

УДК 575.1:577.1

**АНАЛИЗ АЛЛЕЛЬНОГО ПОЛИМОРФИЗМА ДВУХ  
ТЕТРАНУКЛЕОТИДНЫХ ТАНДЕМНЫХ ПОВТОРОВ В ДВУХ ГОРОДСКИХ  
ПОПУЛЯЦИЯХ РОССИИ**

© 1996 г.

*Д.А. ЧИСТЯКОВ, И.А. ЕФРЕМОВ,  
О.Н. ОДИНОКОВА<sup>1</sup>, В.В. НОСИКОВ*

*Государственный научный центр «ГосНИИ генетика», Москва, 113545  
<sup>1</sup>Научно-исследовательский институт медицинской генетики Томского научного центра  
Российской академии медицинских наук, Томск, 634050*

На основе метода полимеразной цепной реакции (ПЦР) проведен анализ аллельного полиморфизма двух тетрануклеотидных tandemных повторов HUMvWFII и D6S366 в выборках из городских популяций Москвы и Томска. Обнаружено семь аллелей в локусе HUMvWFII размером от 154 до 178 п.н. и восемь аллелей длиной от 142 до 170 п.н. в локусе D6S366. Показано отсутствие статистически достоверных различий в частотном распределении аллелей в обеих исследованных популяционных выборках для данных локусов. Распределение частот встречаемости генотипов подчинялось равновесию Харди-Вайнберга; на основе 14 случаев семейного анализа показано менделевское наследование аллелей данных tandemных повторов. Определены параметры информационного содержания полиморфизма исследованных локусов, проведен сравнительный анализ аллельных частот с аналогичными данными по ряду популяций. Предложено использование данных коротких tandemных повторов для идентификации личности и определения спорного отцовства.

*Ключевые слова:* полимеразная цепная реакция, аллельный полиморфизм, русская популяция, микросателлиты

Короткие tandemно повторяющиеся последовательности (STR – short tandem repeats, или микросателлиты) представляют собой один из наиболее распространенных типов повторяющихся последовательностей ДНК. К данному классу обычно относят tandemные повторы простых мотивов, содержащих от одного до десяти нуклеотидов в повторяющейся единице [1]. Для ряда микросателлитов, обычно имеющих размер от 100 до 300 п.н., показан высокий уровень аллельного полиморфизма [2]. К числу высокополиморфных тетрануклеотидных tandemных повторов относятся микросателлиты D6S366 и HUMvWFII.

Пик и соавт. [3] показали наличие в интроне 40 гена фактора фон Виллебранда (vWF – von Willebrand factor gene, хромосомная локализация 12p13.3-p13.2 [4, 5]) полиморфного участка, содержащего множественные копии tandemного повтора АТСТ. С помощью метода полимеразной цепной реакции (ПЦР) [6] в данном локусе выявляется три дистанционированных друг от друга полиморфных микросателлита. Поскольку мутации гена vWF в ряде случаев приводят к развитию болезни фон Виллебранда, изучение данных полиморфизмов может оказаться полезным в плане молекулярной диагностики этого заболевания.

Для микросателлита HUMvWFI, расположенного в 5'-концевой части интрона 40, в европейских популяциях [7–10] обнаружено 11 аллелей длиной от 95 до 131 п.н.

Существует семь аллелей длиной от 154 до 178 п.н. микросателлита HUMvWFII, находящегося в 3'-концевой части полиморфного участка (позиции 2215 ÷ 2380 п.н. в интроне 40) [8, 11]. При амплификации полиморфного участка длиной 744 ÷ 820 п.н., включающего HUMvWFI и HUMvWFII, возможен одновременный анализ этих полиморфных локусов после расщепления продуктов реакции рестриктазой *AluI* [12].

Для микросателлита HUMvWFIII (позиции 1640 ÷ 1794 п.н. в интроне 40) при анализе 100 неродственных британцев показано существование семи аллелей размером от 138 до 162 п.н. [13]. Этот полиморфный STR (обозначается как HUMvWFA31/A) в настоящее время использует Forensic Science Service (Бирмингем, Великобритания) в панели из семи STR для идентификации личности, и при анализе различных расовых групп для микросателлита HUMvWFIII показано существование 12 аллелей [14].

В различных популяционных выборках обнаружено десять аллелей размером от 138 до 174 п.н. микросателлита D6S366 (обозначение по Genome Database D6S366), расположенного на хромосоме 6q21-qter [15]. Данный микросателлит входит в число 13 высокополиморфных и несцепленных между собой STR, предложенных в качестве тест-системы для идентификации личности и определения спорного отцовства в США [15].

Цель данной работы – провести анализ аллельного полиморфизма микросателлитов HUMvWFII и D6S366 в различных выборках из городских популяций России; сравнить частотное распределение аллелей с аналогичными данными по другим популяционным группам; оценить пригодность этих полиморфизмов для идентификационных и других исследований по генетике человека.

#### УСЛОВИЯ ЭКСПЕРИМЕНТА

ДНК-полимераза *Taq*<sup>R</sup> была получена от НПК «Биотех» (Москва). Синтез олигонуклеотидных праймеров был выполнен в Институте биоорганической химии РАН (Москва) на синтезаторе ASM-102U («Биоссет», Новосибирск).

Выделение геномной ДНК из венозной крови человека осуществляли по методике, описанной ранее [16]. Из слюны и пятен крови ДНК выделяли с использованием хелатного полимера Chelex<sup>R</sup>-100 («Bio-Rad Laboratories») [17].

Выборка из неродственных представителей популяции г. Москвы была сформирована из образцов, полученных в травматологических пунктах, институте ревматологии, бюро судебно-медицинской экспертизы и среди сотрудников ГНЦ «ГосНИИгенетика». Выборка из неродственных представителей популяции г. Томска была сформирована из образцов, полученных в родильных домах и от здоровых лиц.

ПЦР проводили на амплификаторе РНС-2 («Technе», Великобритания) в 50 мкл реакционной смеси следующего состава: 67 мМ Трис-НСl, рН 8,8, 16,6 мМ сульфат аммония, 0,01%-ный Твин-20, 1,0 и 1,5 мМ хлорид магния в случае HUMvWFII и D6S366 соответственно, 0,2 мМ каждого dNTP, 2,5 ед. полимеразы *Taq*<sup>R</sup>, 0,01–0,1 мкг геномной ДНК или 20 мкл Chelex<sup>R</sup>-100 ДНК-экстракта. Для амплификации аллелей HUMvWFII и D6S366 использовали по 66 нг каждого из олигонуклеотидных праймеров, последовательность которых приведена в более ранних работах [11, 15] соответственно. 30–35 циклов ПЦР проводили по следующей программе: 94°C/1 мин, 55°C (HUMvWFII) или 65°C (D6S366)/1 мин, 72°C/1 мин, в том числе первая денатурация 5 мин, последний синтез цепи – 7 мин.

Точную идентификацию аллелей в образцах ДНК осуществляли с помощью аллельных «лестниц» для обоих локусов, синтезированных следующим образом: в реакционную смесь объемом 100 мкл, содержащую 10 пкмоль каждого из праймеров, вносили 1 мкл эквимольной смеси полного спектра аллелей после ее разведения в 10<sup>4</sup> раз.

Продукты амплификации анализировали с помощью электрофореза в 12%-ном полиакриламидном геле (ПААГ) (длина 16 см, толщина 0,7 мм, дополнительно

содержал 7%-ный глицерин), нанося в лунки геля 10 мкл реакционной смеси. По окончании электрофореза гель окрашивали серебром [18].

Наблюдаемые частоты встречаемости генотипов исследованных локусов проверяли на отклонение от равновесия Харди-Вайнберга по критериям  $\chi^2$  и G-статистики с помощью компьютерной программы RxC (Rows x Columns) на основе алгоритма, описанного ранее [19]. Данный алгоритм позволяет оценить статистическую значимость отклонений от ожидаемого частотного распределения в случае, когда количество наблюдений по значительному числу классов меньше пяти, и применение стандартного критерия  $\chi^2$  неправомерно. Программу RxC также использовали для сравнения распределения аллельных частот исследованных локусов в различных популяционных выборках.

При проведении расчетов данная программа генерировала 1 000 случайных двумерных массивов такой же размерности и с такой же предельной суммой, что и исходный массив сравниваемого числа наблюдений (частот). Рассчитывались значения критериев  $\chi^2$  и G-статистики для каждого из массивов (далее приведены для исходного массива). Программа оценивала вероятность получения массива (процентная фракция таких массивов) со значениями  $\chi^2$  и G-статистики не меньше исходных. Если значения вероятности низкие (т.е.  $\approx 0,05$ ), то строки и столбцы в массиве не являются независимыми, и данные не достаточно гомогенны, чтобы быть объединенными. Значения больше 0,05 указывают на то, что данные имеют вероятность наблюдения более 5% в случае независимости строк и столбцов, и данные могут быть легитимно объединены.

Ожидаемую гетерозиготность ( $H_{exp}$ ) и следующие параметры информативности полиморфизма исследованных микросателлитов – вероятность случайного совпадения генотипов двух неродственных индивидов (probability of random match, pM), среднюю вероятность исключения исследуемого образца по заданному генотипу (mean exclusion chance, W), информационное содержание полиморфизма (polymorphism information content, PIC) – рассчитывали с помощью компьютерной программы на основе известных алгоритмов, приведенных нами ранее [20].

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В результате генотипирования 151 и 75 неродственных представителей городских популяций Москвы и Томска соответственно, мы показали наличие семи аллелей микросателлита HUMvWFI1 размером от 154 до 178 п.н. (рис. 1, 2). Аллели 11 и 12 оказались наиболее часто встречаемыми в обеих популяционных выборках. Из 28 возможных генотипов обнаружены 22 (Москва) и 16 (Томск), причем гетерозиготы 11/12 были отмечены наиболее часто: 21% наблюдений для Москвы и 27% – для Томска (табл. 1).

Для D6S366 в результате генотипирования 102 (Москва) и 75 (Томск) неродственных человек мы выявили 8 аллелей размером от 142 до 170 п.н. (рис. 3, 4). Аллели 12 и 13 оказались наиболее распространенными в обеих популяционных выборках. Из 36 возможных генотипов выявили 20 (Томск) и 22 (Москва). С наибольшей частотой наблюдались гетерозиготы 12/13: в 21% случаев для Москвы и в 15% случаев для Томска (табл. 2).

Экспериментальное распределение частот встречаемости генотипов по исследованным локусам в обеих популяционных выборках достаточно значимо соответствовало равновесию Харди-Вайнберга (значения вероятности  $0,983 \div 1,000$ , табл. 3), что позволяет говорить об отсутствии внутренней гетерогенности в исследованных городских популяциях.

Наблюдаемое распределение частот встречаемости аллелей и генотипов локуса HUMvWFI1 является весьма сходным в обеих исследованных популяционных выборках (табл. 4). Дополнительно мы провели сравнительный анализ распределения частот встречаемости аллелей этого микросателлита с имеющимися данными по европейским популяционным выборкам. Среди 24 исследованных голландцев, как и среди русских, аллели 11 и 12 обладали наибольшими частотами встречае-

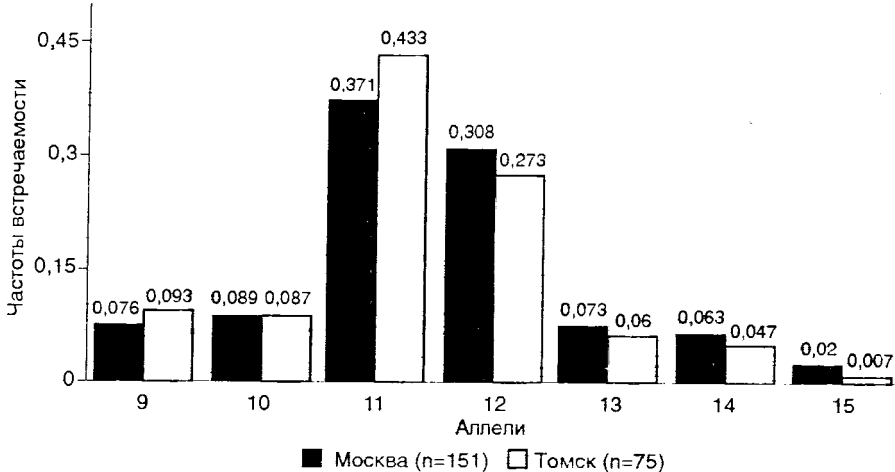


Рис. 1. Частоты встречаемости аллелей в локусе HUMvWFII в городских популяциях Москвы и Томска. Нумерация аллелей отражает число содержащихся в них тандемных повторов.  $n$  – число человек в исследованной выборке

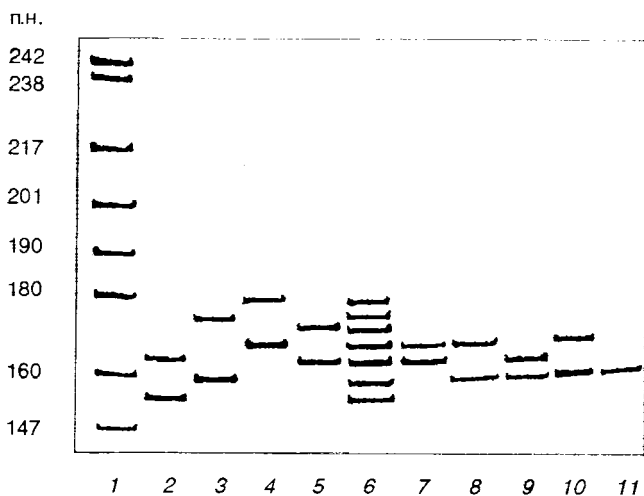


Рис. 2. Разделение в 12%-ном полиакриламидном геле амплифицированных аллелей в локусе HUMvWFII. 1 – фрагменты плазмидной ДНК pBR322, расщепленной рестриктазой *Msp*I; 2–5 – генотипирование четырех неродственных человек: 2 – 9/11, 3 – 10/14, 4 – 12/15, 5 – 11/13; 6 – аллельная «лестница»; 7–11 – семейный анализ, выявляющий следующие генотипы: 7 – 11/12 (отец), 8 – 10/12 (ребенок 1), 9 – 10/11 (ребенок 2), 10 – 10/12 (ребенок 3), 11 – 10/10 (мать)

мости (0,35 и 0,27 соответственно) [11]. Для 44 британцев (Britons) Кумминг и соавт. [8] показали унимодальное распределение частот аллелей: частота встречаемости аллеля 12 составила 0,42 против 0,18 для второго по распространенности аллеля 11.

Как можно видеть из  $R \times C$ -результатов, представленных в табл. 4, различия в распределении частот встречаемости аллелей HUMvWFII между московской, с

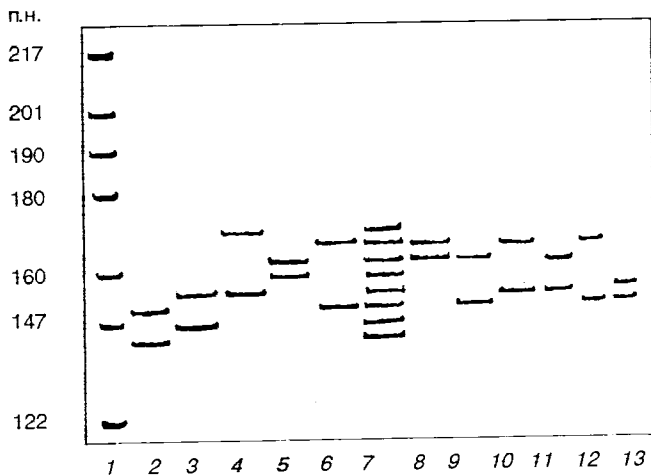


Рис. 3. Разделение в 12%-ном полиакриламидном геле амплифицированных аллелей в локусе D6S366. 1 – фрагменты плазмидной ДНК pBR322, расщепленной рестриктазой *MspI*; 2–6 – генотипирование пяти неродственных человек: 2 – 10/12, 3 – 11/13, 4 – 13/17, 5 – 14/15, 6 – 12/16; 7 – аллельная «лестница»; 8–13 – семейный анализ, выявляющий следующие генотипы: 8 – 15/16 (отец), 9 – 12/15 (ребенок 1), 10 – 13/16 (ребенок 2), 11 – 13/15 (ребенок 3), 12 – 12/16 (ребенок 4), 13 – 12/13 (мать)

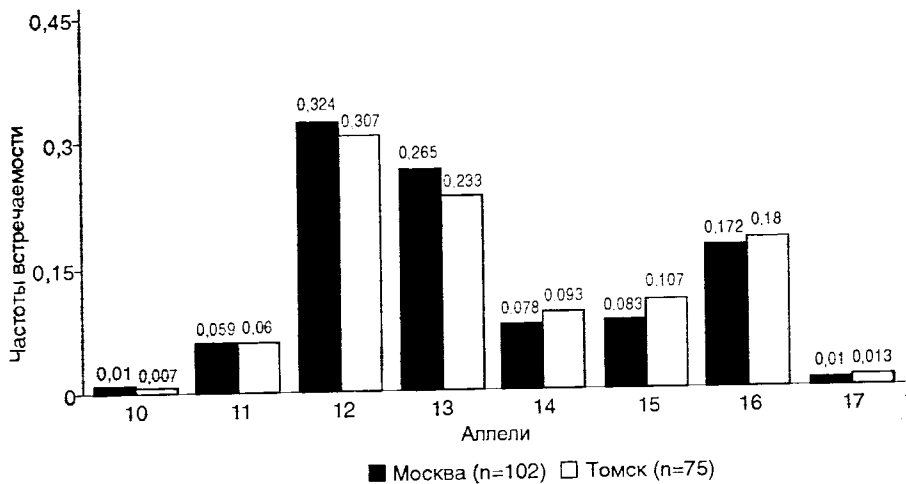


Рис. 4. Частоты встречаемости аллелей в локусе D6S366 в городских популяциях Москвы и Томска. Нумерация аллелей отражает число содержащихся в них тандемных повторов.  $n$  – число человек в исследованной выборке

одной стороны, и томской и голландской популяционными выборками, с другой стороны, не являются статистически значимыми. В то же время, численные значения вероятностей для  $\chi^2$  и G-статистики позволяют предположить достоверные различия в частотном аллельном распределении между русскими и британцами по локусу HUMvWFII.

Мы провели сравнительный анализ распределения частот встречаемости аллелей микросателлита D6S366 в обеих русских популяционных выборках с аналогичными данными по разным расам (рис. 5) [15]. Распределение аллелей среди жителей Москвы и Томска оказалось очень близким и весьма сходным с таковым для европеоидов США (170 человек); для последних, однако, наблюдалось не столь выраженное доминирование аллелей 12 и 13 (частоты встречаемости 0,265 и 0,200 соответственно) и наличие дополнительного очень редкого аллеля 18 длиной 174 п.н. (частота встречаемости 0,003). У латиноамериканцев США (164 человека) выявлены аллели 9 и 18, не обнаруженные в русской популяции, и показано преобладание аллеля 13. У негров США (178 человек) отмечены аллели 9+17 при доминировании аллелей 13 и 14. Среди выходцев из Азии (68 человек) наиболее распространены, как и у европеоидов, оказались аллели 12 и 13, однако было обнаружено лишь 6 аллелей (9, 11-15).

Как можно видеть из результатов, представленных на рис. 5, не обнаружено статистически достоверных различий в распределении частот встречаемости аллелей в локусе D6S366 только между русскими и европеоидами США. Для других

Таблица 1

Частоты встречаемости генотипов HUMvWFII в городских популяциях Москвы и Томска

Генотип	Частота встречаемости			
	Москва		Томск	
	Наблюдаемая	Ожидаемая	Наблюдаемая	Ожидаемая
9-9	0,007	0,006	0,013	0,009
9-10	0,013	0,014	0,013	0,016
9-11	0,053	0,056	0,120	0,081
9-12	0,040	0,047	0,013	0,051
9-13	0,013	0,011	0,013	0,011
9-14	0,013	0,010	0	0,009
9-15	0,007	0,003	0	0,001
10-10	0,026	0,008	0	0,008
10-11	0,033	0,066	0,093	0,075
10-12	0,046	0,055	0,053	0,047
10-13	0,020	0,013	0,013	0,010
10-14	0,013	0,011	0	0,008
10-15	0	0,004	0	0,001
11-11	0,159	0,138	0,147	0,188
11-12	0,212	0,228	0,267	0,237
11-13	0,079	0,054	0,040	0,052
11-14	0,033	0,047	0,053	0,040
11-15	0,013	0,015	0	0,006
12-12	0,106	0,095	0,053	0,075
12-13	0,020	0,045	0,053	0,033
12-14	0,066	0,039	0,040	0,026
12-15	0,020	0,012	0,013	0,004
13-13	0,007	0,005	0	0,004
13-14	0	0,009	0	0,006
13-15	0	0,003	0	0,001
14-14	0	0,004	0	0,002
14-15	0	0,002	0	0,001
15-15	0	0	0	0

расовых выборок результаты R×C-анализа говорят о значимых отличиях от русских.

Мы показали менделевский характер наследования аллелей в локусах HUMvWFI и D6S366 в результате анализа 14 семей, включающих от 1 до 4 детей (рис. 2, 3), мутантных аллелей не было обнаружено.

Результаты проведенного сравнительного анализа аллельных частот позволяют предположить, что для достаточно удаленных и изолированных друг от друга европеоидных популяций существуют реальные различия, количественно сопоста-

Таблица 2

Частоты встречаемости генотипов D6S366 в городских популяциях Москвы и Томска

Генотип	Частота встречаемости			
	Москва		Томск	
	Наблюдаемая	Ожидаемая	Наблюдаемая	Ожидаемая
10-10	0	0	0	0
10-11	0,010	0,001	0	0,001
10-12	0,010	0,006	0	0,004
10-13	0	0,005	0	0,003
10-14	0	0,002	0,013	0,001
10-15	0	0,002	0	0,001
10-16	0	0,003	0	0,002
10-17	0	0	0	0
11-11	0,020	0,003	0	0,004
11-12	0,049	0,038	0,053	0,037
11-13	0,020	0,031	0,040	0,028
11-14	0	0,009	0	0,011
11-15	0	0,010	0	0,013
11-16	0	0,020	0,027	0,022
11-17	0	0,001	0	0,002
12-12	0,078	0,105	0,107	0,094
12-13	0,206	0,171	0,147	0,143
12-14	0,020	0,051	0,053	0,057
12-15	0,049	0,054	0,027	0,065
12-16	0,147	0,111	0,107	0,110
12-17	0,010	0,006	0,013	0,008
13-13	0,039	0,070	0,053	0,054
13-14	0,088	0,042	0,040	0,044
13-15	0,049	0,044	0,067	0,050
13-16	0,078	0,091	0,067	0,084
13-17	0,010	0,005	0	0,006
14-14	0,010	0,006	0	0,009
14-15	0,010	0,013	0,027	0,020
14-16	0,020	0,027	0,053	0,034
14-17	0	0,002	0	0,002
15-15	0,010	0,007	0,013	0,011
15-16	0,039	0,029	0,053	0,038
15-17	0	0,002	0,013	0,003
16-16	0,029	0,029	0,027	0,032
16-17	0	0,003	0	0,005
17-17	0	0	0	0

вимые с межрасовыми: так, численные значения  $R \times C$  вероятностей практически одинаковы для сравниваемых пар русские-британцы (HUMvWFII) и русские-латиноамериканцы Мексики (D6S366). Данное соображение подкрепляется наличием литературных данных о реальных межпопуляционных различиях аллельных частот микро- и мини-сателлитов [21, 22].

Тем не менее, для исследованных локусов показано практически неразличимое распределение частот встречаемости аллелей в выборках из двух удаленных друг от друга городских популяций России. При вероятностных расчетах при идентификации личности можно достаточно уверенно экстраполировать аллельные частоты, определенные для одной репрезентативной выборки, на все русское население без риска существенно завязать потенциал индивидуализации.

Сравнение параметров информативности локусов HUMvWFII и D6S366 (табл. 3) позволяет отметить последний как существенно более информативный, с гетерозиготностью порядка 80% для русской популяции. В целом, значения параметров информативности изученных локусов в популяции г. Москвы сопоставимы с такими же для шести мини-сателлитных локусов, изученных нами ранее [20, 23], что предопределяет информативное использование данных микросателлитов для идентификации личности в русской популяции.

Таблица 3

Значения критериев полиморфизма и тестов  $\chi^2$  и G-статистики на равновесие Харди-Вайнберга для HUMvWFII и D6S366 в городских популяциях Москвы и Томска

Параметры	HUMvWFII		D6S366	
	Москва	Томск	Москва	Томск
$P_{\text{обс}}$	0,695	0,787	0,814	0,800
$H_{\text{экр}}$	0,747	0,720	0,783	0,801
$\rho_M$	0,103	0,121	0,081	0,071
W	0,678	0,632	0,762	0,794
PIC	0,587	0,543	0,626	0,659
$\chi^2$	12,5265	8,4883	14,3408	7,4906
Вероятность ( $\pm$ S.E.)*	0,9900 (0,0031)	0,9930 (0,0026)	0,9830 (0,0041)	1,0000 (0)
G-статистика	14,3040	10,1836	17,5434	9,4731
Вероятность ( $\pm$ S.E.)*	0,9920 (0,0028)	0,9960 (0,0020)	0,9830 (0,0041)	1,0000 (0)

\*Стандартная ошибка

Таблица 4

Сравнение распределений частот встречаемости аллелей в локусе HUMvWFII в московской популяции с другими европеоидными популяциями

Популяция	Русские (Томск)	Голландцы [9]	Британцы [18]
Объем выборки, чел.	75	24	44
$R \times C$ массив	7x2	7x2	7x2
$\chi^2$	3,6024	8,2531	15,6530
Вероятность $\pm$ S.E.*	0,7230 $\pm$ 0,0142	0,1970 $\pm$ 0,0126	0,0130 $\pm$ 0,0036
G-статистика	3,7821	8,1846	16,1609
Вероятность $\pm$ S.E.*	0,7030 $\pm$ 0,0144	0,2460 $\pm$ 0,0136	0,0180 $\pm$ 0,0042

\*Стандартная ошибка





4. Mancuso D.J., Tuley E.A., Westfield L.A., Worall N.K., Shelton-Inloes B.B., Sorace J.M., Alevy Y.G., Sadler J.E. // *J. Biol. Chem.* 1989. V. 264. P. 19514–19527.
5. Ginsburg D., Handin R.I., Bonthron D., Donlon T.A., Bruns G.A., Latt S.A., Orkin S.H. // *Science*. 1985. V. 228. P. 1401–1406.
6. Saiki R.K., Gelfand D.H., Stoffel S., Scharf S.J., Higuchi R., Horn G.T., Mullis K.B., Erlich H.A. // *Science*. 1988. V. 239. P. 487–491.
7. Peake I.R., Bowen D., Bignell P., Liddell M.B., Sadler J.E., Standen G., Bloom A.L. // *Blood*. 1990. V. 76. P. 555–561.
8. Cumming A.M., Armstrong J.G., Pendry K., Burn A.M., Wensley R.T. // *Hum. Genet.* 1992. V. 89. P. 194–198.
9. Мисюрин А.В., Сурин В.Л., Соловьев Г.Я. // *Генетика*. 1994. Т. 30. С. 713–717.
10. Асеев М.В., Скакуи В.Н., Баранов В.С. // *Генетика*. 1995. Т. 31. С. 839–845.
11. Ploos van Amstel H.K., Reitsma P.H. // *Nucleic Acids Res.* 1990. V. 18. P. 4957.
12. Mercier B., Gaucher C., Mazurier C. // *Nucleic Acids Res.* 1991. V. 19. P. 4800.
13. Kimpton C., Walton A., Gill P. // *Hum. Mol. Genet.* 1992. V. 1. P. 287.
14. Urquhart A., Oldroyd N.J., Kimpton C.P., Gill P. // *BioTechniques*. 1995. V. 18. P. 116–121.
15. Hammond H.A., Jin Li, Zhong Y., Caskey C.T., Chakraborty R. // *Am. J. Hum. Genet.* 1994. V. 55. P. 175–189.
16. Johns M.B., Paulus-Thomas J.E. // *Anal. Biochem.* 1989. V. 180. P. 276–278.
17. Walsh P., Metzger D., Higuchi R. // *BioTechniques*. 1991. V. 10. P. 506–513.
18. Budowle B., Chakraborty R., Giusti A.M., Eisenberg A.J., Allen R.C. // *Am. J. Hum. Genet.* 1991. V. 48. P. 137–144.
19. Roff D.A., Bentzen P. // *Mol. Biol. Evol.* 1989. V. 6. P. 539–545.
20. Ефремов И.А., Чистяков Д.А., Носиков В.В. // *Молекуляр. биология*. 1996. Т. 30. С. 307–318.
21. Deka R., Chakraborty R., Ferrell R. // *Genomics*. 1991. V. 11. P. 83–92.
22. Wall W.J., Williamson R., Petrou M., Papaioannou D., Parkin B.H. // *Hum. Mol. Genet.* 1993. V. 2. P. 1023–1029.
23. Чистяков Д.А., Газрилов Д.К., Овчинников И.В., Носиков В.В. // *Молекуляр. биология*. 1993. Т. 27. С. 1304–1314.