

УДК 575.174:577.21

ВАРИАБЕЛЬНОСТЬ И ИДЕНТИФИКАЦИОННЫЙ ПОТЕНЦИАЛ 60 X-ХРОМОСОМНЫХ SNPs В ДВУХ ПОПУЛЯЦИЯХ КОРЕННОГО НАСЕЛЕНИЯ СИБИРИ

© 2016 г. В. А. Степанов^{1,2}, К. В. Вагайцева^{1,2}, В. Н. Харьков^{1,2},
А. А. Чередниченко¹, Л. И. Минайчева¹, А. В. Бочарова¹

¹Научно-исследовательский институт медицинской генетики, Томск 634050
e-mail: vadim.stepanov@medgenetics.ru

²Национальный исследовательский Томский государственный университет,
лаборатория онтогенетики человека, Томск 634050

Поступила в редакцию 23.07.2015 г.

Описано генетическое разнообразие 60 X-хромосомных однонуклеотидных маркеров (панель XSNPid) в популяциях сибирских татар и тувинцев. Выявлена близость двух популяций по частотам аллелей и относительно низкий уровень их генетической дифференциации ($G_{ST} = 0.021$). Показан высокий дискриминационный потенциал панели XSNPid в исследованных популяциях. Вероятность случайного совпадения мультилокусных генотипов у неродственных мужчин (MP, matching probability) составляет 1.12×10^{-18} у сибирских татар и 7.77×10^{-16} у тувинцев. Для женщин этот показатель еще на несколько порядков ниже — 1.51×10^{-25} у сибирских татар и 1.83×10^{-23} у тувинцев.

Ключевые слова: генетическое разнообразие, популяции народов Сибири, X-хромосомные маркеры, ДНК-идентификация.

DOI: 10.7868/S0016675816030152

X-хромосомные генетические маркеры человека активно используются в популяционной и медицинской генетике, а также представляют большой интерес для криминалистики и судебной медицины. В этой сфере маркеры X-хромосомы применяются в основном при тестировании биологического родства в сложных случаях, однако при достаточной информативности они способны решить и стандартные задачи индивидуальной идентификации. В отличие от аутосомных маркеров для ДНК-идентификации, для которых существуют принятые в США и Европе стандартные наборы из 13–20 микросателлитных (STR) локусов, общепринятые стандарты для X-хромосомных систем отсутствуют. С появлением новых технологий генотипирования и массового параллельного секвенирования STR-маркеры постепенно уступают место более распространенным и более простым в генотипировании однонуклеотидным маркерам (SNPs) [1, 2]. Для достижения сопоставимого с STR дискриминационного потенциала число SNP маркеров должно быть в 3–4 раза больше [3]. Панели X-хромосомных SNPs (XSNP), создаваемые на базе традиционных технологий (ПЦР в реальном времени, капиллярный гель-электрофорез), не достигали уровня информативности систем, основанных на STR-маркерах,

из-за малой емкости мультиплексирования в этих технологиях [4–7]. Недавно мы разработали панель из 66 несцепленных XSNP (XSNPid), генотипирование которой осуществляется в составе двух мультиплексов на 36 и 30 маркеров методом MALDI-TOF-масс-спектрометрии (матрично-активированной лазерной десорбции/ионизации с измерением времени пролета) [8].

Задача настоящей работы — описать генетическое разнообразие по панели XSNPid в двух популяциях коренного населения Сибири (тувинцев и сибирских татар) и оценить дискриминационный потенциал панели для ДНК-идентификации на примере этих популяций.

Выборка тувинцев ($N = 95$) была собрана в г. Кызыл, сибирских татар ($N = 76$) — в пос. Эушта Томской области. В состав обеих выборок включались только неродственные индивиды, не метисированные в трех поколениях, подписавшие информированное согласие на анонимное использование образцов ДНК для научных исследований. ДНК выделяли из лимфоцитов периферической крови стандартным методом. Генотипирование осуществляли с помощью мультиплексной ПЦР и разделения молекул ДНК по массе методом тандемной масс-спектрометрии на анализа-

Показатели генетического разнообразия и дискриминационного потенциала системы 60 XSNP-маркеров в популяциях тувинцев и сибирских татар

Показатель	Тувинцы	Сибирские татары
Среднее генное разнообразие (H_e)	0.4354	0.4808
Вероятность случайного совпадения генотипов для женщин (MPf, matching probability, females)	1.83×10^{-23}	1.51×10^{-25}
Вероятность случайного совпадения генотипов для мужчин (MPm, matching probability, males)	7.77×10^{-16}	1.12×10^{-18}
Дискриминационный потенциал для женщин (PDf, power of discrimination, females)	0.99999999999999999998	0.99999999999999999998
Дискриминационный потенциал для мужчин (PDm, power of discrimination, males)	0.9999999999999992	0.9999999999999998

торе “Sequenom MassARRAY 4” (“Sequenom”, США) по протоколам, описанным ранее [9]. Статистический анализ частот генов и генотипов проводили в пакетах программ Arlequin и Haploview [10, 11]. Идентификационные характеристики системы X-сцепленных однонуклеотидных маркеров оценивали с помощью стандартных показателей, используемых в криминалистике и судебной медицине, как описано ранее [12–14].

Шесть из 66 маркеров были исключены из статистического анализа частот аллелей и криминалистических показателей. Три из них (rs2130835, rs4825220 и rs4830049) показали низкую эффективность генотипирования хотя бы в одной из популяций (call rate ниже 85%), тогда как для остальных 63 маркеров было прочтено более 98% генотипов. Один из SNPs (rs225067) показал достоверное, с учетом поправки Бонферрони, отклонение от равновесия Харди–Вайнберга у женщин в популяции сибирских татар (уровень значимости точного теста $P < 0.000001$). При попарном анализе неравновесия по сцеплению с помощью точного теста было обнаружено четыре пары маркеров с участием повторяющихся SNPs, демонстрирующих значимое отклонение от независимого наследования (rs2404797–rs7888207 у тувинцев; rs2404797–rs5934683, rs2411976–rs5937091, rs2411976–rs 916208 у сибирских татар). Два маркера (rs2404797 и rs2411976), один из которых фигурировал в каждой паре сцепленных SNPs, также были удалены из массива данных.

Список маркеров, частоты аллелей, показатели генетического разнообразия для каждого из маркеров и полокусные криминалистические показатели приведены в приложении, доступном он-лайн (<http://www.medgenetics.ru/UserFile/File/Doc/Publications/2015/XSNP-Prilogenie-2015.pdf>) или по запросу у авторов. Средние значения генного разнообразия и суммарные значения дискриминационного потенциала приведены в таблице.

Две исследованные популяции коренного населения Сибири в целом имеют близкий спектр аллельных частот 60 X-хромосомных SNPs. Разница в частотах аллелей варьирует по локусам от 0 до 38%, в среднем находясь на уровне около 10%. Близость двух популяций по частотам аллелей отражается и в относительно низком значении коэффициента генетической дифференциации ($G_{ST} = 0.021$). При этом популяция сибирских татар демонстрирует существенно большее среднее генное разнообразие ($H_e = 0.481$) по сравнению с тувинцами ($H_e = 0.435$, см. таблицу).

Для оценки возможности использования исследованной системы маркеров для ДНК-идентификации определяли стандартные популяционно-статистические показатели, характеризующие идентификационный потенциал системы маркеров. Важнейший из этих показателей – вероятность дискриминации неродственных индивидов (PD, power of discrimination), представляющая собой вероятность несовпадения генотипов у неродственных индивидов. Из-за особенностей

передачи X-хромосом в ряду поколений идентификационный потенциал X-хромосомных маркеров существенно отличается у женщин и мужчин, поэтому дискриминационный потенциал рассчитывают отдельно для женщин (PDF) и мужчин (PDM) (см. таблицу).

Показатели идентификационного потенциала исследованной системы маркеров для ДНК-идентификации свидетельствуют о том, что панель XSNPid демонстрирует очень высокую дискриминационную способность в изученных популяциях. Вероятность случайного совпадения мультилокусных генотипов у неродственных мужчин (MP, matching probability) составляет 1.12×10^{-18} у сибирских татар и 7.77×10^{-16} у тувинцев. Для женщин этот показатель еще на несколько порядков ниже: 1.51×10^{-25} у сибирских татар и 1.83×10^{-23} у тувинцев.

Следует отметить, что достигнутые уровни дискриминации индивидов по панели 60 X-сцепленных SNPs для мужчин не ниже, чем у используемых сейчас американского стандарта CODIS (Combined DNA Index System), европейского стандарта ESS (European Standard Set), основанных на STR-маркерах, и коммерческих наборов для их типирования. Ранее мы показали, что вероятность случайного совпадения генотипов в популяциях населения России по стандартному набору STR-локусов варьирует в пределах 1×10^{-16} – 5.60×10^{-17} [14, 15]. Для женщин дискриминационный потенциал панели XSNPid на несколько порядков превышает таковой для стандартных микросателлитных систем.

Таким образом, в настоящей работе мы описали генетическое разнообразие двух популяций коренного населения Сибири по панели X-сцепленных SNP-маркеров, обнаружили высокий уровень генетического разнообразия в этих популяциях и показали высокий идентификационный потенциал панели XSNPid для индивидуальной ДНК-идентификации.

Работа выполнена при поддержке Госконтракта № 14.604.21.0019 от 17 июня 2014 года в рамках ФЦП “Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014–2020 годы”.

СПИСОК ЛИТЕРАТУРЫ

1. Yang Y., Xie B., Yan Y. Application of Next-generation sequencing technology in forensic science // *Genomics. Proteomics. Bioinformatics*. 2014. V. 12. P. 190–197.
2. Churchill J.D., Chang J., Ge J. et al. Blind study evaluation illustrates utility of the Ion PGM system for use in human identity DNA typing // *Croat. Med. J.* 2015. V. 56. P. 218–229.
3. Krawczak M. Informativity assessment for biallelic single nucleotide polymorphisms // *Electrophoresis*. 1999. V. 20. P. 1676–1681.
4. Li C., Zhang S., Zhao S., et al. Analysis of 14 highly informative SNP markers on X chromosome by TaqMan SNP genotyping assay // *Forensic Sci. Int.: Genetics*. 2010. V. 4. P. e145–e148.
5. Tomas C., Sanchez J.J., Castro J.A. et al. Forensic usefulness of a 25 X-chromosome single-nucleotide polymorphism set // *Transfusion*. 2010. V. 50. P. 2258–2265.
6. Pereira V., Tomas C., Amorim A. et al. Study of 25 X-chromosome SNPs in Portuguese // *Forensic Sci. Int.: Genetics*. 2011. V. 5. № 4. P. 336–338.
7. Вагайцева К.В., Харьков В.Н., Черпинская К.В. и др. Генетическая вариабельность X-сцепленных STR-маркеров в популяциях Сибири // *Мол. биология*. 2015. Т. 49. № 2. С. 305–312.
8. Stepanov V.A., Vagaitseva K.V., Kharkov V.N. Multiplex X-SNP system for forensic genetics // *Program and Abstracts. 9th ISABS conference on Forensic, Anthropologic Genetics and Mayo Clinic Lectures in Individualized Medicine*. June 22–26. Bol, Island of Brac, Croatia, 2015. P. 86.
9. Степанов В.А., Трифонова Е.А. Мультиплексное генотипирование однонуклеотидных полиморфных маркеров методом MALDI-TOF масс-спектрометрии: частоты 56 SNP в генах иммунного ответа в популяциях человека // *Мол. биология*. 2013. Т. 47. № 6. С. 976–986.
10. Excoffier L., Laval G., Schneider S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis // *Evol. Bioinform. Online*. 2005. V. 1. P. 47–50.
11. Barrett J.C., Fry B., Maller J., Daly M.J. Haploview: analysis and visualization of LD and haplotype maps // *Bioinformatics*. 2005. V. 21. № 2. P. 263–265.
12. Степанов В.А. Этногеномика населения Сибири и Средней Азии. Томск: Печатная мануфактура, 2002. 242 с.
13. Степанов В.А., Харьков В.Н., Трифонова Е.А., Марусин А.В. Методы статистического анализа в популяционной и эволюционной генетике человека: Учеб.-методич. пособие. Сер. “Наследственность и здоровье”. Вып. 12. Томск: Печатная мануфактура, 2014. 100 с.
14. Степанов В.А., Балановский О.П., Мельников А.В. и др. Характеристика популяций Российской Федерации по панели пятнадцати локусов, используемых для ДНК-идентификации и в судебно-медицинской экспертизе // *Acta Naturae*. 2011. Т. 3. № 2. С. 59–71.
15. Stepanov V.A., Melnikov A.V., Lash-Zavada A.Y. et al. Genetic variability of 15 autosomal STR loci in Russian populations // *Legal Med.* 2010. V. 12. № 5. P. 256–258.

Variability and Identification Power 60 X-Chromosome in Two Native Siberian Populations

V. A. Stepanov^{a, b}, K. V. Vagaitseva^{a, b}, V. N. Kharkov^{a, b}, A. A. Cherednichenko^a,
L. I. Minaicheva^a, and A. V. Bocharova^a

^aResearch Institute for Medical Genetics, Tomsk, 634050 Russia

e-mail: vadim.stepanov@medgenetics.ru

^bLaboratory of Human Ontogenetics, National Research Tomsk State University, Tomsk, 634050 Russia

Genetic diversity of 60 X-chromosome single nucleotide polymorphisms (XSNPId panel) in populations of Siberian Tatars and Tuvians is described. A close spectrum of allele frequencies and a low level of their genetic differentiation ($G_{st} = 0.021$) is revealed. High discriminating power of the XSNPId panel in populations under study is demonstrated. The random matching probability (MP) of multilocus genotypes in males is 1.12×10^{-18} in Siberian Tatars and 7.77×10^{-16} in Tuvians. In females, MP is several orders of magnitude lower: 1.51×10^{-25} in Siberian Tatars and 1.83×10^{-23} in Tuvians.

Keywords: genetic diversity, Siberian populations, X-chromosome markers, DNA identification.