
ГЕНОМИКА. ТРАНСКРИПТОМИКА. ПРОТЕОМИКА

УДК 577.21:575.17:340.6

ГЕНЕТИЧЕСКАЯ ВАРИАБЕЛЬНОСТЬ И ДИСКРИМИНИРУЮЩИЙ ПОТЕНЦИАЛ ЧЕТЫРЕХ МИКРОСАТЕЛЛИТНЫХ ЛОКУСОВ ДНК В РУССКОЙ ПОПУЛЯЦИИ

© 2005 г. Д. А. Шорохова, В. А. Степанов^{1*}, Ю. Д. Удовенко, В. П. Новоселов, В. П. Пузырев¹

Государственное учреждение здравоохранения Новосибирская областная судебно-медицинская экспертиза,
Новосибирск, 630087

¹Государственное учреждение Научно-исследовательский институт медицинской генетики Томского научного центра Сибирского отделения Российской академии медицинских наук, Томск, 634050

Поступила в редакцию 19.05.2005 г.

Исследовано распределение аллелей и генотипов четырех STR-локусов (LPL, vWA, FES/FPS и F13B), применяемых в судебно-медицинской практике, среди 200 представителей русского населения, проживающего на территории Сибири. Охарактеризовано генетическое и молекулярное разнообразие полиморфных систем в сравнении с европеоидным населением США. Показан высокий дискриминирующий потенциал системы из четырех исследованных STR-локусов ($PD = 0.99975$). Сравнительный анализ генетического разнообразия исследованных локусов у русского населения и у белых американцев свидетельствует о значимых генетических различиях двух популяций, что делает неприемлемым использование данных по частотам этих локусов у европеоидного населения США в судебно-медицинских экспертизах, проводимых в России. Эти данные могут быть использованы в качестве референтных в работе судебно-медицинских лабораторий Сибири.

Ключевые слова: популяции человека, микросателлиты, ДНК-идентификация, вариабельность генома.

GENETIC VARIABILITY AND DISCRIMINATING POWER OF FOUR MICROSATELLITE LOCI IN RUSSIAN POPULATION, by D. A. Shorokhova, V. A. Stepanov^{1*}, Yu. D. Udovenko, V. P. Novoselov, V. P. Puzyrev¹ (Novosibirsk Regional Forensic Medicine Bureau, Novosibirsk, 630087 Russia; ¹Research Institute for Medical Genetics, Tomsk Scientific Centrem Siberian Branch of the Russian Academy of Medical Sciences, Tomsk, 634050 Russia, *e-mail: vadimst@img.tsu.ru). Allele and genotype frequencies of 4 STR loci (LPL, vWA, FES/FPS и F13B), used in forensic medicine, were analyzed in Russian Siberian population. Genetic and molecular diversity of these polymorphic systems were characterized in comparison with US Caucasoid population. High discriminating power ($PD = 0.99975$) of the system of four studied STR loci was shown. Comparative analysis of genetic diversity in Russian population and Caucasoid US population revealed the significant differences between two populations and demonstrated that STR frequency data for US population should not be used for forensic expertise in Russia. The data obtained in the current investigation may be used as reference data for forensic medicine laboratories in Siberia.

Key words: Human populations, microsatellites, DNA identification, genome diversity.

Короткие tandemно повторяющиеся последовательности (STR – Short Tandem Repeats), или микросателлиты, находят все более широкое применение в фундаментальных (популяционная генетика, генетическое картирование) и прикладных (судебно-медицинской практика) отраслях генетических исследований [1–7]. К классу микросателлитов относят tandemные повторы, содержащие от двух до десяти нуклеотидов в повторяющейся единице.

Число копий таких повторов часто варьирует, изменяясь от одного до нескольких десятков. Наибольшее применение получили STR-локусы с

длиной повторяющейся единицы в четыре п.н. Интервал длин таких микросателлитных рядов в различных аллелях составляет всего 100–300 п.н., что увеличивает вероятность их обнаружения даже в сильно деградированной ДНК, зачастую используемой для идентификации личности в судебно-медицинской экспертизе.

Наследуются аллельные варианты микросателлитных повторов по законам Менделя, и это делает их пригодными для применения в качестве маркеров при идентификации личности и определении кровного родства между индивидами. При решении такого рода вопросов достоверность выводов зависит от информативности исследованных полиморфных локусов, которую можно оце-

*Эл. почта: vadimst@img.tsu.ru

Таблица 1. Частоты встречаемости аллелей STR-локуса F13B у русских и белых американцев

	Русские (настоящая работа) $N = 200$	Европеоиды США [8] $N = 205$		
Номер аллеля	частота встречаемости	число наблюдений	частота встречаемости	число наблюдений
6	0.0775	31	0.100	41
7	0.015	6	0.020	8
8	0.25	100	0.259	106
9	0.21	84	0.215	88
10	0.445	178	0.402	165
11	0.000	0	0.002	1
12	0.0025	1	0.002	1
Σ	1.000	400	1.000	410

Примечание. N – число индивидов в выборке.

Таблица 2. Значения частот встречаемости аллелей STR-локуса LPL у русских и белых американцев

	Русские (настоящая работа) $N = 200$	Европеоиды США [8] $N = 200$		
Номер аллеля	частота встречаемости	число наблюдений	частота встречаемости	число наблюдений
8	0.000	0	0.002	1
9	0.057	23	0.047	19
10	0.473	189	0.412	168
11	0.242	97	0.287	117
12	0.210	84	0.203	83
13	0.018	7	0.049	20
Σ	1.000	400	1.000	408

Примечание. N – число индивидов в выборке.

Таблица 3. Значения частот встречаемости аллелей STR-локуса vWA у русских и белых американцев

	Русские (настоящая работа) $N = 200$	Европеоиды США [8] $N = 213$		
Номер аллеля	частота встречаемости	число наблюдений	частота встречаемости	число наблюдений
14	0.102	41	0.131	56
15	0.127	51	0.082	35
16	0.208	83	0.211	90
17	0.277	111	0.265	113
18	0.208	83	0.202	86
19	0.065	26	0.087	37
20	0.013	5	0.021	9
Σ	1.000	400	1.000	426

Примечание. N – число индивидов в выборке.

нить на основе популяционных данных о его генетическом разнообразии (частотах встречаемости его аллелей и дисперсии микросателлитных повторов). Точность оценки будет тем выше, чем ближе соответствие популяционных данных и данных по распределению аллелей в регионе, обслуживаемом судебно-медицинской лабораторией.

Опубликованы данные о частотах встречаемости аллелей STR, которые используются в судебно-медицинской экспертизе для популяций европеоидов, проживающих на территории США [8], Западной Европы [9] и, по отдельным локусам, на европейской части России [10, 11]. Сведений о распределении аллелей данных полиморфных локусов среди населения большей части России нет. Мы определили частоты встречаемости аллелей STR-локусов LPL, vWA, FES/FPS и F13B в выборке индивидов из русского населения, не являющихся родственниками и проживающих на территории Сибири. Кроме того, мы сравнили эти частоты с данными о частотах встречаемости этих же локусов у европеоидов США и оценили параметры пригодности этих локусов для судебно-медицинских исследований.

УСЛОВИЯ ЭКСПЕРИМЕНТА

Геномную ДНК выделяли из венозной крови человека при помощи набора реагентов *DIAAtom™DNAPrep* ("IsoGene", Москва). Раствор ДНК в воде хранили при -20°C .

Концентрацию ДНК определяли с помощью электрофореза в 1%-ном агарозном геле. В качестве маркера использовали ДНК фага лямбда в различной концентрации.

Выборка из 200 образцов крови неродственных представителей русского населения, проживающего на территории Урала, Сибири и северного Казахстана, получена в ходе экспертной практики Областного бюро судебно-медицинской экспертизы (Новосибирск).

Типирование STR-локусов LPL, vWA, F13B, FES/FPS в препаратах ДНК проводили в монолокусном формате с помощью ПЦР согласно методическим указаниям руководства "Использование индивидуализирующих систем на основе полиморфизма длин амплифицированных фрагментов (ПДАФ) ДНК в судебно-медицинской экспертизе идентификации личности и установления родства" (утверждено Минздравом РФ 19.01.1999 г.). ПЦР проводили на амплификаторе "Терцик" ("ДНК-Технология", Москва) по протоколу, используя наборы *GenePrint™ STR Systems* ("Promega", США) и термостабильную ДНК-полимеразу *Taq* производства фирмы "Сибэнзим" (Новосибирск).

Продукты амплификации фракционировали электрофоретически в нативном 8%-ном поли-

Таблица 4. Значения частот встречаемости аллелей STR-локуса FES/FPS у русских и белых американцев

Номер аллеля	Русские (настоящая работа) $N = 200$		Европеоиды США [8] $N = 213$	
	частота встречаемости	число наблюдений	частота встречаемости	число наблюдений
8	0.005	2	0.014	6
9	0.005	2	0.007	3
10	0.2475	99	0.284	121
11	0.4525	181	0.439	187
12	0.2325	93	0.225	96
13	0.055	22	0.028	12
14	0.0025	1	0.002	1
Σ	1.000	400	1.000	426

Примечание. N – число индивидов в выборке.

акриламидном геле (ПААГ), гель окрашивали бромистым этидием. Размеры амплифицированных фрагментов геномной ДНК определяли, сравнивая их с внешними стандартами молекулярных масс и локус-специфичными аллельными маркерами путем компьютерной интерполяции по программе LabWorks™ (“UVP”, США).

Ожидаемые и наблюдаемые уровни гетерозиготности рассчитывали стандартным способом. Отклонение наблюдаемых частот встречаемости генотипов от равновесного распределения Харди–Вайнберга рассчитывали с помощью точного метода Гуо и Томпсона [12]. Кроме показателей генетического разнообразия, основанных на частотах аллелей (уровни гетерозиготности), использовали более адекватные для микросателлитов показатели, основанные на анализе молекулярной дисперсии (дисперсия числа повторов, разброс повторов). Для определения генетических расстояний

Таблица 5. Значения частот встречаемости генотипов четырех STR-локусов

п/№	LPL		FES/FPS		F13B		vWA		
	генотипы	частота встречаемости							
1	9/9	1	0.005	8/11	2	0.01	6/6	4	0.02
2	9/10	14	0.07	9/10	1	0.005	6/8	5	0.025
3	9/11	2	0.01	9/12	1	0.005	6/9	7	0.035
4	9/12	5	0.025	10/10	16	0.08	6/10	10	0.05
5	10/10	55	0.275	10/11	36	0.18	6/12	1	0.005
6	10/11	34	0.17	10/12	24	0.12	7/7	1	0.005
7	10/12	28	0.14	10/13	6	0.03	7/9	1	0.005
8	10/13	3	0.015	11/11	45	0.225	7/10	3	0.015
9	11/11	16	0.08	11/12	47	0.235	8/8	11	0.055
10	11/12	26	0.13	11/13	6	0.03	8/9	22	0.11
11	11/13	3	0.015	12/12	7	0.035	8/10	51	0.255
12	12/12	12	0.06	12/13	6	0.03	9/9	7	0.035
13	12/13	1	0.005	12/14	1	0.005	9/10	40	0.2
				13/13	2	0.01	10/10	37	0.185
14									
15									
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									
Σ	200	1.0		200	1.0		200	1.0	

Таблица 6. Популяционно-генетические характеристики микросателлитных локусов F13B, LPL, VWA, FES/FPS у русских и белых американцев

Характеристики	Русские	Европеоиды США [8]	Всего
F13B			
Число хромосом ($2N$)	400	410	810
Число аллелей (K)	6	7	7
Наблюданое число генотипов (G_o)	14		
Наблюданая гетерозиготность (H_o)	0.700	0.751	0.725
Ожидаемая гетерозиготность (H_e)	0.689	0.715	0.703
Дисперсия числа повторов (V)	1.475	1.641	1.558
Разброс числа повторов (R)	6	6	6
Вероятность случайного совпадения генотипов (pM)	0.148		
Дискриминирующий потенциал (PD)	0.852		
LPL			
Число хромосом ($2N$)	400	408	808
Число аллелей (K)	5	6	6
Наблюданое число генотипов (G_o)	13		
Наблюданая гетерозиготность (H_o)	0.580	0.770	0.676
Ожидаемая гетерозиготность (H_e)	0.670	0.702	0.688
Дисперсия числа повторов (V)	0.867	0.978	0.922
Разброс числа повторов (R)	4	5	5
Вероятность случайного совпадения генотипов (pM)	0.162		
Дискриминирующий потенциал (PD)	0.838		
VWA			
Число хромосом ($2N$)	400	426	826
Число аллелей (K)	7	7	7
Наблюданое число генотипов (G_o)	26		
Наблюданая гетерозиготность (H_o)	0.795	0.821	0.828
Ожидаемая гетерозиготность (H_e)	0.806	0.812	0.810
Дисперсия числа повторов (V)	2.069	2.358	2.217
Разброс числа повторов (R)	6	6	6
Вероятность случайного совпадения генотипов (pM)	0.065		
Дискриминирующий потенциал (PD)	0.935		
FES/FPS			
Число хромосом ($2N$)	256	426	682
Число аллелей (K)	7	7	7
Наблюданое число генотипов (G_o)	14		
Наблюданая гетерозиготность (H_o)	0.650	0.699	0.680
Ожидаемая гетерозиготность (H_e)	0.677	0.675	0.677
Дисперсия числа повторов (V)	0.783	0.797	0.794
Разброс числа повторов (R)	6	6	6
Вероятность случайного совпадения генотипов (pM)	0.160		
Дискриминирующий потенциал (PD)	0.840		
По совокупности локусов			
Ожидаемая гетерозиготность (H_e)	0.710	0.726	0.719
Вероятность случайного совпадения генотипов (pM)	0.00025		
Дискриминирующий потенциал (PD)	0.99975		

Примечание. В графе “Всего” приведены значения показателей для суммарной выборки, объединяющей популяции русских и европеоидов США.

Таблица 7. Генетическая вариабельность STR-локусов F13B, LPL, vWA, FES/FPS в суммарной выборке

Локус	Гетерозиготность (H_e)	Дисперсия числа повторов (V)	Число аллелей (K)	Разброс повторов (R)
F13B	0.703	1.561	7	6
LPL	0.688	0.926	6	5
vWA	0.810	2.217	7	6
FES/FPS	0.676	0.794	7	6
Среднее на локус	0.719	1.374	6.750	5.750

между популяциями использовали показатель генетической дистанции Ddm [13], основанный на пошаговой мутационной модели, которая более соответствует природе накопления мутаций в микросателлитных локусах, чем меры генетического расстояния, основанные на анализе частот аллелей. Распределение частот в популяциях сравнивали, применяя критерий гетерогенности χ^2 с поправкой на редкие варианты по Животовскому [14].

Используемые в судебно-медицинской практике показатели вероятности случайного совпадения генотипов (рM) и дискриминирующий потенциал (PD) рассчитывали по стандартным алгоритмам [15–16]. При помощи показателя рM оценивали, какова вероятность случайного совпадения генотипов по конкретному маркеру или группе маркеров у двух неродственных индивидов, а обратный показатель PD – вероятность обнаружения (дискриминации) отличия двух неродственных индивидов по маркеру/группе маркеров.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В выборке ДНК неродственных индивидов, проживавших на момент проведения исследования на территории Сибири, мы нашли 7 аллелей микросателлита FES/FPS, 5 аллелей локуса LPL, 6 аллелей локуса F13B и 7 аллелей локуса vWA. Значения частот встречаемости аллелей и генотипов в исследованной нами выборке представлены в таблицах 1–5. Среди обнаруженных аллелей в локусах LPL, vWA, FES/FPS и F13B наиболее часто встречаются аллели с числом повторов, рав-

ным 10 (частота встречаемости 0.473), 17 (0.277), 11 (0.4525), 10 (0.445) соответственно. Среди индивидов, проживающих на территории Сибири, более распространены генотипы, которые содержат комбинации аллелей по локусам LPL – 10/10 (0.275), vWA – 16/17 (0.13) и 17/18 (0.135), FES/FPS – 11/12 (0.235) и локусу F13B – 8/10 (0.255). Распределение генотипов по всем четырем локусам соответствует равновесию Харди–Вайнберга ($P = 0.7130 \pm 0.0143$ для LPL; 0.9230 ± 0.0084 для FES/FPS; 0.9110 ± 0.0090 для F13B; и 1.000 для vWA).

Способ наследования аллелей в локусах LPL, vWA, FES/FPS и F13B следует законам Менделя. При анализе 50 семей мутантных аллелей не обнаружено. Сравнение популяционно-генетических характеристик микросателлитных локусов F13B, LPL, vWA, FES/FPS в исследованной нами выборке и у европеоидов США представлено в табл. 6. Показатели генетической вариабельности и информативности в суммарной выборке представлены в табл. 7. По степени генетической вариабельности из четырех изученных STR-локусов можно выделить два более вариабельных – F13B и vWA (ожидаемая гетерозиготность более 0.7 и дисперсия числа повторов более 1.5) и два менее вариабельных – LPL и FES/FPS (ожидаемая гетерозиготность в суммарной выборке менее 0.7 и дисперсией числа повторов менее 1). В целом, система из четырех микросателлитов характеризуется высоким уровнем генетического разнообразия как среди русского населения (средний уровень ожидаемой гетерозиготности 0.710 и средняя дисперсия числа повторов 1.299, табл. 8), так и у европеоидов США (0.726 и 1.344 соответственно). При этом показатели разнообразия в обеих группах очень близки.

Дискриминационные возможности системы из четырех STR-локусов также достаточно высоки и примерно одинаковы для русских восточной части России и для европеоидов США (табл. 5). Вероятность случайного совпадения генотипов в выборке русских по совокупности исследованных микросателлитных локусов составляет 0.00025, а дискриминирующий потенциал системы превышает 99.9% (0.99975).

Несмотря на сходство показателей разнообразия и дискриминационных возможностей в двух сравниваемых популяциях, структура разнообра-

Таблица 8. Генетическая вариабельность в популяциях русских и европеоидов США по совокупности 4 микросателлитных локусов (F13B, LPL, vWA, FES/FPS)

Популяция	Наблюдаемая гетерозиготность (H_o)	Ожидаемая гетерозиготность (H_e)	Дисперсия числа повторов (V)	Число аллелей (K)	Всего аллелей	Разброс числа повторов (R)
Русские	0.693	0.710	1.299	6.750	27	5.75
Европеоиды США	0.760	0.726	1.334	6.250	25	5.50

Таблица 9. Генетические расстояния Ddm и генетическая гетерогенность между русскими и европеоидами США

Локус	Ddm	χ^2 (d.f.)
F13B	0.015	0.048 (d.f. = 0.06), $p < 0.01$
LPL	0.017	0.192 (d.f. = 5.95), $p < 0.01$
VWA	0.004	2.211 (d.f. = 1.085), $p > 0.05$
FES/FPS	0.017	0.183 (d.f. = 0.106), $p < 0.05$
По совокупности локусов	0.013 ± 0.003	0.204 (d.f. = 0.119), $p < 0.05$

зия у русских и европеоидов США различна. Об этом свидетельствуют результаты попарного сравнения популяций по частотам аллелей, а также показатель генетических расстояний Ddm у русских и европеоидов США по системе из четырех STR-локусов (табл. 9). Распределение частот аллелей всех STR-маркеров в двух популяциях достоверно различается, за исключением аллеля vWA (табл. 9). Достоверно превышает пороговое значение и критерий гетерогенности двух популяций по сумме локусов.

Показатель генетического расстояния Ddm, основанный на дисперсии числа повторов, между русскими и европеоидами США по совокупности STR-маркеров составляет 0.013 ± 0.003 , что свидетельствует о существенных генетических различиях двух популяций по этой системе маркеров (табл. 9). При этом основной вклад в различия русских и белых американцев вносят менее вариабельные локусы LPL и FES/FPS, генетическое расстояние по которым достоверно отличается от 0. По более вариабельным аллелям F13B и VWA русские и белые американцы отличаются в меньшей степени.

Таким образом, результаты сравнительного анализа генетической вариабельности четырех STR-локусов, применяемых в практике судебной медицины, свидетельствуют о достоверных генетических различиях популяций русских и европеоидов США. Очевидна неприемлемость использования в судебно-медицинских экспертизах, проводимых в России, расчетов, основанных на частотах аллелей, полученных на других популяциях, в частности, на белых американцах. Адекватное использование STR-систем в судебной экспертизе возможно только при использовании референтных баз по частотам аллелей, которые составлены на основании данных по исследованию образцов из той же популяции, или, по крайней мере, той же этнической или территориальной группы населения, что и образцы, являющиеся предметом экспертных исследований. Именно в качестве таких референтных наши данные могут быть использованы в деятельности судебно-медицинских лабораторий Сибири.

Работа получила финансовую поддержку Российского фонда фундаментальных исследований (03-04-49021, В.С), гранта Президента Российской Федерации (МД-88.2003.04, В.С) и Фонда “Ведущие научные школы” (НШ-840.2003.4, В.П.).

СПИСОК ЛИТЕРАТУРЫ

- Gill P., Ivanov P.L., Kimpton C. et al. 1994. Identification of the remains of the Romanov family by DNA analysis. *Nature Genet.* **6**, 130–135.
- Deka R., Shriver M.D., Yu L.M et al. 1995b. Intra- and interpopulation diversity at short tandem repeat loci in diverse populations of the world. *Electrophoresis.* **16**, 1659–1664.
- Goldstein D.B., Ruiz Linares A., Cavalli-Sforza L.L., Feldman M.W. 1996b. Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl. Acad. Sci. USA.* **93**, 6723–6727.
- Destro-Bisol G., Maviglia R., Caglia A. et al. 1999. Estimating European admixture in African Americans by using microsatellites and a microsatellite haplotype (CD4/Alu). *Hum. Genet.* **104**, 149–157.
- Bosch E., Calafell F., Santos F.R. et al. 1999. Variation in short tandem repeats is deeply structured by genetic background in the human Y chromosome. *Am. J. Hum. Genet.* **65**, 1623–1638.
- Степанов В.А. 2002. Этногеномика населения Северной Евразии. Томск: Печатная мануфактура, 244 с.
- Jobling M.A., Gill P. 2004. Encoded evidence: DNA in forensic analysis. *Nature Rev. Genet.* **5**, 739–751
- Promega Corp. 2002. Genetic identity reference information. Population statistics. <http://www.promega.com>.
- Huckenbeck W., Scheil H.-G. 2003. The distribution of the human DNA-PCR polymorphisms. <http://www.uni-duesseldorf.de>.
- Корниенко И.В., Земская Е.Ю., Фролова С.А., Якушев В.В., Иванов П.Л. (2002) Исследование аллельного полиморфизма молекулярно-генетических индивидуализирующих систем на основе тетрапланкотидных tandemных повторов LPL, vWA и TH01 среди населения России. *Судебно-медицинская экспертиза.* **45**, 12–14.
- Пушкарев В.П., Рахманина Л.В., Новиков П.И., Иванов П.Л. 2004. Исследование с помощью капиллярного электрофореза аллельного разнообразия микросателлитных локусов D16S539, F13B, FESFPS, TH01 и TPOX у европеоидов Уральского региона России. *Судебно-медицинская экспертиза.* **47**, 23–28.
- Guo S., Thomson E. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics.* **48**, 361–372.
- Goldstein D.B., Ruiz Linares A., Cavalli-Sforza L.L., Feldman M.W. 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics.* **139**, 463–471.
- Животовский Л.А. Популяционная биometрия. М: Наука, 1991. 271 с.
- Krueger J., Fuhrmann W., Liechte K.H., Steffens S. 1968. Zur Verwendung des Polymorphismus der sauren Erythrozytenphosphatase bei der Vaterschaftsbegutachtung. *Dtsch. Z. Gerichtl. Med.* **12**, 127–146.
- Jones D.A. 1972. Blood samples: probability of discrimination. *J. Forensic Sci. Soc.* **12**, 355–359.