

# DEVELOPMENT OF MULTIPLEX GENOTYPING METHOD OF POLYMORPHIC MARKERS FOR GENES INVOLVED IN HUMAN ADAPTATION TO COLD CLIMATE

Vadim A. Stepanov<sup>a,b,\*</sup>, Ksenia V. Vagajceva<sup>a,b</sup>, Anna V. Bocharova<sup>a</sup>, Vladimir N. Khar'kov<sup>a,b</sup>

<sup>a</sup>Research Institute of Medical Genetics, Tomsk National Research Medical Center, Ushaiki Nab. 10, Tomsk, 634050 Russian Federation

<sup>b</sup>National Research Tomsk State University, Lenin Ave. 36, Tomsk, 634050 Russian Federation

\* e-mail: vadim.stepanov@medgenetics.ru

Received May 23, 2016;

Accepted in revised form July 26, 2016;

Published December 30, 2016

**Abstract:** Dispersal of the human population out of the place of origin in East Africa over the globe proceeded rapidly (on the evolutionary scale) and was associated with change of climatic zones and associated changes in habitat parameters - temperature, humidity, insolation, and infectious load. The aim of the study is to develop a method for genotyping of genetic markers associated with adaptation to climate according to the literature data and functional analysis of genes; and identification of signals of adaptation to cold climate in two indigenous Siberian populations. In the course of study, genes and genetic markers were selected, which show reliable signals of natural selection in populations living in cold arctic and subarctic climates in previously published papers and which are involved in biological processes having a cold adaptation potential. A panel of 28 single nucleotide markers (SNP) was selected, and a method of their multiplex genotyping was developed based on multiplex PCR and separating DNA fragments by MALDI-TOF mass spectrometry. Allele frequencies of 28 SNPs in two indigenous Siberian populations (Yakuts and Kets) were determined. A low level of intrapopulation diversity in these populations and significant genetic differences between them were found. Loci under natural selection conditions were detected by analyzing the distribution of the observed  $F_{st}$  values in comparison with the expected distribution, obtained in the simulation calculations based on the hierarchical island model of population structure. The possible role of selection ( $p < 0.1$ ) in differentiation of populations between allele frequencies was determined for 2 markers - rs133036 in *MKLI* gene and rs2305508 in *CPT1A* gene, which are candidate in terms of climate change adaptation.

**Keywords:** Single nucleotide polymorphism, MALDI-TOF mass spectrometry, Siberian populations, adaptation to climate, natural selection

DOI: 10.21603/2500-1418-2016-1-2-92-101.

## INTRODUCTION

Dispersal of the human population out of the place of origin in East Africa over the globe proceeded rapidly (on the evolutionary scale), over the past 50–100 thousand years, and was associated with change of climatic zones from tropical to temperate and Arctic and associated changes in habitat parameters - temperature, humidity, insolation, and infectious load. Modern humans entered Eurasian arctic and sub-arctic zones relatively recently (on the evolutionary scale) – 30–40 thousand years ago. Human survival in cold climates could involve mechanisms and processes of different scale and nature – short-term acclimatization, long-term adaptation, and socio-cultural changes. Correlation and role of these mechanisms in survival in extreme climates remain unclear, but phenotypic features of contemporary indigenous populations of northern Eurasia and characteristics of their gene pool suggest a possible role of directional selection in

adaptation of indigenous populations to extreme climatic conditions [1, 2].

A number of current studies aimed at finding signals of natural selection in the human genome and gene pools of individual populations, revealed a large number of polymorphic positions in the genome which were under the influence of selection during human dispersal [3–6]. Some of these markers could be associated with adaptation of Eurasian North population genetic structure to cold climate as a result of natural selection [6].

The above papers are based on genome-wide approaches and on populations of different regions of the world that do not include, as a rule, samples of indigenous populations of Siberia and Northern Asia. The main goal of the study is to select the most significant genetic markers associated with adaptation to cold climate (low temperature) according to genome-wide studies; to develop a method of their multiplex analysis and pilot genotyping of the marker panel in samples from two indigenous Siberian populations.

---

**Please cite this article in press as:** Stepanov V.A., Vagajceva K.V., Bocharova A.V., Khar'kov V.N. Development of multiplex genotyping method of polymorphic markers for genes involved in human adaptation to cold climate. *Science Evolution*, 2016, vol. 1, no. 2, pp. 92–101. doi: 10.21603/2500-1418-2016-1-2-92-101.

**Copyright** © 2016, KemsU. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for any purpose, even commercially, provided the original work is properly cited and states its license. This article is published with open access at <http://science-evolution.ru/>.

## MATERIALS AND METHODS

We used two samples of indigenous Siberian populations - Yakuts and Kets. Yakuts live in the territory of the Republic of Sakha (Yakutia) and speak Turkic language of the Altaic language family. The sample of 102 individuals was collected in the village Byadi, Ust-Aldan region of Yakutia. Kets living in the upper course of the Yenisei River, are one of the smallest indigenous Siberian populations and speak an isolated language, which is the only survived language of the Yenisei family [7, 8]. Sample of Kets (N = 48) was collected in the village of Kellog in Turukhansk District of the Krasnoyarsk Territory. The samples included unrelated individuals non-admixed in three generations, who have signed the informed consent for the study. The study was approved by the ethical committee of the Institute of Medical Genetics.

DNA was isolated by standard methods. Primers were selected using «Sequenom Assay Design» software suite as explained below. Genotyping was performed by multiplex PCR coupled with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF mass spectrometry), with the use of «Sequenom MassARRAY 4» mass spectrometer as described earlier [9]. Multiplex PCR to obtain amplicons containing the analyzed SNPs, was carried out in 96-well plates in a volume of 5 microliters with the use of «Thermo Scientific» and «Applied Biosystems» amplifiers in the presence of 5 ng of genomic DNA matrix and a mixture of PCR primers of the corresponding multiplex. PCR procedure: initial denaturation at 94°C (5 min), then 42 cycles of amplification under the following conditions: denaturation at 94°C (20 s), annealing at 56°C (40 seconds), and elongation at 72°C (60 s), after which samples were incubated for 5 min at 72°C.

Statistical analysis of the results obtained was carried out with the use of standard modern population genetic approaches, implemented in «Arlequin» software suite [10] and described previously [11]. Compliance of genotypes distribution to Hardy-Weinberg equilibrium was assessed using an exact Guo and Thompson test. Intra- and interpopulation genetic diversity was assessed by the method of molecular dispersion (AMOVA). Molecular dispersion analysis means that the total molecular diversity (total dispersion) in the group of populations is made up of independent and additive components characterizing dispersion of alleles (haplotypes) within populations, dispersion, which is due to molecular differences within population groups and dispersion, which is due to differences between population groups. The coefficient of gene differentiation, obtained in AMOVA, is an analogue of  $F_{st}$ .

Loci under natural selection were detected by analyzing the distribution of the observed  $F_{st}$  values in comparison with the expected distribution, obtained in the simulation calculations (20,000 simulations) based on the hierarchical island model of the population structure. The model assumes that distribution of locus-specific genetic differentiation values ( $F_{st}$ ) is a function of interpopulation heterozygosity, formed by factors considered

in the island hierarchical model (isolation by distance and gene flow). Loci, which  $F_{st}$  values are on the edges of the distribution are distribution outliers; level of interpopulation differences for such loci was formed by other factors (in particular, by natural selection) against the background of migration and genetic drift [12, 13].

## RESULTS AND DISCUSSION

### *Selection of markers associated with adaptation to cold climate (low temperature) according to the literature.*

Physiological mechanisms of adaptation to cold climate of Northern Eurasia, whose main characteristics are extremely low winter temperatures and a wide range of winter and summer temperatures, affect biological processes such as metabolism of lipids and carbohydrates, thermoregulation, blood pressure regulation, regulation of smooth muscle contractility, response to stress etc. Obviously, genes involved in these biological processes, are prime candidates for a role of targets for natural selection during human dispersal in the North. Several recent studies have identified some genetic systems, genes and markers correlating with climatic parameters (such as temperature) or demonstrating selection signals in northern populations [3–6].

During the first phase, we chose genes and genetic markers, which: a) demonstrate reliable natural selection signals in populations living in cold arctic and subarctic climates according to the published studies, including whole-genome and genome-wide studies [3–6] and are involved in biological processes mentioned above, which are potentially adaptive to cold. Such genes include genes of thermoregulation (*UCP1*, *UCP2*, *UCP3* uncoupling protein genes), genes of response to temperature stress (*HTR2A*, *SLC52A3*, *MYOF*), energy metabolism genes (*THADA*, *CPT1A*, *GATA3*, *TGFB2*, *GNGT1*, *ITPR3*, *GNG2*, *LEPR*), lipid metabolism genes (*LPA*, *LIPT2*), genes of regulation of smooth muscle contractility (*PRKG1*, *MKLI*), blood pressure regulation genes (*ASIC2*), signal transduction and cell-cell interactions genes (*ITGB8*, *CNTNAP2*, *NRG1*, *RPTOR*, *MAPK1*) as well as some others. At the first phase, about 100 SNPs in more than 40 genes were selected.

### *Development of a multiplex SNP genotyping method.*

One of the efficient modern methods of genotyping of panels of markers, consisting of dozens of SNPs, is the method of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF mass spectrometry). Previously, we have developed a number of multiplex panels of SNPs, based on MALDI-TOF, with capacity of 30–70 markers for the purposes of human population genetics, associative genetic research and DNA identification [9, 14–16]. Based on this experience, we chose mass spectrometry of DNA molecules in order to develop a method of multiplex genotyping of markers associated with adaptation to cold climate.

We formed a multiplex of 28 SNPs in 23 genome regions out of 100 SNPs sampled at the first phase with the use of «Sequenom Assay Design» software suite available on-line at website [www.sequenom.co](http://www.sequenom.co). List of selected markers, their genes (genomic regions) and their characteristics are given in Table 1.

**Table 1.** Markers included in the multiplex and characteristic of their genomic regions

No.	Reference number	Substitution	Chromosome	Chromosome position	Gene	Position in gene	Gene name	Gene product function
1	rs10509653	C/T	10	93366834	MYOF	intron	mioferlin	mioferlin belongs to a family of ferlin-1-like proteins, contains multiple C2 calcium-binding domains and C-terminal transmembrane domain responsible for protein membrane binding. Mioferlin is required at a confluence of myoblasts with muscle fibers in processes of rhabdomyogenesis (differentiation and growth of striated skeletal muscle tissue). Mioferlin protein gene knockout in mice leads to muscle atrophy.
2	rs10887778	C/T	10	88093685	LOC105378415	intron	RNA-encoding gene LOC105378415	uncharacterized gene
3	rs12446160	A/G	16	48324457	LONP2	intron	peroxisomal lon protease 2	Lon protease protein-2 belongs to the mitochondrial ATP-dependent serine proteases. It may play a role in the response to heat shock.
4	rs12946049	C/T	17	80620869	RPTOR	intron	regulatory protein associated with the mammalian rapamycin target, complex 1	protein - a component of signaling pathway that regulates cell growth in response to the presence of nutrients and insulin in blood..
5	rs12979860	C/T	19	39248147	IFNL4	intron	interferon, lambda 4	interferon lambda 4 protein, which binds with receptors and starts a cascade of Jak-STAT phosphorylation reactions and ultimately modulates activity of a group of interferon-stimulated genes (ISGs), which leads to the response of cells in response to viral action, mainly related to the family of rotaviruses.
6	rs133036	C/T	22	40616434	MKL1	intron	gene of acute megakaryoblastic leukemia (translocation) 1	this protein is a cofactor of «whey response protein». Translocation [t (1; 22) (p13; q13);] chromosome 1 to chromosome 22 often occurring at acute megakaryoblastic leukemia, affect MKL1 and RBM15 genes and leads to formation of RBM15/MKL1 pathological fusional protein.
7	rs1513687	C/T	18	39724791	LINC00669	intron	extended intergenic non-protein coding RNA 669	unknown function
8	rs16984239	A/C	2	18053180	KCNS3 – RDH14	intergenic	potassium voltage-gated channel modifier subfamily S, member 3 Retinol dehydrogenase 14	KCNS3 gene encodes a protein - subunit of potassium channel. Modulates channel activity and reduces ion flow. RDH14 gene encodes a protein that exhibits oxidoreducing catalytic activity towards retinoids.

**Table 1.** *Continued.* Markers included in the multiplex and characteristic of their genomic regions

No.	Reference number	Substitution	Chromosome	Chromosome position	Gene	Position in gene	Gene name	Gene product function
9	rs1800592	A/G	4	140572807	UCP1	intergenic	thermogenin, uncoupling protein 1, RB-1	Encodes a protein found in the mitochondria of brown adipose tissue adipocytes. Used as the main mechanism of heat production in neonates and hibernating mammals.
10	rs1800849	C/T	11	74009120	UCP3	5' UTR	thermogenin, uncoupling protein 3, RB-3 (RB-1 homolog)	Encodes the mitochondrial transport protein that produces a current of protons across the inner mitochondrial membrane, thereby disconnecting the oxidative phosphorylation. As a result, the energy is dissipated as heat. It may play a role in the modulation of a respiratory process in tissues. Participates in thermogenesis and energy balance.
11	rs1805490	A/G	12	13588137	GRIN2B	intron	glutamate ionotropic receptor, N-methyl-D aspartate 2B	gene encoding NR2B-subunit of NMDA-glutamate receptor. NR2B subunit participates in the mechanisms of memory and chronic pain.
12	rs2015865	A/C	22	43393211	LOC105373053 / LOC105373054	intergenic	RNA-encoding gene LOC105373053 / RNA-encoding gene LOC105373054 (RNA gene)	uncharacterized gene
13	rs2193045	A/G	12	68140740	HNRNPA1P70 / LOC100509370	intergenic	heterogeneous nuclear ribonucleoprotein A1 pseudogene 70 / ribosomal protein L21, mitochondrial pseudogene	uncharacterized gene
14	rs2216163	C/T	12	68137176	HNRNPA1P70 / LOC100509370	intergenic	heterogeneous nuclear ribonucleoprotein A1 pseudogene 70 / ribosomal protein L21, mitochondrial pseudogene	uncharacterized gene
15	rs2273428	C/T	6	69933126	COL19A1	intron	/ collagen, type XIX, $\alpha$ 1	collagen (type XIX, $\alpha$ 1) is a type of fibril-associated collagens (FACIT).
16	rs2283792	G/T	22	21776836	MAPK1	intron	mitogen-activated protein kinase 1	protein of kinase family that respond to extracellular stimuli (mitogens) and regulates various cellular processes (gene expression, division, differentiation and apoptosis).
17	rs2298432	A/C	22	21768900	MAPK1	intron	mitogen-activated protein kinase 1	See no.16
18	rs2301727	C/T	7	20381596	ITGB8	intron	$\alpha$ V $\beta$ 8 integrin beta subunit,	membrane protein, glycoprotein of integrin superfamily, is a receptor for fibronectin.

**Table 1.** *End.* Markers included in the multiplex and characteric of their genomic regions

No.	Reference number	Substitution	Chromosome	Chromosome position	Gene	Position in gene	Gene name	Gene product function
19	rs2305508	A/G	11	68782078	CPT1A	intron	hepatic carnitine palmitoyltransferase 1A	a protein that plays a role in mitochondrial fatty acid beta-oxidation, namely catalyzes the transfer of the acyl group of long-chain fatty acid conjugates onto carnitine.
20	rs3741135	C/T	11	74002915	UCP3	intron	thermogenin, uncoupling protein 3, RB-3 (RB-1 homolog)	See no.10
21	rs4507607	A/G	6	134088661	SLC2A12 / NR_037938.1	intergenic	family of 2 glucose transport proteins, isoform 12 High Mobility Group AT-Hook 1 Pseudogene 7 / High Mobility Group AT-Hook 1 Pseudogene 7 /	encodes SLC2A12 membrane protein, responsible for the glucose transport across the cell membrane. Is expressed mainly in neurons (which is believed to be the main glucose transporter isoform) and in placenta / Unknown function
22	rs4930248	C/T	11	68798436	CPT1A	intron	hepatic carnitine palmitoyltransferase 1A	See no.16
23	rs4944925	A/G	11	74648484	POLD3	intron	DNA polymerase auxiliary subunit, Delta 3.	a protein that is required for optimal DNA polymerase delta activity, involved in nuclear genome DNA replication, as well as in formation of the DNA ligase 1 complex.
24	rs608343	A/G	11	68429362	LRP5	intron	low-density lipoprotein receptor-related protein 5.	This gene encodes a transmembrane low-density lipoprotein receptor that binds and internalizes ligands in the process of receptor-mediated endocytosis This protein plays a key role in skeletal homeostasis and many bone density related diseases are caused by mutations in this gene Mutations in this gene also cause familial exudative vitreoretinopathy.
25	rs6724627	A/G	2	170651752	MYO3B	intron	myosin IIIB	This gene encodes a fibrillar protein, one of the main components of contractile muscle fibers – myofibril. It has ATPase activity.
26	rs809812	A/G	10	93448022	MYOF	intron	mioferlin	See no.1
27	rs892878	C/T	2	137114290	THSD7B	intron	thrombospondin, type I, domain containing 7B	one of the adhesive proteins that regulate interactions between cells and extracellular matrix. After synthesis, thrombospondin-1 molecules are combined into homotrimers interacting with multiple receptors.
28	rs908394	A/G	12	59119000	LOC105369791	intron	RNA-encoding gene LOC105369791	uncharacterized gene

Further we selected two PCR primers (forward and reverse) and extension primer for IPLEX reaction for each marker. Multiplex PCR to obtain amplicons containing analyzed SNPs, was carried out with the use of «PCR Accessory Set» reagent kit, which is a part of the kit for genotyping by MALDI-TOF mass spectrometry from «Agena Bioscience» (USA). Subsequent stages of the experiment – SAP reaction, IPLEX reaction, transfer of samples to spectro-chip, samples ionization and mass spectra analysis was carried out as described earlier [9]. SAP-reaction is the dephosphorylation of deoxynucleotide triphosphates (dNTPs) not included in amplicons with alkaline phosphatase (SAP). IPLEX reaction involves multiplex extension of prolongin primers by adding a modified dideoxytriphosphate (ddNTP) with modified mass, complementary to nucleotide located in polymorphic position of each SNP. The result is a mixture of short PCR products corresponding to alleles represented in multiplex of polymorphic markers. Reaction mixture purification of salts accumulated in the course of multiplex PCR and iPLEX reaction was performed by cationic resin «SpectroCLEAN» («Agena Bioscience», Germany) treatment. Sample transfer to spectro-chip and ionization of samples was carried out at «Sequenom MassARRAY 4». Real-time mass spectra analysis, primary processing and experiment results registration were carried out with the use of «MassARRAY TYPER 4.0» software suite («Agena Bioscience», Germany).

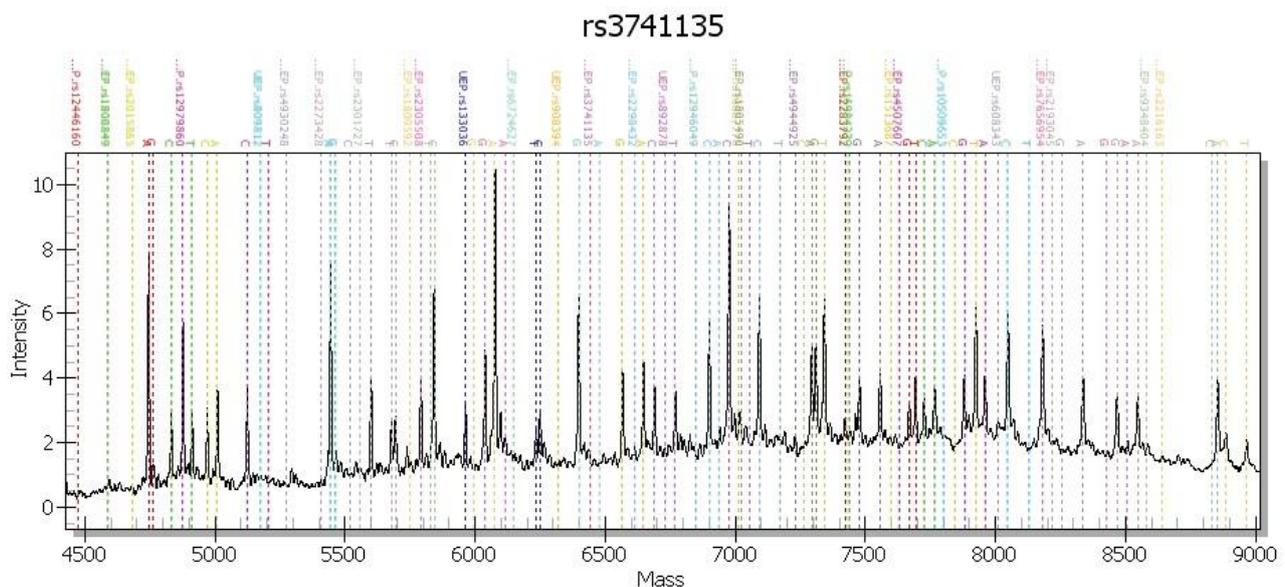
An example of mass spectrum for multiplex consisting of 28 SNPs is shown in Fig. 1., Fig. 2 shows a fragment of mass spectrum of three samples, demonstrating three different genotypes for one SNP (rs3741135). Fig. 3 shows clustering of samples between three different genotypes at genotyping of rs3741135.

**Allele frequencies of SNPs, localized in climate adaptation candidate genes and genetic diversity in populations of Yakuts and Kets.**

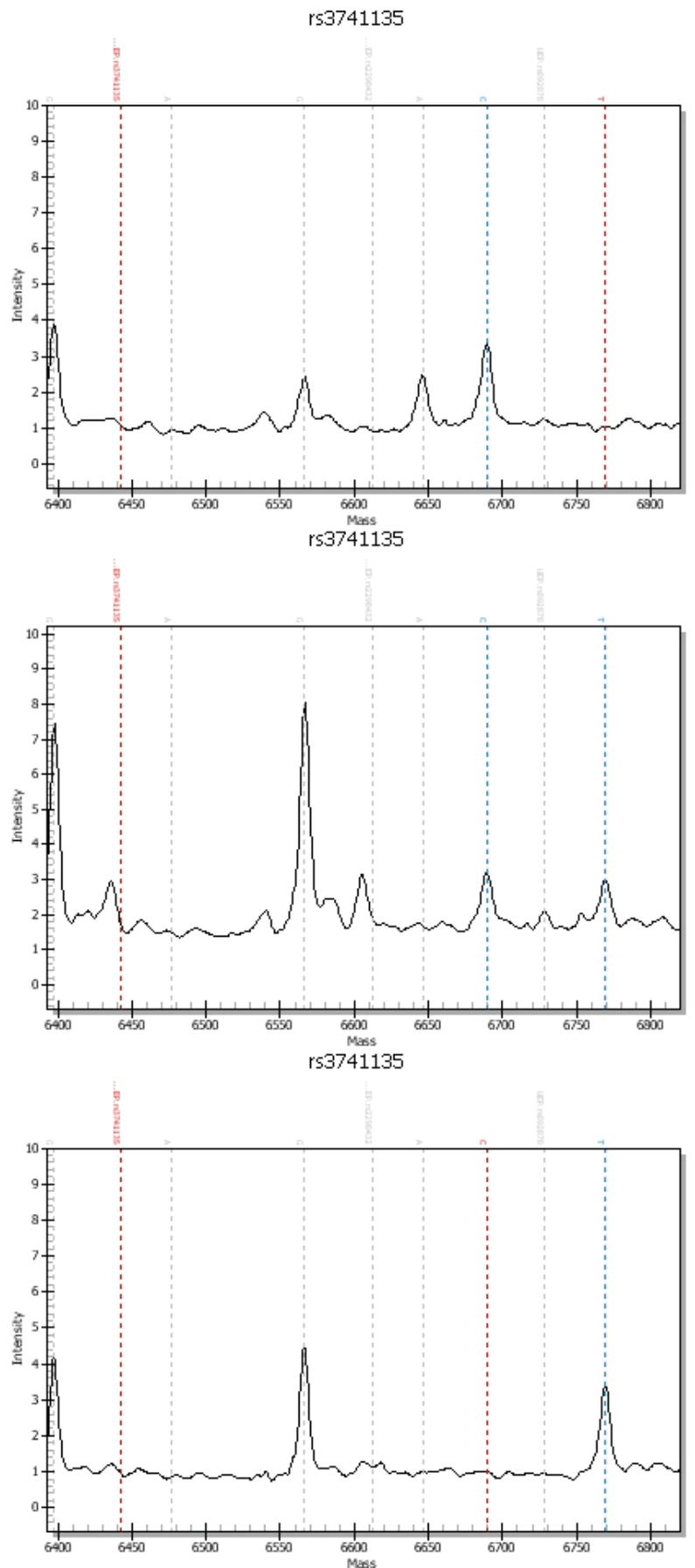
Allele frequencies of 28 SNPs, potentially associated with adaptation to cold climate (low temperature) in populations of Yakuts and Ket are listed in Table 2. Genotype frequencies are available from the authors upon request. Distribution of genotypes in both populations complies with the Hardy-Weinberg equilibrium except for rs6724627 and rs809812 for Kets and rs908394 for Yakuts. A number of deviations from the equilibrium (3 of 56 distributions) does not exceed the expected number of random deviations.

Both studied populations demonstrate sufficiently low intrapopulation genetic diversity – 0.36–0.37 (Table 2) that corresponds with data for other systems of genetic markers for these populations [17–19]. Analysis of genetic differentiation of Kets and Yakuts by 28 markers revealed significant differences between populations with respect to allele frequencies. The overall level of genetic differentiation was 4.3% ( $F_{ST} = 0.04313$ ,  $p = 0.00000$ ). Locus-specific values of genetic difference index range from -0.09 for rs2216163 in the gene region of *HNRNPA1P70 / LOC100509370* to 0.34 for rs908394 in the *LOC105369791* locus (Table 2). In addition to last one, a very high genetic heterogeneity (10%) was found for rs133036 in *MKLI* gene, rs1800849 in *UCP3* gene and rs2305508 in *CPTIA* gene.

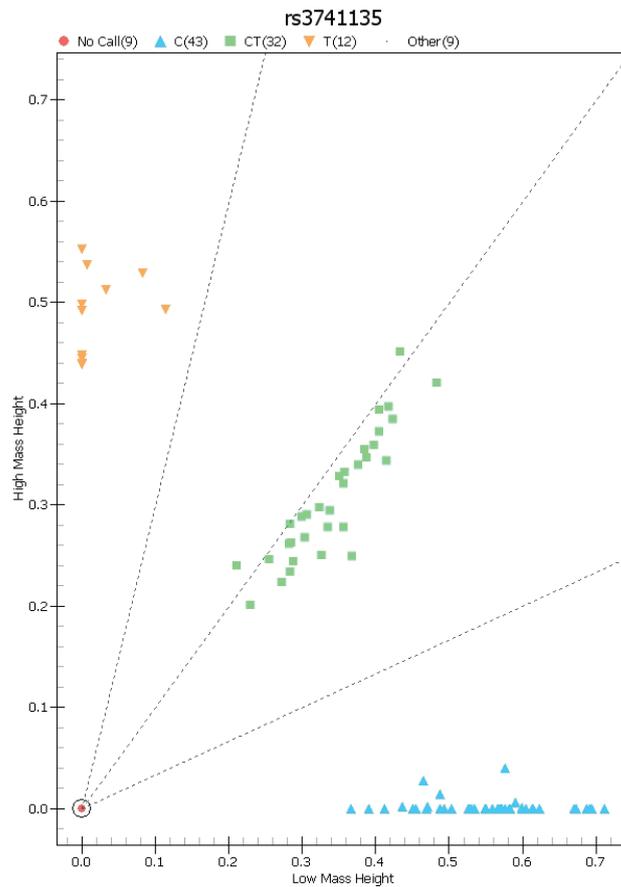
Loci under natural selection were detected by analyzing the distribution of the observed values of  $F_{ST}$  in comparison with the expected distribution, obtained in the simulation calculations (20.000 simulations) based on the hierarchical island model of the population structure [13, 20–22]. The possible role of selection ( $p < 0.1$ ) in differentiation of allele frequencies between populations was revealed for 2 markers - rs133036 in *MKLI* gene ( $p = 0.01$ ) and rs2305508 in *CPTIA* gene ( $p = 0.09$ ). The functional characteristic of one of these genes, *CPTIA*, whose product is involved in the processes of thermoregulation and energy metabolism, makes it an obvious target for selection in northern



**Fig. 1.** Primary mass spectrum for multiplex 1. The mass spectrum may have a total of 84 peaks (3 peaks for each SNP) - prolonged primer peak, peak 1 allele (prolonging primer + nucleotide corresponding to allele 1), allele peak 2 (prolonging primer + nucleotide corresponding to allele 2).



**Fig. 2.** Three different genotypes for one rs3741135. Homozygote CC (top), heterozygote CT (in the middle), homozygote TT (at the bottom). Positions of the peaks for rs3741135 presence in the sample are highlighted in blue dotted line. Brown color dotted line indicates the prolonging primer peak position and peak positions for rs3741135 alleles missing in the sample.



**Fig. 3.** Samples clustering between three different genotypes for rs3741135 genotyping. Blue triangles – CC homozygotes, green squares – CT heterozygotes, brown triangles – TT homozygotes The zero position sample (red circle) – the control sample without DNA.

**Table 2.** Allele frequencies of 28 SNPs, genetic diversity (average expected heterozygosity,  $H_e$ ) and genetic differentiation ( $F_{ST}$ ) parameters and significance level of inter-population differences ( $F_{ST} P$ ) in the Ket and Yakut populations

SNP	Ancestral allele	Kets	Yakuts	$F_{ST}$	$F_{ST} P$
rs10509653	T	0.3617	0.2245	0.03907	0.01466
rs10887778	C	0.5000	0.5153	-0.00793	1.00000
rs12446160	G	0.0326	0.0459	-0.00551	1.00000
rs12946049	T	0.5682	0.6237	-0.00166	0.58065
rs12979860	T	0.2021	0.0918	0.04490	0.02053
rs133036	T	0.4651	0.8093	0.23766	0.00000
rs1513687	C	0.4468	0.5567	0.01639	0.10166
rs16984239	C	0.7174	0.5266	0.06524	0.00293
rs1800592	A	0.7234	0.5714	0.03963	0.02248
rs1800849	C	0.6522	0.3939	0.11586	0.00000
rs1805490	G	0.5652	0.6134	-0.00346	0.70381
rs2015865	C	0.6848	0.7350	-0.00189	0.60606
rs2193045	A	0.6196	0.5306	0.00777	0.19844
rs2216163	C	0.3804	0.4643	0.00592	0.26491
rs2273428	T	0.5745	0.5206	-0.00130	0.57771
rs2283792	G	0.4111	0.2418	0.05850	0.00098
rs2298432	C	0.5543	0.4439	0.01655	0.09775
rs2301727	C	0.9130	0.8061	0.03306	0.03910
rs2305508	A	0.5213	0.7474	0.10261	0.00000
rs3741135	C	0.6522	0.4750	0.05188	0.01369

**Table 2.** *End.* Allele frequencies of 28 SNPs, genetic diversity (average expected heterozygosity,  $H_e$ ) and genetic differentiation ( $F_{ST}$ ) parameters and significance level of inter-population differences ( $F_{ST}P$ ) in the Ket and Yakut populations.

SNP	Ancestral allele	Kets	Yakuts	$F_{ST}$	$F_{ST}P$
rs4507607	A	0.0638	0.0859	-0.00465	1.00000
rs4930248	C	0.1170	0.3300	0.10243	0.00000
rs4944925	G	0.9239	0.9175	-0.00714	1.00000
rs608343	G	0.1413	0.2263	0.01472	0.11046
rs6724627	A	0.1889	0.0765	0.05264	0.01369
rs809812	G	0.1087	0.1378	-0.00411	1.00000
rs892878	C	0.6596	0.8163	0.05844	0.00293
rs908394	A	0.1304	0.0408	0.04992	0.03812
$H_e$		0.3721	0.3609	0.04313	0.00000

populations, whereas a mechanism of possible adaptive importance of variants of the second gene, *MKLI*, is not clear.

Thus, this paper presents a new panel of genetic markers associated with adaptation of human populations to low temperatures and a cold climate. The multiplex panel genotyping method was developed, which was tested in the analysis of allele frequencies and genetic diversity in two indigenous Siberian populations – Kets and Yakuts. Comparison of the populations studied revealed a high level of inter-

population differences against the background of relatively low intra-population diversity. The possible role of natural selection in formation of the allele frequency spectrum through markers of two genes involved in mechanisms that ensure resistance to low temperatures is revealed.

#### ACKNOWLEDGEMENT

The research was supported by Russian Foundation for Basic Research (grant no. 15–04–02442) and TSU Competitiveness program.

#### REFERENCES

- Stepanov V.A. Geneticheskoe raznoobrazie i bolezni cheloveka. [Evolution of genetic diversity and human diseases]. *Genetika* [Russian Journal of Genetics], 2016. vol. 52, no. 7, pp. 746–756.
- Stepanov V.A. Geneticheskoe raznoobrazie populyatsiy cheloveka i problemy evolyutsionnoy meditsiny. [Genetic diversity of human populations and problems of evolutionary medicine]. *Genetika cheloveka i patologiya. Problemy evolyutsionnoy meditsiny* [Human genetics and pathology. Problems of evolutionary medicine], 2014, no. 10, pp. 7–17.
- Lappalainen T., Salmela E., Andersen P.M., Dahlman-Wright K., Sistonen P., Savontaus M-L, Schreiber S., Lahermo P., Kere J. Genomic landscape of positive natural selection in Northern European populations. *European Journal of Human Genetics*, 2010, no. 18, pp. 471–478. doi: 10.1038/ejhg.2009.184.
- Rasmussen M., Li Y., Lindgreen S., Pedersen J.S., Albrechtsen A., Moltke I., Metspalu M., Metspalu E., Kivisild T., Gupta R., Bertalan M. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature*, 2010, no. 463, pp. 757–762. doi: 10.1038/nature08835.
- Hancock A.M., Witonsky D.B., Alkorta-Aranburu G., Beall C.M., Gebremedhin A., Sukernik R. Adaptations to Climate-Mediated Selective Pressures in Humans. *PLoS Genet*, 2011, vol. 7, no. 4: e1001375. doi:10.1371/journal.pgen.1001375.
- Cardona A., Pagani L., Antao T., Lawson D.J., Eichstaedt C.A., et al. Genome-Wide Analysis of Cold Adaptation in Indigenous Siberian Populations. *PLoS ONE*, 2014, vol. 9, no. 5: e98076. doi: 10.1371/journal.pone.0098076.
- Ruhlen M. *A guide to the World languages*. Vol. 1. Classification. Stanford: Stanford University Press. 1991. 464p.
- Bruk S.I. *Naselenie mira. Etno-demograficheskiy spravochnik* [Population of the World. Ethno- Demographic Handbook]. Moscow: Nauka Publ., 1986. 832 p.
- V.A. Stepanov, E.A. Trifonova. Mul'tipleksnoe genotipirovanie odnonukleotidnykh polimorfnykh markerov metodom MALDI-TOF mass-spektrometrii: chastoty 56 SNP v genakh immunnogo otveta v populyatsiyakh cheloveka [Multiplex SNP genotyping by MALDI-TOF mass spectrometry: Frequencies of 56 immune response gene SNPs in human populations]. *Molekulyarnaya Biologiya* [Molecular Biology], 2013, vol. 47, no. 6, pp. 976–986.
- Excoffier L., Lischer H.E. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 2010, vol. 10, pp. 564–567.
- Stepanov V.A., Khar'kov V.N., Trifonova E.A., Marusin A.V. Metody statisticheskogo analiza v populyatsionnoy i evolyutsionnoy genetike cheloveka [Methods of Statistical Analysis in Human Population and Evolutionary Genetics]. *Uchebno-metodicheskoe posobie Seriya «Nasledstvennost' i zdorov'e»* [Study Guide Series: «Human heredity and Health»]. Tomsk: Pechatnaya Manufaktura Publ., 2014. 100 p.
- Beaumont M.A., Nichols R.A. Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society London B*, 1996, vol. 263, pp. 1619–1626.
- Excoffier L., Hofer T., Foll M. Detecting loci under selection in a hierarchically structured population. *Heredity*, 2009, vol. 103, pp. 285–298.
- Stepanov V., Vagaitseva K., Kharkov V., Cherednichenko A., Bocharova A., Berezina G., Svyatova G. Forensic and population genetic characteristics of 62 X chromosome SNPs revealed by multiplex PCR and MALDI-TOF mass spectrometry genotyping in 4 North Eurasian populations. *Legal Medicine*, 2016, vol. 18, pp. 66–71. doi: 10.1016/j.legalmed.2015.12.008.

15. Stepanov V.A., Vagaitseva K.V., Kharkov V.N., Cherednichenko A.A., Bocharova A.V. Panel' odnonukleotidnykh stseplennykh s Kh-polimorfnykh markerov dlya DNK-identifikatsii (Xsnpid) na osnove mul'tipleksnogo genotipirovaniya s ispol'zovaniem metodov mnogolokusnoy PTsR i mass-spektrometrii MALDI-TOF [Panel of X-linked single-nucleotide polymorphic markers for DNA identification (XSNPid) based on multiplex genotyping by multilocus PCR and MALDI-TOF mass spectrometry]. *Molekulyarnaya biologiya* [Molecular Biology], 2016, vol. 50, no. 3, pp. 445–456.
16. Stepanov V.A., Vagaitseva K.V., Kharkov V.N., Cherednichenko A.A., Minajcheva L.I., Bocharova A.V. Variabel'nost' i identifikatsionnyy potentsial Kh-khromosomnykh SNPs v dvukh populyatsiyakh koren'nogo naseleniya Sibiri [Variability and identification power of 60 X-chromosome in two native Siberian populations]. *Genetika* [Russian Journal of Genetics], 2016, vol. 52, no. 4, pp. 493–496.
17. Kharkov V.N., Stepanov V.A., Medvedeva O.F., Spiridonova M.G., Maksimova N.R., Nogovicyna A.N., Puzyrev V.P. Proiskhozhdenie yakutov: analiz gaplotipov Y-khromosomy [The origin of Yakuts: Analysis of the Y-chromosome haplotypes]. *Molekulyarnaya biologiya* [Molecular biology], 2008, vol. 42, no. 2, pp. 226–237.
18. Khitrinskaya I.Yu., Khar'kov V.N. Voevoda M.I., Stepanov V.A. Geneticheskoe raznoobrazie i vzaimootnosheniya populyatsiy Severnoy Evrazii po polimorfnykh insertsiyam Alu-elementa [Genetic diversity and relationships of populations of northern Eurasia by polymorphic Alu insertions]. *Molekulyarnaya biologiya* [Molecular Biology], 2014, vol. 48, no. 1, pp. 69–80.
19. Flegontov P., Changmai P., Zidkova A., Logacheva M.D., Flegontova O., Nikolsky YuV, Starostin G., Stepanova V.V., Travinsky I.V., Triska M., Triska P., Tatarinova T.V. Genomic study of the Ket: a Paleo-Eskimo-related ethnic group with significant ancient North Eurasian ancestry. *Scientific Reports*, 2016. doi: 10.1038/srep20768.
20. Schlotterer C. A microsatellite-based multilocus screen for the identification of local selective sweeps. *Genetics*, 2002, vol. 160, no. 2, pp. 753–763.
21. Beaumont M.A., Balding D.J. Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol*, 2004, vol. 13, pp. 969–980.
22. Foll M, Gaggiotti O.E. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 2008, vol. 180, pp. 977–993.

---

**Vadim A. Stepanov**

Dr.Sci.(Biol.), Professor, Head of the Research Institute of Medical Genetics, Tomsk, Russian Federation

**Ksenia V. Vagajceva**

Cand.Sci.(Biol.), Junior research associate of Laboratory of Evolutionary Genetics, Research Institute of Medical Genetics, Tomsk, Russian Federation

**Anna V. Bocharova**

Junior research associate of Laboratory of Evolutionary Genetics, Research Institute of Medical Genetics, Tomsk, Russian Federation

**Vladimir N. Khar'kov**

Dr.Sci.(Biol.), Leading research worker of Laboratory of Evolutionary Genetics, Research Institute of Medical Genetics, Tomsk, Russian Federation