

На правах рукописи

**БАЛАНОВСКИЙ
Олег Павлович**

**ИЗМЕНЧИВОСТЬ ГЕНОФОНДА В ПРОСТРАНСТВЕ И ВРЕМЕНИ:
СИНТЕЗ ДАННЫХ О ГЕНОГЕОГРАФИИ
МИТОХОНДРИАЛЬНОЙ ДНК И Y-ХРОМОСОМЫ**

03.02.07 – генетика

03.01.03 – молекулярная биология

**АВТОРЕФЕРАТ
диссертации на соискание ученой степени
доктора биологических наук**

Москва – 2012

Работа выполнена в Федеральном государственном бюджетном учреждении
«Медико-генетический научный центр» Российской академии медицинских наук.

Научные консультанты:

доктор биологических наук, академик Эстонской АН **Виллемс Рихард Лео-Энделевич**;
доктор биологических наук, член-корреспондент РАН **Янковский Николай Казимирович**.

Официальные оппоненты:

Степанов Вадим Анатольевич, доктор биологических наук, профессор
Федеральное государственное бюджетное учреждение «Научно-исследовательский
институт медицинской генетики» Сибирского отделения Российской академии
медицинских наук, заместитель директора по научным вопросам;

Спицын Виктор Алексеевич, доктор биологических наук, профессор
Федеральное государственное бюджетное учреждение «Медико-генетический научный
центр» Российской академии медицинских наук, заведующий лабораторией
экологической генетики;

Захаров-Гезехус Илья Артемьевич, доктор биологических наук, профессор, член-
корреспондент Российской академии наук
Федеральное государственное бюджетное учреждение науки Институт общей генетики
им. Н.И. Вавилова Российской академии наук, советник РАН.

Ведущая организация:

Федеральное государственное бюджетное образовательное учреждение высшего
профессионального образования «Московский государственный университет имени
М.В. Ломоносова», биологический факультет, кафедра генетики.

Защита состоится «__» _____ 2012 г. в __ часов на заседании Диссертационного ученого
совета Д 001.016.01 при Федеральном государственном бюджетном учреждении «Медико-
генетический научный центр» Российской академии медицинских наук (115478, Москва, ул.
Москворечье, д.1).

С диссертацией можно ознакомиться в библиотеке Федерального государственного
бюджетного учреждения «Медико-генетический научный центр» Российской академии
медицинских наук по адресу: 115478, Москва, ул. Москворечье, д.1.

Автореферат разослан «_____» _____ **2012 г.**

Учёный секретарь диссертационного совета Д 001.016.01
по защите докторских и кандидатских диссертаций,
доктор медицинских наук, профессор

Зинченко Рена Абульфазовна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. Генетическое изучение популяций человека на протяжении десятилетий привлекает внимание многих исследователей, среди которых виднейшие представители отечественной и мировой биологии: А.С. Серебровский, Н.К. Кольцов, В.В. Бунак, Г.Ф. Дебец, Ю.П. Алтухов, Ю.Г. Рычков, L.L. Cavalli-Sforza, W. Bodmer, а также представители смежных дисциплин – археологии, лингвистики, прикладной математики, палеогеографии. В последние двадцать лет популяционная генетика переживает новый подъем, связанный с использованием нерекombинирующих систем: митохондриальной ДНК (мтДНК) и Y-хромосомы. Ярко выраженная межпопуляционная вариабельность этих маркеров, высокая скорость накопления мутаций, возможность проследить их постепенное «наслоение», не разбиваемое рекомбинационными событиями – все это сделало мтДНК и Y-хромосому эффективными инструментами для прослеживания древних миграций человечества, изучения генофонда современных популяций и факторов их микроэволюции.

Эти нерекombинирующие генетические системы используются преимущественно для реконструкции истории популяций, тогда как аутосомные маркеры чаще исследуются в связи с задачами геномной медицины (Пузырев и др., 2000, 2004; Спицын и др., 2000, 2006; Степанов, 2003; Янковский и др., 2005; Марусин и др., 2006, 2007; Воевода и др., 2006; Vorinskaya et al., 2009; Stepanov et al., 2010a; Боровкова и др., 2010; Баранов, 2011) и ДНК-идентификации (Степанов и др., 2003а,б; Шорохова и др., 2005; Verbenko et al., 2006; Деренко и др., 2007; Zhivotovsky et al., 2007, 2009a,b; Stepanov et al., 2010b). В самые последние годы набирает силу и анализ панелей из сотен тысяч SNP маркеров (главным образом, аутосомных). Достигая точного определения генетического сходства популяций, такой анализ из-за высокой стоимости генотипирования проводится лишь для немногих популяций и небольших (10-20 человек) выборок, что может смещать оценки. Изучение же мтДНК и Y-хромосомы проводится на обширных выборках десятками известных лабораторий для почти всех народов мира со всех континентов.

Особенно подробно изучен генофонд Европы, ставший для европейской науки полигоном для проверки генетических концепций и методов. Вопрос о палеолитическом или неолитическом времени формирования основных черт европейского генофонда является, пожалуй, наиболее широко обсуждаемым вопросом в популяционной генетике человека. Об этом свидетельствует число статей на эту тему (в т.ч. в журналах Nature и Science) и авторитет их авторов (Ammerman, Cavalli-Sforza, 1987; Cavalli-Sforza et al., 1994; Richards et al., 1996; Richards et al., 2000; Semino et al., 2000; Barbujani, Bertorelle, 2002; Chikhi et al., 2002; Haak et al., 2005; Bramanti et al., 2009 и т.д.). Однако работы, обобщающие изменчивость мтДНК и Y-хромосомы в Европе, последний раз проводились десятилетие назад (Richards et al., 2000; Rosser et al., 2000; Semino et al., 2000; Richards et al., 2002), когда впервые появилась возможность получить целостное представление о генофонде Европы по однородительским ДНК маркерам. Подавляющее большинство последовавших крупных работ и по мтДНК (Helgason et al., 2001, 2003; Meinila et al., 2001; Malyarchuk et al., 2002, 2003, 2004, 2006, 2008; 2011; Бермишева и др., 2002; Orekhov et al., 1999; Pfeiffer et al., 1999; Pereira et al., 2004; Tambets et al., 2004; Goodacre et al., 2005; Falchi et al., 2006; Grzybowski et al., 2007; Lappalainen et al., 2008; Alvarez-Iglesias et al., 2009; Santos et al., 2009; Garcia et al., 2010; Karachanak et al., 2011) и по Y-хромосоме (Behar et al., 2003; Харьков и др., 2004, 2005; Cinnioglu et al., 2003; Di Giacomo et al., 2003; Brion et al., 2005; Flores et al., 2003, 2004; Tambets et al., 2004; Alonso et al., 2005; Goncalves et al., 2005; Kayser et al., 2005; Pericic et al., 2005; Capelli et al., 2006, 2007; Lappalainen et al., 2006, 2008; Adams et al., 2008; Balanovsky et al., 2008, 2011;

Battaglia et al., 2008; Fechner et al., 2008; Varzari et al., 2009; King et al., 2011; Yunusbaev et al., 2011; многие другие работы) ставило задачей изучение отдельных регионов Европы. Ряд работ был посвящен также отдельным гаплогруппам (Di Giacomo et al., 2004; Cruciani et al., 2007, 2010; Myres et al., 2010; Underhill et al., 2010; Mendez et al., 2011; Onofri et al., 2008; Derenko et al., 2006, 2007, 2010; Rootsi et al., 2007; Tofanelli et al., 2009 и другие работы). Но обобщающий анализ в масштабе всей Европы после работ 2000 года не проводился.

О популяциях Северной Азии также накоплены подробные данные и по маркерам Y-хромосомы (Karafet et al., 1997, 2001; Степанов, Пузырев, 2000а,б,в; Степанов и др., 2001; Lell et al., 2002; Пузырев и др., 2003; Tambets et al., 2004; Харьков и др., 2005, 2007а,б, 2008, 2009, 2011; Stepanov, Kharkov, 2006; Derenko et al., 2006; Lessig et al., 2008; Dulik et al., 2011), и по мтДНК (Schurr et al., 1999; Derenko et al., 2000, 2001, 2003, 2007; Деренко и др., 2002, 2004; Pakendorf et al., 2003, 2005; Fedorova et al., 2003; Zakharov et al., 2004; Tajima et al., 2004; Бермишева и др., 2005; Starikovskaya et al., 2005; Zlojutro et al., 2006; Volodko et al., 2008; Gokcumen et al., 2008; Rubicz et al., 2010). Эти данные и опыт их обобщения (Степанов, 2002; Степанов и др., 2006; Derenko et al., 2006; Деренко, 2009; Харьков, 2012) позволяют изучить взаимодействие европейского и азиатского генофондов, а обширные данные по другим регионам мира делают возможным анализ и в глобальном масштабе. Поэтому представляется своевременным и актуальным данное исследование, в котором сделана попытка на современном этапе изучения генофонда собрать воедино, проанализировать и подытожить как многочисленные собственные данные, так и составленные под руководством автора базы данных, включающие опубликованные данные не только о генофондах народов Европы и Азии, но и о мировой изменчивости мтДНК и Y-хромосомы.

Создание таких баз данных актуально и само по себе: их отсутствие чрезвычайно затрудняет сравнительный анализ генофондов. Поэтому созданные базы данных заслужили признание научного сообщества (Behar et al., 2007, 2009; Quintano-Murci et al., 2008, 2010; Naak et al., 2010; Haber et al., 2011) и позволили провести анализ в географическом масштабе не только Европы, но и Евразии.

Учитывая сложность реконструкции древних миграций по данным о современных популяциях, большое значение приобретает анализ ДНК палеоантропологического материала из археологических раскопок: исследование палео ДНК позволяет проследить изменчивость генофонда не только в пространстве, но и во времени. Поэтому актуален анализ древней ДНК, который реализован в данной работе для пяти временных срезов (на протяжении 7000 лет): эпох мезолита, неолита, раннего металла, античности и Нового времени. Генетические исследования приобретают все большее значение для смежных наук о человеке - не только физической антропологии и археологии, но и лингвистики. Поэтому актуален параллельный анализ лингвистических и генетических данных, выполненный для двух наиболее крупных лингвистических семей Европы - индоевропейской и северокавказской. И, несомненно, особую актуальность для российской науки имеет исследование русского генофонда, восточнославянских и северокавказских народов, составившее основной каркас работы.

Цель исследования. Изучить пространственную изменчивость и этапы формирования генофондов на примере народонаселения Европы, проанализировать геногеографические закономерности в масштабе Евразии и сравнить роль географического соседства и лингвистического родства в структурировании генофондов, основываясь на данных о Y-хромосоме и митохондриальной ДНК.

Задачи исследования.

1. Создать унифицированные базы данных по изменчивости Y-хромосомы и митохондриальной ДНК в народонаселении мира, объединяющие результаты собственных исследований с литературными данными.
2. Изучить пространственную изменчивость Y-хромосомы и митохондриальной ДНК на трех иерархических уровнях: в популяциях восточных славян, Европы, Евразии.
3. Провести картографический анализ митохондриального генофонда Европы и выявить факторы, определяющие пространственную изменчивость внутрипопуляционного разнообразия.
4. Определить соотношение межпопуляционной изменчивости у европейских народов, говорящих на языках различных лингвистических групп.
5. Создать картографический атлас распространения гаплогрупп Y-хромосомы и митохондриальной ДНК в населении Евразии; выявить объективное деление на региональные генофонды; определить степень влияния центральноазиатских популяций на генофонд восточных славян и Европы в целом.
6. На примере подразделенного генофонда народов Кавказа реконструировать древо родства популяций по SNP-гаплогруппам и по STR-гаплотипам Y-хромосомы и сопоставить генетические, лингвистические и исторические датировки разделения родственных народов.
7. Оценить роль лингвистического и географического фактора в структурировании генофонда Европы по Y-хромосоме (на примере северокавказской семьи) и по митохондриальной ДНК (на примере индоевропейской языковой семьи).
8. Исследовать изменчивость европейского генофонда во времени с использованием данных по древней ДНК эпох мезолита, неолита, раннего металла, античности и Нового времени.

Научная новизна.

Впервые созданы обширные базы данных по изменчивости Y-хромосомы (более 112 тысяч образцов) и мтДНК (более 132 тысяч образцов) в коренных популяциях мира.

Впервые по маркерам Y-хромосомы выявлена структурированность европейских популяций по этническому признаку: формирование четких «этнических облаков».

Впервые выявлена ведущая роль лингвистического фактора в структурировании митохондриального генофонда Европы: кластеризация по лингвистическим группам.

Впервые показано, что финно-угорские популяции представляют основную часть межпопуляционного разнообразия Европы, в то время как их внутрипопуляционное разнообразие минимально; выдвинута гипотеза о роли дрейфа генов в формировании такого соотношения внутри- и межпопуляционного разнообразия.

Впервые обнаружен широтный тренд убывания к северу гаплотипического разнообразия мтДНК населения Европы и его связь с грузом наследственной патологии.

Впервые созданы атласы изменчивости генофонда Евразии по широкому спектру гаплогрупп Y-хромосомы и мтДНК, благодаря которым выявлены основные генетические границы в ареале Евразии.

Впервые показана двусоставность русского генофонда по маркерам Y-хромосомы, формирующая широтный тренд и связываемая с восточнославянским и палеоевропейским компонентами, на основе которых сформировался русский генофонд.

Впервые по маркерам Y-хромосомы и мтДНК доказана незначительность влияния миграций из Центральной Азии на генофонд славян и других народов Европы.

Впервые выявлен параллелизм в генетическом и лингвистическом разнообразии на Кавказе и близость генетических и лингвистических датировок популяционных событий. Такой анализ впервые основан на количественных лингвистических расстояниях.

Впервые по данным о древней ДНК показано ближневосточное происхождение первого неолитического населения Европы.

Впервые по генетическим данным подтверждена значительная доля монголоидного (или «уралоидного») компонента в генофонде севера Восточной Европы эпохи мезолита.

Впервые доказано отсутствие преемственности между населением эпохи раннего металла и современным коренным населением Кольского полуострова.

Практическая значимость.

Полученные результаты расширяют и систематизируют знания о генофонде восточных славян, народов Северного Кавказа, Сибири и Евразии в целом. Для медико-генетических исследований имеют значение обнаруженные тренды в гаплотипической изменчивости мтДНК, поскольку зоны низкого разнообразия скоррелированы с отягощенностью аутосомно-рецессивной патологией. Выявленные генетические границы в генофонде Евразии также позволяют прогнозировать пределы распространения редких наследственных нозологий. Данные о генотипах индивидуальных образцов позволяют оптимизировать формирование контрольной выборки, идентичной выборке пациентов.

Для судебно-медицинской экспертизы полученные обширные данные об изменчивости мтДНК и Y-хромосомы важны для создания референсных баз данных и установления вероятного региона происхождения неизвестного лица по образцу ДНК.

Данное исследование интенсифицировало сотрудничество российских генетиков с российскими и зарубежными лингвистами (Balanovsky et al., 2011), археологами (Haak et al., 2010), палеоклиматологами (Балановский, 2008), палеоантропологами (Der Sarkisyan, in preparation), представителями других гуманитарных и естественных наук. Результаты использовались в лекциях студентам, аспирантам, учителям, школьникам, многократно описывались на телевидении, радио и в печати для популяризации научных знаний.

Результаты исследования используются в работе российских и зарубежных организаций: ФГБУ «Медико-генетический научный центр» РАМН, ФГБУН Институт общей генетики им. Н.И. Вавилова РАН, ФГБУН Институт молекулярной биологии РАН, ФГБУН Институт географии РАН, Институт антропологии МГУ им. М.В. Ломоносова, Институт языкознания РАН, Российский государственный гуманитарный университет, Адыгейский, Белгородский, Забайкальский, Казанский, Кемеровский, Мордовский, Харьковский, Забайкальский государственные университеты, Следственный комитет РФ, Российский центр судебно-медицинской экспертизы Минздрава РФ, Эстонский биоцентр, Национальное географическое общество США, Институт Пастера (Франция), Австралийский центр древней ДНК, Центр здравоохранения Рамбам (Израиль), Ливанский Американский университет, Университет Мадурая (Индия), Университет Помпея Фабра (Испания), Евразийский национальный университет им. Л.Н. Гумилева (Казахстан), НИИ молекулярной биологии и медицины при Минздраве Киргизии, Институт молекулярной биологии Национальной академии наук Армении, Центр судебной экспертизы Министерства юстиции Казахстана, Институт искусствоведения, этнографии и фольклора НАН Беларуси, Монгольская академия медицинских наук.

Основные положения, выносимые на защиту.

1. Изменчивость Y-хромосомы в европейских популяциях в высокой степени структурирована, причем не только по географическому, но и по этническому принципу: популяции одного народа, во-первых, обладают значительным сходством

друг с другом и, во-вторых, отдалены от генофондов других народов. Важны нарушения этой закономерности: исключение из первого правила - выраженное различие между северными и южными русскими популяциями, исключение из второго правила - значительное «перекрывание» славянских генофондов.

2. В генофонде народонаселения Восточной Европы по маркерам и Y-хромосомы, и митохондриальной ДНК не обнаруживается сколь либо существенное влияние миграций центральноазиатских кочевников, о чем свидетельствуют картографические атласы, основанные на созданных обширных базах данных по распространению гаплогрупп Y-хромосомы и митохондриальной ДНК у народов мира.
3. Для изменчивости митохондриальной ДНК в населении Европы характерны три закономерности. Во-первых, уменьшение внутривидового разнообразия к северу, что объясняется усилением дрейфа генов в северных популяциях вследствие снижения продуктивности территорий и падения плотности населения. Во-вторых, те же факторы стали причиной повышения межпопуляционного разнообразия мтДНК у финно-угорских популяций северо-востока Европы. В-третьих, наряду с этими географическими закономерностями, имеется связь структуры митохондриального генофонда с лингвистической классификацией: генофонды популяций кластеризуются в соответствии с их лингвистической принадлежностью.
4. Параллелизм между генетической и лингвистической структурой народонаселения ярко проявляется в четко подразделенных популяциях коренных народов Северного Кавказа. Каждой лингвистической ветви соответствует своя доминирующая гаплогруппа Y-хромосомы, а многие кластеры гаплотипов в пределах гаплогрупп высоко (в среднем 95%) специфичны для отдельных народов. Генетические датировки возраста кластеров хорошо согласуются с лингвистическими датировками разделения соответствующих языков и с историческими данными.
5. Созданные картографические атласы, основанные на совокупности собственных результатов и разработанных автором баз данных, объективно описывают изменчивость Y-хромосомы и митохондриальной ДНК в Евразии. Выявлена основная генетическая граница, разделяющая Евразию на западный и восточный субгенофонды: граница начинается на Кавказе, проходит через южный Урал, северный Казахстан, Южную Сибирь и далее следует вдоль Енисея по Средней Сибири.
6. О ближневосточном происхождении первых неолитических популяций Европы (культура линейно-ленточной керамики) свидетельствует сравнение данных по древней ДНК с созданной базой данных о современном генофонде. В предшествующий период мезолита восточноевропейское население (данные по Южному Оленьему острову) обладало большим, чем сейчас, сходством с населением Западной Сибири; часть последующих миграций из Сибири угасла, не оставив следов в современном генофонде (данные по Большому Оленьему острову).

Апробация работы.

Основные результаты были доложены на International Congress of the European Anthropological association "Anthropology and Society" (Чехия, 2003); Третьих Антропологических чтениях памяти академика В.П. Алексеева "Экология и демография человека в прошлом и настоящем" (Москва, 2004); Второй международной (совместно с посольством Великобритании) школе-семинаре "Проблемы генетической безопасности: научные инновации и их интерпретация" (Москва, 2004); III Съезде Вавиловского общества генетиков и селекционеров (Москва, 2004); семинаре группы член-корр. РАН С.А. Старостина в Российском государственном гуманитарном университете (Москва,

2004); семинаре отдела эволюционной биологии Эстонского биоцентра (Эстония, 2005); конференции "Актуальные вопросы сравнительно-исторического языкознания и дальнейшее родство языков" (Москва, 2006); конференции "Путь на север: инициальное заселение человеком Арктики и Субарктики" (Москва, 2007); 5th ISABS Conference in Forensic Genetics and Molecular Anthropology (Хорватия, 2007); семинаре экспертов-биологов ЭКЦ МВД, ГУВД, УВД субъектов РФ "Использование современных технологий и оборудования при производстве экспертиз тканей и выделений человека" (Красноярск, 2008); Всероссийской конференции "Адаптация как фактор формирования антропологического своеобразия древнего и современного населения Евразии" памяти акад. Т.И. Алексеевой (Москва, 2008); V Съезде Вавиловского общества генетиков и селекционеров (Москва, 2009); Международной конференции "Человек: его биологическая и социальная история" (Москва, 2009); конференции молодых ученых ФГБУ «МГНЦ» РАМН (Москва, 2009); 3rd workshop "Foundations of Europe - Prehistoric roots of Europe" (Испания, 2009); VI Съезде Российского общества медицинских генетиков (Ростов-на-Дону, 2010); Всероссийской научной школе для молодежи "Горизонты нанобиотехнологии" (Звенигород, 2009); VIII Курчатовской молодежной научной школе (Москва, 2010); 5й научной школе молодых ученых по экологической генетике "Экологическая генетика человека" (С-Петербург, 2010); XVII Савёловских чтениях "Генеалогия и генетика" (Москва, 2010); семинаре Государственного Дарвиновского музея "Эволюция человека сегодня"; Международной конференции «Проблемы популяционной и общей генетики» (Москва, 2011); на шести ежегодных международных конференциях The Genographic Project (США, 2005; ЮАР, 2006; Китай, 2007; Эстония, 2008; Австралия, 2010; США, 2011).

Личный вклад автора. Автор участвовал в планировании и проведении экспедиций по обследованию населения Восточной Европы, Кавказа, Южной Сибири, Казахстана, Киргизии, Монголии, Таджикистана, руководил шестью экспедиционными грантами РФФИ, лично генотипировал маркеры мтДНК и Y-хромосомы в восточнославянских популяциях, руководил генотипированием Y-хромосомы в населении Кавказа, Сибири, Центральной Азии, руководил созданием баз данных и программного обеспечения геногеографического анализа, лично разрабатывал технологии картографического анализа генофонда и создавал картографические атласы изменчивости мтДНК и Y-хромосомы, лично проводил весь статистический, картографический, филогенетический анализ, выполнил интерпретацию и описание результатов, руководил кандидатскими диссертациями, в исследованиях древней ДНК участвовал в экспериментальной работе, выполнил статистический анализ и интерпретацию результатов. Большинство публикаций, приведенных в списке работ по теме диссертации, написаны лично автором. Суммарно личный вклад автора составляет более 80%.

Публикации: Основные результаты исследования опубликованы в 37 научных работах, в том числе 29 статей в ведущих рецензируемых научных журналах, рекомендованных ВАК РФ для защиты диссертаций, а также в монографии.

Внедрение результатов работы. Результаты диссертационного исследования внедрены в научно-педагогическую практику Биологического факультета Московского государственного университета им. М.В. Ломоносова (Москва, Ленинские горы, д.1, стр. 12), Харьковского национального университета им. В.Н. Каразина (г. Харьков, пл. Свободы, 4), ФГБОУ ВПО «Кемеровский государственный университет» (г. Кемерово, ул. Красная, д.6), ФГБОУ ВПО «Адыгейский государственный университет» (г. Майкоп,

ул. Первомайская, д. 208), ФГБОУ ВПО «Амурский гуманитарно-педагогический государственный университет» (г. Комсомольск-на-Амуре, ул. Кирова, д.17, корп. 2).

Структура и объем работы. Работа изложена на 357 страницах и содержит 6 глав, введение, выводы, список литературы (348 источников, из них 215 зарубежных). Работа иллюстрирована 30 таблицами и 99 рисунками (включая карты).

МАТЕРИАЛЫ И МЕТОДЫ

Материалом исследования послужили многочисленные собственные данные по изменчивости Y-хромосомы (табл. 1), мтДНК (табл. 2) и созданные под руководством автора базы данных на основе опубликованных источников (табл. 3). Информация по древней ДНК получена совместно с Австралийским центром древней ДНК (табл. 4).

СОБСТВЕННЫЕ ДАННЫЕ. Перечень популяций, по которым получены собственные данные, представлен в таблицах 1 и 2. Биологический материал (образцы венозной крови) собран в ходе многолетних экспедиций (1998-2010 гг.), проведенных под руководством Е.В. Балановской и автора данной работы коллективом лаборатории популяционной генетики человека ФГБУ «МГНЦ» РАМН в сотрудничестве с Адыгейским, Белгородским, Кемеровским государственными университетами, научными учреждениями Армении, Белоруссии, Казахстана, Киргизии, Монголии, Таджикистана и Украины в рамках грантов РФФИ, РГНФ и международного проекта Genographic. Во всех экспедициях обследовались только индивиды, которые согласно родословным на протяжении трех поколений (пробанд, его родители и бабушки-дедушки) принадлежали данной популяции (в пределах административных границ данного и соседних районов) и данной этнической группе. Родственники до третьей степени родства тщательно исключались. Все обследованные дали письменное информированное согласие под контролем Этической комиссии ФГБУ «МГНЦ» РАМН.

По Y-хромосоме изучены популяции русские (N=1353), кавказские (N=1695), Сибири и Центральной Азии (N=1241); выборка превысила 4000 образцов (табл. 1).

По митохондриальной ДНК охарактеризованы русские, украинские, белорусские и северокавказские популяции, суммарно более 3000 образцов (табл. 2).

Древняя ДНК эпох мезолита, неолита, раннего металла, античности и Нового времени из разных регионов Европы (табл. 3) генотипирована по маркерам мтДНК в Австралийском центре древней ДНК, в том числе при участии автора.

БАЗЫ ДАННЫХ. Вторая часть материалов – базы данных по мтДНК (MURKA) и по Y-хромосоме (Y-base) о населении мира на основе опубликованных источников, созданные под руководством автора.

В течение последних 20 лет сошла целая лавина статей, несущих сведения по изменчивости мтДНК в популяциях мира. Наша база данных («**MURKA: база данных по митохондриальной ДНК и интегрированное программное обеспечение**»), основные составители: О.П. Балановский, В.В. Запороженко, А.С. Пшеничнов) росла также лавинообразно. В 2004 году в базе имелась информация о 33 000 образцов мтДНК; в 2007 году – 62 000; в 2008 году – 95 000; в 2010 году – 135 000. На текущий момент база содержит информацию о 206 000 образцах. Из них 8 870 - полные мтДНК геномы, 15 000 - данные по отдельным гаплогруппам, 50 000 – данные общественных участников проекта «Генография» (нет привязки к конкретным популяциям) и 132 600 образцов – обычные популяционные данные о коренном населении, которые и использовались в исследовании. Важной чертой базы является подробная характеристика популяций.

Таблица 1. Популяции, изученные по маркерам Y-хромосомы

Популяция*	Народ	Географическая локализация	N	Долгота	Широта
Пинега	русские	Архангельская обл., Пинежский р-н	114	46.53	63.43
Мезень	русские	Архангельская обл., Лешуконский р-н	54	45.74	64.9
Красноборск	русские	Архангельская обл., Красноборский и Ленский р-ны	91	45.94	61.56
Вологда	русские	Вологодская обл, разные р-ны	121	39.9	59.23
Кашин	русские	Тверская обл., Кашинский р-н	73	37.61	57.36
Порхов	русские	Псковская обл., Порховский р-н	57	29.56	57.77
Остров	русские	Псковская обл., Островский р-н	75	28.32	57.35
Смоленск	русские	Смоленская обл., Рославльский р-н	107	32.88	53.93
Кострома (Унжа)	русские	Костромская обл., Мантуровский и Межевской р-ны	52	44.77	58.33
Русские Белгород	русские	Белгородская обл., Яковлевский, Красненский, Прохоровский р-ны	143	36.48	50.78
Ливны	русские	Орловская обл., Ливненский р-н	110	37.59	52.4
Пристенъ	русские	Курская обл., Пристенский р-н	45	36.71	51.23
Репьевка	русские	Воронежская обл., Репьевский р-н	96	38.65	51.08
Кубанские казаки	русские	Адыгея, Майкопский р-н	90	40.17	44.51
Терские казаки	русские	Кабардино-Балкария, Майский, Прохладненский р-ны	125	44.0	43.7
Шапсуги	шапсуги	Краснодарский край, Туапсинский, Лазаревский р-ны	100	39.1	44.2
Абхазы	абхазы	Абхазия, разные р-ны	58	41.10	43.1
Черкесы	черкесы	Карачаево-Черкесия, Хабезский, Прикубанский р-ны	142	41.7	43.8
Осетины	осетины	Северная Осетия, Дигорский, Ирафский, Алагирский, Правобережный, Пригородный р-ны; Южная Осетия.	359	44.0	42.8
Ингуши	ингуши	Ингушетия, Малгобекский р-н	143	45.0	43.1
Чеченцы	чеченцы	Ингушетия, Малгобекский р-н; Чечня, Ачхой-Мартановский р-н; Дагестан, Казбековский, Хасавюртовский, Новолакский р-ны	330	46.0	43.2
Аварцы	аварцы	Дагестан, Унцукульский, Гунибский р-ны	115	46.9	42.5
Даргинцы	даргинцы	Дагестан, Акушинский, Дахадаевский р-ны	101	47.2	42.2
Кубачинцы	кубачинцы	Дагестан, Дахадаевский р-н	65	47.6	42.1
Кайтагцы	кайтагцы	Дагестан, Кайтагский р-н	33	47.63	42.2
Лезгины	лезгины	Дагестан, Ахтынский р-н	81	47.70	41.5
Армяне	армяне	Краснодарский край и Адыгея, разные р-ны	115	40.1	44.1
Армяне-амшены	армяне	Краснодарский край (мигранты из Турции)	53	-	-
Кумандинцы, челканцы, тубалары	северные алтайцы	Алтайский край, Солтонский р-н Алтай, Турочакский и Чойский р-ны	202	52.2	87.42
Алтай-кижи, теленгиты	южные алтайцы	Алтай, Онгудайский, Улаганский, Кош-Агачский, Чойский, Турочакский р-ны	207	50.03	88.5
Шорцы (горные и абаканские)	шорцы	Кемеровская обл., Таштагольский р-н Хакасия, Таштыпский р-н	139	52.75	88
Качинцы, сагайцы, койбалы, кызыльцы	хакасы	Хакасия, Аскизский, Бейский, Орджоникидзевский, Таштыпский, Ширинский р-ны	185	53	90
Казахи Алтая	казахи	Казахстан, Восточно-Казахстанская обл., Катон-Карагайский р-н.	122	49.16	85.6
Киргизы северные	киргизы	Киргизия, Нарынский, Таласский р-ны	140	41.42	76
Киргизы Памира	киргизы	Таджикистан,, Мургабский р-н	106	38.15	73.95
Монголы Алтая	монголы	Монголия, аймаки Байан-Улгий, Ховд, Гоби-Алтай	140	48	96
Суммарно	20 этносов	6 стран, 39 областей, 81 район	4289	-	-

Примечание: * дано условное название популяции, далее использованное в таблицах и тексте.

Таблица 2. Популяции, изученные по маркерам митохондриальной ДНК

Популяция*	Народ	Географическая локализация	№	Долгота	Широта
Пинега	русские	Архангельская обл., Пинежский р-н	144	46.53	63.43
Смоленск	русские	Смоленская обл., Рославльский р-н	147	32.88	53.93
Кострома (Унжа)	русские	Костромская обл., Мантуровский р-н	79	44.77	58.33
Белгород	русские	Белгородская обл., три р-на	148	36.48	50.78
Боровский	русские	Калужская, Боровский р-н	70	36.5	55
Барятинский	русские	Калужская обл., Барятинский р-н	75	34.53	54
Болховский	русские	Орловская обл., Болховский р-н	76	36	53.43
Черемисиновский	русские	Курская обл., Черемисиновский р-н	62	37.25	51.87
Михайловский	русские	Рязанская обл., Михайловский р-н	82	39	54.22
Спасск-Рязанский	русские	Рязанская обл., Спасск-Рязанский р-н	86	40.53	54.4
Петровский	русские	Тамбовская обл., Петровский р-н	76	40.25	52.63
Кубанские казаки	русские	Адыгея, Майкопский р-н	132	40.17	44.51
Терские казаки	русские	Кабардино-Балкария, два р-на	124	44.0	43.7
Столин	белорусы	Брестская обл., Столинский р-н	89	26.77	51.97
Светлогорск	белорусы	Гомельская обл., Светлогорский р-н	71	29.72	52.62
Витебск	белорусы	Витебская обл.	100	29	55
Брест	белорусы	Брестская обл.	104	24	52
Гомель	белорусы	Гомельская обл.	121	30	52
Украинцы Белгород	украинцы	Белгородская обл., два р-на	95	35.75	50.4
Черкассы	украинцы	Черкасская обл.	179	32.07	49.43
Хмельницкая	украинцы	Хмельницкая обл., Старо-Конст. р-н	179	27.3	49.7
Украинцы западные	украинцы	Львовская и Ивано-Франковская обл.	157	24	49.25
Лезгины	лезгины	Дагестан, Ахтынский р-н	69	47.70	41.5
Кайтагцы	кайтагцы	Дагестан, Кайтагский р-н	32	47.63	42.2
Кубачинцы	кубачинцы	Дагестан, Дахадаевский р-н	55	47.6	42.1
Аварцы	аварцы	Дагестан, Унцукульский, Гунибский	81	46.9	42.5
Даргинцы	даргинцы	Дагестан, Акушинский, Дахадаевский	95	47.2	42.2
Чеченцы	чеченцы	Чечня, Дагестан, Ингушетия	239	46.0	43.2
Осетины	осетины	Северная Осетия, пять р-нов	267	44.0	42.8
Суммарно	10 этносов	23 области, 40 районов	3234	-	-

Примечание: * дано условное название популяции, далее использованное в таблицах и тексте.

Таблица 3. Созданные базы данных по Y-хромосоме и мтДНК

База данных	Число образцов	Число популяций	Число источников	Размещение	Число образцов в зарубежных аналогах
Y-хромосома <i>Y-base</i>	112 400	2 474	238	www.genofond.ru	97 575 www.yhrd.org
мтДНК <i>MURKA</i>	132 600	2 101	497	Локальная сеть ЛПГЧ ФГБУ «МГНЦ» РАМН	17 321 www.empop.org

Для каждой популяции указывается название, объем выборки, административная принадлежность (страна, провинция), этническая и лингвистическая принадлежность, географические координаты, ссылки на источник данных и т.д. База содержит как частоты гаплогрупп, так и сиквенсы (ГВС1, ГВС2 или же полные сиквенсы мтДНК) с

указанием начальной и конечной позиции секвенирования. Ряд характеристик БД в сравнении с ведущими зарубежными аналогами приведен в табл. 3.

Массовое изучение изменчивости Y-хромосомы началось позже, чем мтДНК, но благодаря большей межпопуляционной изменчивости Y-хромосома привлекает еще большее внимание исследователей (табл. 3). В базе данных «**Y-base: изменчивость Y-хромосомы у народов мира**» (основные составители: О.П. Балановский, А.С. Пшеничнов, Р.С. Сычев) большое внимание уделено формализованному описанию популяций для обеспечения эффективного поиска. Главной проблемой при создании БД были различия в наборах SNP-маркеров, используемых разными авторами и большое число SNP-маркеров, определяющих более 400 гаплогрупп. Эта проблема решена путем реализации триады: 1) формализация представления исходных данных (дефенирующий SNP-маркер, а не часто меняющееся название гаплогруппы); 2) создание формализованного дерева гаплогрупп Y-хромосомы; 3) программная реализация алгоритма, определяющего для каждого образца четкий «статус» каждой гаплогруппы – либо принадлежность к гаплогруппе, либо исключение из гаплогруппы, либо невозможность определения гаплогруппы по опубликованным экспериментальным данным. При усреднении этих данных по всем образцам, входящим в популяционную выборку, для каждой популяции автоматически вычислялась, соответственно, либо частота гаплогруппы, либо проставлялась нулевая частота гаплогруппы, либо указывалось, что частота данной гаплогруппы в популяции неизвестна.

Созданные базы данных превосходят зарубежные аналоги как подробностью описания популяций, так и объемом данных (табл. 3). Базы данных не только послужили одним из столпов, на которых выстроено данное исследование, но уже используются для решения широкого круга научных и практических задач. Например, с их помощью верифицирован алгоритм автоматического определения гаплогрупп (Behar et al., 2007), реконструированы асимметричные потоки генов в популяцию цветного населения ЮАР (Quintana-Murci et al., 2008), выявлено древнее родство африканских пигмеев и бантуязычных народов (Quintana-Murci et al., 2010), изучена ассоциация сердечно-сосудистой патологии с гаплогруппами мтДНК (Haber et al., 2011).

Таблица 4. Изученные популяции древней ДНК

Популяция	Эпоха	Датировка	Локализация	N
Южный Олений остров	мезолит	7.5 тыс. лет назад	Карелия	9
Деренбург	неолит	7.0 – 7.5 тыс. лет назад	Германия	21
Большой Олений остров	эпоха раннего металла	3.5 тыс. лет назад	Мурманская обл.	23
Скифы Дона	античное время	2.5 тыс. лет назад (VI-III вв. до н.э.)	Ростовская обл.	16
Чальмны-Варрэ (саамы)	Новое время	0.2 тыс. лет назад (XVIII век н.э.)	Мурманская обл.	42

МЕТОДЫ ГЕНОТИПИРОВАНИЯ

Выделение ДНК из образцов венозной крови проводилось фенол-хлороформным методом. Концентрация ДНК определялась спектрофотометрически и/или в ходе ПЦР в реальном времени с помощью набора Quantifiler Human DNA Kit (Applied Biosystems).

Генотипирование Y-хромосомы проведено с использованием 40 SNP маркеров. В перечне рядом с каждой гаплогруппой представлен детерминирующий ее SNP маркер:

E(M96), E1b1b1(M35), F(M89), G1(M285), G(M201), G2a(P15), G2a1(P16), G2a1a(P18), G2a3b1(P303), G2a3a(M406), I(M170), I1(M253), I2a(P37.2), J(M304), J1(M267), J1e(P58), J2(M172), J2a4b(M67), J2b(M12), J2a4a(M47), J2a4b1(M92), K(M9), L(M20), L1(M27), L2(M317), L3(M357), N1(LLY22), N1c1(M178), O(M175), O3(M122), P(M45), Q(M70), R(M207), R1(M173), R1a(SRY1532.2), R1a1a(M198), R1a1a7(M458), R1b(M343), R1b1b2(M269), T(M70). Генотипирование русских популяций проведено для части маркеров методом ПЦР-ПДРФ, а для остальных - секвенированием, как описано в (Balanovsky et al., 2008). Генотипирование кавказских и сибирских популяций проведено методом TaqMan на приборе ABI 7900 (Applied Biosystems) как описано в (Balanovsky et al., 2011). Анализ 17 STR локусов Y-хромосомы (DYS389I, DYS389II, DYS390, DYS456, DYS19, DYS385a, DYS385b, DYS458, DYS437, DYS438, DYS448, GATA_H4, DYS391, DYS392, DYS393, DYS439, DYS635) проведен на секвенаторе Applied Biosystems 3130xl набором Y-filer (Applied Biosystems). Анализ русских популяций проведен на базе Эстонского биоцентра, а кавказских и сибирских – на базе ЛПГЧ ФГБУ «МГНЦ» РАМН.

Генотипирование мтДНК включило секвенирование ГВС1 (позиции 16024-16400) и ПДРФ-анализ (или TaqMan анализ) мутаций в кодирующем регионе, определяющих гаплогруппы H, V, HV, U, U4, K, R, T, J, J1, N, N1b, W, X, I, M, A, C, Z, D, D5. Секвенирование проводилось на капиллярных секвенаторах ABI 377, ABI 3130xl, Megabase 2000 с использованием секвенирующих наборов DYEnamic™ ET terminator или BigDye Terminator v3.1. Анализ украинских, белорусских и части русских популяций выполнен на базе Эстонского биоцентра, а южных русских групп и северокавказских популяций – на базе ЛПГЧ ФГБУ «МГНЦ» РАМН.

МЕТОДЫ АНАЛИЗА ДАННЫХ

Статистический анализ полиморфизма Y-хромосомы и мтДНК выполнен на двух уровнях организации вариантов ДНК - гаплогрупп и гаплотипов. На уровне частот гаплогрупп с помощью программы DJ genetic (Balanovsky et al., 2008) рассчитаны генетические расстояния (Nei, 1975). Матрицы расстояний были визуализированы методами многомерного шкалирования и кластерного анализа в программе Statistica 6.0 (StatSoft. Inc, 2001). С помощью той же программы проведен и анализ главных компонент. Генетическая гетерогенность этносов и лингвистических групп (различия по частотам гаплогрупп между локальными популяциями в пределах этносов или между этносами в пределах лингвистических групп) рассчитывалась как средние генетические расстояния между популяциями, методом G_{ST} (Nei, 1975) с помощью оригинальной программы и методом AMOVA с помощью программы Arlequin 3.11 (Schneider и др., 2000). Сравнительный анализ эффективности географической и лингвистической классификаций популяций для определения сходства их генофондов также проведен методом AMOVA. В той же программе (Arlequin 3.11) выполнен тест Мантеля для оценки общих и частных коэффициентов корреляций между матрицами генетических, географических и лингвистических расстояний. Необходимые для последних двух видов анализа географические расстояния рассчитаны по географическим координатам популяций (табл. 1, 2) в программе DistGeo (www.genofond.ru). Матрицы лингвистических расстояний между языками (для индоевропейской и северокавказской лингвистических семей), полученные по стословному списку М. Сводеша (Swadesh, 1955) с исключением заимствованных слов по методу С.А. Старостина (Старостин, 1989), были любезно предоставлены член-корр. РАН А.В. Дыбо и проф. О.А. Мудраком.

Анализ на уровне гаплотипов состоял в расчете гаплотипического разнообразия мтДНК (с использованием MS Excel 7.0) и в построении филогенетических сетей по STR-гаплотипам Y-хромосомы. Сети построены в программе Network 4.1.1.2 (Fluxus Technology Ltd., www.fluxus-engineering.com) на основе алгоритма reduced median и отредактированы в Network Publisher (Fluxus Engineering, Clare, U.K.). Возраст кластеров гаплотипов определялся с помощью показателя ρ , как описано в (Forster et al., 1996; Saillard et al., 2000). Для датировок возраста кластеров использовались две скорости мутирования STR маркеров: «эволюционная» скорость (Zhivotovsky et al., 2004) и «генеалогическая» скорость, полученная при прямых сравнениях пар отец-сын (Gusmao et al., 2005; Sánchez-Diz et al., 2008; Ge et al., 2009).

Картографический анализ проведен с помощью программы GeneGeo, разрабатываемой в течение ряда лет под руководством автора. Построены геногеографические карты пяти видов: 1) карты распространения отдельных гаплогрупп Y-хромосомы и мтДНК; 2) суммарные карты распространения гаплогрупп; 3) карты гаплотипического разнообразия; 4) карты генетических расстояний; 5) карты генетических границ (карты межпопуляционной изменчивости). Все карты созданы методом средневзвешенной интерполяции со степенью весовой функции, равной 3. Выявление генетических границ на географической карте проводилось также и другим методом, с помощью программы Barrier 2.2 (Manni and Guerard, 2004).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

1. ИЗМЕНЧИВОСТЬ Y-ХРОМОСОМЫ В ЕВРОПЕ

РАСПРОСТРАНЕНИЕ ГАПЛОГРУПП Y-ХРОМОСОМЫ. Собственные данные по частотам гаплогрупп Y-хромосомы в русских популяциях (Balanovsky et al., 2008) приведены в табл. 5, по кавказским популяциям (Balanovsky et al., 2011) – в табл. 6, по сибирским и центрально-азиатским популяциям (Балаганская и др., 2011) – в табл. 7. На территории Европы суммарно изучено 35 популяций (N=3048 человек). По этим данным охарактеризованы два модельных объекта - генофонд русского народа (крупнейший народ Европы по численности и занимаемому ареалу) и народов Северного Кавказа. Данные по другим народам Европы взяты из созданной базы данных Y-base. Суммарно проанализированы данные о 585 популяциях Европы (N=38 000 образцов). Отметим, что предыдущее исследование, преследовавшее ту же цель (Rosser et al., 2000), проводилось более 10 лет назад по суммарной выборке, в десять раз меньшей (N=3616).

Созданы подробные карты распространения каждой гаплогруппы Y-хромосомы. Обнаружены четкие тренды изменчивости: гаплогруппа **R1a** доминирует у балто-славянских популяций Восточной Европы, **N1c** на северо-востоке Европы, **N1b** на северном Урале, **I1** на севере Европы, **I2a** на Балканах, **R1b** почти во всей Западной Европе. В статье большого коллектива с участием автора (Myres et al., 2010) гаплогруппа **R1b-M269** была дополнительно подразделена на ряд субгаплогрупп, многие из которых специфичны для той или иной территории. Ряд гаплогрупп (**G2**, **J2**, **J1**, **E-M35**) преобладает на Ближнем Востоке и Анатолии, но они также нередки и в соседних регионах Южной Европы. Таким образом, основной чертой структурированности европейского генофонда по Y-хромосоме оказывается четкое выделение географических зон, в каждой из которых преобладает своя гаплогруппа. Столь четкую структуру генофонда Европы - «принцип пазлов» - иллюстрирует рис. 1.

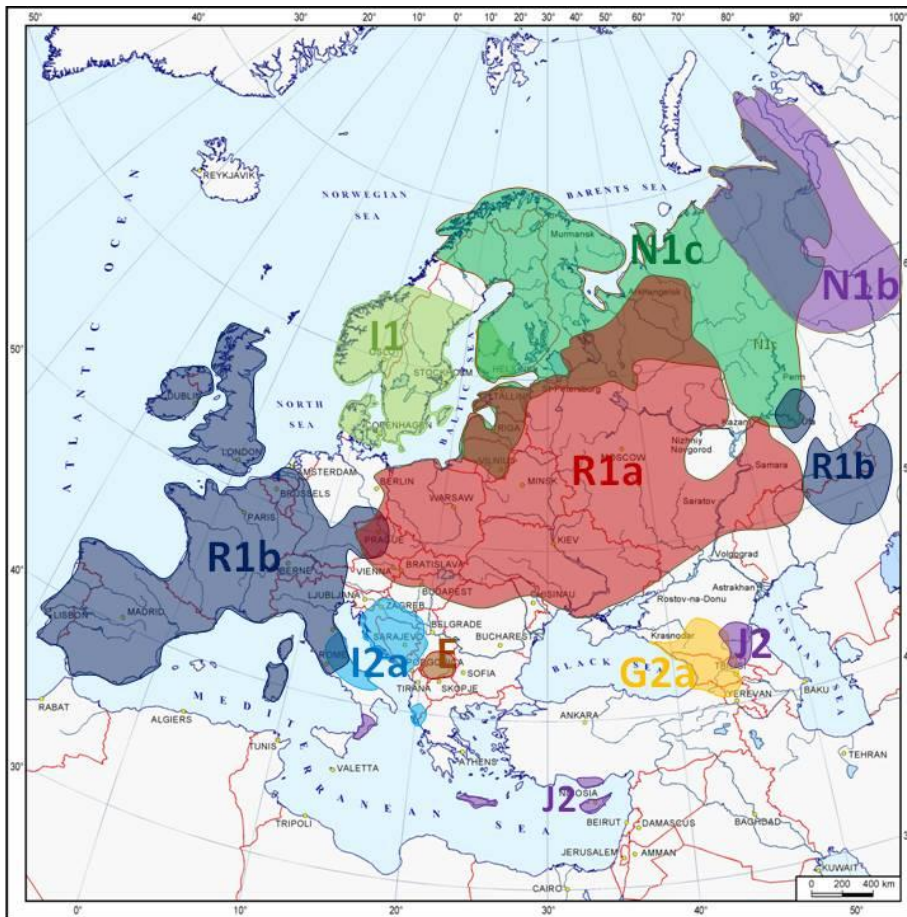


Рис.1. Структура генофонда Европы по Y-хромосоме: «принцип пазлов».

Выделены территории, где частота гаплогруппы составляет более трети генофонда (>35%). Этот ареал показан своим цветом для каждой гаплогруппы. В каждой географической части Европы доминирует одна гаплогруппа, редко встречающаяся в остальных частях.

ИЗМЕНЧИВОСТЬ НА РАЗНЫХ УРОВНЯХ ПОПУЛЯЦИОННОЙ ИЕРАРХИИ.

Обнаружение четких географических закономерностей, с одной стороны, развивает тенденции, показанные ранее (Rosser et al., 2000). С другой стороны, основной вывод этих авторов состоял в том, что географическое соседство, а не лингвистическое родство задает структуру генофонда. Результаты же нашего исследования, подтверждая важность географического фактора, указывают, что доминирует структурированность по лингвистическому принципу. Это следует из результатов многомерного шкалирования (рис. 2). В этот вид анализа включены те народы, для которых имелись данные по четырем или более региональным популяциям с объемом выборки не менее 70 человек. В данном случае этнический фактор совпадает с лингвистическим, поскольку этническая принадлежность популяции определялась по языку. График показывает, что во всех случаях популяции, относящиеся к одному народу, кластеризуются совместно, образуя «этнические облака», почти не перекрывающиеся друг с другом (рис. 2).

Два исключения касаются славянских групп. Во-первых, северные русские популяции не присоединяются к южным, а образуют отдельный кластер. Во-вторых, популяции поляков, южных русских и украинцев формируют не три отдельных, а одно единое «этническое облако», демонстрируя общность генофонда. Таким образом, славяноязычные популяции, занимающие половину территории Европы, оказываются генетически не различными (как должно бы быть в случае доминирования в структуре генофонда географического фактора), а сходными, в согласии с лингвистическими данными о близком родстве восточнославянских и западнославянских языков.

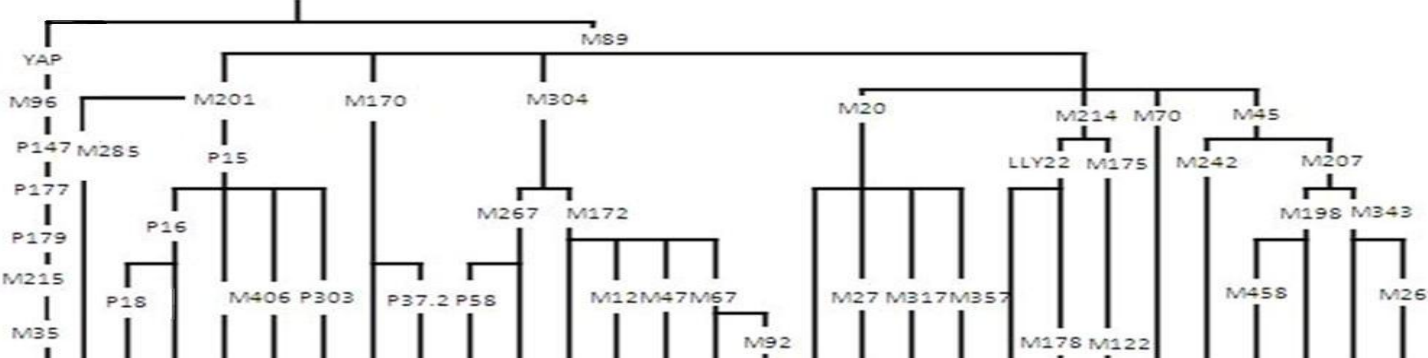
Таблица 5. Частоты (в процентах) гаплогрупп Y-хромосомы в русских популяциях

Популяция	N	C	E-M78	E-M123	F* ¹	G1* ²	G2a	I*	II	I2a1	I2a2a	J1	J2	K* ³	T	N* ⁴	N1b	N1c	Q	R1a	R1b1a1	R1b1a2	R2
			M130	M78	M123	M89	M285	P15	M170	M253	P37	M223	M267	M172	M9	M70	M231	P43	TAT	M242	SRY1532	M73	M269
Мезень	54	0	0	0	1.9	0	0	0	0	0	0	0	0	0	0	0	7.4	46.3	0	44.4	0	0	0
Пинега	114	0	0	0	0	0	0	0	0.9	4.4	0	0.9	0	0	0	0	15.8	23.7	0.9	39.5	0	14	0
Красноборск	91	0	0	0	1.1	0	2.2	3.3	12.1	9.9	0	0	5.5	0	0	0	3.3	36.3	0	19.8	0	6.6	0
Вологда	121	0	0.8	0	0	1.7	0.8	0	11.6	8.3	1.7	0	0.8	0	0	0	3.3	35.5	1.7	33.1	0	0.8	0
<i>северные русские</i>	380	0	0.2	0	0.8	0.4	0.8	0.8	6.2	5.7	0.4	0.2	1.6	0	0	0	7.5	35.5	0.7	34.2	0	5.4	0
Унжа	52	0	5.8	0	1.9	0	0	0	11.5	11.5	3.8	3.8	3.8	0	0	0	0	13.5	0	32.7	1.9	9.6	0
Кашин	73	0	4.1	0	1.4	0	0	0	2.7	8.2	4.1	0	4.1	0	0	0	0	11	1.4	56.2	0	6.8	0
Порхов	57	0	1.8	1.8	0	0	0	1.8	3.5	10.5	0	0	0	3.5	0	1.8	1.8	15.8	0	52.6	0	5.3	0
Остров	75	0	4	0	0	0	0	1.3	6.7	9.3	0	1.3	1.3	0	0	0	0	28	0	45.3	0	2.7	0
Рославль	107	0.9	7.5	0	0.9	0	0	0	1.9	10.3	0	0	2.8	1.9	2.8	0	0.9	13.1	0	45.8	0	11.2	0
<i>центральные русские</i>	364	0.2	4.6	0.4	0.8	0	0	0.6	5.3	10	1.6	1	2.4	1.1	0.6	0.4	0.5	16.3	0.3	46.5	0.4	7.1	0
Ливны	110	0	0.9	0	0	0	0	0	8.2	13.6	0.9	0	0.9	0	2.7	0	0.9	4.5	0.9	62.7	0	3.6	0
Пристенъ	45	0	2.2	0	0	0	0	0	4.4	17.8	0	0	2.2	0	2.2	0	0	13.3	0	55.6	0	2.2	0
Репьевка	96	0	1	0	1	0	2.1	1	3.1	16.7	0	2.1	1	0	0	0	0	6.3	0	59.4	0	5.2	1
Белгород	143	0.7	0.7	0	0	0	0.7	0	3.5	12.6	0.7	0	4.2	0.7	1.4	0	0.7	11.9	0	59.4	0	2.8	0
Кубанские казаки	90	0	3.3	0	0	0	1.1	1.1	4.4	16.5	2.2	0	4.4	0	1.1	0	1.1	6.6	1.1	47.3	0	8.8	0
Терские казаки	125	0	2.4	0	0	0	19.2	0	5.4	7.6	0	3.2	5.6	4.0	0.8	0	0	6.4	0.8	38.4	0	5.6	0
<i>южные русские</i>	609	0.1	1.75	0	0.166	0	3.9	1.3	4.8	14.2	0.6	0.9	3.0	0.8	1.4	0	0.5	8.2	0.5	53.8	0	3.8	0.16

Примечания:

¹ К гаплогруппе F* относятся образцы с мутацией, выявляемой маркером M89, и с предковыми вариантами M9, M201, M170, 12f2.² К гаплогруппе G1* относятся образцы с мутацией, выявляемой маркером M285, и с предковым вариантом P20.³ К гаплогруппе K* относятся образцы с мутацией, выявляемой маркером M9, и с предковыми вариантами M70, 92R7, M231.⁴ К гаплогруппе N* относятся образцы с мутацией, выявляемой маркером M231, и с предковыми вариантами TAT, P43, M128.

Таблица 6. Частоты (в процентах) гаплогрупп Y-хромосомы в популяциях Кавказа



Популяция	N	1	0	0	0	1	0	86	0	0	0	0	6	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	3	0	0
ШАПСУТИ	100	1	0	0	0	1	0	86	0	0	0	0	6	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	3	0	0	
АБХАЗЫ	58	2	0	12	0	24	0	21	0	0	0	0	5	0	0	7	2	0	0	3	0	0	0	0	0	2	0	0	10	0	12	
ЧЕРКЕСЫ	142	1	0	9.2	0	1	0	30	1	0	1	5	12	1	1.4	6	2	0	1	0	0	1	2	0	1	1	6.3	13	2	3		
ОСЕТИНЫ	359	0	0	64	3	0	0	3	0	0	0	2	7	0	0	10	0	0	0	1	0	0	0	0	0	0	2	0	1	1	7	
ИНГУШИ	143	0	0	1	0	0	0	0	1	0	0	3	1	0	0	87	0	0	0	0	3	0	0	0	0	0	0	0	4	0	0	
ЧЕЧЕНЦЫ	330	0	0	3	0	0	0	2	0	0	0	21	1	0	0	55	0	0	0	0	7	0	0	0	0	3	0	4	0	2		
АВАРЦЫ	115	0	0	1	0	0	1	9	1	1	1	58	5	0	0	1	0	0	0	3	0	2	0	0	0	1	0	2	0	15		
ДАРГИНЦЫ	101	0	0	0	0	1	0	1	0	0	1	69	0	0	0	1	0	0	0	0	0	0	0	0	3	0	0	0	22	0	2	
КУБАЧИНЦЫ	65	0	0	0	0	0	0	0	0	2	0	99	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
КАЙТАГЦЫ	33	0	0	0	0	0	0	0	0	0	0	85	3	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	0	6	
ЛЕЗГИНЫ	81	1	0	1	0	12	0	0	0	0	0	44	3	0	0	0	0	0	0	1	0	0	1	0	3	0	0	4	0	30		
		E1b1b1	G1	G2a1a	G2a1	G2a*	G2a3a	G2a3b1	I*	I2a	I1e	I1*	I2*	I2b	I2a4a	I2a4b*	I2a4b1	I*	L1	L2	L3	N1*	N1c1	O3	T	Q	R1a1a7	R1a1a*	R1b	R1b1b2		

Таблица 7. Частоты (в процентах) гаплогрупп Y-хромосомы в популяциях Сибири и Центральной Азии

	N	C	D	E	G	H	I	J	L	N1b	N1c	O	Q	R1a	R1b1b1	R1b1b2	T	Другие
		M130	M174	M96	M201	M69	M170	M304	M20	P43	M178	M175	M242	M198	M73	M269	M70	
АЛТАЙЦЫ СЕВЕРНЫЕ	202	0	0	0	0	0	2	1	0	9	2	0	26	30	18	0	0	13
АЛТАЙЦЫ ЮЖНЫЕ	207	12	10	0	0	0	1	3	1	5	5	6	5	51	0	1	1	1
ШОРЦЫ	139	0	0	1	0	1	0	0	0	3	31	1	0	57	7	0	0	0
ХАКАСЫ	185	0	0	0	0	0	0	0	0	40	18	0	16	26	2	0	0	1
КАЗАХИ	122	34	1	0	24	0	0	3	0	1	4	22	2	7	1	0	0	2
КИРГИЗЫ	246	33	0	1	1	0	0	1	1	12	1	4	1	42	3	0	0	0

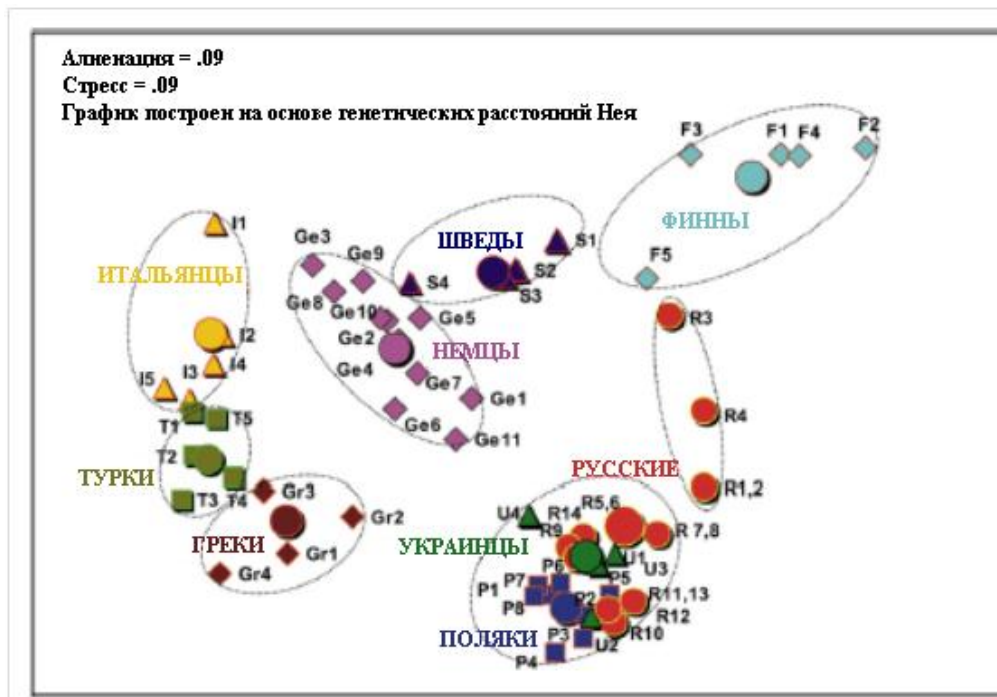


Рис.2. Генетические взаимоотношения европейских популяций по Y-хромосоме.

Популяции, относящиеся к одному народу, показаны однотипными значками одного цвета.

Поскольку этнический фактор оказался столь существенным для структурирования европейского генофонда, был проведен анализ межпопуляционной изменчивости на двух уровнях популяционной иерархии: а) **внутриэтническом** (различия популяций в пределах одного народа, т.е. оценка гетерогенности этнического генофонда); б) **межэтническом** (различия между усредненными этническими частотами гаплогрупп разных народов Европы). Таблица 8 показывает, что в пределах Европы наибольшая гетерогенность характерна для финнов, хорватов, русских и итальянцев; шведы и немцы демонстрируют умеренную гетерогенность; другие народы (греки, турки, поляки, белорусы, украинцы) характеризуются сниженной в масштабе Европы гетерогенностью. В среднем, гетерогенность типичного европейского народа составляет 0.03 ($G_{ST}=0.029$, $F_{ST}=0.027$). Таким образом, по гаплогруппам Y-хромосомы для населения Европы различия между популяциями одного этноса значительно меньше, чем различия этносов между собой (межэтнический уровень $G_{ST}=0.15$, $F_{ST}=0.17$). Общая величина межпопуляционных различий по Y-хромосоме (суммарно внутриэтнический и межэтнический уровни) составляет для Европы около 0.19. Это является дополнительным подтверждением резко выраженных различий внутри Европы по Y-хромосоме (что видно и на рис. 1, 2), поскольку по другим генетическим системам межпопуляционная изменчивость Европы на порядок ниже: $G_{ST}=0.02$ (Балановская, Балановский, 2007).

Аналогичный анализ этнической гетерогенности проведен для народов Южной Сибири и Центральной Азии (табл. 8). Суммарные межпопуляционные различия в этом регионе (0.20) оказались практически такими же, как и для популяций Европы (0.19). Но, в отличие от Европы, основная часть изменчивости приходится на изменчивость внутри этнотерриториальных объединений (0.15), а на различия между ними остается только 0.06. Такой результат объясняется тем, что этнотерриториальные объединения, выделяемые традиционно (хакасы, алтайцы), включают ряд малых народностей, генофонды которых значительно различаются между собой (Балаганская и др., 2011).

Таблица 8. Закономерности изменчивости Y-хромосомы на внутриэтническом и межэтническом уровнях популяционной иерархии

	N_{POP}	N	N_{HG}	Средние генетические расстояния	G_{ST} ($G_{ST} * 10^2$)	AMOVA ($F_{ST} * 10^2$)
ЕВРОПА						
Финны	5	107	6	0.13	8.2	8.1
Хорваты	5	100	8	0.18	8	7.4
Русские	10	123	8	0.14	5.2	4.5
Итальянцы	5	105	8	0.13	3.9	5.0
Шведы	4	76	12	0.13	2.7	1.4
Немцы	11	110	9	0.07	2.1	1.5
Греки	4	91	8	0.10	1.4	1.1
Турки	5	87	26(8)	0.17 (0.04)	1.3(0.8)	1.0
Поляки	8	114	7	0.02	1.1	0.3
Украинцы	4	102	10	0.02	0.9	0.4
Межэтнический уровень	10		6		14.9	16.6
Внутриэтнический уровень	6		6		2.9	2.7
Внутрипопуляционный уровень	61		6		82.2	80.7
СИБИРЬ И ЦЕНТРАЛЬНАЯ АЗИЯ						
Северные алтайцы	3	67	30	0.80	15.1	20.0
Южные алтайцы	2	104	30	0.07	3.9	3.2
Шорцы	2	70	30	0.09	4.9	4.7
Хакасы	4	62	30	0.85	18.2	30.3
Киргизы	2	123	30	0.15	2.9	5.1
Межэтнический уровень	5		30		-	5,5
Внутриэтнический уровень	3		30		-	14,7
Внутрипопуляционный уровень	13		30		-	79,8

Примечания: N_{POP} – число региональных популяций, изученных для данного народа; N – средний размер выборки региональных популяций; N_{HG} – число основных гаплогрупп, включенных в анализ. Приведены значения гетерогенности, рассчитанные тремя разными методами. Различия в гетерогенности турок объясняются высоким филогенетическим разрешением, с которым изучены эти популяции (26 гаплогрупп). Если эти гаплогруппы объединить в 8 (до уровня, на котором изучены другие народы), то средние расстояния падают с 0,17 to 0,04, однако показатель G_{ST} более стабилен, снижаясь с 1,3 до 0.8.

2. ИЗМЕНЧИВОСТЬ МИТОХОНДРИАЛЬНОЙ ДНК В ЕВРОПЕ

Изменчивость мтДНК изучена в 22 популяциях русских, украинцев и белорусов, охватывающих все части обширного ареала восточных славян (собственные данные, табл. 2). Частоты гаплогрупп приведены в таблице 9. Данные представлены также на нашем сайте www.genofond.ru и в работе (Балановский и др., 2009, 2011).

Использование нашей базы данных MURKA расширяет число доступных для анализа популяций (табл. 3). Такое обилие данных позволило провести строгий отбор популяций по следующим критериям: а) объем выборки не менее 70 (меньшие выборки были исключены или объединены); б) надежное определение гаплогрупп (подтвержденных ПДРФ-анализом или однозначно прогнозируемых по ГВС1); в) наличие четкой географической привязки в описании выборки. Был также проведен и строгий отбор гаплогрупп: в анализ включены только гаплогруппы, средняя частота которых в Европе не менее 1% (критерий полиморфизма).

Таблица 9. Частоты гаплогрупп мтДНК у восточнославянских популяций

гаплогруппа	северные русские	южные русские	казаки	белорусы	украинцы
A	0.002	0.000	0.000	0.000	0.003
B	0.000	0.000	0.000	0.000	0.000
C	0.006	0.008	0.015	0.006	0.007
D	0.017	0.009	0.008	0.003	0.002
F	0.000	0.003	0.000	0.003	0.000
G	0.000	0.003	0.000	0.009	0.002
Z	0.006	0.009	0.000	0.000	0.000
L	0.000	0.000	0.000	0.000	0.001
M1	0.003	0.005	0.000	0.000	0.000
preHV	0.000	0.000	0.000	0.006	0.003
H	0.470	0.389	0.394	0.389	0.384
HV	0.008	0.005	0.015	0.015	0.035
I	0.000	0.017	0.038	0.020	0.027
J	0.063	0.078	0.106	0.087	0.086
T*	0.006	0.008	0.000	0.005	0.008
T1	0.030	0.023	0.053	0.027	0.030
T2	0.066	0.086	0.061	0.049	0.077
N1a	0.000	0.003	0.000	0.000	0.000
N1b	0.000	0.007	0.000	0.014	0.011
R*	0.021	0.010	0.008	0.010	0.003
U1a	0.000	0.008	0.000	0.000	0.001
U1b	0.000	0.000	0.023	0.000	0.000
U2	0.016	0.024	0.015	0.024	0.015
U3	0.000	0.021	0.023	0.020	0.011
U4	0.045	0.042	0.023	0.030	0.052
U5a	0.038	0.070	0.053	0.067	0.055
U5b	0.035	0.017	0.030	0.024	0.043
U6	0.000	0.000	0.000	0.000	0.000
U7	0.008	0.002	0.008	0.011	0.000
U8	0.011	0.005	0.000	0.008	0.003
K	0.077	0.036	0.068	0.021	0.049
V	0.019	0.060	0.015	0.062	0.050
W	0.023	0.023	0.008	0.038	0.026
X	0.003	0.009	0.023	0.003	0.012
прочие	0.028	0.020	0.015	0.052	0.004

Перечисленным критериям удовлетворяли 118 популяций (средний объем выборки N=200, суммарно 23500 образцов) и 35 гаплогрупп (табл. 8).

КЛАСТЕРИЗАЦИЯ ПОПУЛЯЦИЙ ПО ЧАСТОТАМ ГАПЛОГРУПП. При анализе методом многомерного шкалирования (рис. 3А) четко выделились четыре кластера. Первый кластер включил только саамов, что не удивительно, учитывая их генетическое своеобразие (Cavalli-Sforza et al., 1994; Tambets et al., 2004). Второй кластер включил те популяции восточных рубежей Европы, у которых повышена частота восточно-евразийских гаплогрупп. Третий кластер включил популяции Передней Азии и Кавказа. Все остальные популяции с основной территории Европы (от Волги до Пиренейского полуострова) вошли в четвертый «паневропейский» кластер, небольшие размеры которого на графике свидетельствуют о низкой межпопуляционной изменчивости. Эти результаты подтверждают гомогенность генофонда Европы (Simoni et al., 2000), но указывают на своеобразие генофондов Приуралья и Передней Азии.

При отдельном анализе только тех популяций, которые вошли в «паневропейский» кластер (рис. 3Б), обнаружилась обособленность популяций северо-востока Европы (главным образом, финно-угорских). Однако подавляющее большинство популяций по-прежнему группировалось в единый гомогенный кластер, который можно назвать «узко-европейским» (рис. 3Б).

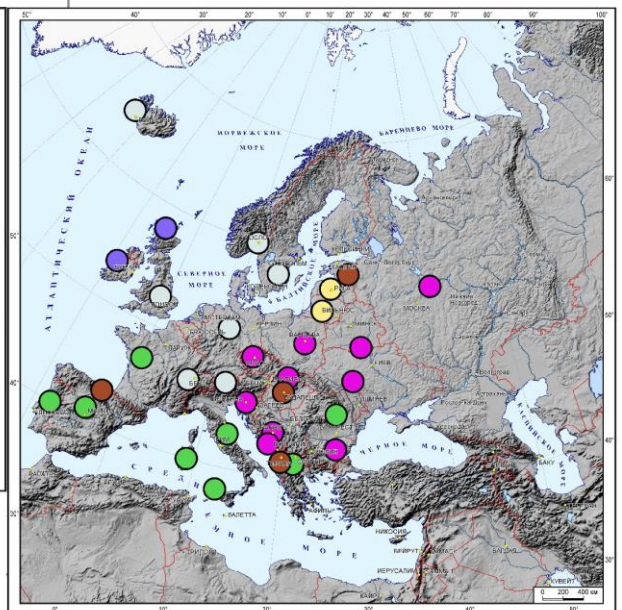
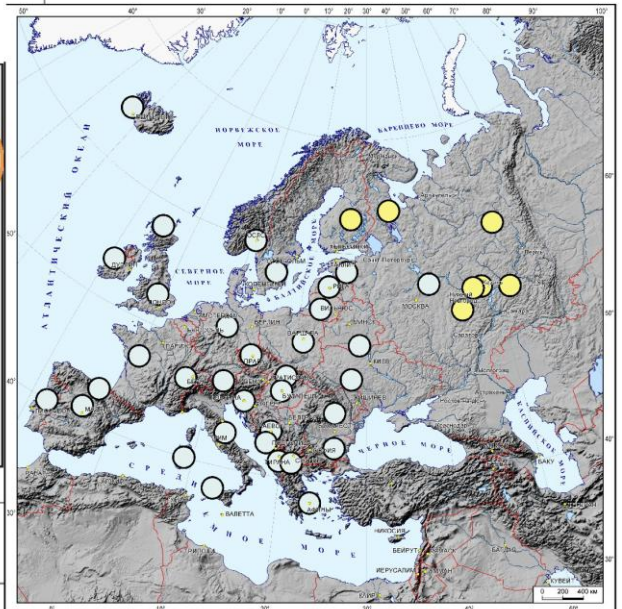
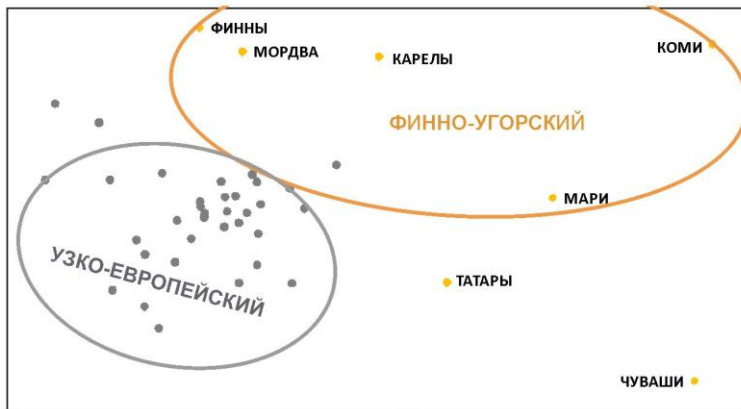
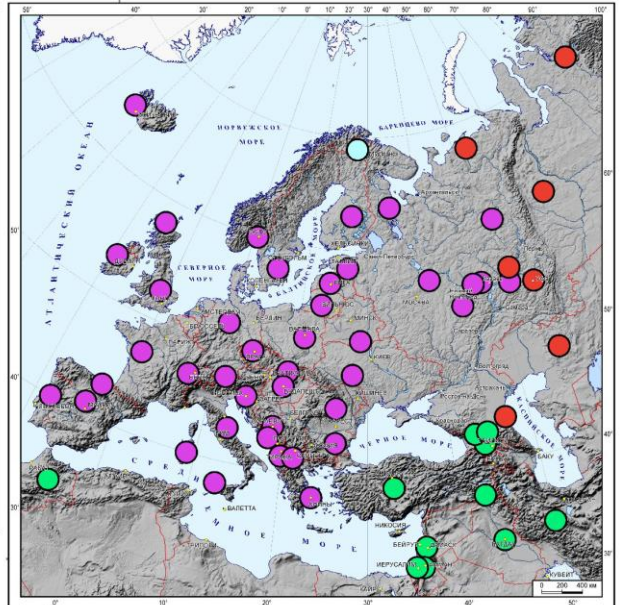
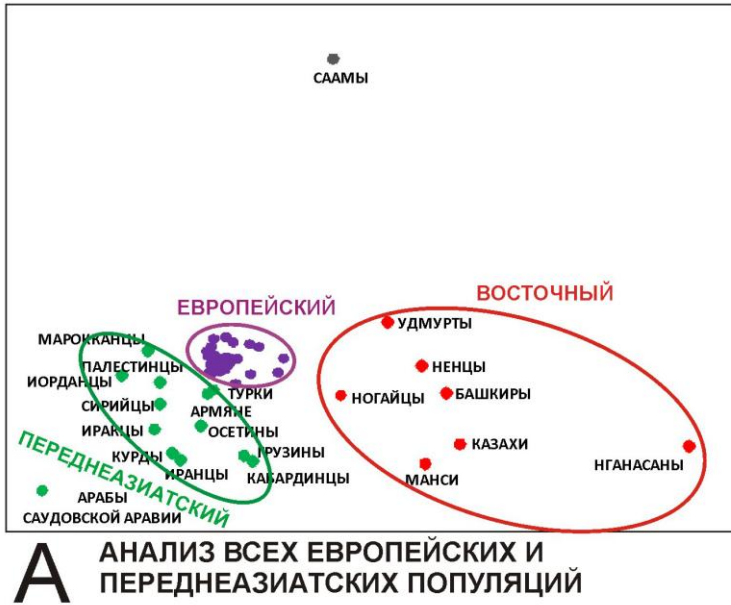


Рис. 3. Генетические взаимоотношения европейских популяций по мтДНК.

На картах соответствующим цветом показано географическое положение популяций, анализируемых на графиках.

Наличие обширной базы данных позволило провести отдельный анализ и этого кластера. В результате (рис. 3В) такого «рассмотрения под большим увеличением» в пределах «узко-европейского» кластера выявился целый ряд кластеров второго порядка. Популяции, говорящие на языках одной лингвистической группы, расположились в генетическом пространстве рядом друг с другом. Так, популяции германской, славянской, романской и балтской языковых групп формируют каждая отдельную группу и в генетическом пространстве. Учитывая, что население Шотландии лишь исторически недавно сменило язык, их генетическое сходство с ирландцами обрисовывает общность митохондриального генофонда и кельтской группы (рис. 3В).

В обнаруженной закономерности генофонда Европы (популяции одной языковой группы характеризуются сходными митохондриальными генофондами) отмечаются два исключения: румыны и аромуны (лингвистически принадлежат к романской группе, но генетически обособлены), а также эстонцы и венгры (лингвистически финно-угорская группа, но генетически примыкают к «славянскому» кластеру).

РОЛЬ ЛИНГВИСТИЧЕСКОГО И ГЕОГРАФИЧЕСКОГО ФАКТОРОВ. Значимость географического и лингвистического факторов оценивалась во многих работах (Rosser et al., 2000; Nasidze et al., 2004; Кутуев, 2010; Спицын и др., 2009). Впервые обнаруженная значимость лингвистической принадлежности популяций для структурирования митохондриального генофонда Европы подтверждена и другим методом: проведен анализ связи генетических, географических и лингвистических расстояний между популяциями Европы (тест Мантеля, табл. 10). Выявлен примерно одинаковый уровень корреляции генетических расстояний и с географическими, и с лингвистическими расстояниями.

Таблица 10. Связь лингвистики, географии и генетики в генофонде индоевропейских и северокавказских народов Европы

Корреляция между параметрами	мтДНК, индоевропейские народы	Y-хромосома, северокавказские народы
Генетика и лингвистика*	0.31	0.64***
Генетика и география*	0.37	0.60***
Генетика и лингвистика (без влияния географии)**	0.19	0.34
Генетика и география (без влияния лингвистики)**	0.29	0.21

Обозначения: *коэффициент парной корреляции; **частной корреляции; ***достоверность $p < 0.01$.

ИЗМЕНЧИВОСТЬ В ПРЕДЕЛАХ ЛИНГВИСТИЧЕСКИХ ГРУПП. Для уточнения выявленной связи генетических кластеров популяций с лингвистическими группами была рассчитана межпопуляционная изменчивость (гетерогенность) в пределах каждой из лингвистических групп Европы (табл. 11, рис. 4). Оказалось, что митохондриальные генофонды всех лингвистических групп индоевропейской семьи характеризуются умеренной гетерогенностью (от 0.01 до 0.04), тогда как гетерогенность финно-угорской группы уральской языковой семьи значительно выше (0.42 до коррекции, 0.07 после исключения саамов и удмуртов). Это согласуется с обособленностью генофонда финно-угорских популяций и их дисперсным расположением на графике многомерного шкалирования (рис. 3Б). Значительная гетерогенность тюркских народов Европы (0.24) может быть вызвана предполагаемой для многих из них (турки, чуваша и др.) сменой автохтонного языка на тюркский без значительных изменений генофонда.

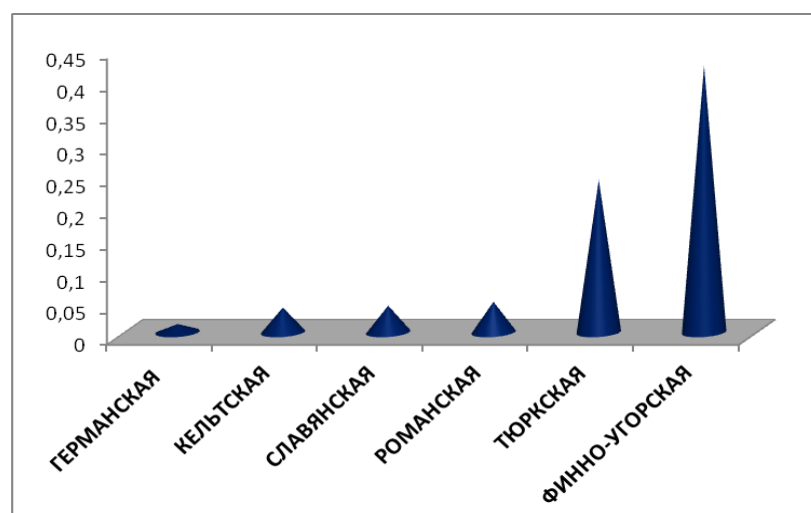


Рис.4. Сравнение лингвистических групп Европы по гетерогенности их генофонда.

Таблица 11. Гетерогенность генофондов лингвистических групп Европы по мтДНК

Лингвистическая группа	Число популяций	Средние генетические расстояния	Расстояния после коррекции*	Исключенные обособленные популяции
Германская	14	0.013	0.013	-
Кельтская	7	0.039	0.039	-
Славянская	27	0.042	0.034	Хорваты о-ва Крк, белорусы Бобруйска
Романская	27	0.048	0.044	Кантабрия
Тюркская	5	0.240	0.240	-
Финно-угорская	9	0.420	0.070	Саамы, удмурты

*В некоторых случаях одна-две обособленные популяции (например, островные популяции хорватов) за счет больших генетических расстояний до остальных популяций завышали уровень гетерогенности всей группы. Поэтому при коррекции исключались из расчета те популяции, среднее расстояние до которых более чем в 2 раза превосходило средний уровень для данной языковой группы.

ГАПЛОТИПИЧЕСКОЕ РАЗНООБРАЗИЕ мтДНК. Все выше описанные виды анализа проведены по частотам гаплогрупп. Для увеличения разрешающей способности проведен также анализ и по отдельным гаплотипам. Уровень гаплотипического разнообразия рассчитан для 51 восточноевропейских популяций как по собственным данным (табл. 2), так и по созданной базе данных (табл. 3). Столь представительный анализ гаплотипического разнообразия мтДНК выполнен впервые в мире.

Картографирование этого показателя демонстрирует закономерное убывание к северу: от 0.99 у популяций Причерноморья до 0.95 и ниже на берегах Северного Ледовитого океана (рис. 5). Снижение внутривнутрипопуляционного разнообразия является индикатором действия дрейфа генов, особенно сильного в небольших по размеру и изолированных популяциях севера Европы. Эту зону (рис. 5) населяют в основном финно-угорские народы, межпопуляционное разнообразие которых велико (табл. 11, рис. 4). Итак, низкое внутривнутрипопуляционное и высокое межпопуляционное разнообразие оказываются двумя сторонами одной медали, обусловленными действием дрейфа генов (эффект Валунда) в подразделенном финно-угорском генофонде.

Северные области способны поддерживать меньшую численность населения, к тому же северные популяции более изолированы за счет меньшей проходимости территории. Такая эколого-географическая трактовка подтверждается корреляцией карты разнообразия (рис. 5) с картой плотности населения, которая также снижается к северу.



Рис.5. Карта гаплотипического разнообразия мтДНК в Восточной Европе.

Можно предположить причинно-следственную цепь взаимосвязанных изменений от юга к северу сразу нескольких параметров: географической широты (возрастает), продуктивности среды (убывает), плотности населения (убывает), эффективного размера популяции (убывает), изоляции субпопуляций (возрастает), дрейфа генов (возрастает), гаплотипического разнообразия мтДНК (убывает).

Согласно теории популяционной генетики интенсивный дрейф генов должен приводить к возрастанию случайного инбридинга, а, следовательно, и генетического груза популяции. И действительно, выявлена достоверная корреляция ($r=0.4$) между гаплотипическим разнообразием мтДНК и отягощенностью популяций аутосомно-рецессивной наследственной патологией (Балановский и др., 2011). Это подтверждает роль дрейфа генов как ключевого фактора в формировании северных генофондов.

3. ОСОБЕННОСТИ ГЕНОФОНДА ВОСТОЧНЫХ СЛАВЯН

Структура генофонда изучена как на уровне Европы и Евразии в целом, так и на более низких иерархических уровнях. В качестве такого модельного объекта выбраны восточные славяне (наиболее многочисленная группа населения России и сопредельных стран Европы), а среди восточных славян особое внимание уделено русскому этносу.

МИТОХОНДРИАЛЬНЫЙ ГЕНОФОНД ВОСТОЧНЫХ СЛАВЯН. Исследованию изменчивости мтДНК у восточных славян в их историческом ареале посвящен ряд работ (Orekhov et al., 1999; Malyarchuk et al., 2001., 2002., 2004; Belyaeva et al., 2003; Kornienko et al., 2004; Grechanina et al., 2006; Grzybowski et al., 2007; Балановский и др, 2010; Morozova et al., 2011), различавшихся подходами к формированию выборок. Поскольку для близкородственных генофондов этот вопрос особенно важен, то в анализ включены главным образом собственные данные. Результаты анализа частот гаплогрупп у 22

восточнославянских популяций (N=2 396 образцов) и окружающих их неславянских народов представлены на рис. 6. Во избежание перегрузки графика, на нем оставлены лишь популяции (от Балканского полуострова до Урала), имеющие непосредственное отношение к этногенезу восточной ветви славян. Выявленная структура генофонда сводится к трем кластерам. В «приуральский» кластер вошли восточные финно-угорские (марийцы, коми) и тюркские народы (татары, чувашаи). В «восточно-европейский» кластер вошли западные и восточные славяне, балты, а также балканские народы. В «северо-европейский» кластер вошли западные финно-угорские народы и северные русские популяции.

Таким образом, выявляется генетически гомогенная балто-славянская общность, к которой тяготеют также и балканские популяции («восточно-европейский» кластер). От этой общности отличаются финно-угорские и тюркские народы Восточной Европы, генофонд которых, в свою очередь, подразделяется на западную и восточную части (кластеры 1 и 3). Такая картина могла сформироваться в ходе славянской колонизации Восточно-Европейской равнины (Происхождение и этническая история русского народа, 1965; Алексева, 1973; Дерябин, 1999):двигающееся с запада славянское население несло с собой «восточно-европейский» митохондриальный генофонд, который замещал ранее существовавший на этих территориях автохтонный «приуральский» генофонд, разделяя оставшиеся не ассимилированными группы на западную и восточную части.

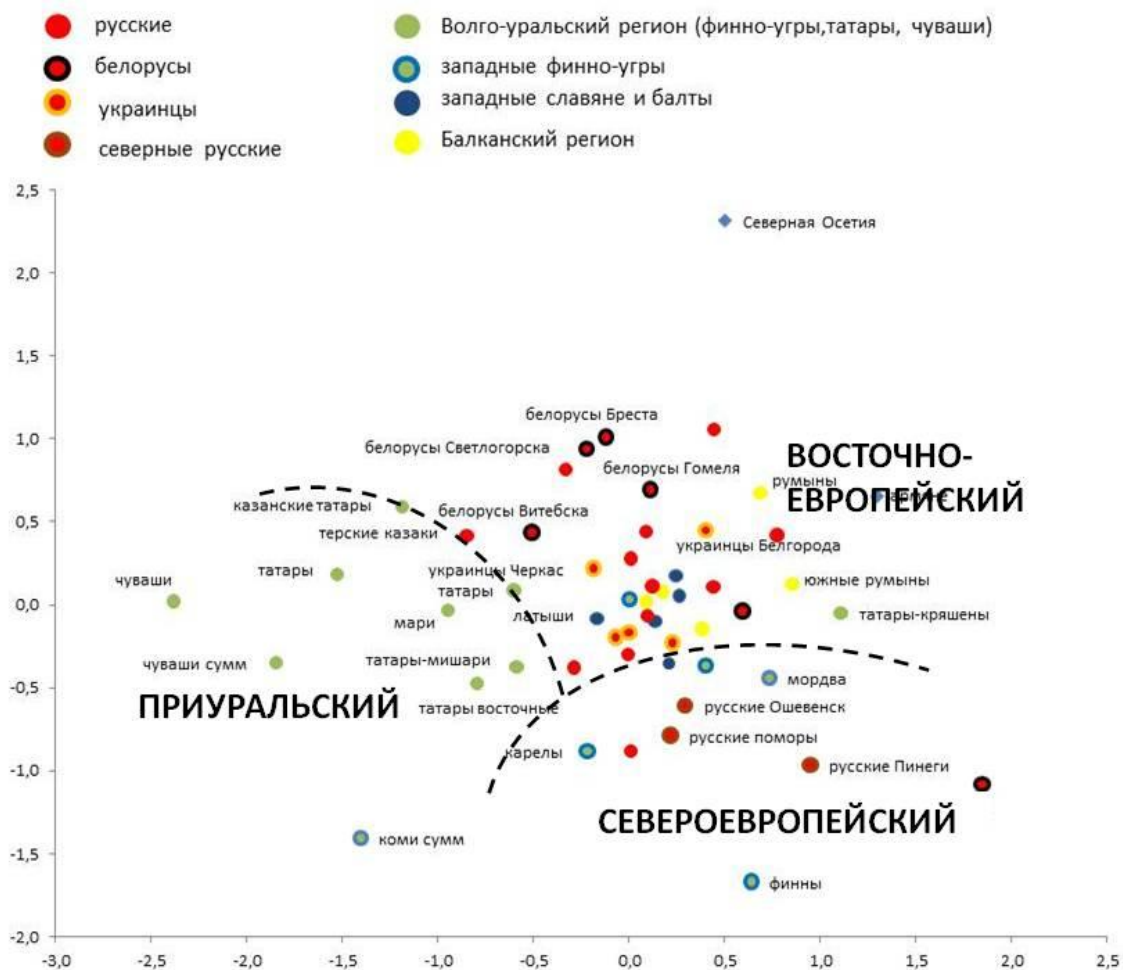


Рис.6. Структура митохондриального генофонда восточных славян и их соседей.

ГЕНОФОНД ВОСТОЧНЫХ СЛАВЯН ПО Y-ХРОМОСОМЕ. Анализ Y-хромосомы обрисовал в целом аналогичную картину: восточнославянские популяции оказываются сходными между собой и с западнославянскими популяциями, при этом прослеживается своеобразие северных русских популяций (Balanovsky et al., 2008).

Различия в результатах, полученных по мтДНК и Y-хромосоме, невелики и связаны главным образом с тем, что мтДНК в целом более гомогенна в Европе и поэтому в меньшей степени, чем Y-хромосома, выявляет различия между генофондами. Например, по мтДНК южные славяне присоединяются к остальной славянской общности, хотя в генетическом пространстве располагаются на ее окраине (рис. 6). Но результаты по Y-хромосоме (разрешающая способность которой выше) уточняют картину: четко видно, что южные славяне отделены от восточных и западных славян (Balanovsky et al., 2008).

Ряд закономерностей выявлен для генофондов украинцев и белорусов. Литературные данные об этих популяциях (Rosser et al., 2000; Semino et al., 2000; Behar et al., 2003; Харьков и др., 2004, 2005; Battaglia et al., 2008) получены по трудно сопоставимым панелям SNP маркеров, поэтому мы опирались главным образом на собственные данные. Для украинцев показано сходство с более северными группами (Польша, Белоруссии и юга России) и некоторое своеобразие западных украинцев. Обнаружена гомогенность генофонда белорусов, сквозь которую все же пробиваются слабые различия между северными и южными группами (Балановский, Тегак, 2008).

СТРУКТУРА РУССКОГО ГЕНОФОНДА. Более высокая разрешающая способность Y-хромосомы по сравнению с мтДНК обнаруживается и при анализе русского генофонда. По мтДНК его структура выявляется лишь на уровне тенденций: в изменении частот некоторых гаплогрупп с юга на север, в отличии окраинных популяций казаков и поморов (Балановский и др., 2009). Близкие результаты получены и другими авторами (Malyarchuk et al., 2002, 2004; Morozova et al., 2011). При анализе же Y-хромосомы изменчивость большинства гаплогрупп близка к клинальной, что подтверждено анализом автокорреляций (Balanovsky et al., 2008). Информативность Y-хромосомы проявляется и при многомерном анализе: центральные и южные популяции оказываются генетически сходными, а северные популяции - своеобразными (рис. 7).

Кардинальные различия северных и южных русских проявляются и при анализе карт главных компонент изменчивости различных признаков: гаплогрупп Y-хромосомы и классических маркеров (Balanovsky et al., 2008), фамилий (Балановский и др., 2001), антропологических признаков (Балановская, Балановский, 2007). Взаимное подтверждение результатов по четырем системам признаков (принцип «полисистемного подхода») не оставляет сомнений в реальности обнаруженной закономерности широтной изменчивости русского генофонда.

Первая причина высокой межпопуляционной изменчивости на Русском Севере - сниженная плотность населения и меньший эффективный размер популяций (Ревазов и др., 1979; Медико-генетическое описание населения Адыгеи, 1997). Поэтому дрейф генов в северных русских популяциях действует сильнее. Тем самым снова обнаруживается параллелизм с изменчивостью мтДНК (с описанным выше широтным трендом гаплотипического разнообразия и связанных с ним параметров). Но отличия северных русских популяций вызваны не только действием дрейфа генов: в этом случае северные русские группы разместились бы случайным образом вокруг южных, без систематического сдвига в одну сторону. Поэтому второй причиной могут быть особенности происхождения северных русских. Это побудило нас провести специальное исследование генофонда Русского Севера (Балановская и др., 2011).

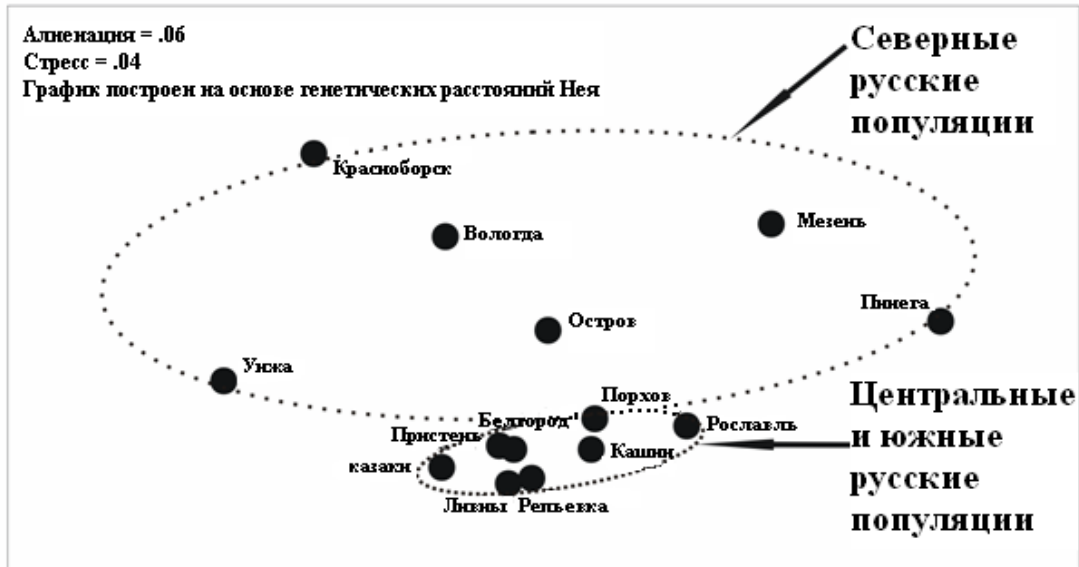


Рис. 7. Структура русского генофонда по Y-хромосоме (многомерное шкалирование)
Выделение кластеров подтверждено анализом AMOVA: при выделении групп, показанных овалами, изменчивость между группами (5.8%) оказывается в 6 раз выше, чем внутри групп (1.5%).

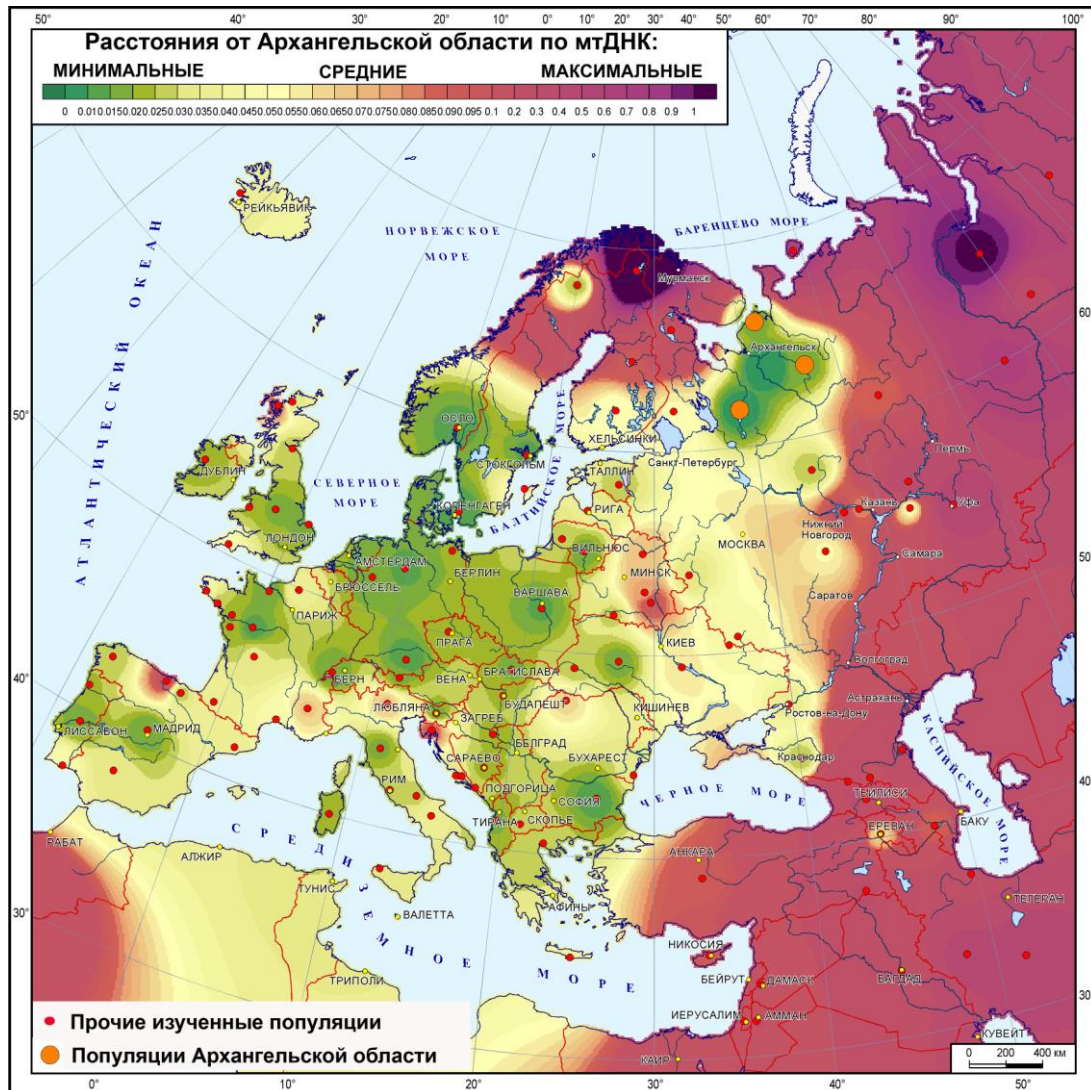


Рис.8. Карта генетических расстояний от северных русских популяций (мтДНК)
Обозначения: зеленые тона – популяции, генетически близкие к генофонду северных русских;
красные тона – популяции, генетически далекие от генофонда северных русских.

ПРОИСХОЖДЕНИЕ СЕВЕРНЫХ РУССКИХ ПОПУЛЯЦИЙ. Анализ популяций Русского Севера проведен по совокупности данных о полиморфизме Y-хромосомы, мтДНК, аутосомных маркеров и фамилий.

По данным о Y-хромосоме северные русские вошли в обширный «северный» кластер вместе с популяциями балтов (латышей и литовцев), финноязычных народов (коми, финнов, эстонцев, карелов), а также германоязычных шведов. Сходство Русского Севера с географически отдаленными балтами более выражено, чем с финно-уграми. Этот пестрый в лингвистическом отношении (славяне, балты, финно-угры, германцы), но географически единый (от Балтики до Печоры) континуум популяций мог сформироваться еще до выделения балто-славянских и финских лингвистических групп.

По митохондриальной ДНК Русский Север обнаруживает близость с еще более отдаленными популяциями Европы (рис. 8). Русский Север оказался генетически близким к норвежцам, немцам, полякам, литовцам и т.д., и при этом генетически отличным от финноязычных популяций. Анализ аутосомного маркера CCR5del32, изменчивость которого активно изучалась многими российскими коллективами (Libert et al., 1998; Сломинский и др., 1997; Limborska et al., 2002; Кожекбаева и др., 2004) дал сходные результаты (Балановская и др. 2011).

В целом, обнаруженное сходство северных русских с популяциями Северной и Центральной Европы позволяет выдвинуть гипотезу о сохранении на этих территориях древнего генофонда Европы, восходящего, возможно, еще к эпохе мезолита. Этот палеоевропейский генофонд унаследован как северными древнерусскими популяциями, так и частью финноязычных племен, расселившихся здесь до славянской колонизации, а также балтоязычными и германоязычными популяциями. Тем самым роль финно-угорского субстрата при подробном изучении оказывается не столь однозначной, а сама проблема уходит в большую древность, чем эпоха славянской экспансии.

4. КОЭВОЛЮЦИЯ ГЕНОФОНДОВ и ЯЗЫКОВ

Анализ взаимосвязи генетической, географической и лингвистической изменчивости популяций проведен на примере Кавказа. Попытки решить эту проблему традиционны для популяционной генетики человека (Cavalli-Sforza et al., 1994; Rosser et al., 2000; Рычков, Ящук, 1980, 1983, 1985). Но в большинстве регионов мира сложные паттерны миграций, ассимиляций, смены языка усложняют реконструкцию популяционных процессов. Такую реконструкцию легче осуществить на Кавказе, где горный рельеф препятствовал интенсивному смешению популяций и внешним миграциям. Это создавало условия как для постепенного формирования лингвистического разнообразия путем дифференциации из единого праязыка северокавказской семьи, так и для формирования четкой популяционной структуры генофонда. Эта четкость наглядно проявилась при рассмотрении спектров гаплогрупп в кавказских популяциях (рис. 9).

Налицо тройное соответствие: четыре группы популяций, разделяемые основными генетическими границами, принадлежат четырем различным лингвистическим группам и характеризуются преобладанием одной из четырех различных гаплогрупп (рис. 9). Впервые показанное существование этих четырех лингвистико-географических регионов Кавказа (Дибирова и др., 2010; Balanovsky et al., 2011) ставит нас перед вопросом: какой из факторов – географический или лингвистический – являлся определяющим при формировании такой генетической структуры? Для объективного количественного ответа был проведен анализ тремя методами.

Во-первых (табл. 10), рассчитаны корреляции между матрицами генетических, географических и лингвистических расстояний (аналогично описанному выше анализу генофонда Европы по мтДНК). Связь генетических расстояний с лингвистическими ($r=0.64$) оказалась несколько выше, чем с географическими ($r=0.60$), эта тенденция сохраняется и после вычета тесной ($r=0.78$) связи между лингвистикой и географией.

Во-вторых, эффективность лингвистической классификации в объяснении структуры кавказского генофонда подтверждена методом AMOVA. При географической группировке популяций различия между группами охватили только 15% межпопуляционной изменчивости, тогда как при лингвистической группировке - почти в два раза больше (27%). Этот метод позволил включить в анализ не только народы северокавказской семьи, но и осетин, говорящих на языке индоевропейской семьи.

В-третьих, кластерный анализ показал, что генетическая дендрограмма северокавказских популяций является почти точным отражением лингвистической классификации северокавказских языков (рис. 10).

Такое четкое соответствие генетической и лингвистической реконструкций не было выявлено ранее ни в одном из регионов мира. В ранних работах и на Кавказе не удавалось выявить достоверные корреляции - по всей видимости, из-за недостаточного уровня филогенетического разрешения (Nasidze et al., 2004). В более поздних работах также делался вывод о примате географии над лингвистикой на Кавказе (Кутуев, 2010; Yunusbaev et al., 2011). Недооценка лингвистического фактора в этом случае могла быть вызвана применением «пошаговых» лингвистических расстояний, а не точных количественных показателей сходства языков, как в нашем исследовании.

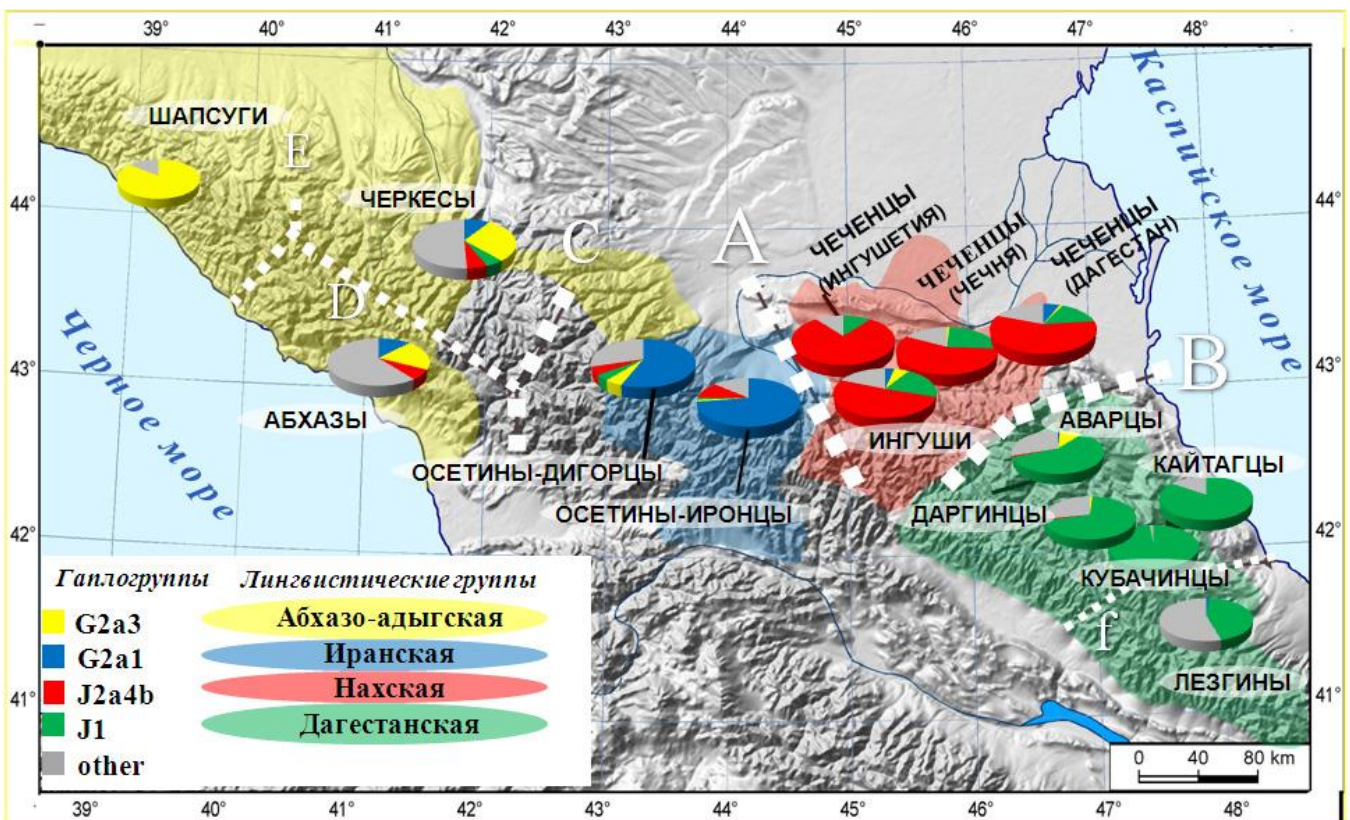
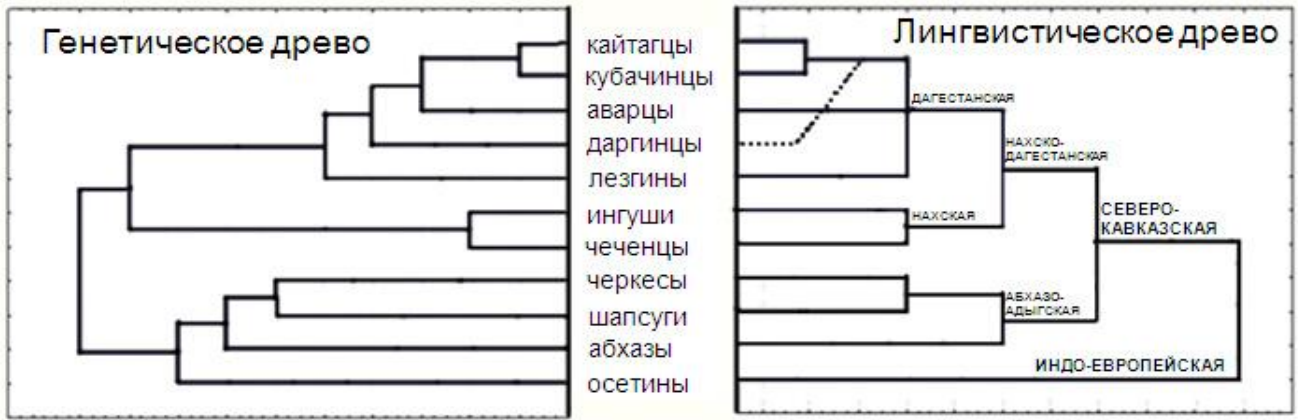


Рис.9. Структура генофонда Кавказа: лингвистико-географические регионы

Каждая изученная популяция показана круговой диаграммой, размер секторов разного цвета отражает частоту соответствующей гаплогруппы Y-хромосомы согласно легенде. Белые пунктирные линии обозначают объективные генетические границы, полученные с помощью программы Barrier 2.2 (толщина линии - значимость границы). Цветным фоном показаны ареалы лингвистических групп.



Ри

с. 10. Сравнение генетической кластеризации и лингвистической классификации.

Убедившись, что все три метода указывают на ведущую роль лингвистического фактора в структурировании кавказского генофонда на уровне гаплогрупп Y-хромосомы, рассмотрим изменчивость на уровне STR-гаплотипов. Например, на филогенетической сети гаплогруппы **G2a1a-P18** (рис. 11а) четко выделяется кластер **α**, характерный для осетин-иронцев, и кластер **β**, характерный для осетин-дигорцев. Генетические датировки обоих кластеров близки (1400 лет при использовании «генеалогической» скорости мутирования). Эти генетические данные свидетельствуют, что разделение осетинской прапопуляции на субэтноты иронцев и дигорцев произошло не позднее чем 1400 лет назад: если бы разделение произошло позже этой даты возникновения кластеров, то в обоих кластерах поровну встречались бы и иронцы, и дигорцы. По лингвистическим датировкам разделение иронского и дигорского диалектов произошло 1300 лет назад. По данным историков разделение на иронцев и дигорцев произошло тоже не позднее 1300 лет назад. Таким образом, обнаружено хорошее совпадение генетических, лингвистических и исторических датировок.

Аналогичные филогенетические сети созданы еще для 10 гаплогрупп (Balanovsky et al., 2011). Все они демонстрируют важную особенность изменчивости Y-хромосомы на Кавказе: существование четких кластеров гаплотипов, многие из которых приурочены к одним популяциям и отсутствуют в других. Обилие таких популяционно-специфичных кластеров (их обнаружено несколько десятков) позволяет впервые совместить генетическую и лингвистическую реконструкции микроэволюции населения Кавказа. На рис. 12 приведено лингвистическое древо, иллюстрирующее постепенное разделение северокавказского праязыка на отдельные языки. На ветви этого лингвистического дерева наложены генетические кластеры гаплотипов, специфичные для соответствующих ветвей (популяций и их групп). Можно видеть параллелизм лингвистической и генетической эволюции. Отметим, что использование «генеалогической» скорости мутирования STR маркеров Y-хромосомы дает лучшее согласие с лингвистическими датировками, чем применение «эволюционной» скорости.

Это позволяет сделать вывод, что структура генофонда Кавказа развивалась параллельно с формированием многообразия северокавказских языков. Связующим звеном между генетической и лингвистической эволюцией служили демографические события в истории популяций - подразделение на дочерние группы. Горный рельеф способствовал изолированности дочерних групп, в которых возникали специфичные для них гаплотипы. Когда дочерняя популяция, в свою очередь, делилась, то генофонды всех «внучатых» групп наследовали характерный кластер гаплотипов. Новые же возникающие гаплотипы

оказывались специфичны уже для этих популяций третьего порядка. Те же самые демографические события, которые приводили к изоляции популяций, способствовали формированию большого числа иерархически родственных языков. Тем самым одни и те же события в демографической истории популяций оставляли свои следы в двух независимых регистрирующих системах: в генофонде и в языке.

Рис. 11А. Филогенетическая сеть гаплогруппы G2a1a-P18 на Кавказе

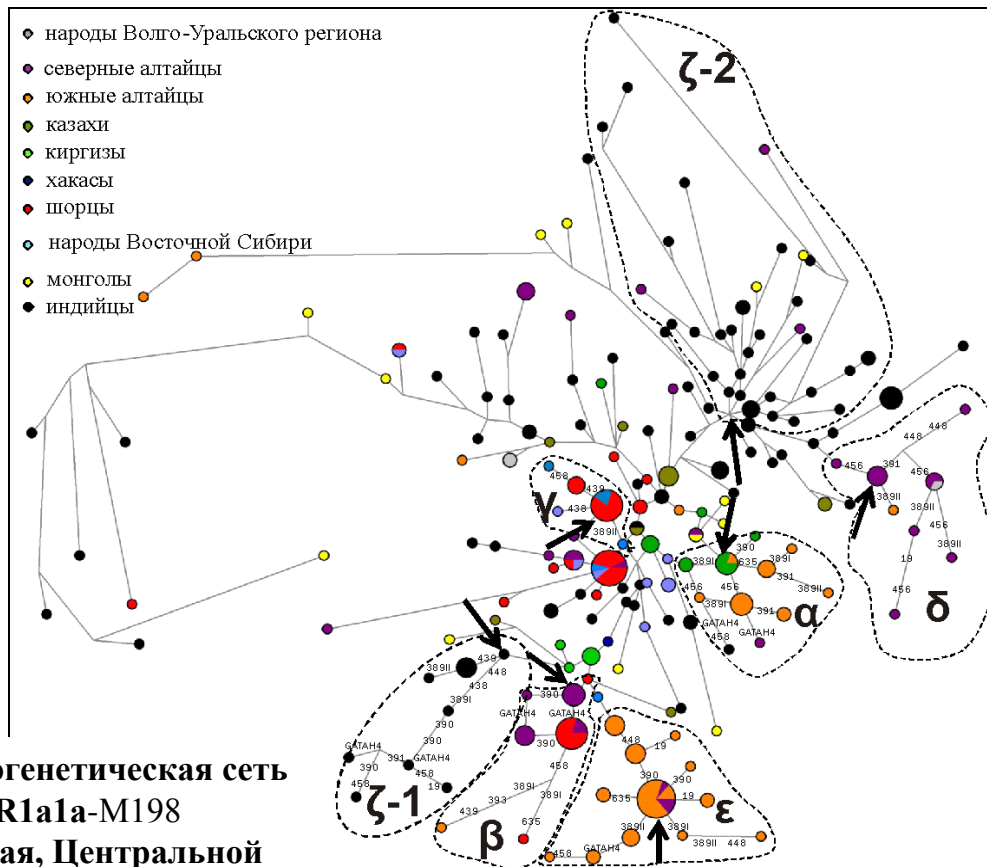
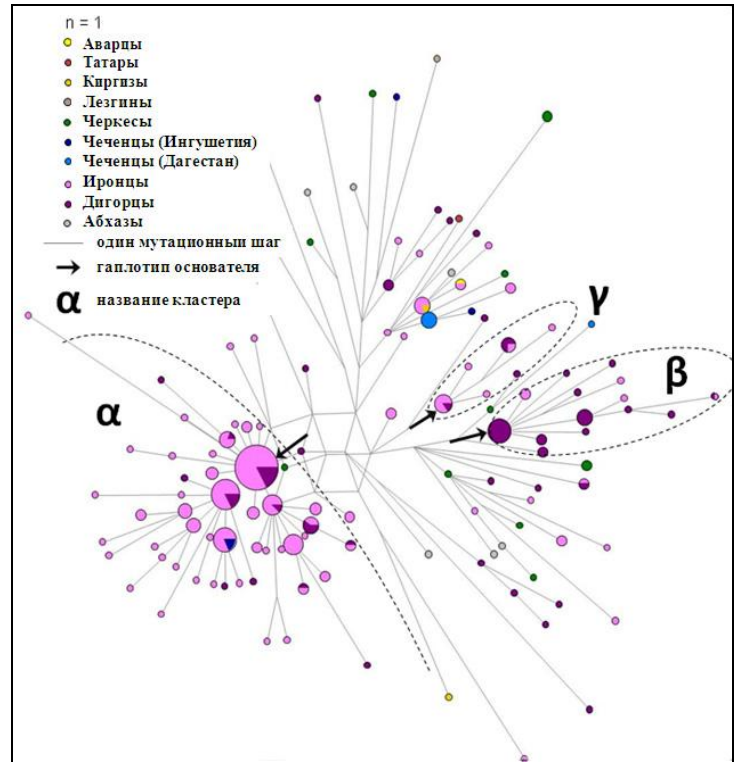


Рис. 11Б. Филогенетическая сеть гаплогруппы R1a1a-M198 у народов Алтая, Центральной Азии и Индии.

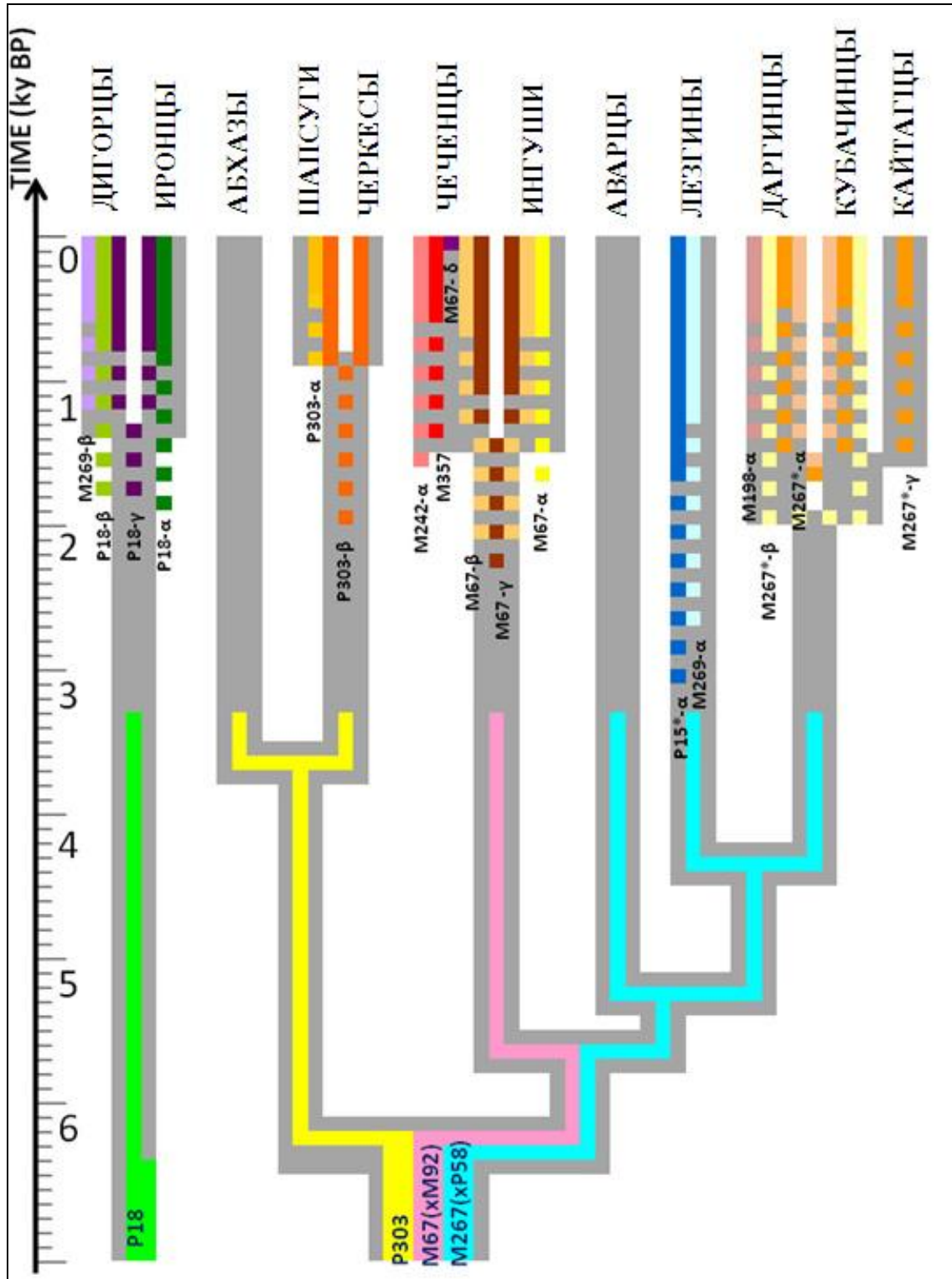


Рис. 12. Комбинированная схема микроэволюции популяций Кавказа

Серым цветом показано лингвистическое древо родства северокавказских языков; датировка ветвлений дана на шкале времени слева. На верхнюю часть этого лингвистического древа цветными линиями нанесены кластеры STR-гаплотипов Y-хромосомы, специфичные для народов, говорящих на соответствующих языках. Пунктирная часть каждой линии указывает доверительный интервал возраста кластера. Для каждого кластера указано его обозначение (например, **G2a1a-P18- α** , структура которого приведена на рис. 11А). На нижнюю часть древа цветными линиями нанесены SNP-гаплогруппы, характерные для групп народов. Во избежание перекрывания с линиями кластеров линии гаплогрупп кончаются 3300 лет назад.

5. АТЛАСЫ ИЗМЕНЧИВОСТИ Y-ХРОМОСОМЫ и мтДНК В ЕВРАЗИИ

В филогеографических исследованиях карты распространения гаплогрупп являются обязательным инструментом (Cruciani et al., 2007, 2010; Semino et al., 2004; Myres et al., 2010; Rootsi et al., 2007; Tofanelli et al., 2009; Underhill et al., 2010). Но, несмотря на такое признание значимости отдельной карты, отсутствуют систематические собрания карт - картографические атласы. Этот пробел объясняется отсутствием систематизированных глобальных баз данных. Единственная известная нам успешная попытка картографирования широкого спектра гаплогрупп была осуществлена в работе (Степанов, 2001). Но за десятилетие, прошедшее со времени этой работы, накоплен огромный массив новых данных, открыты новые гаплогруппы. Что же касается митохондриальной ДНК, то картографические атласы отсутствуют. В связи с большой научной значимостью такого подхода автором предпринято масштабное картографическое исследование: созданы Атлас изменчивости Y-хромосомы и Атлас изменчивости митохондриальной ДНК. Атласы основываются на созданных базах данных (Y-base и MURKA). Каждый атлас состоит из карт частот гаплогрупп, суммарных карт (карты суммарной частоты западно-евразийских, центральноазиатских и других типов гаплогрупп) и карт генетических границ.

АТЛАС ИЗМЕНЧИВОСТИ Y-ХРОМОСОМЫ. Созданный Атлас включает карты распространения 79 гаплогрупп. Они были отобраны следующим образом. Из 485 гаплогрупп, присутствующих в последней версии филогенетического Y-древа (www.isogg.org), для 243 гаплогрупп в научной литературе имеются данные об их частотах в популяциях мира. Из них были отобраны все гаплогруппы, обнаруженные на территории евразийского континента (включая острова) с частотой выше 1% хотя бы в одной-двух популяциях (далее курсивом выделены те дочерние субгаплогруппы, суммарное распространение которых не полностью составляет ареал «родительской» гаплогруппы, выделенной полужирным шрифтом): A3b2-M127, B2-M182, C1-M105, C2-M38, **C3-M217**, C3c-M48, C5-M356, D1-M15, D2-12f2.2, D3a-P47, E1b1a1-DYS271, E1b1b1a1a-V12, E1b1b1a1b-L142.1, E1b1b1a1c-V22, E1b1b1a1d-V65, E1b1b1b1-M81, E1b1b1c-M123, E2-M75, F2-M427, G1-M285, G2a1-P16, G2a2-M286, G2a3a-M406, G2a3b1-P303, G2c-L183, H1-M52, H2-Apt, I1-L118, I2a-P37.2, I2b-L35, **J1-L255**, J1b-M365.1, J1c3-P58, **J2-L228**, J2a-L152, J2a3a-M322, J2a3b-M67, J2a3e-M339, J2a3h-L207.1, J2b-L282, **L-M11**, L1a-M27, L1b-M317, L1c-M357, M1a-P34, N1a-M128, N1b-P43, N1c-M46, **O1-MSY2.2**, O1a1-M307.1, O1a2-M103, O2a-PK4, O2b-022454, **O3-M122**, O3a1c-002611, O3a2b-M7, O3a2c-P164, **Q-M242**, Q1a1-M120, Q1a2-M143, Q1a3-L528, Q1b-L275, **R1a1-L120**, R1a1a1e-PK5, R1a1a1f-M434, R1a1a1g-M458, R1b-L10, R1b1a1-M478, **R1b1a2-L265**, R1b1a2a1a1a-M405, R1b1a2a1a1a1-M467, R1b1a2a1a1b-P312, R1b1a2a1a1b2a-M153, R1b1a2a1a1b2b1-M167, R1b1a2a1a1b4b-M222, R1b1c-V88, R2a-L266, S-M230, T-L206. При анализе генетических границ учитывались только «родительские» гаплогруппы. Карта генетических границ представляет собой карту межпопуляционного разнообразия G_{ST} , построенную по технологии плывущего окна (Балановская, Нурбаев, 1995) с размером окна 1000x1000 км. Максимальное разнообразие ожидается в зонах резкого перепада частот сразу нескольких гаплогрупп, то есть на границах между контрастными генофондами.

В качестве примера на рис. 13 приведена одна из 79 карт атласа – распространение гаплогруппы **G1-M285**. Обнаруживается ее приуроченность к центральным областям Евразии: максимальные частоты в северном Казахстане, повышенные частоты на Иранском нагорье и узкий коридор низких частот в Средней Азии, соединяющий эти области: необычный паттерн, явно несущий следы исторических событий (а не «диффузии»

гаплогруппы из места ее возникновения). Возможно, **G1-M285** была характерна для носителей иранской группы языков, ранее доминировавшей как в горных (Иран, Таджикистан), так и в степных областях Евразии.

Карта генетических границ (рис. 14) выявляет основную зону разлома евразийского генофонда, которая проходит не по Уралу, служащему географической границей между Европой и Азией, а протянулась наискосок через весь континент. Она состоит из нескольких отрезков: кавказского, южно-уральского, северо-казахстанского и среднесибирского (вдоль Енисея). В районе Алтая к ней примыкает с юга другая генетическая граница, отделяющая Среднюю Азию от населения западных районов Китая. Третья генетическая граница совпадает с восточным участком Гималаев. Четвертая (последняя из основных границ, выявленных на территории Евразии) является «морской», разделяя генофонды приамурских народов и японцев. Таким образом, зоны наибольшей изменчивости на стыках контрастных генофондов связаны как с географическими барьерами (горные хребты, морские проливы), так и с историческими контактными зонами (степная полоса Евразии).

АТЛАС ИЗМЕНЧИВОСТИ МИТОХОНДРИАЛЬНОЙ ДНК. Атлас состоит из карт 82 гаплогрупп: A11, A2, A4, A5, A8, B4_16261, B4b1, B4c1b, B4c2, B4g, B5a, B5b, B6, C4a1a, C4a2, C4b3, C5, D, D4b1, D4e1D2, D4o, D5, E, F1a1, F1a4, F1b, F2a, F3a, F3b, G, HV, H, I, J, K, L1b, L2a, M1, M7a, M7b, M7c, M8a, M9a"b, M10, M12a"b, M13, M20, M21, M32, M33, M71, M74, N1a, N1b, N9a, Q, R0a, R2, R22, R5, R6, R9b, R9c, T1, T2, U1, U2a, U2b, U2c, U2e, U3, U4, U5, U5a, U5b, U6, U7, V, W, X, Y, Z. Этот перечень включает гаплогруппы, средние частоты которых в изученных евразийских популяциях превышают 0,1%.

Не останавливаясь на картах отдельных гаплогрупп, перейдем к их типологии. Выделение восточно-евразийских и западно-евразийских гаплогрупп давно стало традиционным, но отнесение конкретной гаплогруппы к тому или другому типу оставляется на усмотрение исследователей, между взглядами которых имеются определенные расхождения. Поэтому автором разработан метод объективной классификации гаплогрупп на их географические континуумы. На первом этапе частоты гаплогрупп в популяциях анализируются методом главных компонент, но на график выводятся не наблюдения (популяции), а переменные (гаплогруппы). Проявившееся на графике объективное объединение гаплогрупп в кластеры (континуумы) означает их скоррелированность друг с другом и следование общему географическому паттерну. Поэтому на втором этапе строятся карты суммарной частоты гаплогрупп, попавших в один континуум, и по его географической приуроченности дается название.

Применение этого подхода к митохондриальному генофонду Евразии показано на рис. 15 и 16. На графике выделились три объективных континуума (рис. 15). Из них континуум 1 и континуум 2 соответствуют западно-евразийским и восточно-евразийским гаплогруппам. Но новый обнаруженный континуум 3 обладает особым географическим паттерном (рис. 16). Его можно назвать «южно-китайским», поскольку его максимальные частоты сосредоточены вокруг Южно-Китайского моря (в Индокитае, Южном Китае, Индонезии, Филиппинах).

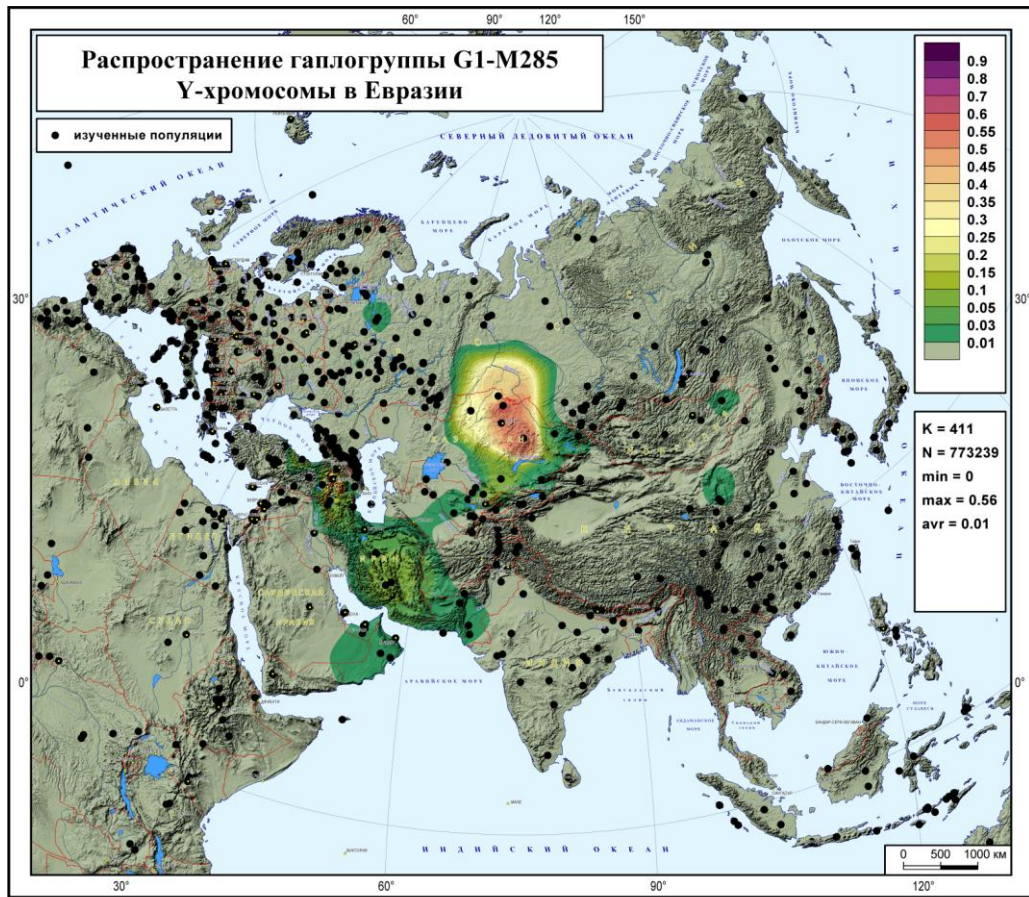


Рис. 13. Карта распространения гаплогруппы G1-M285 Y-хромосомы в Евразии.

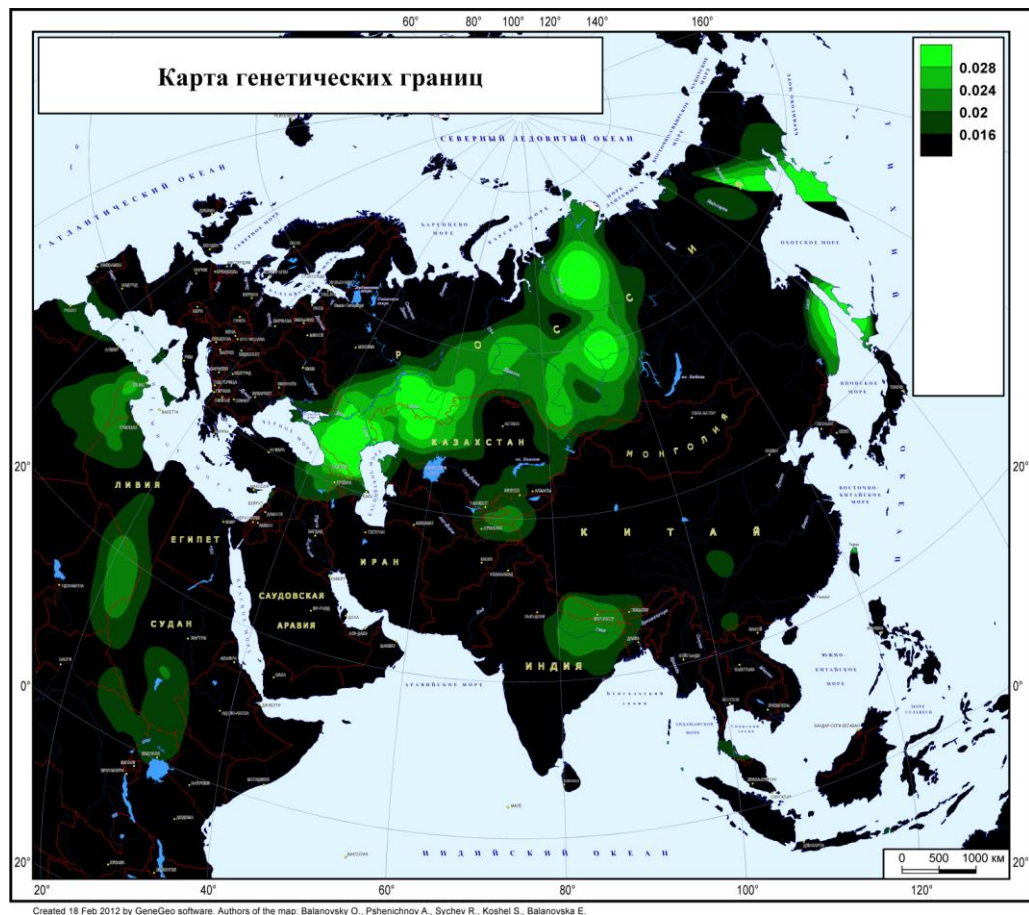
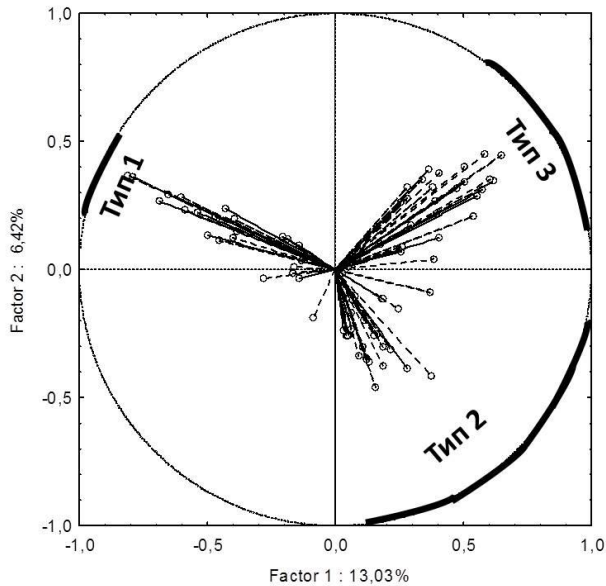


Рис. 14. Структура генофонда Евразии: локализация генетических границ.
(обобщенная карта межпопуляционной изменчивости 79 гаплогрупп Y-хромосомы)



На рисунке представлен результат анализа главных компонент изменчивости частот гаплогрупп мтДНК в популяциях Евразии. В отличие от обычного для популяционных исследований применения главных компонент, на график выведены не наблюдения (популяции), а переменные (гаплогруппы). Кружками показаны все 82 анализируемые гаплогруппы, а пунктирные линии соединяют их с началом координат. Четко выделились три континуума (типа) гаплогрупп. Жирные дуги показывают интервал значений двух первых главных компонент, принадлежащий каждому континууму.

Рис. 15. Объективная географическая классификация гаплогрупп мтДНК.

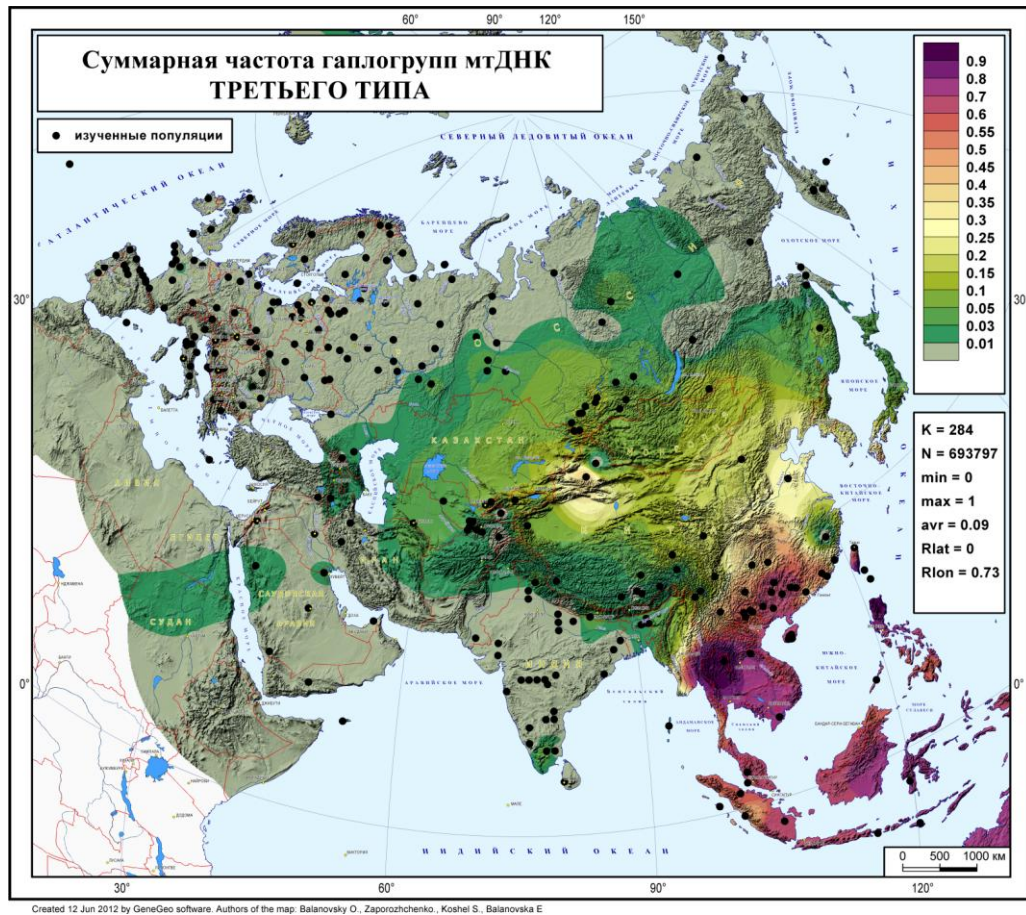


Рис. 16. Карта суммарной частоты гаплогрупп мтДНК «южно-китайского» континуума Евразии

ВЗАИМОДЕЙСТВИЕ ГЕНОФОНДОВ ЕВРОПЫ и АЗИИ. Из исторических источников хорошо известны мощные миграции из Азии, включая скифов в VII веке до н.э., гуннов в IV веке н.э. и монголов в XIII веке н.э. Эти три волны (с интервалом в тысячелетие) и множество менее известных миграций могли приносить в Европу «азиатские» гены по степной полосе, причем описаны генетические особенности по крайней мере последней волны (Zerjal et al., 2003; Захаров, 2003, 2010; Деренко и др., 2007). Для выявления

генетических следов этих миграций с использованием БД Y-base и собственных данных (табл. 1) автором картографировано распространение гаплогрупп Y-хромосомы в степной полосе Евразии (www.genofond.ru; Балаганская и др., 2011).

Карты выявляют ряд восточно-евразийских гаплогрупп (**C3c**, **C(xC3c)**, **D2**, **O**), распространявшихся по степной полосе. Карта их суммарной частоты (рис. 17) показывает абсолютное преобладание в Центральной Азии и сниженную частоту в Казахстане и Средней Азии. Частоты невелики в горах Алтая и Тянь-Шаня, но повышены в степях между ними, формируя необычную «перетяжку» ареала этих гаплогрупп. Такой паттерн изменчивости указывает, что центрально-азиатское население распространялось на запад по степям через исторически известный проход («Джунгарские ворота»). Эта азиатская волна стихает на равнинах Западного Казахстана и Зауралья, так что в Европе центрально-азиатские гаплогруппы почти отсутствуют. Единственное пятно их высоких частот в Европе (рис. 17) - калмыки, мигрировавшие из Центральной Азии исторически недавно. Можно сделать вывод, что монгольские и другие миграции из Центральной Азии не привели к значительным изменениям генофонда Европы. Этот вывод подтверждается и анализом мтДНК (Балановская, Балановский, 2007). Так, суммарная доля восточно-евразийских гаплогрупп мтДНК у восточных славян не достигает даже двух процентов (табл. 2).

Совсем другой паттерн характерен для гаплогрупп **N1c**, **N1b** и **Q**, которые мы предлагаем называть северо-евразийскими. Их суммарная частота велика и в Сибири, и в Восточной Европе, но приурочены они не к степной, а к лесной полосе.

Карта западно-евразийской гаплогруппы **R1a** показывает ее доминирование в Восточной Европе и высокие частоты также в горах Алтая и Тянь-Шаня (где снижены частоты восточно-евразийских гаплогрупп). Такая «альтернативность» гаплогруппы **R1a** и восточно-евразийских гаплогрупп создает необычный кольцевой ареал **R1a**. Филогенетическая сеть изменчивости этой гаплогруппы (рис. 11Б) подтверждает существование древних кластеров STR-гаплотипов **R1a** в южносибирских популяциях. Можно предполагать, что первоначально **R1a** была распространена по степной полосе вплоть до Монголии. Эта зона согласуется с ареалом скифо-сарматских групп.

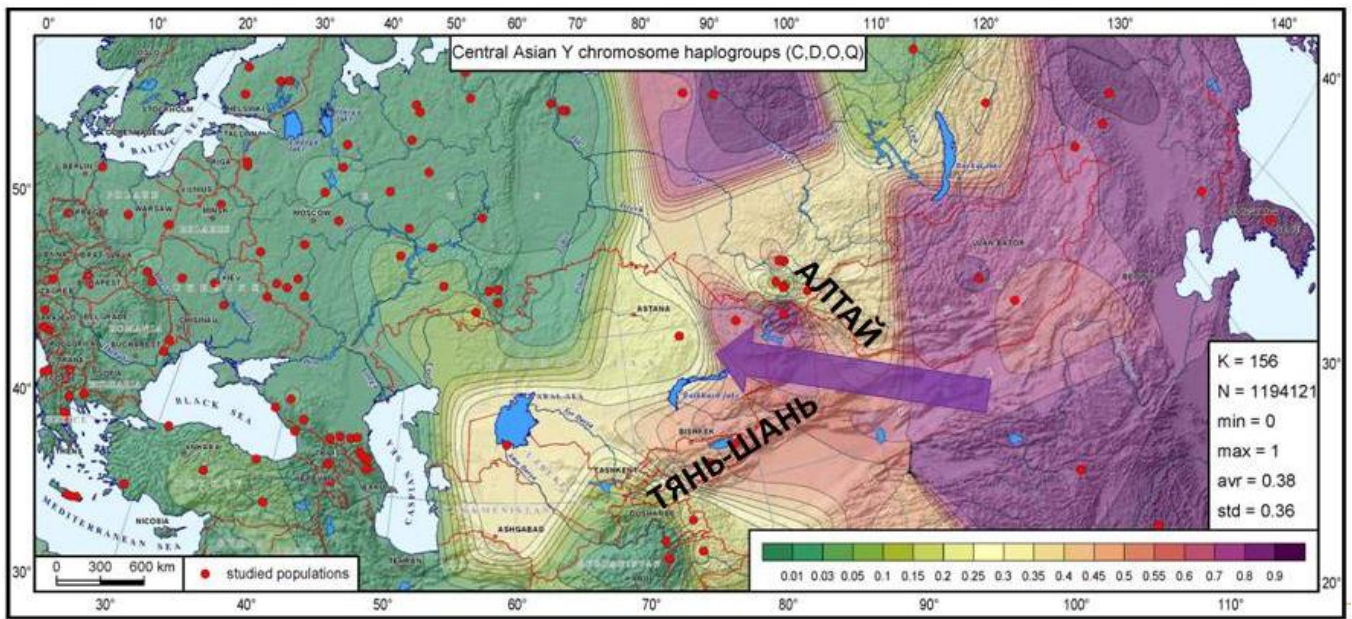


Рис. 17. Карта суммарной частоты восточно-евразийских гаплогрупп Y-хромосомы. Суммированы частоты гаплогрупп **C**, **D**, **O**, **Q**, наиболее частых в популяциях Центральной Азии. Надписи показывают положение горных хребтов Алтая и Тянь-Шаня. Стрелкой отмечено предполагаемое направление миграции через «Джунгарские ворота».

Более того, гаплогруппа **R1a** найдена в древней ДНК из родственной скифам курганной культуры (Keyzer et al., 2009). Позднее (с начала 1 тыс. н.э.) экспансия тюркских народов распространила монголоидные черты (Исмагулов, 1982) и восточно-евразийские гаплогруппы по *стенной* полосе, тогда как в *горах* сохранился предшествующий генофонд, маркируемый гаплогруппой **R1a**.

6. ФОРМИРОВАНИЕ ГЕНОФОНДА ПО ДАННЫМ О ДРЕВНЕЙ ДНК

Целью изучения генофондов является не только описание современной генетической структуры, но и реконструкция путей и механизмов ее формирования, то есть прослеживание изменчивости генофонда не только в географическом пространстве, но и во времени. Поэтому столь важны данные по древней ДНК – они предоставляют прямые данные о генофонде прошлых эпох, свободные от допущений, которые неизбежно сопутствуют реконструкции прошлых этапов истории генофонда по его современной структуре. Лишь количественные (объемы выборок) и качественные (проблема контаминации) ограничения в работе с древней ДНК мешают этому направлению произвести революцию в популяционной генетике человека (Paabo et al., 1989; Handt et al., 1996; Cooper et al., 2000, 2001; Noonan et al., 2006; Brotherton et al., 2007; Григоренко и др., 2009; Burbano et al., 2010).

В данном исследовании изучена древняя ДНК пяти популяций (табл. 3), покрывающих временной интервал в 7 тысяч лет (от мезолита до Нового времени) и представляющих различные регионы Европы (северо-восток, юго-восток, центр, север). Проблема контаминации решалась сотрудничеством с общепризнанной лабораторией древней ДНК (Australian center for ancient DNA), возглавляемой проф. Alan Cooper (Австралия), с независимым подтверждением результатов в лаборатории Johannes Gutenberg University, возглавляемой проф. Kurt Alt (Германия). Анализ полученных результатов по древней ДНК проведен в сравнении с созданной автором базой данных по современному населению (табл. 4), репрезентативность и обширность которой позволила отчасти компенсировать ограниченный объем «древних» выборок.

МЕЗОЛИТ СЕВЕРНОЙ ЕВРОПЫ. Южный Олений остров (7500 лет назад, табл. 3) является опорным памятником мезолита северной Европы. Обнаруженный спектр гаплогрупп мтДНК (самые частые **U4** и **C**, присутствуют также **U2e**, **U5a**, **J**, **H**) позволяет сделать два вывода. Первый касается сходства с другими мезолитическими популяциями Европы (Bramanti et al., 2009), в особенности по преобладанию гаплогруппы **U**. Во-вторых, выявленный баланс западно-евразийских и восточно-евразийских гаплогрупп характерен для современных популяций Западной Сибири и отчасти Урала. На графике главных компонент также четко выделились европейский, ближневосточный, уральский и сибирский кластеры, причем олениостровцы примыкают к уральскому. Карта генетических расстояний (рис. 18) тоже указывает, что современное население Урала и особенно Западной Сибири наиболее сходно с мезолитическим населением Северо-Восточной Европы. Учитывая многолетнюю антропологическую дискуссию о частичной монголоидности олениостровцев, можно заключить, что в мезолите миграции между Европой, Уралом и прилежащими районами Сибири были намного интенсивнее, чем в последующие эпохи. Вероятно, население востока Европы, Урала, Западной Сибири представляло единый континуум, что согласуется с зоной распространения уральских языков и отчасти с теорией «уральской расы» (Бунак, 1980; Моисеев, 1999; Перевозчиков, 2003).

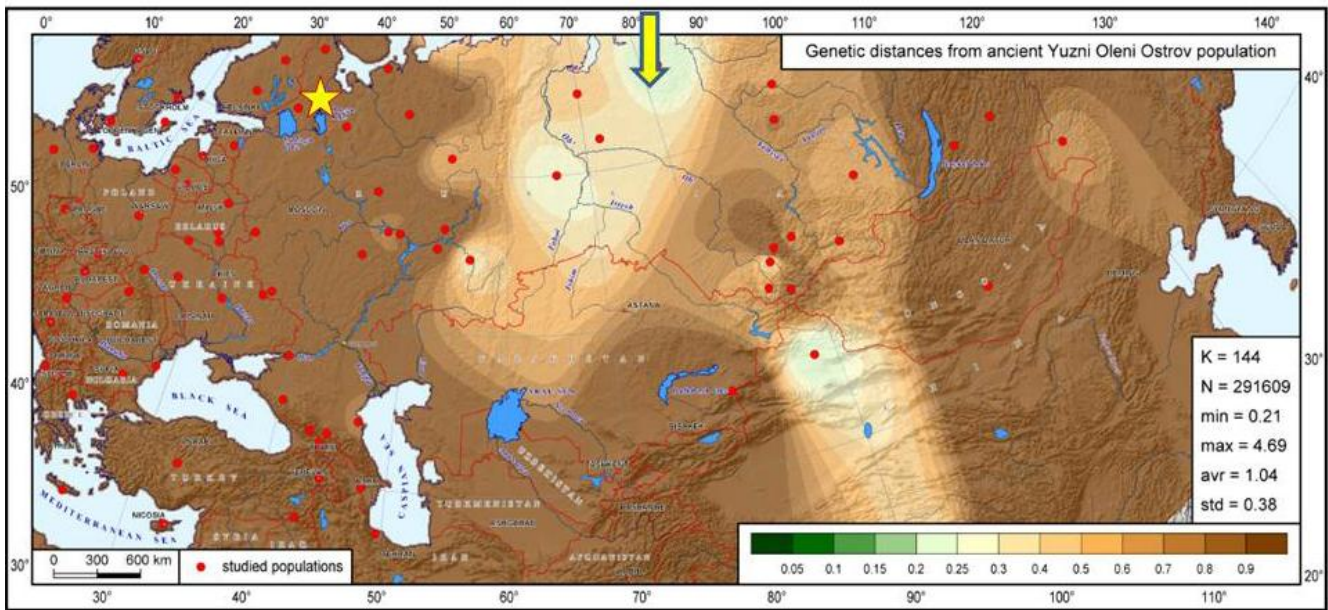


Рис. 18. Карта генетических расстояний от мезолитической популяции Южного Оленьего Острова до современного генофонда.

Звездочкой отмечено положение изученной древней популяции, а стрелкой – зона наибольшего сходства в современном генофонде.

НЕОЛИТ ЦЕНТРАЛЬНОЙ ЕВРОПЫ. Примерно в то же время, когда на северо-востоке Европы существовал этот мезолитический памятник, Центральная Европа уже была охвачена процессом неолитизации: возникла первая обширная земледельческая культура Европы - линейно-ленточной керамики (табл. 3). Предыдущее исследование древней ДНК носителей этой культуры (Haak et al., 2005) показало их резкое отличие от современного генофонда Европы (преобладание гаплогруппы **N1a**). Эта работа ответила на вопрос “что произошло с генофондом первых земледельцев после их появления в Европе?” (растворение в автохтонном генофонде). Чтобы ответить на следующий вопрос “откуда появились первые земледельцы?”, в нашем исследовании (Haak, Balanovsky, et al., 2010) проведен анализ дополнительных 22 образцов древней ДНК (табл. 3) и сопоставление с большим массивом современных популяций из нашей базы данных.

При анализе на уровне гаплогрупп были рассчитаны и картографированы генетические расстояния от популяции европейского неолита до современных популяций. Карта (рис. 19) демонстрирует, что наименьшие расстояния локализуются в северной части Ближнего Востока (северная Месопотамия, Закавказье, восточная Анатолия). Анализ на уровне гаплотипов также обнаружил наибольший процент совпадений в ближневосточных, а не в современных европейских популяциях. Хотя генофонд Ближнего Востока мог с неолита несколько измениться, все имеющиеся данные указывают, что именно эта зона была наиболее вероятным источником миграции для населения культуры линейно-ленточной керамики.

Этот результат имеет прямое отношение к многолетней дискуссии о происхождении генофонда Европы. Теория демической диффузии (Ammerman, Cavalli-Sforza, 1984) предполагает неолитическую волну заселения, заменившую палеолитический генофонд. Теория культурной диффузии, поддержанная данными по современной мтДНК (Richards et al., 2000), предполагает заимствование навыков земледелия без миграций земледельцев и сохранение палеолитического генофонда. Наши новые данные указывают на промежуточную схему: неолитическая миграция с Ближнего Востока имела место, но мигрантов было не столь много, чтобы кардинально изменить европейский генофонд.

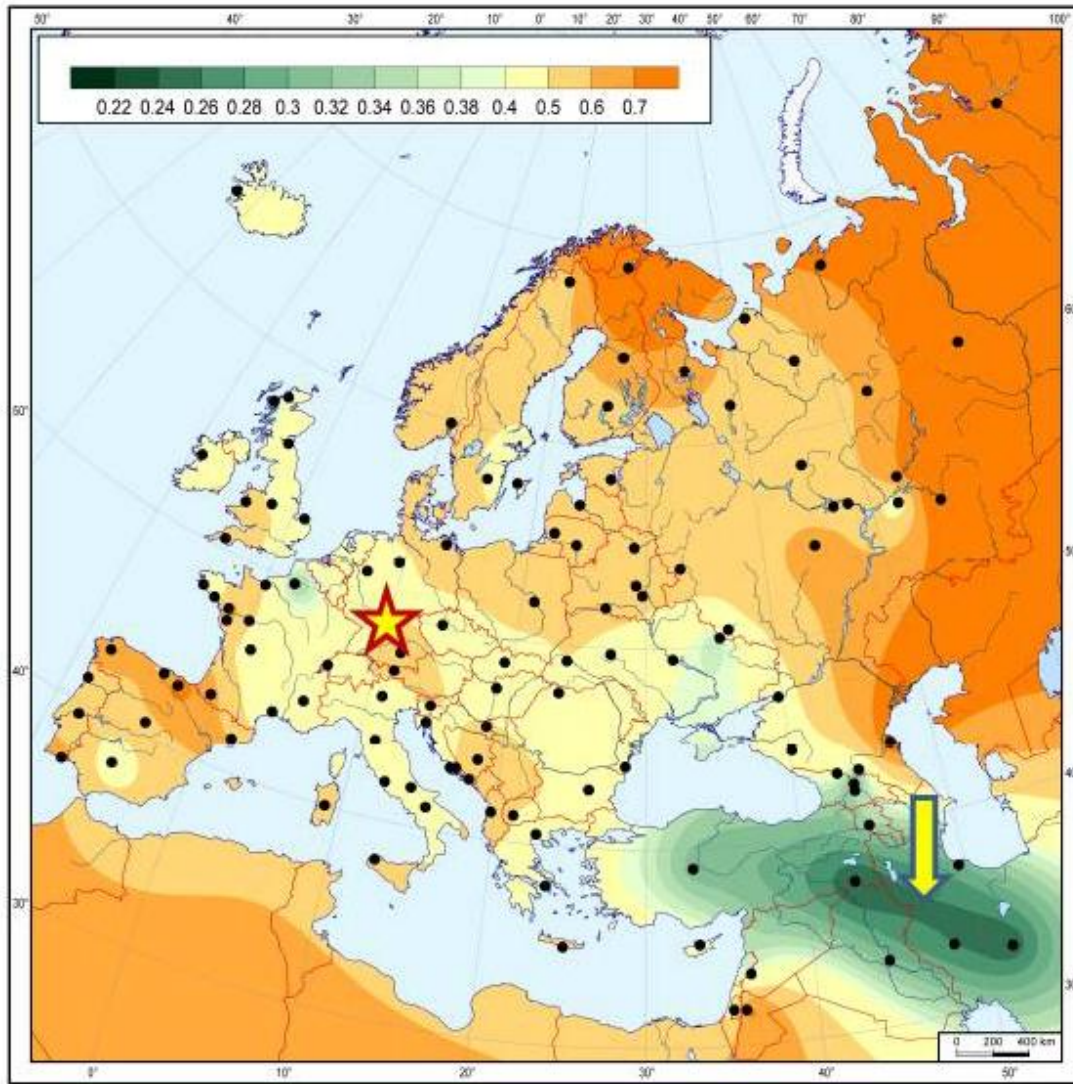


Рис. 19. Карта генетических расстояний от неолитической популяции Центральной Европы до 118 современных популяций Европы и Азии. Звездочкой отмечено положение изученной древней популяции, а стрелкой – зона наибольшего сходства в современном генофонде.

ЭПОХА РАННЕГО МЕТАЛЛА СЕВЕРНОЙ ЕВРОПЫ. Значительно более поздняя (Большой Олений Остров, 3500 лет назад, табл. 3) популяция севера Европы так же, как и мезолитический Южный Олений остров, несет в своем генофонде свидетельства сибирского влияния: выявлены гаплогруппы мтДНК - **C, U5a, D, Z, U4, T**. Однако, в отличие от мезолита (7500 лет назад), в этом более позднем населении Европы восточно-евразийские гаплогруппы преобладают. Вероятным источником миграции (согласно картам генетических расстояний до современных популяций) была уже не Западная, а Средняя Сибирь.

Территорию, населенную 3500 лет назад этой древней популяцией, сейчас занимают саамы. Но их спектр гаплогрупп полностью отличен. Стабильность генофонда саамов и их отличие от древней «большеостровской» популяции подтверждены данными по саамским образцам XVIII века (табл. 3). Можно заключить, что популяция эпохи раннего металла (Большого Оленьего острова) сформировалась в результате миграции из Сибири (вероятно, вдоль арктического побережья по тундровой зоне), но в дальнейшем исчезла, не оказав влияния на современный генофонд этой территории (Балановский и др., 2008).

СКИФЫ ЕВРОПЫ. Античное время (около 2500 лет назад) представлено в нашем исследовании данными по европейским скифам (табл. 3). Обнаруженный спектр гаплогрупп

мтДНК (в порядке убывания **T, U5a, H, I, D, A, C, F, U2e, U7**) свидетельствует о большом генетическом разнообразии и о преобладании западно-евразийских гаплогрупп. Другие авторы (Lalueza-Fox et al., 2003) обнаружили сходный генофонд у родственных скифам культур на территории Казахстана. Это указывает, что в скифскую эпоху «западно-евразийский» митохондриальный генофонд распространялся много дальше на восток, чем сейчас.

МОДЕЛЬ ЭТАПОВ ФОРМИРОВАНИЯ ГЕНОФОНДА ЕВРОПЫ. Анализ основных черт генофонда Европы на основе собственных и литературных данных по древней ДНК позволяет предложить следующую модель этапов его формирования.

А) Верхнепалеолитический этап состоял в первоначальном заселении с Ближнего Востока (около 40 тыс. лет назад) и затем в сосредоточении населения в южных рефугиумах в период максимального оледенения (26-18 тыс. лет назад).

Б) В мезолите популяции, снова распространившиеся на освободившиеся из-под ледника территории, характеризовались преобладанием гаплогруппы **U** мтДНК. Северо-восточная часть Европы в это время (7,5 тыс. лет назад) была генетически сходна с близлежащими частями Урала и Сибири, население которых также мигрировало вслед за отступавшим скандинавским ледниковым щитом (гаплогруппы **C** и **U4**).

В) Начиная с 7 тыс. лет назад по Европе начинают расселяться земледельческие неолитические культуры, мигрировавшие с Ближнего Востока через Балканы и характеризовавшиеся гаплогруппой **N1a** мтДНК. Затем их генофонд растворился в массе автохтонного населения, а навыки земледелия стали передаваться уже путем культурных заимствований. После этого (отчасти благодаря миграциям с Ближнего Востока), митохондриальный генофонд Европы претерпел существенные изменения: частоты гаплогрупп мтДНК **J, T** и особенно **H** резко выросли за счет повсеместного уменьшения частоты гаплогруппы **U** и исчезновения гаплогруппы **C** на востоке Европы.

Г) В эпоху раннего металла (3,5 – 2 тыс. лет назад) миграции из Сибири в Европу носили ограниченный характер, проникали по тундровой (возможно, лесной) полосе и почти не влияли на европейский генофонд. Миграции же из Европы в Сибирь, напротив, приобрели массовый характер и шли по степной полосе.

Д) В историческое время, степные миграции из Центральной Азии изменили лишь пограничный «буферный» генофонд (например, на территории Казахстана), но, за несколькими исключениями (калмыки, ногайцы), практически не повлияли на генофонд Европы. Зато прослеживается результат внутриевропейской миграции славян из Центральной в Восточную Европу с сохранением на севере Восточной Европы палеоевропейского генофонда.

Данная модель не претендует на полноту. Например, центр внимания смещен на восточную половину Европы. Ряд закономерностей можно интерпретировать не только в терминах миграций, но и в терминах «генетических общностей»: в мезолите Восточная Европа, Урал и Западная Сибирь могли обладать общим генофондом. Наконец, картина, вырисовывающаяся по генетическим данным, пока менее точна, чем синхронные реконструкции археологов или лингвистов и, несомненно, будет в дальнейшем корректироваться. Тем не менее, всеобщий интерес (в том числе генетиков, археологов и лингвистов) к формированию генофонда Европы оправдывает нашу попытку предложить эту комплексную модель, основанную в значительной мере на результатах данного исследования.

ВЫВОДЫ

1. Созданы базы данных по изменчивости Y-хромосомы (112 400 образцов из 2 474 популяций мира) и митохондриальной ДНК (132 600 образцов из 2 100 популяций мира). Включение собственных данных (4 289 образцов Y-хромосомы и 3 234 образца мтДНК) заполнило пробелы изученности генофондов населения Европы и Азии.
2. Изменчивость Y-хромосомы в народонаселении Европы структурирована не только по географической близости, но и по этнической принадлежности: популяции одного народа образуют в генетическом пространстве компактные «этнические облака». Количественно это выражается в резком преобладании межэтнической изменчивости ($G_{ST}=0.15$) над внутриэтнической ($G_{ST}=0.03$). Исключение составляют восточные и западные славяне, обладающие сходным генофондом вопреки обширности их ареала.
3. Впервые обнаружена неоднородность митохондриального генофонда народов Европы, находящая параллели в лингвистическом древе индоевропейских языков. Структурированность европейского генофонда подтверждается графиком многомерного шкалирования (наличие славянского, германского, кельтского, романского и балтского кластеров), а также связью генетических расстояний не только с географическими (коэффициент корреляции $r=0.37$), но и с лингвистическими расстояниями ($r=0.31$).
4. Впервые показана клинальная изменчивость гаплотипического разнообразия митохондриальной ДНК с юга на север Европы. Это убывание внутрипопуляционного разнообразия (с 0.99 до 0.95) объясняется действием дрейфа генов, связанным со сниженной плотностью населения на северо-востоке Европы.
5. Максимум межпопуляционных различий митохондриальной ДНК при сравнении лингвистических групп Европы выявлен у финно-угорских популяций, населяющих северо-восточную Европу. Средние генетические расстояния между популяциями финно-угорской группы ($d=0.42$) значительно превышают различия между популяциями германской ($d=0.01$), кельтской ($d=0.04$), славянской ($d=0.04$), романской ($d=0.05$) и даже тюркской ($d=0.24$) групп. Обнаруженное сочетание максимальной межпопуляционной и минимальной внутрипопуляционной изменчивости у финно-угорских народов является двумя сторонами одной медали - эффектов дрейфа генов в этих популяциях.
6. Выявлена четкая широтная структура русского генофонда. Карты главных компонент изменчивости гаплогрупп Y-хромосомы, график многомерного шкалирования и сравнение с другими типами маркеров подтверждают эту закономерность. Южные и центральные русские популяции характеризуются невысокими частотами гаплогрупп N1c-M178 и N1b-R43 и преобладанием гаплогруппы R1a1-M198 (более 50%). Для северных русских популяций и по Y-хромосоме, и по мтДНК характерно сходство с широким кругом народов Северной и Центральной Европы, что может быть связано с сохранением в этом ареале генофонда палеоевропейского населения.
7. Созданный картографический атлас изменчивости Y-хромосомы в народонаселении Евразии включил карты 79 гаплогрупп и выявил генетические границы - объективные географические зоны максимальной изменчивости на картах межпопуляционной изменчивости всей совокупности гаплогрупп. Основная генетическая граница разделяет генофонды Западной и Восточной Евразии и проходит от Кавказа до Средней Сибири.
8. Созданный картографический атлас изменчивости митохондриальной ДНК в народонаселении Евразии включил карты 82 гаплогрупп и, благодаря предложенному

методу объективного выделения географических типов гаплогрупп, выявил не только традиционные «западно-евразийский» и «восточно-евразийский» типы, но и третий новый тип, который можно назвать «южно-китайским».

9. Анализ взаимодействия европеоидного и монголоидного населения в обширной зоне степной полосы Евразии, проведенный с помощью картографического анализа, выявил лишь незначительное влияние центральноазиатского генофонда, ограниченное юго-восточными степными районами Европы. В русских популяциях заметный (выше 1-2%) «монгольский» компонент не детектируется ни по Y-хромосоме, ни по мтДНК.

10. Впервые показана параллельная эволюция генофонда и языков Северного Кавказа. Лингвистическое дерево практически совпадает с генетическими реконструкциями родства 11 кавказских народов, основанными как на частотах гаплогрупп, так и на STR гаплотипах 1500 образцов. Глоттохронология распада языков хорошо согласуется с генетическими датировками расхождения этносов. Наилучшее согласие датировок получено при использовании «генеалогической» скорости мутирования Y-STR маркеров.

11. Анализ древней ДНК проведен для пяти популяций, покрывающих временной диапазон в 7 тысяч лет. Впервые проведенный анализ древней ДНК мезолитических популяций северо-востока Европы подтвердил палеоантропологические данные о существенном монголоидном или «уралоидном» компоненте в их генофонде (Южный Оленеостровский могильник). Последующая миграция из Сибири угадала без значимого влияния на современный генофонд (эпоха раннего металла, Большой Олений остров). Анализ древней ДНК населения неолитической культуры линейно-ленточной керамики впервые доказал ближневосточное происхождение первых земледельцев Центральной Европы.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи, опубликованные в изданиях, рекомендованных ВАК.

1. **Balanovsky O.**, Rootsi S., Pshenichnov A., Kivisild T., Churnosov M., Evseeva I., Pocheshkhova E., Boldyreva M., Yankovsky N., Balanovska E., Villems R. Two sources of the Russian patrilineal heritage in their Eurasian context // *American Journal of Human Genetics*. 2008. V.82(1). P. 236-250.

2. **Balanovsky O.**, Dibirova Kh., Dybo A., Mudrak O., Frolova S., E. Pocheshkhova, Haber M., Platt D., Schurr T., Haak W., Kuznetsova M., Radzhabov M., Balaganskaya O., Romanov A., Zakharova T., Baranova E., Hernanz D. F. S., Zalloua P., Koshel S., Ruhlen M., Renfrew C., Wells R. S., Tyler-Smith Ch., Balanovska E., The Genographic Consortium. Parallel Evolution of Genes and Languages in the Caucasus Region // *Molecular biology and evolution*. 2011. V.28(10). P. 2905-2920.

3. Haak W., **Balanovsky O.**, Sanchez J.J., Koshel S., Zaporozhchenko V., Adler C.J., Der Sarkissian C.S.I., Brandt G., Schwarz C., Nicklisch N., Dresely V., Fritsch B., Balanovska E., Villems R., Meller H., Alt K.W., Cooper A. & The Genographic Consortium. Ancient DNA from European Early Neolithic Farmers Reveals Their Near Eastern Affinities // *PLoS Biology*. 2010. V.8. P. 1-16.

4. **Балановский О.П.**, Кошель С.М., Запорожченко В.В., Пшеничнов А.С., Фролова С.А., Кузнецова М.А., Баранова Е.Е., Теучеж И.Э., Кузнецова А.А., Ромашкина М.В., Утевская О.М., Чурносков М.И., Виллемс Р., Балановская Е.В. Эколого-генетический мониторинг в популяциях человека: гетерозиготность, гаплотипическое разнообразие мтДНК и генетический груз // *Генетика*. 2011. Т.47. №.11. С. 1523–1535.

5. **Балановский О.П.**, Дибирова Х.Д., Романов А.Г., Утевская О.М., Шанько А.В., Баранова Е.Г., Почешхова Э.А. Взаимодействие генофондов народов Кавказа и восточных славян по

данным о полиморфизме Y хромосомы // Вестник Московского университета. Серия XXIII. Антропология. 2011. №1. С. 69-75.

6. **Балановский О.П.**, Пшеничнов А.С., Фролова С.А., Васинская О.А., Дибирова Х.Д., Кузнецова М.А., Кошель С.М., Запорожченко В., Чурносков М.И., Атраментова Л.А., Утевская О., Тегако О.В., Почешхова Э.А., Микулич А.И., Виллемс Р., Балановская Е.В. Основные черты митохондриального генофонда восточных славян // Медицинская генетика. 2010. Т.9. №1 С.29-37.

7. **Балановский О.П.**, Бужилова А.П., Балановская Е.В. Русский генофонд. Геногеография фамилий // Генетика. 2001. Т. 37. № 7. С. 974-990.

8. Дибирова Х.Д., Балановская Е.В., Кузнецова М.А., Фролова С.А., Васинская О.А., Почешхова Э.А., Запорожченко В.В., Дружинина Е.Г., Пшеничнов А.С., Раджабов М.О., Теучеж И.Э., Схалыхо Р.А., Захарова Т.А., Евсеева И.В., Дубинецкая Е., **Балановский О.П.** Генетический рельеф Кавказа: четыре лингвистико-географических региона по данным о полиморфизме Y хромосомы // Медицинская генетика. 2010. Т.9. №10. С. 9-18.

9. Соловьева Д.С., Балановская Е.В., Кузнецова М.А., Васинская О.А., Фролова С.А., Почешхова Э.А., Евсеева И.В., Болдырева М.Н., **Балановский О.П.** Русский генофонд: геногеография ALU-инсерций (ACE, APOA1, B65, PV92, TRA25) // Молекулярная биология, 2010. Т. 44. №3. С.447-455

10. Балановская Е.В., Пежемский Д.В., Романов А.Г., Баранова Е.Е., Ромашкина М.В., Агджоян А.Т., Балаганский А.Г., Евсеева И.В., Виллемс Р., **Балановский О.П.** Генофонд Русского Севера: славяне? Финны? Палеоевропейцы? // Вестник Московского Университета. Серия XXIII "Антропология". 2011. №3. С 27-58.

11. Балаганская О.А., Балановская Е.В., Лавряшина М.Б., Исакова Ж.Т., Сабитов Ж.М., Фролова С.А., Романов А.Г., Дибирова Х.Д., Кузнецова М.А., Захарова Т.А., Pitchappan R., Урасин В.М., Балаганский А.Г., Баранова Е.Г., **Балановский О.П.** Полиморфизм Y хромосомы у тюркоязычного населения Алтае-Саян, Тянь-Шаня и Памира в контексте взаимодействия генофондов западной и восточной Евразии // Медицинская генетика. 2011. Т. 10. № 3. С. 12-22

12. Балаганская О.А., **Балановский О.П.**, Лавряшина М.Б., Кузнецова М.А., Романов А.Г., Дибирова Х.Д., Фролова С.А., Захарова Т.А., Баранова Е.Е., Сабитов Ж., Нимадава П., Балановская Е.В. Генетическая структура по маркерам Y хромосомы народов Алтая (России, Казахстана, Монголии) // Вестник Московского университета. Серия XXIII. Антропология. 2011. С. 25-39.

13. Myres N.M., Rootsi S., Lin A.A., Jārve M., King R.J., Kutuev I., Cabrera V.M., Khusnutdinova E.K., Pshenichnov A., Yunusbayev B., **Balanovsky O.**, Balanovska E., Rudan P., Baldovic M., Herrera R.J., Chiaroni J., Cristofaro J.D., VILLEMS R., Kivisild T. and Underhill P.A. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe // European Journal of Human Genetics. 2011. V.19. №1. P 95-101.

14. Javed A., Mele M., Pybus M., Zalloua P., Haber M., Comas D., Netea M.G., **Balanovsky O.**, Balanovska E., Jin L., Yang Y., ArunKumar G.P., Pitchappan R., Bertranpetit J., Calafell F., Parida L., The Genographic Consortium. Recombination Networks as Genetic Markers in A Human Variation Study of the Old World // Human Genetics. 2012. DOI 10.1007/s00439-011-1104-8.

15. Tambets K, Rootsi S, Kivisild T, Help H, Serk P, Loogvali E-L, Tolk H-V, Reidla M, Metspalu E, Pliss L, **Balanovsky O**, Pshenichnov A, Balanovska E, Gubina M, Zhadanov S, Osipova L, Damba L, Voevoda M, Kutuev I, Bermisheva M, Khusnutdinova E, Gusar V, Grechanina E, Parik J, Pennarun E, Chaventre A, Moisan J-P, Barac L, Pericic M, Rudan P, Terzic R, Mikarezi I, Krumina A, Baumanis V, Beckman L, VILLEMS R. The western and eastern roots of the extreme European genetic outliers - the origin of mtDNAs and Y-chromosomes of the Saami // American Journal of Human Genetics. 2004. V. 74. № 4. P. 661-82.

16. Rootsi S, Magri C, Kivisild T, Benuzzi G, Help H, Bermisheva M, Kutuev I, Barac L, Pericic M, **Balanovsky O**, Pshenichnov A, Dion D, Grobei M, Zhivotovsky L, Battaglia V, Achilli A, Al-Zahery A, Parik J, King R, Cinnioglu C, Khusnutdinova E, Rudan P, Balanovska E, Scheffrahn W, Simonescu M, Brehm A, Goncalves R, Rosa A, Moisan J-P, Chaventre A, Ferak V, Furedi S, Oefner P, Shen P, Beckman L, Mikerezi I, Terzic R, Primorac D, Cambon-Thomsen A, Krumina A, Torroni A, Underhill P, Santachiara-Benerecetti S, Villems R, Semino O. Phylogeography of Y-Chromosome Haplogroup I Reveals Distinct Domains of Prehistoric Gene Flow in Europe // *American Journal of Human Genetics*. 2004. V. 75. № 1. P. 128-37.

17. Loogvali EL, Roostalu U, Malyarchuk BA, Derenko MV, Kivisild T, Metspalu E, Tambets K, Reidla M, Tolk HV, Parik J, Pennarun E, Laos S, Lunkina A, Golubenko M, Barac L, Pericic M, **Balanovsky OP**, Gusar V, Khusnutdinova EK, Stepanov V, Puzyrev V, Rudan P, Balanovska EV, Grechanina E, Richard C, Moisan JP, Chaventre A, Anagnou NP, Pappa KI, Michalodimitrakis EN, Claustres M, Golge M, Mikerezi I, Usanga E, Villems R. Disuniting Uniformity: A Pied Cladistic Canvas of mtDNA Haplogroup H in Eurasia // *Molecular biology and evolution*. 2004. V. 21. № 11. P. 2012–2021.

18. Behar DM, Metspalu E, Kivisild T, Achilli A, Hadid Y, Tzur S, Pereira L, Amorim A, Quintana-Murci L, Majamaa K, Herrstadt C, Howell N, **Balanovsky O**, Kutuev I, Pshenichnov A, Gurwitz D, Bonne-Tamir B, Torroni A, Villems R, Skorecki K. The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event // *American Journal of Human Genetics*. 2006. V. 78. № 3. P. 487-497.

19. Behar DM, Rosset S, Blue-Smith J, **Balanovsky O**, Tzur S, Comas D, Mitchell RJ, Quintana-Murci L, Tyler-Smith C, Wells RS; & The Genographic Consortium. The Genographic project public participation mitochondrial DNA database // *PLoS Genetics*. 2007. V. 3. № 6. P. 104.

20. Quintana-Murci L., Quach H., Harmant C., Luca F., Massonnet B., Patin E., Sica L., Mougouma-Daouda P., Comas D., Tzur S., **Balanovsky O.**, Kidd KK., Kidd JR., van der Veen L., Hombert JM., Gessain A., Verdu P., Froment A., Bahuchet S., Heyer E., Dausset J., Salas A., Behar DM. Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers // *Proceedings of the National Academy of Sciences of the USA*. 2008. V. 5. № 105(5). P. 1596-1601.

21. Quintana-Murci L., Harmant C., Quach H., **Balanovsky O.**, Zaporozhchenko V., Bormans C., van Helden P.D., Hoal E.G., Behar D.M. Strong maternal Khoisan contribution to the South African coloured population: a case of gender-biased admixture // *American Journal of Human Genetics*. 2010. V. 9. № 86(4). P. 611-620.

22. Underhill P., Myres N., Rootsi S., Metspalu M., Zhivotovsky L., King R., Lin A., Chow C., Semino O., Battaglia V., Kutuev I., Jarve M., Chaubey G., Ayub Q., Mohyuddin A., Mehdi Q., Sengupta S., Rogaev E., Khusnutdinova E., Pshenichnov A., **Balanovsky O.**, Balanovska E., Jeran N., Augustin D., Baldovic M., Herrera R., Thangaraj K., Singh V., Singh L., Majumder P., Rudan P., Primorac D., Villems R., Kivisild T. Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a // *European Journal of Human Genetics*. 2010. V. 18(4). P. 479-484.

23. Melé M., Javed A., Pybus M., Zalloua P., Haber M., Comas D., Netea M.G., **Balanovsky O.**, Balanovska E., Jin L., Yang Y., Pitchappan R.M., Arunkumar G., Parida L., Calafell F., Bertranpetit J. & The Genographic Consortium . Recombination Gives a New Insight in the Effective Population Size and the History of the Old World Human Populations // *Molecular biology and evolution*. 2012. V. 29, № 1. P. 25-30.2011.

24. Haber M., Youhanna S.C., **Balanovsky O.**, Saade S., Martínez-Cruz B., Ghassibe-Sabbagh M., Shasha N., Osman R., El Bayeh H., Koshel S., Zaporozhchenko V., Balanovska E., Soria-Hernanz

D.F., Platt D.E., Zalloua P.A. mtDNA Lineages Reveal Coronary Artery Disease-Associated Structures in the Lebanese Population // *Annals of Human Genetics*. 2012. V. 76, № 1. P. 1-8.

25. Балановская Е.В., Нурбаев С.Д., **Балановский О.П.**, Почешхова Э.А., Боровинских А.А., Гинтер Е.К. Геногеографический анализ подразделенной популяции. I. Генофонд адыгов в системе кавказских генофондов // *Генетика*. 1999. Т. 35. №6. С.818-830.

26. Болдырева М.Н., Алексеев Л.П., Хаитов Р.М., Гуськова И.А., Богатова О.В., Янкевич Т.Э., Хромова Н.А., Балановская Е.В., **Балановский О.П.**, Пшеничнов А.С., Целинская И.Н., Кашенин М.Н., Ганичева Л.Л., Поздеева О.С., Евсеева И.В., Сароянц Л.В. HLA-генетическое разнообразие населения России и СНГ. I. Русские // *Иммунология*. 2005. №5. Т26. С. 260-263

27. Балановская Е.В., Соловьева Д.С., **Балановский О.П.**, Чурносков М.И., Сорокина И.Н., Евсеева И.В., Аболмасов Н.Н., Почешхова Э.А., Серегин Ю.А., Пшеничнов А.С. "Фамильные портреты" пяти русских регионов // *Медицинская генетика*. 2005. № 1. С. 2-10.

28. Балановская Е.В., Романов А.Г., **Балановский О.П.** Однофамильцы или родственники? Подходы к изучению связи между гаплогруппами Y-хромосомы и фамилиями // *Молекулярная биология*. 2011. Т.45. №3. С.1-13.

29. Behar D.M., Yunusbayev B., Metspalu M., Metspalu E., Rosset S., Parik J., Rootsi S., Chaubey G., Kutuev I., Yudkovsky G., Khusnutdinova E.K., **Balanovsky O.**, Semino O., Pereira L., Comas D., Gurwitz D., Bonne-Tamir B., Parfitt T., Hammer M.F., Skorecki K., Villems R. The genome-wide structure of the Jewish people // *Nature*. 2010. V. 8. № 466(7303). P. 238-242.

Статьи в других изданиях и монография.

30. **Балановский О.П.** Генофонд высоких широт Евразии или откуда пришли саамы? Путь на север: окружающая среда и самые ранние обитатели Арктики и Субарктики (материалы международной конференции) / Под ред. А.А. Величко, С.А. Васильева. Москва, Институт географии РАН. 2008. С. 277-282.

31. **Балановский О.П.**, Тегак О.В. Генофонд белорусов по данным о трех типах генетических маркеров - аутосомных, митохондриальных, Y хромосомы. "Актуальные вопросы антропологии". Минск. 2008. Т. 2. С. 53-65.

32. **Balanovsky O.** Human Genetics and Neolithic Dispersals: What's New? The East European Plain on the Eve of Agriculture / edited by P.M. Dolukhanov, G.R. Sarson and A.M. Shukurov / *BAR International Series* 1964. 2009. P. 235-246.

33. **Балановский О.П.** Восток Европы в свете современной генетики. Человек в культурной и природной среде: труды Третьих антропологических чтений к 75-летию со дня рождения академика В.П. Алексеева. Москва, Наука. 2007. С. 267-274.

34. **Балановский О.П.** Пространственная структура русского генофонда (сравнительный геногеографический анализ) // *Антропология на пороге III тысячелетия. Избранные доклады*. Москва. 2004. Т.1. С. 390- 416.

35. **Балановский О.П.** Пути формирования генофонда Европы по данным об изменчивости Y хромосомы // "Человек: его биологическая и социальная история". 2010. Т. 1. С. 207-213.

36. **Balanovsky O.**, Pocheshkhova E, Pshenichnov A, Solovieva D, Kuznetsova M, Voronko O, Churnosov M, Tegako O, Atramentova L, Lavryashina M, Evseeva I, Borinska S, Boldyreva M, Dubova N, Balanovska E. Is spatial distribution of the HIV-1-resistant CCR5Delta32 allele formed by ecological factors? // *Journal of Physiological Anthropology and Applied Human Science*. 2005. V. 24. № 4. P. 375-382.

37. Балановская Е.В., **Балановский О.П.** Русский генофонд на Русской равнине. 2007. М.: Луч, 416 с.

Список сокращений:

БД - база данных;
 ГВС1, ГВС2 – первый и второй гипервариабельный сегменты мтДНК;
 ЛПГЧ – лаборатория популяционной генетики человека ФГБУ «МГНЦ» РАМН;
 мтДНК – митохондриальная ДНК человека;
 ПДРФ – полиморфизм длин рестрикционных фрагментов;
 SNP – однонуклеотидный полиморфизм (single nucleotide polymorphism);
 STR – короткие тандемные повторы (short tandem repeats);
 Y-base - база данных по изменчивости Y-хромосомы в популяциях мира;
 MURKA - база данных по изменчивости мтДНК в популяциях мира;
 GeneGeo – программное обеспечение картографо-статистического анализа генофонда.

Благодарности: Автор глубоко признателен своим учителям Е.В. Балановской и Р. Виллемсу, которые ввели его в область геногеографии и филогеографии; своему научному консультанту Н.К. Янковскому, который в течение многих лет стимулировал написание диссертационной работы; благодарит за содействие коллективы Эстонского биоцентра и Медико-генетического научного центра РАМН; признателен А.С. Пшеничному, В.В. Запорожченко, Р.С. Сычеву, взявших на себя труд наполнения баз данных по мтДНК и Y-хромосоме; рад случаю упомянуть своих прекрасных аспирантов Х.Д. Дибирову и О.А. Балаганскую; благодарит коллег Э.А. Почешхову, Л.А. Атраментову, М.Б. Лавряшину, М.И. Чурносова, W. Naak, С. Der Sarkisyan, А.В. Дыбо, О.А. Мудрака, S. Roots, В.И. Хартановича, А.П. Бужилову и многих других, в сотрудничестве с которыми собирались образцы или разрабатывались вопросы их анализа. Исследования стали возможны благодаря грантам, выполнявшимся под руководством автора: трем инициативным и шести экспедиционным грантам РФФИ, гранту РГНФ, гранту программы «Молекулярная и клеточная биология» и гранту международного проекта Genographic.