

<https://doi.org/10.25557/2073-7998.2023.02.52-54>

Сигналы направленного отбора в популяциях коренного населения Средней Азии

Колесников Н.А., Харьков В.Н., Зарубин А.А., Степанов В.А.

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук

634050, г. Томск, ул. Набережная р. Ушайки, д. 10

Популяции коренных этносов Средней Азии представляют значительный интерес для популяционной геномики по причине специфичности их генофондов, развивавшихся в различных генетико-демографических условиях. Данные о сигналах направленного отбора являются важным дополнением к существующим данным о эволюции генофондов и механизмах генетической адаптации населения Евразии. Мы использовали массив генотипов по 1779819 SNP в выборке из 128 человек, включающей 5 популяций населения Средней Азии, для поиска сигналов направленного отбора с помощью тестов на протяженную гомозиготность гаплотипов (nSL, iHS). Обнаружено, что все исследованные популяции сильно отличаются друг от друга по составу генов, которые демонстрируют влияние отбора. Наибольшее число значимых сигналов естественного отбора выявлено в популяциях киргизов и узбеков. Среди локусов генома, несущих наиболее выраженные сигналы направленного отбора в популяциях выделяются гены *CUEDC1*, *ERC1*, *FAM20C*, *MRC1*.

Ключевые слова: популяционная генетика, Средняя Азия, естественный отбор, адаптация, nSL, iHS.

Для цитирования: Колесников Н.А., Харьков В.Н., Зарубин А.А., Степанов В.А. Сигналы направленного отбора в популяциях коренного населения Средней Азии. *Медицинская генетика* 2023; 22(2): 52-54.

Автор для корреспонденции: Колесников Н.А.; **e-mail:** nikita.kolesnikov@medgenetics.ru

Финансирование. Исследование выполнено за счет гранта Российского научного фонда № 22-64-00060, <https://rscf.ru/project/22-64-00060/>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила: 10.02.2023

Signals of directed selection in the indigenous populations of Central Asia

Kolesnikov N.A., Kharkov V.N., Zarubin A.A., Stepanov V.A.

Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences

10, Nab. r. Ushaiki, Tomsk, 634050, Russian Federation

The populations of the indigenous ethnic groups of Central Asia are of considerable interest for population genomics, due to the specificity of their gene pools, which developed under various genetic and demographic conditions. The data on directional selection signals is an important addition to the existing data on the evolution of gene pools and the mechanisms of genetic adaptation of the Eurasian population. We used an array of genotypes for 1779819 SNPs in a sample of 128 people, including 5 populations of the population of Central Asia, to search for directional selection signals using tests for extended haplotype homozygosity (nSL, iHS). For the populations of Central Asia, it was found that all the studied populations differ greatly from each other in the composition of genes that demonstrate the influence of selection. The largest number of significant signals of natural selection was found in the populations of the Kyrgyz and Uzbeks. The *CUEDC1*, *ERC1*, *FAM20C*, and *MRC1* genes are among the genome loci that carry the most pronounced signals of directional selection in populations.

Keywords: population genetics, Central Asia, natural selection, adaptation, nSL, iHS.

For citation: Kolesnikov N.A., Kharkov V.N., Zarubin A.A., Stepanov V.A. Signals of directed selection in the indigenous populations of Central Asia. *Medical genetics [Meditsinskaya genetika]* 2023; 22(2): 52-54. (In Russ.)

Corresponding author: N.A. Kolesnikov; **e-mail:** nikita.kolesnikov@medgenetics.ru

Funding. The study was supported by the Russian Science Foundation grant No. 22-64-00060, <https://rscf.ru/project/22-64-00060/>

Conflict of interest. The authors declare no conflicts of interest.

Accepted: 10.02.2023

Введение

Основные миграции человека современного вида из Африки произошли в течение последних 60-50 тыс. лет [1]. В результате этих миграций люди столкнулись с новой средой с различным климатом, патогенами и продуктами питания и адаптировались к этим новым условиям посредством естественного отбора. Исследование популяций Средней Азии необходимо для выяснения генетических особенностей и естественного отбора в аспекте адаптации генофонда популяций к факторам окружающей среды, изменения которых произошли после миграции современных людей из Африки.

Цель и задачи: провести поиск сигналов естественного отбора в популяциях Средней Азии с использованием биоинформатического анализа.

Методы

Использован массив генотипов по 1779819 SNP в выборке из 128 человек, включающей 5 популяций населения Средней Азии, для поиска сигналов направленного отбора с помощью тестов на протяженную гомозиготность гаплотипов (nSL, iHS). При биоинформатической обработке полученных данных из анализа были исключены SNP с более чем 10% отсутствующих генотипов. Были исключены все позиции делеций и инсерций. Данные были предварительно отфильтрованы по минимальной частоте редкого аллеля (MAF, minor allele frequency > 0,01). В итоге после фильтрации исходных данных по 1779819 маркерам в финальный массив данных вошло 886889 аутоматических SNP. Кроме этого, были исключены образцы, имеющие более 5% пропущенных SNP, а также метисы первого, второго и третьего поколений.

Популяции северных киргизов (N = 28) (северные киргизы представлены жителями г. Бишкек и села Кегеты Чуйского района Республики Киргизия; материал собран сотрудником Киргизской государственной медицинской академии канд. биол. наук Ж.О. Солтобаевой), узбеков (N = 22), казахов (M – младший жуз N = 29, C – старший жуз N = 26) и дунган (N = 23) относятся к популяциям Средней Азии (N = 128). Образцы ДНК входят в состав ДНК-банка НИИ медицинской генетики.

Поиск сигналов направленного отбора на основе анализа данных по высокоплотному набору SNP проводили в пакете SelScan с помощью геномного теста на протяженную гомозиготность гаплотипов (nSL –

number of Segregating sites by Length и iHS – integrated haplotype score) [2]. Ожидаемая доля ложных отклонений FDR (False discovery rate) была больше или равна 0,05.

Результаты

Результаты поиска сигналов, сформированных под действием направленного отбора в изучаемых популяциях, с помощью тестов nSL и iHS приведены в **таблице**. Всего было найдено 74 SNP ($FDR \geq 0,05$), попадающих в функциональные участки 135 генов, и 52 SNP – на межгенные участки для nSL. По тесту iHS найдено 276 SNP ($FDR \geq 0,05$), попадающих в функциональные участки 51 генов, и 170 SNP – на межгенные участки.

Сигналы направленного отбора в генах *CUEDC1*, *ERC1*, *FAM20C*, *MRC1*, выявленные для четырех из пяти исследуемых популяций по тесту iHS (для гена *CUEDC1* для всех за исключением узбеков и *ERC1* только для киргизов по тесту nSL), вероятно, имеют важное значение для адаптации исследуемых популяций. Для генов *CUEDC1* и *ERC1* ранее отмечались сигналы отбора по тесту nSL в популяциях Сибири [3]. *FAM20C* кодирует секреторную киназу, которая фосфорилирует белки секреторного пути в мотивах Ser-x-Glu / pSer и играет ключевую роль в биоминерализации костей и зубов [4]. Помимо своей роли в биоминерализации, он также играет роль в гомеостазе липидов, заживлении ран, миграции и адгезии клеток. Ген *MRC1* кодирует белок, представляющий собой рецептор маннозы C-типа 1, распознающий концевые остатки маннозы, N-Ацетилглюкозамина и фукозы на гликановых цепочках белков [5], расположенных на поверхности некоторых микроорганизмов, и играет роль во врожденном и приобретенном иммунитете. Кроме этого, маннозный рецептор участвует в выведении гликопротеинов из кровотока, включая сульфатированные гликопептидные гормоны и гликопротеины, высвобождающиеся в результате патологических процессов [6].

Отмечены сигналы направленного отбора в генах, отвечающих за пигментацию. В популяции дунган были отмечены сигналы в гене *EDAR* (рецептор эктодисплазии А). Этот ген кодирует рецептор эктодисплазии А, члена семейства рецепторов фактора некроза опухоли (*TNFR*). Он необходим для развития волос, зубов и других производных эктодермы. Многочисленные исследования сообщают о доказательствах избирательного охвата этого гена в популяциях Восточной Азии [7, 8]. У казахов старшего жуза найдены сигналы отбора в гене *MLPH* (Melanophillin). Белок, кодируемый этим геном,

Таблица. Результаты поиска сигналов направленного отбора в популяциях Средней Азии с помощью теста iHS и nSL
Table. Results of the Search for Directional Selection Signals in the Populations of Central Asia Using the iHS and nSL Tests

Популяции	Результаты поиска сигналов направленного отбора с помощью теста iHS и nSL	
	iHS	nSL
Казахи (М)	<i>ACOT7, ARHGAP42, COLEC11, CSMD1, CUEDC1, ELMO1, EPB41L3, ERC1, FSTL1, GRM6, KCNH1, LINC01229, LOC100420587, LOC101927822, LOC101928416, ME2, MPPED1, MRC1, MYO1B, NPHP4, SP3, SULT1C4, TEF, TMEM117, TRABD2A, UM OD</i>	<i>CLMP, CUEDC1, CYP1B1-AS1, DLG2, EPB41L3, LINC01229, LOC100420587, LOC101927822, NPHP4, PXDNL, SDK1, TRABD2A</i>
Казахи (С)	<i>AGFG1, APOL1, COLEC11, CUEDC1, DAB1, ELMO1, ERC1, FAM20C, FHIT, GABRA5, GALNT9, GAS6-AS1, IMMP2L, LINC01497, LIPA, MIR548F3, MIR646HG, MLPH, MRC1, TMEM236, NAV2, POLE2, PPP2R5C, RNASEH1, RNF39, STK10, TMEM255B, TWIST2</i>	<i>BACE2, CUEDC1, DAB1, ELMO1, LIMCH1, LINC00682, NAV2, PPP2R5E, PRKN, RHBDD1, SPRYD3, STK10, TM6IM6, WWOX</i>
Узбеки	<i>ABCA4, ANO3, APOL1, ATP2C1, CDH13, CDIP1, CPXM2, DMRT1, ELAVL3, ELMO1, ERC1, EXOC6B, FAM20C, FBXW8, GALNTL6, HMOX2, HYDIN, IFT140, ISX-AS1, JAKMIP2-AS1, LINC01844, LOC102724084, LOC400655, LOC730100, MPPED1, MRC1, TMEM236, MSRI, MYO10, NAALADL2, NMRAL1, NPAS3, NPSRI-AS1, NPTN, NRG3, NRXN1, OTOGL, PAAFI, PRKAR1B, PRMT8, PRSS1, PSMA8, RAS A3, RIPK1, SETD5, SH3RF1, SLC12A1, SLC30A9, SPATA13, STARD13, TLE4, TMEM255B, TSBP1-AS1, VAV2, VWA3A, ZNF717</i>	<i>ANO3, ATP13A3, C5orf66, CNN1, CYB561A3, DMRT1, ELMO1, NRG3, STARD13</i>
Киргизы	<i>BDH1, CD226, CEACAM18, CTIF, CUEDC1, DDX55, DNAH10, DNAH6, DNAJC5B, DYK, EFCAB6-AS1, ERC1, FAM20C, FYTDD1, HIPK2, HS2ST1, ISX-AS1, LINC02299, LOC101927822, LOC440982, MPPED1, MRC1, PIKFYVE, PTPRG, RNASEH1, RNASEH1-AS1, SLC2A9, SLC30A9, SRP54-AS1, STK10, TMEM117, TMEM255B, TRABD2A, TLL1</i>	<i>BDH1, CADM3, CDH13, COL21A1, CUEDC1, ERC1, FBXO25, LAMA1, LIMCH1, LINC02147, LOC101927822, PPP2R5C, RNF14, STK10, TRABD2A</i>
Дунгане	<i>ADI1, CADPS, CCDC138, CCHCR1, CENPT, CPEB4, CREB5, CUEDC1, D2HGDH, DOCK7, DOCK8, EDAR, EFCAB6-AS1, FAM20C, GH2, GRID1, HDAC7, HSPC324, IFI30, KCNK10, KCNQ1OT1, LIMS1, LOC102724659, LRRC4C, MEG8, MRC1, NRN1L, NUTF2, PAPA2, PCDHGA1, PCDHGA2, PCDHGA3, PCDHGA4, PCDHGA5, PCDHGA6, PCDHGA7, PCDHGA8, PCDHGB1, PCDHGB2, PCDHGB3, PCDHGB4, PDE3A, PPT2-EGFL8, RANBP10, RNF14, SCHLAP1, SIK3, SMC6, SPTLC2, SULT1C4, SV2B, TAF11, THAP4, TMEM117, TMEM236, TMEM255B, TSBP1-AS1, TSNAIPI, ULK4, VPS37C, XKR6, ZBTB38, ZKSCAN7, ZNF322</i>	<i>BDH1, CENPT, CUEDC1, CYP1B1-AS1, EDAR, LINC00448, PRKN, SHC3, SV2B, TSNAIPI, XKR6, ZNF322</i>

представляет собой эффекторный белок, участвующий в миграции меланосом. Ранее уже сообщалось о наличии сигнатур селекции в области *MLPH* [8, 9].

Выводы

В результате данного исследования выявлено существенное число сигналов направленного естественного отбора, вероятно, повлиявших на генетическую структуру популяций Средней Азии в процессе долговременной адаптации к климато-географическим условиям обитания в резко континентальном климате.

Литература/References

- Henn B.M., Cavalli-Sforza L.L., Feldman M.W. The great human expansion. *Proc Natl Acad Sci U S A.* 2012;109(44):17758.
- Szpiech Z.A., Hernandez R.D. Selscan: An efficient multithreaded program to perform EHH-based scans for positive selection. *Molecular Biology and Evolution.* 2014;31(10):2824–2827.
- Kolesnikov N.A., Kharkov V.N., Zarubin A.A., et al. Signals of Directed Selection in the Indigenous Populations of Siberia. *Russian Journal of Genetics.* 2022;58(4):473–477.
- Tagliabracci V.S., Engel J.L., Wen J. et al. Secreted Kinase Phosphorylates Extracellular Proteins That Regulate Biomineralization. *Science.* 2012;336(6085):1150.
- Schlesinger P.H., Doebber T.W., Mandell B.F., et al. Plasma clearance of glycoproteins with terminal mannose and N-acetylglucosamine by liver non-parenchymal cells. Studies with beta-glucuronidase, N-acetyl-beta-D-glucosaminidase, ribonuclease B and agalacto-orosomucoid. *Biochemical Journal.* 1978;176(1):103.
- Lee S.J., Evers S., Roeder D., et al. Mannose receptor-mediated regulation of serum glycoprotein homeostasis. *Science.* 2002;295(5561):1898–1901.
- Tang K., Thornton K.R., Stoneking M. A New Approach for Using Genome Scans to Detect Recent Positive Selection in the Human Genome. *PLOS Biology.* 2007; 5(7):e171.
- Myles S., Somel M., Tang K., et al. Identifying genes underlying skin pigmentation differences among human populations. *Human Genetics.* 2007;120(5):613–621.
- Pickrell J.K., Coop G., Novembre J., et al. Signals of recent positive selection in a worldwide sample of human populations. *Genome Research.* 2009;19(5):826.