

ОСОБЕННОСТИ РАСПРЕДЕЛЕНИЯ РЕГИОНОВ ВЫСОКОЙ ГОМОЗИГОТНОСТИ В ГЕНОМАХ ПРЕДСТАВИТЕЛЕЙ КОРЕННОГО НАСЕЛЕНИЯ СЕВЕРНОЙ ЕВРАЗИИ

© 2019 г. Н. А. Колесников¹, *, В. Н. Харьков¹, А. А. Зарубин¹, В. А. Степанов¹

¹Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, 634050 Россия

*e-mail: nikita.kolesnikov@medgenetics.ru

Поступила в редакцию 22.01.2019 г.

После доработки 19.03.2019 г.

Принята к публикации 15.04.2019 г.

Популяции коренных этносов Северной Евразии представляют значительный интерес для популяционной геномики как в силу относительно слабой их изученности с привлечением современных геномных технологий, так и по причине специфичности их генофондов, развивавшихся в различных генетико-демографических условиях. Мы использовали данные генотипов по 242179 аутосомным SNP у 876 человек для поиска регионов высокой гомозиготности протяженностью более 1 Мб. Для популяций Сибири обнаружено большее общее количество и большая суммарная длина регионов высокой гомозиготности, чем для других исследованных популяций Северной Евразии.

Ключевые слова: популяционная генетика, Сибирь, регионы высокой гомозиготности, популяции человека, Северная Евразия.

DOI: 10.1134/S0016675819100072

Генетико-демографические процессы и естественный отбор в популяциях человека оказывают влияние на структурированность генетического разнообразия в геномах отдельных индивидов и популяциях в целом. Они, в частности, приводят к тому, что формируются особенности распределения регионов высокой гомозиготности (англ. ROH – runs of homozygosity) между отдельными людьми и популяциями, что обеспечивает ценный и часто неиспользуемый ресурс для изучения генетического разнообразия человечества и его эволюционной истории [1, 2].

Число и общая длина ROH у каждого человека показывают значительные различия между отдельными людьми и группами населения. Сумма длин коротких и средних ROH у индивида возрастает с увеличением расстояния популяций от Восточной Африки, коррелируя с процессами расселения человека [1]. Напротив, общий суммарный размер длинных ROH показывает большие межиндивидуальные вариации, которые, вероятно, отражают события недавнего инбридинга, причем более высокие значения чаще встречаются в популяциях с известными высокими частотами кровнородственных браков. Регионы высокой гомозиготности также связаны с эффективным размером популяции, небольшие популяции имеют тенденцию к накоплению большего числа ROH по срав-

нению с более многочисленными группами населения. Инбридинг добавляет небольшое количество очень длинных ROH, тогда как эффект бутылочного горлышка увеличивает количество ROH, которые, как правило, остаются относительно короткими. Наконец, наибольшая длина и количество ROH наблюдаются в популяциях с относительно небольшим эффективным размером в прошлом и недавним инбридингом [3, 4].

Территория Северной Евразии обладает уникальными географическими характеристиками с высоким уровнем этнической и лингвистической дифференциации населения, а также в этом регионе исторически имело место переплетение коренных и пришлых генетических элементов, проникших сюда с соседних территорий. Популяции коренного населения Сибири характеризуются, как правило, небольшой численностью и значительной географической изоляцией от других популяций, что находит отражение в структуре их генетического разнообразия [5–9]. Задачей настоящей работы была оценка распределения регионов высокой гомозиготности в популяциях различных территорий Северной Евразии.

Мы использовали генотипы 876 индивидуумов по 242179 аутосомным SNP, опубликованные в работе Triska et al. [5]. Популяции коренного населения Сибири представлены кетами ($N = 26$),

Таблица 1. Средняя суммарная длина ROH (Mb), а также доля человек с одним или несколькими ROH в исследуемых популяциях по диапазонам длин (Mb) для каждого региона

Диапазон длин	Сибирь	Кавказ	Средняя Азия	Европа	Волго-Уральский регион
1–2	15.14	4.40	3.69	6.65	7.45
	100%	100%	100%	100%	100%
2–3	11.40	3.24	2.34	3.46	5.09
	100%	100%	100%	100%	100%
3–4	6.99	1.41	1.25	1.58	2.80
	74%	48%	35%	33%	30%
4–5	5.61	1.03	0.64	1.09	2.23
	66%	37%	17%	16%	15%
5–6	4.75	0.88	0.51	0.51	1.65
	51%	18%	8%	20%	8%
6–8	5.78	1.07	0.62	1.04	1.71
	42%	23%	11%	18%	13%
8–10	3.93	1.16	0.55	0.59	1.12
	33%	11%	6%	10%	6%
>10	14.27	1.58	4.09	2.59	4.01
	43%	22%	6%	10%	12%
Σ	67.87	14.76	13.70	17.50	26.05

тувинцами ($N = 44$), хантами ($N = 29$), бурятами ($N = 44$), якутами ($N = 41$). Волго-Уральский географический регион представлен популяциями коми ($N = 32$), чувашей ($N = 28$), башкир ($N = 34$), мордвы ($N = 33$), татар ($N = 33$). Популяции Европы представлены карелами ($N = 26$), белорусами ($N = 28$), молдаванами ($N = 32$), русскими ($N = 77$), украинцами ($N = 29$). Популяции карачаевцев ($N = 22$), чеченцев ($N = 18$), адыгейцев ($N = 31$), мегрелов ($N = 34$), абхазов ($N = 26$), кабардинцев ($N = 35$), черкесов ($N = 32$) относятся к Кавказскому региону, а популяции киргизов ($N = 28$), узбеков ($N = 35$) и казахов ($N = 44$) – к Средней Азии. Метисы первого поколения были исключены из анализа.

ROH были идентифицированы с помощью программы Runs of Homozygosity, реализованной в PLINK версии 1.9 [10] с использованием стандартных настроек. Для сравнения данных полученные результаты были обобщены путем расчета средних значений количества ROH и совокупной

длины ROH на человека для каждой популяции. Статистическую значимость межпопуляционных различий оценивали при помощи теста Уилкоксона (уровень значимости = 0.05).

Медианное число ROH на индивидуальный геном в исследованных образцах составило 7, граница первого квартиля 4, граница третьего квартиля 12. Медианная длина регионов 15 Mb, граница первого квартиля 7.5 Mb, третьего – 35.2 Mb. Для популяций Сибири медианное число ROH составило 19, с границей для первого квартиля 13 и третьего 27, медианная длина регионов 55.7 Mb с границами 38 и 86.1 Mb для первого и третьего квартиля соответственно.

Средняя суммарная длина ROH во всех диапазонах в популяциях коренного населения Сибири значительно превышает все остальные географические регионы (табл. 1). Средняя суммарная длина ROH в популяциях Волго-Уральского региона выше в диапазонах длин 1–2, 2–3, 3–4, 4–5, 5–6, 6–8 Mb, а также по сумме всех диапазонов по

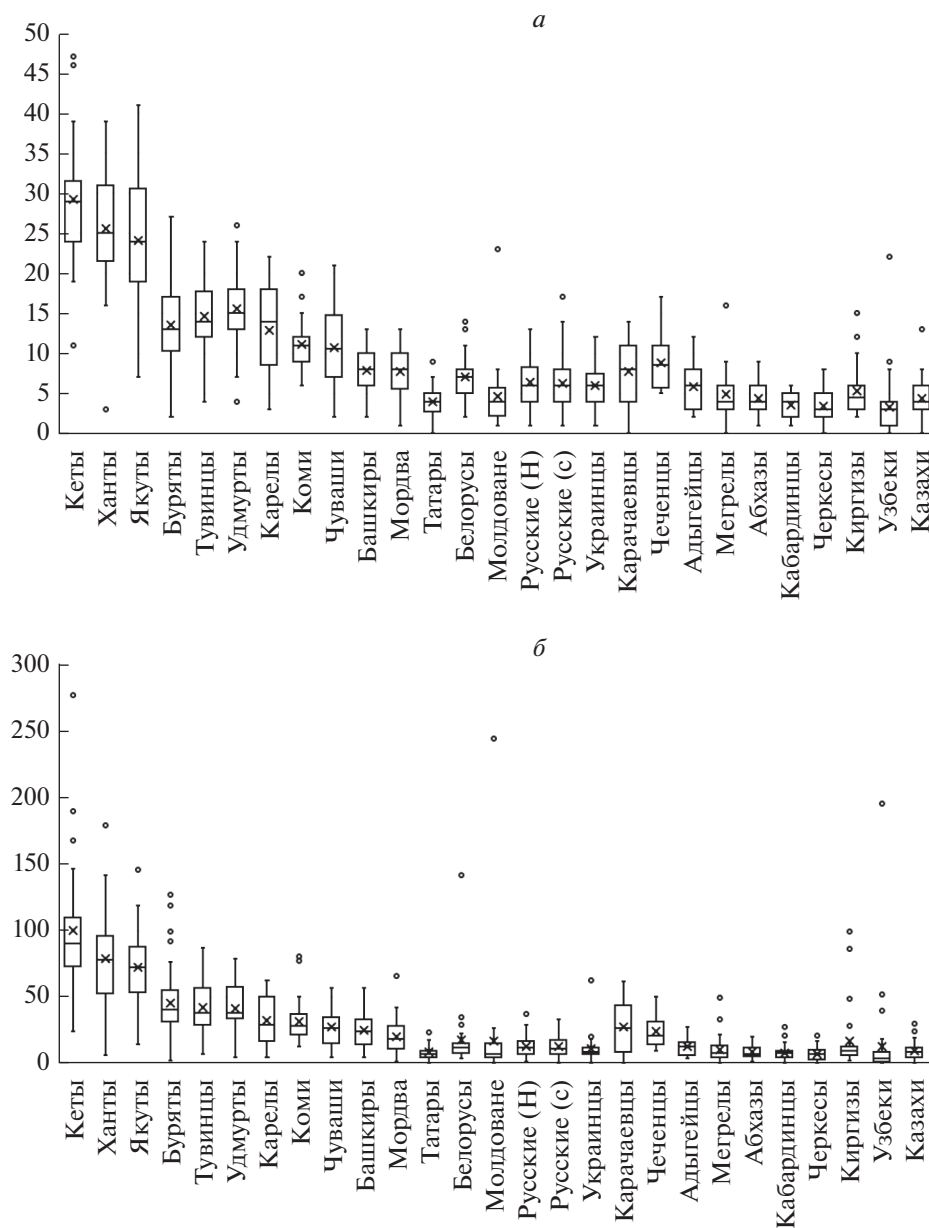


Рис. 1. Количество ROH в исследуемых популяциях (а). Суммарные длины ROH (Mb) в исследуемых популяциях (б).

сравнению с популяциями Европы, Средней Азии и Кавказа (табл. 1). Для популяций Средней Азии в диапазоне более 10 Mb средняя суммарная длина ROH выше, чем у популяций кавказского и европейского регионов, за счет популяций киргизов, близкородственных по происхождению к южносибирским этносами узбеков.

Популяции коренного населения Сибири по среднему количеству и средней суммарной длине ROH образуют две группы: в первую группу входят якуты, кеты и ханты, во вторую — тувинцы и буряты (рис. 1). Для кетов наблюдается наибольшее количество и длина ROH, что, вероятно, связано с сильным инбридингом, малой эффектив-

ной численностью и географической изоляцией популяции. В популяциях кетов, хантов и бурятов наблюдается существенный вклад более протяженных регионов высокой гомозиготности, что также свидетельствует о недавнем сильном инбридинге (табл. 2). Ранее по маркерам Y-хромосомы у бурятов было обнаружено относительно низкое генетическое разнообразие по сравнению с другими популяциями Южной Сибири и сильные эффекты основателя по двум основным гаплогруппам [11]. Для якутов, несмотря на большой эффективный размер популяции, также наблюдается большое количество и длина ROH, что, вероятно, связано с сильным эффектом основателя,

Таблица 2. Среднее суммарное количество и длина RОН по диапазонам (Мб) в исследуемых популяциях

Популяции	Среднее суммарное количество по диапазонам (Мб)										Средняя суммарная длина по диапазонам (Мб)									
	1–2	2–3	3–4	4–5	5–6	6–8	8–10	>10	Σ	1–2	2–3	3–4	4–5	5–6	6–8	8–10	>10	Σ		
Карелы	7.38	2.85	1.04	0.62	0.27	0.42	0.15	0.19	12.90	10.68	7.07	3.52	2.72	1.47	2.91	1.38	2.90	32.66		
Абхазы	2.58	1.15	0.31	0.15	0.04	0.08	0.00	0.00	4.30	3.82	2.73	1.04	0.72	0.20	0.57	0.00	0.00	9.08		
Адыгейцы	3.61	1.26	0.42	0.16	0.29	0.06	0.00	0.03	5.80	5.37	3.06	1.39	0.73	1.62	0.44	0.00	0.37	12.97		
Башкиры	3.65	1.79	0.79	0.50	0.41	0.26	0.18	0.32	7.90	5.22	4.35	2.77	2.28	2.24	1.88	1.59	5.28	25.62		
Белорусы	4.75	1.25	0.43	0.25	0.00	0.11	0.07	0.14	7.00	7.03	2.98	1.48	1.08	0.00	0.71	0.59	3.83	17.70		
Буряты	7.05	2.48	1.14	0.82	0.50	0.48	0.32	0.75	13.50	10.15	6.03	3.98	3.70	2.71	3.26	2.76	13.28	45.86		
Кабардинцы	2.14	0.91	0.14	0.03	0.09	0.03	0.09	0.09	3.50	3.01	2.15	0.49	0.12	0.49	0.20	0.74	1.12	8.31		
Казахи	2.62	1.09	0.31	0.13	0.09	0.04	0.02	0.04	4.40	3.85	2.61	1.06	0.57	0.50	0.32	0.22	0.51	9.86		
Карачаевцы	3.26	1.70	0.65	0.39	0.39	0.43	0.57	0.39	7.80	4.75	4.12	2.26	1.73	2.12	2.96	5.14	4.94	29.29		
Кеты	14.04	6.50	2.54	1.73	1.23	1.15	0.77	1.31	29.30	20.69	15.66	8.63	7.70	6.67	7.87	6.88	25.90	100		
Киргизы	3.00	1.21	0.36	0.21	0.07	0.11	0.07	0.21	5.30	4.35	2.92	1.18	0.93	0.39	0.75	0.64	5.64	16.80		
Кочи	6.28	2.34	0.72	0.66	0.19	0.31	0.19	0.44	11.10	9.07	5.71	2.47	2.85	1.02	2.12	1.61	6.52	31.37		
Мегрелы	2.92	1.28	0.36	0.11	0.11	0.03	0.06	0.03	4.90	4.40	3.20	1.29	0.52	0.62	0.21	0.53	0.88	11.64		
Молдаване	2.59	0.72	0.38	0.13	0.19	0.19	0.09	0.34	4.60	3.68	1.76	1.25	0.60	1.06	1.38	0.84	6.30	16.87		
Мордва	4.61	1.64	0.52	0.45	0.15	0.12	0.06	0.24	7.80	6.67	4.04	1.75	2.03	0.83	0.83	0.51	4.06	20.71		
Русские (Н)	3.89	1.58	0.39	0.21	0.08	0.11	0.03	0.03	6.30	5.68	3.84	1.33	0.91	0.41	0.71	0.24	0.29	13.39		
Русские (с)	4.36	1.15	0.28	0.21	0.03	0.08	0.05	0.05	6.20	6.31	2.80	0.95	0.93	0.14	0.54	0.48	0.67	12.81		
Татары	2.50	0.94	0.12	0.06	0.06	0.18	0.03	0.03	3.90	3.71	2.39	0.41	0.28	0.34	1.33	0.26	0.32	9.04		
Тувинцы	7.64	3.30	1.34	0.80	0.30	0.66	0.20	0.41	14.60	11.20	7.90	4.70	3.60	1.60	4.51	1.76	7.11	42.38		
Удмурты	7.96	3.52	1.68	0.76	0.64	0.44	0.28	0.24	15.50	11.24	8.55	6.02	3.38	3.45	2.94	2.42	3.80	41.79		
Узбеки	1.64	0.51	0.38	0.08	0.10	0.10	0.08	0.31	3.20	2.79	1.43	1.47	0.39	0.64	0.79	0.80	6.10	14.43		
Украинцы	4.55	1.00	0.28	0.07	0.00	0.00	0.00	0.07	6.00	6.51	2.30	0.94	0.30	0.00	0.00	0.00	1.54	11.59		
Ханты	12.79	5.41	2.52	1.28	1.07	1.00	0.48	0.97	25.50	18.33	13.00	8.71	5.75	5.84	6.84	4.41	16.27	79.14		
Черкесы	2.03	0.91	0.18	0.03	0.18	0.09	0.00	0.03	3.50	2.94	2.28	0.60	0.13	1.01	0.67	0.00	0.32	7.96		
Чеченцы	4.39	2.11	0.78	0.72	0.00	0.33	0.17	0.28	8.80	6.29	4.97	2.66	3.21	0.00	2.29	1.47	3.19	24.08		
Чуваши	6.04	2.32	0.96	0.57	0.36	0.18	0.04	0.29	10.80	8.81	5.47	3.37	2.56	2.00	1.17	0.31	4.10	27.79		
Якуты	10.66	5.93	2.56	1.61	1.27	0.95	0.44	0.63	24.0	15.36	14.43	8.91	7.28	6.95	6.43	3.82	8.79	71.97		

Примечание. Русские (Н) — выборка русских из Новосибирска, (с) — русские-староверы.

который выявляется на маркерах мтДНК и Y-хромосомы [12].

Среди европейских популяций значительно отличаются карелы, показывая большее количество и длину, а также максимальный разброс ROH (рис. 1), что, вероятно, свидетельствует о наличии у них эффекта основателя. Карелы этнически и генетически близки к финнам, у которых также обнаружен очень сильный эффект основателя по маркерам мтДНК и Y-хромосомы [13]. По сравнению с другими исследованными европейскими этносами у карелов меньше эффективная численность популяции, следовательно, ожидается высокая степень гомозиготизации генома по сравнению с более гомогенными и обладающими высокой эффективной численностью популяциями, такими как восточные славяне, украинцы, белорусы и молдаване.

По доле образцов с одним и более ROH длиной больше 10 Mb по разным регионам выделяются три различные группы, которые явно связаны с эндогамией и изоляцией: большая часть сибирских образцов имеет длинные ROH (43% имеют ROH > 10 Mb), далее идут образцы Кавказского региона (22% имеют ROH > 10 Mb) и для среднеазиатского, европейского и Волго-Уральского регионов составляют соответственно 6, 10 и 12% (табл. 1).

Таким образом, в популяциях коренного населения Сибири количество и суммарная длина ROH значительно выше, чем в остальных исследуемых популяциях Северной Евразии, за исключением карелов и удмуртов, не показывающих значимых отличий с бурятами и тувинцами. Вероятно, это связано с генетико-демографической историей популяций, генетическое разнообразие которых складывалось под действием эффекта основателя, инбридинга, миграций, изоляцией состоянием.

Работа выполнена в рамках Научно-технической программы союзного государства “Разработка инновационных геногеографических и геномных технологий идентификации личности и индивидуальных особенностей человека на основе изучения генофондов регионов Союзного государства (ДНК-идентификация)”, госконтракт № 011-17 от 26.09.2017 г., а также поддержана грантом РФФИ № 18-29-13045 “Популяционная геномика и транскриптомика человека: поиск сигналов не-нейтральной эволюции”.

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального и/или национального комитета по исследовательской этике и Хельсинкской декларации 1964 г. и ее последующим изменениям или сопоставимым нормам этики.

От каждого из включенных в исследование участников было получено информированное добровольное согласие.

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. *Pemberton T.J., Absher D., Feldman M.W. et al.* Genomic patterns of homozygosity in worldwide human populations // *Am. J. Hum. Genet.* 2012. V. 8. № 91. P. 275–292. <https://doi.org/10.1016/j.ajhg.2012.06.014>
2. *Ropers H.H.* New perspectives for the elucidation of genetic disorders // *Am. J. Hum. Genet.* 2007. V. 81. P. 199–207. <https://doi.org/10.1086/520679>
3. *Ceballos F.C., Joshi P.K., Clark D.W. et al.* Runs of homozygosity: windows into population history and trait architecture // *Nat. Rev. Genet.* 2018. V. 19(4). P. 220–234. <https://doi.org/10.1038/nrg.2017.109>
4. *Triska P., Chekanov N., Stepanov V. et al.* Between Lake Baikal and the Baltic Sea: Genomic history of the gateway to Europe // *BMC Genetics.* 2017. V. 18. (Suppl. 1). P. 110. <https://doi.org/10.1186/s12863-017-0578-3>
5. *Степанов В.А., Пузырев В.П.* Микросателлитные гаплогруппы Y-хромосомы демонстрируют отсутствие подразделенности и наличие нескольких компонентов в мужском генофонде тувинцев // *Генетика.* 2000. Т. 36. № 3. С. 377–384.
6. *Степанов В.А., Пузырев В.П.* Анализ аллельных частот семи микросателлитных локусов Y-хромосомы в трех популяциях тувинцев // *Генетика.* 2000. Т. 36. № 2. С. 241–248.
7. *Степанов В.А., Харьков В.Н., Пузырев В.П.* Эволюция и филогеография линий Y-хромосомы человека // *Информацион. вестн. ВОГиС.* 2006. Т. 10. № 1. С. 57–73.
8. *Хитринская И.Ю., Степанов В.А., Пузырев В.П. и др.* Генетическое своеобразие население Якутии по данным аутосомных локусов // *Мол. биол.* 2003. Т. 37. № 2. С. 234–239.
9. *Степанов В.А.* Этногеномика населения Северной Евразии. Томск: Печатная мануфактура, 2002. 244 с.
10. *Purcel Sh., Chang Ch.* PLINK [1.9]. URL: www.cog-genomics.org/plink/1.9/.
11. *Харьков В.Н., Хамина К.В., Медведева О.Ф. и др.* Генофонд бурят: клинальная изменчивость и территориальная подразделенность по маркерам Y-хромосомы // *Генетика.* 2014. Т. 50. № 2. С. 203–213. <https://doi.org/10.1134/S1022795413110082>
12. *Харьков В.Н., Степанов В.А., Медведева О.Ф. и др.* Происхождение якутов: анализ гаплотипов Y-хромосомы // *Мол. биол.* 2008. Т. 42. № 2. С. 226–237.
13. *Sajantila A., Salem A.H., Savolainen P. et al.* Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population // *PNAS USA.* 1996. V. 93. № 21. P. 12035–12039.

Characteristics of Genomic Distribution of Runs of Homozygosity in the Indigenous Population of Northern Eurasia

N. A. Kolesnikov^{a,*}, V. N. Kharkov^a, A. A. Zarubin^a, and V. A. Stepanov^a

^a*Tomsk National Research Medical Center of the Russian Academy of Sciences,
Research Institute of Medical Genetics, Tomsk, 634050 Russia*

**e-mail: nikita.kolesnikov@medgenetics.ru*

Populations of the indigenous ethnic groups of Northern Eurasia are of considerable interest for population genomics, both because they were relatively poorly studied with the use of modern genomic technologies, and due to the specificity of their gene pools that developed in different genetic-demographic conditions. We used genotype data on 242 179 autosomal SNPs for 894 individuals, to search for regions with runs of homozygosity more than 1 Mb. For Siberian populations, number of runs of homozygosity, and their total length was higher than for other populations of Northern Eurasia.

Keywords: population genetics, Siberia, ROH, runs of homozygosity, human populations, North Eurasia.