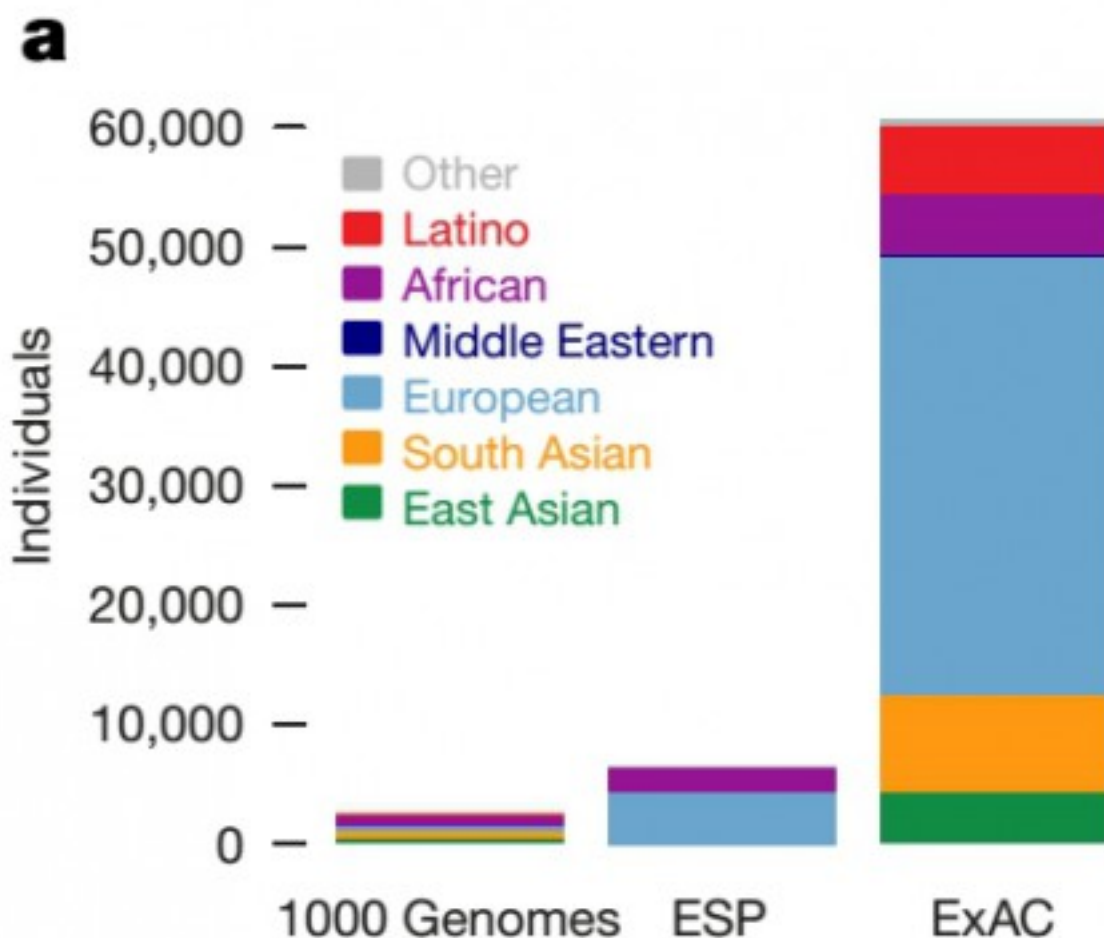


60 тысяч экзомов в каталог генетического разнообразия

Проект по секвенированию более 60 тысяч экзомов (часть генома, кодирующая белки) в популяциях на разных континентах выявил гены, устойчивые к мутированию, показал, сколько носимых нами мутаций полностью блокируют синтез белка, а также значительно приблизил специалистов к пониманию природы редких заболеваний.

Опубликованы результаты проекта по секвенированию экзомов (часть генома, кодирующая белки), над которым работал международный консорциум — [Exome Aggregation Consortium \(ExAC\)](#), ведущие организации консорциума – Broad Institute (относящийся к Массачусетскому технологическому институту) и Гарвардский университет. В рамках этого проекта было секвенировано более 60 тысяч экзомов из разных популяций мира, эти данные находятся [в открытом доступе на сайте ExAC](#). Основная публикация по результатам проекта [вышла в журнале Nature](#) (ведущий автор Daniel MacArthur) и еще две – в журналах [Genetics in Medicine](#) (ведущий автор Roddy Walsh) и [Nature Genetics](#) (ведущий автор Douglas Ruderfer). К сожалению, среди участников консорциума ExAC, представляющих 126 научных организаций, нет российских ученых, так же как среди анализируемых экзомов нет российских.

Прочитанные 60 706 экзомов принадлежат людям из разных популяций пяти континентальных групп: Европы, Африки, Южной Азии, Восточной Азии и Латинской Америки. В пределах Европы авторы отделили финнов от всех остальных европейцев из-за того, что финская популяция прошла через узкое бутылочное горлышко (и, соответственно, снижение генетического разнообразия). В результате они получили исключительно обширную и детальную базы данных по глобальному генетическому разнообразию, которая на порядок превышает имеющиеся до сих пор.



Размер базы данных по экзомам для различных континентальных групп (показаны разными цветами), полученных в проектах

«1000 геномов», ESP (Exome Sequencing Project) и ExAC).

b

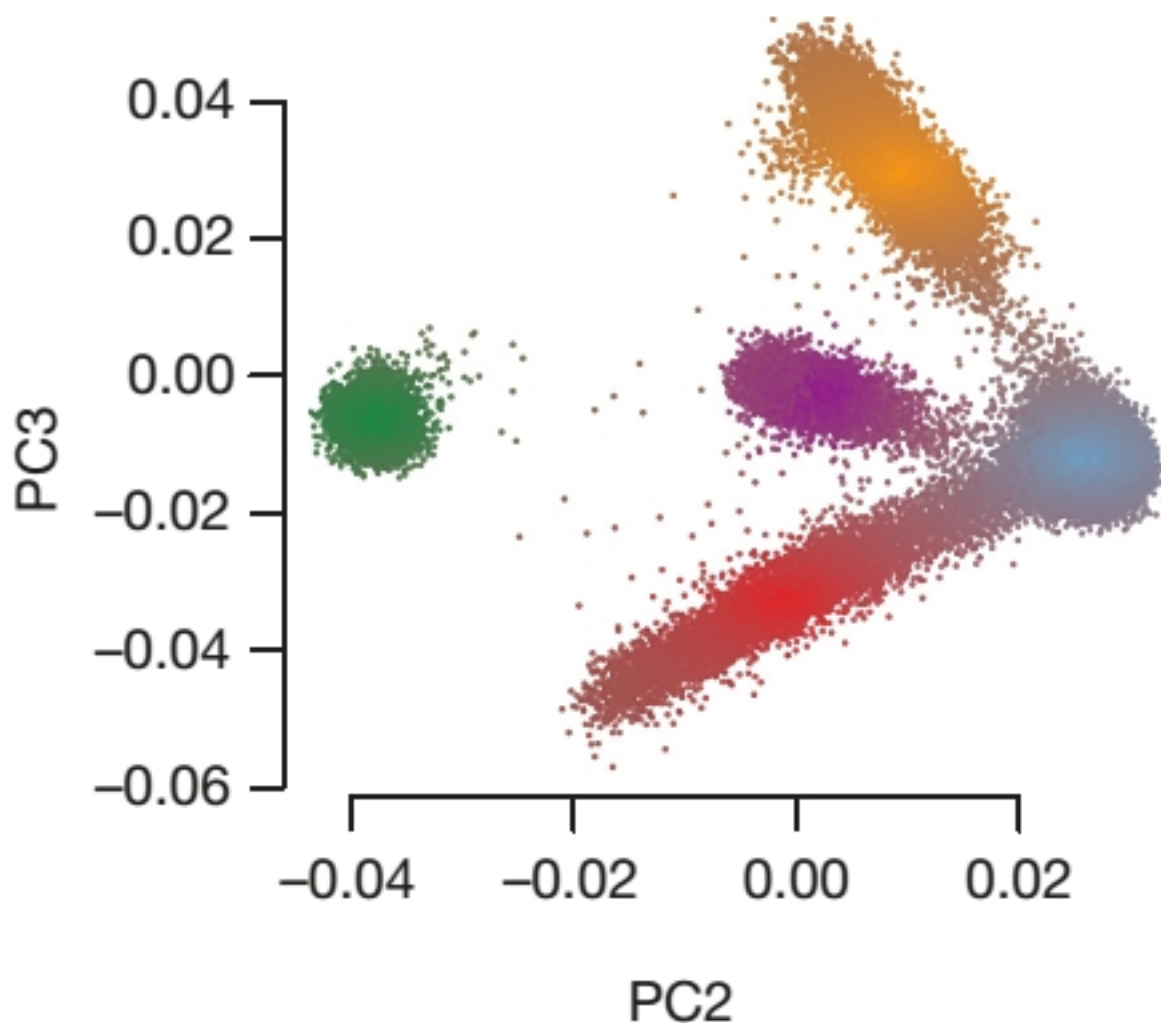
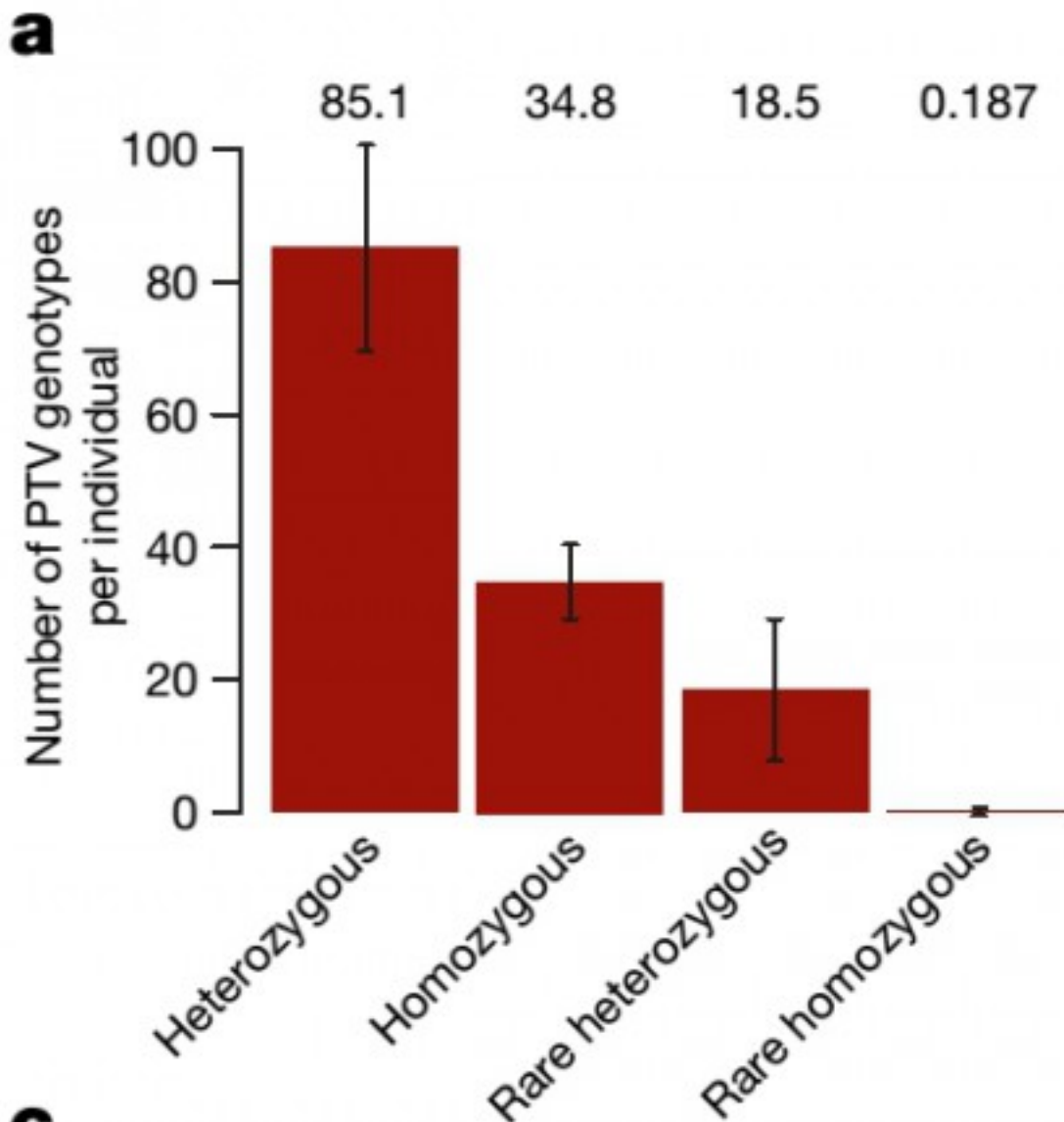


График анализа главных компонент (PCA) группирует исследуемые экзомы в пять континентальных групп: голубой цвет – Европа, фиолетовый – Африка, желтый – Южная Азия, зеленый – Восточная Азия, красный – Латинская Америка.

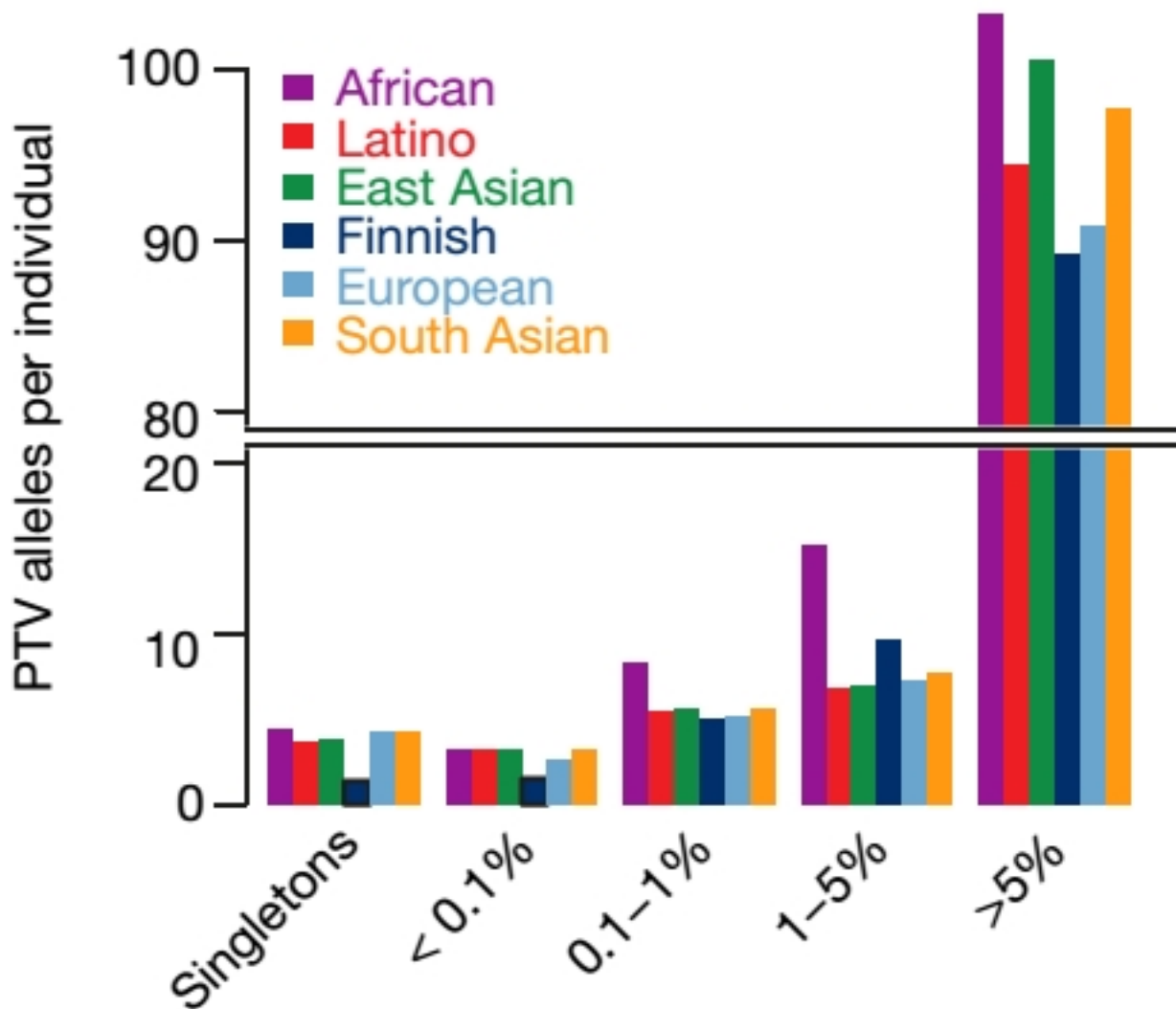
Благодаря большому числу данных удалось добиться «высокого разрешения» каталога геномного разнообразия и открыть закономерности, которые было невозможно увидеть на меньшем материале. Ученые описали более 7,4 млн кодирующих белки генетических вариантов, встречающихся в популяциях мира, в среднем одну однонуклеотидную замену (SNP) на каждые восемь нуклеотидов. Описали в том числе мультиаллельные локусы (включающие множество вариантов – аллелей — одного гена). А также те, которые никогда ранее не были описаны из-за своей исключительной редкости – они встречаются у одного человека из 10 тысяч и реже.

Впервые удалось описать мутации, которые независимо возникали множество раз, так называемые рекуррентные мутации. Описаны также целые классы генов, которые отличаются постоянством – в них возникает меньше мутаций, чем должно было

быть, исходя из расчетов. Причина в том, что мутации этих генов имеют фатальные последствия для жизнеспособности организма. Описаны мутации, полностью подавляющие синтез белка с данного гена (protein-truncating variants), причем в 72% случаев науке еще не известна патология, к которой такая мутация приводит. Но эти мутации не всегда оказываются фатальными. Как показано на рисунке, каждый индивид в среднем имеет 85 таких мутаций в гетерозиготном состоянии (на одной хромосоме), когда они не проявляются, так как мутантный ген дублируется здоровой копией, и 35 – в гомозиготном состоянии (на обеих хромосомах), когда они проявляются, реально блокируя синтез белка.



Каждый индивид носит некоторое число мутаций, которые полностью выключают синтез белка с гена гена (protein-truncating variants): в гетерозиготном состоянии в среднем 85,1 мутацию (первый столбик), в гомозиготном состоянии в среднем 34,8 мутаций (второй столбик). Третий и четвертый столбики указывают на соответственное число редких мутаций.

b

Число аллелей, выключающих синтез белка (protein-truncating variants) в различных континентальных группах (показаны разными цветами), внизу указаны частоты мутаций.

Новая экзомная база данных значительно повышает возможность оценивать потенциальную патогенность каких-то генетических вариантов, ассоциировать те или иные мутации с развитием заболеваний. Особенно много дает база ExAC для понимания природы редких заболеваний. Теперь врачи-генетики смогут сравнить с базой ExAC мутацию, которая имеется у пациента, страдающего от той или иной редкой болезни. А также понять и механизм развития этой болезни – связь между геном и патологией.

текст Надежды Маркиной