

Limborska S.A.<sup>1</sup>  
Slominsky P.A.<sup>1</sup>  
Nurbaev S.D.<sup>2</sup>  
Balanovskaya E.V.<sup>2</sup>  
Shadrina M.I.<sup>1</sup>  
Popova S.N.<sup>1</sup>  
Belyaeva O.V.<sup>1</sup>  
Pogoda T.V.<sup>1</sup>  
Balanovsky O.P.<sup>1</sup>  
Verbenko D.A.<sup>1</sup>  
Bulaeva K.B.<sup>3</sup>  
Khusnutdinova E.K.<sup>4</sup>  
Spitsyn V.A.<sup>2</sup>  
Mikulich A.I.<sup>5</sup>

## East European Lowland as an Area of Long-time Interaction Between Caucasoid and Mongoloid Peoples.

Most of the population of Eastern Europe inhabit an area of great anthropological interest, because of the contact between Caucasoid and Mongoloid anthropological types. We have analyzed normal variability in minisatellite and microsatellite loci in some East European population. Different synthetic maps were constructed using reliability theory to evaluate the degree of accuracy. Comparison of the synthetic maps for DNA with classical markers has revealed a high level of correlation. All the data obtained show the diverse influence of both anthropological types in forming the gene pool of the Eastern European peoples.

*1 Institute of Molecular Genetics,  
Russian Academy of Sciences,  
123182 Moscow, Russia;*

*2 Medical Genetics Scientific Center,  
107012 Moscow, Russia;*

*3 Institute of General Genetics,  
Russian Academy of Sciences,  
117809 Moscow, Russia;*

*4 Bashkir Science Center of Ural  
Branch of Russian Academy of  
Sciences, 450054 Ufa, Russia;*

*5 Institute of Ethnography, 220015  
Minsk, Byelarus*

*Keywords: East Europe, Caucasoid, Mongoloid,  
Ethnogenesis, DNA diversity.*

Most of the populations of Eastern Europe inhabit an area of great anthropological interest, because of the contact between Caucasoid and Mongoloid anthropological types. This makes it possible to analyse the interaction of both components during ethnogenesis of European peoples. Outbreeding does not always lead to a simple mixing of anthropologically contrasting groups, but gene recombination can lead to unusual new combinations of traits.

Comparisons with nearest neighbours and distant populations permits the evaluation of the role played by Caucasoid and Mongoloid components in the origin and ethnogenesis of populations studied. We have analysed normal variability in minisatellite

and microsatellite loci in some East European populations. There are Caucasoid populations (Byelorussians, Russians, Moldavians, Dagestan peoples) and also admixture populations with a different level of Mongoloid components (Mordvinians, Udmurts, Komi, Mari, Chuvashes, Tatars, Bashkirs, and Nogai people).

Different synthetic maps were constructed using a reliability theory to evaluate the degree of accuracy. Comparison of the synthetic maps for DNA with classical markers has revealed a high level of correlation. All the data obtained show the diverse influence of both anthropological types in forming the gene pool of the Eastern European peoples.

DNA samples from each population studied were tested for polymorphisms by PCR analysis of APOB minisatellite (Deka et al. 1992, Khushutdinova et al. 1995, Pogoda et al. 1995, Spitsyn et al. 1997), DM microsatellite (Slominsky et al. 2000, Zerylnick et al. 1995), CAn 685 microsatellite (Belyaeva et al. 1997, 1998) and CCR5 insertion/deletion polymorphisms (Galeeva et al. 1998, Martinson 1997, Slominsky et al. 1997). Figure 1 shows the map of ethnic areas in East European Lowland (Russian Plain). One can see a lot of peoples living there. We designate only those peoples, that were analysed in this study.

Figure 2 shows the linguistic classification of populations studied. There are 28 populations from peoples, that belong to four different linguistic families. First of all we have calculated the genetic distances. But the question arose what should be the reference population, that is, from which population should the genetic distances be calculated. We decided to evaluate genetic distances from total mean allele frequencies. Mean gene frequencies for each people were calculated, then for representatives of each linguistic family, and finally for all four families (that is total mean frequencies). So genetic distances were calculated from these total mean frequencies. These data are shown in Figure 3 as a map of genetic distances. The white colour designates minimal genetic distances, the black colour, maximal distances. One can see that the total mean frequencies are in the Southern part of East European Lowland, namely between the rivers of the Don and the Volga. Genetically closer to them are the populations inhabiting the wide region between Lake Onega on the one hand and Caucasus foothills and the Caspian Sea on the other hand. More genetically distant are populations from South Urals, North-European part of Russia and the western part of East European Lowland.

To construct this map we have used data on 28 populations studied. Data for the most part of the area have been obtained by interpolation. The question arises, how reliable is this extrapolation? To answer this question, we have used the map of reliability for DNA markers (fig.4). The map was constructed on the basis of the reliability theory used in engineering. It shows the estimated reliability in each point of space (Nurbaev & Balanovskaya 1998).

The map of reliability of the genetic information allows one to sort the data for reliability at any degree of accuracy. The black colour designates the area with the probability of results being above 99%. The more populations analysed in this area, the higher the level of probability. The grey colour shows the area with over 95% probability of data shown. At the 95% level of probability we regard our data as reliable, all other areas were omitted from consideration, as one can see on the Figure 3.

Then we have analysed synthetic maps of principal components. Figure 5 shows the mapping of the first principal component in East European Lowland. These results, too, are reliable at 95% or higher. You can see two extremes: one in the North-East, another in the South-West of the Lowland. This is amazing, but the same results were revealed (Balanovskaya et al. , 1994; Nurbaev & Balanovskaya, 1997) with the use of classical markers. 100 alleles of 34 loci have been analysed and revealed almost the same extremes. Comparison of the first principal component maps for DNA and classical markers shows that correlation between the two maps is 0.834.

Apparently, the first principal component in both cases shows the interaction of Caucasoid and Northern Mongoloid peoples beginning from a very early period.

Cavalli-Sforza (1998) has analysed the first five principal components of classical polymorphisms over entire Europe. Interestingly, our results on the first principal component coincide with Cavalli-Sforza's data on the second principal component. He treats his data as a result of the spread of Uralic language speakers from the region around the Northern Urals at some very early period.

Figure 6 shows the second principal component of DNA polymorphisms of our four loci. Here, too, one can see two extremes, but their positions are different: one is in the West, another in the South-East. There is also a high correlation with classical markers. The correlation coefficient is 0.572.

The second principal component in both cases apparently shows the interaction of Caucasoid and Southern Mongoloid peoples from the early times.

Also, there is a resemblance with Cavalli-Sforza's map for all of Europe but with his third principal component map. He regards it as being a result of the spread of pastoral nomads (in the Bronze Age, about 5000 years ago).

Figure 7 shows the map of third principal component of DNA polymorphism of our four loci. One can see four extremes. The same picture emerges in the case of classical markers (Nurbaev, Balanovskaya 1997). The correlation coefficient is 0.485. The third principal component shows a complex genetic landscape. On these maps, the gene pool of the population of East European Lowland is a valley in which opposite influences are at equilibrium.

Figure 8 shows the space of the principal components. These results were derived from the geographic space to the principal component space due to Nurbaev and

Balanovskaya (1997). Six peoples were taken as an example. They are Byelorussians, Russians, Daghestan peoples, Mordva, Mari, Bashkirs. All these populations are placed in the space of first and the second principal components. The ethnic "centroids" are also positioned in the same space. The number of points is directly proportional to the ethnic area. The large Russian area is represented by a large number of points, smaller areas of other ethnic groups, by a smaller number of points. Each ethnic group forms an "ethnic cloud". Some of ethnic clouds overlap, which can be explained by their close genetic relationship.

In Figure 9 one can see GST and heterozygosity values for DNA markers and classical markers. Interestingly, GST levels are the same in both cases.

All data, then, show the diverse effects of various ethnic factors on the gene pool of the East European peoples. Synthetic maps, especially the first two, apparently reveal intensive contacts between Mongoloid and Caucasoid populations in the East European Lowland over several millennia from a very early period. It can be assumed that Uralic languages originated within a group of prehistoric people who had settled not far from the Arctic region near the boundary between Europe and Asia. This group may have been genetically related to the Mongoloids, especially east of the Urals. The spread of the Uralic speakers to the west and southwest into areas inhabited by Caucasoid popu-

**LINGUISTIC CLASSIFICATION OF PEOPLES STUDIED**

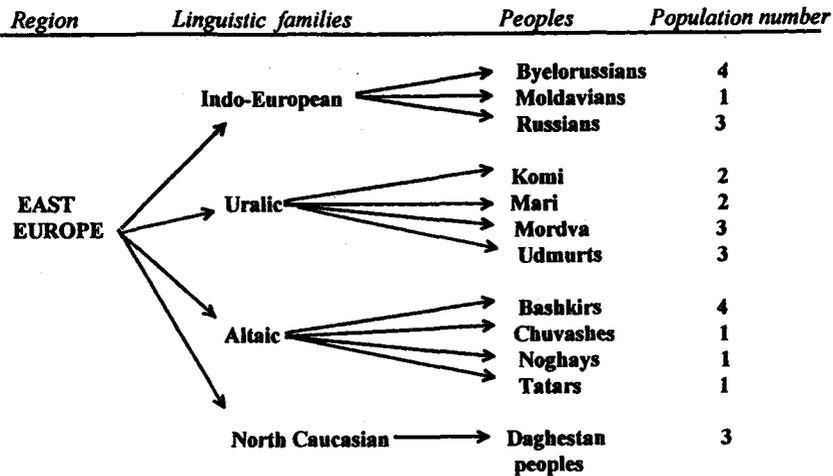


Figure 2. Linguistic classification of peoples studied.

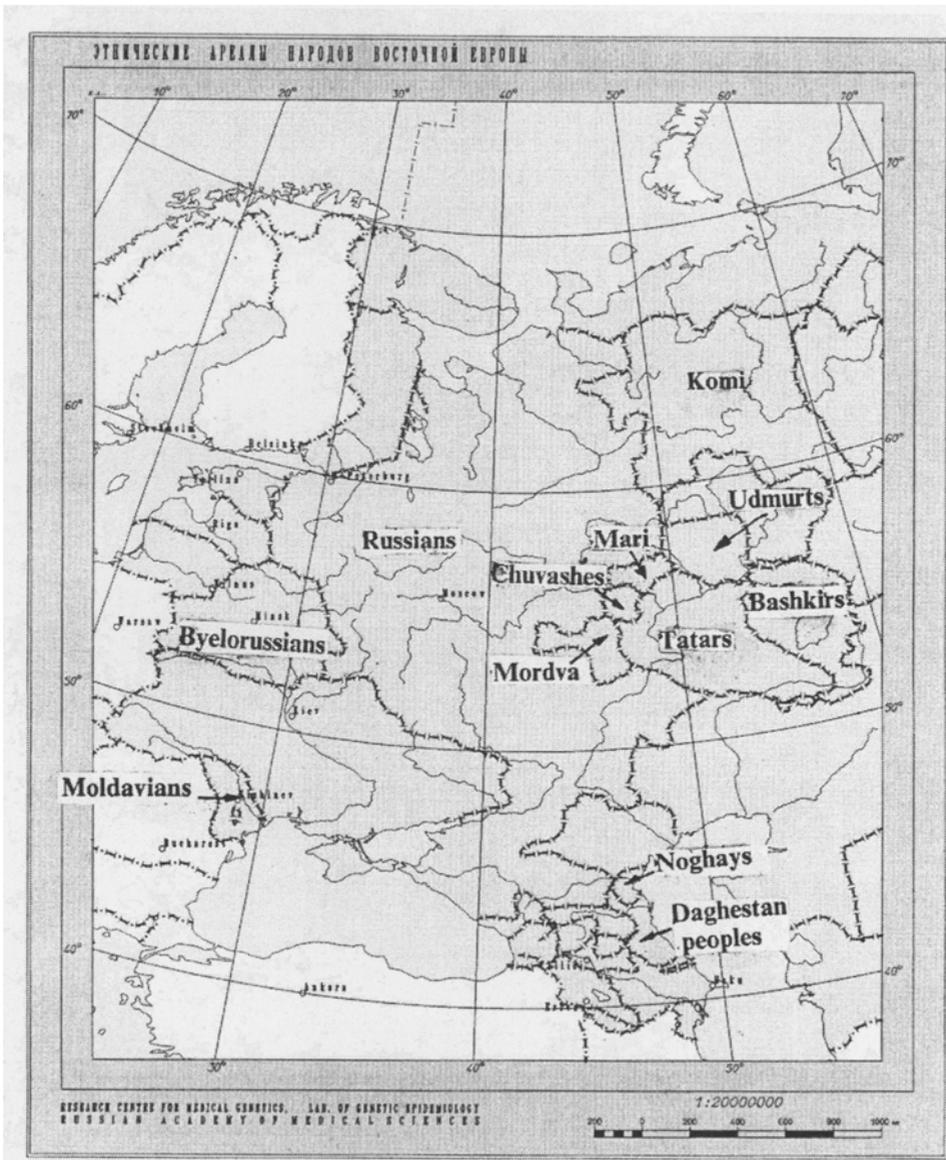


Figure 1. Map of ethnic areas in East European Lowland (Russian Plain). Only peoples analysed in this study are designated.

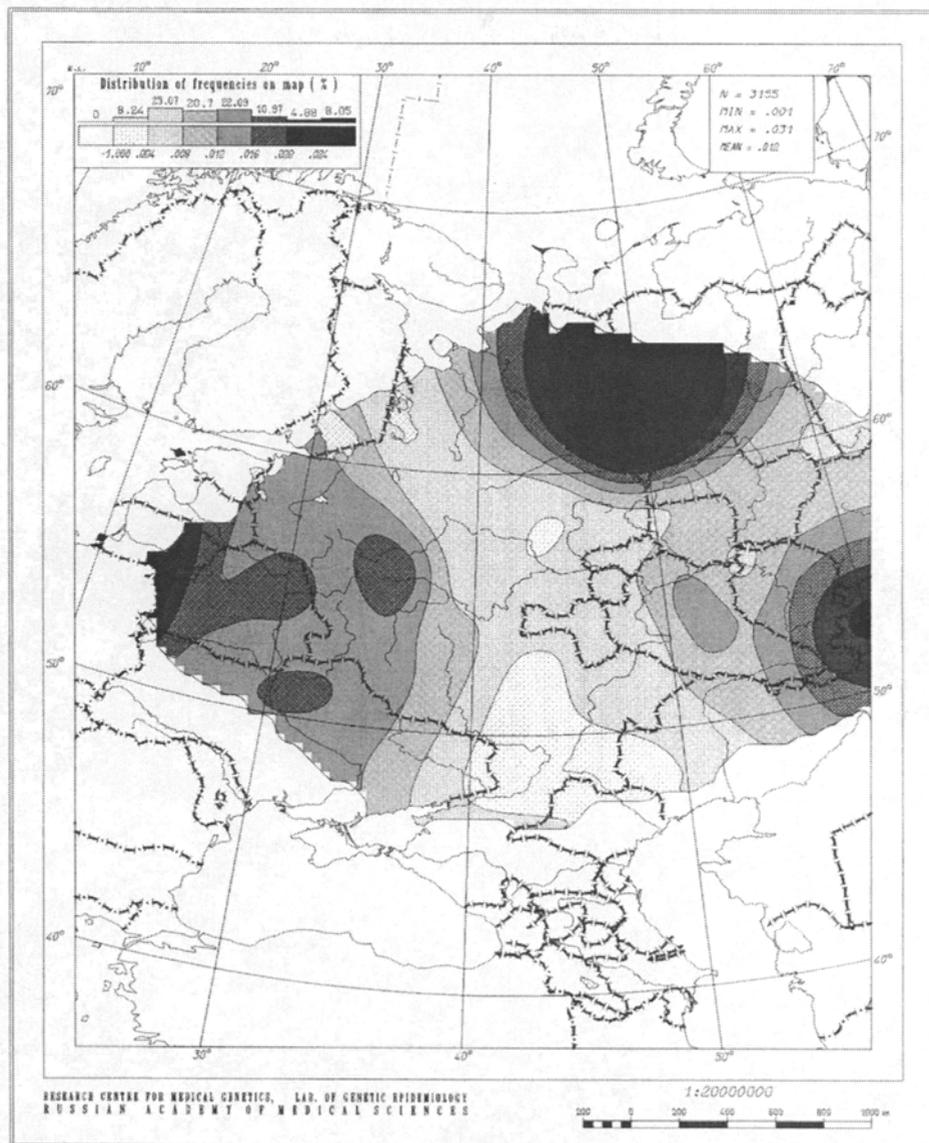


Figure 3. Map of genetic distances. The areas with probabilities of statistically significant data less 95% have excluded from consideration (see the text and legend to Fig. 4).

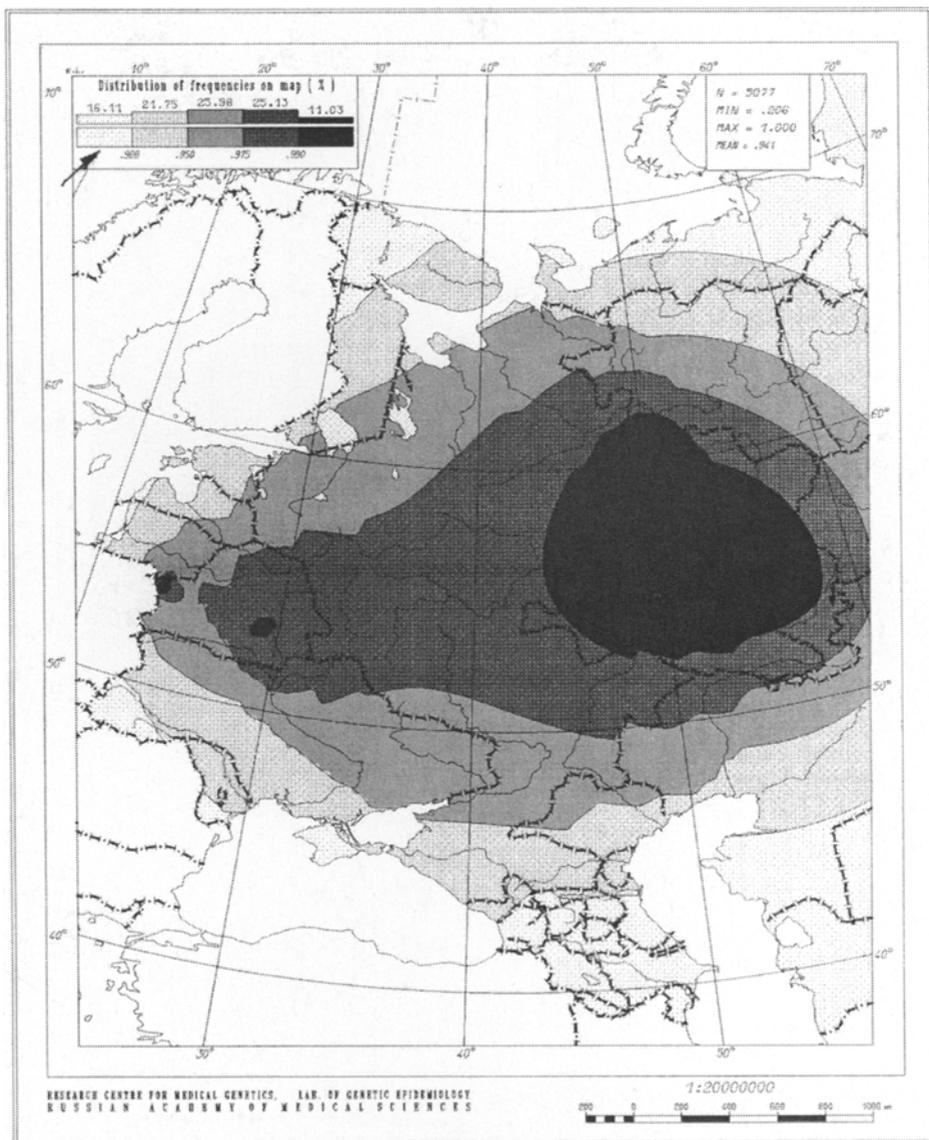


Figure 4. Map of reliability for DNA marker results. The scale (designated by arrow) indicates the probabilities of statistically significant estimations at the given levels of reliability. The higher the reliability, the darker the shade of the corresponding area in the map.

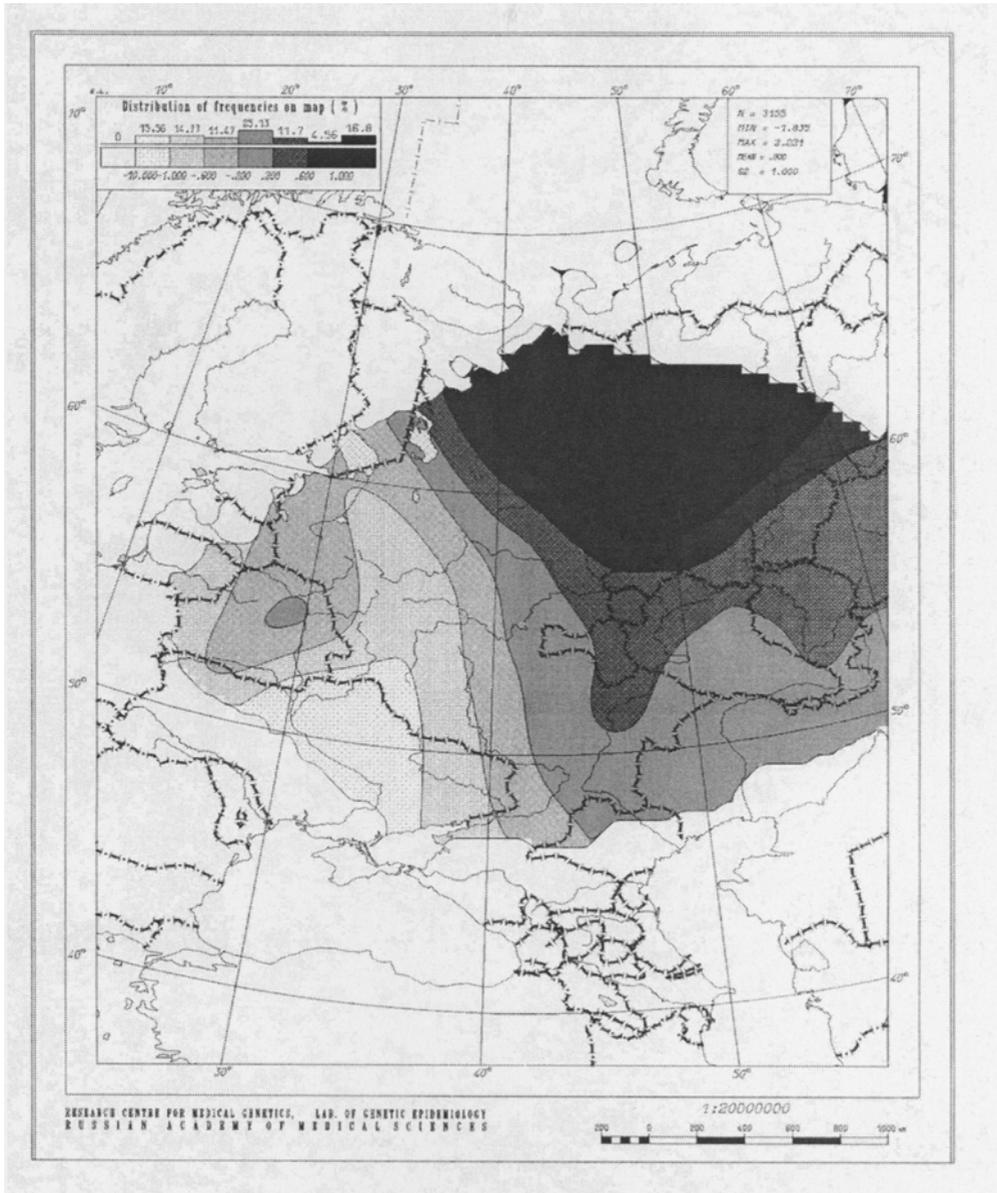


Figure 5. Map of the first principal component of the gene pool of East European population. The first principal component accounts for about 32% of the total genetic variation.

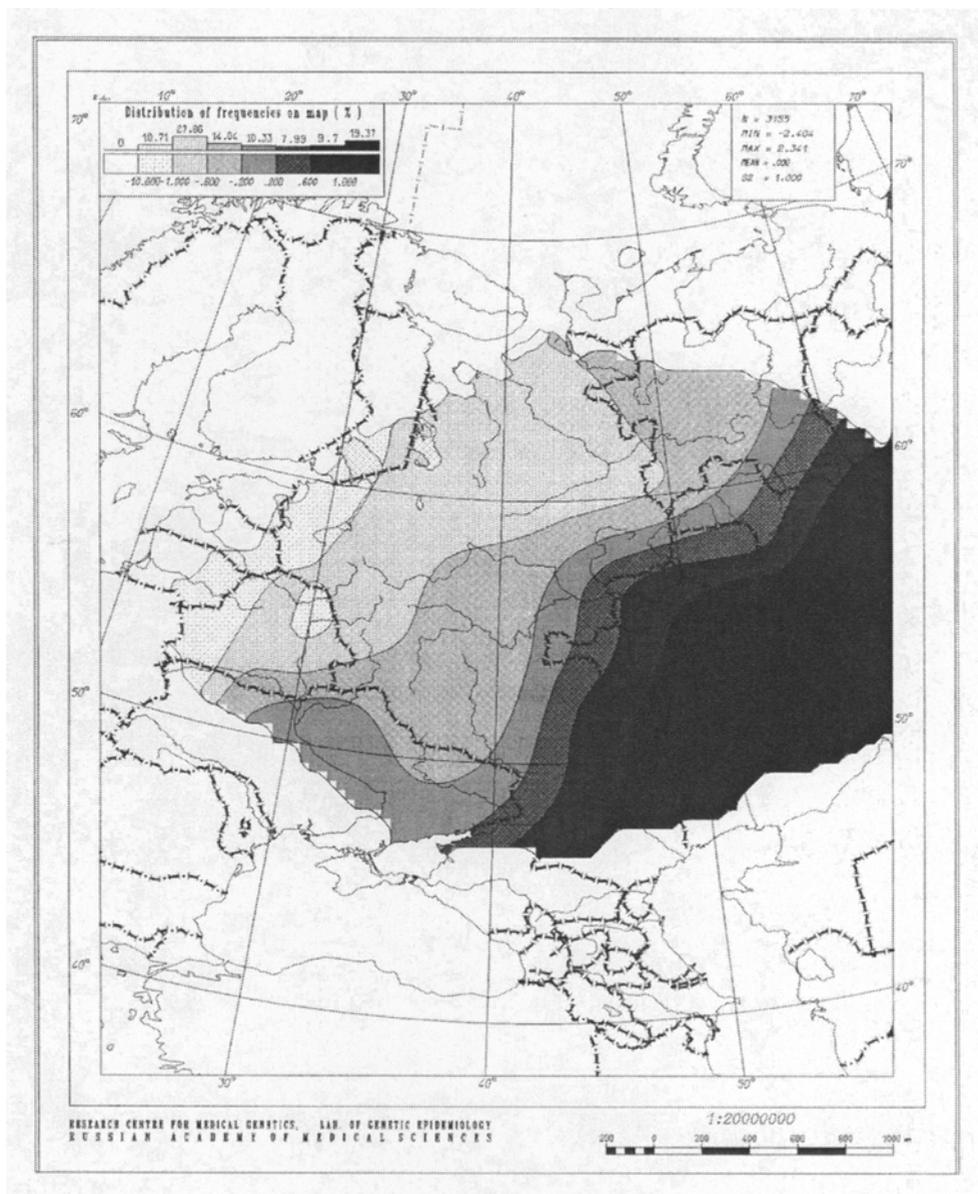


Figure 6. Map of the second principal component of the gene pool of East European population. The second principal component accounts for about 28% of the total genetic variation.

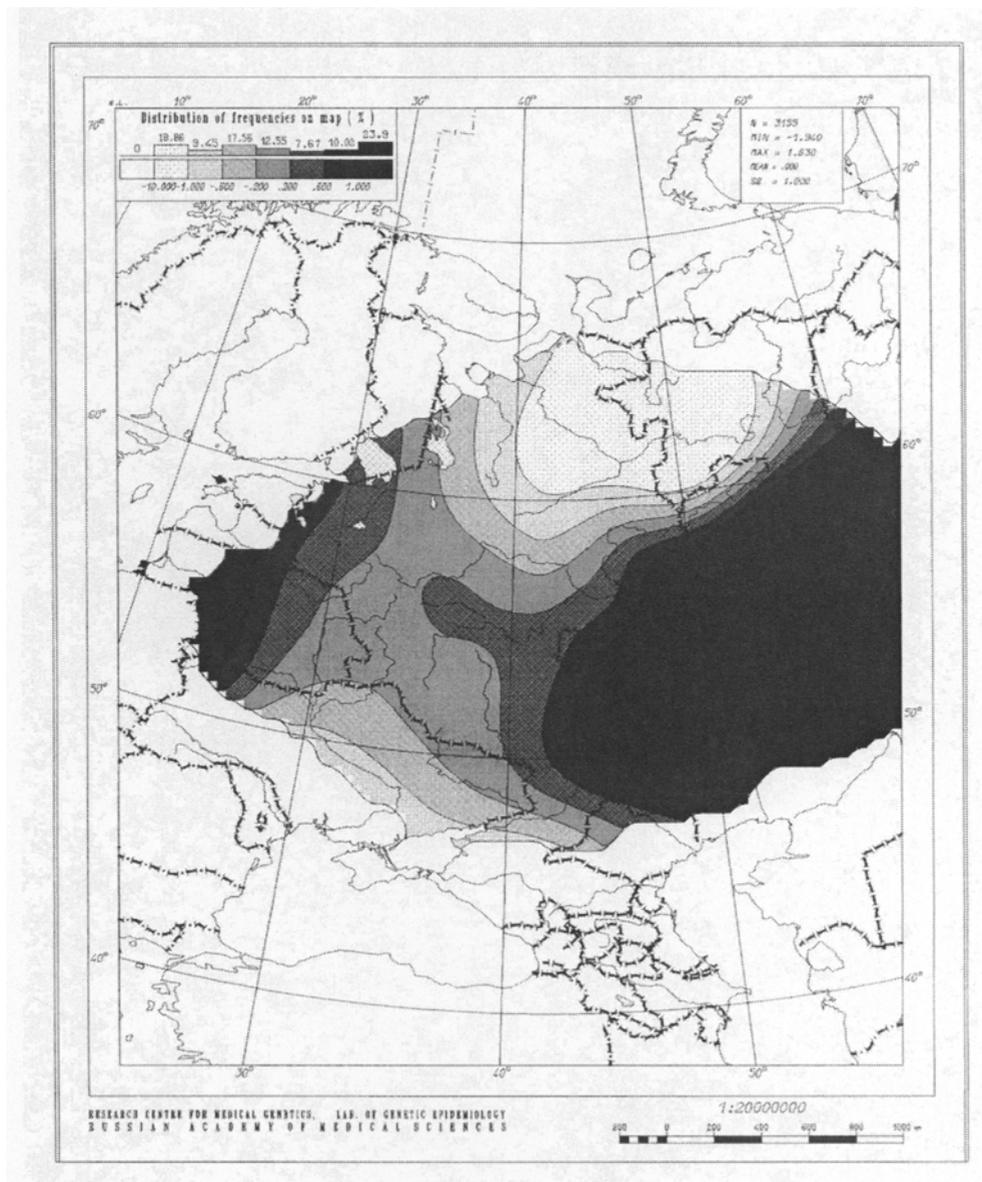


Figure 7. Map of the third principal component of the gene pool of East European population. The third principal component accounts for about 8,5% of the total genetic variation.

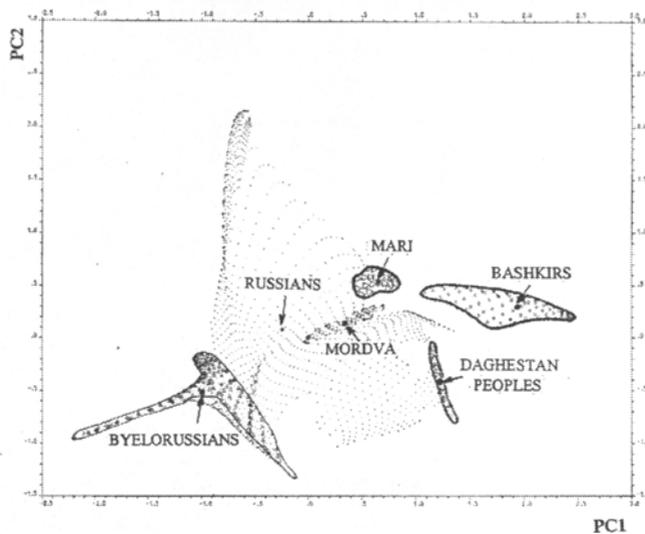


Figure 8. Space of the first two principal components (data from the synthetic maps).

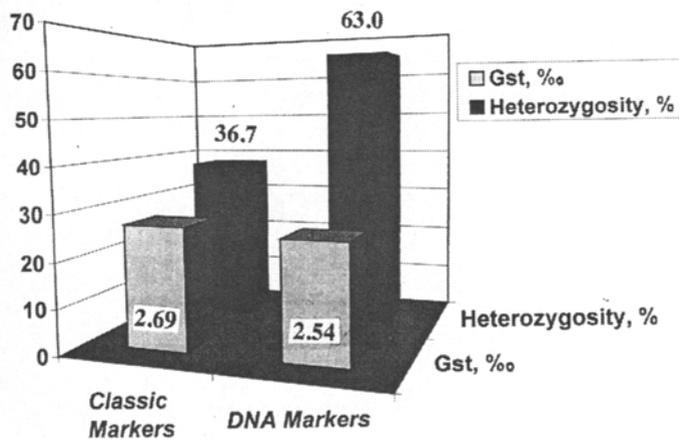


Figure 9. Comparison of GST and heterozygosity for DNA and classical markers.

lations resulted in a variably intense hybridisation in several regions. Another important migration event revealed by the synthetic map was the spread of pastoral nomads who originally inhabited the Lower Don-Volga interfluves, the area north of the Caucasus and the Black Sea. The third synthetic map provides indications of other possible expansions from various regions of the East European Lowland. Further studies using a larger number of DNA markers will doubtlessly reveal a more detailed picture of population relationships from the early periods up to the present.

## References.

- Balanovskaya, E.V., Nurbaev, S.D., Rychkov, Yu.G. (1994). Computer technology of gene-geographic investigation of gene pool. *Genetika, Russia*, 30, 1538-55.
- Balanovskaya, E.V., Nurbaev, S.D. (1997). A computer- rided method for gene-geographic study of a population gene pool in the space of principal components. *Russian Journal of Genetics, Russia*, 33,1703-18.
- Belyaeva, O.V., Shabrova, E.V., Prosnjak, M.I., Spitsyn, E.A., Khushnutdinova, E.K., Mikulich, A.I., Limborska, S.A. (1997). New highly polymorphic marker (CA)<sub>n</sub> from DNA library of human chromosome 19. *Molecular Biology, Russia*, 31, 795-800.
- Belyaeva, O.V., Balanovsky, O.P., Ashworth, L.K., Lebedev, Yu.B., Spitsyn, V.A., Khushnutdinova, E.K., Mikulich, A.I., Guseva, I.A., Erdes, Sh., Limborska, S.A. (1998). Fine mapping of a polymorphic CA repeat marker on human chromosome 19 and its use in population studies. *American Journal of Human Genetics, USA*, 63.
- Cavalli-Sforza, L.L. (1998). DNA revolution in population genetics. *Trends in Genetics* 14, 60-65.
- Deka, R., Chakraborty, R., DeCroo, S., Rothhammer, F., Barton, S.A., Ferrell, R.E. (1992). Characteristic of polymorphism at a VNTR locus 3' to the ApoB gene in five human populations. *American Journal of Human Genetics, USA*, 51, 1325-33.
- Galeeva, A.R., Khushnutdinova, E.K., Slominsky, P.A., Limborska, S.A. (1998). Spread of CCR5 gene 32-basepair deletion in population of Volga-Ural Region. *Russian Journal of Genetics, Russia*, 34, 1160-62.
- Khushnutdinova, E.K., Pogoda, T.V., Prosnjak, M.I., Khidiyatova, I.M., Raphicov, K.C., Limborska, S.A. (1995). Allele frequencies of VNTR ApoB locus in Bashkir and Komi populations. *Russian Journal of Genetics, Russia*, 31, 995-1000.
- Martinson, J.J, Chapman, N.H., Rees, D.C., Liu, U.T., Clegg, J.B. (1997). Global distribution of the CCR5 gene 32-basepair deletion. *Nature Genetics, Great Britain*, 16, 100-103.
- Nurbaev, S.D., Balanovskaya, E.V. (1998). A computer- rided method for gene-geographic study of a population gene pool. *Russian Journal of Genetics, Russia*, 34, 825-38.
- Pogoda, T.V., Nikonova, A.L., Kolosova, T.V., Lyudvikova, E.K., Perova, N.V., Limborska, S.A. (1996). Allelic variants of the genes of ApoB and ApoC<sub>2</sub> in patients with coronary heart disease and in healthy individuals from the Moscow population. *Russian Journal of Genetics, Russia*, 31, 1001-09.

- Slominsky, P.A., Schadrina, M.I., Spitsyn, V.A., Khushnutdinova, E.K., Mikulich, A.I., Limborska, S.A. (1997). Simple and rapid method for deletion of CCR5 chemokine receptor gene deletion. Russian Journal of Genetics, Russia, 33, 1596-98.
- Slominsky, P.A., Popova, S.N., Schadrina, M.I., Pomazanova, M.A., Lomova, T.Yu, Khushnutdinova, E.K., Guseva, I.A., Erdes, Sh., Bulaeva, K.B., Mikulich, A.I., Spitsyn, V.A., Limborska, S.A (2000). Russian Journal of Genetics, Russia, 36,980-985.
- Spitsyn, V.A., Knorte, M.V., Pogoda, T.V., Schadrina, M.I., Slominsky, P.A., Limborska, S.A. (1997). ApoB VNTR polymorphism in Udmurt populations. Russian Journal of Genetics, Russia, 33, 278-82.
- Zerylnick, M., Torroni, A., Sherman, S.L., Warren, S.T. (1995) Normal variation at the Myotonic dystrophy locus in global human populations. American Journal of Human Genetics, USA, 56, 123-130.

*Received: June 10, 2003      Accepted: December 20, 2003*