

Программа структурно-функционального изучения генома человека не только открыла увлекательные перспективы исследования наследственности — этого фундаментального свойства живого мира, но породила трудные задачи технического и концептуального характера.

Центральное положение в программе занимает глобальное секвенирование генома человека. Это поистине гигантская задача: здесь речь идет о нуклеотидной последовательности в $3 \cdot 10^9$ пар нуклеотидов — таков совокупный размер ДНК 23 хромосом гаплоидной клетки человека. Для выполнения подобной аналитической работы уже недостаточно ручных методов секвенирования по Сэнгеру или Максаму — Гилберту — один исследователь в течение года способен секвенировать 50 тыс., а 100 тыс. нуклеотидов уже составляют потолок человеческой работоспособности. Современные автоматические секвенаторы, дорогие и сложные приборы, обладают высокой производительностью, но не устраняют всех затруднений. Конечно, возможно их дальнейшее техническое усовершенствование, но более вероятно создание каких-то иных концепций секвенирования с последующей разработкой на их основе новых автоматов; поиски в этом направлении только начинаются.

Положение с секвенированием может служить иллюстрацией ситуации, возникшей в программе по изучению генома человека. Суть этой ситуации заключается в том, что задача, поставленная программой, в принципе и в перспективе доступна решению, но пока исследователь еще не располагает нужными для этого средствами. А вместе с тем существуют убежденность и уверенность в том, что целенаправленные усилия в итоге приведут к желаемой цели. Однако, если бы мы не поставили себе задачу секвенировать геном человека в обозримый промежуток времени, скажем, в течение 15 лет, то работу можно было бы выполнить и имеющимися средствами без поисков новых. Правда, для этого понадобились бы, вероятно, столетия, и потому такой вариант упомянут не как практически приемлемый, а для уяснения положения.

В конце концов последовательность в 3 млрд нуклеотидов нам нужна не сама по себе, но для суждений и знаний более высокого порядка. Раскрытые в результате секвенирования последовательности нуклеотидов генома должны быть сопоставлены с известными физиологическими процессами организма и связаны, прежде всего, с белковыми веществами, несущими структурную и каталитическую функции. До сих пор остаются загадочными процессы развития многоклеточных организмов, человека в том числе, неясны пути и механизмы эволюции живых существ, и, нужно думать, секвенирование генома, если и не даст окончательного решения этих проблем, то существенно их продвинет.

Особенное значение в дальнейшей судьбе человека приобретают сейчас проблемы экологии. Существует реальная опасность неотвратимого роста генетического груза человека, т. е. отягощения опасными для него мутациями вследствие действия вредных факторов окружающей среды, возникновение которых вызвано его же собственной неразумной и недальновидной хозяйственной деятельностью. Для предотвращения этих губительных последствий, кроме оздоровления окружающей среды, необходимы слежение за накоплением мутаций, развитие диагностики, профилактики и лечения генетических болезней. А это, по современным представлениям, немыслимо без знания структуры генома человека. И, таким образом, мы убеждаемся в том, что программа изучения генома служит не только фундаментальным целям, но и благу человека.

Воспитанные на коротких последовательностях транспортных рибонуклеиновых кислот и доступных зрительному обозрению относительно коротких фрагментов ДНК мы до сих пор мало задумывались о наших возможностях в отношении огромных последовательностей, которые доставит нам программа секвенирования генома человека. Здесь приходится сталкиваться с задачей, которая до сих пор не возникала. Секвенировать геном мы можем даже существующими методами, хотя крайне медленно. Но объем информации, которая поступит в наше распоряжение в результате секвенирования, достигнет такого масштаба, что расшифровать и осмыслить ее человек окажется неспособным ни сейчас, ни в будущем — это лежит за пределами его природных возможностей.

Действительно, последовательность оснований человеческого генома в окончательном виде составит 3 млрд нуклеотидов, а на промежуточных этапах она будет еще больше за счет неизбежного повторения одних и тех же последовательностей разными исследователями, перекрытия и других причин. Следует прибавить к этому последовательности иных геномов, которые будут изучены параллельно с геномом человека. Для хранения этой информации понадобятся огромные объемы памяти. Еще более трудные задачи возникнут при анализе генома. Можно назвать здесь некоторые очевидные задачи типа отыскания последовательностей по заданным признакам, их сравнения, присвоения определенным структурам функциональных характеристик и т. п.

Намечается проблема архитектуры генома человека. Сейчас мы ориентируемся на последовательности, кодирующие белки, что естественно. Но эта часть составляет только 3—5 % генома, а остальное принадлежит ДНК с неизвестной или неопределенной функцией, которую некоторые именуют унизительным термином «хлам»; последнее вряд ли справедливо.

Несмотря на щедрое разнообразие, природа, когда рассматриваются отдельные явления, оказывается крайне экономной. А если это так, то естественно возникает предположение о плане построения генома человека, его архитектуре, связанной с функциональными свойствами. Едва ли этот план может быть разгадан с помощью обычных логических приемов.

Решение вопроса может быть найдено только в самой широкой компьютеризации исследований генома человека, в информатике. Этот взгляд стал в последнее время господствующим, и многие считают, что если техника секвенирования вступила на надежный путь, то информатика стоит в самом начале пути и даже еще не сделала своего выбора.

Состояние дел в данной области отражают статьи, собранные в этом и частично следующем номерах журнала «Биополимеры и клетка». Я желаю их авторам, посвятившим себя трудной задаче компьютерной поддержки расшифровки генома человека, дальнейших успехов.

© Академик АН СССР А. А. БАЕВ, 1990