

## ПАНЕЛЬ ОДНОНУКЛЕОТИДНЫХ СЦЕПЛЕННЫХ С X-ХРОМОСОМОЙ ПОЛИМОРФНЫХ МАРКЕРОВ ДЛЯ ДНК-ИДЕНТИФИКАЦИИ (XSNPid) НА ОСНОВЕ МУЛЬТИПЛЕКСНОГО ГЕНОТИПИРОВАНИЯ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МНОГОЛОКУСНОЙ ПЦР И МАСС-СПЕКТРОМЕТРИИ MALDI-TOF

© 2016 г. В. А. Степанов<sup>1,2\*</sup>, К. В. Вагайцева<sup>1,2</sup>, В. Н. Харьков<sup>1,2</sup>,  
А. А. Чередниченко<sup>1</sup>, А. В. Бочарова<sup>1</sup>

<sup>1</sup>Научно-исследовательский институт медицинской генетики, Томск, 634050

<sup>2</sup>Национальный исследовательский Томский государственный университет,  
лаборатория онтогенетики человека, Томск, 634050

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

Поступила в редакцию 28.07.2015 г.

Принята к печати 26.10.2015 г.

Генетические маркеры человека, сцепленные с X-хромосомой (кратко – X-сцепленные), применяются в области популяционной генетики, медицинской генетики, а также для ДНК-идентификации индивида в криминалистике и судебной медицине. Нами предложена панель XSNPid, состоящая из 66 несцепленных однонуклеотидных маркеров X-хромосомы, и разработан протокол их мультиплексного генотипирования с помощью методов многолокусной ПЦР и масс-спектрометрии MALDI-TOF. Панель XSNPid генотируется в составе двух мультиплексов (на 36 и 30 маркеров). Разработанный протокол обеспечивает эффективное чтение генотипов: доля определяемых генотипов составляет 98.29%. Для русской популяции характерен высокий уровень генного разнообразия (0.461) по X-сцепленным SNP, включенным в панель. 63 из 66 маркеров, обеспечивающих высокую эффективность генотипирования и независимое наследование, пригодны для целей ДНК-идентификации. Для панели XSNPid характерна очень высокая дискриминационная способность при исследовании русской популяции. Вероятность совпадения генотипов у двух неродственных индивидов составляет  $9 \times 10^{-27}$  для женщин и  $2 \times 10^{-18}$  для мужчин. При сравнении с другими ближайшими аналогами наборов X-хромосомных SNP, панель XSNPid обладает, кроме более высокой дискриминирующей способности, также большей емкостью мультиплексов, что делает ее более экономически выгодной и требует меньших затрат времени при использовании. Панель XSNPid – удобный инструмент не только для индивидуальной ДНК-идентификации, но и для популяционно-генетических исследований.

**Ключевые слова:** однонуклеотидные полиморфные маркеры (SNP), X-хромосома, масс-спектрометрия MALDI-TOF, мультиплексное генотипирование, ДНК-идентификация, популяционная генетика, генетическое разнообразие, популяции человека.

**DOI:** 10.7868/S0026898416030150

Генетические маркеры X-хромосомы (далее для краткости – X-сцепленные) используют в популяционной генетике человека, медицинской генетике, в криминалистике и судебной медицине. В популяционных исследованиях данные по генетической вариативности X-хромосомы дополняют картину генетического разнообразия и генетической дифференциации популяций, которая может быть получена при помощи одно-

родительских линий (по маркерам Y-хромосомы и мтДНК), а также панелей аутосомных полиморфных локусов [1–4]. В медицинской генетике маркеры X-хромосомы применяют для прямой и косвенной диагностики X-сцепленных болезней [5] и для поисков генетической компоненты распространенных заболеваний. По данным каталога широкогеномных ассоциативных исследований (GWAS), на X-хромосоме имеется около 150 SNP

Принятые сокращения: MALDI-TOF (Matrix Assisted Laser Desorption/Ionisation-Time Of Flight) – метод матрично-активированной лазерной десорбции/ионизации с измерением времени пролета; SNP (Single Nucleotide Polymorphism) – однонуклеотидный полиморфизм.

(Single Nucleotide Polymorphism), ассоциированных с различными распространенными хроническими заболеваниями [6]. В криминалистике и судебной медицине маркеры X-хромосомы применяются, в основном, при тестировании биологического родства. Особенности передачи X-хромосом в ряду поколений (мужчины наследуют одну из X-хромосом своей матери; женщина — одну хромосому матери и одну хромосому, унаследованную от бабушки по отцовской линии) наиболее полезны в сложных случаях определения родства. Например, при определении отцовства, когда ДНК предполагаемого отца не доступна; при определении отцовства между близкими родственниками; при определении родства по мужской линии, когда в родословной доступны только женщины [7]. Помимо сложных случаев определения родства, генетические маркеры X-хромосомы, обладая достаточной информативностью, способствуют решению и стандартных задач индивидуальной идентификации [8].

Для ДНК-идентификации традиционно используют микросателлитные маркеры (STR), в том числе X-сцепленные [4, 9–14]. Эти мультиаллельные маркеры характеризуются высокой гетерозиготностью, и поэтому относительно небольшого их числа (15–20) достаточно для индивидуальной идентификации человека [9]. Однонуклеотидные маркеры (SNP) менее вариабельны, однако гораздо более распространены в геноме человека. Для достижения сопоставимого с STR дискриминационного потенциала число SNP маркеров должно быть в 3–4 раза больше [15]. Панели X-сцепленных SNP, создаваемые на базе традиционных технологий (ПЦР в реальном времени, капиллярный гель-электрофорез), не достигают уровня информативности систем, основанных на аутосомных STR-маркерах [16–18]. Появление технологий массового параллельного секвенирования и новых методов генотипирования делает очевидной перспективу замещения STR-маркеров для ДНК-идентификации системами, основанными на SNP [19–21]. В ряду технологий генотипирования, оптимальных для среднemasштабных задач (типирование десятков или сотен маркеров в сотнях или тысячах образцов), одно из лидирующих мест занимает масс-спектрометрия методом MALDI-TOF (матрично-активированной лазерной десорбции/ионизации с измерением времени пролета) [22, 23].

Цель настоящего исследования — разработка панели X-сцепленных однонуклеотидных полиморфных маркеров для популяционно-генетического анализа и индивидуальной ДНК-идентификации населения России, а также разработка метода их мультиплексного генотипирования на основе многолокусной ПЦР и MALDI-TOF-масс-спектрометрии. Создание информативной панели на основе XSNP позволит изучать и генетическую из-

менчивость в популяциях, и проводить ДНК-идентификацию личности, — при сохранении преимуществ использования X-хромосомных маркеров в случаях сложного родства.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Материалы и экспериментальные процедуры.** Изучали образцы ДНК 180 русских жителей г. Томска (94 мужчин и 86 женщин). SNP-маркеры на X-хромосоме подбирали, используя базу данных NCBI [24], ориентируясь на равномерное расположение маркеров вдоль хромосомы с расстоянием между двумя соседними маркерами не менее 1 млн.п.н., на отсутствие сцепления между ними в европеоидных популяциях из проекта HarMap [25] и полагая, что уровень генетического разнообразия (ожидаемой гетерозиготности ( $H_e$ )) составляет не более 0.40 в этих популяциях.

Из отобранных на этом этапе более 200 SNP (при использовании программного обеспечения “Sequenom Assay Design”, доступного в режиме он-лайн на сайте [www.sequenom.com](http://www.sequenom.com)) сформировали два мультиплекса с максимальным числом совместимых в одном мультиплексе маркеров. В состав первого вошли 36 SNP (W1, далее 36-плекс), в состав второго — 30 (далее W2, 30-плекс). Состав мультиплексов приведен в табл. 1. К каждому маркеру сконструировали по два ПЦР-прайма (прямой и обратный), а также пролонгирующий праймер для iPLEX-реакции. Праймеры синтезированы в компании “Евроген” (Россия). Последовательности праймеров доступны по запросу у авторов.

Многолокусную ПЦР для наработки ампликонов, которые содержат исследуемые SNP, проводили отдельно для каждого мультиплекса в 96-луночных планшетах, объем 5 мкл (амплификаторы “Thermo Scientific”, Германия, и “Applied Biosystems”, США), в присутствии 5 нг геномной ДНК-матрицы и смеси ПЦР-праймеров соответствующего мультиплекса. Для ПЦР использовали реагенты “PCR Accessory Set”, входящие в состав набора для генотипирования методом масс-спектрометрии MALDI-TOF (“Agena Bioscience”, США). Состав ПЦР-смеси: 2 mM MgCl<sub>2</sub>, по 500 мкмоль каждого из dNTP, 0.1 мкМ смесь праймеров, 1 ед. ДНК-полимеразы. Схема ПЦР: начальная денатурация при 94°C (5 мин), затем 42 цикла амплификации в следующих условиях: денатурация при 94°C (20 с), отжиг при 56°C (40 с), элонгация при 72°C (60 с), после чего пробы инкубировали 5 мин при 72°C.

Последующие стадии эксперимента (SAP-реакция, iPLEX-реакция, перенос образцов на спектро-чип, ионизация образцов и анализ масс-спектров) проводили, как описано ранее [22]. SAP-реакция — это дефосфорилирование щелоч-

**Таблица 1.** Частоты аллелей, уровень генетического разнообразия и дискриминационный потенциал X-сцепленных SNP в составе панели XSNPId

| №  | SNP        | Позиция     | MAF   | Аллели | Мульти-плекс | He     | PDf    | PDm    |
|----|------------|-------------|-------|--------|--------------|--------|--------|--------|
| 1  | rs2694742  | 3 127 322   | 0.494 | G:A    | W1           | 0.4999 | 0.6250 | 0.5000 |
| 2  | rs1405303  | 4 120 689   | 0.411 | A:C    | W1           | 0.4842 | 0.6175 | 0.4856 |
| 3  | rs4826682  | 5 119 276   | 0.328 | C:A    | W1           | 0.4408 | 0.5892 | 0.4394 |
| 4  | rs5962008  | 6 325 520   | 0.311 | A:G    | W1           | 0.4286 | 0.5777 | 0.4232 |
| 5  | rs2130835  | 7 470 499   | —     | —      | W1           | —      | —      | —      |
| 6  | rs2404797  | 8 795 378   | 0.441 | A:G    | W1           | 0.4930 | 0.6213 | 0.4928 |
| 7  | rs5934683  | 9 751 474   | 0.378 | C:T    | W1           | 0.4702 | 0.6082 | 0.4692 |
| 8  | rs7888207  | 11 916 455  | 0.491 | G:T    | W2           | 0.4998 | 0.6250 | 0.5000 |
| 9  | rs952076   | 13 946 956  | 0.357 | G:A    | W2           | 0.4591 | 0.5995 | 0.4550 |
| 10 | rs2317327  | 15 407 061  | 0.419 | A:G    | W2           | 0.4869 | 0.6192 | 0.4888 |
| 11 | rs4484871  | 22 751 065  | 0.482 | A:G    | W2           | 0.4994 | 0.6244 | 0.4988 |
| 12 | rs1351260  | 26 948 596  | 0.427 | C:T    | W1           | 0.4893 | 0.6241 | 0.4925 |
| 13 | rs1389433  | 28 128 729  | 0.286 | G:T    | W2           | 0.4084 | 0.5692 | 0.4118 |
| 14 | rs225067   | 29 149 024  | 0.314 | A:G    | W1           | 0.4308 | 0.5924 | 0.4442 |
| 15 | rs4454452  | 30 985 342  | 0.232 | G:A    | W2           | 0.3564 | 0.5222 | 0.3564 |
| 16 | rs3005641  | 34 029 930  | 0.4   | T:C    | W1           | 0.4800 | 0.6130 | 0.4775 |
| 17 | rs761913   | 37 859 510  | 0.306 | G:T    | W1           | 0.4247 | 0.5805 | 0.4270 |
| 18 | rs5963641  | 39 216 082  | 0.452 | T:G    | W2           | 0.4954 | 0.6230 | 0.4961 |
| 19 | rs5917990  | 40 387 891  | 0.375 | C:A    | W2           | 0.4688 | 0.6054 | 0.4646 |
| 20 | rs6609159  | 41 573 566  | 0.497 | T:C    | W1           | 0.5000 | 0.6250 | 0.5000 |
| 21 | rs205847   | 42 791 946  | 0.444 | A:G    | W1           | 0.4937 | 0.6207 | 0.4916 |
| 22 | rs766117   | 43 816 206  | 0.405 | G:A    | W2           | 0.4820 | 0.6159 | 0.4827 |
| 23 | rs5953326  | 49 373 567  | 0.483 | C:T    | W1           | 0.4994 | 0.6246 | 0.4992 |
| 24 | rs5915291  | 50 379 075  | 0.226 | G:T    | W2           | 0.3498 | 0.5319 | 0.3669 |
| 25 | rs4131729  | 51 697 194  | 0.274 | T:C    | W1           | 0.3978 | 0.5583 | 0.3978 |
| 26 | rs11799030 | 53 056 518  | 0.453 | T:C    | W1           | 0.4956 | 0.6230 | 0.4961 |
| 27 | rs4826609  | 54 765 913  | 0.383 | A:G    | W1           | 0.4726 | 0.6107 | 0.4736 |
| 28 | rs6624701  | 63 749 154  | 0.194 | G:T    | W1           | 0.3127 | 0.4889 | 0.3224 |
| 29 | rs471205   | 66 238 317  | 0.196 | G:A    | W1           | 0.3152 | 0.4851 | 0.3188 |
| 30 | rs5919529  | 67 358 208  | 0.268 | C:T    | W2           | 0.3924 | 0.5625 | 0.4032 |
| 31 | rs5937091  | 70 745 323  | 0.206 | A:G    | W1           | 0.3271 | 0.5053 | 0.3387 |
| 32 | rs2207739  | 75 644 692  | 0.259 | A:G    | W2           | 0.3838 | 0.5450 | 0.3819 |
| 33 | rs2411976  | 78 383 858  | 0.419 | A:G    | W1           | 0.4869 | 0.6195 | 0.4893 |
| 34 | rs5969528  | 81 668 873  | 0.394 | A:G    | W1           | 0.4775 | 0.6113 | 0.4745 |
| 35 | rs5922869  | 83 049 688  | 0.262 | C:T    | W2           | 0.3867 | 0.5561 | 0.3951 |
| 36 | rs5968597  | 84 946 832  | 0.381 | A:G    | W2           | 0.4717 | 0.6094 | 0.4712 |
| 37 | rs222108   | 86 910 110  | 0.341 | T:C    | W1           | 0.4494 | 0.6250 | 0.5000 |
| 38 | rs1474970  | 90 394 689  | 0.357 | T:C    | W2           | 0.4591 | 0.6034 | 0.4614 |
| 39 | rs5941047  | 91 431 385  | 0.425 | T:C    | W1           | 0.4888 | 0.6198 | 0.4899 |
| 40 | rs5949581  | 94 756 278  | 0.335 | G:A    | W2           | 0.4456 | 0.5942 | 0.4469 |
| 41 | rs5921682  | 100 130 437 | 0.476 | G:A    | W2           | 0.4988 | 0.6248 | 0.4996 |
| 42 | rs4898334  | 101 387 968 | 0.447 | G:A    | W1           | 0.4944 | 0.6226 | 0.4952 |
| 43 | rs5945770  | 102 594 936 | 0.483 | T:C    | W1           | 0.4994 | 0.6249 | 0.4998 |

Таблица 1. Окончание

| №  | SNP       | Позиция   | MAF   | Аллели | Мульти-плекс | He     | PDf    | PDm    |
|----|-----------|-----------|-------|--------|--------------|--------|--------|--------|
| 44 | rs1285715 | 106308416 | 0.476 | G:T    | W2           | 0.4988 | 0.6248 | 0.4996 |
| 45 | rs5973840 | 107319029 | 0.461 | T:C    | W1           | 0.4970 | 0.6226 | 0.4952 |
| 46 | rs5974348 | 112218701 | 0.483 | T:C    | W1           | 0.4994 | 0.6246 | 0.4991 |
| 47 | rs7058109 | 113228559 | 0.317 | A:G    | W1           | 0.4330 | 0.5853 | 0.4338 |
| 48 | rs9329406 | 115740985 | 0.411 | A:G    | W1           | 0.4842 | 0.6171 | 0.4849 |
| 49 | rs217937  | 118511843 | 0.406 | T:C    | W1           | 0.4823 | 0.6153 | 0.4816 |
| 50 | rs5909923 | 121978166 | 0.5   | C:C    | W1           | 0.5000 | 0.6250 | 0.5000 |
| 51 | rs5977571 | 124496989 | 0.497 | G:A    | W2           | 0.5000 | 0.6250 | 0.5000 |
| 52 | rs5974708 | 125675579 | 0.444 | T:C    | W1           | 0.4937 | 0.6211 | 0.4923 |
| 53 | rs4830049 | 126759219 | 0.363 | T:C    | W1           | 0.4625 | 0.6051 | 0.4641 |
| 54 | rs916208  | 127941251 | 0.491 | G:A    | W2           | 0.4998 | 0.6246 | 0.4992 |
| 55 | rs926640  | 129462353 | 0.429 | G:T    | W2           | 0.4899 | 0.6190 | 0.4884 |
| 56 | rs2797125 | 130815979 | 0.372 | G:A    | W1           | 0.4672 | 0.6085 | 0.4697 |
| 57 | rs17391   | 131910697 | 0.389 | G:A    | W2           | 0.4754 | 0.6120 | 0.4758 |
| 58 | rs5977991 | 133429960 | 0.405 | A:G    | W2           | 0.4820 | 0.6178 | 0.4862 |
| 59 | rs5975695 | 135268469 | 0.408 | C:T    | W1           | 0.4831 | 0.6175 | 0.4856 |
| 60 | rs4825220 | 139214478 | 0.383 | G:T    | W2           | 0.4726 | 0.6184 | 0.4872 |
| 61 | rs4825002 | 140430391 | 0.423 | A:G    | W2           | 0.4881 | 0.6110 | 0.4740 |
| 62 | rs2869922 | 141490609 | 0.437 | C:T    | W2           | 0.4921 | 0.6197 | 0.4896 |
| 63 | rs4825213 | 143435373 | —     | —      | W2           | —      | —      | —      |
| 64 | rs1781486 | 144674621 | 0.274 | C:T    | W2           | 0.3978 | 0.5679 | 0.4101 |
| 65 | rs2504169 | 146358005 | 0.5   | G:G    | W2           | 0.5000 | 0.6250 | 0.5000 |
| 66 | rs614511  | 149537834 | 0.47  | G:A    | W2           | 0.4982 | 0.6245 | 0.4990 |

Обозначения: позиция — позиция нуклеотида согласно референсной последовательности генома человека, сборка 38.2 (GRCh38.p2); MAF — частота редкого аллеля; аллели — указаны частый и редкий аллели; мультиплекс — номер мультиплекса; He — генное разнообразие; PDf — дискриминационный потенциал для женщин; PDm — дискриминационный потенциал для мужчин.

ной фосфатазой (SAP) не включенных в ампликоны дезоксирибонуклеотидтрифосфатов (dNTP). В процессе iPLEX-реакции происходит мультиплексное удлинение пролонгирующих праймеров путем включения такого модифицированного дидезоксирибонуклеотидтрифосфата (ddNTP) с измененной массой, который комплементарен нуклеотиду, находящемуся в полиморфной позиции каждого из SNP. В результате, образуется смесь коротких ПЦР-продуктов, соответствующих аллелям представленных в мультиплексе полиморфных маркеров. Реакционную смесь очищали от солей, накопившихся при мультиплексной ПЦР и iPLEX-реакции, обрабатывая катионной смолой “SpectroCLEAN” (“Agena Bioscience”, Германия). Образцы перенесли на спектро-чип и ионизировали на приборе “Sequenom MassARRAY 4”. Для анализа масс-спектров в реальном времени, а также для первичной обработки и документирования результатов опыта использовали программное обеспе-

чение “MassARRAY TYPHER 4.0” (“Agena Bioscience”, Германия).

Для верификации генотипов брали несколько образцов (по 2–3 каждого их трех генотипов) и проводили терминирующее секвенирование по Сэнгеру с флуоресцентно-мечеными терминаторами каждого SNP, с последующим капиллярным гель-электрофорезом на генетическом анализаторе “ABI PRISM 3730” (“Life Technologies”). Всего секвенировано около 600 образцов. Расхождений в характеристиках генотипов, полученных по результатам MALDI-TOF и путем прямого секвенирования, не обнаружено.

**Статистическая обработка результатов.** Применяли стандартные методы популяционной генетики и математической статистики. Неравновесие по сцеплению между парами и группами маркеров оценивали при помощи коэффициента  $D'$ , предложенного Левонтином, и коэффициента корреляции Пирсона ( $r^2$ ) в пакете “Haploview”,

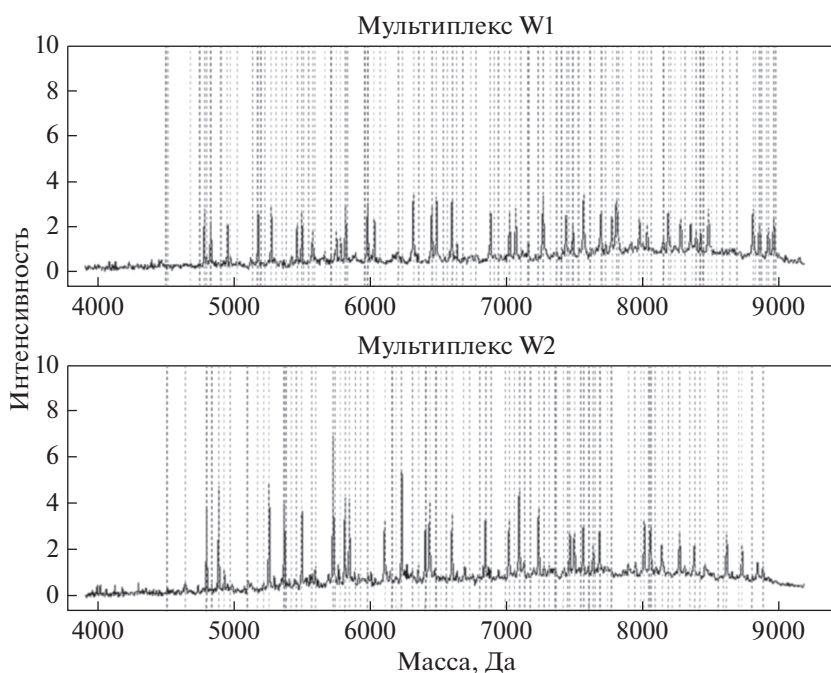


Рис. 1. Масс-спектры мультиплексов W1 и W2 панели XSNPId в одном из образцов.

как описано ранее [26–28], а также при помощи точного теста с использованием цепей Маркова, реализованного в пакете “Arlequin” [26, 29, 30]. Идентификационные характеристики системы X-сцепленных однонуклеотидных маркеров оценивали, используя стандартные для криминалистики и судебной медицины показатели [9, 28].

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

### Анализ масс-спектров и генотипирование

На рис. 1 приведены примеры масс-спектров мультиплексов W1 и W2 панели XSNPId одного из образцов. Анализируемые молекулы распределены по массе от 4500 до 9000 Да. При этом в 36-плексе W1 детектируется 108 возможных продуктов — по три молекулы на каждый SNP, включая исходный пролонгирующий праймер, и два продукта iPLEX-реакции, состоящие из пролонгирующего праймера с присоединенным нуклеотидом, соответствующим одному из двух альтернативных аллелей. В мультиплексе W2, состоящем из 30 SNP, определяется 90 возможных продуктов. Позиции пиков масс-спектра всех возможных продуктов показаны вертикальными прерывистыми линиями. В данном образце (мужчина) успешно прошла амплификация и элонгация всех 66 маркеров, и наблюдается 36 пиков специфических продуктов iPLEX-реакции (36 SNP в гемизиготном состоянии) для мультиплекса W1 и 30 пиков (30 SNP в гемизиготе) — для второго мультиплекса. При этом высота пиков для непролонгирован-

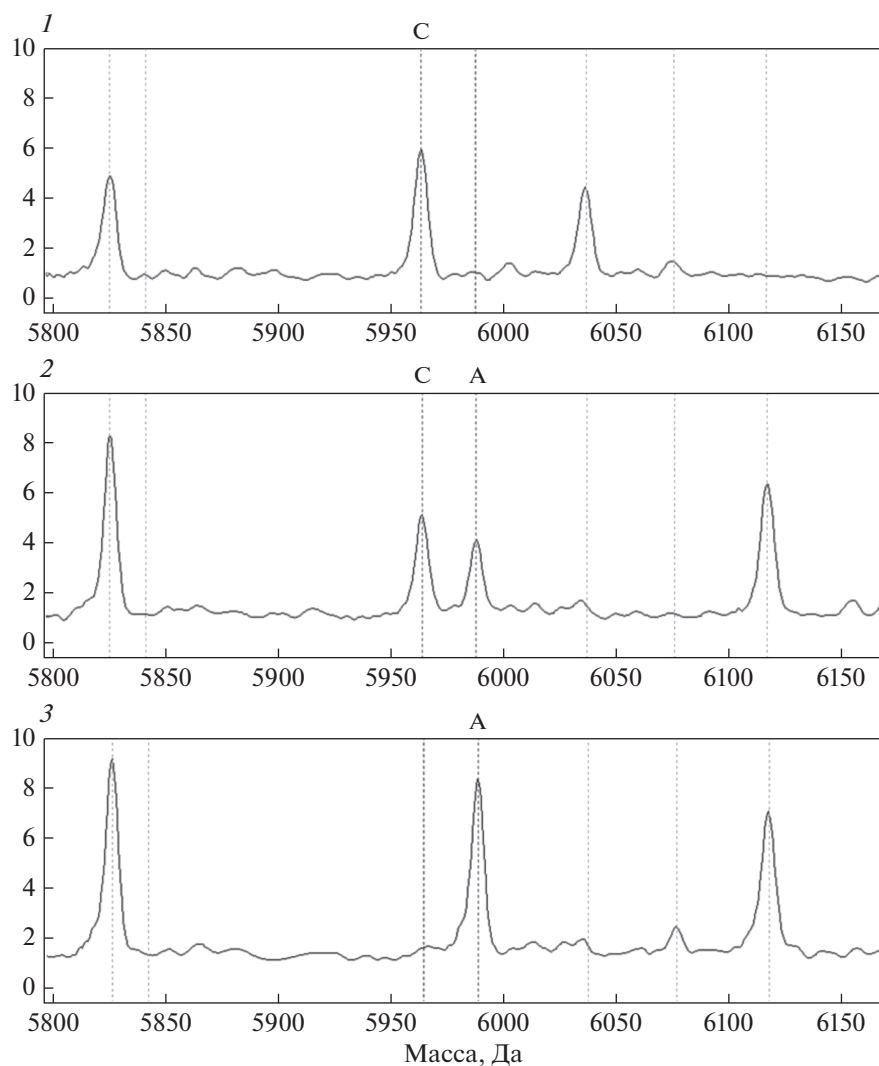
ных праймеров не превышает фоновой интенсивности сигнала, что свидетельствует о высокой эффективности iPLEX-реакции и о включении большей части молекул праймеров в специфичный iPLEX-продукт.

На рис. 2 представлен пример генотипирования конкретного SNP на фрагменте масс-спектра реакции W1 в трех образцах с тремя различными генотипами по rs1405303. Пик для аллеля С соответствует продукту iPLEX-реакции с массой 5963 Да, пик аллеля А — продукту с массой 5988 Да. Различия в 15 Да между специфичными для аллелей молекулами позволяет легко идентифицировать аллели и генотипы даже при ручном анализе масс-спектра.

Автоматическое чтение генотипов на основе мультиплексного масс-спектра в пакете “MassARRAY TYPER 4.0” опирается на определение соотношения высоты пиков продуктов iPLEX-реакции двух аллелей. Пример кластеризации образцов в соответствии с генотипами для rs217937 в реакции приведен на рис. 3. В данном случае успешно определены генотипы всех 95 образцов. Один образец без пиков (аллели не найдены), находящийся в нулевой точке осей координат, соответствует контрольному образцу без ДНК.

### Эффективность генотипирования и оценка характеристик панели XSNPId

Качество разработанной панели маркеров (чувствительность, специфичность, воспроиз-



**Рис. 2.** Фрагмент масс-спектра реакции W1 (36-плекса) для трех образцов с тремя различными генотипами по rs1405303. Пик для аллеля С соответствует продукту iPLEX-реакции массой 5963 Да, пик аллеля А – продукту массой 5988 Да. 1 – С; 2 – СА; 3 – А.

димось, точность генотипирования) оценивали, а генотипы верифицировали – согласно рекомендациям Научной рабочей группы по методам анализа ДНК (SWGAM) для валидации новых методов ДНК-идентификации [31], а также согласно указаниям компании “Promega” по внутренней валидации маркерных систем для судебно-медицинских лабораторий [32].

Чувствительность панели оценивали по доле определенных генотипов из всех возможных (“call rate”) при различных концентрациях ДНК-матрицы (от 20 нг до 500 пг). В случае двух из включенных в панель SNP (rs2130835 из W1 (36-плекса) и rs4825213 из W2 (30-плекса)) наблюдается низкое значение “call-rate” (менее 60%), и они были исключены из дальнейшего анализа. При стандартном для нашего протокола количестве матрицы

(5 нг геномной ДНК) “call rate” для мультиплекса W1 составляет 99.27%, для мультиплекса W2 – 97.12 всех возможных генотипов. Суммарное значение показателя “call rate” для 64 SNP составляет 98.29% (определено 11323 генотипов из 11520).

Для оценки возможности использования нашего подхода в криминалистике, когда доступно небольшое количество материала, мы тестировали эффективность генотипирования при различных концентрациях ДНК, меньших, чем стандартная для нашего протокола. При уменьшении количества матрицы в реакции до 1 нг значение “call rate” падает до 89% – в случае мультиплекса W1 и до 87% – в случае мультиплекса W2. При содержании ДНК в реакции 500 пг доля полученных генотипов уменьшается до 83%. Таким образом, максимальная чувствительность панели наблю-

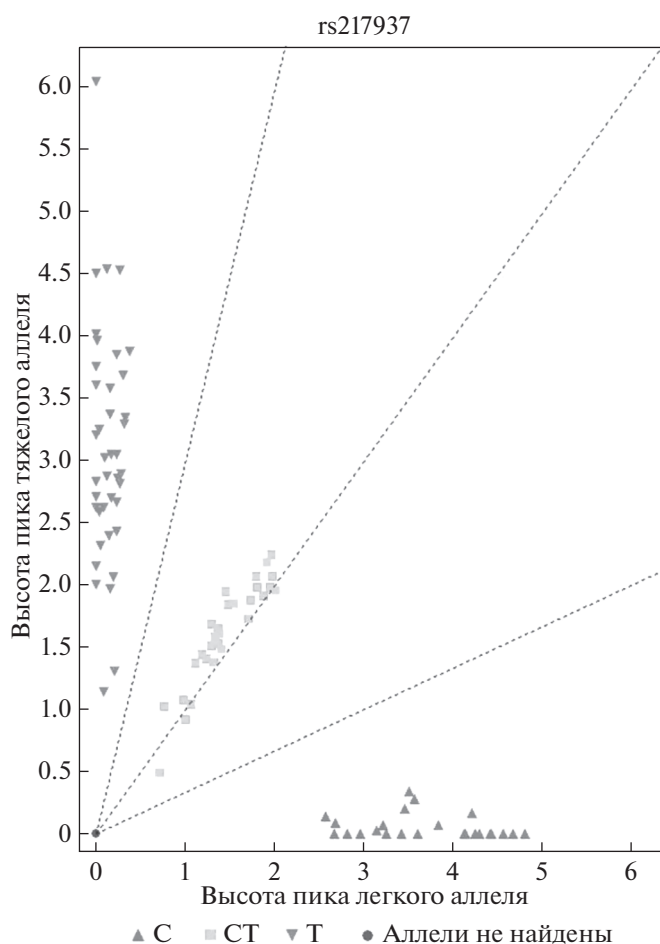


Рис. 3. Кластеризация образцов по генотипам в программе “MassARRAY TYPER 4.0” на примере rs217937.

дается при анализе 10 нг ДНК-матрицы (по 5 нг на каждый мультиплекс).

Специфичность тест-системы оценивали путем определения отношения интенсивности пиков ПЦР-продукта к неспецифичным пикам, при учете значения соотношения сигнала к шуму (SNR) в 10 контрольных образцах, которое рассчитывается для каждого пика в программном обеспечении “MassARRAY TYPER 4.0”. Значения SNR колебались в пределах 9.1–35.9, т.е. доля специфичного продукта составляет как минимум 90% для любого из пиков. Тем самым все пики, соответствующие аллелям включенных в панель SNP, определяются однозначно, и специфичность панели составляет 100%.

Точность генотипирования оценивали путем верификации генотипов прямым секвенированием (как описано в разделе “Экспериментальная часть”); она равна 100%.

Воспроизводимость результатов оценивали по результатам серии повторного генотипирования 10 контрольных образцов ДНК по всем маркерам,

входящим в тест-систему. Все генотипы воспроизводились в 100% случаев.

#### Анализ частот аллелей и неравновесия по сцеплению

Частоты аллелей и уровень генетического разнообразия 64 исследованных SNP приведены в табл. 1. Наименьшая частота редкого аллеля в русской популяции характерна для rs6624701 (0.194), наибольшая — для rs5909923 и rs2504169 (0.5). Генное разнообразие ( $H_e$ ) варьирует от 0.31 до 0.5. Средний для 64 SNP уровень генного разнообразия составляет 0.461. Уровень генетической вариативности русской популяции по исследованным маркерам соответствует уровню разнообразия в европеоидных популяциях (из проекта HarMap, 0.447–0.462) и превышает таковой в монголоидных и негроидных популяциях (0.310–0.419). Отклонений от равновесия Харди-Вайнберга в распределении генотипов у женщин ни по одному из маркеров не наблюдается.

X-хромосомные маркеры для панели XSNPId подбирали исходя из их равномерного распределения по хромосоме и отсутствия сцепления. Физическая дистанция между соседними маркерами варьирует от 993 тыс. до 8.9 млн.п.н. (среднее расстояние между двумя соседними SNP — 2.252 т.п.н.). Анализ неравновесия по сцеплению в пакете “Haploview” показал, что возможно сцепление в девяти парах маркеров ( $LOD$ -балл  $> 2$ ;  $D'$  = 0.237–0.507). Однако SNP ни в составе пар, ни в более протяженных конфигурациях не формируют блоки неравновесия по сцеплению. Все пары, показавшие возможное сцепление, состоят из маркеров, расположенных в нескольких десятках миллионах пар нуклеотидов друг от друга, и характеризуются низкими значениями коэффициента корреляции Пирсона ( $r^2$  = 0.016–0.073). Точный тест на неравновесие по сцеплению с учетом поправки Бонферрони на множественные сравнения показывает, что лишь одно парное сочетание маркеров (rs1351260–rs2411976) характеризуется значимой величиной неравновесия по сцеплению ( $p$  = 0.00000). Расстояние на хромосоме между этими маркерами составляет 43.79 млн.п.н., и физическое сцепление между ними — в отсутствие сцепления с промежуточными маркерами — можно исключить. Вероятно, гаметическое неравновесие между этими двумя удаленными SNP объясняется особенностями выборки. Тем не менее мы решили исключить один из входящих в пару маркеров, rs2411976, характеризующийся меньшей частотой минорного аллеля, из дальнейшего анализа дифференцирующего потенциала панели X-сцепленных SNP.





технологий и прогрессом в описании генетической вариабельности в геноме человека – очевидно, приведут к замене STR-маркеров на более простые и технологичные системы. Панели SNP-маркеров для криминалистики и судебной медицины, хотя пока не входят в официальные стандарты, активно разрабатываются в мире. Сравнительные характеристики некоторых панелей для ДНК-идентификации, основанные на разных технологических платформах, приведены в табл. 3.

Первые из разработанных SNP-панелей включали 40–50 SNP, которые также обеспечивают ве-

роятность идентификации, сравнимую со стандартными STR-панелями, однако технологичность их использования была несовершенной. Например, панель из 49 SNP, предложенная ранее [34, 35], использует технологию Genplex, основанную, как и фрагментный анализа STR, на капиллярном гель-электрофорезе. Разработанный ранее набор 45 несцепленных аутосомных SNP обеспечивает вероятность идентификации ( $PI$ ) в пределах от  $10^{-16}$  до  $10^{-19}$ , однако основан на моноплексной амплификации с помощью TaqMan-проб [36, 37].

**Таблица 3.** Сравнительные характеристики некоторых панелей для ДНК-идентификации

| Панель                          | Тип маркеров             | Число маркеров* | Технологическая платформа                        | Вероятность совпадения генотипов ( $PI$ ) | Ссылка           |
|---------------------------------|--------------------------|-----------------|--|---|------------------|
| CODIS                           | aSTR                     | 14/13           | Фрагментный анализ/капиллярный гель-электрофорез | $5.02 \times 10^{-16}$                    | [44]             |
| ESS                             | aSTR                     | 17/16           | Фрагментный анализ/капиллярный гель-электрофорез | $3.04 \times 10^{-16}$                    | [44]             |
| PowerPlex 16                    | aSTR                     | 16/15           | Фрагментный анализ/капиллярный гель-электрофорез | $2.81 \times 10^{-17}$                    | [9]              |
| PowerPlex Fusion                | aSTR                     | 24/23           | Фрагментный анализ/капиллярный гель-электрофорез | $6.58 \times 10^{-29}$                    | [44]             |
| SNPforID                        | aSNP                     | 52              | SBE/капиллярный гель-электрофорез                | $10^{-17} \times 10^{-20}$                | [38]             |
| Genplex 49-plex                 | aSNP                     | 49              | Genplex/капиллярный гель-электрофорез            | $10^{-17} \times 10^{-19}$                | [35]             |
| К. Кидд                         | aSNP                     | 45              | TaqMan   | $10^{-16} \times 10^{-19}$                | [36, 37]         |
| iPLEX Pro Sample ID Panel       | aSNP                     | 52/44           | MALDI-TOF MS                                     | —**                                       | [39]             |
| HID-Ion AmpliSeq Identity Panel | aSNP + YSNP              | 124/90          | MPS  | $1 \times 10^{-31} - 1 \times 10^{-33}$   | [41]             |
| ForenSeq DNA Signature Prep Kit | aSTR + YSTR + XSTR + SNP | 230/94          | MPS  | н.д.                                      | [42]             |
| 53 XSNP                         | XSNP                     | 67/52           | MALDI-TOF MS                                     | $6.9 \times 10^{-16} - 1 \times 10^{-19}$ | [43]             |
| XSNPid                          | XSNP                     | 66/63           | MALDI-TOF MS                                     | $2 \times 10^{-18} - 9 \times 10^{-27}$   | Настоящая работа |

\* Указано общее число маркеров и число маркеров для ДНК-идентификации. Некоторые панели включают также дополнительные маркеры для определения пола, оценки количества ДНК и других целей.

\*\* Не предназначена для ДНК-идентификации в криминалистике и судебной медицине. Производитель позиционирует панель как систему для идентификации образцов ДНК в биобанках. Сокращения: aSTR – аутосомные микросателлитные маркеры; MALDI-TOF MS – масс-спектрометрия методом матрично-активированной лазерной десорбции/ионизации с измерением времени пролета; MPS – массовое параллельное секвенирование; SBE – удлинение на один нуклеотид (single base extension); XSTR – X-сцепленные STR; YSTR – Y-сцепленные STR; н.д. – нет данных.

Панель “SNPforID”, предложенная европейским консорциумом “EU GROWTH”, включает 52 маркера и генотипируется в одной мультиплексной ПЦР и двух реакциях дотраивания одного нуклеотида (SBE), как и наша панель. Финальная детекция продуктов проводится с помощью капиллярного гель-электрофореза [38]. Компания “Sequenom” на основе панели “SNPforID” разработала панель “iPLEX Sample ID panel”, включив в нее 47 из 52 SNP [39]. Эта панель, которая генотипируется в одной многолокусной реакции, представляет собой первый набор для идентификации, использующий технологию MALDI-TOF-масс-спектрометрии ДНК. Панель “iPLEX Sample ID panel”, однако, предназначена для идентификации образцов в биологических банках, а не для ДНК-идентификации в криминалистике. Потенциально, она может быть применена и для идентификации образцов в криминальных случаях, но, как показано позднее, величина “call rate” панели резко снижается при анализе деградированных или загрязненных образцов и при количестве матрицы менее 10 нг [40]. В отличие от результатов упомянутой работы, наши данные свидетельствуют о высокой эффективности нашей панели XSNPid при стандартном количестве ДНК в 5 нг, а также и достаточно достоверную эффективность в случае использования существенно меньших количеств ДНК (0.5–1 нг).

Появление относительно недорогих приборов, использующих технологию массового параллельного секвенирования (MPS), привело к попыткам разработать панели SNP-маркеров для ДНК-идентификации на основе MPS. Сейчас на рынке имеется две разработки – “HID-Ion AmpliSeq Identity Panel”, ориентированная на персональные MPS-секвенизаторы компании “Life Technologies” [21, 41], и “ForenSeq DNA Signature Prep Kit” для приборов компании “Illumina” [42]. Первая из них состоит из 124 SNP-маркеров, включая 48 SNP из панели “SNPforID”, 43 маркера из панели K. Кидда и 34 Y-хромосомных SNP, позволяющих определить принадлежность образца к Y-хромосомной кладе верхнего уровня. Совокупная вероятность совпадения генотипов в этой панели варьирует – для разных популяций – в пределах от  $1 \times 10^{-31}$  до  $1 \times 10^{-33}$  [39] (табл. 2).

Вторая панель, “ForenSeq DNA Signature Prep Kit”, включает 230 разнообразных маркеров (аутосомные, Y- и X-сцепленные STR, 94 SNP для идентификации, SNP для определения биогеографического происхождения и предсказания некоторых фенотипов).

Несмотря на существенный прогресс в разработке подходов к ДНК-идентификации и на появление SNP-панелей для этих целей, ниша для X-хромосомных систем остается свободной как по спектру задач, которые можно решать с ее помощью, так и по состоянию практических разра-

боток. Панели XSNP, создаваемые на базе традиционных технологий (ПЦР в реальном времени, капиллярный гель-электрофорез), включают не более 25 маркеров и не достигают уровня информативности систем, основанных на STR-маркерах [16–18]. Наиболее близким аналогом нашей разработки является панель из 52 XSNP-маркеров, предложенная китайскими авторами [43] (табл. 2). Эта панель также основана на масс-спектрометрии MALDI-TOF, однако она имеет небольшую емкость мультиплексов (17–18 маркеров) и предполагает проведение четырех отдельных мультиплексных ПЦР и четырех – iPLEX-реакций. В нашей же панели большее число маркеров собрано лишь в два мультиплекса. Примечательно, что наша панель пересекается с упомянутой [43] лишь по одному маркеру (rs471205). Вероятность дискриминации неродственных индивидов при использовании нашей панели также на несколько порядков выше, чем при использовании китайской разработки. В их панели вероятность случайного совпадения генотипов ( $PI$ ) для женщин составляет  $1 \times 10^{-19}$  (в нашей панели –  $9 \times 10^{-27}$ ), а для мужчин –  $6.9 \times 10^{-16}$  (в нашей –  $2 \times 10^{-18}$ ).

Таким образом, представленная в настоящей работе панель X-сцепленных однонуклеотидных маркеров превосходит все имеющиеся аналоги (кроме основанной на MPS панели “HID-Ion AmpliSeq Identity Panel”) по дискриминирующей способности при индивидуальной ДНК-идентификации у женщин, а у мужчин превышает существующие показатели стандартных STR-систем. Следует, однако, заметить, что порог дискриминирующего потенциала, который может быть достигнут при использовании маркеров стандартов CODIS и ESS ( $10^{-16}$ ), на практике достаточен для подавляющего большинства задач ДНК-идентификации. При сравнении с ближайшими аналогами наборов X-хромосомных SNP, панель XSNPid – помимо более высокой дискриминирующей способности – обладает также большей емкостью мультиплексов, что делает ее более дешевой и требующей меньших затрат времени. Панель XSNPid может оказаться универсально применимой, поскольку обеспечивает высокую дискриминирующую способность в анализе, по крайней мере, основных мировых популяций европейского и азиатского происхождения и может быть использована не только в России.

Учитывая показатели чувствительности панели и потребность в относительно больших количествах ДНК для анализа, панель не может быть рекомендована к использованию для решения криминалистических задач (например, для анализа ДНК из экстремально малых количеств биологического материала с места преступления), однако она вполне может быть применена в тех случаях, когда биологического материала доста-

точно (например, при массовых катастрофах, терактах, в военных конфликтах).

Панель XSNPId может быть использована не только для индивидуальной идентификации, но и для сложных случаев определения родства. Высокий уровень генетического разнообразия маркеров, входящих в панель, и технологичность ее использования делают ее удобным инструментом также для популяционно-генетических исследований, пример которых будет представлен в одной из наших следующих публикаций.

Работа получила финансовую поддержку 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. Степанов В.А. 2002. Этногеномика населения Сибири и Средней Азии. Томск: Печатная Мануфактура. 242 с.
2. Степанов В.А., Харьков В.Н., Пузырев В.П. 2006. Эволюция и филогеография линий Y-хромосомы человека. *Вестник ВОГиС*. **10**, 57–73.
3. Хитринская И.Ю., Харьков В.Н., Степанов В.А. 2010. Генетическое разнообразие X-хромосомы в популяциях коренных этносов Сибири: структура неравновесия по сцеплению и филогеография гаплотипов локуса ZFX. *Молекуляр. биология*. **44**, 804–815.
4. Вагайцева К. В., Харьков В. Н., Черпинская К. В., Хитринская И.Ю., Степанов В.А. 2015. Генетическая вариабельность X-сцепленных STR-маркеров в популяциях Сибири. *Молекуляр. биология*. **49**, 305–312.
5. Dobyns W.B., Filauo A., Tomson B., Chan A.S., Ho A.W., Ting N.T., Oosterwijk J.C., Ober C. 2004. Inheritance of most X-linked traits is not dominant or recessive, just X-linked. *Am. J. Med. Genetics*. **129**, 136–143.
6. Burdett T., Hall P.N., Hasting E., Hindorff L.A., Junkins H.A., Klemm A.K., MacArthur J., Manolio T.A., Morales J., Parkinson H., Welter D. 2015. A Catalog of Published Genome-Wide Association Studies. Available at: [www.genome.gov/gwastudies](http://www.genome.gov/gwastudies). Accessed [date of access: 18.07.2015].
7. Szibor R. 2007. X-chromosomal markers: past, present and future. *Forensic Sci. Int. Genet.* **1**, 93–99.
8. Lancia M., Severini S., Coletti A., Margiotta G., Dobosz M., Carnevali E. 2011. Using X-chromosomal markers in rape investigation. *Forensic Sci. Int. Genet.* **3**, e55–e56.
9. Степанов В.А., Балановский О.П., Мельников А.В., Харьков В.Н., Тяжелова Т.В., Ахметова В.Л., Жукова О.В., Шнейдер Ю.В., Шильникова Н.Н., Боринская С.А., Марусин А.В., Спиридонова М.Г., Симонова К.В., Хитринская И.Ю., Раджабов М.О., Романов А.Г., Штыгашева О.В., Кошель С.М., Балановская Е.В., Рыбакова А.В., Хуснутдинова Э.К., Пузырев В.П., Янковский Н.К. 2011. Характеристика популяций Российской Федерации по панели пятнадцати локусов, используемых для ДНК-идентификации и в судебно-медицинской экспертизе. *Acta Naturae*. **3**, 59–71.
10. Stepanov V.A., Melnikov A.V., Lash-Zavada A.Y., Kharkov V.N., Borinskaya S.A., Tyazhelova T.V., Zhukova O.V., Schneider Y.V., Shil'nikova I.N., Puzyrev V.P., Rybakova A.A., Yankovsky N.K. 2010. Genetic variability of 15 autosomal STR loci in Russian populations. *Legal Medicine*. **12**, 256–258.
11. Шорохова Д.А., Степанов В.А., Удовенко Ю.Д., Новоселов В.П., Пузырев В.П. 2005. Генетическая вариабельность и дискриминирующий потенциал четырех микросателлитных локусов в русской популяции. *Молекуляр. биология*. **39**, 965–970.
12. Costa A., Morais P., Vieira da Silva C., Matos S., Marques dos Santos R., Espinheira R., Costa Santos J., Amorim A. 2014. X-chromosome STR markers data in a Cabo Verde immigrant population of Lisboa. *Mol. Biol. Rep.* **41**, 2559–2569.
13. Gill P., Ivanov P.L., Kimpton C., Piercy R., Benson N., Tully G., Evett I., Hagelberg E., Sullivan K. 1994. Identification of the remains of the Romanov family by the DNA analysis. *Nat. Genet.* **6**, 130–135.
14. Gill P., Werrett D.J., Budowle B., Guerrieri, R. 2004. An assessment of whether SNPs will replace STRs in national DNA databases—joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDM). *Sci. Justice*. **44**, 51–53.
15. Krawczak M. 1999. Informativity assessment for biallelic single nucleotide polymorphisms. *Electrophoresis*. **20**, 1676–1681.
16. Li C., Zhang S., Zhao S., Liu Y., Lin Y. 2010. Analysis of 14 highly informative SNP markers on X chromosome by TaqMan SNP genotyping assay. *Forensic Sci. Int. Genetics*. **4**, e145–e148.
17. Tomas C., Sanchez J.J., Castro J.A., Børsting C., Morling N. 2010. Forensic usefulness of a 25 X-chromosome single-nucleotide polymorphism set. *Transfusion*. **50**, 2258–2265.
18. Pereira V., Tomas C., Amorim A., Morling N., Gusmao L., Prata M.J. 2011. Study of 25 X-chromosome SNPs in Portuguese. *Forensic Sci. Int. Genetics*. **5**, 336–338.
19. Berglund E.C., Kiialainen A., Syvänen A.-C. 2011. Next-generation sequencing technologies and applications for human genetic history and forensics. *Investigative Genetics*. **2**, 23.
20. Yaran Yang, Bingbing Xie, Jiangwei Yan. 2014. Application of Next-generation Sequencing Technology in Forensic Science. *Genom. Proteom. Bioinformatics*. **12**, 190–197.
21. Churchill J.D., Chang J., Ge J., Rajagopalan S.C., Wootton C.W. 2015. Blind study evaluation illustrates utility of the Ion PGM system for use in human identity DNA typing. *Croat. Med. J.* **56**, 218–229.
22. Степанов В.А., Трифонова Е.А. 2013. Мультиплексное генотипирование однонуклеотидных полиморфных маркеров методом MALDI-TOF масс-спектрометрии: частоты 56 SNP в генах иммунного ответа в популяциях человека. *Молекуляр. биология*. **47**, 976–986.
23. Cole R.B. (Ed.). 2012. Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicalities, and Biological Applications, Second Edition. John Wiley and Sons.
24. <http://www.ncbi.nlm.nih.gov/projects/SNP/>
25. <http://hapmap.ncbi.nlm.nih.gov/>

26. Guo S., Thompson E. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*. **48**, 361–372.
27. Трифонова Е.А., Еремина Е.Р., Урнов Ф.Д., Степанов В.А. 2012. Генетическое разнообразие и структура неравновесия по сцеплению гена MTHFR в популяциях Северной Евразии. *Acta Naturae*. **4**, 80–96.
28. Степанов В.А., Харьков В.Н., Трифонова Е.А., Марусин А.В. 2014. *Методы статистического анализа в популяционной и эволюционной генетике человека. Учебно-методическое пособие. Серия “Наследственность и здоровье”. Вып. 12.* Томск: Печатная мануфактура. 100 с.
29. Slatkin M. 1994. Linkage disequilibrium in growing and stable populations. *Genetics*. **137**, 331–336.
30. Excoffier L., Lischer H. 2011. Arleqion ver. 3.5. An Integrated Software Package for Population Genetics Data Analysis. Swiss Institute of Bioinformatics.
31. Scientific Working Group on DNA Analysis Methods. 2012. Validation Guidelines for DNA Analysis Methods. 13 p.
32. Promega. 2013. Internal Validation Guide of Autosomal STR Systems for Forensic Laboratories. Reference Manual. 4 P.
33. <https://worldwide.promega.com/products/genetic-identity/str-analysis-for-forensic-and-paternity-testing/powerplex-fusion-str-kits/>
34. Sanchez J.J., Phillips C., Borsting C., Balogh K., Bogus M., Fondevila M., Harrison C.D., Musgrave-Brown E., Salas A., Syndercombe-Court D., Schneider P.M., Carracedo A., Morling N. 2006. A multiplex assay with 52 single nucleotide polymorphisms for human identification. *Electrophoresis*. **27**, 1713–1724.
35. Philips C., Fang R., Ballard D., Fondevila M., Harrison C., Hyland F., Musgrave-Brown E., Proff C., Ramos-Luis E., Sobrino B., Carracedo A., Furtado M.R., Syndercombe-Court D., Schneider P.M.; SNPforID Consortium. 2007. Evaluation of the Genplex SNP typing system and a 49plex forensic marker panel. *Forensic Sci. Int. Genetics*. **1**, 180–185.
36. Pakstis A.J., Speed W.C., Fang R., Hyland F.C.L., Furtado M.R., Kidd J.R., Kidd K.K. 2010. SNPs for a universal individual identification panel. *Hum Genet*. **127**, 315–324.
37. Kidd K.K. 2011. Population Genetics of SNPs for Forensic Purposes (Updated). NIJ Final Report. 103 p.
38. Musgrave-Brown E., Ballard D., Balogh K., Bender K., Berger B., Bogus M., Borsting C., Brion M., Fondevila M., Harrison C., Oguzturun C., Parson W., Phillips C., Proff C., Ramos-Luis E., Sanchez J.J., Sánchez Diz P., Sobrino Rey B., Stradmann-Bellinghausen B., Thacker C., Carracedo A., Morling N., Scheithauer R., Schneider P.M., Syndercombe-Court D. 2007. Forensic validation of the SNPforID 52plex assay. *Forensic Sci. Int. Genet.* **1**, 186–90.
39. Agena Bioscience. 2014. iPLEX Pro Sample ID Brochure. 4 P.
40. Johansen P., Andersen J.D., Borsting C., Morling N. 2013. Evaluation of the iPLEX Sample ID Plus Panel designed for the Sequenom MassARRAY system. A SNP typing assay developed for human identification and sample tracking based on the SNPforID panel. *Forensic Sci. Int. Genet.* **7**, 482–487.
41. Life technologies. 2014. HID-Ion AmpliSeq Identity Panel Data Sheet. 2 P.
42. Illumina. 2014. ForenSeq DNA Signature Prep Kit Data Sheet. 4 P.
43. Li Y., Liu Y., Lin Y. 2015. Typing of 67 SNP Loci on X Chromosome by PCR and MALDI-TOF MS. *Res. Genetics*. **2015**, Article ID 374688, 9 pages.
44. Promega. 2013. PowerPlex® Fusion Systems Data Sheet. 5 P.

## PANEL OF X-LINKED SINGLE NUCLEOTIDE POLYMORPHIC MARKERS FOR DNA IDENTIFICATION (XSNPId) BASED ON MULTIPLEX GENOTYPING BY MULTI-LOCUS PCR AND MALDI-TOF MASS SPECTROMETRY

V. A. Stepanov<sup>a,\*</sup>, K. V. Vagaitseva<sup>a,b</sup>, V. N. Kharkov<sup>a,b</sup>, A. A. Cherednichenko<sup>a</sup>, A. V. Bocharova<sup>a</sup>

<sup>a</sup>Research Institute for Medical Genetics, Tomsk, 634050 Russia

<sup>b</sup>National Research Tomsk State University, Tomsk, 634050 Russia

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

X-linked genetic markers in humans are used in population genetics, medical genetics and in forensic DNA identification. XSNPId panel, consisted of 66 highly polymorphic unlinked X-chromosomal markers, and protocol of its genotyping by multi-locus PCR and MALDI-TOF mass spectrometry have been developed. XSNPId panel is genotyped in two multiplexes (36 and 30 markers in each). Genotyping protocol developed in the current study provides effective genotypes reading. Average call rate was 98.29%. High level of gene diversity (0.461) for X-linked SNPs included into the panel was revealed in a Russian population. 63 out of 66 SNPs, demonstrated high genotypic efficiency and independent inheritance, are suitable for DNA identification purposes. XSNPId panel shows very high discriminating power in a Russian population. Probability of genotypes identity in two random unrelated individuals is  $9 \times 10^{-27}$  in females and  $2 \times 10^{-18}$  in males. Comparing to closest XSNP analogues, XSNPId panel besides higher discriminating power is also characterized by higher multiplexing capacity, being therefore more effective and less time-consuming. XSNPId panel in addition to individual DNA identification purposes can be a useful tool for population genetic studies.

**Keywords:** single nucleotide polymorphisms (SNPs), X-chromosome, MALDI-TOF mass-spectrometry, multiplex genotyping, DNA identification, population genetics, genetic diversity, human population.