fukami, K. (2005). A novel phospholipase C, PLC(eta)2, is a neuron-specific isozyme. *J Biol Chem*, Vol.280, No.32, (Aug 12), pp. 29128-29134, issn 0021-9258
- Nakamura, Y., Carlson, M., Krapcho, K. & White, R. (1988). Isolation and mapping of a polymorphic DNA sequence (pMCT118) on chromosome 1p D1S80. *Nucleic Acids Res*, Vol.16, No.19, (Oct 11), pp. 9364, issn 0305-1048
- Nei, M. (1972). Genetic Distance between Populations. *The American Naturalist*, Vol.106, No.949, pp. 283-292
- Orekhov, V., Poltoraus, A., Zhivotovsky, L.A., Spitsyn, V., Ivanov, P. & Yankovsky, N. (1999). Mitochondrial DNA sequence diversity in Russians. *FEBS Lett*, Vol.445, No.1, (Feb 19), pp. 197-201, issn 0014-5793
- Paquette, J., Giannoukakis, N., Polychronakos, C., Vafiadis, P. & Deal, C. (1998). The INS 5' variable number of tandem repeats is associated with IGF2 expression in humans. *J Biol Chem*, Vol.273, No.23, (Jun 5), pp. 14158-14164, issn 0021-9258
- Pena, S.D. & Chakraborty, R. (1994). Paternity testing in the DNA era. *Trends Genet*, Vol.10, No.6, (Jun), pp. 204-209, issn 0168-9525
- Poltl, R., Luckenbach, C., Reinhold, J., Fimmers, R. & Ritter, H. (1996). Comparison of German population data on the apoB-HVR locus with other Caucasian, Asian and black populations. *Forensic Sci Int*, Vol.80, No.3, (Jul 12), pp. 221-227, issn 0379-0738
- Popova, S.N., Mikulich, A.I., Slominskii, P.A., Shadrina, M.I., Pomazanova, M.A. & Limborskaia, S.A. (1999). Polymorphism of the (CTG)<sub>n</sub> repeat in the myotonin protein kinase (DM) gene in Belarussian populations: analysis of interethnic heterogeneity. *Genetika*, Vol.35, No.7, (Jul), pp. 994-997, issn 0016-6758
- Popova, S.N., Slominsky, P.A., Pocheshnova, E.A., Balanovskaya, E.V., Tarskaya, L.A., Bebyakova, N.A., Bets, L.V., Ivanov, V.P., Livshits, L.A., Khusnutdinova, E.K., Spitsyn, V.A. & Limborska, S.A. (2001). Polymorphism of trinucleotide repeats in loci DM, DRPLA and SCA1 in East European populations. *Eur J Hum Genet*, Vol.9, No.11, (Nov), pp. 829-835, issn 1018-4813
- Popova, S.N., Slominskii, P.A., Galushkin, S.N., Tarskaia, L.A., Spitsyn, V.A., Guseva, I.A. & Limborskaia, S.A. (2002). Analysis of the allele polymorphism of (CTG)<sub>n</sub> and (GAG)<sub>n</sub> triplet repeats in DM, DRPLA, and SCA1 genes in various populations of Russia. *Genetika*, Vol.38, No.11, (Nov), pp. 1549-1553, issn 0016-6758
- Proudfoot, N.J., Gil, A. & Maniatis, T. (1982). The structure of the human zeta-globin gene and a closely linked, nearly identical pseudogene. *Cell*, Vol.31, No.3 Pt 2, (Dec), pp. 553-563, issn 0092-8674
- Renges, H.H., Peacock, R., Dunning, A.M., Talmud, P. & Humphries, S.E. (1992). Genetic relationship between the 3'-VNTR and diallelic apolipoprotein B gene polymorphisms: haplotype analysis in individuals of European and south Asian origin. *Ann Hum Genet*, Vol.56, No.Pt 1, (Jan), pp. 11-33, issn 0003-4800

- Ryskov, A.P., Tokarskaia, O.N., Verbovaia, L.V., Dzhincharadze, A.G. & Gintsburg, A.L. (1988). Genomic fingerprinting of microorganisms: its use as a hybridization probe of phage M13 DNA. *Genetika*, Vol.24, No.7, (Jul), pp. 1310-1313, issn 0016-6758
- Ryskov, A.P., Faizov, T., Alimov, A.M. & Romanova, E.A. (1990). Genomic fingerprinting: new possibilities in determining the species identity of *Brucella*. *Genetika*, Vol.26, No.1, (Jan), pp. 130-133, issn 0016-6758
- Sajantila, A., Budowle, B., Strom, M., Johnsson, V., Lukka, M., Peltonen, L. & Ehnholm, C. (1992). PCR amplification of alleles at the DIS80 locus: comparison of a Finnish and a North American Caucasian population sample, and forensic casework evaluation. *Am J Hum Genet*, Vol.50, No.4, (Apr), pp. 816-825, issn 0002-9297
- Sajantila, A., Lukka, M. & Syvanen, A.C. (1999). Experimentally observed germline mutations at human micro- and minisatellite loci. *Eur J Hum Genet*, Vol.7, No.2, (Feb-Mar), pp. 263-266, issn 1018-4813
- Schlotterer, C. (1998). Microsatellites, In: *Molecular genetics analysis of populations*, Hoelzel, A.R., pp. 237, IRL Press, Oxford Univ. Press, ISBN 9780199636358, London
- Sedov, V.V. (1979). *Origin and early history of Slavs* (Book in Russian). Nauka, Moscow
- Semenova, S.K., Romanova, E.A. & Ryskov, A.P. (1996). Genetic differentiation of helminths on the basis of data of polymerase chain reaction using random primers. *Genetika*, Vol.32, No.2, (Feb), pp. 304-309, issn 0016-6758
- Semina, E.V., Bukina, A.M., Startseva, E.A., Limborskaia, S.A. & Ginter, E.K. (1993). Genetic distances between various ethnic populations calculated on the basis of polymorphism of DNA detected by the hypervariable phage M13 DNA probe. *Genetika*, Vol.29, No.10, (Oct), pp. 1612-1619, issn 0016-6758
- Shabrova, E.V., Khusnutdinova, E.K., Tarskaia, L.A., Mikulich, A.I., Abolmasov, N.N. & Limborska, S.A. (2004). DNA diversity of human populations from Eastern Europe and Siberia studied by multilocus DNA fingerprinting. *Mol Genet Genomics*, Vol.271, No.3, (Apr), pp. 291-297, issn 1617-4615
- Shabrova, E.V., Limborska, S.A. & Ryskov, A.P. (2006). Multilocus DNA fingerprinting-genotyping based on micro and minisatellite polymorphisms, In: *Focus on DNA fingerprinting research*, Read, M.M. Nova Publishers, ISBN 9781594549533, New York.
- Song, J., Yoon, Y., Park, K.U., Park, J., Hong, Y.J., Hong, S.H. & Kim, J.Q. (2003). Genotype-specific influence on nitric oxide synthase gene expression, protein concentrations, and enzyme activity in cultured human endothelial cells. *Clin Chem*, Vol.49, No.6 Pt 1, (Jun), pp. 847-852, issn 0009-9147
- Spitsyn, V.A., Khorste, M.V., Pogoda, T.V., Slominsky, P.A., Nurbaev, S.D., Agapova, R.K. & Limborska, S.A. (2000). Apolipoprotein B 3'-VNTR polymorphism in the Udmurt population. *Hum Hered*, Vol.50, No.4, (Jul-Aug), pp. 224-226, issn 0001-5652
- Stepanov, V.A., Spiridonova, M.G. & Puzyrev, V.P. (2003). Comparative phylogenetic study of native north Eurasian populations from a panel of autosomal microsatellite loci. *Genetika*, Vol.39, No.11, (Nov), pp. 1564-1572, issn 0016-6758
- Tanaka, T. (2005). International HapMap project. *Nihon Rinsho*, Vol.63 Suppl 12, (Dec), pp. 29-34, issn 0047-1852
- Vassart, G., Georges, M., Monsieur, R., Brocas, H., Lequarre, A.S. & Christophe, D. (1987). A sequence in M13 phage detects hypervariable minisatellites in human and animal DNA. *Science*, Vol.235, No.4789, (Feb 6), pp. 683-684, issn 0036-8075
- Verbenko, D.A., Pogoda, T.V., Spitsyn, V.A., Mikulich, A.I., Bets, L.V., Bebyakova, N.A., Ivanov, V.P., Abolmasov, N.N., Pocheshkhova, E.A., Balanovskaya, E.V., Tarskaya,

- L.A., Sorensen, M.V. & Limborska, S.A. (2003). Apolipoprotein B 3'-VNTR polymorphism in Eastern European populations. *Eur J Hum Genet*, Vol.11, No.6, (Jun), pp. 444-451, issn 1018-4813
- Verbenko, D.A., Kekeeva, T.V., Pogoda, T.V., Khusnutdinova, E.K., Mikulich, A.I., Kravchenko, S.A., Livshits, L.A., Bebyakova, N.A. & Limborska, S.A. (2003). Allele frequencies for D1S80 (pMCT118) locus in some East European populations. *J Forensic Sci*, Vol.48, No.1, (Jan), pp. 207-208, issn 0022-1198
- Verbenko, D.A., Pocheshkhova, E.A., Balanovskaya, E.V., Marshanija, E.Z., Kvitzinija, P.K. & Limborska, S.A. (2004). Polymorphisms of D1 S80 and 3'ApoB minisatellite loci in Northern Caucasus populations. *J Forensic Sci*, Vol.49, No.1, (Jan), pp. 178-180, issn 0022-1198
- Verbenko, D.A., Knjazev, A.N., Mikulich, A.I., Khusnutdinova, E.K., Bebyakova, N.A. & Limborska, S.A. (2005). Variability of the 3'APOB minisatellite locus in Eastern Slavonic populations. *Hum Hered*, Vol.60, No.1, pp. 10-18, issn 0001-5652
- Verbenko, D.A., Slominsky, P.A., Spitsyn, V.A., Bebyakova, N.A., Khusnutdinova, E.K., Mikulich, A.I., Tarskaia, L.A., Sorensen, M.V., Ivanov, V.P., Bets, L.V. & Limborska, S.A. (2006). Polymorphisms at locus D1S80 and other hypervariable regions in the analysis of Eastern European ethnic group relationships. *Ann Hum Biol*, Vol.33, No.5-6, (Sep-Dec), pp. 570-584, issn 0301-4460
- Verbenko, D.A., Limborskaia, S.A. (2008). Human minisatellite markers: D1S80 locus in population studies. *Mol Gen Mikrobiol Virusol*, No. 2, pp. 3-11.
- Vergnaud, G. & Denoeud, F. (2000). Minisatellites: mutability and genome architecture. *Genome Res*, Vol.10, No.7, (Jul), pp. 899-907, issn 1088-9051
- Verstrepen, K.J., Jansen, A., Lewitter, F. & Fink, G.R. (2005). Intragenic tandem repeats generate functional variability. *Nat Genet*, Vol.37, No.9, (Sep), pp. 986-990, issn 1061-4036
- Walsh, S.J. & Eckhoff, C. (2007). Australian Aboriginal population genetics at the D1S80 VNTR locus. *Ann Hum Biol*, Vol.34, No.5, (Sep-Oct), pp. 557-565, issn 0301-4460
- Watanabe, G. & Shimizu, K. (2002). DNA sequence analysis of long PCR amplified products at the D1S80 locus. *Leg Med (Tokyo)*, Vol.4, No.1, (Mar), pp. 37-39, issn 1344-6223
- Weir, B.S. (1996). *Genetic data analysis* (1st), Sinauer Associates, Inc, ISBN 0878939075, Sunderland
- Weller, P., Jeffreys, A.J., Wilson, V. & Blanchetot, A. (1984). Organization of the human myoglobin gene. *EMBO J*, Vol.3, No.2, (Feb), pp. 439-446, issn 0261-4189
- Wyman, A.R. & White, R. (1980). A highly polymorphic locus in human DNA. *Proc Natl Acad Sci U S A*, Vol.77, No.11, (Nov), pp. 6754-6758, issn 0027-8424
- Yeh, FC, Yang, R-C & Boyle, T. (1999) *POPGENE version 1.32*, Available from: <http://www.ualberta.ca/~fyeh/>
- Zago, M.A., Silva Junior, W.A., Tavella, M.H., Santos, S.E., Guerreiro, J.F. & Figueiredo, M.S. (1996). Interpopulational and intrapopulational genetic diversity of Amerindians as revealed by six variable number of tandem repeats. *Hum Hered*, Vol.46, No.5, (Sep-Oct), pp. 274-289, issn 0001-5652
- Zhang, M.X., Ou, H., Shen, Y.H., Wang, J., Coselli, J. & Wang, X.L. (2005). Regulation of endothelial nitric oxide synthase by small RNA. *Proc Natl Acad Sci U S A*, Vol.102, No.47, (Nov 22), pp. 16967-16972, issn 0027-8424
- Zhivotovskii, L.A. (2006). Population aspects of forensic genetics. *Genetika*, Vol.42, No.10, (Oct), pp. 1426-1436, issn 0016-6758

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen