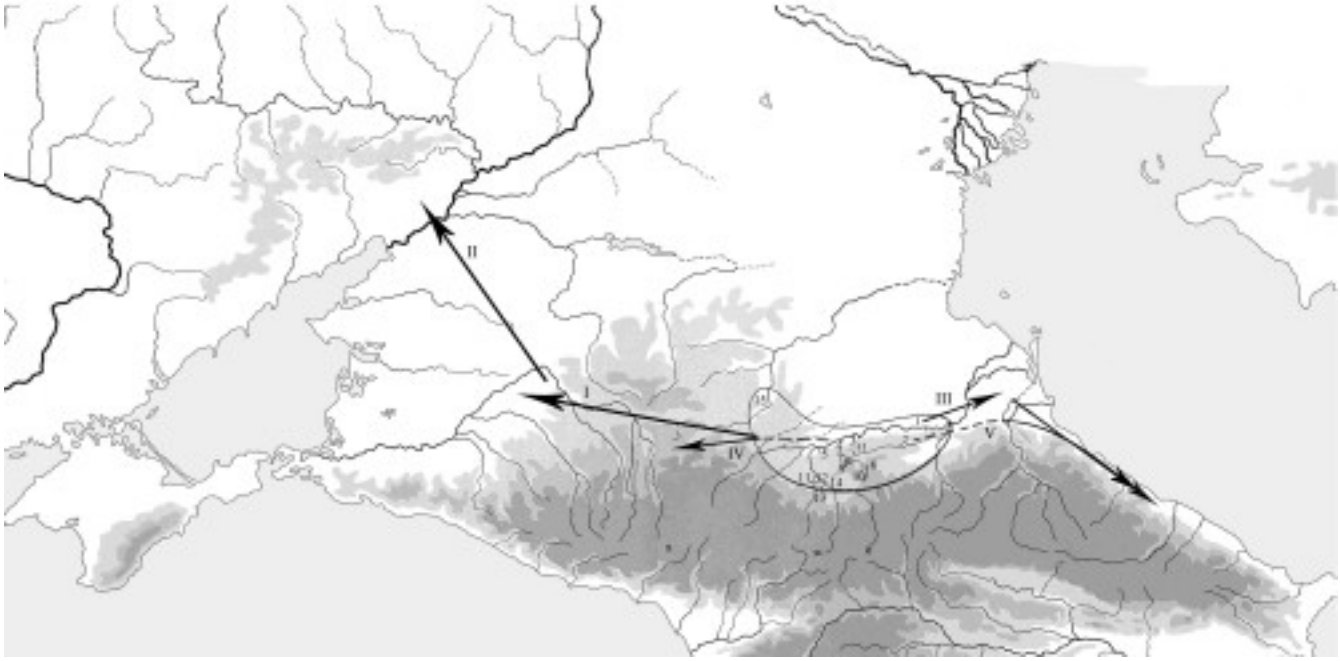
Поиск генетического наследия алан в современном населении Кавказа

Методами полногеномного анализа ученые проверяли гипотезу о сохранении наследия тысячелетней истории аланских племен в генофондах современных народов Кавказа. С использованием метода ADMIXTURE они оценили вклад разных предковых компонент в геномы 29 популяций. Оказалось, что среди «претендентов» на аланское наследие – народов, ныне живущих в ареале аланских племен, — нет единой предковой компоненты, которую можно было бы идентифицировать как генетический след алан. Наибольший вклад в геномы предполагаемых «наследников» алан внесли четыре предковых компоненты: условно «осетинская», «нахская», «абхазо-адыгская»и «карачаево-ногайская». Вклад «осетинской» предковой компоненты выше у южных обществ осетин, и некоторые из них могли в большей мере сохранить аланский след. Но для более определенного ответа на этот вопрос необходим анализ древней ДНК. В работе также исследован фармакогенетический ландшафт популяций Северного Кавказа, значительно отличающийся от генетического ландшафта по селективно-нейтральным маркерам.

Вопрос о том, оставили ли аланы след в генофондах современных народов Кавказа и в каких именно, относится не только к фундаментальной науке. В течение многих лет эта проблема вызывает довольно острую общественную дискуссию с участием представителей разных народов, считающих себя наследниками алан. Сегодня очевидна роль генетики, которая в сотрудничестве с археологией, антропологией, историей и лингвистикой может сказать существенное слово в реконструкции событий прошлого, причем ее возможности намного повысились с развитием полногеномных технологий. Такой анализ был проведен в работе междисциплинарного коллектива под руководством профессора Е.В. Балановской, опубликованной в журнале «Вестник Московского государственного университета, серия Антропология».

Аланы – ираноязычные кочевые племена, появляются на Северном Кавказе в начале 1-го тысячелетия н.э., ко второй половине II в. н.э. складывается аланская археологическая культура, включающая керамику, погребальные комплексы, множество больших (от 0,5 до 1,2 км^2) городищ со сложной структурой и четко выделенной цитаделью. Экспансии алан на обширную территорию региона способствовал благоприятный для животноводства и земледелия климат и, соответственно, рост численности и плотности населения. В конце IV-V вв. н.э. началась эпоха Великого переселения народов, а климат стал более засушливым. Это привело к деградации аланских городищ на равнинах, и аланы начинают более интенсивно осваивать предгорные районы. Аланская культура прослеживается на Северном Кавказе почти в течение тысячелетия, после периода Золотой орды потомки алан сохранились в горных районах. Согласно одной из убедительных научных гипотез аланы и ассимилированные ими горные племена сформировали в средние века осетинскую народность. Средневековые осетины сохранили некоторые культурные аланские традиции и «иранскую» речь; палеоантропологические данные указывают на преобладание аланского антропологического варианта в раннесредневековом населении Северной Осетии. Однако и другие народы Северного Кавказа из ареала аланских племен – балкарцы, ингуши, карачаевцы, чеченцы (перечислены в алфавитном порядке) — могут быть их генетическими наследниками

До сих пор вопрос аланского генетического наследия не исследовался методами полногеномного анализа генофондов современного населения Северного Кавказа. Целью данной работы стал поиск общего генетического компонента, кандидатного на роль аланского следа, в аутосомных геномах современных популяций Северного Кавказа (осетин, балкарцев, карачаевцев, ингушей и чеченцев), которые проживают в основном ареале аланских племен. Наибольшее внимание авторы уделили генофонду осетин, вероятных лингвистических наследников алан, но в сравнительный анализ вовлекли и другие народы.

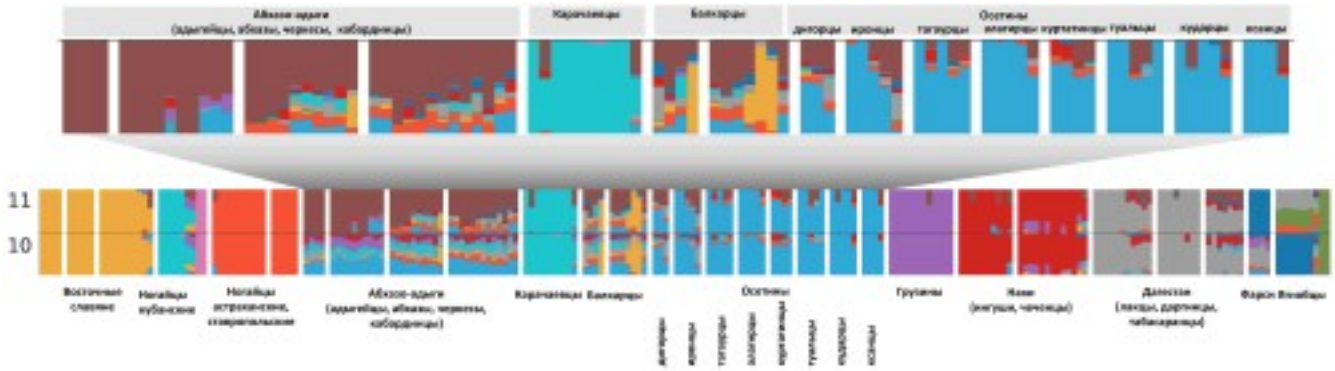


«Метрополия» аланской культуры Северного Кавказа раннего этапа (II-IV в. н.э.) и миграции ее носителей (приводится по [Малашев, 2022]). Миграции: I – в Прикубанье (первая половина III в. н.э.); II – на Нижний Дон (первая половина – середина III в. н.э.); III – в Терско-Сулакское междуречье (первая половина – середина III в. н.э.); IV – в Кисловодскую котловину (начало IV в. н.э.); V – в Южный Дагестан с населением Терско-Сулакского междуречья (первая половина – середина IV в. н.э.).

ДНК для анализа, выделенная из собранных коллективом образцов венозной крови, хранящихся в Биобанке Северной Евразии, включила 207 геномов из 29 популяций. Наиболее детально были представлены осетины – 22 генома северных (алагирцы, дигорцы, иронцы, куртатинцы, тагаурцы) и 14 геномов южных (ксанцы, кударцы, туальцы) осетин. В выборку включены абхазы, адыгейцы, балкарцы (11 популяций), даргинцы, ингуши (5 популяций), кабардинцы, казаки кубанские, карачаевцы (7 популяций), ногайцы (астраханские, кубанские, ставропольские), лазы Грузии, лакцы, южные русские, табасаранцы, фарсиязычные популяции Дагестана, черкесы, чеченцы (7 популяций), ягнобцы Таджикистана. Генотипирование выполняли по полногеномной панели 4.5 млн SNP маркеров с помощью биочипа Infinium OmniExome BeadChip Kit (Illumina, США) на приборе iScan (Illumina, США).

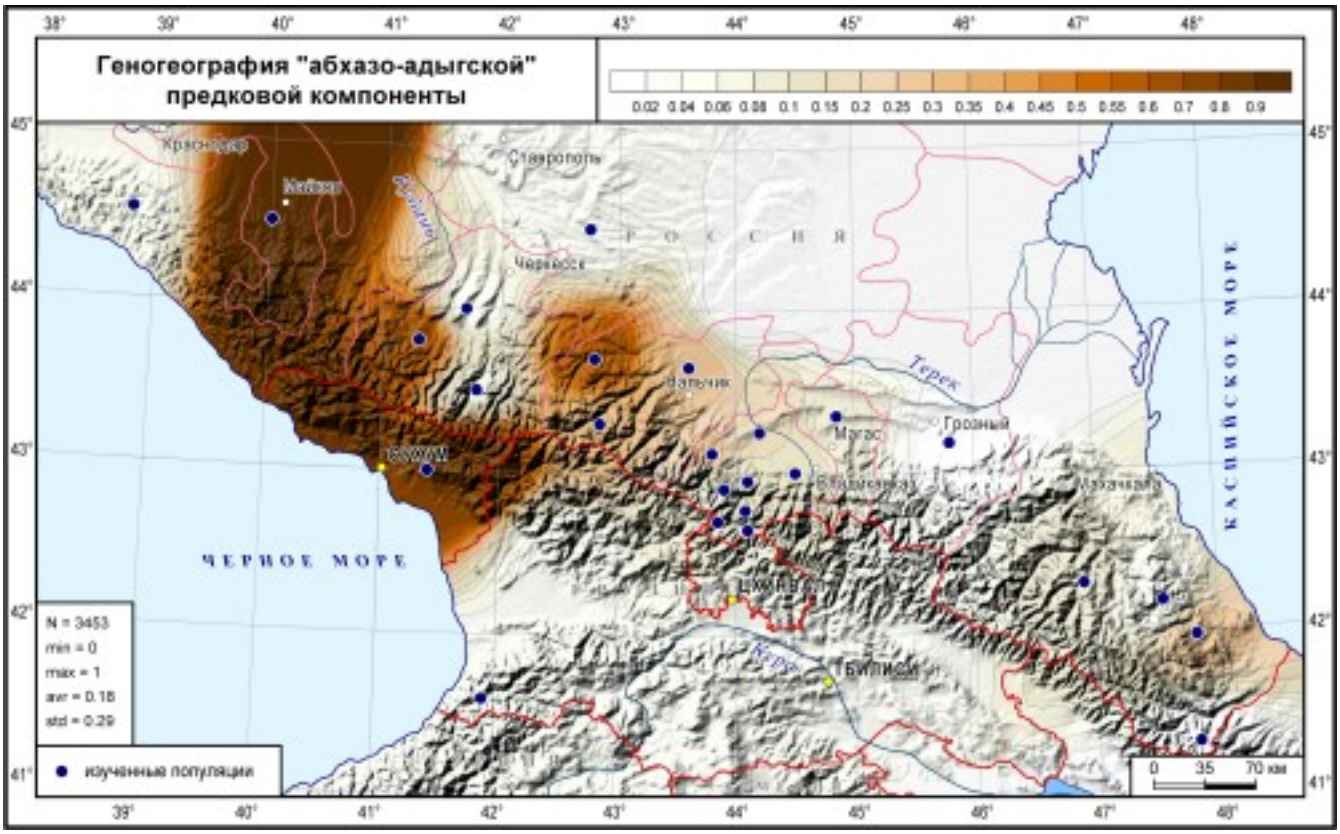
Основным методом анализа стал ADMIXTURE — моделирование вклада (в%) в каждый индивидуальный геном заданного числа предковых компонент K. При увеличении числа K программа создает независимые модели, выявляя все более детальные предковые компоненты. Обобщив данные о геномах из одной популяции, получают средний вклад (в %) в генофонд современной популяции каждой предковой компоненты K. В данном исследовании авторы создавали 14 независимых моделей (от K=2 до K=15), и для каждого K повторили расчет 10 раз, итого создав 140 независимых моделей предковых компонент для совокупности 207 геномов. Результаты моделирования ADMIXTURE затем картографировали при помощи программного пакета «GeneGeo». В работу входила и еще одна задача – оценить фармакогенетический статус населения региона. Для этого использованы данные о ранее описанных исследователями 42 ключевых фармакогенетических ДНК-маркерах (генов ADME; генов, кодирующих фармакодинамические мишени лекарственных средств; генов, кодирующих компоненты системы гемостаза). По этим данным были созданы фармакогенетические карты.

Анализируя результаты моделирования ADMIXTURE, авторы пришли к выводу, что оптимальная по детализации и устойчивости структура генофонда Северного Кавказа проявляется при числе предковых компонент K=10 и K=11.

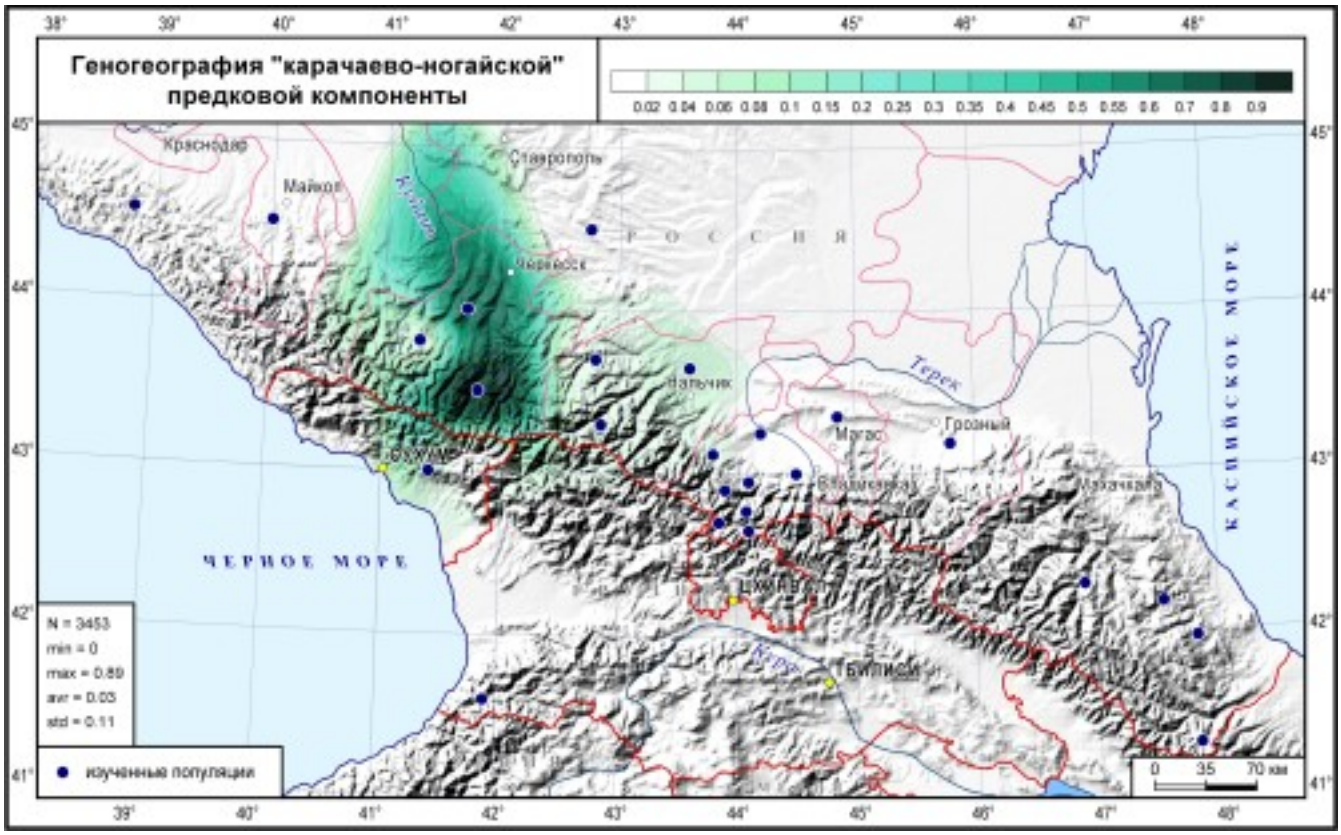


Спектр предковых компонент ADMIXTURE для популяций Кавказа и групп сравнения при K=10-11 (фрагмент А, внизу), подробнее для Западного и Центрального Кавказа при K=11 (фрагмент Б, вверху).

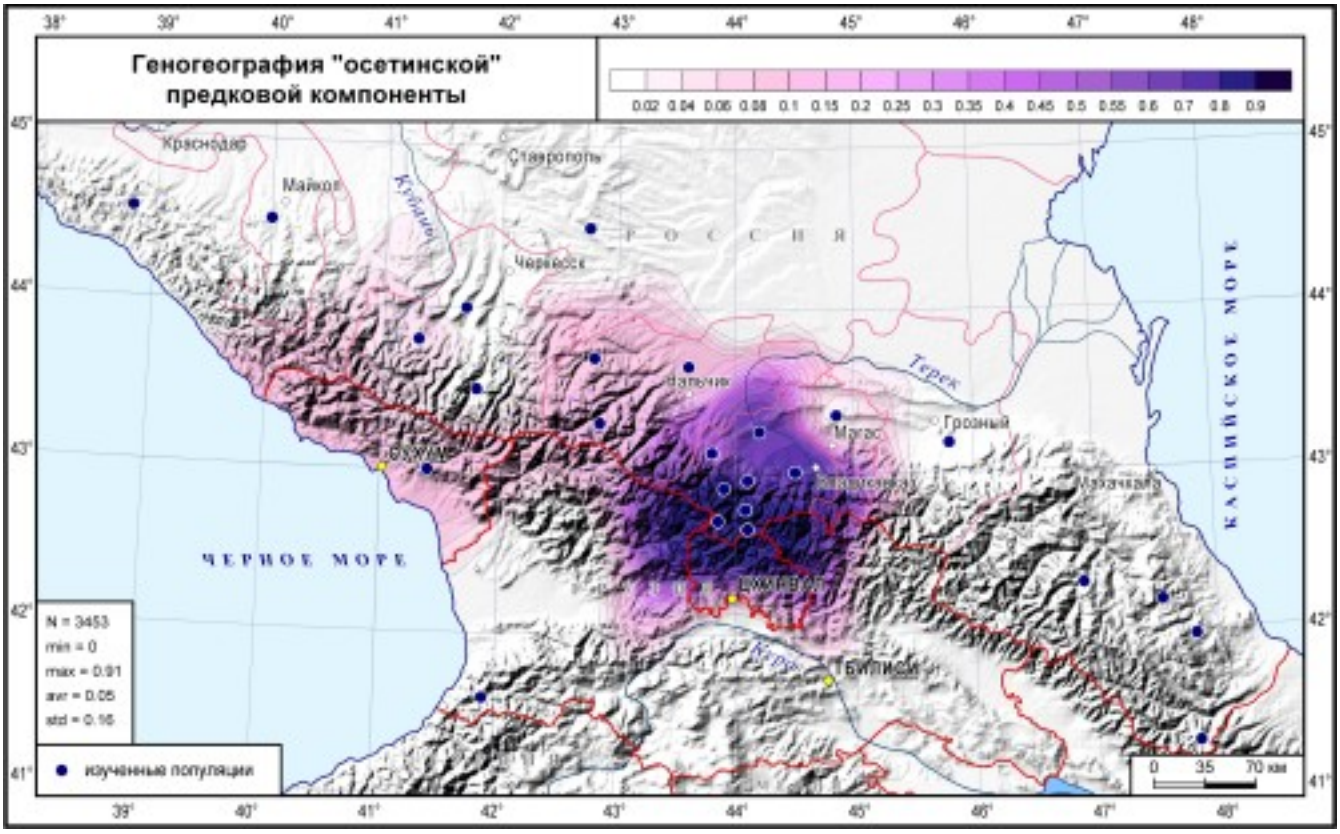
«Абхазо-адыгская» компонента (коричневый цвет на графике ADMIXTURE) составляет три четверти генофонда абхазо-адыгских народов — абхазов, адыгейцев, кабардинцев, черкесов. При этом она единственная обнаружена в генофондах практически всех анализируемых популяций (за исключением фарсиязычных популяций Дагестана, ягнобцев Таджикистана и южных русских). Она составляет треть генофонда балкарцев, значительную часть генофондов дигорцев (26%) и иронцев (19%), варьируя в остальных обществах осетин в интервале 7-12%. Ее вклад снижается у нахских народов (5%), снова возрастая в Дагестане (13%). Обширность ареала этой компоненты демонстрирует карта: доминирование на Западном Кавказе и невысокие частоты на Центральном и Восточном Кавказе. Иными словами, «абхазо-адыгская» предковая компонента обнаруживается во всех популяциях Северного Кавказа, возможно, фиксируя генетические следы общей для них прапопуляции автохтонного населения Северного Кавказа и частично след их более позднего расселения на восток.



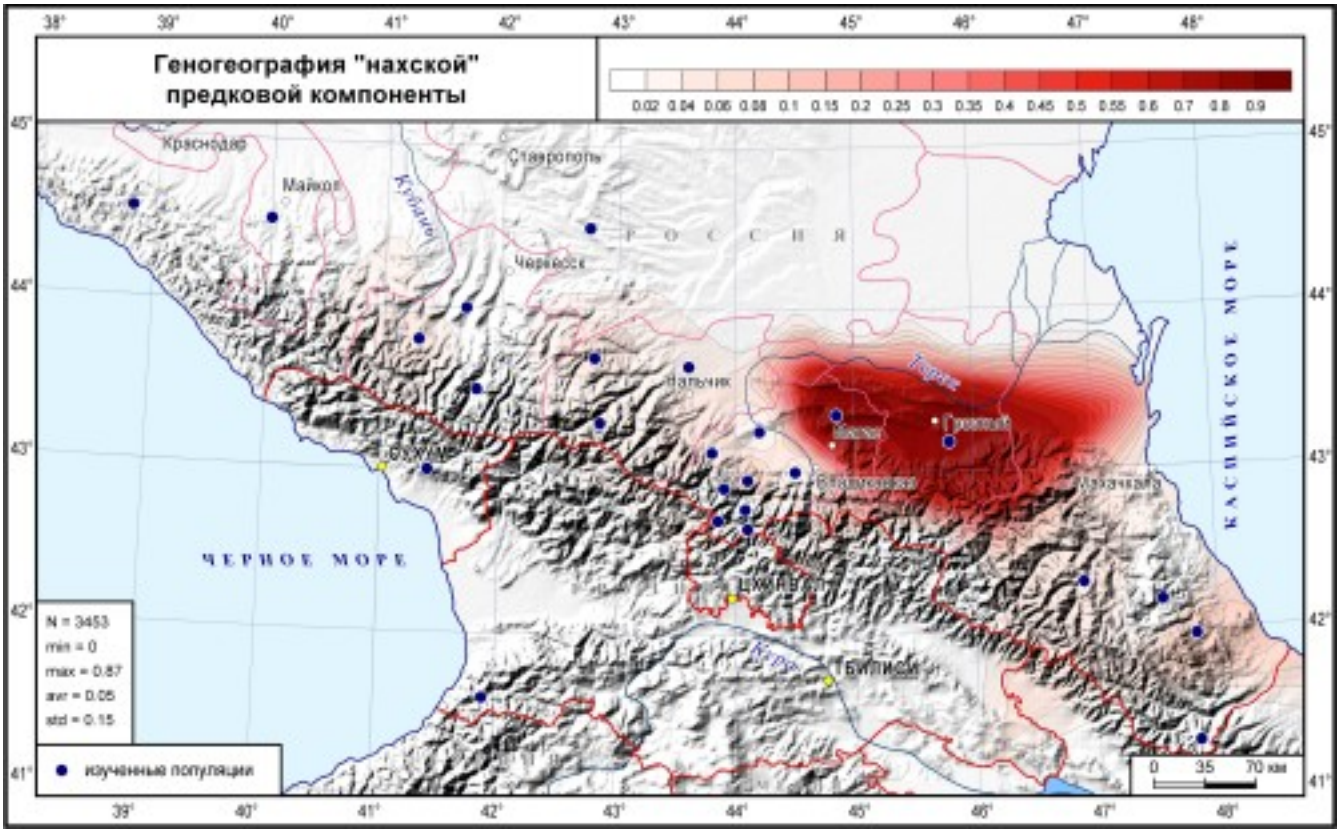
«Карачаево-ногайская» компонента (светло-голубой цвет на графике ADMIXTURE) доминирует у карачаевцев (89%) и у кубанских ногайцев (62%). Карта показывает, что ареал ее компактен и как бы встроен в ареал «абхазо-адыгской» компоненты. У географических соседей карачаевцев – черкесов и равнинных популяций балкарцев – ее вклад невелик (5-6%), в других популяций Западного и Центрального Кавказа она обнаружена лишь у отдельных индивидов.



«Осетинская» компонента (ярко-голубой цвет на графике ADMIXTURE, карта) претендует на второе место по распространенности, охватывая не только этнический ареал осетин, но и с невысокими частотами область Западного и Центрального Кавказа. Ее вклад в среднем составляет три четверти генофонда осетин (76%), но значительно варьирует в разных обществах от максимальных значений 90-91% у алагирцев и ксанцев до 64-69% у иронцев и куртатинцев и даже 45% у дигорцев. В других генофондах ее вклад тяготеет к Западному Кавказу, составляя 9-10% генофондов абхазов, балкарцев и кабардинцев, 6% у черкесов и лишь 3% у ингушей и 1% у чеченцев Центрального Кавказа.



«Нахская» компонента (красный цвет на графике ADMIXTURE, карта) почти столь же обширна по влиянию. Очень высок ее вклад в генофонды ингушей (87%) и чеченцев (83%), но область небольших и малых ее частот распространяется на популяции от Западного до Восточного Кавказа. Ее вклад в генофонды осетин-куртатинцев и табасаранцев Дагестана достигает 10%, в генофонды осетин-дигорцев и лакцев Дагестана 6%, в генофонды Центрального Кавказа (иронцев, тагаурцев) и народов Западного Кавказа (кабардинцев, черкесов, балкарцев) 3-4%.



«Дагестанская» компонента (светло-серый цвет на графике ADMIXTURE) в среднем составляет три четверти (77%) общего генофонда народов Дагестана, достигая максимума у даргинцев и практически отсутствуя у соседей — нахских народов Центрального Кавказа. Но отдельные включения ее обнаруживаются в Осетии (8-9% у дигорцев и иронцев) и на Западном Кавказе у балкарцев (11%).

«Степная» компонента (оранжевый цвет на графике ADMIXTURE) составляет практически весь генофонд ставропольских и астраханских ногайцев. Удивительно, что у кубанских ногайцев ее вклад составил лишь 1%. По мнению авторов, возможно, именно «степным» влиянием можно объяснить заметный вклад этой компоненты в генофонды Карачаево-Черкесии и Кабардино-Балкарии (11-12% у кабардинцев и черкесов, 9% у балкарцев), а также у северных осетин (3-8%.)

«Грузинская» компонента (сиреневый цвет на графике ADMIXTURE) кроме лазов и имеретинцев Грузии (96%) с заметным вкладом обнаружена у чеченцев (11%) и абхазов (5%).

«Фарси» компонента (синий цвет на графике ADMIXTURE) кроме фарсиязычной популяции Дагестана (включенной в анализ в качестве представителя юго-западной ветви иранских языков), выявлена также с 6% вкладом у табасаранцев и балкарцев.

«Ягнобская» компонента (темно-зеленый цвет на графике ADMIXTURE) отсутствует как у осетин, так и у других претендентов на аланский след. Геномы ягнобцев горного Таджикистана были включены в анализ именно потому, что лишь они и осетины являются сейчас единственными живыми представителями северо-восточной ветви иранских языков, в которую ранее входили и сарматские языки, и аланский.

«Частная ногайская» компонента (бирюзовый цвет, на графике она почти не видна) ярко проявилась только в двух геномах кубанских ногайцев. Но ее небольшой вклад в отдельные геномы других народов обнаруживается повсеместно: у всех народов Западного Кавказа, в некоторых обществах осетин, у чеченцев, табасаранцев, лакцев и кубанских казаков. Видимо, он фиксирует уже третий, хотя и неяркий, вариант тех прапопуляций, пришедших из евразийской степи, которые вместе с «карачаево-ногайской» и «степной» участвовали в сложении генофондов ногайцев и карачаевцев.

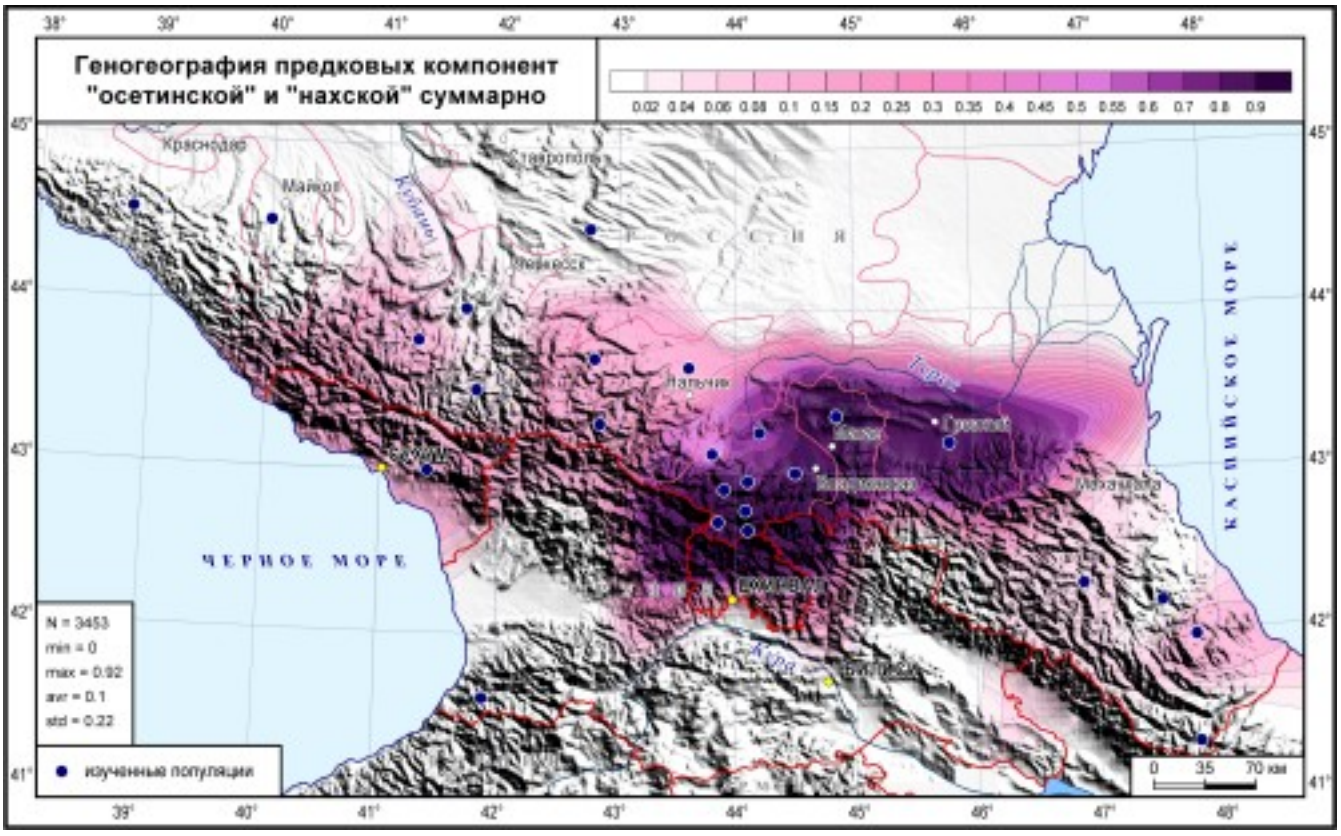
«Славянская» предковая компонента (желтый цвет на графике ADMIXTURE) отражает наиболее поздний вклад славянского населения в генофонды коренного населения Кавказа. Конечно же, ее вклад максимален у южных русских (100%) и кубанских казаков (93%). Однако она также составляет четверть генофонда балкарцев, причем максимальна в популяциях горных балкарцев (27%), в то время как у кабардинцев и черкесов ее вклад очень мал (2-4%), а у карачаевцев и во всех иных популяциях также практически отсутствует.

Полученные результаты – и особенно наглядно геногеографические карты — четко показывают, что в современном населении Кавказа отсутствует единственная предковая компонента, которую можно было бы прямо связать с аланским генетическим наследием: ареалы всех четырех преобладающих предковых компонент («абхазо-адыгской», «карачаево-ногайской» «осетинской» и «нахской» прапопуляций) складываются друг с другом как паззл, соединяясь на периферии своих ареалов. Более того, суммарно эти четыре компоненты составляют основную часть генофондов тех современных народов, у которых предполагается наличие «аланского генетического следа»: общий вклад четырех «прапопуляций» составляет практически весь генофонд ингушей (99%), карачаевцев (98%), осетин (95%), чеченцев (85%) и лишь половину генофонда балкарцев (49%).

Придя к такому выводу, авторы работы переходят на более детальный уровень анализа – оценивают вклады разных предковых компонент в индивидуальных геномах. Дело в том, что наличие даже заметного вклада некоей предковой компоненты в генофонде популяции может не иметь отношения к древней истории популяции, а оказаться связано лишь с недавними отдельными мигрантами: в этом случае данная компонента будет преобладать в геномах лишь некоторых индивидов и полностью отсутствовать у остальных. И напротив, даже небольшой вклад предковой компоненты в генофонд популяции может отражать важный этап ее генетической истории, некий древний пласт ее генофонда, если он прослеживается в большинстве геномов популяции пусть даже с небольшим вкладом.

С этих позиций «абхазо-адыгская» компонента, представленная практически у всех балкарцев (91% геномов), среди 50% геномов осетин, 30% ингушей, 20% карачаевцев, 8% чеченцев, а также у половины геномов Дагестана, предположительно отражает вклад некоего автохтонного населения Кавказа. «Карачаево-ногайская» компонента, присутствует у всех карачаевцев, у 36% геномов балкарцев, 22% осетин, 15% чеченцев, но отсутствует у ингушей; она характерна почти для всех кубанских ногайцев и при этом отсутствует у астраханских и ставропольских ногайцев. Можно предположить, что эта компонента отражает след одной из миграций тюркоязычных популяций золотордынского времени, которые могли резко различаться по своим генетическим портретам. «Осетинская» компонента присутствует у всех осетин, у 73% балкарцев, 30% ингушей, 15% чеченцев 10% карачаевцев и у 8% среди народов Дагестана. Наличие ее среди заметного числа геномов всех пяти этносов, предположительно, имеющих аланский генетический след, позволяет рассматривать ее как наиболее вероятный кандидат на него. «Нахская» компонента присутствует во всех геномах ингушей и чеченцев, у 73% геномов балкарцев, 42% осетин, 10% карачаевцев и у 38% геномов народов Дагестана. То есть и эта компонента оставила внушительный вклад в геномах всех пяти этносов, потому она с не меньшим основанием может служить кандидатом на аланский генетический след.

Это позволяет выдвинуть гипотезу, что обе — «осетинская» и «нахская» предковые компоненты – каждая по-своему и лишь отчасти отражают аланский генетический след. Рассмотрение числа индивидуальных геномов, в которых обнаруживается вклад и «осетинской,» и «нахской» компонент, демонстрирует такой градиент: осетины, ингуши, чеченцы, балкарцы, карачаевцы, народы Дагестана. А суммарно вклад «осетинской» и «нахской» компонент (карта) в генофонды Северного Кавказа выявляет иной градиент: ингуши (91%), чеченцы (83%), осетины (80%), кабардинцы (14%), балкарцы (13%), абхазы, табасаранцы, черкесы (10%), лакцы (6%), карачаевцы, ногайцы кубанские (2%).



Итак, анализ вклада предковых компонент в генофонды современных народов Северного Кавказа, проживающих в основном ареале аланских племен, не обнаружил генетический след одной предковой прапопуляции. Однако две предковые компоненты — «осетинская» и «нахская» — дают наибольший вклад как в индивидуальные геномы, так и в популяционные генофонды основных кандидатов на звание «наследников» алан – осетин, ингушей и чеченцев.

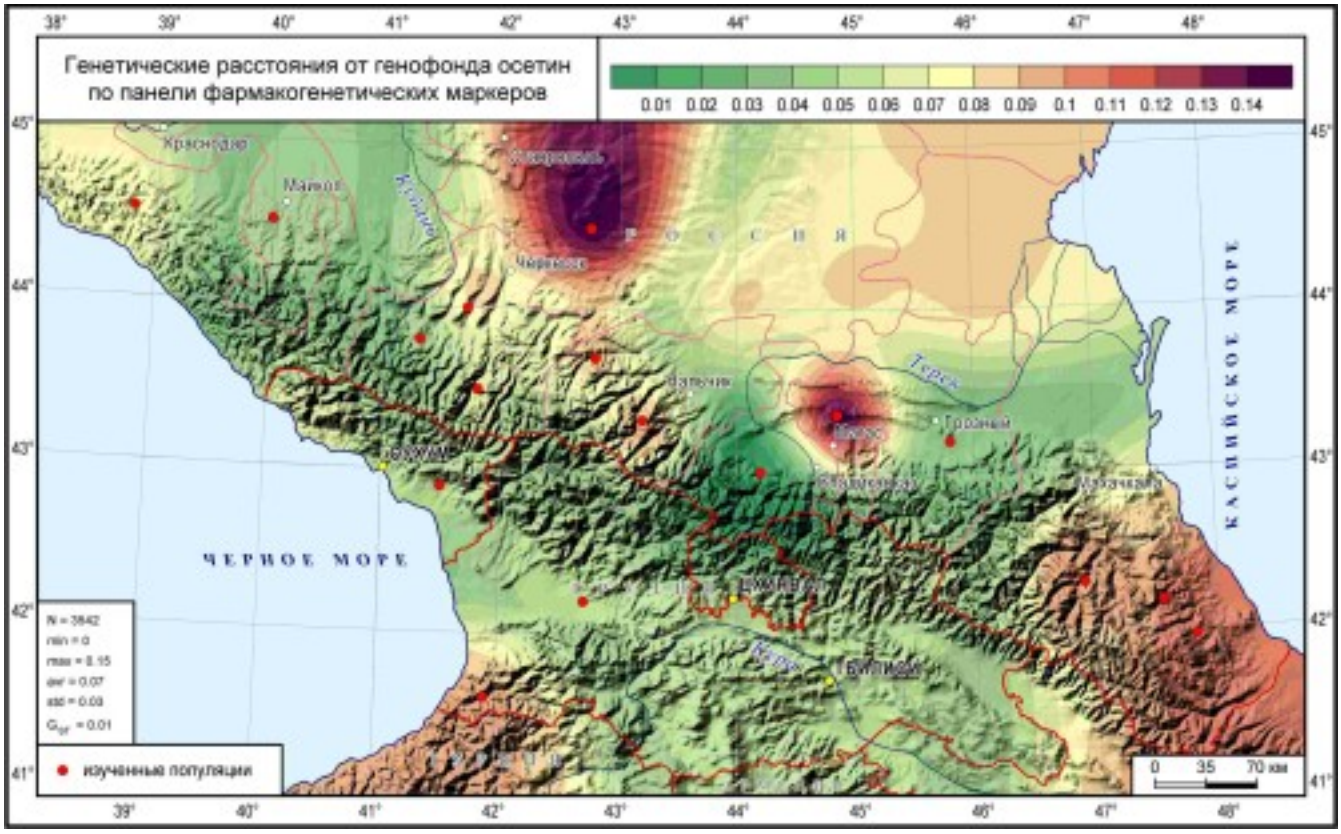
Генофонд балкарцев предстает как наиболее сложносоставной — с наибольшим вкладом «абхазо-адыгской» компоненты (треть генофонда) и с небольшим суммарным вкладом «осетинской» и «нахской» компонент (лишь 13%). Генофонд карачаевцев в основном состоит из предковой компоненты, общей с ногайцами, отражая след одной из миграций тюркоязычных популяций золотордынского времени.

Геногеографический анализ распределения вклада основных кандидатов на аланский генетический след («осетинской» и «нахской» компонент), выявленная неоднородность осетинских обществ и лингвистическое сходство осетин с аланами поддерживает гипотезу, что именно среди осетинских обществ наиболее вероятно найти аланский генетический след. При этом нельзя отвергать гипотезу, что обе предковые компоненты — «осетинская» и «нахская» — могут отражать генетическое наследие алан, хотя, как отмечают исследователи, в этом случае приходится допустить совокупность маловероятных генетических событий.

В любом случае, не подтвердилась исходная гипотеза исследования – наличие общего аланского следа в генофондах тех народов, которые проживают в ареале аланской «метрополии» и которых принято считать потомками алан. Возможно, рассуждают авторы, этот след в современных генофондах настолько слаб, что лишь анализ древней ДНК сможет его выявить. С другой стороны, можно допустить крайне мало вероятную гипотезу, что сами аланы были генетически весьма гетерогенны, и именно две прапопуляции («осетинская» и «нахская») действительно отражают аланское наследие. Более точный ответ можно будет получить лишь при анализе древней ДНК индивидов из разных частей аланского ареала и из разных хронологических срезов.

Вторая задача – исследование фармакогенетического статуса популяций Северного Кавказа, привела к следующим результатам. Была построена карта генетических расстояний Нея (d) по совокупности 42 наиболее значимых фармакогенетических маркеров от популяции осетин до остальных популяций изучаемого региона. Надо подчеркнуть, что, в отличие от селективно-нейтральных ДНК-маркеров, которые рассматривались выше, ДНК-маркеры, связанные со здоровьем человека, могли быть мишенью естественного отбора, поэтому их распределение в популяциях может отличаться.

Так и вышло – для фармакогенетических маркеров обнаружены иные паттерны генетического сходства, чем для селективно-нейтральных маркеров. Так, выявляется наибольшая близость фармакогенетического статуса осетин и чеченцев. На втором месте по степени сходства (зеленоватые тона) находятся черкесы и адыгейцы. Далее по степени снижения сходства фармакогенетических статусов: абхазы, карачаевцы, кубанские ногайцы, кабардинцы, балкарцы, народы Дагестана. Наибольшие фармакогенетические отличия от осетин обнаружены у ставропольских ногайцев и, что крайне неожиданно, у ингушей. В целом ареал наибольшего сходства с осетинами тянется с востока на запад вдоль Кавказского хребта. Наиболее ярко выражено сходство осетин по фармакогенетическому статусу с чеченцами и резкие отличия от ингушей, что ярко контрастирует с данными, полученными в рамках метода ADMIXTURE, где генофонды чеченцев и ингушей почти неотличимы друг от друга. Это показывает, что для развития персонифицированной медицины необходим специальный тщательный анализ фармакогенетического профиля популяций.



Карта средних генетических расстояний по фармакогенетическим маркерам от осетин до популяций Кавказа. Минимальные генетические расстояния (близость к генофонду осетин) обозначены зелеными тонами; максимальные генетические расстояния (несходство с генофондом осетин) обозначены красно-фиолетовыми тонами.

Источник:

Балановская Е.В., Агджоян А.Т. Горин И.О., Петрушенко В.С., Пылёв В.Ю, Кулемин Н.А., Маркина Н.В., Черкасов Н.А., Кошель С.М., Почешхова Э.А. В поисках аланского следа: генетическая история Северного Кавказа по полногеномным данным об аутосомно м генофонде // Вестник Московского государственного университета, серия 23 Антропология. 2022. №3. с. 48-62. **DOI:** 10.32521/2074-8132.2022.3.048-062