

УДК 631.523.5;340.64

ИНФОРМАТИВНОСТЬ СТАНДАРТНЫХ АУТОСОМНЫХ STR В ПОПУЛЯЦИЯХ КОРЕННЫХ НАРОДОВ РОССИИ ДЛЯ ОПРЕДЕЛЕНИЯ РОДСТВА ПЕРВОЙ СТЕПЕНИ

© К. В. Вагайцева^{1,*}, М. Е. Лопаткина¹, Н. А. Колесников¹, В. Н. Харьков¹, В. А. Степанов¹

¹Научно-исследовательский институт медицинской генетики, Томский национальный
исследовательский медицинский центр Российской академии наук, Томск, 634050 Россия

*e-mail: kseniya.simonova@medgenetics.ru

Поступила в редакцию 03.11.2023 г.

После доработки 08.12.2023 г.

Принята к публикации 14.12.2023 г.

Действие различных генетико-демографических факторов на формирование генофонда привело к снижению уровня генетического разнообразия в малых изолированных популяциях коренных народов Российской Федерации. Такое снижение влияет на уровень информативности системы генетических маркеров при установлении родственных отношений. В нашем исследовании была проведена оценка эффективности двух стандартных наборов аутосомных STR (13 CODIS, 20 CODIS, Combined DNA Index System) для генетической экспертизы родства “родитель – ребенок” в четырех высоко инбридных популяциях РФ и популяции русских, с использованием двух типов референсных частот. Результаты проведенного исследования подтвердили предположение, что уровень идентификационной информативности стандартных аутосомных маркеров в высокоинбридных популяциях из Сибири и Дагестана ниже, чем в популяции русских. Совокупная информативность маркеров нового стандарта CODIS превосходит требуемые в приказе министерства здравоохранения и социального развития РФ (№ 346н) пороговые значения. При этом вероятность ложноположительного результата возрастает с увеличением коэффициента инбридинга в популяции.

Ключевые слова: ДНК-идентификация, популяционная генетика, уровень идентификационной информативности, определение родства, генетическая экспертиза, референсная база данных.

DOI: 10.31857/S0016675824050078 **EDN:** CJBLZB

Уровень идентификационной информативности генетических маркеров напрямую связан с уровнем генетического разнообразия в популяции. Причина этого заключается в том, что вероятность случайного совпадения генотипов анализируемого объекта и сравниваемого с ним образца ДНК тем меньше, чем больше вариантов аллелей встречается в популяции. При этом уровень генетического разнообразия варьирует между популяциями за счет действия различных генетико-демографических процессов. Популяции, проживающие на наибольшем генетическом расстоянии от места происхождения вида, обладают наименьшим уровнем гетерозиготности. Самый высокий уровень генетического разнообразия характерен для народов Африки, за которыми следуют популяции из Ближнего Востока, Европы, Центральной и Южной Азии. У населения Восточной Азии уровень генетического разнообразия еще ниже, а популяциях

жителей тихоокеанских островов и коренных американцев, находящихся на наибольшем географическом удалении по путям миграции, пройденным в ходе эволюции человека, являются наименее гетерозиготными [1]. Тенденция к снижению уровня генетического разнообразия показана на большом числе нейтральных генетических маркеров, в том числе на аутосомных STR, традиционно используемых для ДНК-идентификации [2].

Одна из последних публикаций, описывающих изменение уровня генетического разнообразия коренных народов России – работа Колесникова Н.А. и др. [3], где с помощью анализа широкогеномных данных была проведена оценка коэффициента инбридинга для 76 популяций коренного населения Российской Федерации и стран ближнего зарубежья. Коэффициент геномного инбридинга отражает долю генома, расположенную в регионе высокой гомозиготности. Таким образом,

наибольший коэффициент геномного инбридинга характерен для популяций с наименьшим уровнем генетического разнообразия. Опубликованные результаты отражают общую тенденцию уменьшения уровня генетического разнообразия с запада на восток Евразии, за исключением популяций Дагестана. Согласно этим данным популяции Кавказа и Средней Азии обладают минимальными значениями коэффициента инбридинга. Популяции Волго-Уральского региона занимают промежуточное положение. Высокий уровень коэффициента инбридинга характерен для популяций Сибири и Дагестана, что вероятнее всего связано со значительной территориальной разделенностью популяций, проживающих в данных регионах, за счет рельефа местности, географических расстояний и климатических условий. Максимальное значение коэффициента инбридинга наблюдается в популяциях гинухцев и гунзибцев (Дагестан), причем показатель для гинухцев является одним из самых высоких в мире и схож лишь с уровнем инбридинга в изолированной популяции индейцев каритиана из бразильской Амазонии [3].

Наглядно снижение уровня идентификационной информативности и значимость выбранной референсной группы при определении родственных отношений продемонстрирована в двух зарубежных работах. В работе R. Rohlf с соавт. [4] уровень идентификационной информативности стандартных аутомомных STR-маркеров оценивался для пяти популяций: африканцы, европейцы США, латиноамериканцы, вьетнамцы, навахо. Анализ проводился *in silico*, то есть для анализа случаи родства генерировались искусственно, на основе популяционных частот. Полученные в этом исследовании результаты подтвердили тот факт, что для определенных локусов уровень идентификационной информативности ниже в группах с низкой дисперсией распределения частот, что приводит к снижению различимости родственных отношений, даже при использовании точных референсных данных. Когда порог принятия решения выбран таким образом, что вероятность идентификации родственников второй степени достаточно высока в популяциях, где уровень генетического разнообразия ниже, чем в используемых референсных частотах, вероятность ложноположительной семейной идентификации выше. Это может усугубляться неизвестными популяционными различиями, например мутациями в сайтах связывания праймеров [4].

В работе V. Zvenigorosky с соавт. [5], на реальных случаях родства, оценивалась эффективность и надежность генетической экспертизы в популяциях якутов (Сибирь) и бенинцев (Африка) при использовании трех типов референсных групп: частоты аллелей в выборке неродственных индивидов (а), частоты аллелей в тотальной выборке

анализируемой популяции (б), усредненные частоты аллелей по мировым данным (в). В данном эксперименте авторы собрали выборки индивидов с известными родственными отношениями разной степени и оценили информативность экспертизы проведенной с помощью 21 аутомомного STR, генотипирование осуществлялось наборами Investigator 24plex (Qiagen) и Globalfiler (Thermo Fisher). Оценивалось два показателя: вероятность обнаружения родства (эффективность) и вероятность ложноположительного результата (надежность). В результате было показано, что в якутской популяции экспертиза «родитель-ребенок» одинаково эффективна и надежна при выборе всех трех типов референсных групп. При этом надежность экспертизы, проводимой для родственников первой степени, в якутской популяции ниже, чем в популяции бенинцев. Эффективность выявления родственной связи между братьями/сестрами оказалась несколько выше при использовании референсных групп якутов (а, б), а в случае определения родства сводных братьев/сестер эффективнее оказалась усредненная референсная выборка (в). Однако надежность такой экспертизы в обоих случаях была выше при использовании референсных частот якутской популяции (а, б). Авторы отмечали, что в популяциях, для которых идентификационная информативность стандартной панели маркеров снижена и/или отсутствуют референсные значения частот, эксперт может быть не в состоянии отличить пару полных братьев и сестер от пары полусибсов [5].

Также о влиянии референсной группы на результаты экспертизы по установлению родственных отношений между индивидами из бурятской популяции ранее сообщалось в публикации Исаевой Е.Н. и др., где показано, что использование некорректной референсной выборки (европеоиды северной Америки) завысило показатели вероятности отцовства [6].

В настоящем исследовании была проведена оценка эффективности двух стандартных наборов аутомомных STR для генетической экспертизы родства первой степени в четырех высоко инбредных популяциях РФ и в популяции русских. Первый набор включает локусы стандарта CODIS (CODIS, Combined DNA Index System), действовавшего до 2017 г. и включавшего 13 аутомомных STR. Второй набор маркеров, анализируемый в данном исследовании — расширенный вариант стандарта CODIS, действующий с января 2017 г. [7].

МАТЕРИАЛЫ И МЕТОДЫ

Для оценки уровня идентификационной информативности системы аутомомных STR в высоко инбредных популяциях РФ мы провели оценку двух криминалистических показателей,

вероятности отцовства (Paternity Probability, **PP**), а также индекса отцовства (Paternity Index, **PI**). Частоты аллелей и генотипы, используемые для расчета, были получены ранее в рамках научно-технической программы Союзного государства “Разработка инновационных геногеографических и геномных технологий идентификации личности и индивидуальных особенностей человека на основе изучения генофондов регионов Союзного государства” (“ДНК-идентификация”), патент КОДИС-Р RU 2021622454 11.11.2021. Для анализа было выбрано пять популяций коренных народов РФ с разным уровнем геномного инбридинга: две популяции из Дагестана (андийцы ($F_{\text{РОН}} = 0.0578$), гинухцы (0.1437)), две популяции Сибири – тувинцы (0.0591) и якуты (0.0723), а также одна популяция русских (0.0377).

Оценка показателей проводилась *in silico*, на основе популяционных данных. В рамках настоящей работы было сгенерировано по 1002 полных трио (для экспертизы доступны образцы предполагаемого родителя, ребенка и родителя, родство с которым не подвергается сомнению), и дуо (для экспертизы доступны образцы предполагаемого родителя и ребенка) для всех пяти популяций. При генерации (случайным образом) случаев из индивидов, составляющих популяционную выборку, мы сформировали по 334 искусственные пары. Для каждой такой пары было сгенерировано три потомка, с учетом того, что родители с равной вероятностью передают своему потомству по одному аллелю в каждом локусе. Частоты аллелей STR-локусов в сгенерированных таким образом выборках статистически значимо не отличались от частот аллелей в первоначальной популяции (критерий хи-квадрат Пирсона, уровень значимости $p < 0.05$).

Помимо оценки криминалистических показателей в сгенерированных случаях был проведен поиск потенциальных родителей среди представителей соответствующих популяций для каждого ребенка. Реальные родители были исключены из анализа.

Расчеты **PP** и **PI** проводили согласно стандартным формулам [8] для двух наборов аутосомных STR (13 локусов CODIS: CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, TH01, TPOX, vWA, и 20 локусов CODIS: CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, TH01, TPOX, vWA, D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433, D22S1045). Для каждого случая оценка проводилась двумя способами: с использованием корректной референсной группы и с использованием средних по России референсных частот [9]. Порог принятия решения о вероятном родстве для трио равнялся $PI \geq 1000$, для дуо $PI \geq 400$. Все программы, используемые

для моделирования и расчетов, были написаны на языке Python.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для того чтобы оценить, достаточно ли информативны стандарты аутосомных STR для экспертизы на родство в высокоинбредных популяциях России, мы оценили значения **PP** и **PI** для 10020 случаев. На основе популяционных данных были сгенерированы 1002 трио (родитель, предполагаемый родитель, ребенок) и 1002 дуо (предполагаемый родитель, ребенок) для пяти популяций с разным уровнем значения коэффициента $F_{\text{РОН}}$ [3]: русские (0.0377), андийцы (0.0578), тувинцы (0.0591), якуты (0.0723), гинухцы (0.1437). В каждой таблице результатов, для удобства анализа популяции расположены в порядке возрастания уровня геномного инбридинга.

При анализе полных трио средние значения индекса отцовства и вероятности отцовства в пяти популяциях превысили пороговые значения, представленные в Приказе Министерства здравоохранения и социального развития РФ №3 46н (табл. 1, 2). Однако при анализе 13 аутосомных STR в популяции гинухцев в 21 (2.1%) сгенерированном случае полных трио и 618 (61.8%) дуо показатели не достигли порога принятия решения ($PI < 1000$, $PI < 400$ соответственно). В популяции якутов, при анализе с помощью 13 аутосомных STR, было выявлено 11 неубедительных результатов (1.1%) экспертизы трио и 428 (42.8%) дуо. В популяции тувинцев три случая трио и 213 случаев дуо не достигли требуемого порога при анализе с помощью 13 STR. При использовании расширенного набора маркеров (20 STR) случаи, когда криминалистические показатели не достигли принятого порогового значения, были выявлены только при экспертизе 16 дуо случаев в популяции гинухцев. Также показано, что использование усредненных референсных частот в высокоинбредных популяциях завышает значения показателей (табл. 3).

Выводы о достаточной информативности напрямую зависят от выбранного порога принятия решения о наличии родственных отношений между анализируемыми индивидами. Например, в Китае для подтверждения родства порог принятия решения составляет 10000 (**PI**) [11]. Если мы возьмем это пороговое значение для оценки информативности в нашем исследовании, будут признаны неубедительными результаты для гораздо большего числа экспертиз (табл. 3). Отметим также, что полученные значения доли экспертиз, не достигших требуемого порога принятия решения, занижены, поскольку генерация “детей” проводилась без учета возможности мутации в анализируемых локусах.

Таким образом, из полученных результатов следует, что информативность расширенного набора

Таблица 1. Средние значения и стандартное отклонение показателей идентификационной информативности в случаях экспертизы полных трио

Популяция	Набор	Точная референсная группа		Референсная группа, россияне	
		<i>PP</i> (sd)	<i>PI</i> (sd)	<i>PP</i> (sd)	<i>PI</i> (sd)
Рус.	13 маркеров	0.999992 (0.00002)	21763079847 (332918895769)	0.999993 (0.00002)	12253095569618800 (183235273388355000)
Анд.		0.999993 (0.00002)	535289816 (6693901441)	0.999995 (0.00001)	1283830736984920000 (26676491095381200000)
Тув.		0.99998 (0.00008)	66681911 (848332027)	0.999994 (0.00003)	5109514684 (128794392739)
Як.		0.99992 (0.0003)	9289861106 (148171240278)	0.999991 (0.00002)	47828972895203200000 (1000459335519990000000)
Гин.		0.9998 (0.0003)	1960950 (13500794)	0.999996 (0.00001)	413973204216057 (5224276537948190)
Рус.	20 маркеров	0.99999996 (0.00000002)	1157225212913760 (14282006053073000)	0.99999996 (0.00000001)	6763881348692140000000000 (1048682404250870000000000000)
Анд.		0.99999999 (0.0000001)	126710591330963 (2449234403926140)	0.99999997 (0.00000003)	4918512278695320000000000000000 (1061373650778270000000000000000000)
Тув.		0.99999998 (0.00000007)	847376637300 (9951259215426)	0.99999998 (0.000000009)	273387848232467 (8137984535706190)
Як.		0.9999998 (0.0000007)	68531967847183 (1167928870429900)	0.99999997 (0.00000001)	2916363053944620000000000000 (5993798576654140000000000000)
Гин.		0.999998 (0.000006)	30066518080 (256025790550)	0.999999991 (0.000000004)	9259553187056510000 (1265526481348840000000)

Примечание. Рус. – русские, Анд. – андийцы, Тув. – тувинцы, Як. – якуты, Гин. – гинухцы. В скобках указаны значения стандартного отклонения.

стандартных аутосомных STR превосходит требуемые в РФ на настоящий момент пороговые значения. Однако стоит помнить о том, что, несмотря на высокую информативность современных коммерческих мультиплексов, вероятность ошибочного заключения сохраняется. Так, в литературе описан случай, когда генетическая экспертиза, проведенная с помощью 46 аутосомных маркеров STR, получила ложноположительный результат при установлении родственных отношений между предполагаемой матерью и ребенком [12]. В публикации Ковтуна П.А и др. при установлении родственных отношений между предполагаемыми родителями и детьми ложноположительный результат был получен при анализе набора из 15 аутосомных STR [13]. Во всех трех описанных случаях предполагаемый родитель и ребенок имели общие аллели по каждому анализируемому локусу.

В связи с этим для каждого сгенерированного ребенка мы провели поиск индивидов в популяции, для которых результат вероятностных расчетов превысил бы значения порога принятия решения. Пара, на основе генотипов которой генерировался генетический профиль ребенка, из поиска была исключена. Таким образом, была рассчитана доля ложноположительных результатов при

условии, что у предполагаемого родителя и ребенка отсутствуют несовпадения аллелей (табл. 4). Из таблицы видно, что процент таких ложноположительных результатов растет при увеличении уровня инбридинга в популяции, не вписывается в эту тенденцию только популяция андийцев.

В целом, результаты настоящего исследования подтвердили предположение, что уровень идентификационной информативности стандартных аутосомных маркеров в высокоинбредных популяциях Сибири и Дагестана ниже, чем в популяции русских. Совокупная информативность маркеров нового стандарта CODIS превосходит требуемые в Приказе Министерства здравоохранения и социального развития РФ (№ 346н) пороговые значения. При этом вероятность ложноположительного результата возрастает с увеличением коэффициента инбридинга в популяции.

Данные результаты правомерны для анализируемой нами модели популяции, однако нужно отметить, что в данном варианте расчетов при генерации случаев мы приняли допущение, что в популяции отсутствует ассортативность. При генерации второго поколения и поиске ложноположительных связей не учитывалась возможность возникновения случайных мутаций. Поиск

Таблица 2. Средние значения и стандартное отклонение показателей идентификационной информативности в случаях экспертизы "один родитель — один ребенок" / "родитель — ребенок"

Популяция	Набор	Точная референсная группа		Референсная группа, россияне	
		<i>PP</i> (sd)	<i>PI</i> (sd)	<i>PP</i> (sd)	<i>PI</i> (sd)
Рус.	13 маркеров	0.9992 (0.001)	7975 (34066)	0.9993 (0.0008)	25682 (291796)
Анд.		0.9993 (0.0009)	12397 (101573)	0.9996 (0.0006)	19635497 (404092850)
Тув.		0.998 (0.002)	3044 (10955)	0.9995 (0.0006)	32121 (403925)
Як.		0.996 (0.004)	3145 (14838)	0.9995 (0.0007)	128815 (1209796)
Гин.		0.994 (0.006)	828 (2774)	0.9996 (0.0005)	58369871 (640640448)
Рус.	20 маркеров	0.999995 (0.000009)	181614399 (3617645858)	0.999996 (0.000008)	73406389580528600 (1869619268014010000)
Анд.		0.999992 (0.00001)	4099054 (40093714)	0.999997 (0.000005)	8501869983 (134260262672)
Тув.		0.99997 (0.00005)	503869 (2546337)	0.999997 (0.000006)	13146120 (143938396)
Як.		0.9999 (0.0002)	1085648 (23218559)	0.999997 (0.000005)	2864383370 (52431704081)
Гин.		0.9996 (0.0006)	55653 (417673)	0.999998 (0.000003)	51318486753 (680675090558)

Примечание. Рус. — русские, Анд. — андийцы, Тув. — тувинцы, Як. — якуты, Гин. — гинухцы. В скобках указаны значения стандартного отклонения.

Таблица 3. Количество анализируемых случаев, не преодолевших порог принятия решения

	Популяция	13 STR		20 STR	
		трио	дуо	трио	дуо
$PI > 400$ (дуо), $PI > 1000$ (трио)	Русские	0 (0)	259 (35)	0 (0)	0 (0)
	Андийцы	0 (0)	48 (17)	0 (0)	0 (0)
	Тувинцы	3 (0)	213 (18)	0 (0)	0 (0)
	Якуты	11 (0)	428 (131)	0 (0)	0 (0)
	Гинухцы	21 (0)	618 (9)	0 (0)	16 (0)
$PI > 10000$ (дуо, трио)	Русские	8 (4)	863 (855)	0 (0)	2 (1)
	Андийцы	8 (5)	816 (713)	0 (0)	4 (0)
	Тувинцы	41 (7)	957 (809)	0 (0)	54 (1)
	Якуты	149 (17)	954 (766)	0 (0)	249 (0)
	Гинухцы	202 (2)	994 (609)	1 (0)	611 (0)

Примечание. В скобках указаны значения, полученные при использовании средних референсных частот.

Таблица 4. Процент экспертиз, показавших ложноположительные родственные связи (случаи, где у предполагаемого родителя и ребенка были несовпадения аллелей по 1–2 локусам, не учитывались)

Популяция	13 STR		20 STR	
	трио	дуо	трио	дуо
Русские	0 (0)	0.04 (0.05)	0 (0)	0.003 (0.003)
Андийцы	0 (0)	0.035 (0.047)	0 (0)	0 (0)
Тувинцы	0 (0)	0.05 (0.11)	0 (0)	0.004 (0.004)
Якуты	0.009 (0.01)	0.079 (0.44)	0 (0)	0.01 (0.01)
Гинухцы	0 (0)	0.027 (0.23)	0 (0)	0.019 (0.023)

Примечание. В скобках указаны значения, полученные при использовании средних референсных частот.

ложноположительных вариантов предполагаемых родителей осуществлялся только в одном поколении и только среди индивидов, не являющихся родственниками. Внесение соответствующих поправок в модель приведет к увеличению доли ложноположительных и неубедительных результатов.

Согласно полученным в настоящей работе данным, для проведения экспертизы родства первой степени в популяциях Дагестана и Сибири рекомендуется выбирать наборы для мультиплексного анализа, включающие расширенный состав стандартных маркеров (20 и более). В случаях, когда для определения родства используются наборы с меньшим числом маркеров, необходимо использовать точные референсные данные — это позволит избежать завышенных значений криминалистических показателей. При этом референсные частоты должны быть получены на однородных выборках, то есть включающих только индивидов, принадлежащих к конкретной этно-территориальной группе. Более того, на наш взгляд, полученные результаты подчеркивают необходимость использования дополнительных типов генетических маркеров (X, Y, mtДНК) при проведении генетической экспертизы, касающейся определения родственных отношений в дуо случаях.

Исследование выполнено за счет гранта Российского научного фонда №23-74-10058, <https://rscf.ru/project/23-74-10058/>.

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Rosenberg N.A., Kang J.T.L. Genetic diversity and societally important disparities // *Genetics*. 2015. V. 201. № 1. P. 1–12. doi: 10.1534/genetics.115.176750
2. Silva N.M., Pereira L., Poloni E. S., Currat M. Human neutral genetic variation and forensic STR data // *PLoS One*. 2012. V. 7. № 11. <https://doi.org/10.1371/journal.pone.0049666>
3. Колесников Н.А., Харьков В.Н., Зарубин и др. Особенности геномного распределения регионов высокой гомозиготности у коренного населения Северной Евразии на индивидуальном и популяционном уровнях на основе анализа SNP высокой плотности // *Генетика*. 2021. Т. 57. № 11. С. 1261–1275. <https://doi.org/10.31857/S0016675821110059>
4. Rohlf R.V., Fullerton S.M., Weir B.S. Familial identification: Population structure and relationship distinguishability // *PLoS Genet*. 2012. V. 8. № 2. <https://doi.org/10.1371/journal.pgen.1002469>
5. Zvenigorosky V., Sabbagh A., Gonzalez A. et al. The limitations of kinship determinations using STR data in ill-defined populations // *Int. J. Legal. Med*. 2020. V. 134. №. 6. P. 1981–1990. <https://doi.org/10.1007/s00414-020-02298-w>
6. Исаева Е.Н., Исаев Ю.С., Семинский И.Ж. Значение ДНК маркеров при судебно-медицинской экспертизе родства у этнических бурят Байкальского региона Восточной Сибири // *Сиб. мед. журнал (Иркутск)*. 2011. Т. 104. № 5. С. 24–27. Combined DNA Index System (CODIS): Overview // FBI Resources for Law Enforcement. URL: <https://le.fbi.gov/science-and-lab/biometrics-and-fingerprints/codis#Overview> (Accessed 11 Dec 2023).

7. *Культин А.Ю.* Экспертная оценка и вероятностно-статистическая обработка результатов исследования ДНК при установлении биологического родства. М.: ЭКЦ МВД России, 2011. 132 с.
8. *Культин А.Ю., Полякова А.В.* Генетическое разнообразие 38 аутосомных STR-маркеров в российской популяции // Судебная экспертиза Беларуси. 2023. Т. 1. № 16. С. 69–73.
9. *Культин А.Ю., Полякова А.В., Константинов С.В. и др.* Частоты встречаемости аллелей аутосомных STR-локусов у жителей России. Справочник. М.: ЭКЦ МВД России, 2022. 276 с.
10. *Gao H., Wang C., Zhang R., Wu H. et al.* Application of CPI cutoff value based on parentage testing of duos and trios typed by four autosomal kits // PLoS One. 2019. V. 14. № 11. <https://doi.org/10.1371/journal.pone.0225174>. eCollection 2019
11. *Li L., Ge J., Zhang S.* Maternity exclusion with a very high autosomal STRs kinship index // Int. J. Legal Med. 2012. V. 126. № 4. P. 645–648. <https://doi.org/10.1007/s00414-012-0668-8>
12. *Ковтун П.А., Куклев М.Ю., Лапенков М.И., Плахина Н.В.* Недостаточность аутосомных STR-маркеров для достоверного установления родства в дутах родитель – ребенок // Судебно-медицинская экспертиза. 2013. Т. 56. № 6. С. 17–22.

Identifiable Information Value of Standard Autosomal STR in Indigenous Populations of Russia for Determination of First-Degree Relatedness

K. V. Vagaitseva¹, *, M. E. Lopatkina¹, N. A. Kolesnikov¹, V. N. Kharkov¹, V. A. Stepanov¹

¹Research Institute of Medical Genetics, Tomsk National Research Medical Center, Tomsk, 634050 Russia

*e-mail: kseniya.simonova@medgenetics.ru

The effects of the demographic history of mankind have led to the fact that the indigenous peoples of Dagestan and Siberia are inferior in terms of genetic diversity to the populations of Europe, which affects the level of identification informativeness of standard forensic autosomal markers in these populations. In our study, we evaluated the effectiveness of two standard sets of autosomal STRs (13 CODIS, 20 CODIS, Combined DNA Index System) for the genetic testing of parent-child relatedness in four highly inbred populations of the Russian Federation and the Russian population, using two types of reference frequencies. The results of the study confirmed the assumption that the level of identification informativity of standard autosomal markers in highly inbred populations of Siberia and Dagestan is lower than in the Russian population. The total information content of markers of the new CODIS standard exceeds the threshold values required in the order of the Ministry of Health and Social Development of the Russian Federation (No. 346n). At the same time, the probability of a false positive result increases with an increase in the inbreeding coefficient in the population.

Keywords: forensic genetics, population genetics, level of identification informativity, kinship determination, genetic test, reference database.