



# ВСЕ ДАННЫЕ МИРА В ОДНОМ ЯЙЦЕ

Как можно использовать  
молекулу ДНК для  
хранения и генерирования  
гигантских объемов  
информации

*Джеймс Далман*

## ОСНОВНЫЕ ПОЛОЖЕНИЯ

- ДНК обладает множеством свойств, благодаря которым она идеально подходит для хранения не только генетического кода, но и любой другой информации. В настоящее время ДНК пока неспособна заменить обычные электронные накопители, например жесткие диски.
- Однако по мере совершенствования методов секвенирования специалисты в таких областях, как химическая инженерия, начинают использовать ДНК в качестве молекулярного записывающего устройства, которое позволяет генерировать данные с невиданной ранее скоростью.
- Таким образом, ДНК используется не только для «чтения», но и для «записи» информации. Это поразительное научное достижение способно в значительной степени ускорить разработку лекарств и создание новых методов терапии.

## ОБ АВТОРЕ

**Джеймс Далман** (James E. Dahlman) — научный сотрудник Отделения биомедицинской инженерии им. Уоллеса Коултера в Технологическом институте Джорджии и Университета Эмори. Его лаборатория ведет исследования в следующих областях: транспорт лекарственных веществ в ткани организма, нанотехнологии, геномика и редактирование генома.



несколько миллиардов лет тому назад, задолго до того момента, как человек создал жесткий диск, на гребне эволюции возникла молекула ДНК, способная хранить самую, пожалуй, ценную информацию — генетический код. ДНК настолько хорошо справилась с этой задачей, что теперь она присутствует во всех известных живых организмах, населяющих Землю. Благодаря современным технологическим достижениям, которые позволяют нам легко «читать» ДНК и даже «записывать» в нее информацию, человечество научилось использовать эту древнюю молекулу для хранения вообще любой информации, которую в эпоху больших данных человечество с нарастающей скоростью создает и накапливает.

В настоящее время специалисты широко обсуждают вопрос о репрофилировании ДНК. Они хотят, чтобы эта молекула уместила в себя всю накопленную человечеством информацию точно так же, как она хранит в себе генетический код. В конце концов, современные запоминающие устройства, на которых информация представлена, как известно, в виде двоичного кода, уже работают на пределе возможностей. Вот, например, недавно возникла следующая проблема, затронувшая вопрос о безопасном хранении данных: социальная сеть *Myspace* (некогда самая популярная) объявила о том, что все персональные данные пользователей, накопленные за десять лет, могут быть безвозвратно потеряны в ходе переноса серверов. Вся уязвимость и отсталость существующих технологий хранения информации отчетливо проявляется при решении задачи долгосрочной защиты данных (например, в ходе перезагрузки веб-сайта после перерыва в работе). Таким образом, перед нами возникает вопрос: где хранить всю нашу информацию? Кроме того, хранение данных подразумевает еще и значительные энергозатраты.

Преодолеть все эти трудности нам помогут свойства ДНК. Во-первых, структура двойной спирали

этой молекулы идеально подходит для хранения информации, поскольку (по правилу комплементарности) если мы знаем последовательность нуклеотидов одной нити ДНК, то автоматически можем определить последовательность другой нити. Кроме того, ДНК способна сохранять свою стабильность на протяжении очень-очень долгого времени, а это означает гарантию целостности и точности представления информации, заключенной в этой молекуле. Например, в 2017 г. ученые подробно исследовали ДНК, выделенную из человеческих останков, возраст которых составил 8,1 тыс. лет. Заметим, что все это время они хранились отнюдь не в идеальных условиях. Таким образом, при хранении в холодной и сухой среде молекула ДНК с большой долей вероятности способна сохранять свои свойства на протяжении десятков тысяч лет. Итак, ДНК способна сохранять свою стабильность на протяжении весьма долгого времени, а это означает гарантию целостности и точности представления информации, заключенной в этой молекуле.

Но, возможно, самое поразительное свойство двойной спирали заключается в том, что она умеет компактно складываться, образуя тем самым

очень плотную структуру. Так, каждая отдельная клетка человека содержит ядро диаметром примерно 0,00001 м. Если бы ДНК, извлеченную из ядра, вытянули в длину, то она бы достигла 2 м. Кроме того, если все молекулы ДНК человека связать вместе и вытянуть в одну линию, то ее общая длина достигла 100 трлн м. В 2014 г. ученые подсчитали, что в одном грамме ДНК можно хранить 455 эксабайт информации. Такая плотность хранения данных примерно в миллион раз выше, чем у жестких дисков.

Итак, молекула ДНК — это, по сути, потенциальное средство хранения информации. Однако прежде чем она заменит традиционные жесткие диски, нам предстоит решить множество не только научных, но и экономических задач, к тому же учесть и этические вопросы. Между тем благодаря своим достоинствам ДНК быстро становится одной из самых востребованных форм информационных технологий. Например, ученые уже смогли приспособить эту молекулу в качестве носителя, на который были записаны старые голливудские фильмы, — они теперь хранятся не на хрупкой микропленке, а прямо в генетическом коде. А совсем недавно ДНК стали использовать в качестве инструмента для разработки более безопасных методов генной терапии и противоопухолевых препаратов, а также для создания, возможно, первого в истории человечества средства, позволяющего вести непрерывный скрининг состояния живого организма. В этой передовой области технологий возникают и другие идеи: скажем, молекуле ДНК уготована роль не только долгосрочного хранилища информации — ее можно использовать еще и для формирования данных. И она способна делать это с невиданной ранее скоростью. И все потому, что в отличие от других молекул ДНК более эластична, что позволяет нам в значительной мере наращивать объемы данных, сократив при этом затраты на их хранение.

### Увеличиваем количество наночастиц

В последние годы ученые все чаще используют ДНК в качестве молекулярного регистратора, предназначенного для наблюдения за экспериментом и анализа его результатов. При этом во многих случаях специалисты стали применять так называемое ДНК-штрихкодирование: чтобы отследить результаты отдельного эксперимента и их проанализировать, ученые используют заранее известную последовательность ДНК в качестве молекулярного маркера. Например, с результатом эксперимента номер один связана нуклеотидная последовательность АСТАТС, с результатом эксперимента номер два — ТСТГАТ и т.д.

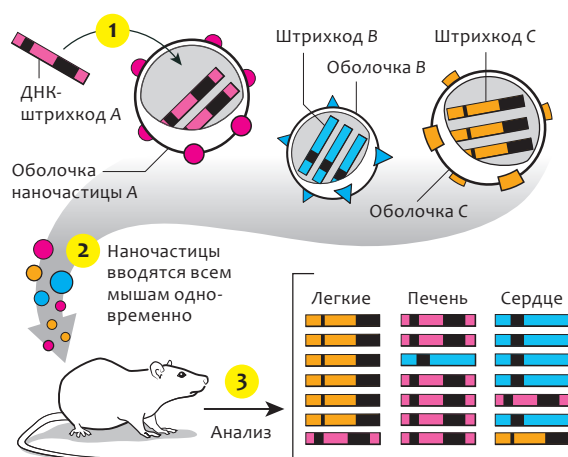
ДНК-штрихкодирование появилось в начале 1990-х гг. Его предложили в качестве способа наблюдения за химическими реакциями Ричард

Лернер (Richard Lerner) и ныне покойный Сидни Бреннер (Sydney Brenner), работавшие тогда в Научно-исследовательском институте Скриппса. Идея была новаторской, но она опередила свое время: ведь на тот момент еще не были разработаны технологии, позволяющие легко и без значительных финансовых затрат считать ДНК. Потенциал ДНК-штрихкодирования реализовали только после того, как было сделано множество открытий в таких областях, как химия нуклеотидов и микрогидродинамика, — все вместе они привели к созданию передовых методов секвенирования. Крупный прорыв произошел в 2005 г., когда в рамках одного из экспериментов, длившегося четыре часа, ученым удалось прочесть 25 млн пар нуклеотидов ДНК.

Передовые методы секвенирования быстро совершенствуются; теперь нам вполне по силам с легкостью прочитать одновременно миллионы последовательностей нуклеотидов ДНК, а это значит, что мы можем провести одновременно

## Отслеживание наночастиц по ДНК-штрихкоду

ДНК-штрихкодирование позволяет ученым эффективно тестировать наночастицы, предназначенные для доставки лекарственных препаратов к различным тканям организма. Раньше данный процесс был трудоемким и занимал много времени. Однако теперь появилась возможность тестировать одновременно сотни частиц различного типа. Во время тестирования уникальный ДНК-штрихкод помещают под оболочки наночастиц разных видов **1**. Эти наночастицы предназначены для доставки лекарственных средств напрямую в поврежденные клетки. В процессе экспериментальной проверки ученые вводят одновременно большое количество наночастиц **2**. Затем проводится сканирование клеток на предмет выявления типа наночастиц, попадающих в ткани организма **3**. Эта процедура позволяет быстро определить оптимальную конструкцию наночастицы, способной с заданной точностью доставлять препараты в пораженные ткани и при этом минимизировать побочные эффекты.



SOURCE: "A DIRECT COMPARISON OF IN VITRO AND IN VIVO NUCLEIC ACID DELIVERY MEDIATED BY HUNDREDS OF NANOPARTICLES REVEALS A WEAK CORRELATION," BY VALERIA PALOMONZAK ET AL., IN NANO LETTERS, VOL. 13, NO. 3, MARCH 14, 2014; ILLUSTRATION BY JEN CHRISTENSEN

несколько тысяч экспериментов и проанализировать их результаты. Анализ экспериментов с применением ДНК-штрихкодирования и современных методов секвенирования — это одна из форм управления данными: вместо того чтобы проверять по одной гипотезе за раз, ученые теперь могут составлять, скажем, по 20 тыс. прогнозов и проверять их все сразу с целью выявления правильных.

Первыми широко использовать ДНК-штрихкодирование стали биологи. По мере роста доступности этой технологии ее стали все чаще применять для проведения экспериментов и в других научных областях, включая химическую технологию и материаловедение. Например, в моей лаборатории в Технологическом институте Джорджии инженеры используют ДНК-штрихкодирование для совершенствования конструкции и функций наночастиц, предназначенных для безопасной доставки препаратов к поврежденным клеткам. Кто-то, наверное, скажет, что описанная нами нанотехнология, опирающаяся в основном на инструментарий физики и химической технологии, не имеет никакого отношения к ДНК. Но, как бы то ни было, именно с помощью ДНК удалось анализировать и хранить данные любого вида, вот почему эту молекулу выгодно использовать как инструмент организации данных.

Специалисты-нанотехнологи зачастую сталкиваются со следующей непростой задачей: эксперимент, призванный найти эффективные методы лечения, куда проще планировать, чем проводить и анализировать его результаты. Это происходит по той причине, что характеристики каждой из наночастиц — форма, размер, заряд, химический состав — обуславливают способность наночастицы доставлять генетические препараты в поврежденные клетки. Кроме того, перечисленные факторы друг с другом еще и взаимодействуют, из-за чего трудно предсказать, какая из наночастиц точнее всех доставит препарат прямо в цель. Значит, приходится отслеживать каждую наночастицу в отдельности! И, как нас уверяют разработчики наночастиц, создаваемых для препаратов на основе РНК (а это крупные фармацевтические компании), для проведения столь утомительного тестирования необходимы гигантские инвестиции в сотни миллионов долларов.

Вот тут-то молекула ДНК и может проявить себя во всей своей красе — в смысле ее способности хранить информацию. Теперь мы получили



*Благодаря двойной спирали ДНК превращается в идеальный инструмент, с помощью которого можно хранить информацию. Но ДНК еще не способна прийти на смену традиционным жестким дискам.*

возможность увеличить количество наночастиц, подлежащих тестированию; для этого надо создать несколько тысяч таких наночастиц, наделив их различными химическими свойствами (скажем, сделав их в виде больших положительно заряженных сфер или маленьких треугольников, обладающих нейтральным зарядом) и снабдив каждую такую наночастицу своим уникальным ДНК-штрихкодом.

В итоге получаем, например, следующую картину: наночастица номер один, обладающая химической структурой номер один, несет ДНК-штрихкод номер один; наночастица номер два, обладающая химической структурой под номером два, снабжена ДНК-штрихкодом номер два. Процесс штрихкодирования мы повторяем многократно, создавая тем самым множество различных наночастиц, у каждой из которых свой уникальный молекулярный ДНК-маркер. Затем мы вводим сотни таких наночастиц в поврежденные клетки; и, наконец, для того чтобы определить ту самую наночастицу, которая точнее всех доставила препарат в цель, мы используем секвенирование ДНК для выявления штрихкодов внутри клеток.

Размах, с которым проводятся подобные эксперименты, — явление для наномедицины совершенно новое. Обычно в моей области исследований при проведении экспериментов традиционными методами можно было довольствоваться только наблюдением за пятью наночастицами. К концу 2019 г. моя лаборатория уже надеется выяснить,



как ведут себя 500 различных наночастиц, как они осуществляют генную терапию в отношении 40 различных видов клеток, а это эквивалентно проведению одновременно 20 тыс. экспериментов.

В результате мы должны были разработать аналитические способы контроля качества данных; кроме того, при проверке полученных результатов статистическими методами нам потребовалась помощь. В первую очередь нам удалось определить, насколько качественно результаты одного из экспериментов предсказывают доставку препаратов к поврежденным тканям в другом эксперименте. Убедившись в надежности больших наборов данных, мы перешли к статистическим методам; наша цель заключалась в том, чтобы понять, как влияют некоторые признаки наночастиц (скажем, размеры) на доставку к тканям-мишеням. Мы установили, что качество доставки наночастиц обусловлено не их размером, а химическим составом. С помощью ДНК-штрихкодирования мы надеемся в ближайшем будущем предложить безопасную и менее затратную генную терапию. Одна из задач состоит в том, чтобы создать такие наночастицы, которые бы обеспечили специфическую генную терапию и при этом могли бы эффективнее бороться с раковой опухолью; одновременно с этим мы хотим снизить побочные эффекты, такие как тошнота и выпадение волос, сопутствующие традиционным методам лечения онкологических заболеваний.

И здесь нам удалось достичь некоторых успехов. В 2018 г., опираясь на достаточно большой массив данных, полученных в ходе экспериментов с применением ДНК-штрихкодирования, мы быстро выявили новые наночастицы, обеспечивающие генную терапию эндотелиальных клеток, которыми выстланы кровеносные сосуды, а также несколько видов иммунных клеток, управляющих иммунным ответом человеческого организма. Это открытие поможет изменить подходы к методам лечения — теперь появляется возможность регулирования активности белков в иммунных клетках, которые в настоящее время не поддаются медикаментозному лечению (заметим, что белки малочувствительны к воздействию низкомолекулярных лекарственных веществ или антител). К нашим разработкам проявили интерес специалисты в области генной терапии, о чем свидетельствуют работы, опубликованные в 2018 и 2019 гг. в таких научных изданиях, как *Proceedings of the National Academy of Sciences USA*, *Advanced Materials* и *Journal of the American Chemical Society*. Кроме того, мы создали компанию GuideRx, специализирующуюся на ДНК-штрихкодировании; главная задача компании — разработка методов безопасной генной терапии.

К настоящему моменту ДНК-штрихкодирование стало явлением настолько распространенным, что появились различные его модификации даже

в пределах одной области исследований. Один из примеров — онкобиология; ее задача — выяснить, каким образом генетические мутации вызывают рак, и создать новые фармакологические средства против онкологических заболеваний. Попутно заметим, что в настоящее время серьезную проблему представляет собой устойчивость к лекарственным препаратам: поначалу организм пациента зачастую хорошо реагирует на лекарство, но потом положительное воздействие лекарственных средств снижается, поскольку они теряют способность уничтожать раковые клетки.

Чтобы разобраться в причинах подобной резистентности к лекарственным средствам, ученые из лаборатории Тодда Голуба (Todd Golub) при Гарвардском университете воспользовались ДНК-штрихкодированием. В 2016 г. они с помощью вируса вставили ДНК-штрихкод непосредственно в геном раковых клеток. При этом раковой клетке вида *A* соответствовала штрихкод-последовательность *A*; раковой клетке вида *B* — штрихкод *B* и т.д. Ученые смешали разные виды клеток, поместили их в чашку Петри и обработали противоопухолевым препаратом.

Если препарат разрушает раковую клетку или замедляет ее рост, то клетка перестает делиться. И наоборот, если клетка приобретает устойчивость к лекарству (резистентность), то она быстро делится. Следовательно, со временем относительное количество штрихкодовых меток вида *A* увеличивается, если клетки вида *A* становятся невосприимчивы к лекарству. И, наоборот, число последовательностей вида *A* уменьшается, если лекарство уничтожает эти клетки. После секвенирования всех штрихкодов, взятых из выживших клеток, лаборатория количественно оценила, насколько эффективно все виды клеток одновременно реагировали на препарат.

В конце того же года лаборатория Монте Уинслоу (Monte Winslow) при Стэнфордском университете использовала клеточные линии поджелудочной железы, помеченные ДНК-штрихкодом, для идентификации лекарств, способных предотвращать распространение раковых клеток (метастазирование). С помощью вируса ученые внедрились штрихкод во все клеточные линии, а затем поместили каждую клеточную линию в предназначенную для нее лунку. Затем все лунки были обработаны противораковым препаратом. Таким образом, препарат под номером один стал ассоциироваться со штрихкодом номер один. Сразу после этого ученые вводили клетки в кровь, а затем определяли, какие из них смогли проникнуть в легкие. Затем, идентифицируя штрихкоды (выявив максимальное их количество или их отсутствие), ученые смогли определить те препараты, которые способствовали метастазированию или же, наоборот, его предотвращали.

В третьем эксперименте ученые из Института Броуда при Массачусетском технологическом институте и Гарвардском университете воспользовались ДНК-штрихкодированием, чтобы выяснить, каким образом гены данного конкретного генома реагируют на определенный вид рака. Сначала ученые вырастили очень большую колонию клеток, которую затем поместили в большую чашку Петри. После этого с помощью генного редактирования все гены данного генома один за другим включали или, наоборот, выключали. Последовательность нуклеотидов данного гена, экспрессия которого была изменена, служила в качестве штрихкода. Затем клетки обрабатывали противораковым препаратом и секвенировали ДНК. Это позволило ученым разобраться в том, каким образом каждый ген данного генома влияет на устойчивость к лекарственным средствам.

При использовании этих методов ДНК действует одновременно и как молекула, генерирующая данные (поскольку эксперименты требуется проводить все одновременно), и как молекула, способная эти данные хранить (поскольку передовые методы секвенирования используются для анализа ДНК-штрихкодов). Выводы обнадеживают: одни и те же методы могут использоваться по отношению к аутоиммунным, неврологическим и сердечно-сосудистым заболеваниям. ДНК-штрихкодирование оказалось очень мощным инструментом; чтобы в этом убедиться, поставьте вместо слова «рак» название любого другого заболевания, а вместо «резистентность» — любую другую реакцию на лекарственное средство, которую ожидает получить исследователь. Таким образом, метод ДНК-штрихкодирования позволяет уже на ранних стадиях значительно упростить разработку лекарств, тем самым ускоряя поиск эффективных методов лечения.

#### «Читать» vs «писать»

В основе ДНК-штрихкодирования лежат методы «чтения» определенных последовательностей ДНК. Однако до недавнего времени мы практически не умели генерировать эти последовательности, то есть «записывать» их. Вообще говоря,

«записывать» ДНК, по моему мнению, означает целенаправленно преобразовывать один вид информации (например, рисунки, фильмы или биологические состояния) в последовательности, которые можно накапливать, чтобы затем их считывать. В основе многих из таких новых технологий записывания лежат методы редактирования генов, которые, в свою очередь, базируются на *CRISPR* (от англ. *clustered regularly interspaced short palindromic repeats* — «короткие палиндромные повторы, регулярно расположенные группами»). Теперь с помощью *CRISPR*, спроектированных определенным образом, ученые научились синтезировать последовательности ДНК.

Некоторые современные достижения науки появились лишь после того, как ученые воспользовались *CRISPR* — механизмом, который в процессе эволюции появился у бактерий и предназначался для защиты от вирусных атак. Остановимся на нем подробнее: когда вирус атакует бактерию, он сначала прикрепляется к ее поверхности, а затем внедряет в клетку бактерии вирусную молекулу ДНК или РНК. В свою очередь, бактерии, чтобы «запомнить» вирус и тем самым предотвратить последующие вирусные атаки, создали механизм *CRISPR*, назначение которого — опознавать вирусную ДНК или РНК, а затем вставлять небольшие фрагменты вирусной ДНК в свой бактериальный геном. Другими словами, бактерия, чтобы себя обезопасить, как бы «записывает» в досье характеристики каждого вируса, который на нее напал.

И вот Сет Шипман (Seth Shipman) из Калифорнийского университета в Сан-Франциско (ранее он работал в лаборатории гарвардского генетика Джорджа Черча) воспользовался методом *CRISPR* и записал изображение человеческой руки прямо в геноме кишечной палочки (*Escherichia coli*). Для этого группа под руководством Шипмана впервые выделила два белка: *Cas1* и *Cas2*. Каждый из них способен захватывать нуклеотид ДНК и вставлять его в геном. Затем исследователи «вмонтировали» в ДНК кишечной палочки последовательности нуклеотидов, кодирующих пиксели, из которых (при секвенировании) и было создано изображение руки. Для этой цели ученым

**В отличие от обычных накопительных устройств ДНК может сохраняться десятки тысяч лет, если, конечно, хранить ее в прохладных и сухих условиях. Кроме того, молекула ДНК выдерживает условия экстремальной жары, которую обычные электронные устройства выдержать не могут**

пришлось сопоставить определенному массиву информации некоторые части молекулы ДНК. Скажем, нуклеотидные звенья А, С, G и T соответствуют определенным цветовым пикселям, в то время как штрихкод-последовательность ДНК отвечает за пространственное расположение данного пикселя в целом изображении.

После этого исследователи провели обратную операцию: после секвенирования ДНК из кишечной палочки (*E. coli*) ученые воспроизвели исходное изображение с точностью, превышавшей 90%. После этого они повторили эксперимент, производя существенное изменение его условий: они добавляли ДНК в разные моменты времени, а затем с помощью некоторого метода выявляли местоположение записанных последовательностей ДНК друг относительно друга. Затем ученые определили, когда именно (то есть в какой момент времени) данный фрагмент был добавлен в геном *E. coli*, и после этого им удалось создать серию изображений, кодирующих фильм. Ученые записали файл в графическом формате (*GIF*), на котором был запечатлен отрывок из знаменитой серии фотографий Эдварда Мейбриджа, созданных в 1878 г., «Лошадь в движении». В одной из работ, опубликованных в 2017 г., было заявлено о том, что с помощью секвенирования бактериального генома удалось воссоздать знаменитые фото Мейбриджа.

А совсем недавно ученые из лаборатории Рэндалла Платта (Randall Platt) при Швейцарской высшей технической школе Цюриха сделали одно очень важное открытие, которое развивает описанные выше подходы: авторы нацелились на мРНК, которая приходится молекуле ДНК, так сказать, одной из двоюродных молекулярных сестер и имеет для нее большое значение. Вместо записи изображений, кодируемых искусственно созданными последовательностями ДНК, ученые воспользовались механизмом *CRISPR*, порожденным различными видами бактерий, и стали напрямую записывать экспрессию генов естественной мРНК бактерий. Множество всех различных мРНК, генерируемых в клетке, отвечает за выработку белков и, следовательно, предопределяет все клеточные функции.

Чтобы записать мРНК, продуцируемую клеткой в разные моменты времени, ученые из лаборатории Платта сначала отсортировали белки типа *CRISPR-Cas*, полученные из множества различных бактериальных штаммов. Этот процесс позволил выявить белки, способные синтезировать цепь ДНК на обычной мРНК, трансформируя ее в геном. Ученые обнаружили, что эту операцию могут прodelывать белки *Cas1* и *Cas2*, находящиеся в бактерии *Fusicatenibacter saccharivorans*. В 2018 г. в ходе серии блестяще поставленных экспериментов с использованием специализированных вирусов исследовательская группа из лаборатории Платта продемонстрировала, что клетки способны хорошо

запоминать такие внешние факторы, как воздействие окислительного стресса, кислотной среды или даже гербицида.

Полученные результаты впечатляют, поскольку они продемонстрировали следующее: гены, полученные в ходе естественной экспрессии клетки в некоторый момент времени, можно записывать прямо в геном для последующего их анализа. Лаборатория Платта продолжает совершенствовать описанную выше технологию, поэтому увеличивается вероятность того, что процесс записи информации на клеточном уровне станет обычным явлением. И, может быть, данная технология наконец даст ученым возможность понять, каким образом клетка становится раковой, как она реагирует на инфекцию не только на небольших временных промежутках, но также с увеличением возраста человека.

### ДНК в роли жесткого диска?

Молекулу ДНК используют теперь в самых разных областях (и число их возрастает) для генерирования, записи и хранения информации. Возникает логичный вопрос: сможет ли ДНК успешно конкурировать с традиционными устройствами хранения данных и, в конечном итоге, обеспечит ли эта молекула сохранность всей информации, накапливаемой человечеством? К сожалению, придется дать отрицательный ответ: на сегодня жесткие диски и флеш-накопители гораздо лучше приспособлены к хранению информации, чем даже самые передовые системы, создаваемые на основе ДНК.

Но у традиционных электронных устройств, как и у любого технического средства, имеются свои ограничения. Например, для них необходимо выделять физическое пространство, они требуют особой внешней среды. Кроме того, даже самые долговечные из них вряд ли проработают более нескольких десятилетий. Таким образом, принимая все это во внимание, мы вскоре обнаружим, что всю информацию, накопленную человечеством к настоящему моменту, традиционным устройствам сохранить будет крайне трудно.

В отличие от обычных накопительных устройств ДНК может сохраняться десятки тысяч лет, если, конечно, хранить ее в прохладных и сухих условиях. В лабораторных условиях, где можно создавать экстремальные температуры, ДНК обычно хранят при температурах  $-20$  и даже  $-80^{\circ}\text{C}$ . Кроме того, молекула ДНК выдерживает условия экстремальной жары, которую обычные электронные устройства выдержать не могут. В 2015 г. ученые из Швейцарской высшей технической школы Цюриха Роберт Грасс (Robert Grass) и Венделин Старк (Wendelin Stark) продемонстрировали, что ДНК, заключенная в оболочку из двуокиси кремния, способна выдерживать температуру до  $+70^{\circ}\text{C}$  в течение недели и при этом в молекуле не происходит



никаких изменений. К настоящему времени уже созданы жесткие диски, способные вмещать до одного терабита на квадратный дюйм, но, согласно некоторым оценкам, оказалось, что всю информацию, порожденную человечеством к настоящему моменту, в принципе можно заключить в ДНК весом 1 кг.

Для того чтобы хранение информации в молекуле ДНК стало обычным явлением, необходимо решить сложнейшие технические задачи. Главное препятствие заключается в том, что хранение информации и ее извлечение — не одно и то же. Скажем, считывание данных с привычного жесткого диска осуществляется практически мгновенно, в то время как их извлечение из ДНК требует секвенирования — а на эту операцию в настоящее время уходит от нескольких минут до целого дня. Несмотря на тот факт, что за последние несколько лет мощность секвенаторов увеличилась, они по сравнению с привычным жестким диском выглядят громоздкими и дорогими.

И это еще не все проблемы, которые необходимо решить, прежде чем методы хранения информации с помощью ДНК действительно покажут всю свою мощь. Кроме того, наше общество обязано признать, что из-за доступности секвенирования ДНК отследить любого человека будет намного проще; к тому же возникнут новые угрозы для безопасности данных. И это не считая того факта, что в США и в других странах мира уже и так имеются проблемы с утечкой информации.

Секвенирование ДНК уже используется в отделениях полиции США, причем какого-то мощного контроля над этой процедурой не заметно. Стражи закона заставляют арестованных (даже тех, кто совершил незначительное преступление) сдавать тест на ДНК, после чего вся эта информация перетекает в большие банки генетических данных. Кое-кто утверждает, что сдачу персональной ДНК можно приравнять к своего рода дактилоскопии XXI в. Однако, согласитесь, разница между отпечатками пальцев и личной ДНК слишком велика. Отпечатки пальцев идентифицируют лишь одного человека; другое дело — ДНК: если ее предоставит кто-то из ваших родственников, то тем самым он распространит информацию, которая имеет отношение к любому члену вашей семьи, включая вас. В Китае под видом внедрения программы здравоохранения были собраны генетические данные почти у 36 млн человек. В эту группу вошли и многие уйгуры — представители мусульманской этнической группы, которая подвергается дискриминации. При этом никто не знает, как именно правительство распорядится полученными данными.

В настоящее время опасения насчет хранения информации в молекуле ДНК связаны с доступностью генетического кода человека; в результате дискуссия стала вращаться вокруг вопроса

о защите прав личности. Однако если другие виды цифровых данных (например, информация из медицинских учреждений, информация о заключенных контрактах, персональные цифровые архивы) действительно будут храниться в ДНК, то возникнет еще больше вопросов по поводу уязвимости информации, хранимой в этой молекуле, — причем с точки зрения как физической защиты устройства, так и кибербезопасности. А поскольку в столь крошечном пространстве, как ДНК, можно хранить гигантские объемы информации, то возникает вопрос: каким образом нужно распределить данные, чтобы избежать их высокой концентрации в одном-единственном месте? И даже если мы упростим процесс извлечения информации из ДНК, все равно всплывет другой вопрос: как с технической точки зрения осуществлять регулярное обращение к данным, заключенным в молекуле ДНК, и одновременно не допустить их взлома или случайной потери?

Нам еще только предстоит найти ответы на довольно трудные вопросы, причем не только научные, но и этические. А это, увы, энтузиазма не прибавляет. Вот тогда я мысленно обращаюсь к примеру братьев Райт — о них я вспоминаю потому, что сам вырос в том же городе, расположенном в штате Огайо, где и они. Как известно, первый полет братьев Райт длился всего 12 с, а его дальность составила 37 м. Но уже через 66 лет человек без всякой помощи компьютеров высадился на Луну. Вспоминая об этих великих свершениях, я проникаюсь оптимизмом и верю, что мы тоже через несколько десятилетий поставим себе на службу силы, таящиеся в ДНК. Не спорю, эта молекула может создать нам новые проблемы. Но пользы от нее будет гораздо больше. ■

*Перевод: И.В. Ногаев*

#### ДОПОЛНИТЕЛЬНЫЕ ИСТОЧНИКИ

- Next-Generation Digital Information Storage in DNA. George M. Church et al. in *Science*, Vol. 337, page 1628; September 28, 2012.
- High-Throughput In Vivo Screen of Functional mRNA Delivery Identifies Nanoparticles for Endothelial Cell Gene Editing. Cory D. Sago et al. in *Proceedings of the National Academy of Sciences USA*, Vol. 115, No. 42, pages E9944-E9952; October 16, 2018.
- Transcriptional Recording by CRISPR Spacer Acquisition from RNA. Florian Schmidt et al. in *Nature*, Vol. 562, pages 380–385; October 18, 2018.
- Tech Turns to Biology as Data Storage Needs Explode. Prachi Patel; *ScientificAmerican.com*, опубликовано онлайн 31.05.2016.