Следы взаимодействия славянских, финноязычных и тюркоязычных народов в геномах русских Волго-Окского междуречья

В популяциях Волго-Окского междуречья и Рязанской области исследовали следы взаимодействия генофондов финноязычных, славянских и тюркоязычных народов методом моделирования предковых компонент геномов (изучено 248 геномов представителей 9 этносов). Оказалось, что из финноязычных народов лишь предковые компоненты этнических групп Мордовии распространены во всех популяциях региона независимо от их языковой принадлежности. Генофонды русских популяций включают 80% собственной предковой компоненты, 19% вклада финноязычных народов и 1% центральноазиатского влияния. В генофонде татар отсутствует собственная предковая компонента – он является комбинацией всех выявленных предковых компонент популяций региона. Обнаружены особенности фармакогенетического ландшафта Волго-Окского междуречья, который резко отличается от ландшафта по селективно-нейтральным аллелям. Наиболее близки по фармакогенетическому статусу к рязанским русским популяции Мордовии, несколько меньше — русские Калужской, Смоленской и Костромской областей. Сходные по селективно-нейтральным геномам рязанские и нижегородские популяции резко различаются по фармакогенетическому статусу.

Генетическая история русского народа складывалась при взаимодействии дославянского финно-угорского населения, проживающего на Восточноевропейской равнине, с пришедшими сюда славянскими племенами, впоследствии древнерусское население испытало нашествие Золотой орды. Таким образом, в современном генофонде отражается взаимодействие трех генетических пластов – дославянского (финно-угорского), славянского и золотоордынского (тюркского). Поиску следов этих генетических пластов посвящено исследование, проведенное в коллективе проф. Е.В.Балановской (Медико-генетический научный центр, Институт общей генетики РАН) и опубликованное в журнале «Вестник РГМУ». Для решения поставленной задачи авторы постарались выбрать такой узловой регион, где можно ожидать наиболее интенсивное взаимодействие финноязычного, славянского и тюркского генетических пластов. Таким регионом, несомненно, является Волго-Окское междуречье в целом и Рязанская область в частности.

Со второй половины I тысячелетия н.э. на эти земли, населенные финноязычными и отчасти балтскими племенами, начинают проникать славянские племена. В начале XI века здесь возникает Муромское княжество, а с середины XII века оно разделяется на два – собственно Муромское и Рязанское со столицей в Старой Рязани. В 1237 г. Рязанское княжество стало первой жертвой монгольского нашествия под предводительством Батыя, набеги и разорения рязанской земли продолжались более 350 лет, в том числе и после того, как Рязанское княжество перешло в подчинение Москве. А еще до нашествия Золотой орды рязанцы ходили в поход на Волжскую Булгарию, так что их взаимодействие с тюркским миром началось еще до XII века. Кроме того, Рязанский край был форпостом, который граничил с Диким полем. История взаимодействия славянских и финноязычных племен еще более долгая.

Несмотря на то, что сегодня развиваются методы анализа ДНК из ископаемых останков, и это позволяет получить непосредственную информацию о древних генофондах, количество такого материала всегда ограничено. Особенно это касается населения, которое, как славяне, практиковало кремацию умерших. Поэтому важным источником для реконструкции генетической истории остаются исследования современного населения. В данной работе поставлена цель выявить в аутосомном генофонде современных русских популяций — в узловом регионе Волго-Окского междуречья — генетический след взаимодействия финноязычных, славянских и тюркоязычных народов. Вторая, более прикладная, задача – изучить фармакогенетический ландшафт региона. Это необходимо для разработки протоколов по подбору оптимальных препаратов и их доз в зависимости от генотипа пациента, что составляет основу персонифицированной медицины. На сегодня основные мировые исследования ориентированы на население Западной Европы и мало применимы к популяциям России, отличающимся огромным генетическим разнообразием.

Были исследованы геномы 20 индивидов из четырех районов Рязанской области, которые, в соответствии с критериями популяционно-генетических исследований, на протяжении не менее трех поколений проживали в данном районе и идентифицировали себя с русскими. Количество образцов в выборке для полногеномного исследования, включающего в себя анализ 4.5 млн SNP, более чем достаточно. ДНК генотипировали по полногеномной панели 4.5 млн SNP маркеров с помощью биочипа Infinium OmniExome BeadChip Kit (Illumina). Для сравнения были привлечены русские популяции из Тверской, Костромской, Смоленской, Калужской, Орловской, Тамбовской и Нижегородской областей. Финноязычные популяции были представлены народами Волго-Уральского региона (этнические группы Мордовии, марийцы и удмурты), а также для сравнения — западной финноязычной ветвью (южными карелами). Тюркоязычные популяции Волго-Уральского региона представляли казанские татары и чуваши, для сравнения привлечены астраханские и ставропольские ногайцы. Для выявления монгольского генетического следа приведены данные о геномах шести родовых групп монголоязычных калмыков.

Основным методом полногеномного исследования стал анализ ADMIXTURE, который на основе моделирования дает количественную оценку вклада разных предковых компонент в геномы. Число предковых компонент k задает исследователь, а

программа строит модель вклада каждой предковой компоненты геном (обозначаемых на диаграмме каждая своим цветом) в каждый индивидуальный. Модели при разных k создаются независимо друг от друга. При усреднении вклада каждой компоненты в геномы оценивается ее вклад в генофонд популяции. Анализ ADMIXTURE провели для 248 геномов из 47 популяций 9 этносов: 104 генома из русских популяций, 81 геном четырех финноязычных народов, 47 геномов трех тюркоязычных народов и 16 геномов монголоязычных калмыков. При этом последовательно задавали величины k от 2 до 14. Наиболее информативными для решения основной задачи оказались две модели: при k=3 и k=7 (Рис. 1).

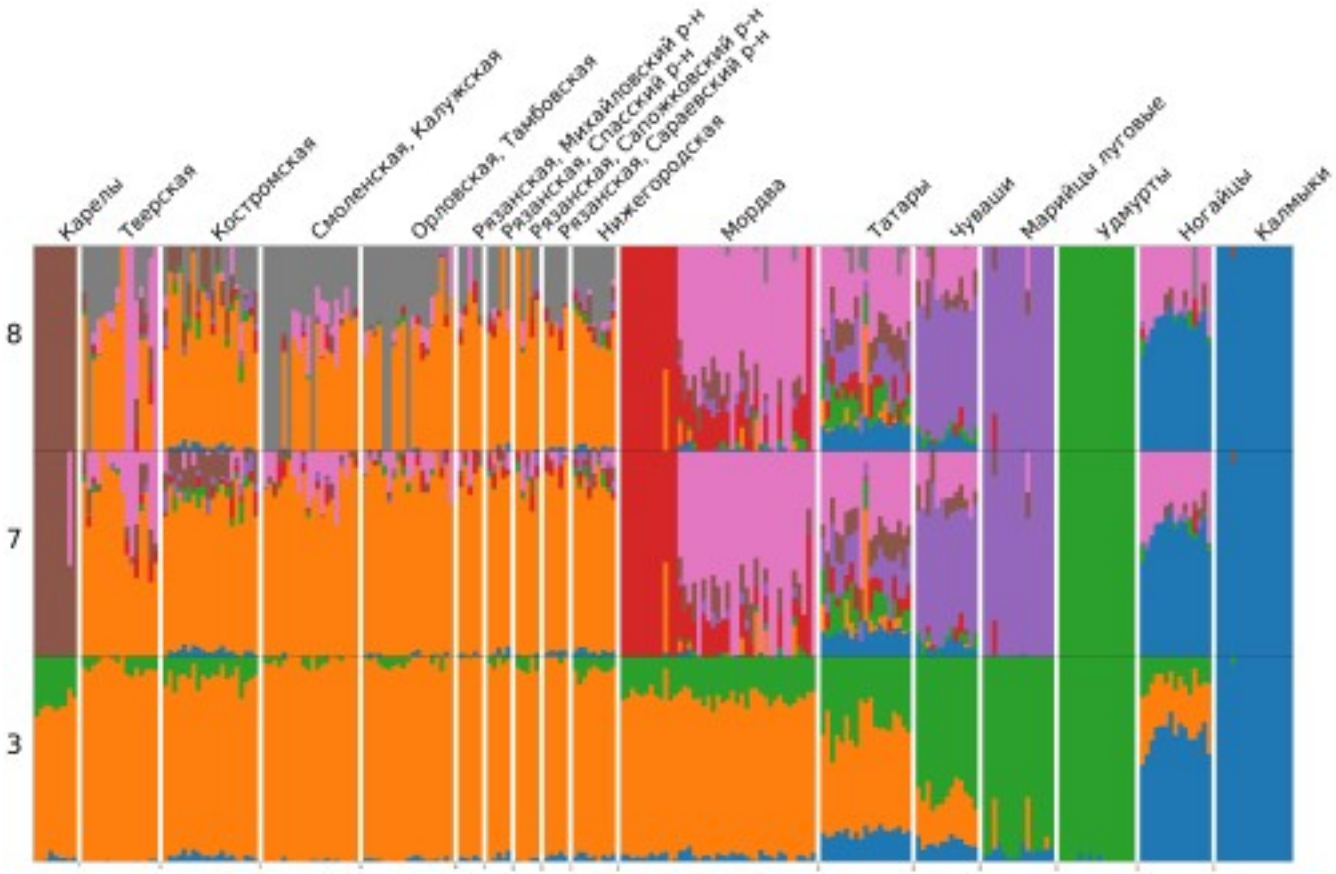
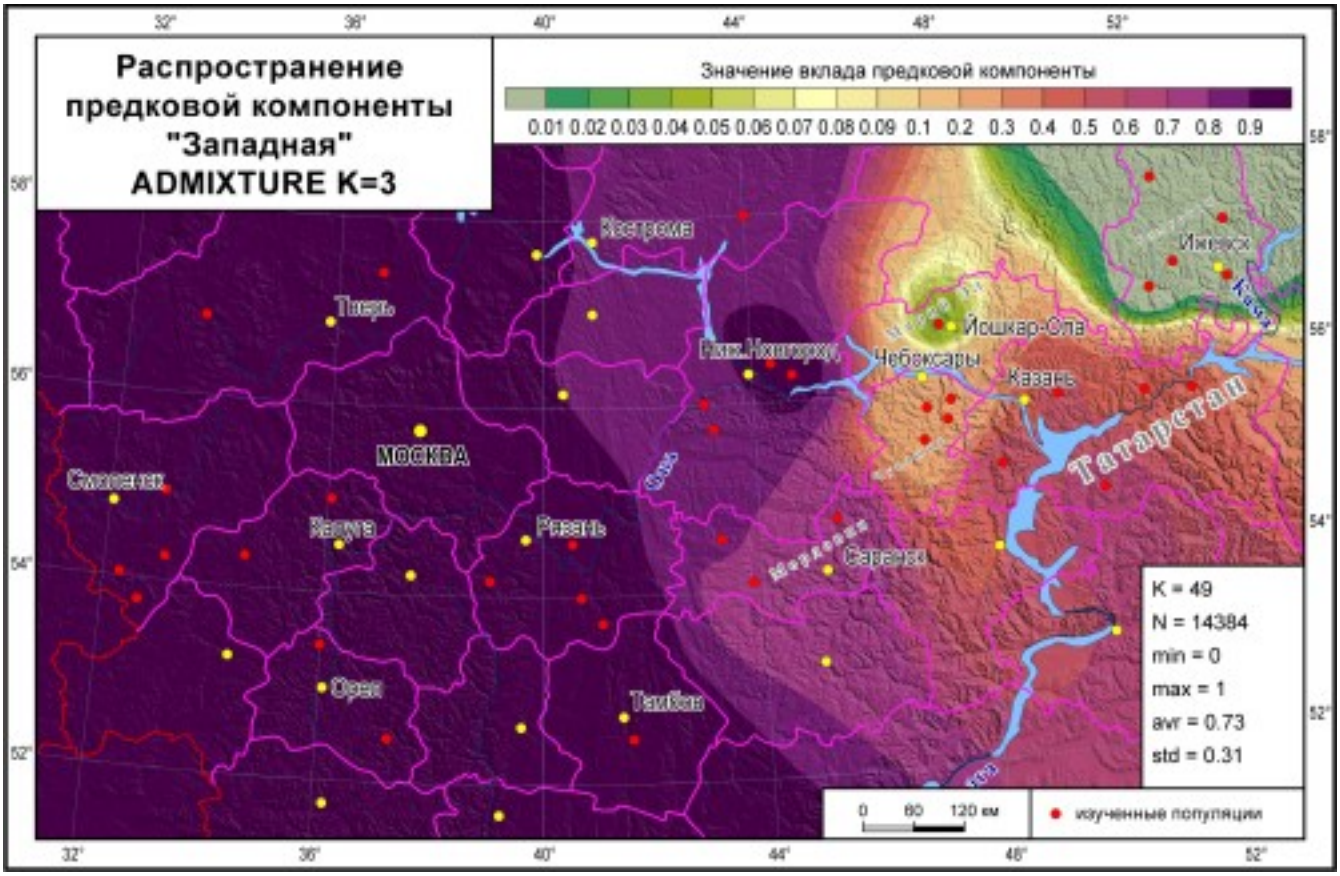
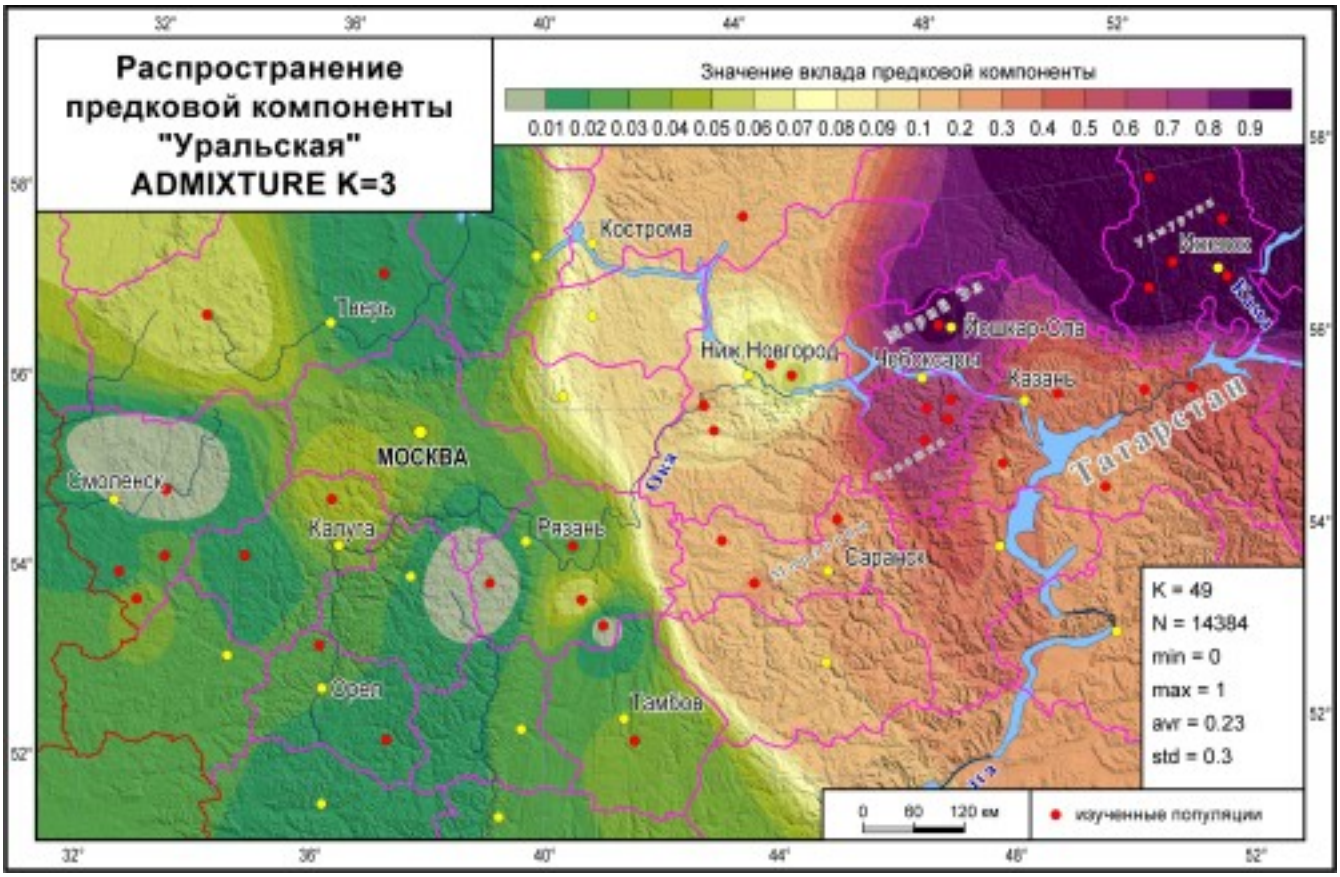
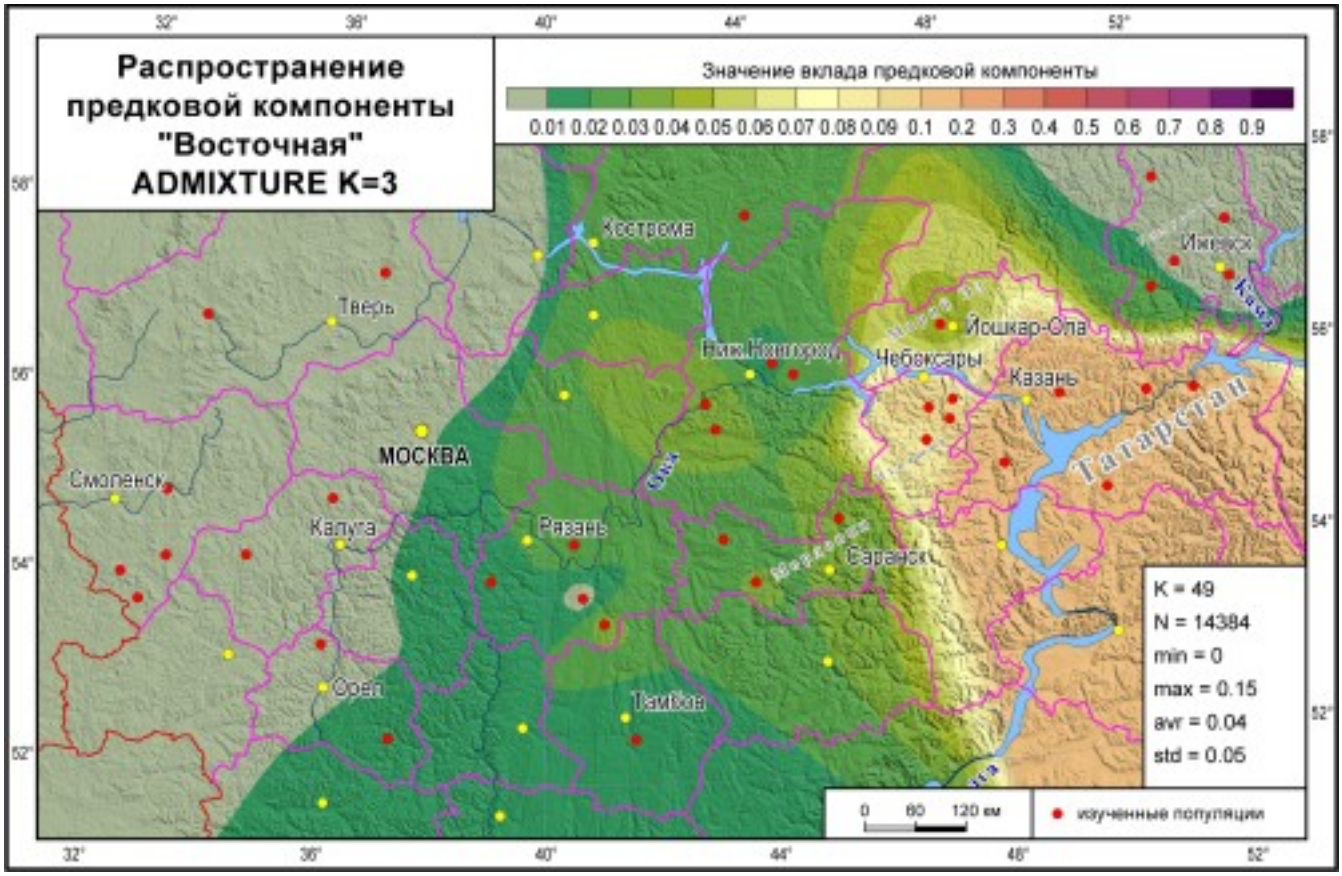


Рис. 1 – Вклад предковых компонент ADMIXTURE в индивидуальные геномы изученных популяций при k=3, k=7, k=8 предковых компонент.

Уровень k=3 демонстрирует три предковых компоненты, которые можно условно назвать «западная», «уральская» и «восточная». Авторы отмечают, что они не укладываются в ложе лингвистических групп. Для анализа пространственной изменчивости предковых компонент созданы геногеографические карты. Их вклад в генофонды популяций (в процентах) соответствует цвету на шкале: от зеленых цветов (малая доля) до фиолетовых (большая доля).


Распространение предковой компоненты "Западная" ADMIXTURE K=3
Значение вклада предковой компоненты
0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
Кострома
Тверь
Ник.Новгород
Йошкар-Ола
Чебоксары
Казань
Татарстан
МОСКВА
Смоленск
Калуга
Рязань
Саранск
Орел
Тамбов
K = 49
N = 14384
min = 0
max = 1
awr = 0.73
std = 0.31
0 60 120 км
изученные популяции


Распространение предковой компоненты "Уральская" ADMIXTURE K=3
Значение вклада предковой компоненты
0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
Кострома
Тверь
Ник.Новгород
Йошкар-Ола
Чебоксары
Казань
Татарстан
МОСКВА
Смоленск
Калуга
Рязань
Саранск
Орел
Тамбов
K = 49
N = 14384
min = 0
max = 1
awr = 0.23
std = 0.3
0 60 120 км
изученные популяции

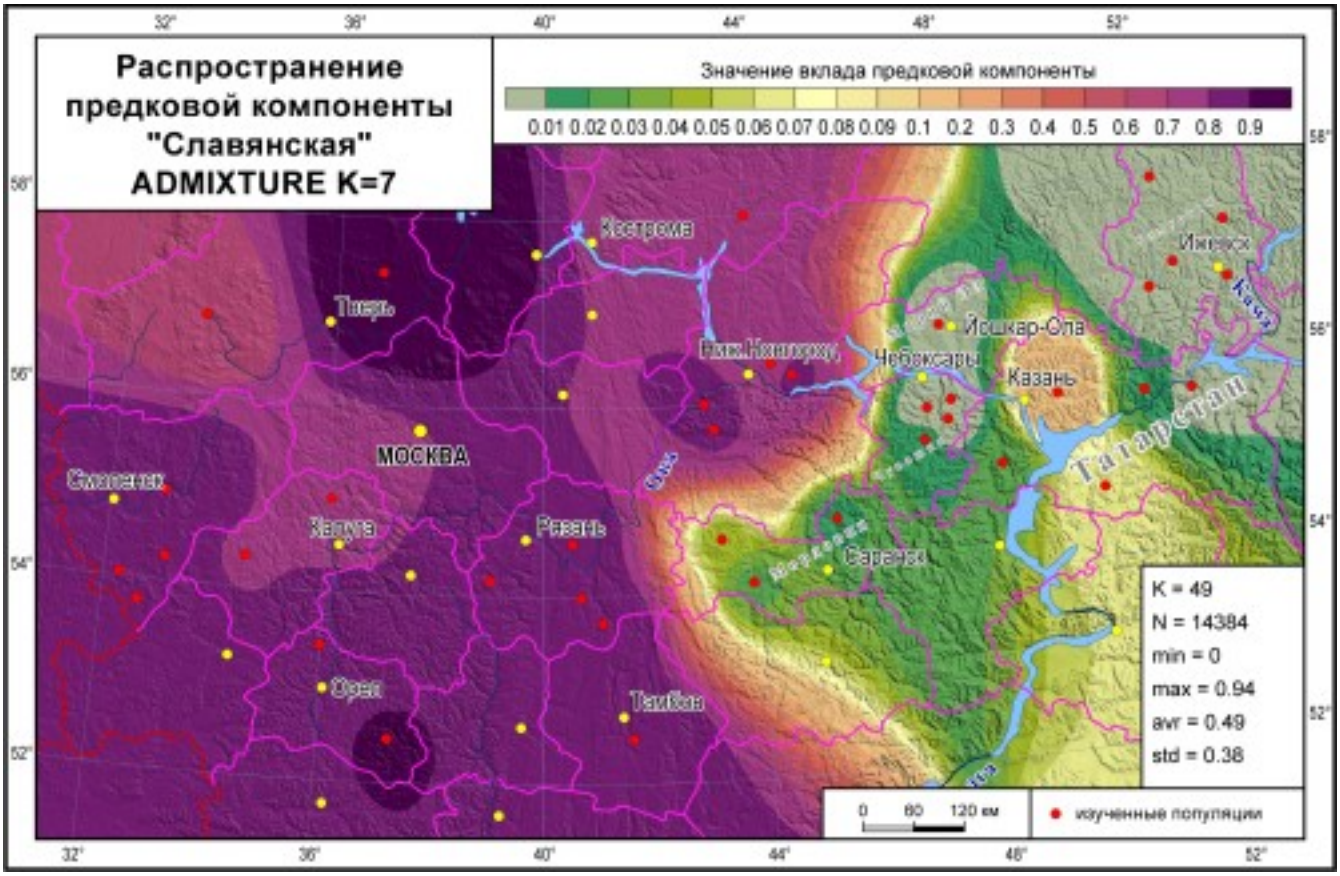


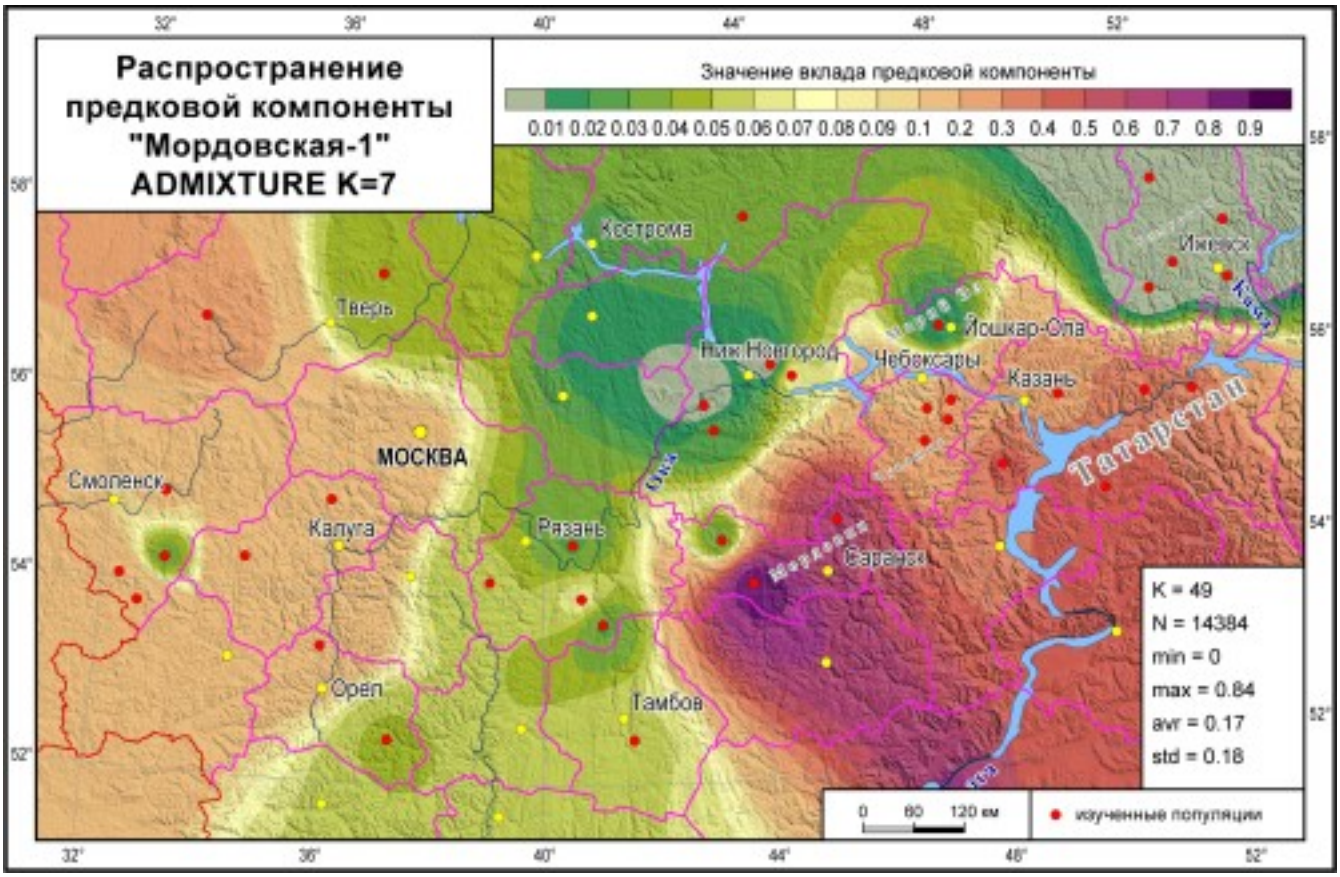
«Западная» предковая компонента (Рис. 2А) преобладает не только во всех русских популяциях (95%), но также и в финноязычных популяциях карел (75%) и Мордовии (78%). Более того, она составляет существенную часть генофондов тюркоязычных народов: у казанских татар более половины генофонда (52%) и около четверти генофонда ногайцев (25%) и чувашей (23%).

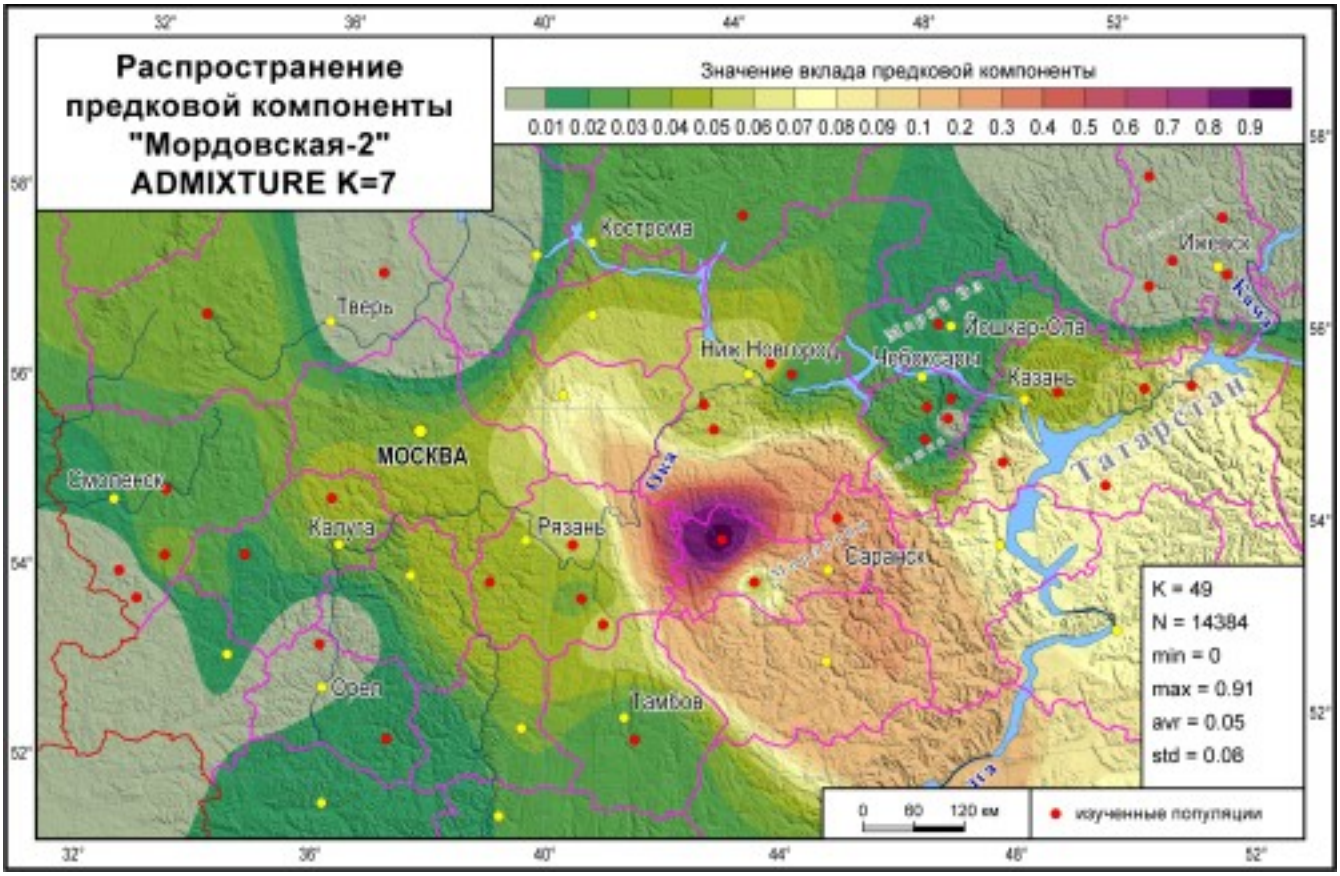
«Уральская» предковая компонента (Рис. 2Б) составляет практически весь генофонд удмуртов (99%) и марийцев (91%). Но высоких значений она достигает также и у тюркоязычных народов, составляя две трети генофонда чувашей (67%) и треть генофонда татар (34%). Меньший, но ощутимый вклад ее выявлен у карел (24%) и в Мордовии (19%). Средний вклад в русские популяции мал (4%) с максимумами в Костромской и Нижегородской областях.

«Восточная» предковая компонента (Рис. 2В) оставляет весь генофонд шести родовых групп калмыков, поэтому ее можно использовать для оценки влияния Центральной Азии на европейский генофонд. Значительной величины она достигает у ногайцев (62%), что подтверждает ее оценку как «центральноазиатской». В Поволжье наибольшее центральноазиатское влияние обнаружено у казанских татар (14%) и чувашей (9%). В остальных популяциях роль «восточной» компоненты мала и не превышает 5%, средний вклад в русские популяции составляет всего 1%.

При k=7 предковых компонент картина меняется. Авторы описывают ее, следуя традиции, условно именуя все компоненты по тем популяциям, в геномах которых выявлен их наибольший вклад. На картах представлено распространение трех наиболее информативных для целей исследования предковых компонент.


Распространение предковой компоненты "Славянская" ADMIXTURE K=7
Значение вклада предковой компоненты
0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
Кострома
Тверь
Йошкар-Ола
Чебоксары
Казань
МОСКВА
Смоленск
Калуга
Рязань
Орел
Тамбов
Татарстан
K = 49
N = 14384
min = 0
max = 0.94
awr = 0.49
std = 0.38
0 60 120 км
изученные популяции


Распространение предковой компоненты "Мордовская-1" ADMIXTURE K=7
Значение вклада предковой компоненты
0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
Кострома
Тверь
Йошкар-Ола
Чебоксары
Казань
МОСКВА
Смоленск
Калуга
Рязань
Орел
Тамбов
Татарстан
K = 49
N = 14384
min = 0
max = 0.84
awr = 0.17
std = 0.18
0 60 120 км
изученные популяции



«Карельская» предковая компонента составляет 94% генофонда карел и редка в остальных популяциях. Ее вклад максимален в генофонде казанских татар (13%) и костромских русских (11%).

«Славянская» предковая компонента (рис. 3А), доминирует во всех русских популяциях (в среднем 81%), в других популяциях практически отсутствует за исключением казанских татар (6%) и Мордовии (3%).

Мордовскую предковую компоненту авторы делят на **«Мордовскую-1»** (рис. 3Б) и **«Мордовскую-2»** (рис. 3В). «Мордовская-1» максимальна среди геномов Мордовии, но широко распространена и в других популяциях. Наибольший вклад она вносит в тюркоязычные популяции: 36% в генофонд казанских татар, 35% — астраханских ногайцев, 20% — чувашей (причем ее вклад значителен практически во всех индивидуальных геномах). В русских популяциях ее вклад невелик (7%), но она обнаружена в 60% рассмотренных геномов русских, максимален вклад в Тверской (19%) и Калужской (16%) областях. В целом, «мордовская-1» компонента составляет постоянную часть генофонда практически всех славяноязычных, тюркоязычных и финноязычных популяций региона. Интересно, что различаются и геномы Мордовии: одна пятая часть геномов содержит 100% «мордовской-1», а другая пятая часть — 100% «мордовской-2» компоненты. Влияние «мордовской-2» компоненты прослеживается во всех популяциях региона (кроме удмуртов), но нигде не превышает 5%. В русских популяциях «мордовская-2» предковая компонента составляет в среднем 3%, но присутствует в 60% индивидуальных геномов русских, наиболее часто на востоке русского ареала – в Костромской, Нижегородской и Рязанской областях (рис. 4).



Рис. 4 – Доля индивидуальных геномов (%), в которых обнаружены предковые компоненты «Мордовская-1» и «Мордовская-2» (русские популяции)

«Марийская» предковая компонента уже при k=5 составляет не только весь генофонд луговых марийцев (96%), но и две трети генофонда чувашей (62%). Из других народов самый значительный вклад ее снова наблюдается у казанских татар (15%).

«Удмуртская» предковая компонента при k=3 выступает как «уральская», а на всех более высоких уровнях составляет 100% генофонда удмуртов. В геномах всех других народов она крайне редка, не превышая 4% генофонда, за исключением казанских татар, у которых ее вклад 10%.

«Калмыцкая» предковая компонента, отражающая центральноазиатское влияние, ранее (при k=3) описана как «восточная». При k=7 она доминирует в генофонде калмыков (100%) и ногайцев (61%). Из других народов ее вклад значителен лишь у казанских татар (12%), а у прочих очень невелик и нигде не превышает 5%.

Что же анализ показал про наиболее интересующий авторов генофонд русских Рязанской области, находящийся на стыке всех влияний? Несмотря на то, что специально собранная выборка охватила четыре географически разных района, представители всех четырех популяций сходны по генетическому профилю. Отличия проявились лишь при оценке вклада финноязычных популяций в их генофонд: в пограничном Сараевском районе их вклад составил 19%, а в других трех районах варьировал от 10 до 13%. Авторы выдвигают гипотезу, что это результат некоторого ослабления на юго-востоке Рязанской земли славянского влияния и роста влияния Дикого поля.

Для того, чтобы оценить своеобразие фармакогенетического ландшафта региона, авторы построили карту генетических расстояний от русских Рязанской области до других популяций по обширной панели фармакогенетических ДНК маркеров, которые были найдены в предыдущих исследованиях. Карта генетических расстояний показывает сходство и различия окружающих популяций с рязанскими русскими по фармакогенетическим ДНК маркерам. Видно, что наибольшее сходство наблюдается с соседними популяциями Мордовии. На втором месте по сходству с рязанцами – русские Калужской, Смоленской и Костромской областей, за ними следуют орловские и тверские популяции. На третьем месте оказались как тамбовские русские, так и восточные соседи — финноязычные марийцы, и тюркоязычные чуваши. Ожидаемо резко отличается фармакогенетический портрет татар и удмуртов. Но самым неожиданным результатом оказался очень высокий уровень фармакогенетических различий между рязанскими и нижегородскими русскими, притом, что по селективно-нейтральным аллелям тех же самых геномов (по которым вычислялись предковые компоненты) они довольно сходны.

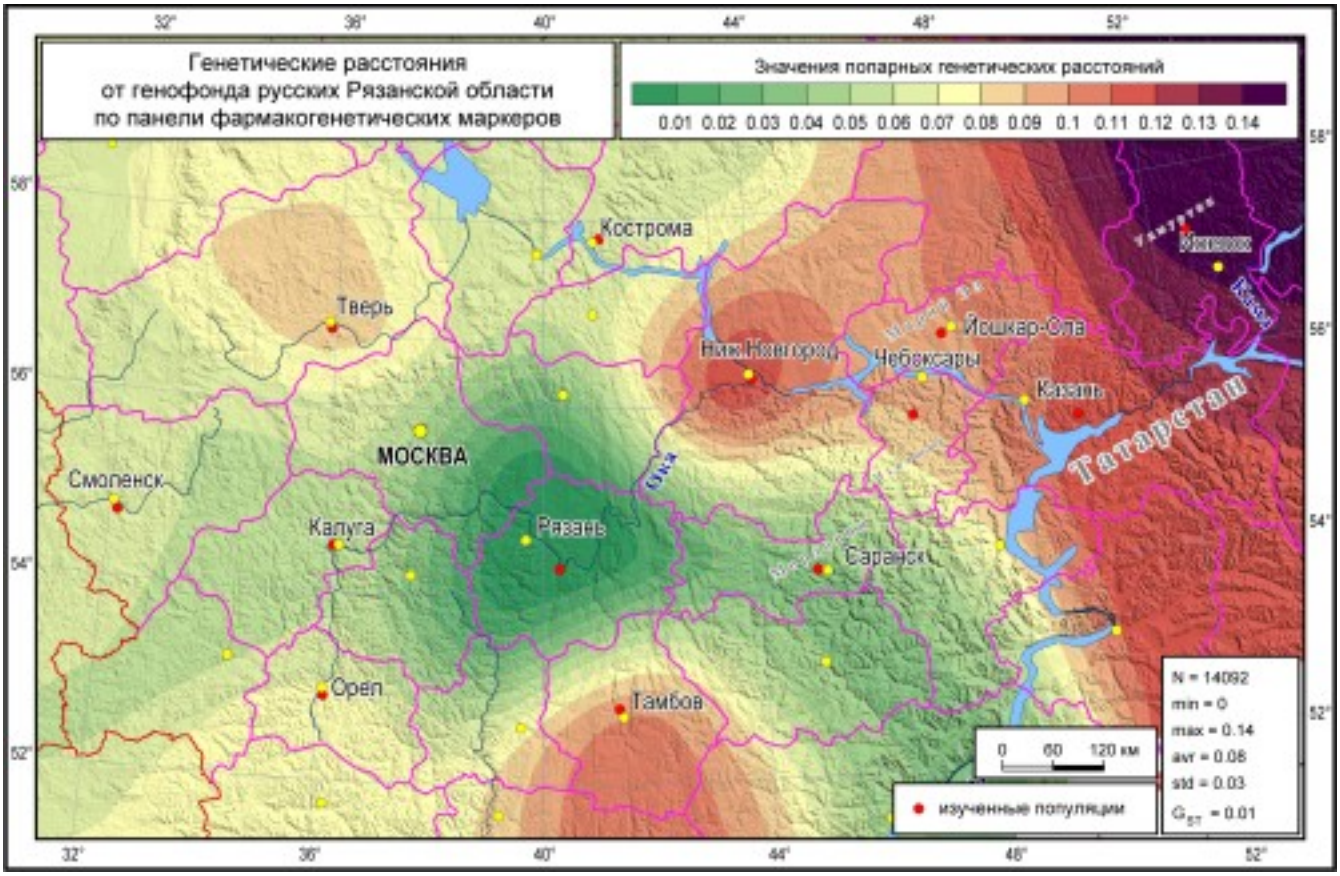


Рис. 5. Фармакогенетический ландшафт региона (генетические расстояния от русских Рязанской области по фармакогенетическим маркерам) Обозначения: минимальные генетические расстояния (близость к генофонду русских Рязанской области) обозначены зелеными тонами; максимальные генетические расстояния (несходство с генофондом русских Рязанской области) обозначены красно-фиолетовыми тонами.

Полученные результаты приводят авторов к нескольким выводам. Поскольку практически во всех популяциях региона вне зависимости от их языковой принадлежности обнаруживаются мордовские предковые компоненты, можно выдвинуть гипотезу, что в генетический портрет дославянского населения региона входили две основные «краски», сохраненные в современном генофонде Мордовии. Генофонды изученных русских популяций на 80% описываются собственной «славянской» предковой компонентой; влияние финноязычных народов оценивается в 19%, а центральноазиатское влияние – всего в 1%. Очень велико оказалось влияние генофондов финноязычных популяций на тюркоязычные народы Поволжья. Генофонд казанских татар показал себя наиболее «составным», включившим все предковые компоненты генофондов других популяций. Поэтому одна из целей исследования – оценить влияние генофонда татар – не достижима методом анализа предковых компонент для данной совокупности геномов из-за отсутствия специфичных черт генофонда казанских татар.

Наконец, анализ фармакогенетических ДНК маркеров выявил своеобразие фармакогенетического ландшафта. Это наглядно показывает, что для выявления регионов со сходными протоколами по подбору лекарственных препаратов и их доз в зависимости от генотипа пациента нельзя полагаться только на сходство и различия по селективно-нейтральным аллелям.

Источник:

Балановская Е.В., Горин И.О., Пономарев Г.Ю., Пылёв В.Ю., Петрушенко В.С., Маркина Н.В., Мамаева А.Д., Ларин А.К., Агджоян А.Т. Следы взаимодействия финноязычного, славянского и тюркоязычного населения в современном генофонде и их отражение в фармакогенетике // Вестник РГМУ. 2022. № 2. Опубликовано online**:** 26.04.2022. DOI: 10.24075/vrgmu.2022.019.

Видеоролик по статье вышел на канале Археология, история, генетика — исследования.

https://www.youtube.com/watch?v=bhjMlRlIA7w