ISSN 1022-7954, Russian Journal of Genetics, 2011, Vol. 47, No. 11, pp. 1353–1363. © Pleiades Publishing, Inc., 2011.
Original Russian Text © O.P. Balanovsky, S.M. Koshel, V.V. Zaporozhchenko, et al., 2011, published in Genetika, 2011, Vol. 47, No. 11, pp. 1523–1535.

HUMAN GENETICS

Genetic Ecological Monitoring in Human Populations: Heterozygosity, mtDNA Haplotype Variation, and Genetic Load

O. P. Balanovsky[a, b], S. M. Koshel[c], V. V. Zaporozhchenko[d], A. S. Pshenichnov[b], S. A. Frolova[b], M. A. Kuznetsova[b], E. E. Baranova[a], I. E. Teuchezh[e, a], A. A. Kuznetsova[f, a], M. V. Romashkina[g, a], O. M. Utevska[h], M. L. Churnosov[i], R. Villems[j], and E. V. Balanovska[b]

[a] *Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119901 Russia*
e-mail: balanovsky@inbox.ru
[b] *Medical Genetic Research Center, Russian Academy of Medical Sciences, Moscow, 115478 Russia*
[c] *Moscow State University, Moscow, 119992 Russia*
[d] *Research Center of Drug Evaluation, Ministry of Health and Social Development of the Russian Federation, Moscow, 117246 Russia*
[e] *Adygei State University, Maikop, 385000 Russia*
[f] *Chita State University, Chita, 672039 Russia*
[g] *Ogarev Mordovian State University, Saransk, 430005 Russia*
[h] *Kharkov State University, Kharkov, 61077 Ukraine*
[i] *Belgorod State University, Belgorod, 308015 Russia*
[j] *Estonian Biocenter, Tartu, 51010 Estonia*

Received May 19, 2011

Abstract—Yu.P. Altukhov suggested that heterozygosity is an indicator of the state of the gene pool. The idea and a linked concept of genetic ecological monitoring were applied to a new dataset on mtDNA variation in East European ethnic groups. Haplotype diversity (an analog of the average heterozygosity) was shown to gradually decrease northwards. Since a similar trend is known for population density, interlinked changes were assumed for a set of parameters, which were ordered to form a causative chain: latitude increases, land productivity decreases, population density decreases, effective population size decreases, isolation of subpopulations increases, genetic drift increases, and mtDNA haplotype diversity decreases. An increase in genetic drift increases the random inbreeding rate and, consequently, the genetic load. This was confirmed by a significant correlation observed between the incidence of autosomal recessive hereditary diseases and mtDNA haplotype diversity. Based on the findings, mtDNA was assumed to provide an informative genetic system for genetic ecological monitoring; e.g., analyzing the ecology-driven changes in the gene pool.

DOI: 10.1134/S1022795411110056

INTRODUCTION

Writing this article for the issue devoted to the memory of Academician Yu.P. Altukhov, we remembered the important step in the development of Russian population genetics that was closely associated with Altukhov. A comparison with modern studies shows that the current step is characterized by a rapid accumulation of experimental data, which are important, but particular results. Altukhov and his colleagues focused more on observing regularities, generalizing, and formulating concepts, which provided a bridge between theoretical constructs and applied studies of gene pool diversity.

As one of his generalizations, Altukhov suggested that heterozygosity is an integral parameter that is associated with both individual lifespan [1, 2] and the state of the population gene pool, noting that both a decrease and an excessive increase in heterozygosity are similarly unfavorable for the normal function of a population [3]. A ratio between heterozygosity (intrapopulation variation) and interpopulation variation was considered to be an important species-specific constant, any bias from the constant suggesting an unfavorable state of the gene pool structure [3]. In this idea, Altukhov's works are closely related to studies by Professor Yu.G. Rychkov, another leader of Russian population genetics, and his school, especially to studies of the distribution of interpopulation variation [5–7] and genetic ecological monitoring [8].

In memory of these prominent researchers and in an attempt to reestablish, at least partly, the theoretical component in Russian population genetics, we analyzed new experimental data on mitochondrial DNA (mtDNA) variation in East European populations. Rather than analyzing the data in the context of historical migrations as is common now (such an analysis was performed previously; e.g., see [9]), we attempted

a theoretical generalization, using the concepts of the important role of heterozygosity and the approaches to genetic ecological monitoring, which had been developed by Altukhov, Rychkov, and their colleagues. In particular, we considered the geographical variation of haplotype diversity (an analog of heterozygosity in the case of mtDNA) and its association with both ecological parameters and the population load of hereditary diseases (genetic load).

The concept of genetic ecological monitoring [8, 10] is based on the idea that the gene pool is a vital resource of the population and is important for its stable existence. These views of the gene pool originated from works of A.S. Serebrovsky, a founder of gene geography, and was developed by Rychkov's school of Russian gene geography. Similar studies are now in progress in the global science; an example is provided by pharmacogenetics, which focuses on the differences in drug responses among populations and is intensely developed now. In a general sense, such studies are a particular field of ecological genetics [11], which focuses on the mechanisms that sustain adaptation of populations to the given environment and manifest themselves in increasing frequencies of the genes that are most favorable in the given conditions. It is important to note that the majority of studies focus on a particular gene that plays a certain role in adaptation of its carrier to natural or artificial environmental conditions. In contrast, Russian population genetics focuses on the gene pool and its structure, that is, on changes in the population distribution of the total gene set. Examples are provided by Altukhov's works on the association between mean heterozygosity (averaged over many genes) and lifespan and on the optimal level of interpopulation differences [1–3].

Although studies focusing on similar problems are many, we could not find a work that where adaptation of human populations is considered on the basis of the structure of the total gene pool, rather than individual genes, and changes in the gene pool are monitored over time. In this work, we summarized our investigations in the field. Mitochondrial DNA was chosen as the DNA system best characterized among all genetic markers.

MATERIALS AND METHODS

Dataset on mtDNA variation. Mitochondrial DNA is one of the genetic systems most intensely studied in current research, and dozens of population genetic studies focusing on mtDNA analysis in particular populations are published annually. Data on mtDNA can additionally be found in papers on forensic medicine, where mtDNA is employed in personal identification. Thus, ample publications with data on an extremely large sample of mtDNA specimens appeared in the past 20 years. Our team have collected and systematized this information in a database for many years.

The resulting database MURKA (mtDNA database and integrated software; Zaporozhchenko, Balanovsky, Pshenichnov, and Balanovska), which accumulated for several years, is intended for further development. The database included information in 33000 mtDNA samples in 2004, 62000 mtDNA samples in 2007, 95000 samples in 2008, and 135000 samples in 2010 and currently has data on 168000 mtDNA samples. An important feature of our database is that the populations under study are characterized in the most thorough and comprehensive manner; the characteristics include the population designation, sample size, administrative location (country and province), ethnicity, language group, geographical coordinates, references to the data sources, etc.

The applications of the database are not restricted to a study of the East European populations, as in this work. Using the database, we analyzed the general regularities at the formation of populations in high-latitude Eurasia and, in particular, the origin of Saami (Lopars) [12]. We found that the north Eurasian populations preserve the separation of western and eastern gene pools, as common for total Eurasia, and that the eastern component comes to prevail considerably in northern regions, extending far into Europe. However, this influence did not reach the Saami, and all main components of their mitochondrial gene pool suggest an East European origin for the ancestors of the Saami.

In collaboration with French researchers, we used the database to study the genetic relationships of two Central African population groups, Bantu-speaking farmers and Pygmy hunter–gatherers [13]. The study showed that mtDNA haplogroup L1c is the most prevalent in both of the groups and that their common ancestral population lived approximately 70000 years ago. The Pygmy and Bantu ancestral populations evolved independently during their more recent history. A recent gene flow was demonstrated for the two Central African populations and proved to be asymmetric, from the Pygmy to Bantu population, but not otherwise. This finding agrees with the ethnographic observations that, although rare, marriages between Bantu males and Pygmy females do occur, whereas opposite marriages are unknown.

The above two examples show that the database has a broad range of applications. As a universal application, we created an atlas of the geography of mtDNA haplogroups in Eurasian populations. Database information was mapped to produce distribution maps for each of the major haplogroups. Several maps of the atlas were published in [14] and are available at our web site (www.genofond.ru). In this work, database information on the prevalence of each mtDNA haplotype in populations of East Europe and adjacent regions was used to calculate the haplotype diversity index. The information included our unpublished data to a considerable proportion, and, consequently, the haplotype diversity indices are summarized in Table 1. Information on other populations shown in the map was obtained from studies reported by other researchers.

Table 1. Levels of mtDNA diversity in populations of East Europe and adjacent regions

Conventional name	Ethnic group	Geographical localization	Sample size (number of individuals)	Longitude (E)	Latitude (N)	Haplotyp diversity index	Reference
Stolin	Belarussians	Brest Region, Stolinskii district	89	26.77	51.97	0.97	Our data
Svetlogorsk	The same	Gomel Region, Svetlogorskii district	71	29.72	52.62	0.959	The same
Vitebsk	"	Vitebsk Region	100	29	55	0.975	"
Brest	"	Brest Region	104	24	52	0.962	"
Gomel	"	Gomel Region	121	30	52	0.967	"
Kuban Kazaks	Russians	Adygeya	132	40.17	44.5	0.98	"
Tersk Kazaks	The same	Kabardino-Balkaria	124	43.5	43.5	0.972	"
Pinega	"	Arkhangelsk Region, Pinezhskii district	144	46.53	63.43	0.953	"
Belgorod Russians	"	Belgorod Region; Yakovlevskii, Krasnenskii, and Prokhorovskii districts	148	36.48	50.78	0.979	"
Smolensk	"	Smolensk Region, Roslovl'skii and Ershichskii districts	147	32.88	53.93	0.981	"
Kostroma (Unzha)	"	Kostroma Region, Manturovskii and Mezhevskoi districts (river Unzha)	79	44.77	58.33	0.958	"
Borovskii	"	Kaluga Region, Borovskii district	70	36.5	55	0.973	"
Baryatinskii	"	Kaluga Region, Baryatinskii district	75	34.53	54	0.968	"
Bolkhovskii	"	Orel Region, Bolkhovskii district	76	36	53.43	0.969	"
Cheremisinovskii	"	Kursk Region, Cheremisinovskii district	62	37.25	51.87	0.963	"
Mikhailovskii	"	Ryazan Region, Mikhailovskii district	82	39	54.22	0.978	"
Spassk-Ryazanskii	"	Ryazan Region, Spassk-Ryazanskii district	86	40.53	54.4	0.979	"
Petrovskii	"	Tambov Region, Petrovskii district	76	40.25	52.63	0.971	"
Chuvash	Chuvash	Tatarstan, Drozhzhanovskii and Kaibitskii districts	71	47.4	54.7	0.96	"
Belgorod Ukrainians	Ukrainians	Belgorod Region, Krasnogvardeiskii and Graivoronskii districts	95	35.75	50.4	0.978	"
Cherkassy	The same	Cherkassy Region	179	32.07	49.43	0.984	"
Khmelnitskaya	"	Khmelnitskaya Region, Starokonstantinovskii district	179	27.3	49.7	0.971	"
West Ukrainians	"	Lvov and Ivano-Frankovsk Regions	157	24	49.25	0.98	"
East Tatars	Tatars	Tatarstan, Aktanyshskii and Sarmanovskii district	127	54.07	55.7	0.979	"

Table 1. (Contd.)

Conventional name	Ethnic group	Geographical localization	Sample size (number of individuals)	Longitude (E)	Latitude (N)	Haplotyp diversity index	Reference
Kazan Tatars	The same	Tatarstan, Arskii and Atninskii districts	131	49.87	56.07	0.984	Our data
Other Tatars	″	Tatarstan, several districts	190	51.38	55.68	0.961	The same
Mishari	″	Tatarstan; Alekseevskii, Drozhzhanovskii, and Kaibitskii districts	98	47.95	54.73	0.97	″
Kryasheny	″	Tatarstan, Mamadyshskii and Drozhzhanovskii districts	62	51	55.75	0.975	″
Turkey	Turks	Turkey	608	33	39	0.991	[30–32]
Armenians	Armenians	Armenia	192	44.53333	40.18333	0.988	[33]
North Ossets	Ossets	North Ossetia	106	44.5	43	0.977	[31]
North Greece	Greeks	Greece	469	22.85	40.53333	0.989	[31, 34, 35]
South Romania	Romanians	Romania	105	28.58333	44.13333	0.968	[34]
Romanians	″	″	92	23.58333	47.63333	0.971	[31]
Bulgarians	Bulgarians	Bulgaria	141	26.13	42.93	0.971	[31, 32]
Poles	Poles	Poland	583	21	52	0.973	[31, 36]
Aukshtaitian Lithuanians	Lithuanians	Ëèòâà	90	24	55	0.965	[37]
Zhemaitian Lithuanians	″	″	90	22	55.5	0.96	[37]
Latvians	Latvians	Latvia	299	24	57	0.968	[38]
South Finland	Finns	Finland	105	26	62	0.959	[39, 40]
North and central Finland	″	″	403	27.46667	64.43333	0.959	[41]
Saami	Saami (Lopars)	Sweden, Norway, Finland, and Russia	434	27.61667	68.93333	0.833	[39, 42–45]
Karels	Karels	Karelia	83	34.3	63.71667	0.952	[42]
Nenets	Nenets	Nenets Autonomic District	69	53.01667	67.61667	0.953	[46]
Komis	Komis	Komis	127	53	61	0.955	[47]
Mansis	Mansis	Khanty–Mansi District	98	63.71667	61.43333	0.963	[48]
Udmurtians	Udmurtians	Udmurtia	109	52.98333	56.55	0.953	[47]
Maris	Maris	Marii El	136	48.1	55.95	0.922	[47]
Mordovians	Mordovians	Mordovia	99	44.46667	54.3	0.964	[47]
Nogaians	Nogaians	Dagestan	206	47	44	0.986	[47]
Bashkirs	Bashkirs	Bashkiria and Perm Regions	207	55.97	54.75	0.984	[47]

Calculation of the haplotype diversity index. The population frequency of a particular haplogroup (or a haplotype) is the simplest characteristic of a mitochondrial gene pool. The second parameter is the haplotype diversity index, which is calculated as unity minus the sum of square frequencies over all haplotypes found in the population:

$$\hat{H} = \frac{n}{n-1}\left(1 - \sum_{i=1}^{k} p_i^2\right),$$

where n is the sample size, k is the number of different haplotypes, and p_i is the frequency of the i-th haplotype.

When only one haplotype is found in a population (i.e., all members of the population are similar in mtDNA), the haplotype diversity index is zero. When all members of the population differ in mtDNA from each other (i.e., not a single haplotype is found in two or more individuals), the haplotype diversity index is unity. It is clear that the haplotype diversity index is calculated in the same manner as the expected heterozygosity for autosomal markers, i.e., the sum of square frequencies corresponds to the total frequency of homozygotes, and unity minus the homozygote frequency is the frequency of heterozygotes, that is, the heterozygosity of the population at the given (single) locus, assuming that Hardy–Weinberg equilibrium is obeyed. The term heterozygosity is applicable only to diploid genetic systems, while mtDNA is haploid. Hence, it is possible to think that the haplotype diversity index for haploid systems is an analog of the heterozygosity for populations corresponding to the Hardy–Weinberg proportions.

In essence, the parameter (heterozygosity or haplotype diversity) is the probability for two alleles (haplotypes) selected from the population at random to be nonidentical. The parameter is theoretical in the case of haploid systems and has practical applications in the case of diploid systems because alleles of diploid systems do unite in one zygote. In view of this, we consider haplotype diversity of the mitochondrial gene pool to be analogous to the heterozygosity of the autosomal gene pool of populations that had been the focus of many Altukhov's works.

Map and statistical analyses. Haplotype diversity was mapped using the GeneGeo program, which was developed in our team. A haplotype diversity map was constructed by the mean weighted interpolation algorithm [15] with a weighting function power of 3.

A framework to construct a map for East Europe was generated using the DigitMap program (Geographical Faculty, Moscow State University).

The resulting map (along with other maps included in the atlas of mtDNA variation in East Europe) is available at www.genofond.ru and is shown in the figure. The map is based on our original data and data reported by other authors. The total primary information used to construct the map (geographical coordinates and haplotype diversity indices) are summarized in Table 1. The sample size varied from 62 to 608 with the mean sample size $N = 160$. Although the samples were far smaller than the corresponding populations (e.g., population sizes of several millions of people are characteristic of the majority of the ethnic groups examined), samples of 100 or more people are considered to be representative in modern studies of mtDNA variation in human populations.

The coefficient of correlation between haplotype diversity index and genetic load was calculated using the program Statistica 6.0 [16].

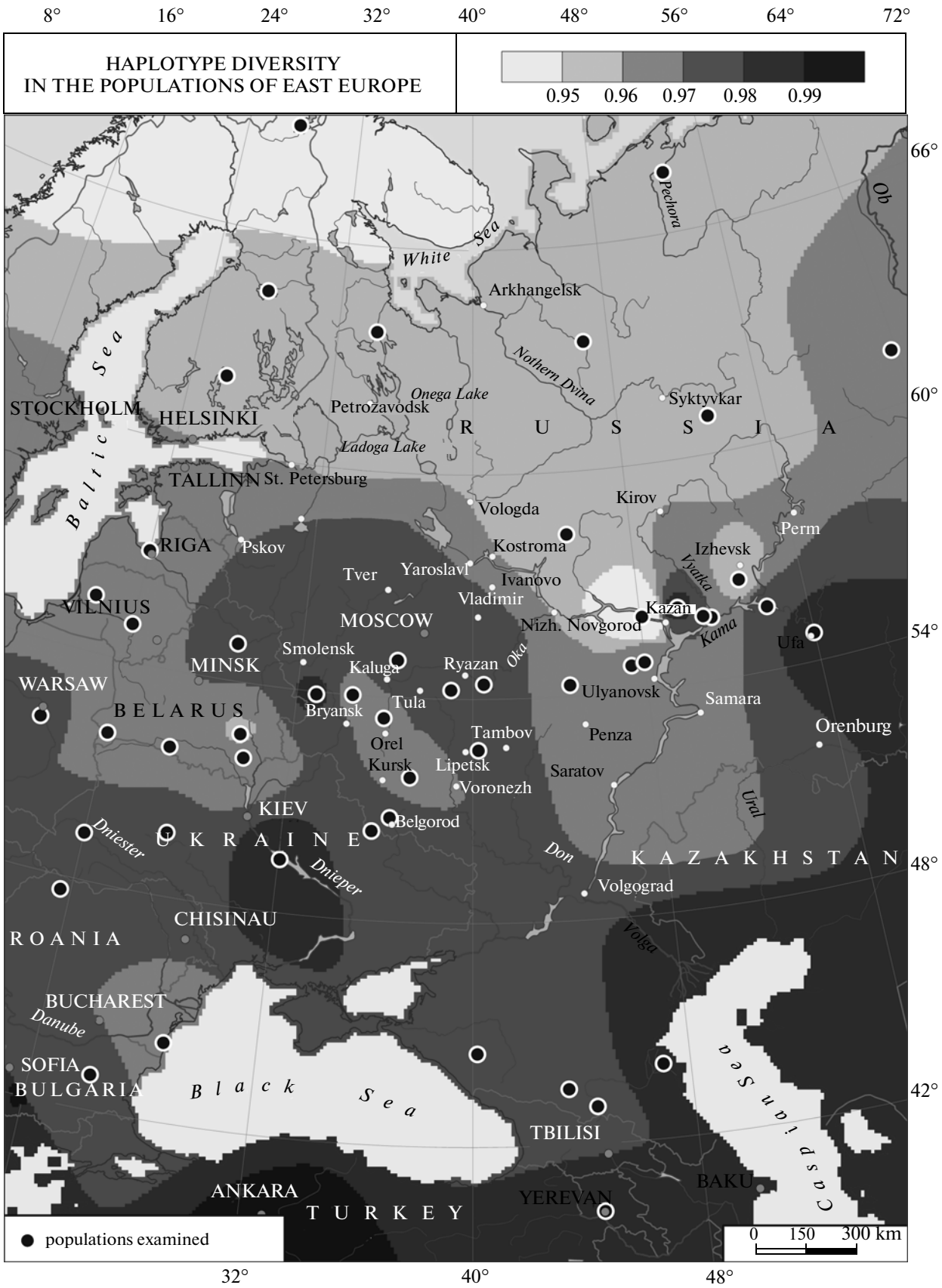
RESULTS AND DISCUSSION

Genetic Changes Determined by the Environment (the Ecological Geographical Factor)

To better understand the complex of intricate interrelated concepts, we think it necessary to take into account that genetic changes may be determined by both ecological factors (the environment) and historical factors (human migrations).

Adaptive changes in the frequencies of individual genes are the subject of intense research, which focus on numerous particular genes as examples [11, 17, 18, etc.]. The majority of the genes are autosomal because autosomal genes are expressed in the phenotype and are subject to selection. Studies of mtDNA are almost totally separate from this field because there are only a few genes in mtDNA. There are examples where natural selection presumably affects the mtDNA haplotypes and haplogroups, but they are too few and particular to believe that mtDNA plays a significant role in population adaptation to the environment.

However, mtDNA may be informative for studying the changes of the total gene pool, rather than individual genes. It is important to note that such changes are not necessarily adaptive; i.e., they do not necessarily improve the state of the population. For instance, a population that has migrated into the ecological conditions of the North is divided into several isolated subpopulations because of the low land productivity. On the one hand, the subpopulations grow more different, and, on the other hand, members of one subpopulation grow more genetically similar through generations (as a result of occasional inbreeding). This change is not adaptive because it does not improve the fitness of individuals or the total population, but it is inevitable in the ecological conditions of the population. Note that the gene pool, which is understood as a set of all genetic variants present, remains unchanged in this case, and the mean allele frequencies in the population system also remain the same. However, genetic variation is redistributed between the intrapopulation and interpopulation levels. As there is no other term (and fol-



Map of mtDNA haplotype diversity in the populations of East Europe.

Table 2. Genetic load and haplotype diversity

Population	Autosomal dominant disorders	Autosomal recessive disorders	X-linked disorders	Total genetic load	Haplotype diversity
Bashkirs	2.73	1.53	0.68	4.94	0.984
Udmurts	3.47	1.56	0.96	5.99	0.953
Chuvashes	1.82	1.25	0.47	3.54	0.967
Maris	2.23	1.43	0.33	3.99	0.922
Russians (Bryansk Region)	1.01	0.81	0.39	2.21	0.981
Russians (Kostroma Region)	1.25	0.9	0.3	2.45	0.958
Russians (Rostov Region)	2	1.34	0.44	3.78	0.976
Russians (Krasnodar Territory)	1.01	0.64	0.35	2	0.98
Russians (Kirov Region)	1.56	1.28	0.36	3.2	
Russians (Tver Region)	0.96	0.72	0.62	2.3	
Adyghes	1.07	1.41	0.5	2.98	

Note: The genetic load indices were taken from [27]. In the studies summarized in [27], the genetic load was determined via a total population screening in selected regions, and the number of families with hereditary disorders was given as the number of cases per 100000 people.

lowing the convention based on Rychkov's and Altukhov's works), we utilize the term gene pool structure, which includes not only the set of the observed alleles, but also the characteristics of their common distribution in a subdivided population and, in particular, the proportions of the variation that correspond to the intrapopulation and interpopulation levels. Examples of the populations that underwent changes in the structure of their gene pools as a result of migration into a low-productive environment are provided by the gene pools of the Saami [12, 19] and the North Russian populations [14]. Altukhov and colleagues [3] observed similar phenomena in several salmon populations.

The genetic consequences of intense fishing are another example described in Altukhov's works. When seine nets with a standard mesh size are used, smaller individuals are more likely to survive, and these individuals have, on average, a higher heterozygosity (because heterozygosity is associated with the body dimensions) [3]. An ecological factor (the appearance of seiners as predators with a particular way of catching fish in the area of a population) thereby changes the mean heterozygosity of populations (selection for heterozygotes).

The above examples illustrate the concept that ecological conditions of a population exert a regular effect not only on the frequencies of individual genes, but also on other genetic characteristics of the population. We analyzed the mtDNA variation in the East European population from this viewpoint.

Ecological Geographical Regularities at mtDNA Variation

In this work, we studied the association of the mitochondrial gene pool with specifics of the geographical environment for East Europe. As a result of the high mutation rate of mtDNA, the haplotype diversity index is extremely high (more than 97%) in the vast majority of global populations. Hence, populations with even a minor decrease of the index are of particular interest.

Our map of haplotype diversity (figure) shows that the index decreases northwards in a regular manner from 0.99 in the southern populations of Black Sear regions to 0.95 or even lower values in Arctic Ocean regions. Such a negative correlation with geographical

latitude was often observed for individual genes, and we were the first to describe it for mtDNA haplotypes.

Lower haplotype diversity may be caused by various factors of the population dynamics, including stabilizing selection, a replacement of the gene pool as a result of migrations, and genetic drift in a subdivided population (Wahlund effect). However, the first two factors are unlikely in the case of the North European gene pool. The mtDNA control region lacks genes and, consequently, is considered to be selectively neutral. As to the factor of migrations, many studies demonstrated that the European mitochondrial gene pool is homogeneous in haplogroup frequencies (northern populations are virtually indistinguishable from southern ones). This fact suggests a lack of mass migrations that are genetically effective enough to change the mtDNA haplotype frequencies in the northern populations of Europe [20]. Hence, it is possible to believe that the lower intrapopulation diversity in north European regions is indicative of genetic drift, which is especially high in the populations of the northern regions of Eastern Europe, which are small in size and isolated to a substantial extent.

This classical explanation is logical and quite sufficient for the purposes of a study of the gene pool structure. However, when ecological genetic monitoring is the purpose, it is necessary to make the next step and to answer the question as to why the total population of northern regions exists in the form of small populations. And why migrations between them are less intense than in the south?

The answer to these questions is to be sought in the ecological geographical conditions of the populations. While southern regions provide a highly productive environment for human populations (Chernozem Band), northern regions belong to the forest zone, which is less favorable for agriculture and, consequently, is only capable of providing food for a smaller population. Only small populations can exist further to the north, in the forest–tundra and tundra zones, and these populations are isolated from each other by distance because the geographical territory is vast and difficult to travel. All of these factors were highly significant in the past centuries and millennia, where the gene pools of the northern populations of Eastern Europe formed. Even now, a comparison of the density and quality of roads between northern and southern regions of European Russia makes it easy to understand why the northern populations are still isolated from each other.

These general considerations are confirmed by direct genetic demographic data. For instance, our study of the endogamy index in one of the southern regions (Belgorod Region) showed that minor changes in migration parameters were caused by changes in the administrative division of regions [21] and that only minor differences in gene pool structure are detectable between different regions of Belgorod Region. At the same time, the available genetic data demonstrate substantial genetic differences for northern Russian populations (Arkhangelsk Region) [22].

Another argument supporting the ecological geographical interpretation of the northward decrease in haplotype diversity is that haplotype diversity is similar in distribution to population density (population density maps are available at http://map-geo.ru/667943.html, http://demoscope.ru/d_f_a/density.html, etc.). The two parameters (population density and haplotype diversity) are maximum in southern regions, intermediate in central regions, and minimum in northern regions of Eastern Europe.

Thus, a common pattern of interrelated northward changes is characteristic for several parameters: geographical latitude increases, land productivity decreases, population density decreases, effective population size decreases, isolation of subpopulations increases, genetic drift increases, and mtDNA haplotype diversity decreases. The parameters are listed in the order of their putative causal relationship.

Thus, mtDNA proved to be a highly efficient tool for monitoring the changes in the geographical structure of the gene pool. Data obtained for mtDNA generally agree with data on other systems of genetic markers (a similar northward decrease in heterozygosity was demonstrated for a set of classical markers [23]). However, data on mtDNA (one genetic system) reveal the trend even better than the set of 33 autosomal genetic markers owing to a higher diversity of variants and a fourfold lower effective size. Moreover, since mtDNA is broadly used in population genetic and forensic studies, ample data on many populations are accumulating. This is an additional factor testifying that mtDNA is a highly efficient and useful system for studying the general parameters of the gene pool structure and, in particular, its ecological genetic aspect.

Haplotype Diversity of mtDNA and Genetic Load

An important aspect is adding morbidity into the above sequence of parameters. The scientific school guided by Academician E.K. Ginter showed that the load of hereditary disorders (a consequence) is associated with the genetic structure of the population, in particular, the random inbreeding rate (a cause) [24–26]. Since random inbreeding is directly associated with the effect of genetic drift, which dramatically decreases haplotype diversity (see above), a close association is possible between mtDNA haplotype diversity and the segregation component of the load of hereditary disorders.

To check this assumption, we compared the two parameters in the same populations (Table 2).

The genetic load was estimated for 11 populations of Russia by a total screening [27]. We calculated the mtDNA variation parameters for seven of these populations. In addition, an eighth population (Bryansk Region) was included in the analysis, and original data on the southernmost part of Smolensk Region (regions

that are adjacent to and historically associated with Bryansk Region) were used as a genetic component in this case.

Based on the data from Table 2, we calculated the coefficients of correlation of the mtDNA haplotype diversity index with each of the four parameters of the genetic load. Only final population genetic and medical genetic data (Table 2) were used for the calculations. The coefficients of correlation ranged from 0.1 to 0.4. The highest correlation (Pearson's correlation coefficient of 0.4) was obtained for autosomal recessive disorders. This is an important finding testifying to the adequacy of our approach because the most direct association with inbreeding (and genetic drift) is expected for recessive genes (homozygosity for these genes as a result of inbreeding determines a hereditary disorder).

The coefficient of correlation $r = 0.4$ is not high. Note, however, that this coefficient is a lower estimate of the actual correlation because the actual genetic load is extremely difficult to quantify. The screening intensity inevitably varies among different regions, and, consequently, the proportion of observed patients with hereditary disorders in the total patient number also differs among populations. The difference increases the noise and, therefore, the coefficient of correlation in statistical comparisons.

Moreover, correlations with this strength are conventionally considered to be informative in studies of the association between parameters differing in nature. For instance, the coefficient of correlation $r = 0.3$ between genetic and geographical distances was used to conclude that the distribution of Y-chromosome haplogroups depends on geographical factors in Europe [28]. Revazov et al. [29] used the same coefficient of correlation $r = 0.3$ to demonstrate that family names are suitable as a quasigenetic marker, which is commonly accepted now.

To summarize, mtDNA haplotype diversity is highly informative for studying the association between the genetic structure of populations and the load of hereditary disorders.

Hence, data on mtDNA (along with other genetic and quasigenetic markers) may be used to predict the genetic load in populations where direct genetic epidemiological data are unavailable.

Tracking Genetic Changes Caused by Historical Factors

The last aspect of ecological genetic monitoring is tracking the changes in the gene pool structure that are caused by population migrations. The migration factor is not tightly associated with the ecological component (because human migrations are a subject of historical, rather than ecological, studies), but this factor is of crucial importance for monitoring. Mass human migrations, which have increased through centuries, are now the key factor that changes the gene pool structure.

Note that it is rather difficult to separate the effects of the ecological and historical factors because the distribution of almost any gene can be explained both from the historical geographical (human migrations) and the ecological genegeographical (the effect of natural selection) viewpoints. An example is provided by the AIDS-protective allele CCR4del32 distribution, which is often considered. Its geographical pattern, which is unusually distinct, was explained by both migrations (Lucotte et al. [49]) and climatic adaptation (Limborska et al. [50]). While a suitable method to distinguish the effects of the ecological and historical factors is unavailable, we analyzed how mtDNA data help to detect the historical factor (human migrations) when the role of the ecological factor is insignificant.

It is known that mtDNA is highly informative for tracking past migrations, which is evident from hundreds of publications where the main steps and paths of human migrations on the planet were reconstructed. As for the possibility of using mtDNA to analyze the migration results that occur now or are expected for the nearest future, we propose migrations be classified into two types. The first type includes migrations of Russians, Ukrainians, and Belorussians within their historical area. The mitochondrial gene pools of these populations are highly similar [9], and migrations between these genetically similar ethnic groups of East Slavs are hardly detectable. However, it is possible to estimate the average intensity of migrations from the average genetic distances between the East Slavonic populations. The distances are 0.049 between the total East Slavonic populations, 0.043 between Belorussian populations, 0.053 between Russian populations, and 0.029 between Ukrainian populations. When the intensity of migrations between populations increases dramatically, the average genetic distance decreases. Thus, repetitive studies make it possible to compare the estimates of the average genetic distances with the distances observed for the indigenous (in three generations) populations and thus allow a genetic monitoring of the total intensity of migrations, which decrease the subdivision of the East Slavonic gene pool. Note that only a principally new level of interpopulation migrations can be tracked using this approach because of the small differences between populations, while minor changes will most likely fail to reach statistical significance.

The other type of migrations consists in the inclusion of people from geographically distant and genetically contrasting regions into the East Slavonic gene pool. The character of the mtDNA haplogroup distribution in ethnic groups of the world makes it possible to track migrations from eastern Eurasia, in particular, Siberia, China, and Southeastern Asia. Migrations from Southwestern Asia are possible to track only by the frequencies of characteristic Midle-East haplogroups, such as J and U7. As for migrations from the Caucasus and Urals, the assessment is now problematic because the mitochondrial gene pools of these

regions are highly similar to the East Slavonic gene pool. The new stage of mtDNA studies on the basis of complete mtDNA sequencing will allow a more detailed differentiation of migrations. It is possible to think, however, that the Y chromosome is the most informative for this purpose because its variants display a far greater geographical specificity than mtDNA variants.

To conclude, mtDNA is highly informative for two out of the four aspects of ecological genetic monitoring, namely, for studying the association of ecological conditions with the structure of the total gene pool and for monitoring such changes. However, mtDNA is low informative for studying adaptation with respect to individual genes and moderately informative for the fourth aspect, a monitoring of the changes in the gene pool that result from human migrations.

ACKNOWLEDGMENTS

We are grateful to R.A. Zinchenko for information on the genetic load in various populations of Russia.

This work was supported by the Russian Human Science Foundation (project no. 06-06-00640a) the Russian Foundation for Basic Research (project no. 10-06-00451), the federal program "Researchers and Teachers of Innovative Russia" (project no. GK 2011-1.4.-501-001-01715)), and the program "Molecular and Cell Biology" of the Presidium of the Russian Academy of Sciences.

REFERENCES

1. Altukhov, Yu.P., Genome Heterozygosity, Sexual Maturation Rate, and Life Expectancy, *Dokl. Akad. Nauk*, 1996, vol. 348, no. 6, pp. 842–845.
2. Altukhov, Yu.P., Genomic Heterozygosity, Metabolic Rate, and Life Span, *Dokl. Akad. Nauk*, 1999, vol. 369, pp. 589–592.
3. *Dinamika populyatsionnykh genofondov pri antropogennykh vozdeistviyakh* (Population Gene Pools Dynamics under Anthropogenic Influence), Altukhov, Yu.P., Ed., Moscow: Nauka, 2004.
4. Altukhov, Yu.P. and Rychkov, Yu.G., Population Systems and Their Structural Components: Genetic Stability and Variation, *Zh. Obshch. Biol.*, 1970, vol. 31, pp. 507–525.
5. Rychkov, Yu.G. and Yashchuk, E.V., Genetics and Ethnogenesis: Historical Regularity of Genetic Differentiation of Human Populations (A Model and Reality), *Vopr. Antropol.*, 1985, no. 75, pp. 3–17.
6. Balanovskaya, E.V. and Rychkov, Yu.G., Ethnogenetics: Ethnogeographic Diversity of the All-World People's Genetic Pool, *Genetika*, 1990, vol. 26, no. 1, pp. 114–121.
7. Balanovskaya, E.V. and Rychkov, Yu.G., Ethnogenetics: Adaptive Structure of the All-World People's Genetic Pool as Revealed with Human Polymorphic Genetic Markers, *Genetika*, 1990, vol. 26, no. 4, pp. 739–748.
8. Rychkov, Yu.G. and Balanovskaya, E.V., The Concept of Ecological-Genetic Monitoring of Russia Population, *Usp. Sovrem. Genet.*, 1996, no. 20, pp. 3–38.
9. Balanovskii, O.P., Pshenichnov, A.S., Frolova, S.A., et al., Main Features of Mitochondrial Gene Pool of Eastern Slavs, *Med. Genet.*, 2010, vol. 9, no. 1, pp. 29–37.
10. Altukhov, Yu.P., Intraspecific Genetic Diversity: Monitoring and Conservation, *Russ. J. Genet.*, 1995, vol. 31, no. 10, pp. 1331–1357.
11. Spitsyn, V.A., *Ekologicheskaya genetika cheloveka* (Human Ecological Genetics), Moscow: Nauka, 2008.
12. Balanovskii, O.P., Gene Pool of High Latitudes of Eurasia, or from Where the Saami Came?, *Put' na sever: okruzhayushchaya sreda i samye rannie obitateli Arktiki i Subarktiki* (A Way to the North: Environment and the Earliest Inhabitants of Arctic and Subarctic. Proc. Int. Conf.), Velichko, A.A. and Vasiliev, S.A., Eds., Moscow: Inst. Geogr. Ross. Akad. Nauk, 2008, pp. 277–282.
13. Quintana-Murci, L., Quach, H., Harmant, C., et al. Maternal Traces of Deep Common Ancestry and Asymmetric Gene Flow between Pygmy Hunter–Gatherers and Bantu-Speaking Farmers, *Proc. Natl. Acad. Sci. USA*, 2008, vol. 105, no. 5, pp. 1596–1601.
14. Balanovskaya, E.V. and Balanovskii, O.P., *Russkii genofond na Russkoi ravnine* (Russian Gene Pool on the Russian Plane), Moscow: Luch, 2007.
15. Serbenyuk, S.N., Koshel', S.M., and Musin, O.R., Methods of Geofields Modeling Using Data in Random Located Points, *Geodez. Kartograf.*, 1990, no. 11, pp. 31–35.
16. *STATISTICA for Windows (Data Analysis Software System), Version 6*, Tulsa: StatSoft, 2001, Available from: www.statsoft.com
17. Borinskaya, S., Kal'ina, N., Marusin, A., et al., Distribution of the Alcohol Dehydrogenase *ADH1B*47His* Allele in Eurasia, *Am. J. Hum. Genet.*, 2009, vol. 84, no. 1, pp. 89–92.
18. Kozlov, A.I., Balanovskaya, E.V., Nurbaev, S.D., and Balanovskii, O.P., Genogeography of Primary Hypolactation in the Old World Populations, *Genetika*, 1998, vol. 34, no. 4, pp. 551–561.
19. Tambets, K., Rootsi, S., Kivisild, T., et al., The Western and Eastern Roots of the Extreme European Genetic Outliers—the Origin of mtDNAs and Y-Chromosomes of the Saami, *Am. J. Hum. Genet.*, 2004, vol. 74, no. 4, pp. 661–682.
20. Balanovskaya, E.V., Gene Pool of Peoples of Northern Eurasia: Internal Structure and Position in the Global Gene Pool, in *Antropologiya na poroge III tysyacheletiya: Izbrannye doklady* (Anthropology on the Edge of 3d Century: Selected Reports), Moscow, 2004, vol. 1, pp. 358–389.
21. Churnosov, M.I., Sorokina, I.N., and Balanovskaya, E.V., The Gene Pool of the Belgorod Region Population: Endogamy Index and Its Changes with Time in Regional Populations, *Russ. J. Genet.*, 2008, vol. 44, no. 8, pp. 1117–1125.
22. Balanovsky, O., Rootsi, S., Pshenichnov, A., et al., Two Sources of the Russian Patrilineal Heritage in Their Eurasian Context, *Am. J. Hum. Genet.*, 2008, vol. 82, no. 1, pp. 236–250.

23. Rychkov, Yu.G., Balanovskaya, E.V., Nurbaev, S.D., and Shneider, Yu.V., Historical Genogeography of Eastern Slaves, in *Vostochnye slavyane i ikh sosedi* (Eastern Slaves and Their Neighbors), Moscow: Nauchnyi mir, 1999, vol. 1, pp. 109–134.

24. Ginter, E.K., Influence of Genetic Structure on the Burden of Hereditary Diseases in Russian Populations, *Vestn. Ross. Akad. Med. Nauk*, 1993, no. 9, pp. 23–26.

25. Zinchenko, R.A., El'chinova, G.I., Balanovskaya, E.V., et al., The Influence of Population Genetic Structure on the Burden of Monogenetic Hereditary Diseases in Russian Populations, *Vestn. Ross. Akad. Med. Nauk*, 2000, no. 5, pp. 5–10.

26. *Nasledstvennye bolezni v populyatsiyakh cheloveka* (Hereditary Diseases in Human Populations), Ginter, E.K., Ed., Moscow: Meditsina, 2002.

27. Zinchenko, R.A. and Ginter, E.K., Specific Features of Medical Genetic Counseling in Different Populations and Ethnic Groups, *Med. Genet.*, 2008, vol. 7, no. 10 (76), pp. 20—29.

28. Rosser, Z.H., Zerjal, T., Hurles, M.E., et al., Y-Chromosomal Diversity in Europe is Clinal and Influenced Primarily by Geography, rather Than by Language, *Am. J. Hum. Genet.*, 2000, vol. 67, pp. 1526–1543.

29. Revazov, A.A., Paradeeva, G.M., and Rusakova, G.I., Suitability of Russian Surnames as a "Quasi-Genetic" Marker, *Genetika*, 1986, vol. 22, no. 4, pp. 699–703.

30. Quintana-Murci, L., Chaix, R., Wells, R.S., et al., Where West Meets East: The Complex mtDNA Landscape of the Southwest and Central Asian Corridor, *Am. J. Hum. Genet.*, 2004, vol. 74, no. 5, pp. 827–845.

31. Richards, M., Macaulay, V., Hickey, E., et al., Tracing European Founder Lineages in the Near Eastern mtDNA Pool, *Am. J. Hum. Genet.*, 2000, vol. 67, no. 5, pp. 1251–1276.

32. Calafell, F., Underhill, P., Tolun, A., et al., From Asia to Europe: Mitochondrial DNA Sequence Variability in Bulgarians and Turks, *Ann. Hum. Genet.*, 1996, vol. 60, pp. 35–49.

33. Metspalu, E., Kivisild, T., Kaldma, K., et al., Mitochondrial DNA Lineages and the History of the Roms (Gypsies), Data Deposited in Genbank, 2003.

34. Bosch, E., Calafell, F., Gonzalez-Neira, A., et al., Paternal and Maternal Lineages in the Balkans Show a Homogeneous Landscape over Linguistic Barriers, except for the Isolated Aromuns, *Ann. Hum. Genet.*, 2006, vol. 70, pp. 459–487.

35. Irwin, J., Saunier, J., Strouss, K., et al., Mitochondrial Control Region Sequences from Northern Greece and Greek Cypriots, *Int. J. Legal Med.*, 2008, vol. 122, no. 1, pp. 87–99.

36. Malyarchuk, B.A., Grzybowski, T., Derenko, M.V., et al., Mitochondrial DNA Variability in Poles and Russians, *Ann. Hum. Genet.*, 2002, vol. 66, pp. 261–283.

37. Kasperaviciute, D., Kucinskas, V. and Stoneking, M., Y Chromosome and Mitochondrial DNA Variation in Lithuanians, *Ann. Hum. Genet.*, 2004, vol. 68, pp. 438–452.

38. Pliss, L., Tambets, K., Loogvali, E.-L., et al., Mitochondrial DNA Portrait of Latvians: Towards the Understanding of the Genetic Structure of Baltic-Speaking Populations, *Ann. Hum. Genet.*, 2006, vol. 70, pp. 439–458.

39. Kittles, R.A., Bergen, A.W., Urbanek, M., et al., Autosomal, Mitochondrial, and Y Chromosome DNA Variation in Finland: Evidence for a Male-Specific Bottleneck, *Am. J. Phys. Anthropol.*, 1999, vol. 108, no. 4, pp. 381–399.

40. Lahermo, P., Sajantila, A., Sistonen, P., et al., The Genetic Relationship between the Finns and the Finnish Saami (Lapps): Analysis of Nuclear DNA and mtDNA, *Am. J. Hum. Genet.*, 1996, vol. 58, no. 6, pp. 1309–1322.

41. Meinila, M., Finnila, S., and Majamaa, K., Evidence for mtDNA Admixture between the Finns and the Saami, *Hum. Hered.*, 2001, vol. 52, no. 3, pp. 160–170.

42. Sajantila, A., Lahermo, P., Anttinen, T., et al., Genes and Languages in Europe: An Analysis of Mitochondrial Lineages, *Genome Res.*, 1995, vol. 5, no. 1, pp. 42–52.

43. Dupuy, B.M. and Olaisen, B., MtDNA Sequences in the Norwegian Saami and Main Populations, in *Advances in Forensic Haemogenetics 6,* Carraudo, A., Brinkmann, B., and Bar, W., Eds., 1996.

44. Delghandi, M., Utsi, E., and Krauss, S., Saami Mitochondrial DNA Reveals Deep Maternal Lineage Clusters, *Hum. Hered.*, 1998, vol. 48, no. 2, pp. 108–114.

45. Tambets, K., Rootsi, S., Kivisild, T., et al., The Western and Eastern Roots of the Saami—the Story of Genetic "Outliers" Told by Mitochondrial DNA and Y Chromosomes, *Am. J. Hum. Genet.*, 2004, vol. 74, no. 4, pp. 661–682.

46. Tonks, S., Winney, B.J., and Evseeva, I., Comparison of Sex-Linked and Autosomal Markers in Orkney and Other North European Populations, Data Deposited in Genbank, 2006.

47. Bermisheva, M., Tambets, K., Villems, R., et al., Diversity of Mitochondrial DNA Haplogroups among Populations in the Volga–Ural Region, *Mol. Biol.*, 2002, vol. 36, no. 6, pp. 905–915.

48. Derbeneva, O.A., Starikovskaya, E.B., Wallace, D.C., et al., Traces of Early Eurasians in the Mansi of Northwest Siberia Revealed by Mitochondrial DNA Analysis, *Am. J. Hum. Genet.*, 2002, vol. 70, pp. 1009–1014.

49. Lucotte, G. and Mercier, G., The Population Distribution of the Met Allele at the PRNP129 Polymorphism (a High Risk Factor for Creutzfeldt-Jakob Disease) in Various Regions of France and in West Europe, *Infect. Genet. Evol.,* 2005, vol. 5, no. 2, pp. 141–144.

50. Limborska, S.A., Balanovsky, O.P., Balanovskaya, E.V., et al., Analysis of CCR5Delta32 Geographic Distribution and Its Correlation with Some Climatic and Geographic Factors, *Hum. Hered.,* 2002, vol. 53, no. 1, pp. 49–54.