

УДК 577.17

## О ГЕНЕТИЧЕСКИХ ХАРАКТЕРИСТИКАХ НАСЕЛЕНИЯ ЦЕНТРАЛЬНОЙ РОССИИ

© И.Н. Сорокина, В.И. Евдокимов, М.И. Чурносков

*Ключевые слова:* Центральная Россия; генетическая структура популяции; урбанизация; фамилии; численность населения.

Охарактеризована популяционно-генетическая структура населения Центральной России по данным антропоники (21556 фамилий, 10 районов шести областей, численность населения 242671 человек). Уровень генетической подразделенности элементарных популяций (район) значительно варьирует ( $0,00002 < f_r^* < 0,00092$ ) и зависит от степени урбанизации популяции, ее численности, количества представленных в ней фамилий. Районы, имеющие в своем составе урбанизованную зону (город), большую численность населения и количество представленных в них фамилий, характеризуются низким уровнем внутренней подразделенности. Районы с высоким уровнем генетической дифференциации отличаются меньшей численностью населения, небольшим числом фамилий и отсутствием урбанизованной зоны.

### ВВЕДЕНИЕ

В настоящее время в мире интенсивно проводятся исследования генетической структуры различных популяций: изолированных и со значимой ролью миграций, однонациональных и многонациональных, существующих в различных климатогеографических зонах. Подробное генетическое описание отдельных популяций создает реальные предпосылки для более глубокого понимания роли инбридинга, дрейфа генов, миграционных процессов в формировании здоровья населения и эволюции популяций [1].

Существует множество различных способов описания генетической структуры популяций с помощью разнообразных параметров. Одним из маркеров, эффективно используемых в популяционной генетике, являются фамилии («квазигенетические маркеры»). Фамилии рассматриваются как аллели одного селективно нейтрального локуса [2] и используются для определения уровня инбридинга и его составляющих в популяции (изонимный метод Кроу и Манжа [3]), оценке генетических расстояний между различными популяциями, расчета показателей разнообразия фамилий и др., описания генетического ландшафта с помощью геногеографических технологий [4].

С использованием данных антропоники к настоящему времени изучена генетическая структура целого ряда популяций как зарубежных [3, 5, 6], так и популяций на пространстве бывшего СССР: русских [2, 4, 7-12], адыгских [13], марийских [14] и других [15, 16]. Однако русский генофонд изучен достаточно фрагментарно. Следует отметить, что генетические факторы, наряду с экологическими, социально-гигиеническими и другими факторами, оказывают значимое влияние на здоровье населения.

В данной работе представлены результаты изучения генетической структуры населения десяти районных популяций шести областей Центральной России на основе распределения частот фамилий.

### МАТЕРИАЛЫ И МЕТОДЫ

Проведено изучение структуры генофонда населения Центральной России с использованием данных антропоники – проанализировано распределение фамилий в 10 районах шести областей Центральной России. Это Михайловский и Спасский районы Рязанской области, Боровский и Бяратинский районы Калужской области, Петровский район Тамбовской области, Болховский и Ливенский районы Орловской области, Пристенский и Черемисиновский районы Курской области, Репьевский район Воронежской области. Источником данных о фамилиях послужили списки избирателей, которые охватили «тотально» все население старше 18 лет. В анализ включены данные об изменчивости 21556 фамилий среди 242671 человека Центральной России. Вычисление популяционно-генетических характеристик осуществлялось на районном уровне (элементарная популяция [12]). Частота фамилий рассчитывалась через отношение числа носителей данной фамилии в изучаемой популяции к общему числу жителей старше 18 лет в данной популяции.

Случайный инбридинг по частотам фамилий оценивался с помощью коэффициента изонимии  $f_r$ , предложенного Grow и Mange [3]. Расчет производился по формуле:

$$f_{r(j)} = I/4,$$

где  $I$  – ожидаемая частота изонимных (однофамильных) браков в  $j$ -й популяции, т. е.  $I = \sum P_i^2$ , где  $P_i$  – частота  $i$ -й фамилии в  $j$ -й субпопуляции, а коэффициент  $1/4$  учитывает передачу фамилий лишь по мужской линии. Случайный инбридинг  $f_r$  (соответствует  $F_{ST}$  Райта) оценивает ожидаемую частоту однофамильных браков в предположении полной панмиксии (элементарная популяция без ассортативности браков). Следует отметить, что данная величина  $f_r$  (район) отражает

Таблица 1

Характеристика распределения фамилий, уровня подразделенности ( $f_r^*$ ), разнообразия фамилий ( $\alpha$ ), энтропии ( $H$ ), избыточности распределения фамилий ( $R$ ) и индекса миграций ( $\nu$ ) в районных популяциях Центральной России

Район, область	Количество населения старше 18 лет	Количество фамилий	Количество фамилий на 1000 человек	$f_r^*$	$\alpha$	$H$	$R$	$\nu$
Барятинский район, Калужская область	5100	1451	285	0,00027	190,65	9,86	19,93	0,1161
Боровский район, Калужская область	41886	9089	217	0,00002	359,25	11,99	21,91	0,0380
Михайловский район, Рязанская область	29425	4485	152	0,00011	150,09	10,91	26,55	0,0324
Спаский район, Рязанская область	26741	4470	167	0,00007	199,51	11,06	24,76	0,0427
Болховский район, Орловская область	15748	2475	157	0,00025	101,38	10,16	27,13	0,0393
Ливенский район, Орловская область	67158	6248	93	0,00030	52,39	10,47	34,72	0,0083
Петровский район, Тамбовская область	16802	2227	133	0,00041	60,89	9,80	30,15	0,0266
Черемисиновский район, Курская область	9266	1311	141	0,00092	33,74	8,99	31,78	0,0251
Пристенский район, Курская область	16198	2227	137	0,00062	45,70	9,49	32,10	0,0201
Репьевский район, Воронежская область**	14347	457**	**	0,00063	**	**	**	**
В среднем	24267	3776	165	0,00036	132,6	10,30	27,67	0,0387

Примечание. \*\* – проанализированы только частые фамилии (число носителей которых в изучаемой популяции было более 4).

случайный инбридинг в каждой отдельной элементарной популяции (район), не являясь характеристикой тотальной популяции (область). При этом величина инбридинга в элементарной популяции (район) уже содержит в себе вклады всех более высоких уровней популяционной системы (область и т. д.). Поэтому для реальной оценки  $f_r^*$  мы вычитали из валовой (прямой) районной оценки  $f_r$  (район) ту величину изменчивости, которая привносится областным уровнем иерархии  $f_r$  (область):  $f_r^* \text{ (район-область)} = \overline{f_r} \text{ (район)} - f_r \text{ (область)}$ , где  $\overline{f_r}$  (район) – это средний уровень инбридинга на уровне элементарной популяции (район), полученный в результате усреднения оценки  $f_r$  по районам, а  $f_r$  (область) – случайный инбридинг в области в целом [13].

По данным о частотах всех фамилий также рассчитаны предложенные Г.И. Ельчиновой индекс миграции ( $\nu$ ), показатель разнообразия фамилий ( $\alpha$ ), энтропия распределения фамилий ( $H$ ), избыточность распределения фамилий ( $R$ ) [8, 14, 15].

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Проведен анализ генетической структуры населения 10 районов шести областей Центральной России на основе данных о фамилиях избирателей.

В районных популяциях Центральной России (табл. 1) число фамилий варьирует в широких пределах. Минимальное количество фамилий выявлено в Черемисиновском районе Курской области (1311 фамилий, при численности 9266 человек), максимальное – в Боровском районе Калужской области (9089 фамилий, 41886 человек).

Среднее число фамилий в районе – 3775. Следует отметить значительную вариабельность (3 раза) такого показателя, как количество фамилий на 1000 человек. Так, в Ливенском районе Орловской области отмечено лишь 93 фамилии на 1000 человек, тогда как в Барятинском районе Калужской области их насчитывается 285.

Уровень подразделенности в районных популяциях Центральной России (табл. 1) в среднем по региону составил  $\overline{f_r^*} = 0,00036$ , варьируя от 0,00002 в Боровском районе Калужской области до 0,00092 в Черемисиновском районе Курской области. Таким образом, наблюдается значительная территориальная вариабельность (46 раз) данного показателя по районным популяциям Центральной России. Следует отметить, что уровень подразделенности в популяциях Центральной России имеет отрицательные корреляционные связи с количеством фамилий (табл. 2). При этом количество фамилий и численность населения достоверно коррелируют между собой ( $\rho = 0,91$ ,  $p < 0,01$ ).

Полученные результаты по уровню подразделенности Центральной России согласуются с полученными ранее данными по Белгородской области [9]: среднее значение подразделенности в районах Белгородской области  $f_r^* = 0,00031$ , варьируя от 0,00007 в Старооскольском районе до 0,00125 в Красненском районе. Значительное колебание аналогичного показателя показано для популяций Костромской области (от 0,0003 до 0,0050) [7], Кировской области (от 0,00021 до 0,00109) [8], Ростовской области (0,00016 до 0,0013) [17].

Таблица 2

Коэффициенты корреляции Спирмена между популяционно-генетическими характеристиками среди населения Центральной России

Сравниваемые показатели	$\rho$	$p$
$a - f_r^*$	-0,93	< 0,01
$H - f_r^*$	-0,95	< 0,001
$R - f_r^*$	-0,72	< 0,05
$\nu - f_r^*$	-0,63	< 0,05
Число фамилий – численность населения	0,91	< 0,001
Число фамилий – $f_r^*$	-0,73	< 0,05
Численность населения – $f_r^*$	-0,47	0,21

Среди исследуемых популяций Центральной России наименьший уровень подразделенности ( $0,00001 \leq f_r^* \leq 0,00030$ ) отмечен в районах Калужской (Боровский, Барятинский районы), Рязанской (Спасский, Михайловский районы) и Орловской (Болховский, Ливенский районы) областей. Для рассматриваемых районов Курской (Пристенский, Черемисиновский районы), Воронежской (Репьевский район) и Тамбовской (Петровский район) областей характерны максимальные значения уровня подразделенности ( $0,00041 \leq f_r^* \leq 0,00092$ ) (табл. 1). Одним из факторов, обуславливающих различия в уровне дифференциации изучаемых популяций Центральной России, может являться их степень урбанизации. Анализ уровня подразделенности районных популяций Центральной России в зависимости от их степени урбанизации подтверждает данное предположение (табл. 3). Популяции, имеющие в своем составе урбанизированную зону (город), имеют значительно более низкий уровень подразделенности ( $f_r^*$  в среднем 0,00015). Тогда как в популяциях, представленных только сельским населением, генетическая дифференциация в 4 раза выше районов с городским населением. Следует отметить и более высокую численность населения, и количество фамилий среди населения в популяциях с урбанизированной зоной (табл. 3). Эти результаты полностью согласуются с полученными нами ранее данными по Белгородской области, где случайный инбридинг для сельского населения в среднем в 2–4 раза выше, чем для городского [10]. Значительные различия в уровне подразделенности среди городского и сельского населения установлены также и для Ростовской [17], Костромской [8], Кировской [8] областей и др. популяций России.

Вариабельность популяционно-генетических характеристик населения Центральной России, продемонстрированная с помощью  $f_r^*$  подтверждается также показателями разнообразия фамилий ( $\alpha$ ), энтропии ( $H$ ), избыточности распределения фамилий ( $R$ ), индекса миграции ( $v$ ) (табл. 1).

Далее нами была проведена оценка эффективности использования этих показателей для корректного описания генетической структуры населения Центральной России. Установлена значимая отрицательная корреляционная связь между случайным инбридингом  $f_r^*$ , с одной стороны, и показателями  $R$ ,  $\alpha$ ,  $H$  и  $v$ , с другой стороны (табл. 2). Максимально высокие отрицательные корреляционные связи ( $\rho = -0,93 \dots -0,95$ ) наблюдались между показателями разнообразия фамилий ( $\alpha$ ), энтропии ( $H$ ) и уровнем подразделенности ( $f_r^*$ ) (табл. 2). Эти результаты свидетельствуют о том, что данные показатели практически полностью соответствуют уровню подразделенности. Аналогичной силы и направленности взаимосвязи между данными показателями ( $\rho = -0,90 \dots -0,96$ ) были установлены нами в Белгородской популяции [10], а также Г.И. Ельчиновой [8] в популяциях Костромской и Кировской областей (на популяциях ранга районов). В данных популяциях Г.И. Ельчинова выявила, что, по крайней мере, два показателя генетической структуры популяции –  $\alpha$  (показатель разнообразия фамилий) и  $H$  (энтропия) дают вполне адекватное описание состояния популяции, которое можно назвать степенью ее панмиксности. В некоторых случаях показатель энтропии дает более точное описание генетической структуры популяций, чем значение  $f_r^*$ , рассчитанное на основе частот фамилий [8]. Среди изучаемых популяций Централь-

ной России обнаруживается значимая отрицательная корреляция средней силы между показателями индекса избыточности распределения фамилий  $R$  и уровнем подразделенности  $f_r^*$ . Однако следует отметить, что в других изученных популяциях РФ данные показатели имеют положительные корреляционные взаимосвязи (Костромская и Кировская области [8], наши данные по Белгородской области [10]). Для районных популяций Центральной России установлена значимая отрицательная корреляционная связь средней силы между случайным инбридингом  $f_r^*$  и индексом миграции  $v$ . Аналогичные результаты получены для Костромской и Кировской областей [8], тогда как в Белгородской области индекс миграций ( $v$ ) не обнаруживал значимых корреляций с уровнем  $f_r^*$ .

В результате проведенного исследования установлено, что для адекватной оценки популяционно-генетической структуры населения Центральной России могут быть использованы два показателя – показатель разнообразия фамилий ( $\alpha$ ) и энтропии ( $H$ ). Лишь эти два показателя имеют однозначно однонаправленные взаимосвязи с  $f_r^*$  в различных группах популяций Центральной России.

Таким образом, полученные результаты свидетельствуют о значительной территориальной вариабельности (46 раз) уровня подразделенности населения в 10 районных популяциях шести областей Центральной России при среднем значении  $f_r^* = 0,00033$ . Наибольший

Таблица 3

Характеристика распределения фамилий и уровня подразделенности ( $f_r^*$ ) в популяциях Центральной России с различной степенью урбанизации

Популяции (район, область)	Средняя численность населения старше 18 лет	Количество фамилий	$f_r^*$
<b>Популяции с урбанизированной зоной</b> (Михайловский и Спасский районы Рязанской области; Болховский и Ливенский районы Орловской области; Боровский район Калужской области)	36192	5353	0,00015
<b>Популяции без урбанизированной зоны</b> (Барятинский район Калужской области; Петровский район Тамбовской области; Черемисиновский и Пристенский районы Курской области; Репьевский район Воронежской области)*	12343	1804	0,00060

*Примечание.* \* – по Репьевскому району количество фамилий в анализе не учитывалось, т. к. в этом районе были изучены только частые фамилии (число носителей которых в изучаемой популяции было более 4).

уровень генетической дифференциации наблюдается в Черемисиновском районе Курской области и Репьевском районе Воронежской области ( $f_r^* > 0,00077$ ), а наименьший – в Спасском и Михайловском районах Рязанской области, Боровском районе Калужской области ( $f_r^* < 0,00012$ ). Значимое влияние на генетическую дифференциацию населения оказывает уровень урбанизации. Показатели разнообразия фамилий ( $\alpha$ ) ( $\rho = -0,93$ ) и энтропии ( $H$ ) ( $\rho = -0,95$ ) практически полностью соответствуют уровню подразделенности ( $f_r^*$ ) и поэтому могут использоваться для корректного описания генетической структуры населения Центральной России. Генетические факторы (уровень подразделенности), наряду с социальными, экономическими и др. факторами оказывают значимое влияние на здоровье населения. Полученные результаты проведения настоящего исследования необходимы для дальнейшего изучения роли популяционно-генетических факторов в распространенности наследственно-детерминированной патологии, и могут послужить основой для планирования генетико-эпидемиологического обследования населения Центральной России.

#### ЛИТЕРАТУРА

1. *Курьшикин А.В., Османов Э.М.* Анализ медико-демографической ситуации как механизм планирования стратегии совершенствования организации медицинской помощи сельскому населению Тульской области // Вестник Тамбовского университета. Серия Естественные и технические науки. Тамбов, 2010. Т. 15. Вып. 2. С. 665-668.
2. *Казаченко Б.Н., Реватов А.А., Тарлычева Л.В., Лавровский В.А.* Использование фамилий для изучения факторов динамики популяционной структуры // Генетика. 1980. Т. 16. № 11. С. 2049-2057.
3. *Crow J.F., Mange A.P.* Measurement of inbreeding from the frequency of marriages between person of the same surname // Eugen. Quart. 1965. V. 12. P. 199-203.
4. *Балановский О.П., Бужилова А.П., Балановская Е.В.* Русский генофонд. Геногеография фамилий // Генетика. 2001. Т. 37. № 7. С. 974-990.
5. *Branco C.C., Mota-Vieira L.* Population structure of Sao Miguel Island, Azores: a surname study // Hum. Biol. 2003. V. 75. № 6. P. 929-939.
6. *Barral I., Rodriguez-Larralde A., Mamolini E. et al.* Elements of the surname structure of Austria // Ann. Human Biol. 2000. V. 27. № 6. P. 607-622.
7. *Гинтер Е.К., Зинченко Р.А., Ельчинова Г.И. и др.* Роль факторов популяционной динамики в распространении наследственной патологии в российских популяциях // Медицинская генетика. 2004. Т. 3. № 12. С. 548-555.
8. *Ельчинова Г.И.* Опыт применения методов популяционно-генетического анализа при изучении популяций России с различной генетико-демографической структурой: автореф. дис. ... д-ра биол. наук. М.: МГНЦ РАМН, 2001. 48 с.
9. *Балановская Е.В., Соловьева Д.С., Балановский О.П. и др.* «Фамильные портреты» пяти русских регионов // Медицинская генетика. 2005. № 1. С. 2-10.
10. *Сорокина И.Н., Балановская Е.В., Чурносов М.И.* Генофонд населения Белгородской области. I. Дифференциация всех районных популяций по данным антропонимики // Генетика. 2007. Т. 43. № 6. С. 841-849.
11. *Сорокина И.Н., Чурносов М.И., Балановская Е.В.* Генофонд населения Белгородской области. II. «Фамильные портреты» в группах районов с разным уровнем подразделенности и роль миграций в их формировании // Генетика. 2007. Т. 43. № 8. С. 1120-1128.
12. *Чурносов М.И., Сорокина И.Н., Балановская Е.В.* Генофонд населения Белгородской области. Динамика индекса эндогамии в районных популяциях // Генетика. 2008. Т. 44. № 8. С. 1117-1125.
13. *Балановская Е.В., Нурбаев С.Д., Балановский О.П. и др.* Геногеографический анализ подразделенной популяции. I. Генофонд адыгов в системе кавказских генофондов // Генетика. 1999. Т. 35. № 6. С. 818-828.
14. *Ельчинова Г.И., Кравчук О.И., Старцева Е.А. и др.* Луговые марийцы: гены, фамилии и миграции // Генетика. 1996. Т. 32. № 10. С. 1421-1422.
15. *Ельчинова Г.И., Тарская Л.А.* Анализ распределения фамилий в Республике Саха (Якутия) // Медицинская генетика. 2006. Т. 5. № 7 (49). С. 21-29.
16. *Кучер А.Н., Данилова А.Л., Конева Л.А., Наговицина А.Н.* Популяционная структура сельских населенных пунктов Республики Саха (Якутия): фамильная структура // Генетика. 2007. Т. 43. № 6. С. 818-826.
17. *Амелина С.С., Шокарев Р.А., Кривенцова Н.В. и др.* Генетико-эпидемиологическое изучение Ростовской области // Медицинская генетика. 2005. Т. 4. № 8. С. 371-377.

Поступила в редакцию 25 сентября 2011 г.

#### Sorokina I.N., Evdokimov V.I., Churnosov M.I. ABOUT GENETIC CHARACTERISTICS OF POPULATION CENTRAL RUSSIA

The population genetic structure of the population of the Central Russia by anthroponimics data (21556 surnames, 10 areas of 6 areas, population of 242671 persons) is characterized. Level  $f_r^*$  elementary populations considerably varies (area) ( $0.00002 < f_r^* < 0.00092$ ) also depends on degree of an urbanization of population, its number, quantity of the surnames presented to it. The areas incorporating urbanization zone (city), the big population and quantity of the surnames presented to them are characterized by low level  $f_r^*$ . Areas with high level of genetic differentiation range with smaller population, a small number of surnames and absence of urbanization zones

*Key words:* Central Russia; genetic structure of population; urbanization; surnames; number of population.