

ОНИ ДАЖЕ НЕ МОГУТ РАСШИФРОВАТЬ ГЕНОМ ЧЕЛОВЕКА!

Алексей
Алексенко,
кандидат
биологических
наук



ГЛЕБ ШЕЛКУНОВ

Сегодня нерасшифрованными остается примерно 5% человеческого генома, и основная часть этой terra incognita — пара десятков больших брешей, каждая размером с геном средней бактерии

В конце минувшего века автор этих строк гулял по хромосоме одной симпатичной плесени в попытках найти центромеру. «Прогулка по хромосоме» — это способ исследования генома, когда от изученных его частей постепенно перемещаются в неизученные. А «центромера» — участок хромосомы, за который клетка при делении тащит ее к одному из полюсов. Не будь центромер, хромосомы распределялись бы случайно и в конце концов потерялись бы. Этой информации должно хватить для понимания нашей истории.

В процессе той прогулки мы забрели в область густых повторов. Дальше идти было некуда. Работу пришлось забросить (то есть опубликовать как есть), и это разочарование стало одним из факторов, которые привели незадачливого исследователя из его лаборатории к нынешнему состоянию научного журналиста и автора этих строк.

Однако первая любовь остается навсегда, и я изредка заглядывал в прессу: как там у них дела с центромерой? В 2005 году объявлено о том, что геном аспергилла (именно он и был той плесенью) полностью расшифрован. Но знаете что? Центромеру в нем так и не нашли.

И тут уместно вспомнить о геноме человека, который, согласно торжественным анонсам, полностью расшифрован еще в 2003-м. Многим кажется, что слово «полностью» означает, что геном и впрямь был расшифрован полностью, но это не так. Центромер — как и кое-чего другого — вы не найдете там до сих пор. Для тех, кто любит читать мелкий шрифт в конце кредитных договоров, участники проекта поясняют: «человеческий геном расшифрован полностью, насколько позволяют существующие технологии». Что же мешает технологиям? То же, что изменило жизнь автора: повторы.

Обычно геном собирают из кусочков путем наложения соседних фрагментов. Возьмем два фрагмента текста: «Он не сумел приготовить свое лицо к тому положению, в ко...» и «...положению, в которое он становился перед женой после открытия его вины». Из них легко собирается целая фраза, а затем, если таких кусочков набрать побольше, компьютер может восстановить полный текст романа «Анна Каренина». Но что получится для большой литературы, не работает в случае треша: полный текст песни «Підманула, підвела» таким способом восстановить не получится. Виной тому повторяющийся припев: если длина известных фрагментов меньше длины этого припева, нет никакого способа узнать, какой куплет за каким следует.

В геномах живых существ — и плесени, и человека — таких повторов предостаточно. Один из примеров — фрагмент длиной в 171 букву, повторяющийся в геноме человека многие десятки раз подряд. В большинстве геномных проектов длина фрагмента секвенирования составляет чуть больше сотни букв, и прорваться через подобную преграду они не могут. Новая техника, вроде представленного в этом году прибора MinION, расшифровывает ДНК кусками до сотни тысяч букв. Однако длина самых больших брешей в геномах млекопитающих превышает этот предел. По оценкам профессора Джорджа Черча из Гарварда (того самого, который недавно сообщил об элегантном способе проследить за судьбой каждой клетки в организме мыши), сегодня нерасшифрованными остаются примерно 5% человеческого генома, и основная часть этой terra incognita — пара десятков больших брешей, каждая размером с геном средней бактерии.

В одном из эпизодов сериала «Теория большого взрыва» официантка Пенни ставит в тупик всезнающего физика Шелдона Купера вопросом о фронтмене группы Radiohead. Ученым порой свойственно даже гордиться незнанием того, что они считают маловажным. Но отсюда один шаг до следующей степени заносчивости: считать маловажным все, чего вы не знаете. И эта опасность всерьез угрожает геномике. Разумеется, гены, кодирующие белки, очень важны, и по счастью они легко расшифровываются. Однако не надо думать, что о бессмысленных повторах в геноме нам знать ничего не нужно. Те же самые центромеры — ключ к анеуплоидиям и другим генетическим расстройствам. Известны случаи, когда тяжелый наследственный дефект развития отличается от здоровья только числом повторяющихся элементов, отделяющих ген от «энхансера» — его выключателя. Да и вообще, знать что-то всегда лучше, чем не знать, даже если речь идет о таком маловажном факте, как то, что вокалиста Radiohead зовут Том Йорк.

Когда они говорят вам, что расшифровали геном человека, они не то чтобы «лгут» — скорее, слегка упрощают ситуацию в угоду нашему невежеству. Но тут есть обратная сторона: уже отпраздновав великую победу разума, мы, широкая публика, просто не поймем, зачем этим ученым нужны еще какие-то деньги. А когда геном все же будет прочитан полностью, от первой до последней буквы — или, как выражаются враги языка, от слова «совсем», — мы не поймем, в честь чего вдруг еще один банкет. А ведь повод для банкета будет весьма достойный.