

УДК 575.174:599.9

СТРУКТУРА ЯДЕРНОГО ГЕНОФОНДА ПОВОЛЖСКИХ ТАТАР (ПО ДАННЫМ АУТОСОМНЫХ МИКРОСАТЕЛЛИТНЫХ ЛОКУСОВ)

О.А. Кравцова

Аннотация

По 11 аутосомным микросателлитным локусам (STR) охарактеризован генофонд современной популяции татар, представленной двумя субэтническими группами казанских татар (Азнакаевский район, $N = 135$) и татар-мишар (Буинский район, $N = 141$). Анализ распределения частот аллелей и генотипов показал значимое отклонение от равновесия Харди-Вайнберга по 6 локусам в популяциях Азнакаевского и Буинского районов, обусловленное недостатком гетерозиготных индивидуумов. При этом не было обнаружено статистически значимых различий в распределении аллелей по 8 локусам, тогда как по 3 STR (FGA, TPOX и D21S11) наблюдаются значимые различия между исследованными группами татар. По данным о распределении частот аллелей 3 локусов, находящихся в равновесии (D3S1358, D5S818, vWA), рассчитаны генетические расстояния по Нею между популяцией татар и некоторыми популяциями мира, тем самым определено их местоположение в системе мирового генофонда.

*Посвящается памяти
А.Н. Аскаровой*

Введение

Вопросы происхождения любого этноса в силу многогранности процесса являются в науке достаточно трудной проблемой. К числу наиболее перспективных подходов для изучения генетической истории народов относится анализ изменчивости высокополиморфных генетических систем [1]. Огромное количество полиморфных маркеров, выявленное при расшифровке генома человека, является мощным инструментом для анализа генофонда, его основных характеристик, динамики, истории и географии. Многочисленные исследования полиморфных систем ядерного и митохондриального геномов привели к развитию нового раздела геномики – этногеномики [2]. На сегодняшний день накоплен большой объем данных по полиморфизму аутосомных микросателлитных локусов, микросателлитов Y-хромосомы и вариабельности митохондриального генома в различных популяциях мира.

Несмотря на это, регион Среднего Поволжья, входящий в состав Волго-Уральского региона, до сих пор является малоизученной территорией. Популяционно-генетические исследования народов Среднего Поволжья в основном носят фрагментарный характер как в отношении изученных популяций, так и в

отношении генетических маркеров. Наиболее изученными в этом плане являются популяции башкир и чувашей [2, 3], тогда как данные о генофонде татар, представляющих основное население Среднего Поволжья, являются неполными и разрозненными.

Исследование генофонда поволжских татар направлено, прежде всего, на выяснение определенного генетического своеобразия, которое формировалось в результате длительного контакта двух рас – европеоидной и монголоидной [4]. Ранее нами был охарактеризован полиморфизм мтДНК в современных популяциях татар, который позволяет проследить формирование популяций по материнской линии [5], однако представляет интерес выявить вклад обоих полов в генетическое разнообразие и формирование современного генофонда татар на основе аутомных микросателлитных локусов (STR).

В связи с вышесказанным, целью данной работы явилась сравнительная характеристика 11 STR-локусов в современных популяциях поволжских татар, представленных двумя субэтническими группами казанских татар и татар-мишерей.

Материалы и методы

Материалом для исследования генофонда современного населения Республики Татарстан (РТ) служили образцы ДНК неродственных индивидов, представляющих две этнографические группы татар: казанские татары и татары-мишары. Выборка казанских татар набрана в Азнакаевском районе РТ ($N = 135$), татары-мишары представлены дисперсной выборкой из Буинского района РТ ($N = 141$). Материал собран в ходе экспедиционных выездов в 2004–2005 гг. Этническую принадлежность выясняли путем индивидуального анкетирования, учитывая данные до третьего поколения.

Тотальный препарат ДНК из 4 мл периферической крови современного населения выделяли стандартным методом фенол/хлороформной экстракции [6].

Генотипирование проводили по 11 аутомным STR-локусов – D3S1358, D5S818, D7S820, D16S539, D21S11, HumTH01, vWA31A, FGA (FIBRA), TPOX, LPL (LIPOL) и CD4 – с использованием праймеров, предложенных ранее [7, 8]. Олигонуклеотиды синтезированы в НПФ «Литех» (г. Москва). ПЦР проводили в 10 мкл реакционной смеси, содержащей около 15–20 нг тотальной ДНК, 200 мМ каждого dNTP и 1 ед термостабильной полимеразы Taq-SE производства ООО «СибЭнзим» (г. Новосибирск).

Разделение продуктов амплификации по аутомным STR-локусам проводили с помощью электрофореза в 6–8% полиакриламидном геле в нативных и денатурирующих условиях (длина геля не менее 10 см).

Частоты аллелей и генотипов рассчитывали путем прямого подсчета с использованием макроса VBA к Microsoft Excel, разработанного Д.С. Тарасовым (кафедра генетики, Казанский государственный университет). Распределение частот аллелей и генотипов исследованных STR-локусов проверяли на отклонение от равновесия Харди – Вайнберга с использованием двустороннего теста методом Guo и Thompson [9] и одностороннего теста на наличие избытка и дефицита гетерозигот [10], ожидаемые и наблюдаемые значения гетерозиготности вычисляли общепринятыми методами, реализованными в программе

Генератор v.3.4 [11]. Показатели генного и генотипического разнообразия рассчитывали по формуле Нея и Таджима [12]. Для определения внутривидовой гетерогенности были рассчитаны расстояния между индивидуумами с помощью программы POPULATIONS (v.1.2.28). Консенсусные дендрогаммы строили с помощью пакета программ Phylip (v.3.61) и Treeview (v.1.6.6). Расчет генетических расстояний между популяциями проводили, используя стандартное генетическое расстояние D_S по Нею [13]. На основе полученных данных по генетическим расстояниям был проведен кластерный анализ методом объединения ближайших соседей (neighbour joining) с помощью программы POPTREE (модифицированной версии NJBAFD v.1.1), статистическая значимость полученных кластеров определялась методом bootstrap-анализа.

Результаты и обсуждение

Анализ 11 аутосомных микросателлитных локусов показал, что распределение частот аллелей в исследованных популяциях не отличается от таковых в мировых популяциях и носит уни- или бимодальный характер распределения (рис. 1). Распределение аллелей в изученных группах представлено на рис. 2.

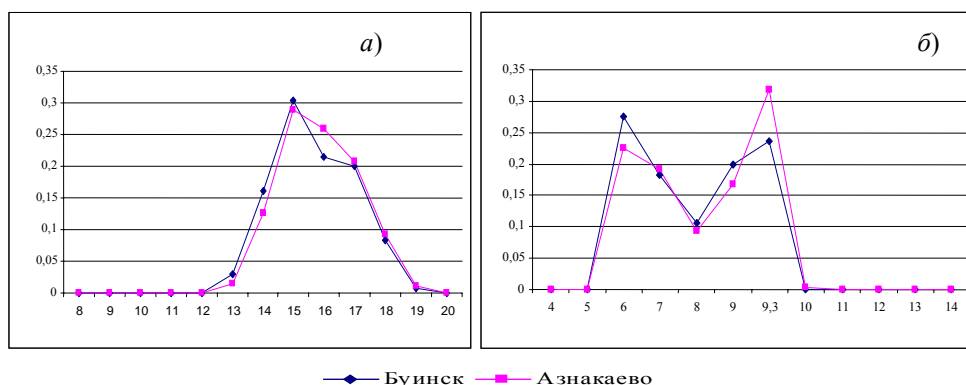


Рис. 1. Унимодальный (а) и бимодальный (б) характер распределения аллелей

Различие в распределении аллелей и генотипов между двумя исследованными популяциями татар по микросателлитам было обнаружено только по 3 локусам: FGA, TPOX и D21S11 (табл. 1). Эти различия могут объясняться эффектом выборочности, так как эти локусы являются наиболее полиморфными системами. В целом, в исследованных популяциях обнаружен высокий «запас» генетического разнообразия по микросателлитным локусам.

Основной характеристикой состояния популяции является подчинение распределения частот аллелей и генотипов равновесию Харди – Вайнберга (РХВ), которое показывает стабильность генетических признаков в исследуемой популяции. При интенсивных процессах метисации или, наоборот, в результате инбридинга, будет наблюдаться сдвиг равновесия, т. е. данная характеристика позволяет проследить процессы, происходящие в каждой конкретной популяции.

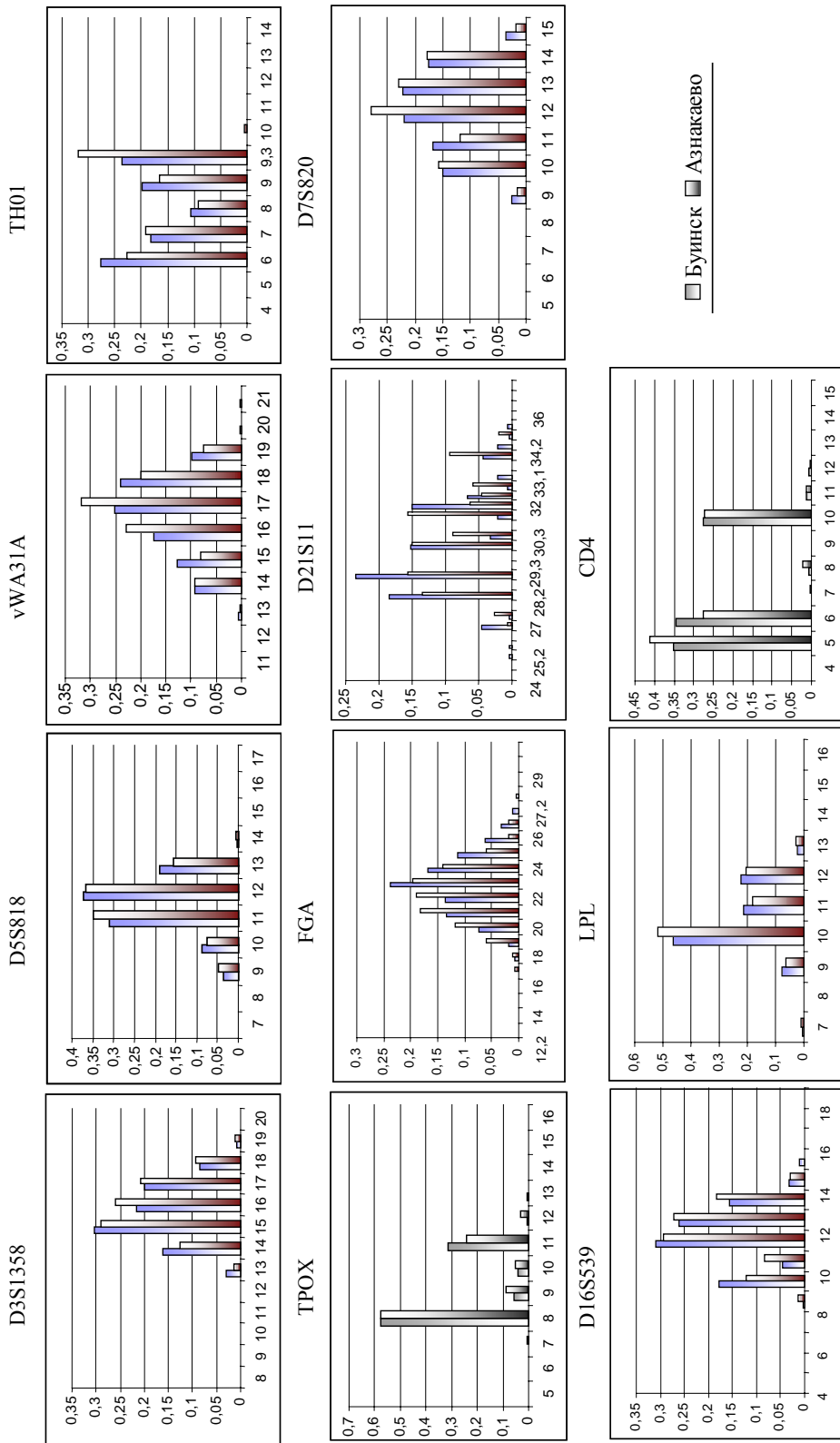


Рис. 2. Распределение частот аллелей STR-локусов в популяциях татар. По оси X – номера аллелей, по оси Y – частота встречаемости

Табл. 1

Показатели аллельной и генотипической дифференциации между популяциями татар Азнакаевского и Буинского р-нов

Локус	Распределение аллелей		Распределение генотипов	
	P-value	S.E.	P-value	S.E.
D3S1358	0.6886	0.0123	0.7125	0.0102
FGA	0.0013*	0.0006	0.0006	0.0003
D5S818	0.8294	0.0078	0.8627	0.0055
TH01	0.2099	0.0137	0.2259	0.0100
TPOX	0.0464	0.0065	0.0497	0.0055
vWA31A	0.1151	0.0114	0.1733	0.0093
D7S820	0.4127	0.0194	0.4762	0.0136
D21S11	0.0000	0.0000	0.0000	0.0000
D16S539	0.0774	0.0073	0.1004	0.0074
LPL	0.7469	0.0097	0.8268	0.0074
CD4	0.3654	0.0148	0.4526	0.0109

* Выделены значимые различия в характере распределения аллелей и генотипов между популяциями (уровень значимости $p < 0.05$).

Табл. 2

Двусторонний точный тест на соответствие равновесию Харди – Вайнберга

Локус	Распределение аллелей		Распределение генотипов	
	P-value	S.E.	P-value	S.E.
D3S1358	0.3997	0.0151	0.0344	0.0060
FGA	0.0170	0.0044	0.0155	0.0044
D5S818	0.2624	0.0154	0.0567	0.0070
TH01	0.0221	0.0053	0.0704	0.0048
TPOX	0.9898	0.0020	0.7195	0.0149
vWA31A	0.9334	0.0060	0.8650	0.0129
D7S820	0.0004	0.0003	0.0461	0.0071
D21S11	0.0000	0.0000	0.0000	0.0000
D16S539	0.0163	0.0030	0.0029	0.0012
LPL	0.0279	0.0040	0.5210	0.0146
CD4	0.4596	0.0248	0.0012	0.0006

Значения P-value приведены для уровня значимости $p < 0.05$. Выделены значения для локусов, отклоняющихся от равновесия.

Нами показано, что по некоторым локусам обе популяции отклоняются от PХВ, причем отклонение обусловлено недостатком гетерозиготных индивидуумов (табл. 2). Отклонение от равновесия наблюдается в обеих популяциях по локусам FGA, D7S820, D21S11, D16S539.

Для населения Азнакаевского р-на неравновесными локусами также являются TH01 и LPL, для жителей Буинского р-на – локусы D3S1358 и CD4. При рассмотрении объединенной выборки татар отклонение от равновесия наблюдается по тем же самым локусам, кроме локуса D3S1358.

Отклонение от PХВ за счет дефицита гетерозигот может наблюдаться в нескольких случаях. Во-первых, исследованные локусы подвержены селективно-

Табл. 3

Хромосомная локализация аутосомных STR-локусов

Локус	Хромосомная локализация
FGA	4q28, 3-й интрон гена α -фибриногена
D5S818	5q23.3-32
TH01	11p15.5, 1-й интрон гена тирозингидроксилазы
TPOX	2p23-pter, 10-й интрон гена пероксидазы
vWA31A	12p-12-pter, 40-й интрон гена фактора фон Виллебрандта
D7S820	7q
D21S11	21q
D16S539	16q22-24
D3S1358	3p21.3
CD4	12p
LPL	8p22, 6-й интрон гена липопротеинлипазы

му отбору; во-вторых, в популяциях происходят процессы инбридинга; в-третьих, имеет место так называемый эффект Валенда, выраженный в генетической подразделенности исследованной популяции на несколько более мелких субпопуляций; в-четвертых, не последнюю роль может играть дрейф генов [14].

По данным ряда авторов, микросателлитные STR-локусы являются селективно нейтральными [15]. Для того чтобы подвергаться отбору, они должны находиться в непосредственной близости от функционального гена. В нашем случае такими локусами являются vWA31A, FGA, TH01, TPOX и LPL (табл. 3).

Как видно из таблицы, 3 из 5 локусов, находящихся внутри интронов функционирующих генов, отклоняются от равновесия, что, по-видимому, может быть связано с селективным отбором. К сожалению, в мировых популяциях нет данных по отклонению этих локусов от равновесия Харди – Вайнберга, поэтому нам трудно сделать какие-либо выводы.

Что касается возможного эффекта инбридинга, то для изучаемых нами популяций не существует прямых доказательств высокого уровня инбриденности.

Таким образом, дефицит гетерозигот в данном случае можно объяснить случайным дрейфом генов и генетической неоднородностью (делением на субпопуляции, или эффектом Валенда) исследуемых популяций. Для подтверждения эффекта Валенда нами были определены генетические расстояния между индивидуумами внутри каждой из выборок, по которым далее был проведен кластерный анализ методом объединения ближайших соседей (рис. 3). Полученные дендрограммы свидетельствуют о внутренней гетерогенности исследуемых популяций татар и делении на несколько субпопуляций.

Описанное выше высокое внутривнутрипопуляционное генетическое разнообразие по исследованным аутосомным микросателлитным локусам отражается в высоком дискриминационном потенциале исследованных локусов. Вероятность случайного совпадения генотипов РМ у двух случайно выбранных неродственных индивидов по отдельным локусам варьирует от 0.0346 по локусу D21S11 до 0.2080 по локусу TPOX в популяции Азнакаевского р-на и от 0.460 (D21S11) до 0.2632 (TPOX) в популяции Буинского р-на. Соответственно, ве-

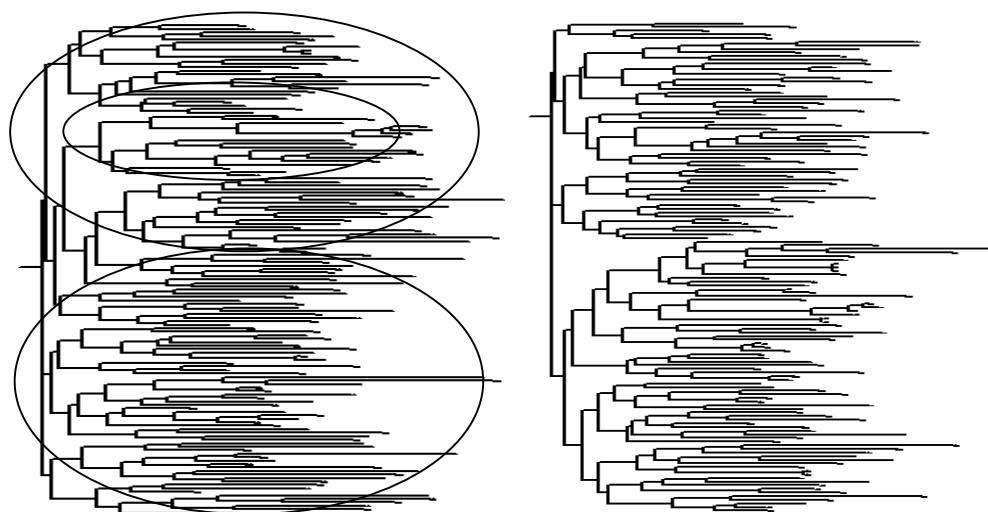


Рис. 3. Консенсусные дендрограммы кластеризации индивидуумов методом объединения ближайших соседей

роятность случайного нахождения двух одинаковых мультилокусных генотипов у неродственных индивидов составляет $5.9 \cdot 10^{-12}$ и $6.3 \cdot 10^{-12}$ для жителей Азнакаевского и Буинского районов соответственно. Отсутствие различий по основным показателям (PM, PD, PE) дает полное право для объединения этих двух выборок при расчете вероятностей в судебно-медицинской экспертизе (табл. 4).

Для оценки межпопуляционных разнообразий были рассчитаны стандартные генетические расстояния по Нею между популяцией татар и некоторыми мировыми популяциями, основанные на частотах аллелей исследованных локусов. Расчет проводился по трем микросателлитам D3S1358, TH01 и vWA. Основными критериями при выборе маркеров являлись: соответствие равновесию Харди – Вайнберга, отсутствие аллельной и генотипической дифференциации между популяциями казанских татар и татар-мишарей, наличие значений о частотах аллелей в мировых популяциях. На основании полученных значений была построена консенсусная дендрограмма, отражающая филогенетические отношения современных популяций татар с некоторыми популяциями мира (рис. 4).

Как видно из рис. 4, по трем локусам наблюдается довольно четкая кластеризация европеоидных и монголоидных популяций. В то же время нам неизвестно влияние ограниченного объема выборки на расчет генетических расстояний, и значения бутстрепа указывают на то, что данные кластеры образуются в силу случайных причин. В любом случае включение дополнительных локусов в построение филогенетических древ позволит повысить разрешающую способность метрики расстояний.

В то же время необходимо отметить существование общего кластера, включающего популяции татар и евреев, которое можно объяснить процессами метисации, происходившими в эпоху становления татарского народа. По историческим данным, болгары, принявшие непосредственное участие в формирова-

Табл. 4

Популяционные характеристики STR-локусов

Локус	Азнакаевский р-н					
	P_{EXACT}	H_{OBS}	H_{EXP}	PM	PD	PE
TH01	0.0704	83	96.78	0.0802	0.9198	0.3904
TPOX	0.7195	81	79.72	0.2632	0.7368	0.2614
LPL	0.521	95	97.04	0.1451	0.8549	0.3888
CD4	0.0012	96	96.45	0.1978	0.8022	0.3883
D7S820	0.0461	111	114.24	0.0675	0.9324	0.5963
0,D16S539	0.0029	96	108.97	0.0877	0.9123	0.1065
D3S1358	0.0344	106	108.33	0.0901	0.9099	0.5512
D5S818	0.0567	98	99.08	0.1397	0.8603	0.4524
vWA31A	0.865	115	115.11	0.0656	0.9344	0.6283
FGA	0.0155	110	110.16	0.0469	0.9561	0.7134
D21S11	0.0000	112	121.01	0.0460	0.9540	0.5886
Локус	Буинский р-н					
	P_{EXACT}	H_{OBS}	H_{EXP}	PM	PD	PE
TH01	0.0221	115	105.11	0.1061	0.8939	0.6985
TPOX	0.9898	83	80.58	0.2080	0.7920	0.3093
LPL	0.0279	74	87.06	0.1727	0.8273	0.2374
CD4	0.4596	89	92.87	0.1606	0.8394	0.3681
D7S820	0.0004	94	107.43	0.0756	0.9243	0.4306
0,D16S539	0.0163	97	105.23	0.0879	0.9167	0.4587
D3S1358	0.3997	94	105.89	0.0783	0.9217	0.4226
D5S818	0.2624	93	96.59	0.1322	0.8678	0.4113
vWA31A	0.9334	112	106.38	0.0825	0.9175	0.6551
FGA	0.0170	115	115.32	0.0491	0.9509	0.6985
D21S11	0.0000	94	118.78	0.0346	0.9654	0.4306

P_{EXACT} – точный тест на соответствие равновесию Харди –Вайнберга;

H_{OBS} – наблюдаемые показатели гетерозиготности;

H_{EXP} – ожидаемые показатели гетерозиготности;

PM – вероятность случайного совпадения;

PD – дискриминирующий потенциал;

PE – индекс исключения.

нии татар Среднего Поволжья, находились некоторое время в составе Хазарского каганата, который характеризовался достаточно высоким процентом иудейского населения. Возможно, тесный контакт этих двух этнических групп нашел отражение в современном генофонде поволжских татар [4].

В заключение хочется отметить, что татарский народ, как и любая другая этническая общность, является продуктом сложного исторического развития. Его этнические основы, по данным археологии, сформировались еще в Волжской Булгарии. Однако многими чертами своей культуры, языка, отчасти и

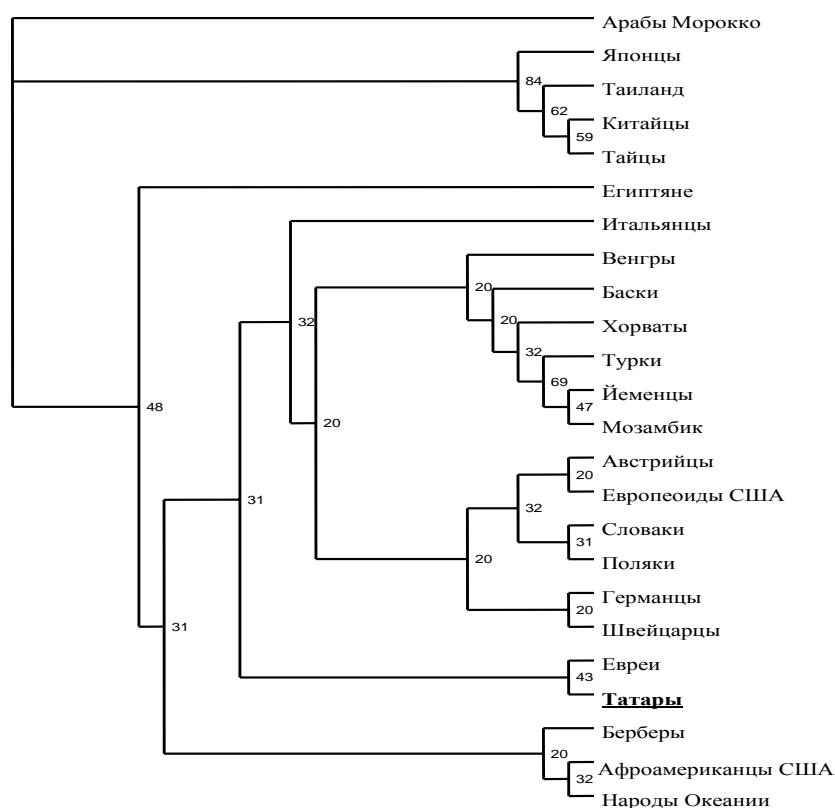


Рис. 4. Консенсусная дендрограмма генетических расстояний, рассчитанная по трем STR-локусам между мировыми популяциями. Цифрами обозначены значения бустрепа (в %)

антропологического типа татарский народ проявляет несомненную близость к тюркоязычным народам (башкирам, казахам, узбекам), формирование которых проходило в тесном контакте двух рас – европеоидной и монголоидной.

Данные, полученные при анализе микросателлитных полиморфных маркеров в современной популяции татар, указывают на большой запас генетической неоднородности, что объясняется интенсивными процессами метисации, происходившими в древности и имеющими место в настоящее время. Выявлено преобладание европеоидного компонента в современном генофонде татар, при этом отмечено присутствие монголоидных черт, особенно в митохондриальном генофонде [5], что согласуется с антропологическими данными.

Summary

O.A. Kravtsova. The structure of nuclear genofound oh the Tatars (based on data of autosomal microsatellite loci).

The allele frequency distribution of 11 short tandem loci (STR) was determined in two Tatar population from Aznakaevo (kazan tatars) and Buinsk (tatar-mishears). A total of 276 individuals were typed. No deviation from the Hardy-Weinberg equilibrium were observed at the 5 STRs whereas 6 loci for both subgroups showed significant departures from Hardy-Weinberg expectations, determined by lack if heterozygous individuals. All loci are highly polymorphic and population differentiation tests showed that the Tatar samples from Azna-

kaevo and Buinsk have significant differences in 3 out of 11 loci (FGA, TPOX и D21S11). Based on allele frequencies of 3 STRs (D3S1358, D5S818, vWA) in the world populations standard genetic distances according to Nei were calculated.

Литература

1. Алтухов Ю.П., Корочкин Л.И., Рычков Ю.Г. Наследственное биохимическое разнообразие в процессах эволюции и индивидуального развития // Генетика. – 1996. – Т. 32, № 11. – С. 1450–1473.
2. Лимборская С.А., Хуснутдинова Э.К., Балановская Е.В. Этногеномика и геногеография народов Восточной Европы. – М.: Наука, 2002. – 261 с.
3. Хуснутдинова Э.К., Викторова Т.В., Ахметова В.Л. и др. Популяционно-генетическая структура чувашей (по данным о восьми ДНК-локусах ядерного генома) // Генетика. – 2003. – Т. 39, № 11. – С. 1550–1563.
4. Газимзянов И.Р. Антропологический облик татар. – М.: Наука, 2001. – 156 с.
5. Кравцова О.А., Аскарлова А.Н. Полиморфизм митохондриальной ДНК в современной популяции Республики Татарстан // Учен. зап. Казан. ун-та. Сер. Естеств. науки. – 2005. – Т. 147, кн. 3. – С. 117–123.
6. Mathew C.G.P. The isolation of high molecular weight eukaryotic DNA. Methods in Molecular Biology. – N. Y., 1984. – P. 311.
7. Ricci U., Sani I., Guarducci S. et al. Infrared fluorescent automated detection of thirteen short tandem repeat polymorphisms and one gender-determining system on the CODIS core system // Electrophoresis. – 2000. – V. 21. – P. 3564–3570.
8. Butler J.M. Recent developments in Y-Short Tandem Repeat and Y-Single Nucleotide Polymorphism analysis // Forensic Sci. Rev. – 2003. – V. 15, No 2. – P. 91–111.
9. Guo S.W., Thompson E.A. Performing the exact test of Hardy-Weinberg proportions for multiple alleles // Biometrics. – 1992. – V. 48. – P. 361–372.
10. Rousset F., Raymond M. Testing heterozygote excess and deficiency // Genetics. – 1995. – V. 140. – P. 1413–1419.
11. Raymond M., Rousset F. An exact test for population differentiation // Evolution. – 1995. – V. 49. – P. 1280–1283.
12. Nei M., Tajima F. DNA polymorphism detectable by restriction endonucleases // DNA Genetics. – 1981. – V. 105. – P. 207–217.
13. Nei M. Molecular evolution genetics. – N. Y.: Columbia Univ. Press, 1987. – 472 p.
14. Солбриг О., Солбриг А. Популяционная биология и эволюция. – М.: Мир, 1982. – 488 с.
15. Степанов В.А. Y-хромосома как основа группоспецифических маркеров у человека // Материалы Второго Съезда Общества биотехнологов России: Москва, 13–15 окт. 2004 г. / Под ред. Р.Г. Василовой. – М.: МАКС Пресс, 2004. – С. 196.

Поступила в редакцию
02.03.07

Кравцова Ольга Александровна – кандидат биологических наук, ассистент кафедры биохимии Казанского государственного университета.

E-mail: okravz@rambler.ru