

# The Gene Pool of the Belgorod Oblast Population: Changes in Population Genetic Relationships during the Past 50 Years

I. N. Sorokina<sup>a</sup>, M. I. Churnosov<sup>a</sup>, and E. V. Balanovska<sup>b</sup>

<sup>a</sup> Belgorod State University, Belgorod, 308015 Russia;  
e-mail: sorokina@bsu.edu.ru

<sup>b</sup> Medical Genetic Research Center, Russian Academy of Medical Sciences, Moscow, 115478 Russia

**Abstract**—The model of the Belgorod oblast population has been used to demonstrate different effects of administrative reforms on microevolution in human populations. For the populations that formerly belonged to Kursk oblast, changes in the regional administrative structure (after which some of them remained in Kursk oblast and others were included into Belgorod oblast) have led to an increase in the genetic distances between them. However, other populations (formerly belonging to Voronezh oblast) have become genetically closer to one another, although these populations now belong to different administrative regions (Belgorod and Voronezh oblasts).

## INTRODUCTION

Estimation of genetic distance between populations is an important part of population genetic studies. Modern population genetics solve these tasks with a sufficient efficiency using various types of traditional markers, including quasi-genetic, immune–biochemical, and molecular genetic ones [1–3]. In addition, a number of novel approaches to estimation of genetic relationships between populations with the use of population demographic markers (vital statistics and marriage migration characteristics) have been suggested in recent years [4–6]. These approaches are implemented in population genetic studies for estimating the differential fitness of populations to the environment and comparative analysis of populations because, in some cases (e.g., studying the changes in genetic relationships between populations during several generations), traditional genetics (classical and DNA-based) and quasi-genetic (surnames) markers can hardly be used, for various reasons [4]. El'chinova et al. [4–6] have developed a metric for genetic distances between populations that is based on the parameters of Malecot's isolation-by-distance model calculated by the standard method. This metric can be used in different variants of multivariate statistics, including cluster analysis, the resultant dendrogram being used for drawing the genetic landscape, as it has been demonstrated using populations of Marii El Republic as a model [6].

We used the distance metric calculated from the parameters of Malecot's isolation-by-distance model to study the changes in the genetic distances between populations of the Central Chernozem region of Russia during the past 50 years.

## MATERIALS AND METHODS

We used the distance metric calculated from the parameters of Malecot's isolation-by-distance model for taxonomic analysis of Central Chernozem populations and their genetic changes during the past 50 years. This time interval was chosen because Belgorod oblast is an "administratively young" region: it was formed in 1954 out of the adjacent raions (administrative districts) of Kursk and Voronezh oblasts. The effect of the changes in the regional administrative structure of central Russia on genetic relationships between the populations involved in them has not been studied until the present time. We studied the populations of two pairs of neighboring districts: (1) Pristen' raion (an administrative district of Kursk oblast) and Prokhorovka raion (a district of Kursk oblast before 1954 and a district of Belgorod oblast afterwards) and (2) Rep'evka raion (a district of Voronezh oblast) and Krasnoe raion (a district of Voronezh oblast before 1954 and a district of Belgorod oblast afterwards) (Fig. 1).

Prokhorovka and Krasnoe raions of Belgorod oblast border on the neighboring Kursk and Voronezh oblasts, respectively, and are located in different parts of the administrative region: Prokhorovka raion, in northern Belgorod oblast; and Krasnoe raion, in northeastern Belgorod oblast. Regarding their administrative subdivision, Prokhorovka raion comprises 19 rural municipalities with a total population of 27800 people, and Krasnoe raion comprises 10 rural municipalities with a total population of 15500 people. Pristen' raion of Kursk oblast comprises 18 rural municipalities with a total population of 10100 people, and Rep'evka raion of Voronezh oblast comprises 11 rural municipalities with a total population of 10100 people.



**Fig. 1.** The map of the rural districts of Belgorod oblast. The analyzed districts are hatched (Prokhorovka and Krasnoe raions of Belgorod oblast, Pristen' raion of Kursk oblast, and Rep'evka raion of Voronezh oblast). The thick line shows the border between Kursk and Voronezh oblasts before 1954, the year when Belgorod oblast was established.

For this study, we selected five “model” municipalities in each district. Detailed analysis of the changes in the marriage migration structure of the analyzed populations by means of Malecot’s isolation-by-distance model was performed in our previous study [7].

We used the method developed by El’chinova [6, 8] to calculate distance matrices (inbred distances) on the basis of Malecot’s isolation-by-distance model. The distance between two populations was taken to be the sum of two elements, the first of which was the difference between Malecot’s local inbreeding ( $a$ ) calculated in the first population and the inbreeding level calculated by Malecot’s formula  $\Phi(d) = ae^{-bd}$  at a distance of  $d$  from the second population, i.e., the inbreeding drop. The second summand was a similar difference between Malecot’s local inbreeding in the second population and the inbreeding calculated by Malecot’s formula at the same distance  $d$  from the second population to the first one but with the parameters of Malecot’s isolation-by-distance model for the second population. The resultant formula was the following:

$$D_{1,2} = [a_1 - a_1 e^{-b_1 d}] + [a_2 - a_2 e^{-b_2 d}],$$

where  $a_i$  and  $b_i$  are the parameters of Malecot’s isolation-by-distance model in the  $i$ th population,  $i = 1, 2$ . Obviously, the longer the geographic distance ( $d$ ) between the populations, the greater the inbreeding drop and, hence, the larger the sum of the two terms [6].

This metric has an obvious advantage of taking into account the effective population sizes and the physical distance between them, the factors that substantially affect the genetic demographic structure of populations and migration characteristics [6]. A drawback of this metric is that its value is substantially affected by the geographic distances between the populations studied, which are included in the equation for the metric. It is a priori evident that, the more geographically remote from each other the compared populations are, the longer will be the genetic distance between them (Malecot’s metric). However, it is also obvious that the genetic distances between populations (calculated, e.g., on the basis of immune–biochemical or molecular genetic markers) will not so “strictly” depend on the geographic distance, because, as demonstrated in a number of studies [8–10], factors other than geographic distances, such as the history of the populations, their positions in the regional administrative structure, and the quality of road infrastructure, also affect genetic distances.

The distance matrices calculated from the parameters of Malecot’s isolation-by-distance model were treated using cluster analysis followed by construction of the dendrograms characterizing the genetic relationships between the populations of rural municipalities. These dendrograms served as the basis for genetic landscape diagrams [6] of the populations studied, which were projections of the dendrogram of the matrix of

distances calculated from the parameters of Malecot's isolation-by-distance model on a geographical map. In these diagrams, elementary populations belonging to the same clusters were grouped to form equidistant figures. Equidistant lines were drawn through a specific genetic distance unit in such a way that the number of grouping levels was no greater than 10, because a larger number of levels makes the diagram too cumbersome for efficient analysis [6].

To estimate the role of geographic distances in subdivision of the population genetic structure in each time period studied, we performed correlation analysis (using Spearman's rank-order correlation coefficient) of distance matrices calculated from Malecot's isolation-by-distance parameters and geographic distances. The geographic distances between the rural municipalities studied were calculated using a map of Belgorod oblast (scale, 1 : 400000).

## RESULTS AND DISCUSSION

We used Malecot's isolation-by-distance model to calculate the distance matrices for the analyzed rural municipalities of the pairs of districts (Krasnoe and Rep'evka raions and Prokhorovka and Pristen' raions) in two periods: the time when Belgorod oblast was established (the 1950s) and recent time (the 1990s, when the current state is assumed to have been reached). The resultant matrices were used for cluster analysis followed by construction of dendrograms (Figs. 2a, 2b, 4a, 4b) and cartographic description of the inbreeding landscape of the populations studied (Figs. 3a, 3b, 5a, 5b).

Figure 2a shows the dendrogram of the distances calculated from the parameters of Malecot's isolation-by-distance model characterizing the genetic relationships between rural municipalities of Rep'evka and Krasnoe raions of Voronezh oblast in the 1950s. As evident from the results of cluster analysis, the Butyrki and Rep'evka municipalities of Rep'evka raion and the Gorki municipality of Krasnoe raion form a single cluster. Another cluster comprised only rural municipalities of Krasnoe raion (Raskhovets, Krasnoe, Gotov'e, and Kamyzino) grouped around the district center.

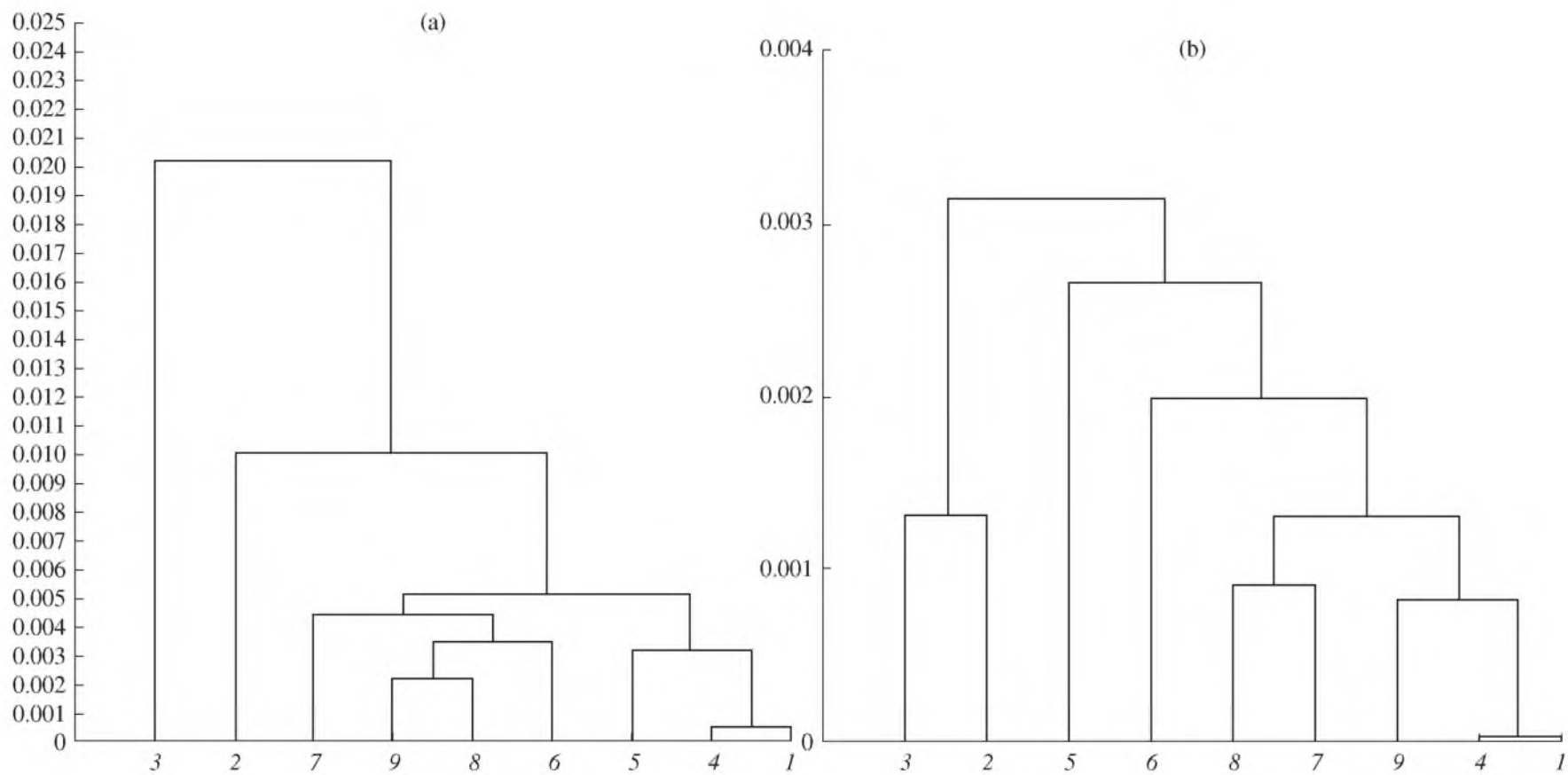
The spatial presentation of the results of cluster analysis (Fig. 3a) also indicate subdivision (separation of the elementary populations studied). The equidistant lines have been drawn without intersection at intervals of 0.001 units of distance calculated from the parameters of Malecot's isolation-by-distance model. Most rural municipalities of Krasnoe raions (four out of five municipalities studied) form a single equidistant figure. The Rep'evka and Butyrki rural municipalities of Rep'evka raion, together with the Gorki municipality of Krasnoe raion, form another equidistant figure. Apparently, the Gorki municipality, which belongs to Krasnoe raion, has been grouped together with rural municipalities of another (Rep'evka) raion because it is

geographically closer to these municipalities of Rep'evka raion than to municipalities of Krasnoe raion where it is located. For example, the distances from the Gorki to the Krasnoe and Raskhovets villages are 13 and 25 km, respectively, and the distances from it to the Rep'evka and Butyrki villages are 8 and 11 km, respectively. The importance of geographic distances between the analyzed rural municipalities of Krasnoe and Rep'evka raions for the population genetic subdivision is confirmed by the results of correlation analysis. Spearman's correlation coefficient between the distance matrix based on the parameters of Malecot's isolation-by-distance model and the matrix of geographic distances is 0.69 ( $p < 0.001$ ).

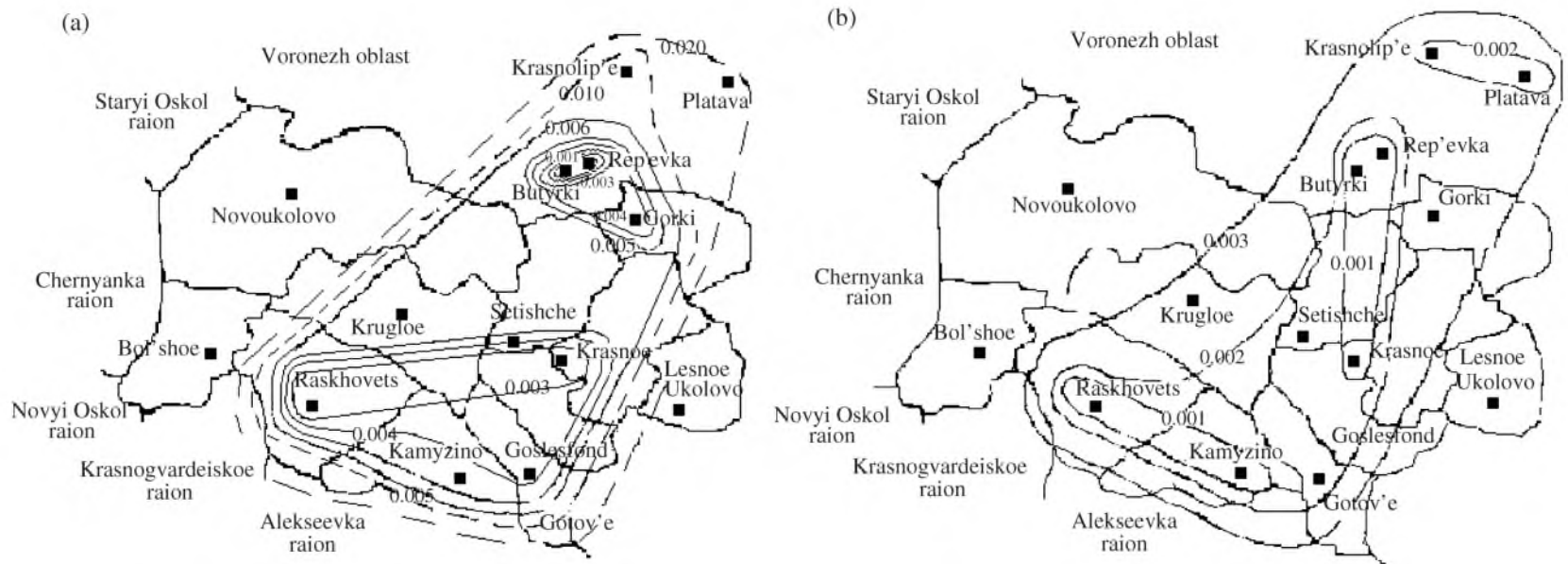
Fifty years have passed since one of the districts (Krasnoe raion) was included into Belgorod oblast, and the relationships between the analyzed rural municipalities estimated from the parameters of Malecot's isolation-by-distance model (Fig. 2b) have considerably changed. First, the absolute distances between the analyzed rural municipalities have decreased (in the 1950s, these municipalities were clustered at a level of 0.010–0.020; by the 1990s, this level decreased to 0.002–0.003). Second, the differentiation between rural municipalities belonging to different districts of two administrative regions has significantly decreased. As can be seen in Fig. 3, the two district administrative centers, Krasnoe of Belgorod oblast and Rep'evka (together with the Butyrki municipality) of Voronezh oblast are the first to be grouped together. Ultimately, six out of nine analyzed rural municipalities, namely, the Butyrki and Rep'evka municipality of Rep'evka raion of Voronezh oblast and the Krasnoe, Kamyzino, Raskhovets, and Gotov'e municipalities of Krasnoe raion of Belgorod oblast, form a single cluster at a level of 0.002.

The inbreeding landscape of the analyzed populations in the 1990s is shown in Fig. 3b. As in the diagram for the 1950s, equidistant lines have been drawn at intervals of 0.001 units of distance calculated from the parameters of Malecot's isolation-by-distance model and do not intersect with one another. Most rural municipalities form a single equidistant figure, although they belong to different administrative districts. The Krasnolip'e and Platava rural municipalities of Rep'evka raion, which are geographically close to each other and remote from other rural municipalities studied, form a separate equidistant figure, which suggests an important role of geographic distances in their differentiation. Ultimately (at a level of 0.003), they join the common group of all other rural municipalities studied. Spearman's correlation coefficient between the matrix of distances calculated from the parameters of Malecot's isolation-by-distance model and the matrix of geographic distances for the 1990s (0.68,  $p < 0.001$ ) is about as high as that for the 1950s.

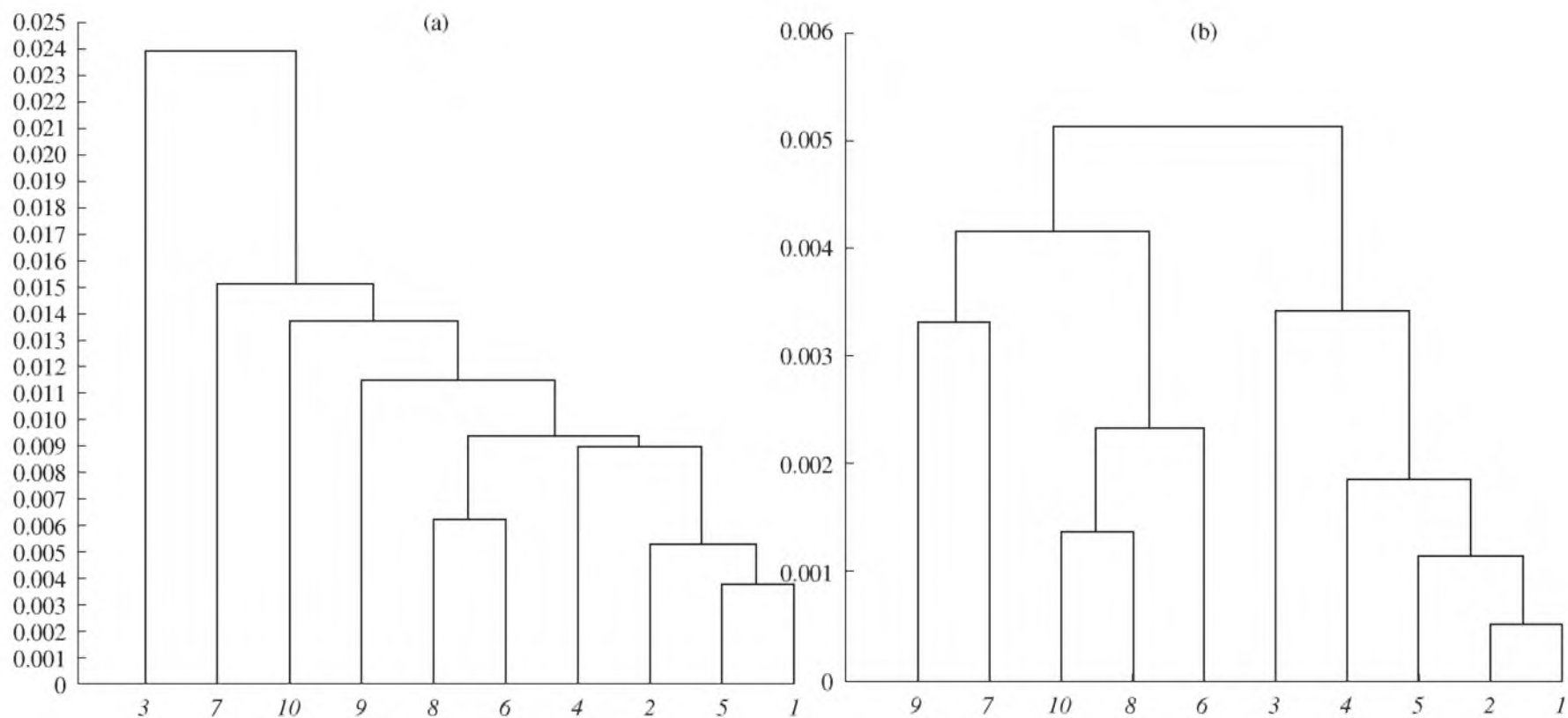
Thus, geographic distances played an important role in the genetic subdivision of the analyzed rural munic-



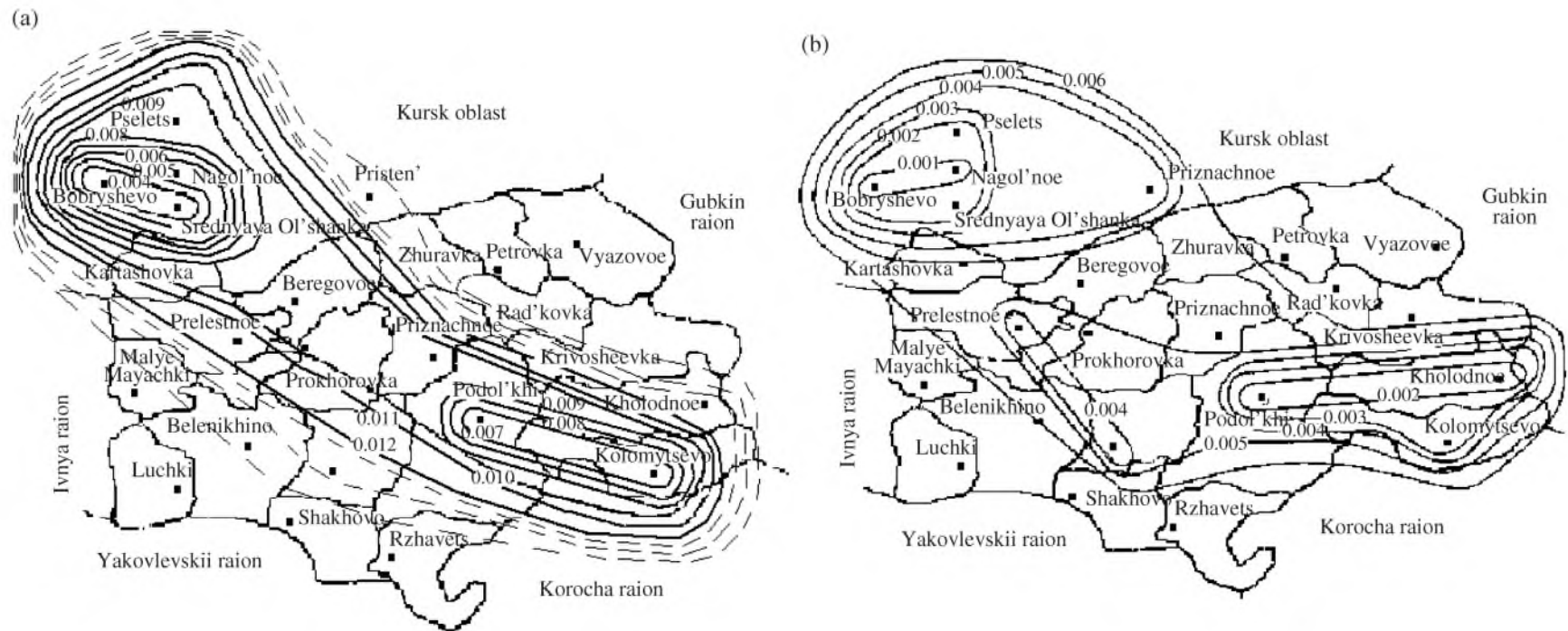
**Fig. 2.** The dendrogram of the distances calculated from the parameters of Malecot's isolation-by-distance model for rural municipalities of Rep'evka and Krasnoe raions (constructed by the median-joining method): (a) the 1950s; (b) the 1990s. Rural municipalities: 1, Butyrki; 2, Krasnolip'e; 3, Platava; 4, Rep'evka (Rep'evka raion); 5, Gorki; 6, Gotov'e; 7, Kamyzino; 8, Raskhovets; 9, Krasnoe (Krasnoe raion).



**Fig. 3.** Equidistant figures constructed on the basis of the distances that were calculated from the parameters of Malecot's isolation-by-distance model for rural municipalities of Rep'evka and Krasnoe raions: (a) the 1950s; (b) the 1990s.



**Fig. 4.** The dendrogram of the distances calculated from the parameters of Malecot's isolation-by-distance model for rural municipalities of Pristen' and Prokhorovka raions (constructed by the median-joining method): (a) the 1950s; (b) the 1990s. Rural municipalities: 1, Bobryshevo; 2, Nagol'noe; 3, Pristen'; 4, Pselets; 5, Srednyaya Ol'shanka (Pristen' raion); 6, Kolomytsevo; 7, Plota; 8, Podol'khi; 9, Prelestnoe; 10, Kholodnoe (Prokhorovka raion).



**Fig. 5.** Equidistant figures constructed on the basis of the distances that were calculated from the parameters of Malecot's isolation-by-distance model for rural municipalities of Pristen' and Prokhorovka raions: (a) the 1950s; (b) the 1990s.

ipalities of Krasnoe and Rep'evka raions in both the 1950s and the 1990s, whereas the role of the fact that they belonged to different administrative districts decreased in the period from the 1950s to the 1990s.

Data on the changes in the genetic relationships between rural municipalities of Prokhorovka and Pristen' raions estimated using Malecot's isolation-by-distance model demonstrated somewhat different tendencies. As evident from the data on the 1950s (Figs. 4a, 5a), all the analyzed rural municipalities could be divided into two clusters. One cluster comprised almost all rural municipalities of Pristen' raion (except for the Pristen' municipality, which is located farther to the north). The center of other cluster is formed only by rural municipalities of Prokhorovka raion (Kolomytsevo and Podol'khi). Other rural municipalities of Prokhorovka raion (Prelestnoe, Kholodnoe, and Plota) and the Pristen' municipality of Pristen' raion joined the above two clusters successively. The geographic distances have not played a noticeable role in the formation of this subdivision of the analyzed teen rural municipalities of Prokhorovka and Pristen' raions (Spearman's correlation coefficient between the matrix of distances calculated from the parameters of Malecot's isolation-by-distance model and the matrix of geographic distances is 0.19,  $p > 0.05$ ).

Note that, in general, these data agree with our data on rural municipalities of Krasnoe and Rep'evka raions in the 1950s, which formed two clusters according to their affiliation with different administrative districts.

By the 1990s, the relationships between rural municipalities of Prokhorovka raion of Belgorod oblast and Pristen' raion of Kursk oblast changed in such a way that the roles of both their location in different administrative districts and geographic distances between them in the formation of the subdivided population structure of the two districts increased. The dendrogram obtained by means of cluster analysis of the matrix of distances calculated from the parameters of Malecot's isolation-by-distance model (Fig. 4b) and the equidistant figures constructed according to it (Fig. 5b) show that the local populations can be grouped into two clusters, as is the case for the data on the 1950s. It is noteworthy that the first cluster comprises only municipalities of Prokhorovka raion of Belgorod oblast (Podol'khi, Kholodnoe, Kolomytsevo, Plota, and Prelestnoe), and the second cluster, only municipalities of Pristen' raion of Kursk oblast (Bobryshevo, Pselets, Nagol'noe, Srednyaya Ol'shanka, Pselets, and Pristen'). Spearman's correlation coefficient between the matrix of distances calculated from the parameters of Malecot's isolation-by-distance model and the matrix of geographic distances for the 1990s is almost four times higher than that for the 1950s. This coefficient is 0.77 ( $p < 0.001$ ), which is similar to the correlation coefficient calculated for rural municipalities of Krasnoe and Rep'evka raions ( $p = 0.68$ ). At the same time, the absolute distances between the analyzed rural

municipalities of Prokhorovka and Pristen' raions, as well as between those of Krasnoe and Rep'evka raions, substantially decreased during the period from the 1950s to the 1990s (from 0.015–0.024 to 0.004–0.005).

In summary, there was distinct differentiation between rural municipalities of Prokhorovka and Pristen' raions belonging to two neighboring administrative regions. The differentiation was related to both their administrative affiliation and geographic location.

El'chinova [6], when describing the genetic landscapes of two districts of Kirov oblast, demonstrated a considerable effect of changes in the regional administrative subdivision on human population structure. El'chinova showed not only differentiation between two districts, but also subdivision within one of them, which the author relate to the fact that this district was formed out of two districts slightly more than one generation lifetime before the study. According to the results of the study [6], changes in the territorial boundaries of some administrative districts of Marii El Republic also substantially affected the population genetic structure.

Thus, the results of our study show that changes in the regional administrative subdivision may affect population structure in different ways and cause different microevolutionary trends. For some populations (those that formerly belonged to Kursk oblast), changes in the regional administrative structure (after which some of them remained in Kursk oblast and others were included into Belgorod oblast) have lead to an increase in the genetic distances between them. However, other populations (formerly belonging to Voronezh oblast) have become genetically closer to one another, although these populations now belong to different administrative regions (Belgorod and Voronezh oblasts).

## ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for the Humanities and Russian Foundation for Basic Research.

## REFERENCES

1. *Nasledstvennye bolezni v populyatsiyakh cheloveka* (Hereditary Diseases in Human Populations), Ginter, E.K., Ed., Moscow: Meditsina, 2002.
2. Balanovskii, O.P., Buzhilova, A.P., and Balanovska, E.V., The Russian Gene Pool: Gene Geography of Surnames, *Russ. J. Genet.*, 2001, vol. 37, no. 7, pp. 974–990.
3. Khusnutdinova, E.K., *Molekulyarnaya etnogenetika narodov Volgo-Ural'skogo regiona* (Molecular Ethnogenetics of Populations of the Volga–Ural Region), Ufa, 1999.
4. El'chinova, G.I., Mamedova, R.A., Brusintseva, O.V., and Ginter, E.K., Comparison of Several Russian Populations with Respect to Vital Statistics and Genes Fre-



- quencies, Causing Hereditary Diseases, *Russ. J. Genet.*, 1994, vol. 30, no. 11, pp. 1558–1559.
5. El'chinova, G.I., A Metric Based on Malecot's Parameters of Isolation-by-Distance as a Characteristic of Genetic Similarity of Populations, *Russ. J. Genet.*, 2000, vol. 36, no. 6, pp. 856–858.
  6. El'chinova, G.I., Application of Population Genetic Analysis to the Study of Russian Populations with Different Genetic Demographic Structure:., *Extended Abstract of Doctoral (Biol.) Dissertation*, Moscow: Medical Genetic Research Center of the Russian Academy of Medical Sciences, 2001.
  7. Sorokina, I.N., Balanovska, E.V., and Churnosov, M.I., The Gene Pool of the Belgorod Oblast Population: Malecot's Isolation-by-Distance Parameters, *Russ. J. Genet.*, 2009, vol. 45 no. 3, pp. 383–389.
  8. El'chinova, G.I., Kadoshnikova, M.Yu., and Mamedova, R.A., Revealing the Peculiarities of Population Genetic Structures by the Descriptive Method of "Genetic Landscape", *Russ. J. Genet.*, 1991, vol. 27, no. 11, pp. 1994–2001.
  9. Paradeeva, G.M., Subdivision Factors of Russian Populations of the Nechernozemye Region, *Extended Abstract of Cand. Sci. (Biol.) Dissertation*, Moscow, 1987, p. 22.
  10. Kadoshnikova, M.Yu., Golubtsov, V.I., El'chinova, G.I., et al., The Genetic Demographic Structure of the Adyg Population, *Russ. J. Genet.*, 1991, vol. 27, no. 2, pp. 327–334.