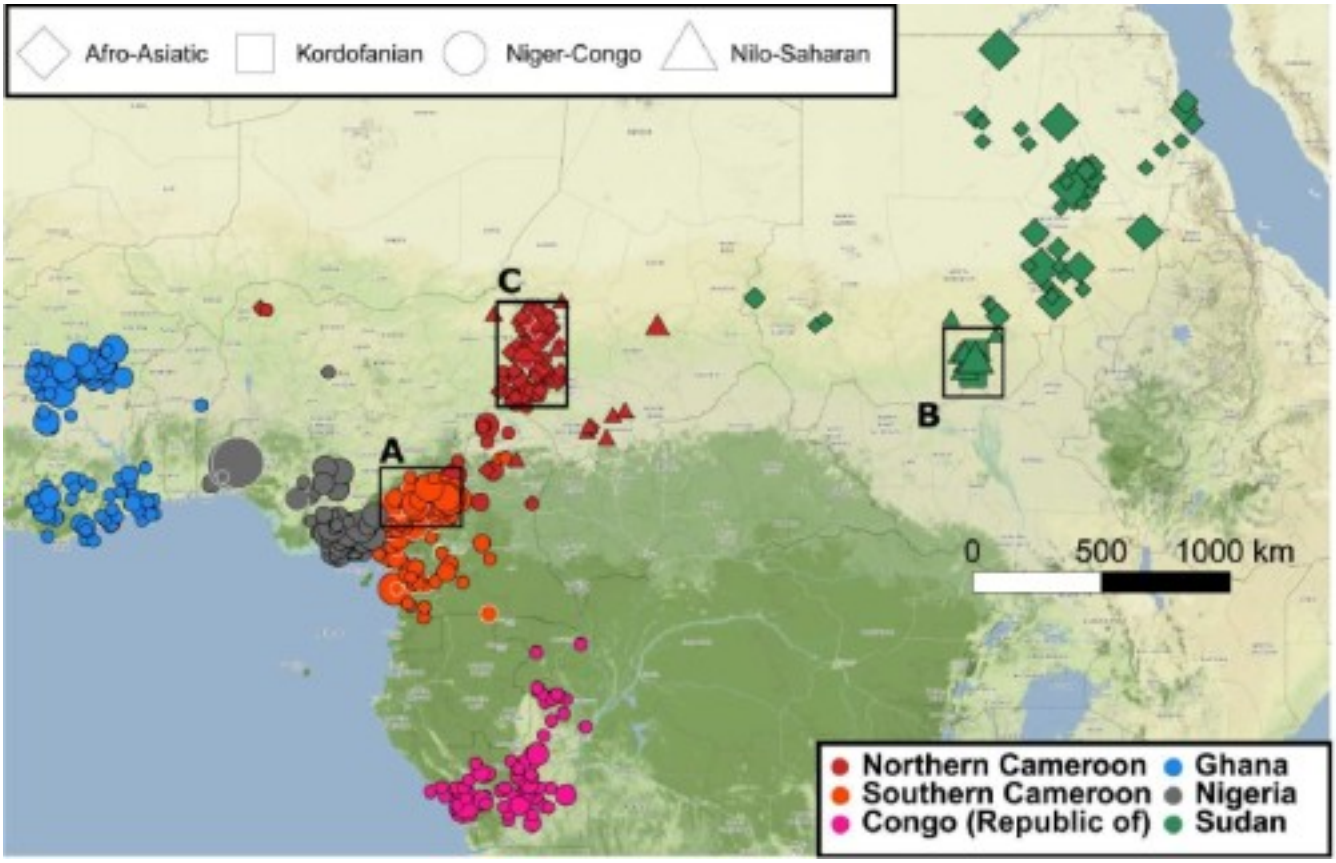
Новые детали к геномному ландшафту Африки

Анализ полногеномных данных, полученных при обследовании более 150 этнических групп пяти стран Африки позволил выявить генетическую структуру африканских популяций на недоступном прежде уровне. Авторы связали их с данными археологии, лингвистики и истории, выявили отражение в геномах демографических и исторических событий. В частности, описали генетический след арабских популяций и проследили экспансию языков банту по африканскому континенту.

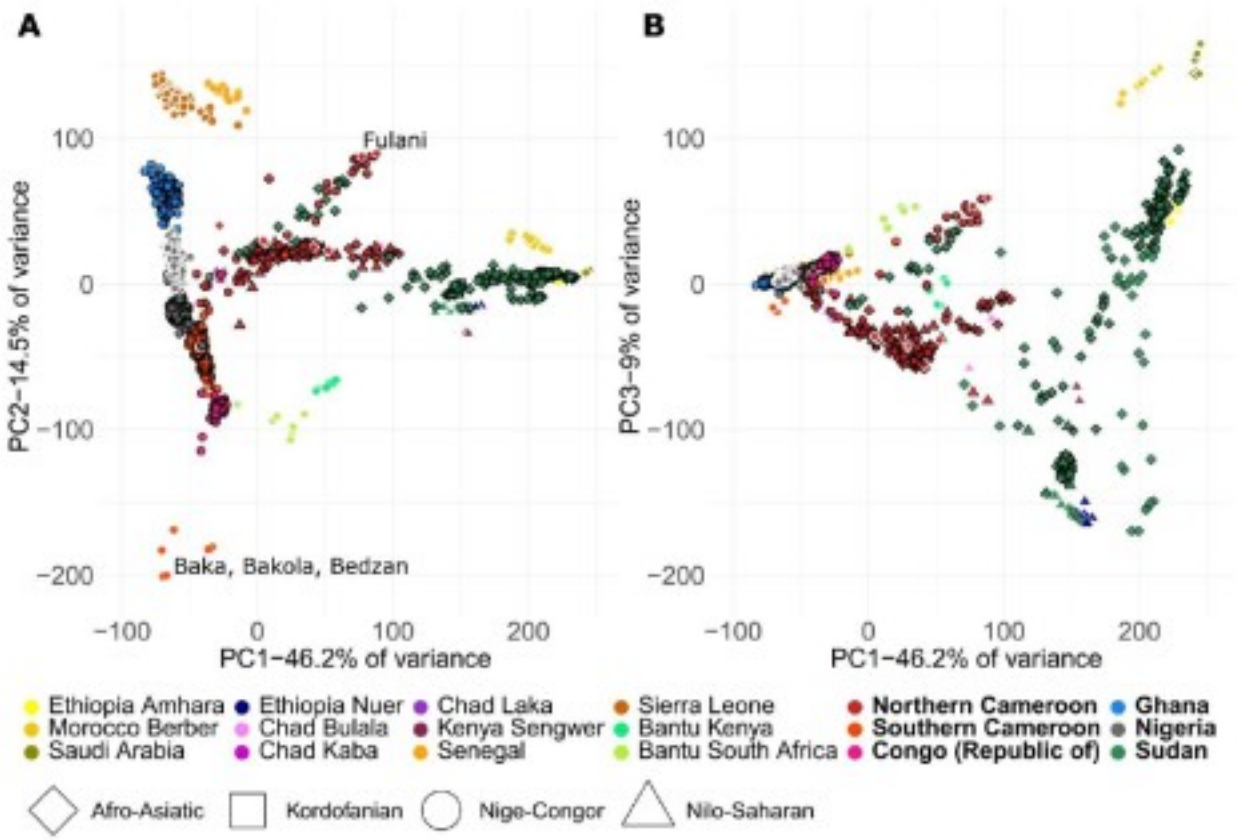
Предыдущие исследования показали, что смешение между географически разными популяциями играло важную роль в формировании геномного разнообразия современного населения Африки. В новой работе, результаты которой опубликованы в журнале Science Advances, генетики анализируют данные по SNP полиморфизму 1387 индивидов из пяти африканских стран: Камеруна, Республики Конго, Ганы, Нигерии и Судана и в дополнение 54 индивидов из трех других стран. Эти данные были собраны в полевых исследованиях от 1997 до 2006 года.

Обследованные представляют 166 этнических групп, относящихся к трем из четырех основных лингвистических ветвей Африки (афро-азиатской, нигеро-конголезской и нило-сахарской), а также к некоторым языковым изолятам в Судане. Эти популяции распределены от востока до запада африканского континента и проживают в разных природных условиях, от долины Нила до гор и дождевых лесов. Эти данные авторы проанализировали вместе с опубликованными генетическим данными по 297 индивидам из современных популяций мира и 20 древним африканским геномам.



Географическое расположение изученных образцов ДНК из Камеруна, Республики Конго, Ганы, Нигерии и Судана (разные цвета). Форма значков соответствует лингвистической принадлежности популяций, размер отражает размер выборки (Bird et al., 2023).

Исследователи провели анализ главных компонент (PCA), основанный на анализе гаплотипов, по 1333 геномам из пяти африканских стран в контексте других геномных данных из Африки и Саудовской Аравии. Анализ продемонстрировал сильную корреляцию между генетикой и географией. По первой компоненте прослеживается клина от Саудовской Аравии к западноафриканским нигеро-конголезским популяциям; вторая компонента отражает географию от Съерра-Леоне до Конго. Охотники-собиратели дождевых лесов Камеруна образуют отдельный кластер, что подтверждается результатами ADMIXTURE.



Основанный на анализе гаплотипов PCA по новым и ранее опубликованным геномным данным. Цвет значков соответствует стране, форма – лингвистической принадлежности популяции (Bird et al., 2023).

В дальнейшем анализе авторы использовали метод генетических расстояний, сравнивая их между парами 1) этнических групп, 2) лингвистических групп и 3) индивидов с определенным местом рождения. Использовали IBD анализ для поиска сегментов общего происхождения в парах геномов из разных популяций; метод fineSTRUCTURE для выявления корреляции между этногеографическим происхождением и генетикой.

Полученные результаты представили картину генетического ландшафта Африки с недоступным прежде разрешением, позволили выявить события смешения, связать генетику с данными археологии, лингвистики и истории, а также проследить в геноме следы известных демографических и политических событий.

В большинстве случаев данные продемонстрировали корреляцию генетики с географией, этничностью и языком в пределах каждой из стран: Камерун, Республика Конго, Гана, Нигерия и Судан. Генетический профиль обычно коррелировал с принадлежностью к основным лингвистически ветвям, а иногда и к малым языковым группам. В Камеруне анализ fineSTRUCTURE различил этнические группы, живущие на расстоянии менее 20 км друг от друга. Вместе с тем, здесь наблюдалась относительно слабая корреляция с географией.

В общую картину не укладывается население Судана, принадлежащее к арабским и нубийским этническим группам и проживающее в долине Нила. В их генетическом разнообразии гораздо большую роль играет география, чем этничность. Авторы объясняют это особенностями долины Нила, который служил коридором для генетических потоков и способствовал смешению.

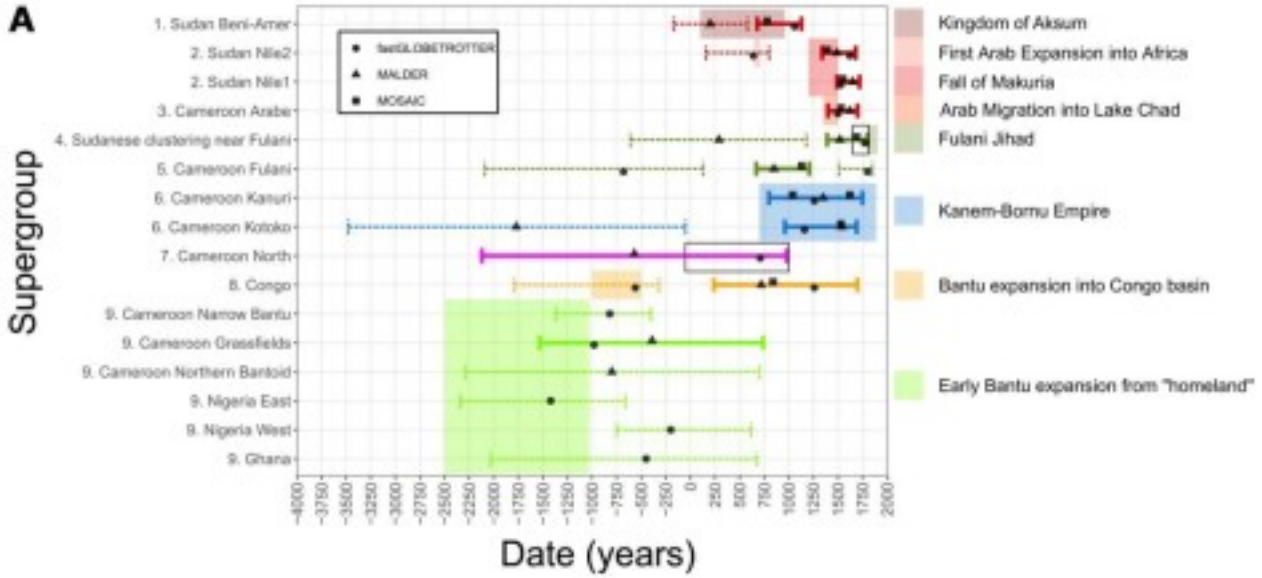
По контрасту с отсутствием генетической структуры в популяциях Судана, проживающих вдоль Нила, авторы описали четкую

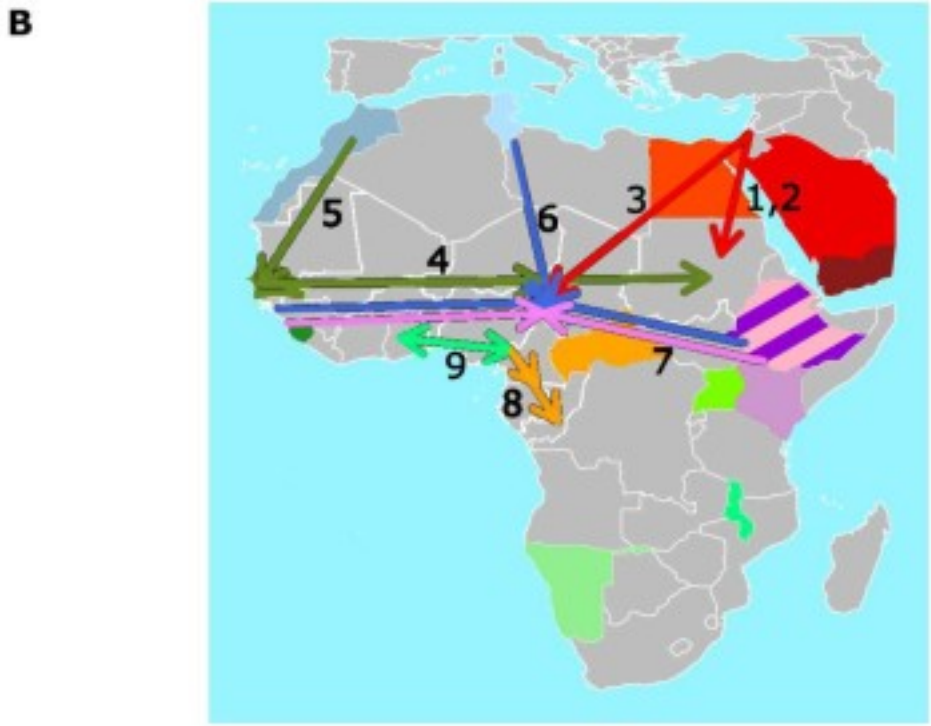
генетическую структуру в популяциях, проживающих в Нубийских горах в Южном Кордофане на территории того же Судана, которая соответствовала этнолингвистическим группам. Выявлены генетические различия между группами, говорящими на кордофанской и говорящими на нило-сахарской ветви языков. При этом в регионе отмечается довольно слабая корреляция между генетикой и географией.

Генетический анализ позволил выявить исторически обусловленное смешение африканских популяций с арабскими. Так, оно обнаружено в Южном Кордофане, регионе Судана, где отражает арабское нашествие в Судан в VII веке н.э. Обнаружено также смешение с арабами на севере Камеруна, датируемое XVI веком, это отсылает к миграциям арабских групп на озеро Чад в середине XIV века.

Вычисление эффективного размера популяций позволило проследить их демографическую историю. Археология говорит, что около 600 лет н.э. популяции в бассейне реки Конго испытали радикальное сокращение, а примерно 800 лет спустя началась их вторичная экспансия. Генетики попытались найти следы этих демографических событий в геномах по изменению эффективного размера популяций. Они показали, что некоторая экспансия началась 60 поколений назад, затем 20 поколений назад она сменилась упадком, вслед за которым экспансия продолжилась. В то же время не было найдено признаков древнего демографического коллапса.

Описаны генетические следы возникновения и распада политических образований, например, империи Канем на севере Камеруна (от 700 до 1800 лет н.э.).





A. Хронология событий смешения, вычисленная по методам fastGLOBETROTTER (кружки), MALDER (треугольник и), and MOSAIC (квадраты). Слева указаны популяции, справа – исторические события, соответствующие эпизодам генетического смешения. B. Карта миграций, которые привели к указанным событиям смешения. Цвет стрелок соответствует цвету линий на рисунке A (Bird et al., 2023).

Генетическими методами авторы проследили одно из важнейших популяционных событий в Африке – экспансию популяций, говорящих на языках банту. На основании данных лингвистики, которые говорят, что экспансия банту-говорящих популяций началась в регионе Камеруна/Нигерии, они выбрали этническую группу бамилеке в качестве оптимального генетического компонента, отражающего этот процесс. Оценивая долю этого компонента в населении других регионов, проследили пути экспансии банту-говорящих популяций по Африке. Полученные данные подтверждают, что банту-говорящие группы сначала мигрировали на юг через Конго, и возможно, южнее, в Анголу. За этим следовала миграция на восток, вероятно, до Малави и Мозамбика. Произошло расщепление банту-говорящих популяций на восточную и западную ветви с охватом почти всего континента. Новые данные говорят в пользу множественных событий смешения мигрирующих банту-говорящих групп с локальными популяциями, они происходили от 560 до н.э. до 1260 н.э.

В целом, полученные результаты помогли более глубоко связать генетические данные с данными археологии, лингвистики и истории. С их учетом последующие GWAS исследования должны будут выявить, как африканские популяции различаются по частоте фармакологически значимых аллелей.

текст Надежды Маркиной

Источник:

Nancy Bird et al. Dense sampling of ethnic groups within African countries reveals fine-scale genetic structure and extensive historical admixture // Science Advances. 29 Mar 2023. Vol 9, Issue 13. doi: 10.1126/sciadv.abq2616.