

От гена к геному

Карта ДНК — не просто техническое достижение. Она позволила по-другому взглянуть на способы работы с генетической информацией. Тесты ДНК не новы: с 1970-х годов генетические анализы часто использовались для диагностики многих наследственных заболеваний и проведения судебных экспертиз, о чем знают поклонники сериала «CSI: Место преступления». Однако методики тестирования позволяли за один раз обработать только какой-то один ген или небольшую группу из нескольких генов. словно исследователи, чьи ресурсы ограничены, специалисты буквально наугад нащупывали цель, прежде чем приступить к анализу ДНК, который мог не дать никаких полезных результатов. Эти тесты были дорогими, долгими и технически сложными, а потому их использовали только в особых случаях: для диагностики сложных наследственных заболеваний или для исследования некоторых опухолей.

К началу XXI века благодаря развитию технологий и наличию карты-справочника у ученых появилась возможность читать и анализировать разом всю ДНК человека. В моду вошли два понятия: *геном* и *геномика*. В современной биологии часто можно встретить термины с суффиксом *-ом*, который означает «целостность». Таким образом, *геном* — это весь генетический материал организма, а *геномика* — наука, изучающая его. Можно сказать, что геномика меняет парадигму: если генетику интересуют отдельные гены, то геномика использует холистический подход, рассматривая все хромосомы в целом.

Инструменты, которые использует геномика, подобны дрону с рентгеновской камерой: когда ученым и врачам надо найти какие-то отклонения в ДНК, им не приходится обследовать каждую отдельную хромосому или ген: они могут быстро просканировать весь геном, чтобы получить общую

картину, а потом прицельно рассмотреть обнаруженную мутацию. Или они могут сравнить ДНК разных людей и найти различия, которые отвечают за индивидуальные особенности, предрасположенность к заболеваниям и вообще все то, что делает каждого из нас уникальным. Благодаря изучению генома целиком, а не отдельных генов, у исследователей появилась возможность наблюдать за неизвестными ранее генетическими механизмами, в работе которых одновременно задействованы сотни генов. В этой книге термины «генетический» и «геномный» будем для удобства использовать как синонимы, однако при этом подразумеваем, что все описываемые прикладные решения основаны на анализе всей ДНК целиком, то есть на геномном анализе.

Геномный подход более перспективный, он заменяет прежние методы изучения и диагностики генетических заболеваний. Кроме того, он создал новый рынок, ориентированный на здоровых людей, которым просто любопытно получить информацию о своей ДНК. Не умаляя достоинств карты, следует, однако, сказать, что она не отражает генетическую вариативность и не объясняет, почему все мы уникальны. Напротив, она максимально обезличена, так как построена на смеси ДНК анонимных доноров: ведь карта задумывалась как образец типичного генома человека — она описывает вид, а не особь. Чтобы сделать геномику персонализированной, надо переходить от *карты* к *картам*, используя уже имеющийся у нас общий справочник как инструмент, который поможет объяснить индивидуальную вариабельность.

Наука о различиях

Иногда я вижу сны, которые могут сниться только генетику. Один из самых частых сюжетов: я хорошо провожу время

с супермоделью, и неожиданно она говорит: «Дорогой, ты бесподобен! Скажи, сколько *моей* ДНК нужно поменять, чтобы стать, как *ты*?»

Как вы понимаете, я далеко не король красоты, однако этот вопрос вполне резонен, поскольку он подводит нас к одной из интереснейших загадок генетики: сколько *моей* ДНК делает меня отличным от супермодели, чемпиона или голливудской звезды? Проведите эксперимент: вспомните кого-нибудь, кто внешне сильно на вас не похож, и попробуйте отгадать, сколько букв в ваших геномах отличаются. Расшифровав и сравнив тысячи индивидуальных ДНК, генетики смогли ответить на этот вопрос. Вы удивитесь, но ДНК двух случайных людей будут совпадать на 99,5–99,9 %, то есть в среднем на тысячу совпадающих букв будет приходиться от одной до пяти различающихся. Если бы я был способен вернуться к самому началу своего эмбрионального развития и знал бы, какие буквы надо поменять, я мог бы сделать из себя генетическую копию Леонардо ДиКаприо, Идриса Эльбы, Скарлетт Йоханссон или любого другого человека, подредактировав небольшую часть своего генома.

Генетическое сходство — это общее правило, которое распространяется и на другие виды. У нас 98 % общих генов с шимпанзе, 85 % — с мышью и 50 % — с бананом. Это кажется невероятным, принимая во внимание разнообразие в лицах и цвете кожи, во внешности и прочих чертах, наблюдаемых у людей. Кто угодно отличит банан от мартышки, мартышку от человека, а меня от супермодели, но на уровне ДНК все выглядит совсем по-другому. Сравнить геномы людей — все равно что играть в игру «Найди 10 отличий»: две индивидуальные последовательности выглядят практически одинаково, и только если посмотреть поближе, можно обнаружить миллионы мелких вариаций — этикие генетические иголки в стоге сена.

Генетики самозабвенно ищут эти иголки, поскольку они-то и объясняют, в чем уникальность каждого человека. Цель персональной генетики — идентифицировать вариации и понять, как они влияют на черты, предрасположенности и здоровье.

Главный «кормилец» индивидуальной и коммерческой геномики — это SNP (Single Nucleotide Polymorphism, произносится как «снип»), или однонуклеотидный полиморфизм. Название сложное, но суть проста: SNP — это такое место в геноме, где буквы у индивидуумов различаются. Например, если в каком-то участке ДНК у меня стоит буква «С», а у Леонардо ДиКаприо — буква «А», это SNP (термин *вариант* — синоним SNP, я буду использовать эти понятия как взаимозаменяемые). Если сравнить ДНК большого количества людей, можно заметить, что одни *варианты* встречаются чаще других. Рассмотрим следующий пример (данные последовательности полностью выдуманы и не являются реальными):

Лео	AGAGCACCATTGCCATGCATTCTAC*				
Серджио	A*TAGCCCCATTGCCATGCATGCTAA				
Скарлетт	AGAGCACCATTGCCATGCATTCTAA				
Идрис	AGAGCACCATTGCCATGCATTCTAA				
Шарлиз	AGAGCCCCATTGCCAC*GCATGCTAA				
Снуп	AGAGCCCCATTGCCATGCATGCTAA				
*Редкий вариант	<table> <tbody> <tr> <td>Snip 1</td> <td>Snip 2</td> </tr> <tr> <td>(аллель А или С)</td> <td>(аллель Т или G)</td> </tr> </tbody> </table>	Snip 1	Snip 2	(аллель А или С)	(аллель Т или G)
Snip 1	Snip 2				
(аллель А или С)	(аллель Т или G)				

Как мы видим, одни участки ДНК различаются часто (SNP 1 и SNP 2), а другие — редко (они помечены звездочкой). Если провести анализ генома целиком и сравнить

данные тысяч разных людей (именно этому было посвящено большинство исследовательских проектов), можно обнаружить сходную картину: одни SNP встречаются чаще других. Благодаря результатам множества исследований нам стали известны участки ДНК, в которых прячутся персональные различия. В рамках таких международных проектов, как *НарМар* и «1000 геномов» (1000 Genomes), ученые сравнили ДНК тысяч людей разных национальностей и составили dbSNP — самый подробный на сегодняшний день каталог SNP, в котором перечислены все известные *варианты* с информацией о частоте встречаемости среди представителей различных народов. SNP составляют очень небольшой, однако исключительно информативный фрагмент ДНК человека, поскольку именно SNP являются самой изменчивой частью генома. Если продолжить аналогию с игрой «Найди 10 отличий», SNP можно рассмотреть как своего рода подсказку, невидимую ручку, которая обводит для нас букву ДНК, показывая, что индивидуальные различия кроются именно здесь⁶.

Есть миллионы более редких генетических *вариантов*, называемых *мутациями*. Разница между мутациями и SNP больше лингвистическая, чем какая-то еще, поскольку и мутации, и *варианты*, и SNP, по сути, означают одно и то же: буквы ДНК, что отличаются у разных индивидуумов. Разница только в частоте, с которой эти явления встречаются в популяции.

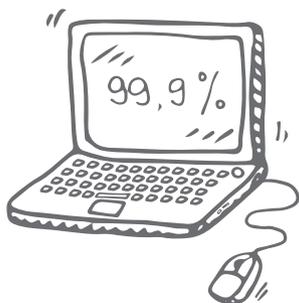
Где находятся SNP? Они могут быть в любом месте генома. Некоторые прячутся внутри генов или рядом с ними, другие — в некодируемых участках хромосом или даже в митохондриальной ДНК. SNP — это главный элемент всех современных исследований ДНК, в том числе в области коммерческой геномики, ведь они позволяют быстро просканировать геном, когда нужно найти индивидуальные генетические различия.

Результаты анализа SNP называют *генотипом*. Этот термин имеет более широкое значение, и, как правило, его применяют, когда говорят о любом наборе информации, полученной из ДНК человека, но в коммерческой геномике под генотипом обычно подразумевают *вариант* или *варианты*, которые нашли у вас. Например, мужчины с генотипом A;G в SNP rs6152 с высокой вероятностью облысеют с возрастом (A;G означает, что у человека в одной аллели буква A, а в другой в том же месте — буква G), в отличие от них обладателям аллелей A;A вряд ли грозит потеря шевелюры.

По мере того как росло число исследований, становилось ясно, что SNP — не единственные источники генетической изменчивости у людей. Кроме них есть *вставки* или *делеции* в последовательности ДНК — индивидуальные различия, вызванные удлинением (*вставка*) или укорачиванием (*делеция*) нитей ДНК до 1000 букв. *Вариация числа копий генов* (Copy Number Variation, CNV) — это тоже вставки или делеции, но они затрагивают последовательности ДНК подлиннее (1000 и более букв), которые могут содержать один или несколько генов⁷. Раньше считалось, что вставки/делеции и вариация числа копий не играют особой роли, однако сейчас ученые полагают, что эти явления вносят не меньший вклад в изменчивость, чем SNP.

Другим важным механизмом изменчивости являются так называемые *эпигенетические модификации* — ряд изменений на высоком уровне, воздействующих на структуру хромосом. Появляется все больше информации, которая подтверждает влияние описанных механизмов на нашу биологию. Мы рассмотрим эту тему подробнее в главе 17.





ГЛАВА 2 ПРИВЕТ, ДРУЖИЦЕ!

Знакомимся с людьми на Relative Finder

Дж. П. за сорок, у него есть семья. На фото в профиле — лицо человека со спокойной улыбкой. В описании сказано, что он живет в Нью-Йорке и работает в крупной медиа-компании. Мы не были с ним знакомы, жили по разные стороны Атлантики, но однажды Дж. П. прислал мне заявку в друзья на 23andMe, и я его добавил. Однако нас объединяли не фильмы, музыка, общая школа или прочие подобные вещи, на базе которых люди обычно знакомятся и общаются в «Фейсбуке», «Твиттере» или «Инстаграме». Нас объединяла ДНК. Сдав тест, я ожидал, что погружусь в тайны своего генома и, возможно, будущего своего здоровья; но вместо этого первое, с чем я столкнулся, стала сгенерированная на основе результатов моего тестирования социальная сеть контактов, которая познакомила меня с неизвестными мне доселе людьми со всего света.

Многое произошло с тех пор, как мой биоматериал покинул пределы родного дома и отправился в Калифорнию. Когда образец добрался до места назначения, мою ДНК извлекли, расшифровали и записали в файл. Дальше

приложение под названием Relative Finder («Поиск родственников») сравнило данные моего файла с данными других клиентов и, как следует из названия, нашло моих родственников.

Оказалось, что мы с Дж. П. — пятиуродные братья. Где-то в Европе (как мы позже выяснили, скорее всего, в Турине) жил наш общий прапрапрадедушка или прапрапрабабушка. Но у каждого человека есть тысячи дальних родственников, и профиль Дж. П. — всего лишь один из профилей растущей онлайн-армии людей, ссылки на которых я вижу на своей странице, — это как список предполагаемых знакомых владельца какого-нибудь популярного профиля в «Фейсбуке». Если предположить, что у каждого из наших прародителей в пятом поколении было по 2,5 ребенка (исследователи считают, что это средний показатель), которые, в свою очередь, тоже оставили потомство, то у каждого из нас наберется орава из 4700 пятиуродных братьев и сестер — ими можно заполнить небольшой концертный зал⁸.

Похоже, что число предлагаемых контактов на моей страничке 23andMe медленно приближается к этой цифре. Все начиналось с нескольких человек, но по мере того как в приложении появляются новые профили, количество моих родственников увеличивается прямо на глазах. Сейчас в моей порядком подросшей «семье» 1800 человек, у каждого есть профиль, и с любым из них я могу связаться, нажав на кнопку «Добавить в друзья». Это множество людей, с которыми я никогда бы не встретился — ни в реальности, ни онлайн, — если бы не система, сравнившая наши геномы. Оказывается, у меня есть четвероуродная сестра Элеонора, которая, если верить профилю (как положено, с фотографией), живет в США и имеет украинские и ирландские корни. Есть пятиуродная сестра Мэри из Филадельфии: ей 90 лет, и она приехала из того же самого сицилий-

ского города, в котором родилась моя мама. У меня есть двоюродная сестра Кэтлин из Канады, зачатая в 1981 году с помощью донора, — она активно разыскивает своего биологического отца. «Я всего лишь хочу познакомиться со своей семьей», — пишет Кэтлин в своем профиле; на фото она обнимает своего сына. В Италии живет мой троюродный брат Аурелио: ему 60 лет, и у нас с ним довольно длинный участок общей ДНК. Я мог бы десятилетиями колесить по миру, если бы решил познакомиться со всеми своими братьями и сестрами. А если хотя бы половина из них стала звать меня на семейные торжества, я бы потратил всю свою оставшуюся жизнь, посещая свадьбы, крещения, первые причастия, семейные ужины в День благодарения и церемонии вручения дипломов. Эта мысль вселяет в меня определенный ужас.

Как относиться ко всем этим псевдородственникам? Кто они мне?

Новая социальная сеть

ДНК-технологии ассоциируются у нас с белыми халатами: исследователями, поглощенными сложной научной работой, врачами, занимающимися предупреждением, изучением и лечением заболеваний, а также криминалистами, с помощью генетических улик раскрывающими убийства. Но Дж. П., Кэтлин, Аурелио и я не работаем ни в полиции, ни в больнице. Мы — обычные люди, которые листают список своих генетических контактов в смартфоне, сидя в вагоне метро или на диване, пока дети смотрят мультики, а на плите варятся макароны к ужину. Мы используем свои геномы как контент для социального взаимодействия в цифровом мире — наподобие фотографий, сообщений или видео, что мы отправляем друг другу в «Фейсбуке» или «Ютьюбе».

Появление геномных социальных сетей свидетельствует об изменении парадигмы того, как мы распоряжаемся своей биологической информацией. Раньше я воспринимал геномный отчет как набор данных, важных прежде всего для диагностико-прогностических медицинских тестов, — то есть нечто очень личное, что может являться предметом обсуждения только с врачом или с близкими родственниками. Теперь же я обнаруживаю себя участником сервиса, где меня просят поделиться информацией о своем геноме с тысячами незнакомых людей. Это полноценная и самая настоящая социальная сеть, построенная на ДНК. Энн Воджицки, одна из основателей 23andMe, сформулировала суть своего проекта в нескольких словах: «Я хочу расшифровать свой геном и на его основе создать сеть контактов»⁹.

Relative Finder никогда не ошибается. Файлы с ДНК клиентов сравниваются между собой, и пользователям сообщается, насколько близкое родство их связывает. Эта система точна с научной точки зрения, в отличие от наших знаний о своей семье, которые со временем стираются. В тот момент, когда программа находит первого родственника пользователя, работа науки заканчивается и начинается работа памяти. Чтобы убедиться, что мы действительно связаны кровными узами, моим двоюродным братьям и мне нужно сравнить фамилии наших дедов, места, где они родились, где жили, — одним словом, вспомнить все то, что мы сами знаем о собственных корнях, и таким образом восстановить ход событий, связывающих одно поколение с другим. После того как мы с Дж. П. добавили друг друга в друзья, каждому из нас стала доступна более подробная информация со страницы другого (совсем как в обычных соцсетях). В профиле Дж. П. — длинный список фамилий и мест рождения известных ему предков: он надеется, что эта информация совпадет с моей или хотя бы всколыхнет что-то в моей