Народы уральской семьи языков: генетические следы общей истории

Впервые популяции уральской языковой семьи систематически изучены с помощью широкогеномного анализа 15 народов – от Северной и Центральной Европы до Сибири. Показано, что генетическое разнообразие уралоязычных популяций во многом определяется географией. В то же время большинство уралоязычных популяций имеет больше генетических сегментов общего происхождения со своими лингвистическими родственниками, чем их географические соседи иных языковых семей. Лексические и генетические (аутосомные) расстояния между уралоязычными популяциями связаны положительной корреляцией. Авторы делают вывод, что распространение уральских языков, по меньшей мере, частично было связано с миграциями населения. И предполагают у уралоязычных популяций общий генетический след сибирского происхождения.

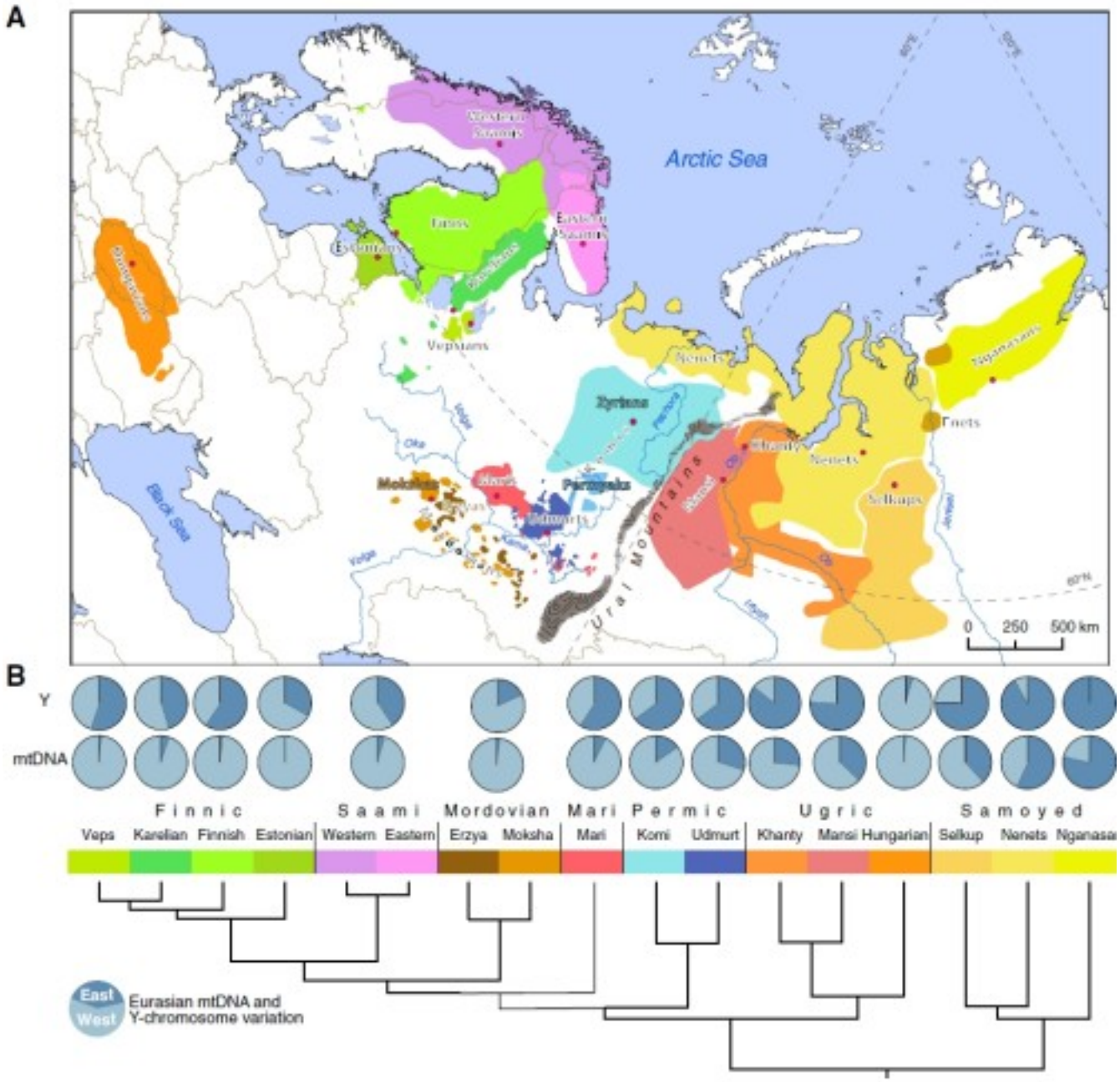
Читайте также ниже, в разделе «Мнения экспертов» экспертный комментарий д.б.н., проф. РАН О.П.Балановского

Уральская языковая семья входит в тройку наиболее распространенных в Северной Евразии (наряду с индоевропейской и тюркской). Лингвисты считают, что языки уральской семьи возникли из единого протоязыка, возрастом от 6000 до 4000 лет, который разделился на две большие ветви – финно-угорских и самодийских языков. Народы, говорящие на языках уральской семьи, расселены по огромной территории — от Балтики до Западной Сибири, от Центральной Европы до полуострова Таймыр. Распространение этих языков и их связь с материальной культурой довольно давно исследуется историческими лингвистам и археологами, но генетическая сторона этого процесса до последнего времени не была изучена с помощью современных, полногеномных, методов.

Поиску генетической общности уралоязычных популяций и генетического вклада в распространение уральских языков посвящена статья, недавно опубликованная в журнале BMC Genome Biology. Основные ее авторы – генетики из Эстонского биоцентра Университета Тарту, в составе международного коллектива соавторов есть и российские ученые из Москвы, Новосибирска, Уфы, Архангельска.

В уральскую семью языков входят несколько групп с определенной географической привязкой. Прибалтийско-финская группа (на этих языках говорят финны, карелы, вепсы, эстонцы) и саамская (саамы) – занимают территорию на северо-западе Европы. Мордовская (эрзя, мокша), марийская (мари), пермская (коми, удмурты) группы распространены в Волжско-Уральском регионе. Носители угорской группы языков проживают в двух географически отдаленных регионах: ханты и манси – в Западной Сибири, а венгры – в Центральной Европе. Наконец, самодийская группа языков (это языки селькупов, ненцев, нганасанов) распространены на территории Центральной и Северо-Восточной Сибири.

Группы языков уральской семьи, лингвистическое дерево и географическое расположение уралоязычных популяций представлены на рисунке. Цвет ареала языка на карте соответствует его цвету на панели групп языков.



Географическое расположение уралоязычных популяций. Цвет на карте соответствует цвету группы языков на панели b. Схематическое представление филогении уральских языков. Круговые диаграммы показывают доли западноевразийского (светло-голубой цвет) и восточноевразийсого (темно-голубой цвет) генетических компонентов в Y-хромосомном и митохондриальном генофондах.

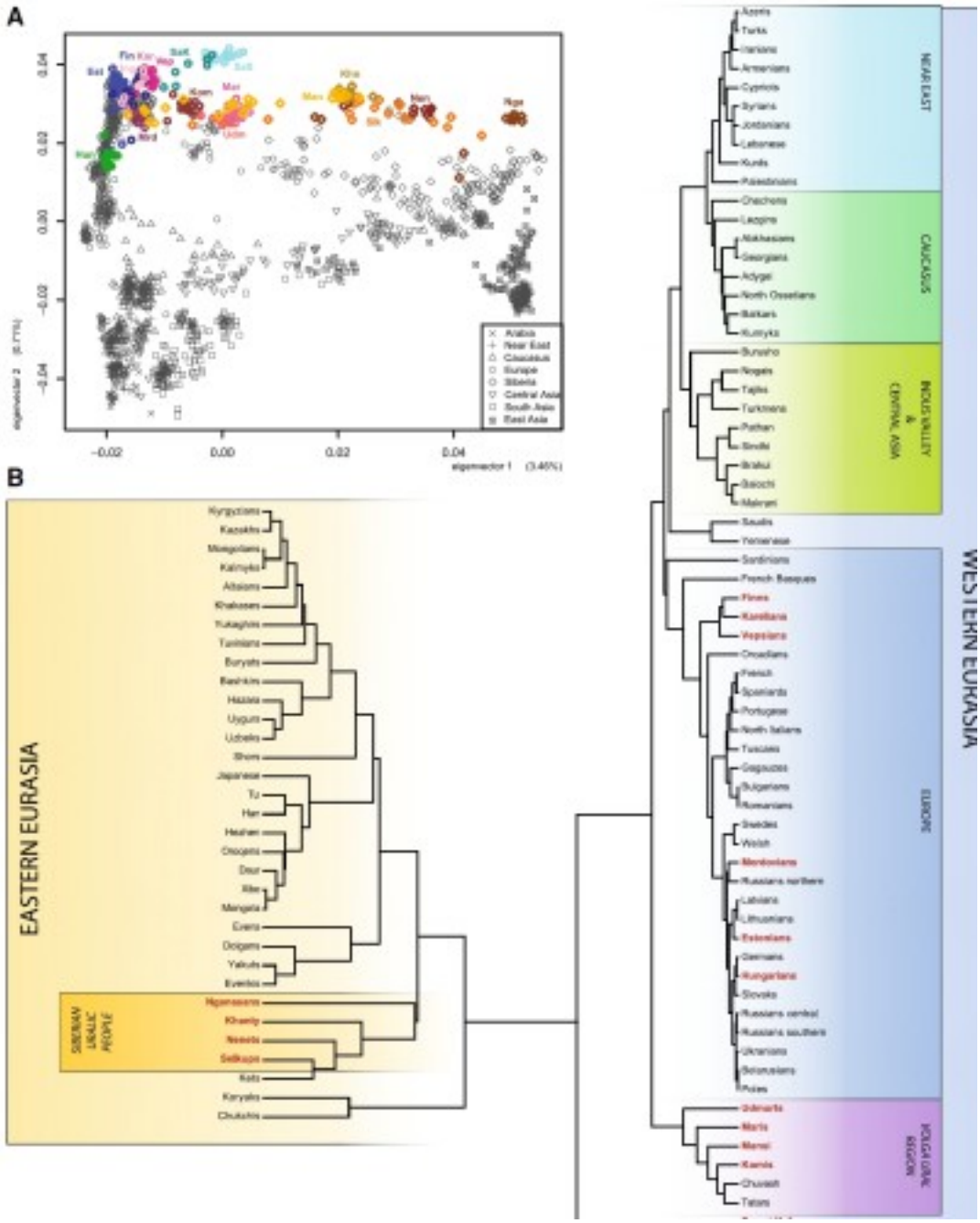
Исследования лингвистов связывают распространение языков уральской семьи с климатическими и культурными изменениями. Можно предположить, что на их распространение повлияли также и демографические события, в частности, миграции населения. Чтобы проверить это, нужно обратиться к генетическому анализу. Предыдущие генетические исследования показали, что демографическая история уралоязычных популяций по материнским и отцовским линиями наследования различна. Митохондриальная ДНК показывала, что соотношение восточных и западных генетических компонентов в уралоязычных популяциях зависит от географии: чем дальше на восток, тем больше восточных линий мтДНК и меньше западных. По Y-хромосоме все иначе, так как больше половины мужчин из популяций уралоязычных народов принадлежит к гаплогруппе N3a, распространенной повсеместно по всей Северной Евразии. Соотношение западных (светло-голубой цвет) и восточных (темно-голубой цвет) линий в Y-хромосомном и митохондриальном генофондах уралоязычных популяций представлено на том же рисунке, на круговых диаграммах.

В данной работе впервые создана и исследована база полногеномных данных (по более чем 500 тысяч позиций однонуклеотидного полиморфизма SNP) для 15 уралоязычных популяций, охватывающих все основные группы языков уральской семьи. Это финны, эстонцы, карелы, вепсы, саамы, мордва, мари, коми, удмурты, ханты, манси, венгры, селькупы, ненцы, нганасаны.

Популяционная структура уралоязычных популяций

Положение уралоязычных популяций в генетическом пространстве Евразии исследователи оценили при помощи анализа главных компонент, заложив в него полногеномные данные по 15 перечисленным выше популяциям. На графике уралоязычные популяции протянулись полосой с запада на восток, расположившись в соответствии с их географическим положением. Это привело авторов к предположению, что главным фактором генетического разнообразия в данном случае выступает география.

Со своими соседями уралоязычные популяции сгруппировались в несколько географических кластеров. Эти кластеры четко выделились на дереве, построенном по величинам генетических расстояний между популяциями.



А. Анализ главных компонент; уралоязычные популяции показаны цветом, в соответствии с группой языков (см. выше). Форма значков соответствует регионам, указанным на легенде. Серым цветом показаны популяции Евразии других языковых семей; В. дерево, построенное по величинам генетических расстояний между популяциями (по аутосомным данным); названия уралоязычных популяций выделены красным. Разными цветами обозначены разные регионы Западной и Восточной Евразии.

Далее с полногеномными данными уралоязычных популяций провели анализ предковых компонентов ADMIXTURE. Авторы отмечают, что при низких значениях числа заданных предковых популяций К по спектру предковых компонентов уралоязычные популяции похожи на своих географических соседей. Но, начиная с К=9 и выше, у них выделяется специфический генетический компонент (его назвали компонентом **k9**, на графиках он показан малиновым цветом). Этот компонент **k9** преимущественно присутствует у носителей уральских языков. Его пик отмечается у носителей западносибирских угорских языков (ханты, манси) и самоедских языков, а также у соседних кетов. Доля **k9** снижается от Западной Сибири к востоку, югу и западу. В уралоязычных популяциях его больше, чем в соседних популяциях других языковых семей. Так в Волго-Уральском регионе **k9** составляет около 40% в финно-угроязычных популяциях и 20% в соседних тюркоязычных популяциях (башкиры, татары, чуваши). У саамов и финнов **k9** снижается до 10-15%, а у эстонцев и у венгров – почти отсутствует. Авторы полагают, что этот компонент **k9** может обозначать некоторую степень генетической общности большинства уралоязычных популяций.

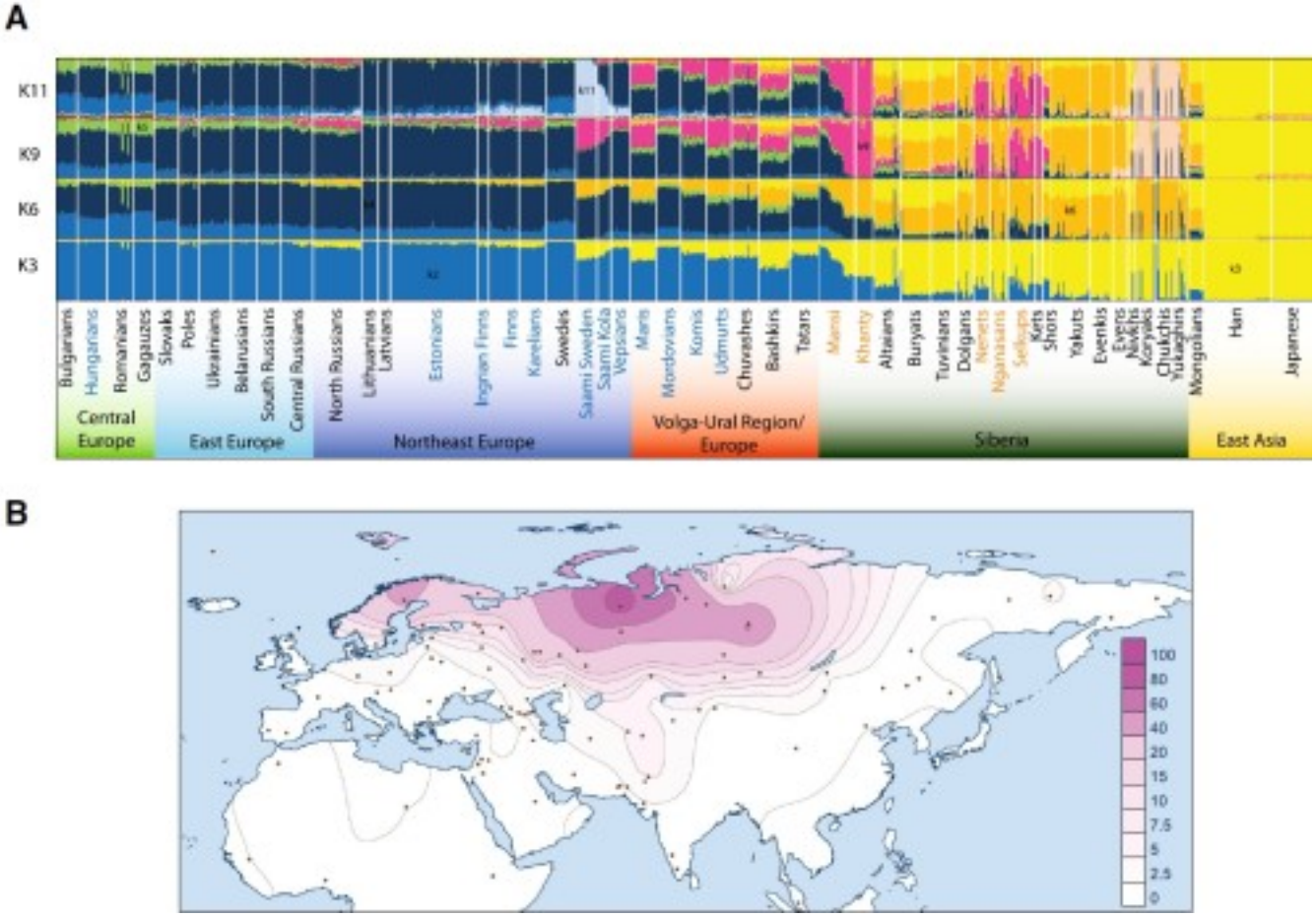
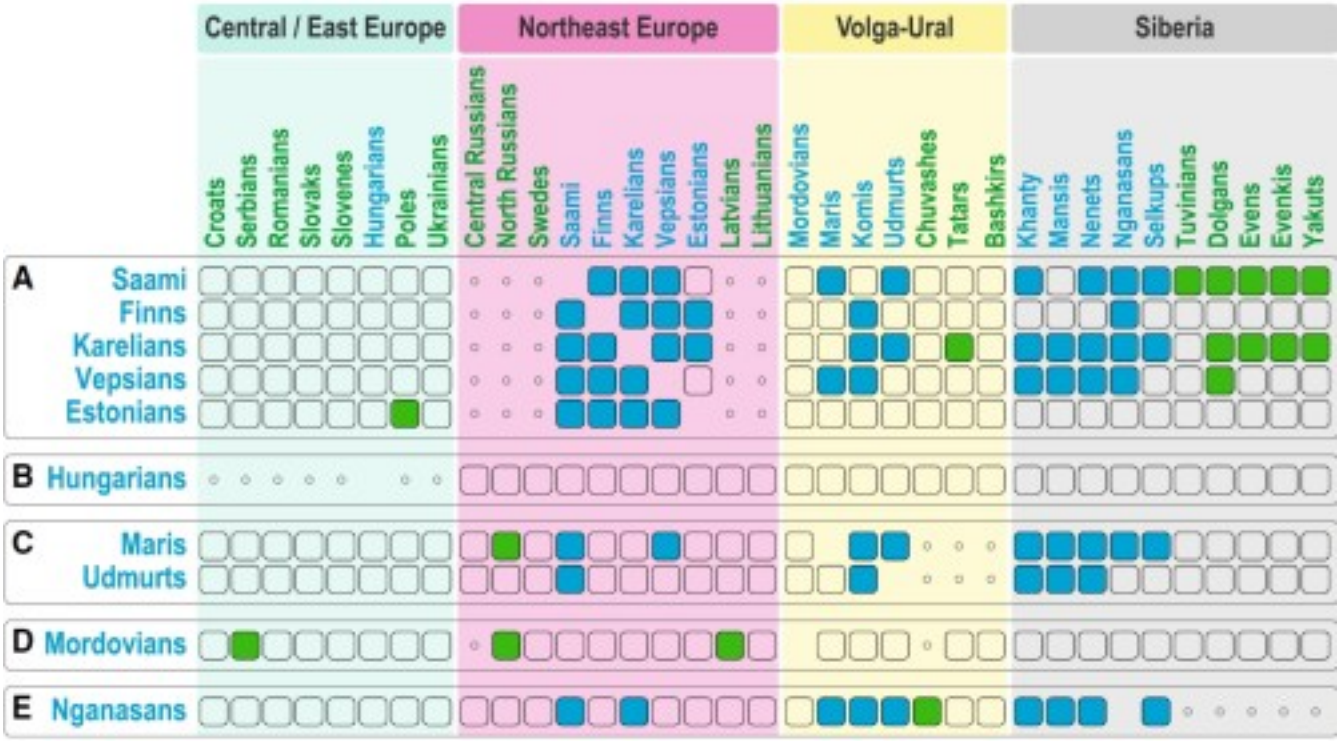


График анализа ADMIXTURE уралоязычных популяций по аутосомным SNP в контексте популяций Евразии (К3, К6, К9, К11). Названия популяций написаны голубым цветом для финно-угорских и оранжевым – для самоедских. Карта частоты компонента

Генетический анализ продолжили сопоставлением в паре геномов из одной и другой популяции сегментов общего происхождения (их называют IBD-сегментами). Задача состояла в том, чтобы сравнить число таких сегментов общего происхождения в парах геномов из двух уралоязычных популяций и в парах геномов из уралоязычной и соседней популяции иной языковой семьи. Результаты анализа авторы представили на рисунке. Уралоязычные популяции здесь указаны голубым цветом, соседние популяции иных языковых семей – зеленым. Большое число общих IBD-сегментов в парах указано на пересечении популяций по горизонтали и вертикали. Большое число IBD-сегментов в парах уралоязычных популяций показано голубыми квадратами; в парах уралоязычных и иных популяций – зелеными квадратами.



Общие IBD сегменты длиной 1-2 cM в парах уралоязычных популяций с уралоязычными и иных языковых семей популяциями. Для каждой уралоязычной популяции, указанной по горизонтали, проведен текст, показывающий, превышает ли число общих IBD сегментов у нее с популяциями, указанными по вертикали, по сравнению с географическими контрольными группами (обозначены мелкими кружочками). Голубым шрифтом обозначены уралоязычные популяции, зеленым шрифтом – популяции иных языковых семей. Большая доля общих IBD сегментов с уралоязычными популяциями обозначена голубым квадратом; большая доля общих IBD сегментов с популяциями иных языковых семей обозначена зеленым квадратом.

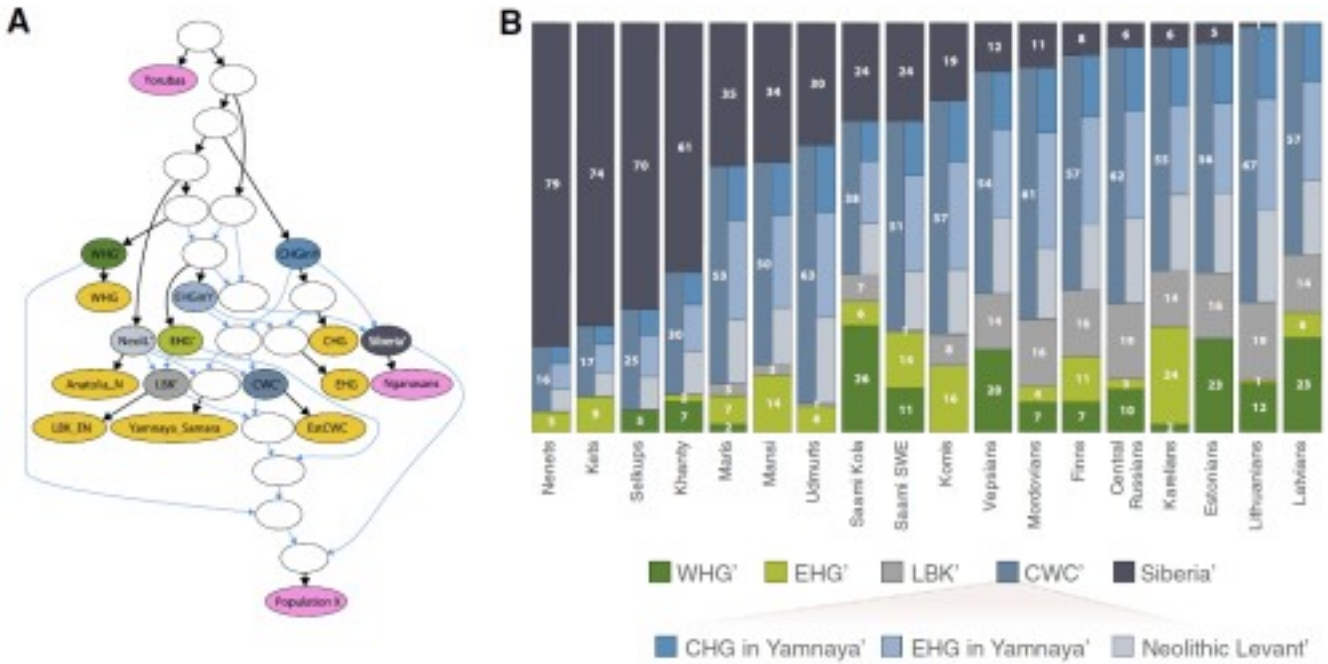
В целом, уралоязычные популяции разделяют больше IBD сегментов друг с другом, чем со своими соседями других лингвистических семей. Так, у мари и удмуртов больше число общих IBD сегментов с хантами и манси, живущими по другую сторону Урала, чем с соседними чувашами, татарами и башкирами. У финнов и саамов число общих IBD сегментов больше с географически отдаленными мари, коми и удмуртами и даже с западносибирскими хантами и манси, чем с географически близкими шведами, латышами, литовцами, северными русскими. Нганасаны разделяют больше IBD сегментов со всеми изученными сибирскими уралоязычными популяциями, большинством уралоязычных популяций в Волго-Уральском регионе и даже с саамами и карелами в Северной Европе, чем со своими соседями других языковых семей.

Авторы расценивают эти результаты как доказательство некоторого общего генетического субстрата среди большинства уралоязычных популяций. Исключение составляют венгры и мордва, которые не показали превышение доли общих IBD сегментов с уралоязычными популяциями. Что касается эстонцев, то они разделяют больше общих IBD сегментов только с соседними уралоязычными популяциями (финны, карелы, вепсы), но не с географически отдаленными.

В работе использованы и другие виды анализа. Так, методом fineSTRUCTURE индивиды были сгруппированы в кластеры, которые оказались соответствующим этническим группам по самоидентификации, а кластеры более высокого порядка соответствовали географическим регионам.

Близость уралоязычных популяций с древними евразийцами

Исследователи вычислили доли предковых компонентов от древних групп населения у всех изученных популяций. Они выделяют компоненты западноевропейских охотников-собирателей (WHG), восточноевропейских охотников-собирателей (EHG), неолитические европейские культуры линейно-ленточной керамики (LBK) и шнуровой керамики (CWC) (которую подразделяют еще на три источника), и, наконец, компонент, который они называют сибирским. Этот сибирский компонент присутствует в сибирских популяциях и почти во всех уралоязычных популяциях, но его доля падает с востока на запад (у латышей – до нуля).



Соотношение предковых генетических компонентов древних популяций в уралоязычных и некоторых соседних с ними популяциях. WHG — западноевропейские охотники-собиратели, EHG — восточноевропейские охотники-собиратели, LBK — культура линейно-ленточной керамики, CWC — культура шнуровой керамики, CHG in Yamnaya – компонент кавказских охотников-собирателей в составе ямного, EHG in Yamnaya компонент восточноевропейских охотников-собирателей в составе ямного, Neolithic Levant – неолит Леванта, Siberia – сибирский компонент.

Связь между лингвистикой, географией и генетикой

Исследователи вычислили корреляцию между лингвистическими, географическими и генетическими данными. Для этого они применили тест Мантеля, заложив в него лексические расстояния между языками, географические расстояния между популяциями и генетические расстояния между популяциями.

Оказалось, что лексические расстояния между уральскими языками положительно коррелировали со всеми типами генетических расстояний. Лексические расстояния (так же, как и генетические расстояния) увеличивались с увеличением географических расстояний. Когда географический фактор зафиксировали, корреляция лексических расстояний с генетическими расстояниями осталась достоверной, но только с вычисленными по аутосомным маркерам (а не по мтДНК и Y-хромосомам).

Итак, полногеномный анализ 15 популяций уральской языковой семьи привел авторов к следующим заключениям:

- В первом приближении характер генетического разнообразия уралоязычных популяций соответствует географии.
- Несмотря на географическое разнообразие, в большинстве уралоязычных популяций присутствует общий генетический компонент, выявляемый методом ADMIXTURE (**k9**). Авторы предполагают его сибирское происхождение.
- Большинство уралоязычных популяций имеют больше сегментов генома общего происхождения со своими лингвистическими родственниками, чем соседние с ними популяции других языковых семей.
- В уралоязычных популяциях найдена достоверная положительная корреляция между лексическими и генетическими данными.

Это приводит исследователей к представлению, что распространение уральских языков, по меньшей мере, частично происходило путем миграций групп населения, что и привело к некоторой генетической общности.

Авторы статьи вносят свой вклад и в представление об источниках формирования современного генофонда Европы. Так, помимо трех главных миграционных волн: в верхнем палеолите (охотники-собиратели), в неолите (ближневосточные земледельцы) и ранней бронзе (степные кочевники – ямники), они выделяют четвертый источник – сибирский компонент.

Именно этому сибирскому компоненту, который распространяется с востока на запад, снижаясь по величине, они отводят,

предположительно, основную роль в распространении уральских языков.

текст Надежды Маркиной

Источник:

Genes reveal traces of common recent demographic history for most of the Uralicspeaking populations

Kristiina Tambets, Bayazit Yunusbayev, Georgi Hudjashov, Anne-Mai Ilumäe, Siiri Rootsi, Terhi Honkola, Outi Vesakoski, Quentin Atkinson, Pontus Skoglund, Alena Kushniarevich, Sergey Litvinov, Maere Reidla, Ene Metspalu, Lehti Saag, Timo Rantanen, Monika Karmin, Jüri Parik, Sergey I. Zhadanov, Marina Gubina, Larisa D. Damba, Marina Bermisheva, Tuuli Reisberg, Khadizhat Dibirova, Irina Evseeva, Mari Nelis, Janis Klovins, Andres Metspalu, Tõnu Esko, Oleg Balanovsky, Elena Balanovska, Elza K. Khusnutdinova, Ludmila P. Osipova, Mikhail Voevoda, Richard Villems, Toomas Kivisild and Mait Metspalu

Genome Biology2018**19**:139

https://doi.org/10.1186/s13059-018-1522-1

отклики в СМИ:

Пресс-релиз по статье, опубликованный на сайте «Открытая наука» https://openscience.news/posts/1436-uchenye-nashli-u-narodov-uralskoy-yazykovoy-semi-geneticheskuyu-obschnost

В передаче «Гранит науки» на «Эхе Москвы» https://echo.msk.ru/programs/granit/2299134-echo/

Пресс-релиз по статье на англ. языке на сайте международных научных новостей EurekAlert

https://www.eurekalert.org/pub_releases/2018-10/arsc-sfc101518.php