Миграционные волны и смешения сформировали неоднородный генетический ландшафт США

Проведено масштабное популяционно генетическое исследование населения Соединенных штатов Америки. Геномные данные 32 с половиной тысяч человек, участников проекта Genographic, использовали для выявления генетического разнообразия, происхождения и картирования разных популяций. По анализу гаплотипов выявлены генетически неоднородные кластеры в разных этнических группах. Показано разнообразие населения испано/латинского происхождения, генетический барьер в пределах афроамериканского населения, высокий уровень гомозиготности в группах азиатского происхождения и географическое распределение американцев европейского происхождения из разных стран.

Население Соединенных штатов Америки – это конгломерат, образованный масштабными миграциями с других континентов и смешением мигрантов с коренным населением Северной Америки. Эти миграции продолжаются в течение последних двух веков, а в пределах территории страны происходит смешение и перемещение разных групп. Все это формирует генетическое разнообразие современных американцев. Предыдущие генетические исследования населения США по большей части охватывали американцев европейского происхождения, афроамериканцев и американцев испано/латинского происхождения, но даже в этих группах генетический ландшафт не достигал высокой степени разрешения. Мало внимания было уделено более малочисленным этническим группам, например, азиатского и ближневосточного происхождения.

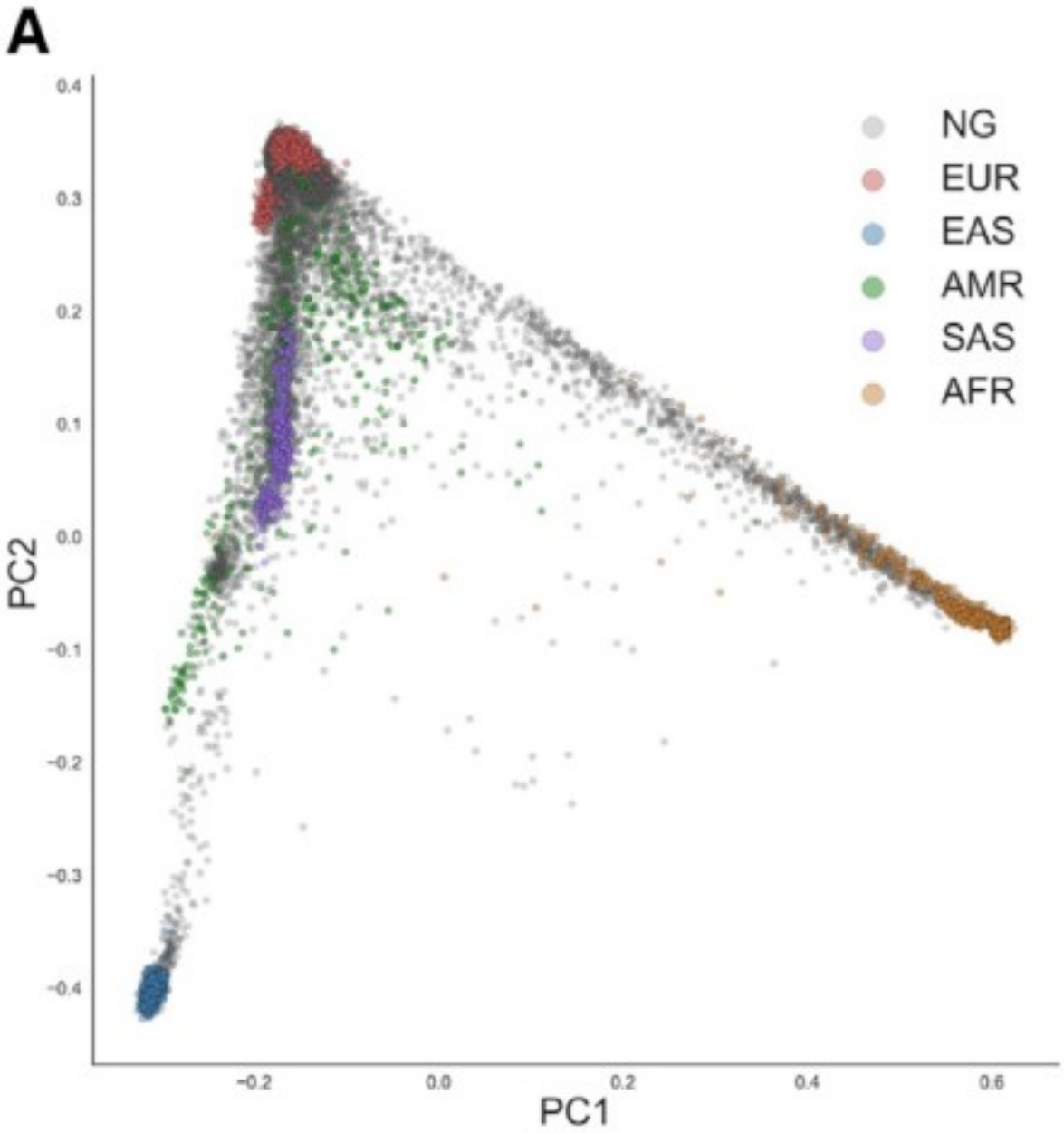
Наиболее полное популяционно-генетическое исследование, выполненное под руководством специалистов из Института Брода Массачусетского технологического института и Гарвардского университета, опубликовано в журнале American Journal of Human Genetics. *Авторы с*обрали полногеномные данные от 32 с половиной тысяч человек, жителей США, которые были обследованы в рамках международного проекта Genographic. Это самый крупный геногеографический проект, начавшийся в 2005 году и охвативший более миллиона человек по всему миру.

Из базы проекта исследователи взяли геномные данные 32 589 человек, проживающих в разных регионах США, охвачены четыре региона страны, обозначенные в государственной переписи США как Юг, Северо-Восток, Средний Запад и Запад. Для этих индивидов объединили геномную информацию (результаты генотипирования по панелям в 150 тысяч SNP), географию мест происхождения участников (места рождения их предков в трех поколениях) и этнические данные по самоидентификации. Для сопоставления использовали базы данных проектов 1000 Genomes Project и Human Genome Diversity Project (HGDP).

Анализируя эти данные, авторы определяли доли различного происхождения для разных групп населения, пропорции смешения и генетическую структуру популяций; на карте генетического ландшафта страны выявляли зоны интенсивных миграций и генетические барьеры. Использовали методы анализа главных компонент, выделение кластеров гаплотипов, анализ identify-by-descent (IBD); для визуализации популяционной структуры применяли метод fineSTRUCTURE, для определения зон генетического обмена и генетических барьеров — метод оценки эффенктивных миграций (EEMS).

Индивиды европейского происхождения составили 78% в когорте (это почти точно соответствует доле «белых» по данным переписи США) и преобладали на Среднем Западе (82,8%). Индивиды смешанного американского происхождения чаще проживали на Западе и Юге (9,7% и 7,8% соответственно). Афроамериканцы наиболее часто проживали на Юге (3,2%), затем – на Северо-Западе (3,0%). Авторы отмечают, что афроамериканцев очень много в штате Калифорния и в то же время почти нет в штатах Айдахо, Мэйн, Небраска, Северная Дакота, Южная Дакота и Вайоминг. Индивиды восточноазиатского происхождения главным образом проживали на Западе (2,1%), а южноазиатского – на Северо-Востоке (1,0%).

Провели анализ главных компонент геномных данных изученной когорты жителей США и геномных данных из проекта 1000 Genomes Project с разных континентов.



Анализ PCA изученной выборки населения США (NG, серые точки) и данных проекта 1000 Genomes Project (цветные точки, соответствуют цвету на легенде) EUR – Европа, EAS – Восточная Азия, AMR — Америка, SAS – Южная Азия, AFR – Африка.

В исследованной когорте жителей США выделили 25 генетических кластеров европейского, испано/латинского, африканского и азиатского происхождения. При сравнении пар геномов по числу сегментов общего происхождения (IBD анализ) внутри каждого кластера их число было выше, чем между кластерами. Для каждой группы кластеров описаны различия и обнаружены определенные закономерности. Так, американцы испано/латинского происхождения образуют четыре генетически отличных кластера, по большей части сосредоточенные в штатах Калифорния, Нью Мехико, Техас и Флорида, и отдельно пуэрториканский кластер. Индивиды в этих кластерах различаются по составу смешанных в них генетических компонентов (индейского, европейского и африканского). Например, в кластере Калифорнии ниже доля европейского генетического компонента и выше доля индейского компонента из Центральной и Южной Америки, чем в кластерах Нью Мехико или Техаса.

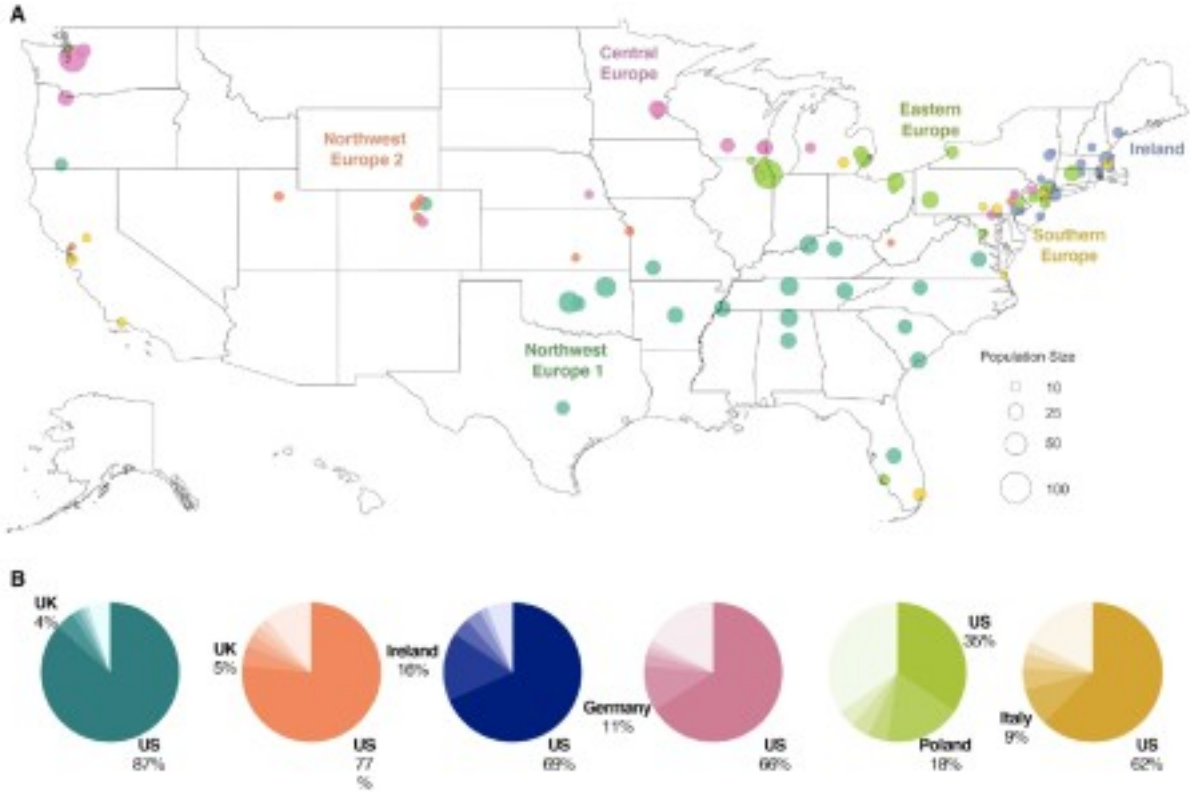
Генетическая структура афроамериканцев складывалась под влиянием особенностей трансатлантической работорговли и последующей расовой сегрегации общества, которая была очень различной в тех или иных штатах. Афроамериканцы генетически образуют два кластера – южный и северный, первый географически привязан к южным штатам, второй – к северным и западным. Средняя доля африканского компонента выше в северном кластере, чем в южном. Исследователи обнаружили генетический барьер, то есть барьер, препятствующий миграциям и потокам генов, проходящий вдоль гор Аппалачи. Этот барьер примерно соответствует границе между штатами, где практиковалось рабовладение, и свободными штатами. Он также разделяет штаты, в которых была узаконена расовая сегрегация, и штаты, где она была запрещена.

Методом EEMS определены зоны генетического обмена (голубые цвета) и генетические барьеры (коричневые цвета) для афроамериканцев.

До сих пор в генетических исследованиях населения США была недооценена группа американцев азиатского происхождения, в последнее время быстро растущая, сейчас она составляет 5% населения. Ученые выявили множество генетически различных южноазиатских и восточноазиатских субкластеров на Западе, Северо-Востоке и Юге. В восточноазиатских субкластерах преобладало происхождение из Китая, Тайваня, Японии и Кореи; в юго-восточноазиатских – с Филиппин. Южноазиатские субкластеры были в основном представлены выходцами из Индии и Пакистана, при этом в них наблюдалась генетическая дифференциация, отражающая популяционную структуру этих стран. Для индивидов, входящих в азиатские кластеры, менее 30% предков в третьем поколении родилось в США, что говорит о недавнем времени миграции.

Исследователи отмечают, что для азиатских, также как для ближневосточных групп населения, характерен повышенный уровень гомозиготности. Они полагают, что в популяциях южноазиатского происхождения большая доля гомозиготности может отражать кастовую систему общества, как например, в Индии. В то же время большая доля гомозиготности в популяциях ближневосточного и юго-восточноазиатского происхождения может отражать практику близкородственных браков в этих регионах.

Что касается американцев европейского происхождения, то авторы выявили неравномерность в распределении по территории США потомков разных европейских популяций. Массовые европейские миграции в США в XIX веке можно разделить на две волны. Мигранты из Северной и Западной Европы образовали первую волну, с 1840-х по 1880-х гг., а во второй волне, с 1880-х по 1910-х гг., преобладали мигранты из Южной и Восточной Европы. Среди американцев европейского происхождения авторы выделили несколько генетических кластеров, названных по регионам Европы либо по отдельным странам. Эти кластеры характеризуются определенным географическим проживанием в регионах США. В каждом кластере, в зависимости от времени его появления, различается доля людей, предки которых в трех поколениях родились в США, и доля людей, дедушки и бабушки которых родились в других странах. На рисунке показано географическое расположение и доля индивидов разного происхождения для двух кластеров Северо-Западная Европа, Центральная Европа, Восточная Европа, Южная Европа и Ирландия.



А. Географическое распределение гаплотипических кластеров американцев европейского происхождения. Каждая точка представляет округ, в котором в наибольшем числе проживают индивиды данного кластера. В. Для каждого кластера показаны доли происхождения (по месту рождения предков в трех поколениях).

Если рассматривать субпопуляции в пределах этих кластеров, то и они характеризуются определенным географическим распределением. Так, потомки финнов и скандинавов в основном проживают на Среднем Западе и в штате Вашингтон, французские канадцы – на Северо-Востоке, акадианцы (переселенцы из Канады) – на Северо-Востоке и в штате Луизиана, потомки итальянцев, греков и евреев ашкенази – в Нью-Йоркской агломерации, поляков – на Юго-Западе.

Необходимость проведения детального популяционно-генетического исследования такой страны как США, очевидна. Помимо демографических и антропологических целей, преследуются и биомедицинские цели. Любые медико-генетические обследования населения, например, на генетические варианты, ассоциированные с заболеваниями, клинические исследования лекарств, исследования лекарственной чувствительности и эффективности, не должны ограничиваться только самой крупной и достаточно гомогенной популяцией американцев европейского происхождения, но включать и другие группы населения, в том числе немногочисленные.

Ученые говорят, что эти результаты – всего лишь верхушка айсберга. По мере того как число участников подобных исследований будет расти, картина популяционной истории США приобретет большую четкость и ясность. По словам руководителя работы Алисии Мартин, представляющей Институт Брода, генетики собираются в будущем довести число геномных данных до 250 тысяч.

О работе, в которой была исследована генетическая история 12 популяций Северной и Южной Америки, можно прочитать на сайте.

текст Надежды Маркиной

Источник:

Chengzhen L. Dai et al. Population Histories of the United States Revealed through Fine-Scale

Migration and Haplotype Analysis // The American Journal of Human Genetics 106, 371–388, March 5, 2020
https://doi.org/10.1016/j.ajhg.2020.02.002