

# Лучше меньше, да лучше

## Чем наш геном отличается от генома червя

**И**звестие о расшифровке человеческого генома было не первым. О том, что получен «черновик» нашего генома, ученые объявили еще в 2001 году. Но именно в 2003-м он был в основном расшифрован. Это касается той части ДНК, которая несет информацию о структуре белков. Кроме нее, в наших хромосомах есть еще некодирующая ДНК, и для чего она предназначена, пока до конца не ясно.

Вообще, расшифровка человеческого генома — это эпохальное событие, и по своему значению для человечества она сопоставима только с использованием атомной энергии и появлением компьютеров. Я считаю, что это три главных достижения науки за последние сто лет. Теперь биологи смогут точно узнать структуру наследственного материала.

После того как ученые, занимающиеся расшифровкой генома, получили последовательность нуклеотидов, первое, что они сделали, — оценили количество генов. Конечно, абсолютно точно подсчитать их было нельзя. По оценке Международного консорциума, включающего сотни исследователей из разных стран, в человеческом геноме приблизительно 32–35 тысяч генов. А еще лет пять тому назад специалисты полагали, что их 70–80 тысяч. Подсознательно нам хотелось, чтобы генов у человека было побольше, — нужно же возвыситься над червем *Caenorhabditis elegans*, у которого их 19 тысяч, и над дрозофилой, у которой их 13 тысяч. Но надо сказать, цифра 32–35 тысяч была получена не экспериментальным, а расчетным путем — методами биоинформатики.

Напомним, что человеческий геном содержит свыше трех миллиардов нуклеотидных пар. Человек, естественно, не может работать с такими числами — на то есть суперкомпьютеры. Как найти в этой последовательности нуклеотидов отдельные гены? Я не буду останавливаться на подробностях, но суть в следующем. Перед тем как браться за подсчеты, биоинформатики делают ряд допущений о типичной структуре гена и формулируют определенные критерии: какую последовательность нуклеотидов можно считать геном. Исходя из этого,

все расшифрованную последовательность нуклеотидов «прогоняют» через созданный алгоритм и получают искомое число. Для того чтобы эту цифру подтвердить, надо заставить работать все гены, то есть выделить кодируемые ими белки. Это задача на несколько десятилетий. Сейчас белковые продукты известны для многих генов, но далеко не для всех.

Что же произошло потом? Биоинформатики, заново изучив гены с известными функциями, пересмотрели критерии гена. И оказалось, что их можно уточнить, поэтому количество генов потребовалось пересчитать снова. Так возникла цифра в 20–25 тысяч генов, которая почти уравнила нас с крошечным червем. Я сразу хочу сказать, что эта цифра не окончательная и количество генов будет меняться еще не раз. Но к этому надо относиться спокойно. «Потеря» человеком приблизительно трети генов — сенсация только для обывателя, а специалисты, занимающиеся этим вопросом, не удивлены.

По количеству генов мы близки к червю, хотя у него всего тысяча клеток, а у нас  $10^{14}$ . Так как же мы в люди-то вышли? Дело в том, что главное — не количество, а качество. Самое важное, сколько информации гены содержат в себе и как они регулируются. Мы отличаемся от других организмов, стоящих ниже нас на эволюционной лестнице, не столько по количеству генов, сколько по их сложности. Весь секрет в том, что «наполненность» гена, его функциональная нагрузка у нас несравненно выше, чем у червя и дрозофилы. Большая часть генов человека устроена таким образом, что позволяет получить из одного гена не один, а много разных белков (см. об этом статью В.В.Велькова в предыдущем номере «Химии и жизни». — *Примеч. ред.*). У таких генов сложная, мозаичная структура. Участки, в которых записана информация о последовательности аминокислот в белках (экзоны), чередуются с участками, в которых такой информации нет, — с интронами. Последние в процессе сплайсинга («вырезания») выбрасываются из информационной РНК (матрицы белкового синтеза), и она становится гораздо короче.

Чтобы понять, что дает такая мозаичная структура, рассмотрим простейший пример. Представим себе очень простой ген, состоящий из трех кодирующих участков А, Б, В, между которыми находятся некодирующие участки. А теперь посмотрим, сколько белков можно получить из этого гена. Если использовать только А — получится короткий белок, который тем не менее будет выполнять какую-то функцию. Можно использовать А и Б, тогда белок окажется длиннее и, очевидно, его функция тоже будет иной. Можно использовать все три участка: А, Б и В и получить еще более длинный белок. Но можно также построить белок из А и В, из Б и В, наконец, только из участка Б или только из В. Если подсчитать все эти сочетания, то окажется, что такого простого гена достаточно, чтобы получить семь разных белков. Но многие гены состоят из десятков и изредка даже сотен таких участков — интронов и экзонов. И количество белков, которое можно из них получить по законам комбинаторики, становится необозримым. Выходит, что из сравнительно небольшого числа сложных генов получается практически неограниченное число белков.

Совсем недавно во всех учебниках по биологии был жирным шрифтом написан постулат: «Один ген — один белок». Но сегодня это правило в корне изменилось. Теперь надо говорить: «Один ген — много белков», а точнее: «Один ген — очень много белков». Это результат того, что произошло в геномике за последние 20 лет.

Так вот, у низших организмов таких сложных генов очень мало. А для человека на сегодняшний день доказано, что по крайней мере половина генов мозаичны. Очевидно, в какой-то момент эволюции матушка-природа решила, что хватит плодить новые гены, и пошла по другому пути — усложнения уже имеющихся, по пути наиболее рационального использования того, что уже создано.

Я еще раз повторяю, что и нынешняя цифра не окончательная. Количество генов может еще уменьшиться, а потом, возможно, увеличится — иными словами, это будет флуктуирующий процесс с постепенным приближением к истине.

*В 2004 году появилась информация, которая сильно озадачила общество. Оказалось, что у человека не 30–35 тысяч генов, как считали после расшифровки генома, а примерно на треть меньше, 20–25 тысяч. Что по этому поводу думают ученые? За разъяснениями наш корреспондент Н.Маркина обратилась к академику РАН Л.Л.Киселеву.*



## ПРОБЛЕМЫ И МЕТОДЫ НАУКИ

**— А не исчезнет ли в результате этого процесса само понятие «ген» как структурная и функциональная единица генома?**

— Об этом можно не беспокоиться. Ген куда не исчезнет, но понятие о нем как о чем-то простом постепенно меняется на другое. Ген не исчезнет так же, как не исчез атом. Мы теперь знаем, что атом устроен крайне сложно, однако от этого он не перестал существовать.

**— Представление о том, что в геноме есть «лишняя», некодирующая ДНК, появилось ведь довольно давно?**

— Да, но раньше считали, что некодирующая ДНК расположена между генами, что есть гены и есть межгенные участки. А сам ген был неприкасаем. Открытие в конце 70-х годов внутригенных некодирующих участков было полной неожиданностью. У бактерий, кстати, их нет вовсе, у них цепочка ДНК и цепочка белка полностью соответствуют друг другу. Однако нельзя сказать, что у высших организмов принцип соответствия структуры гена и структуры белка не работает. Он соблюдается, но не везде, а только на тех участках, которые кодируют последовательность аминокислот, то есть на экзонах.

**— Не следует ли теперь пересчитать все прочие расшифрованные геномы? Вероятно, и там были допущены такие же ошибки и количество генов в них тоже уменьшится?**

— Вопрос справедливый, но ответ на него неоднозначен. Все же геном человека устроен намного сложнее, чем у остальных живых существ. Хотя что касается наших ближайших родственников — человекообразных обезьян, то и у них, вероятно, при тщательном анализе количество генов уменьшится. Наверное, и мышка какое-то количество генов потеряет, но вот дрозофила и *C. elegans* — вряд ли. У простых организмов число генов, скорее всего, существенно пересматривать не придется. Для высокоорганизованных животных эта цифра будет «плавать», но чем организм проще, тем она точнее. Конечно, эту работу выполняют компьютеры, но то, чего человек пока не понимает, не может правильно оценить и программа. И только с развитием наших знаний мы будем приближаться к истине.

Ясно одно: эволюция идет не по пути бесконечного наращивания числа генов, а по пути выжимания из гена всего того, что можно из него выжать. Этот итог исследования генома уже никто не опровергнет.

**— Уточнение количества генов — результат работы того же Международного консорциума, который занимался расшифровкой человеческого генома?**

— Это не совсем так. Международный консорциум поставил перед собой конкретную цель и достиг ее. Теперь огромные фабрики по секвенированию переключились на другие геномы. Но специалисты, анализирующие геном, остались, хотя к ним, ес-

тественно, подключаются и новые люди. Я против того, чтобы называть какие-то имена, поскольку в работе такого рода не нужна персонификация. Это продукт усилий биоинформационного научного сообщества.

И все же хочется сказать, что в России есть две очень мощные школы биоинформатики. Одну, московскую, возглавляют Михаил Сергеевич Гельфанд и Андрей Александрович Миرون — талантливые исследователи с мировыми именами. Вторую, новосибирскую школу возглавляет Николай Александрович Колчанов — заместитель директора Института цитологии и генетики СО РАН. Есть и еще группы — в Пушчине, в МГУ, в московском Институте молекулярной биологии. Так что здесь говорить об отставании российской науки было бы в высшей степени несправедливо.

Я считаю, что наши биоинформатики слишком умны для того, чтобы заниматься подсчетом генов, они работают над более сложными проблемами функциональной геномики. Миرون и Гельфанд чрезвычайно много внесли в познание кодирующих и некодирующих участков генов. Их программа для анализа таких участков признана одной из лучших в мире. Поэтому без ложного патриотизма можно говорить, что Россия достойно представлена в мировом сообществе биоинформатиков — в отличие от секвенирования генома, где наш вклад был ничтожен из-за отсутствия должной финансовой поддержки.

