

Различия по уровню инбридинга в популяциях Дагестана: анализ регионов высокой гомозиготности

Колесников Н.А.¹, Харьков В.Н.¹, Раджабов М.О.², Зарубин А.А.¹,
Литвинов С.С.³, Екомасова Н.В.³, Хуснутдинова Э.К.³, Степанов В.А.¹

- 1 — Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук
634050, г. Томск, Набережная реки Ушайки, 10
- 2 — Дагестанский государственный медицинский университет
367000, Республика Дагестан, г. Махачкала, пл. Ленина, 1
- 3 — Институт биохимии и генетики, Уфимский федеральный исследовательский центр Российской академии наук
450054, Республика Башкортостан, г. Уфа, проспект Октября, 71

Анализ геномных данных (~ 907 750 аутосомных SNP) позволил оценить уровень инбридинга на основе идентификации регионов высокой гомозиготности (FROH) в выборке из 518 человек 23 коренных народов Дагестана. В результате, в популяциях коренного населения Дагестана количество и суммарная длина ROH, а также коэффициент FROH показывают значительное разнообразие как внутри, так и между исследованными популяциями. У гинухцев, гунзибцев, ахвахцев, бежтинцев и цезов наблюдаются одни из самых высоких значений коэффициента инбридинга ($FROH > 1,5$ млн п.н.) по сравнению с остальными мировыми популяциями.

Ключевые слова: популяционная генетика, Дагестан, регионы высокой гомозиготности, популяции человека, инбридинг

Для цитирования: Колесников Н.А., Харьков В.Н., Раджабов М.О., Зарубин А.А., Литвинов С.С., Екомасова Н.В., Хуснутдинова Э.К., Степанов В.А. Различия по уровню инбридинга в популяциях Дагестана: анализ регионов высокой гомозиготности. *Медицинская генетика* 2020; 19(7): 21–22.
DOI: 10.25557/2073-7998.2020.07.21-22

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

Финансирование. Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-34-90101.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила: 20.05.2020

Differences in the level of inbreeding in populations of Dagestan: analysis of runs of homozygosity

Kolesnikov N.A.¹, Kharkov V.N.¹, Radzhabov M.O.², Zarubin A.A.¹, Litvinov S.S.³,
Ekomasova N.V.³, Khusnutdinova E.K.³, Stepanov V.A.¹

- 1 — Research Institute of Medical Genetics, Tomsk National Medical Research Center
Nab. Ushaiki, 10, Tomsk, 634050, Russia
- 2 — Dagestan State Medical University
Lenin sq. 1, Makhachkala, 367000, Russia
- 3 — Institute of Biochemistry and Genetics, Ufa Scientific Center, Russian Academy of Sciences
October prospect 71, Ufa, 450054, Russia

The genomic data analysis (~907750 autosomal SNPs) allowed us to estimate the inbreeding level based on the analysis of runs of homozygosity (Froh) in the group of 518 people from 23 indigenous Dagestan peoples. As a result, the number and total length of ROH, as well as the FROH coefficient, show a significant diversity in the indigenous Dagestan populations both inside and between them. One of the highest values of the inbreeding coefficient ($FROH > 1.5$ Mb) is observed for the Ginukhs, Gunzibians, Akhvahs, Bezhtins, and Tsez, as compared with the other world populations.

Keywords: population genetics, ROH, runs of homozygosity, human populations, Dagestan, inbreeding

For citation: Kolesnikov N.A., Kharkov V.N., Radzhabov M.O., Zarubin A.A., Litvinov S.S., Ekomasova N.V., Khusnutdinova E.K., Stepanov V.A. Differences in the level of inbreeding in populations of Dagestan: analysis of runs of homozygosity. *Medical genetics*. 2020; 19(7): 21–22 (In Rus).
DOI: 10.25557/2073-7998.2020.07.21-22

Corresponding author: Kolesnikov Nikita Alexandrovich; **e-mail:** nikita.kolesnikov@medgenetics.ru

Funding. The reported study was funded by RFBR, project number 19-34-90101

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

Accepted: 20.05.2020

Коэффициент инбридинга F — один из центральных параметров в популяционной генетике. Важным этапом для его определения является обнаружение регионов высокой гомозиготности (RON). Длинные RON, вероятно, происходят от недавнего предка, короткие — от более дальнего предка. Вычисление доли генома человека в виде RON определенной длины (например, > 1 млн п.н., $> 1,5$ млн п.н., и > 5 млн п.н.) предоставляет информацию об уровнях инбридинга относительно популяций N -количества поколений назад. Современные методы исследования на основе высокоплотных чипов позволяют проанализировать генетико-демографические процессы, в том числе и инбридинг на геномном уровне. Для популяций из различных регионов показаны разные значения инбридинга [1]. В отличие от F (коэффициент инбридинга) оценки F_{RON} могут потенциально более точно оценить фактический процент генома, который является аутозиготным, включая аутозиготность, возникшую у очень отдаленного общего предка, например, более 50 поколений назад [2, 3]. Таким образом, уровень инбридинга является информативным в отношении генетико-демографических процессов в популяции, а также может быть связан и распространенностью генетических заболеваний. Следовательно, надежная оценка уровней инбридинга имеет важное значение для генетики популяций человека.

Целью данной работы было оценить степень инбридинга в популяциях народов Дагестана с использованием RON.

Материалы и методы

Данные были полученных с помощью биочипов Infinium Multi-Ethnic Global-8 Kit на выборке из 518 человек из 23 коренных популяций Дагестана: аварцев, даргинцев, лакцев, табасаран, лезгин, агулов, рутульцев, цахуров, арчинцев, андийцев, ахвахцев, тиндинцев, каратинцев, багулалов, ботлихцев, хваршин, цезов, бежтинцев, чамалинцев, гунзибцев, гинухцев, кумыков и караногайцев. Из анализа были исключены образцы, имеющие более 5% пропущенных позиций, а также SNP с отсутствующими генотипами более чем в 10%. Были исключены все позиции делеций и инсерций. В итоге, осталось 907 750 SNP. Весь анализ был выполнен с учетом трех различных наборов RON полученных в PLINK v1.9 [4] идентифицированных с учетом минимальной длины 500 т.п.н., 1,5 млн п.н. и 5 млн п.н., а также трех классов RON полученных с помощью программы Garlic [5], использующей альтернативный подход для поиска RON и имеющей модуль классифи-

кации на основе длин для выявления биологически значимых классов RON. Статистическую значимость межпопуляционных различий оценивали при помощи теста Вилкоксона (уровень значимости = 0,05) используя пакет `wilcox.test` в среде R. Для расчета наблюдаемого процента F_{RON} у каждого индивидуума размеры областей RON, превышающих 1,5 млн п.н. (исключая половые хромосомы) суммировали и затем делили на общую длину аутосом, за исключением центромер (2708,613716 млн п.н. для hg18).

Результаты и выводы

Наибольшим коэффициентом геномного инбридинга для длин $RON > 1,5$ млн п.н. обладают популяции гинухцев ($F_{RON} = 0,1063$) и гунзибцев ($F_{RON} = 0,0913$), что является одним из самых больших показателей в мире. Также выделяются ахвахцы, бежтинцы и цезы (0,0600, 0,0707 и 0,0712 соответственно), не имея между собой статистически значимых отличий, как по длинам, так и по количеству RON на chromosome. Для остальных популяций наблюдается разброс в диапазоне значений F_{RON} от 0,0060 до 0,0485. Коэффициент инбридинга F_{RON} для $RON > 1,5$ млн п.н. (PLINK) показывает высокую корреляцию с F_{RON} для класса длинных RON (Garlic) ($r = 0,9931$), что говорит о том, что порог в 1,5 млн п.н. хорошо характеризует последствия недавнего инбридинга.

Таким образом, в популяциях коренного населения Дагестана количество и суммарная длина RON, а также коэффициент F_{RON} показывают значительное разнообразие как внутри, так и между исследованными популяциями. У гинухцев, гунзибцев, ахвахцев, бежтинцев и цезов, наблюдаются одни из самых высоких значений коэффициента инбридинга ($F_{RON} > 1,5$ млн п.н.) по сравнению с остальными мировыми популяциями.

Литература/References

1. Ceballos F.C., Joshi P.K., Clark D.W., et al. Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics*. 2018 Apr; 19(4): 220–234. 2. McQuillan R., Leutenegger A.L., Abdel-Rahman R., et al. Runs of homozygosity in European populations. *Am J Hum Genet*. 2008 Sep; 83: 359–372.
3. Lemes R.B., Nunes K., Carnavalli J.E.P., et al. Inbreeding estimates in human populations: Applying new approaches to an admixed Brazilian isolate. *PLOS ONE*. 2018 Apr 24; 13(4): e0196360.
4. Purcell S., Neale B., Todd-Brown K., et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*. 2007 Sep; 81: 559–575.
5. Szpiech Z.A., Blant A., Pemberton T.J. GARLIC: Genomic Autozygosity Regions Likelihood-based Inference and Classification. *Bioinformatics*. 2017 Jul 1; 33(13): 2059–2062.