Short Communication

Human Heredity

Hum Hered 2002;53:49–54

Received: September 29, 2000
Revision received: September 17, 2001
Accepted: September 24, 2001

Analysis of CCR5Δ32 Geographic Distribution and Its Correlation with Some Climatic and Geographic Factors

Svetlana A. Limborska[a] Oleg P. Balanovsky[a] Elena V. Balanovskaya[b] Peter A. Slominsky[a] Maria I. Schadrina[a] Ludmila A. Livshits[c] Sergey A. Kravchenko[c] Vladimir M. Pampuha[c] Elza K. Khusnutdinova[d] Victor A. Spitsyn[b]

[a]Institute of Molecular Genetics, Russian Academy of Sciences, Moscow, Russia; [b]Research Center for Medical Genetics, Russian Academy of Medical Sciences, [c]Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, Kiev, Ukraine; [d]Institute of Biochemistry and Genetics, Ufa Scientific Centre of Russian Academy of Sciences, Ufa, Russia

Key Words
CCR5 frequencies · Gene geography · Cartographic models · Climate factors

Abstract
We studied the possible effects of climatic-geographic factors on the world distribution of the mutant allele for the chemokine receptor gene *CCR5*, which has a 32-bp deletion (CCR5Δ32) preventing cell invasion by the primary transmitting strain of HIV-1. New data on *CCR5* polymorphisms in Russian, Ukrainian, and Moldavian populations are presented. All available data on CCR5Δ32 frequencies in the Old World (number of populations n = 77) were used for construction of a geographical gene map to analyze possible correlations between allele frequencies and eight climatic-geographic parameters. A strong positive correlation was found between the allele frequency and latitude (r = 0.72), a strong negative correlation with annual radiation balance (r = –0.66), and a weaker negative correlation with longitude (r = –0.34). Partial correlations were calculated excluding the influence of latitude. The negative correlation between the allele frequency and annual radiation balance decreased (r = –0.42), but remained large and significant. We propose that the existence of correlations between the cline of CCR5Δ32 frequencies and climatic-geographic parameters provides evidence for a possible effect of either natural environmental factors or large-scale population movements on the distribution of this allele.

Introduction

The chemokine receptor CCR5 is encoded by the *CMKBR5* gene located on human chromosome 3p21.3, and represents the major coreceptor for the macrophage-tropic strains of human immunodeficiency virus (HIV-1) [1–7].

Homozygotes for a 32-bp deletion in the gene segment encoding the second extracellular loop of CCR5 (CCR5Δ32) are resistant to infection [7–10]. This mutant allele is either rare or absent in populations from Black Africa and the Far East [11, 12]. The high frequency of the CCR5Δ32 allele in some Caucasian populations raises the question of whether it is the result of random genetic drift or a consequence of selective pressure, possibly driven by

KARGER
Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0001–5652/02/0531–0049$18.50/0

Accessible online at:
www.karger.com/journals/hhe

Dr. S.A. Limborska, PhD
Institute of Molecular Genetics, Russian Academy of Sciences
Kurchatov Square, 2
Moscow, 123182 (Russia)
E-Mail limbor@img.ras.ru

Table 1. Distribution of CCR5Δ32 allele in Old World populations

Population	CCR5Δ32 allele frequency	n	Reference	Population	CCR5Δ32 allele frequency	n	Reference
Russians from Moscow (Russia)	0.139	50	13	Hong Kong	0.000	50	12
Russians from Ryazan (Russia)	0.120	78	14	Taiwan	0.000	83	12
Tatars (Russia)	0.120	40	13	Mongolia	0.000	59	12
Tatars from Almetyevsk (Russia)	0.115	48	this study	Sri Lanka	0.000	37	12
Tatars from Elabuga (Russia)	0.156	45	this study	Burma	0.000	67	12
Tuva republic (Russia)	0.030	50	26	Thailand	0.005	101	12
Udmurts from Malo-Purginski district				Andaman Islands	0.000	24	12
(Russia)	0.115	52	27	Nigeria	0.005	111	12
Bashkirs from Abselilovski district (Russia)	0.025	60	27	Gambia	0.000	56	12
Bashkirs from Ilishevski district (Russia)	0.050	50	27	Ivory Coast	0.000	87	12
Bashkirs from Arkhangelski district				Iceland	0.147	102	12
(Russia)	0.022	46	27	Ireland	0.046	44	12
Bashkirs from Sterlibashevski district				Britain	0.111	283	12
(Russia)	0.051	49	this study	Norwegians from Oslo (Norway)	0.105	100	13
Chuvashes from Morgaushevski district				Cyprus	0.042	84	12
(Russia)	0.051	79	this study	Italy	0.055	91	12
Mordvinians from Saransk (Russia)	0.163	86	13	Italians from Milan (Italy)	0.087	98	13
Mordvinians from Staroshayginski district				Sardinians from Cagliary (Italy)	0.040	100	13
(Russia)	0.059	51	27	Greece	0.024	63	12
Komi Ziryanes from Sisolski district				Hungarians from Budapest (Hungary)	0.086	99	13
(Russia)	0.060	50	27	Turks from Ankara (Turkey)	0.063	104	13
Mari from Zvenigovski district (Russia)	0.096	47	27	Finns from Helsinki (Finland)	0.158	98	13
Daghestan (Russia)	0.064	110	12	French from Lille region (France)	0.109	101	29
Georgia	0.000	50	28	French from Champagne-Ardenne (France)	0.087	276	29
Azerbaijan	0.050	40	28	Population of Nancy and Strasbourg			
Kazakstan	0.060	50	28	(France)	0.111	291	29
Uzbekistan	0.085	40	26	French from Paris (France)	0.129	294	29
Uigur (Russia)	0.040	50	26	Bretons from Bretagne (France)	0.136	107	29
Ukrainians from Lvov (Ukraine)	0.096	88	this study	French from south of France	0.098	326	29
Ukrainians from Kiev (Ukraine)	0.110	83	this study	Catalans from Perpignan region (France)	0.069	102	29
Ukrainians from Lugansk (Ukraine)	0.091	86	this study	Corses from Ajaccio region (France)	0.010	104	29
Crimea Tatars from Crimea (Ukraine)	0.050	90	this study	French from Montpellier (France)	0.101	99	13
Moldavians from Kishinev (Moldavia)	0.080	56	this study	Basques from Biarritz (France)	0.018	111	29
Byelorussians (Byelorussia)	0.160	80	14	Danes from Copenhagen (Denmark)	0.110	100	13
Lithuanians from Vilnius (Lithuania)	0.115	283	13	Belgians from Brussels (Belgium)	0.092	704	13
Saudi Arabia	0.021	241	12	Swedes from Umea (Sweden)	0.142	204	13
Yemen	0.000	34	12	Swedish Saamis from Lapland (Sweden)	0.083	120	13
Gujarat	0.047	32	12	Spaniards from Murcia (Spain)	0.095	100	13
Pakistan	0.029	34	12	Basques from Bilbao (Spain)	0.062	89	13
Sindh	0.017	29	12	Basques (Spain)	0.086	29	12
Punjab	0.015	34	12	Catalonia (Spain)	0.082	49	12
Bengal	0.000	25	12	Portugese from Lisbon (Portugal)	0.064	101	13

an increased resistance to some infectious agents or by other factors. Analysis of tightly linked polymorphic microsatellites demonstrated strong linkage disequilibrium [13]. The results of screening >2,500 individuals from 18 European populations suggest a unique and relatively recent origin for the CCR5Δ32 allele, which probably arose in northeastern European populations. Therefore, its current broad distribution may be due to some selective advantage. Here, we present data on the possible effects of factors related to geography and climate on the distribution of the CCR5Δ32 allele in human populations of the Old World.

Materials and Methods

We analyzed the frequencies of the CCR5Δ32 allele among some ethnic groups of Russia, the Ukraine, and Moldavia by the polymerase chain reaction, as described previously [14]. Two Tartar populations were studied from the small industrial towns of Elabuga and Almetyevsk. Tartars belong to the Turkic language family.

A third population sample was collected in central Russia from representatives of the rural Sterlibashevski region of the Bashkirian Republic. The Bashkirs are one of many Turkic-speaking peoples of eastern Europe, who also dwell around the southern Ural mountains.

The Chuvashes who live in the basin of the Volga River were the fourth population genotyped. The samples studied came from the

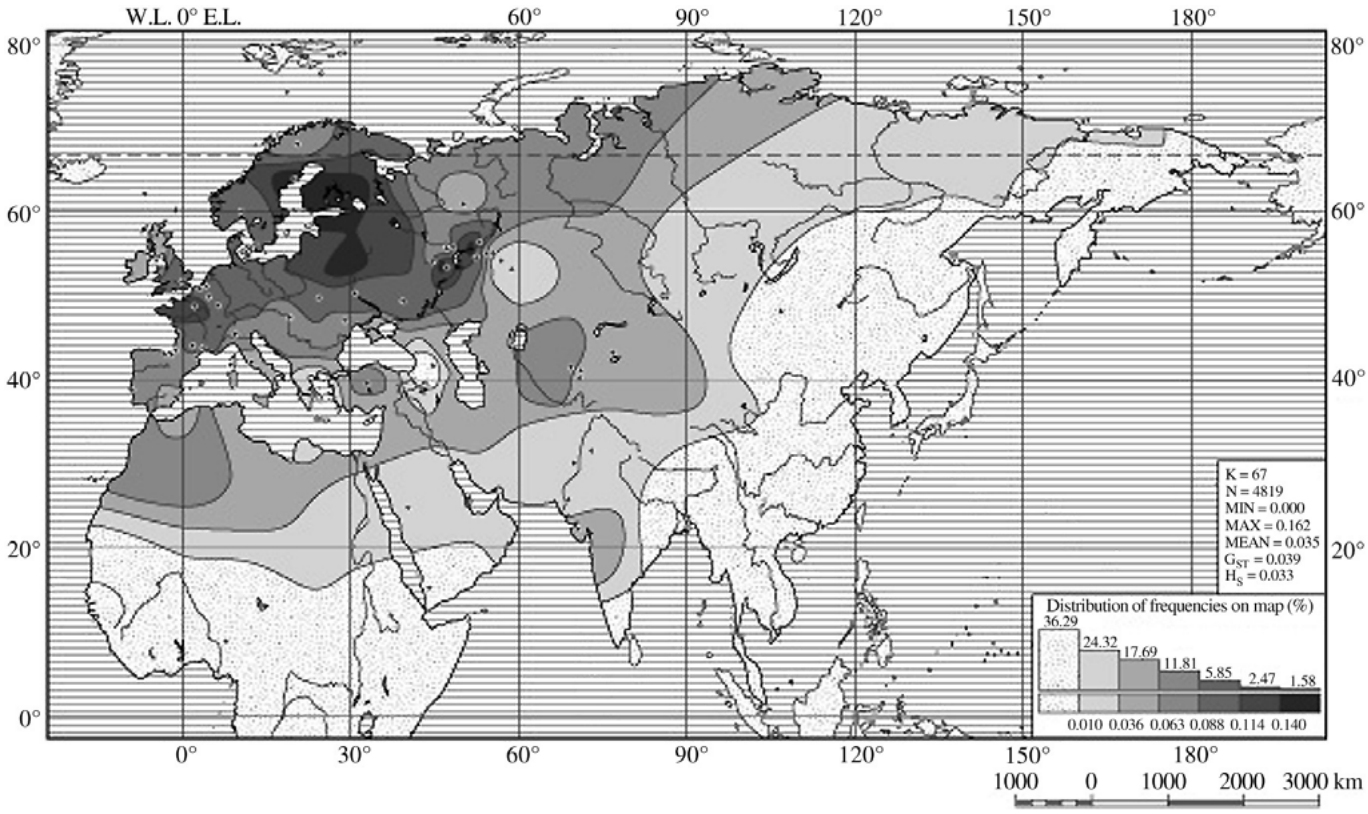

Fig. 1. Distribution of CCR-5 deletion in the Old World. Statistical parameters of the territory where the greatest number of initial data have been obtained: K = number of initial populations (asterisks); N = number of nodes in the regular map grid, characteristic of the mapped territory; MIN, MAX and MEAN refer to gene frequency values for the whole territory; G_{ST} is a measure of standard interpopulational genetic differences (normalized variance). H_s = Average heterozygosity. Bars indicate ranges of gene frequencies, and values above refer to the percentage of the territory characterized by the respective gene frequency.

inhabitants of the rural Morganshevski region of the Chuvashian Republic.

In addition, we analysed one urban Moldavian population from Kishinev, the capital of the Moldavian Republic. Moldavians speak an Indo-European Romance language.

Three Ukrainian population samples were collected in Lvov, Kiev, and Lugansk. The Ukrainian language is also part of the Indo-European family, within the Slavic subfamily. From the viewpoint of classical anthropology, Ukrainians are a rather heterogeneous people. However, they are closely related to Russians and middle-Europeans.

Finally, we investigated a small Tartar population from the Crimean peninsula in the southern part of the Ukraine, who speak a Turkic (Altaic family) language.

Geographic maps were drawn based on the gene frequencies in the Old World populations listed in table 1. A cartographic model of the CCR5Δ32 allele distribution was constructed using gene-geographical methods of interpolation [15–17] similar to those used by Cavalli-Sforza et al. [18] and Sokal et al. [19].

The maps we evaluated are based on a uniform grid covering the regions studied. Each element of the matrix corresponds to a node of the grid, and is calculated from the data from all populations studied within a specified circle. A 5,000-km radius was used for the circle, and neighboring localities beyond a defined radius were ignored. The calculations were performed with the use of orthogonal polynomials. The value of a given trait in each population was assigned a weight coefficient, which was inversely related to the distance between the population and the grid node. The weight coefficient used was equal to 6. Smoothing the maps followed the interpolation and involved averaging in a rectangle of a given size. The size can vary depending on the purpose of smoothing.

In figure 1, the population data are interpolated on the mapped territories by the zeroth-order orthogonal polynomial method. The intensely and lightly hatched zones correspond to territories with high and low frequencies of the CCR5Δ32 allele, respectively.

Data from 77 populations from various parts of the world were used for correlation analysis. Climatic-geographic data (annual radiation balance, average January temperature, average July temperature, total amount of insolation, altitude, annual precipitation rate) were obtained from an atlas [20]. Spearman's rank-order correlation coefficients were computed between CCR5Δ32 frequencies and climatic-geographic variables.

Table 2. Coefficients of correlation between CCR5Δ32 allele frequencies and climatic-geographic parameters

Parameter number	Climatic parameters	Coefficients of rank-order Spearman correlation	Coefficients of partial correlation		
			coefficient values depending on temperature parameters	numbers of parameters which are held constant[a]	coefficient values if latitude is held constant
Temperature parameters					
1	Annual radiation balance (kcal/cm^2 per year)	−0.66***	−0.62	*2, 3, 4*	−0.42
2	Average January temperature (°C)	−0.50***	+0.50	*1, 3, 4*	+0.25
3	Average July temperature (°C)	−0.64***	−0.22	*1, 2, 4*	−0.09
4	Total amount of insolation (kcal/cm^2 per year)	−0.66***	−0.06	*1, 2, 3*	−0.22
Geographical coordinates					
5	Longitude	−0.34**			–
6	Latitude	+0.72***			–
Common parameters					
7	Altitude (m)	−0.26*			−0.34
8	Annual precipitation rate (mm/year)	−0.07			−0.07

Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

[a] *1* designates annual radiation balance; *2* designates average January temperature; *3* designates average July temperature; *4* designates total amount of insolation.

Results

Figure 1 shows a map of the CCR5Δ32 allele distribution, depicting allele variability in the Old World populations. Table 2 presents rank-order pair correlations between the CCR5Δ32 allele frequencies and each climatic-geographic factor addressed. We found a strong positive correlation with latitude (r = 0.72) and a somewhat weaker negative correlation with longitude (r = –0.34). Our data also suggest that the annual radiation balance, total amount of insolation, and the average July temperature also affect the CCR5Δ32 allele frequency and its expansion over the world (r = –0.66, r = –0.66, and r = –0.64, respectively). The average January temperature and altitude have weaker negative effects (r = –0.50 and r = –0.26, respectively). The annual precipitation rate showed no correlation with the frequency of CCR5Δ32 gene distribution in the Old World. Coefficients of partial correlation were also calculated (table 2). The first four factors are closely interrelated, so it was necessary to clarify which of these factors is the most important after the three other factors are excluded. We found that only one factor, the annual radiation balance, still correlated strongly with CCR5Δ32 allele frequencies (r = –0.62). Interestingly, compared with pair correlations, partial correlations between allele frequencies and the average July temperature, and between allele frequencies and the total amount of insolation, decreased markedly, whereas the partial correlation with average January temperature increased.

The temperature parameters (parameters 1–4, table 2) are apparently tied to latitude. Therefore, another range of partial correlations was calculated; i.e., partial correlations without the influence of latitude (table 2, last column). Lower coefficients were calculated for almost all parameters, but the correlation with radiation balance remained higher than for the other parameters. Although this latter correlation decreased (r = –0.42), it was still significant. Thus, the analyses of pair and partial correlations revealed a negative correlation between CCR5Δ32 allele frequencies and the annual radiation balance. This association can also be explained by correlation with other climatic factors not analysed in this study, but tied to the annual radiation balance.

Discussion

The data presented show that the frequency of the CCR5Δ32 allele is highest in the populations of northeastern Europe (fig. 1). Its frequency gradually decreases from the Baltic region in all directions.

There are two alternative explanations for such clines: possible effects of geographically variable selective factors that we cannot yet specify, or some episode of gene flow in the course of the history of the Old World populations. A cline or gradient may reflect adaptation to variable environments, a population expansion at a moment in time, or continuous gene flow between groups that initially differed in allele frequencies. Selection tends to affect single genes, whereas demographic changes determine similar patterns across the genome. In Europe, many genetic markers show broad gradients spanning from the Levant to northern and western Europe [18]. These clines are generally attributed to the effects of a demographic expansion from the Levant in Neolithic times [21, 22]. However, the likely age of CCR5Δ32 is about 2,000 years [13] or less [28]; consequently, its distribution cannot be due to a Neolithic process. Therefore, there are two possibilities: geographically variable selection, or a more recent migration process.

The CCR5Δ32 mutation may have arisen in populations in the north, and then spread to neighboring groups in different proportions, depending largely on geographic distances. Alternatively, taking into consideration the function of this gene and the age of the allele, it may be that CCR5Δ32 protects carriers from infections other than HIV acquired immune-deficiency syndrome (HIV-AIDS), a premise that is difficult to prove at present. The results presented here are consistent with the idea that climatic factors can play a certain selective role, either directly related to the expression of the CCR5Δ32 allele, or to the action of a pathogen against which that allele confers some degree of protection. Future steps in this research program will include attempts to identify potential selective factors, the effects of which can account for the clinal patterns of *CCR5* variation described in this study.

Acknowledgments

We are grateful to two anonymous reviewers whose comments and suggestions significantly improved the manuscript. This work was partly supported by the Russian Foundation for Basic Research and by State Programs 'Human Genome' and 'Frontiers in Genetics'.

References

1 Alrhatib G, Combadiere C, Brader CC, Fend Y, Kennedy PE, Murphy PM, Berger EA: CC-CCR5:A RANTES, MIP1α, MIP1β receptor as a fusion cofactor for macrophage-tropic HIV-1. Science 1996;272:1955–1958.

2 Choe H, Farzan M, Sun Y, Sullivan N, Rollins B, Ponath PD, Wu L, Mackay CR, LaRosa G, Newman W, Gerard N, Gererd C, Sodroski J: The β-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. Cell 1996;85:1135–1148.

3 Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, Di Marzio P, Marmon S, Sutton RE, Hill SM, Davis CB, Peiper SC, Schall TJ, Littman DR, Landau NR: Identification of a major co-receptor for primary isolates of HIV-1. Nature 1996;381:661–666.

4 Doranz BJ, Rucker J, Yi Y, Smyth RJ, Samson M, Peiper S, Parmentier M, Collman RG, Doms RW: A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptor CKR-5, CKR-3 and CKR-2b as fusion cofactors. Cell 1996;85:1149–1158.

5 Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, Cayanan C, Maddon PJ, Koup RA, Moore JP, Paxton WA: HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. Nature 1996; 381:667–673.

6 Feng Y, Broder CC, Kennedy PE, Berder EA: HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science 1996;272:872–877.

7 Samson M, Soularue P, Vassart G, Parmentier M: The genes encoding the human CC-chemokine receptor CC-CKR1 to 5 are clustered in the p21.3-p24 region of chromosome 3. Genomics 1996;36:522–526.

8 Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R, Goedert JJ, Buchbinder SP, Vittinghott E, Gomperts E, Donfield S, Vlahov D, Kaslow R, Saah A, Rinaldo C, Detels R, Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study, O'Brien SJ: Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CCR5 structural gene. Science 1996;273:1856–1862.

9 Huang Y, Paxton WA, Wolinsky SM, Neumann AU, Zhang L, He T, Kang S, Ceradini D, Jin Z, Yazdanbakhsh K, Kunstman K, Erickson D, Dragon E, Landau NR, Phair J, Ho DD, Koup RA: The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. Nat Med 1996;2:1240–1243.

10 Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, Mac-Donald ME, Stuhlmann H, Koup RA, Landau NR: Homozygous defect in HIV-1 coreceptor accounts for resistance to some multiply-exposed individuals to HIV-1 infection. Cell 1996;86:367–377.

11 Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, Saragosti S, Lapoumeroulie C, Cognaux J, Forceille C, Muyldermans G, Verhofstede C, Burtonboy G, Georges M, Imai T, Rana S, Yi Y, Smyth RJ, Collman RG, Doms RW, Vassart G, Parmentier M: Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. Nature 1996;382:722–725.

12 Martinson JJ, Chapman NH, Rees DC, Liu Y-T, Clegg JB: Global distribution of the CCR5 gene 32-basepair deletion. Nat Genet 1997;16: 100–103.

13 Libert F, Cochaus P, Beckman G, Samson M, Aksenova M, Cao A, Czeizel A, Claustres M, Rua C, Ferrari M, Ferrec C, Glover G, Grinde B, Guran S, Kucinskas V, Lavinha J, Mercier B, Ogur G, Peltonen L, Rosatelli C, Schwartz M, Spitsyn V, Timar L, Beckman L, Parmentier M, Vassart G: The ΔCCR5 mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. Hum Mol Genet 1998;7: 399–406.
14 Slominsky PA, Shadrina MI, Spitsyn VA, Mikulich AI, Khusnutdinova EK, Limborska SA: A simple and rapid method for determining a 32-bp deletion in the gene for the chemokine receptor CCR5. Russian J Genet 1997;33: 1596–1598.
15 Balanovskaya EV, Nurbaev SD, Rychkov YuG: Computer technology of gene-geographic analysis of a gene pool. I. Statistical information of a gene-geographic map. Russian J Genet 1994;30:951–965.
16 Balanovskaya EV, Nurbaev SD, Rychkov YuG: Computer technology of gene-geographic analysis of a gene pool. II. Statistical transformation of maps. Russian J Genet 1994;30: 1538–1555.
17 Balanovskaya EV, Nurbaev SD: Computer technology of gene-geographic analysis of a gene pool. III. Derivation of trend surfaces. Russian J Genet 1995;31:536–559.
18 Cavalli-Sforza LL, Menozzi P, Piazza A: History and Geography of Human Genes. Princeton, Princeton University Press 1995, p 1069.
19 Sokal RR, Oden NL, Thomson BA: A problem with synthetic maps. Hum Biol 1999;71:1–13.
20 Physico-geographical World Atlas. Moscow, 1965, p 295.
21 Chikhi L, Destro-Bisol G, Bertorelle G, Pascali V, Barbujani G: Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. Proc Natl Acad Sci USA 1998;95:9053–9058.
22 Sokal RR, Oden NL, Wilson C: Genetic evidence for the spread of agriculture in Europe by demic diffusion. Nature 1991;351:143.
23 Piazza A, Menozzi P, Cavalli-Sforza LL: The HLA-A,B gene frequencies in the world: Migration or selection? Hum Immunol 1980;1:297–304.
24 Piazza A, Menozzi P, Cavalli-Sforza LL: Synthetic gene frequency maps of man and selective effects of climate. Proc Natl Acad Sci USA 1981;78:2638–2642.
25 Calafell F, Bertranpetit J: Mountains and genes: Population history of the Pyrénées. Hum Biol 1994;66:823–842.
26 Aseev MV, Shawi A, Dean M, Baranov VS: Population features of gene mutation frequency of the CKR-5 receptor, determining sensitivity to the AIDS virus. Russian J Genet 1997; 33:1724–1726.
27 Galeeva AR, Khusnutdinova EK, Slominsky PA, Limborska SA: Distribution of the CCR5 chemokine receptor-gene 32-bp deletion in population of Volga-Ural region. Russian J Genet 1998;34:1160–1162.
28 Stephens JC, Reich DE, Goldstein DB, Shin HD, Smith MW, Carrington M, Winkler C, Huttley G, Allikmets R, Schriml L, Gerrard B, Malasky M, Ramos MD, Morlot S, Tzetis M, Oddoux C, Giovine FS, Nasioulas G, Chandler D, Aseev M, Hanson M, Kalaydjieva L, Glavac D, Gasparini P, Kanavakis E, Claustres M, Kambouris M, Ostrer H, Duff G, Baranov V, Sibul H, Metspalu A, Goldman D, Martin N, Duffy D, Schmidtke J, Estivill X, O'Brien SJ, Dean M: Dating the origin of the CCR5-Δ32 AIDS-resistance allele by the coalescence of haplotypes. Am J Hum Genet 1998;62:1507–1515.
29 Lucotte G, Mercier G: Δ32 mutation frequencies of the CCR5 coreceptor in different French regions. Life Sci 1998;321:409–413.