



УДК 577.2

НОВОСТИ В ИССЛЕДОВАНИЯХ ГЕНОВ И ГЕНОМОВ

© 1996 г. Е. Д. Сverdlov

Институт биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН,
117871, Москва, ГСП-7, ул. Миклухо-Маклая, 16/10

Поступила в редакцию 22.12.95 г.

В 1988 г. Национальным научным советом США (National Research Council) была ясно сформулирована идея национального проекта "Геном человека", предполагающего к 2005 г. завершить определение полной последовательности 3 миллиардов нуклеотидных звеньев, кодирующих всю генетическую информацию человека. Введение этого проекта в действие означало, что техническое развитие молекулярных исследований в области наук о жизни достигло качественно нового уровня, позволяющего решать принципиально новые задачи. Комплекс исследований молекулярных основ жизнедеятельности, в который входит и биоорганическая химия, изменил свое лицо идеологически и технически.

Идеологически – переходя:

- от исследований отдельных генов к исследованиям целых геномов, от генетики скрещивания к парасексуальной генетике (термин, определяющий исследования, основанные на введении генетического материала в клетку или в целый организм неполовым путем и создании условий для его последующей передачи по наследству);
- от классической прямой генетики, двигавшейся к идентификации генов от признаков, кодируемых этими генами, к "обратной" генетике, которая сначала идентифицирует фрагмент генома, а уже затем занимается выяснением, какой признак определяет этот фрагмент;
- от работы с изолированными генами или продуктами их экспрессии к изучению их эффектов на уровне целого организма, в котором роль этих продуктов осмысливается в контексте комплекса их природных взаимосвязей.

Технически – переходя:

- к направленному воздействию на генетический аппарат клетки или целого организма, приводящему к его наследственному изменению;
- к исследованиям структур отдельных молекул путем их специфической амплификации;
- к методам эволюции в пробирке: направленному систематическому изменению свойств взаимодействующих молекул, позволяющему достигать их максимального взаимодействия;

- к переходу от традиционных белковых ферментов к искусственным ферментам полинуклеотидной природы;
- к тотальной автоматизации и роботизации экспериментальной работы и к максимальному переносу груза анализа экспериментальных результатов на компьютер;
- к созданию интегральных информационных баз данных, позволяющих быстро сопоставлять структуры новых продуктов с уже существующими и на основании гомологий делать предварительные выводы о их возможной функциональной роли.

Началась эпоха интегральных исследований геномов.

Предлагаемая вниманию читателей подборка новостей дает несколько примеров того, что происходит сегодня в области исследований генов и геномов. Заметка "Секвенирование геномов" дает информацию о первых двух секвенированиях полноразмерных геномов свободноживущих организмов – бактерии *Haemophilus influenzae* и микоплазмы *Mycoplasma genitalium*. Заметка, посвященная идентификации новых генов болезни Альцгеймера, дает пример типичной для нынешнего времени ситуации, когда гены, ответственные за различные заболевания, идентифицируются и приписываются к хромосомам, картируются. Две другие заметки, каждая в своем роде, дают представление о том, как происходит анализ функций идентифицированных генов.

СЕКВЕНИРОВАНИЕ ГЕНОМОВ

Публикация последовательностей геномов *H. influenzae* (штамм Rd) и *M. genitalium* еще раз показала, насколько информативен интегральный анализ геномов. Плодотворность такого подхода стала понятной, еще когда в 1977 г. Фредом Сэнгером был расшифрован первый ДНК-содержащий геном бактериофага φX174, имеющий длину 5386 п. о. С тех пор были установлены структуры полных геномов множества вирусов, и в их числе геном цитомегаловируса длиной 229 т. п. о.,

который до недавнего времени был самым длинным известным геномом. Однако геномы вирусов функционируют в зависимости от клеточной системы регуляции, и с этой точки зрения две опубликованные последовательности являются первыми расшифрованными геномами свободно существующих организмов.

Микоплазма была выбрана, поскольку она имеет минимальный известный геном среди самореплицирующихся организмов (580 т. п. о.) и может дать представление о минимальном наборе генов, необходимых для независимой жизнедеятельности. *H. influenzae* также имеет сравнительно небольшой размер генома – $1.8 \cdot 10^3$ т. п. о. Эта бактерия относится к виду одного из главных патогенов человека, и поэтому можно надеяться, что знание последовательности ее генома поможет в борьбе со многими бактериальными инфекциями.

Методология секвенирования

Поскольку физических карт ни для одного из названных геномов не существовало, секвенирование осуществлялось так называемым шот-ган-способом. Этот способ предполагает дробление генома на короткие перекрывающиеся фрагменты, клонирование этих фрагментов в вектор для секвенирования (например, в бактериофаг M13), секвенирование большого числа случайно выбранных фрагментов и построение непрерывных последовательностей на основе компьютерной идентификации перекрывающихся фрагментов. В случае обсуждаемых двух геномов сумма длин проанализированных коротких фрагментов в 3–9 раз превышала длины анализируемых геномов. Это означает, например, что для *H. influenzae* было проведено 28000 секвенирующих экспериментов. Работа была выполнена в течение 13 месяцев, и цена секвенирования оказалась равной примерно 50 центам за пару оснований. Для микоплазмы секвенирование было осуществлено на восьми секвенирующих машинах пятью сотрудниками, работавшими в течение 2 месяцев. Частота ошибок была сравнительно небольшой: одна на 5000–10000 п. о. Цена секвенирования составляла 30 центов на пару оснований, т.е. \$200000 для всей микоплазмы.

Примерно такое же время занял анализ полученных данных и сравнение с существующими бактериальными последовательностями. Таким образом были идентифицированы вероятные гены, и предположительные опероны и регуляторные области были помещены на генетическую карту.

Стратегии организации геномов

Геном *H. influenzae* содержит 1743 гена с типичным для бактерий малым содержанием некодирующей ДНК. 1007 из них могли бы иметь

определенные функции в клетке, включая метаболизм аминокислот и липидов, биосинтез кофакторов, производство энергии, транспорт, синтез нуклеотидов и белков, репликацию и транскрипцию ДНК. Оказывается, что примерно 10% ДНК типичной бактерии отвечает за функции энергетического метаболизма, 17% – транскрипции и трансляции, 12% – транспортные функции и 8% посвящены кодированию белков оболочки клетки. 736 потенциальных генов, т.е. примерно 40% генома, не имели аналогов среди секвенированных до этого генов из других организмов. Многие необычные вещи выявляются в ходе такого интегрального анализа. Например, у этой бактерии отсутствуют три фермента, участвующие в цикле трикарбоксильных кислот (цикл Кребса). Совершенно неожиданным свойством этой бактерии является способность узнавать и поглощать ДНК, которую другие бактериальные клетки этого же вида оставляют после смерти. Для этого ДНК бактерии “маркирована” и содержится 1465 идентичных копий узнаваемой бактериальными клетками 9-звенной последовательности, вкрапленной в более длинную 29-звенную.

В случае 470 предсказанных кодирующих последовательностей микоплазмы распределение количеств ДНК-информации по различным функциям сильно отличается от предыдущего случая. 90 белков, по-видимому, участвуют в трансляции, а полная репликационная система включает около 30 белков. 140 (30%) из 482 генов кодируют мембранные белки. Это очень много, особенно если учесть, что микоплазма содержит только один тип мембраны. В то время как *Haemophilus* имеет 68 генов биосинтеза аминокислот, микоплазма имеет только один (!). У микоплазмы нет генов цитохромов или цикла лимонной кислоты. Микоплазма не является наиболее эволюционно ранней зубактерией и, по-видимому, эволюционно упростила свою жизнь за счет частичного использования жизненных функций клеток млекопитающих. Все же она отдает почти 5% своего генома на повторяющиеся элементы, кодирующие адгезин, благодаря которому микоплазма прикрепляется к клетке и который, возможно, меняется путем рекомбинаций, что позволяет избегать иммунной реакции клетки. 90 генов, найденных в микоплазме, не найдено у бактерии. Интересно, что микоплазма обходится без транскрипционного фактора для стрессовой реакции. Вообще, *Haemophilus* содержит в 10 раз больше генов, вовлеченных в осуществление регуляторных функций, чем микоплазма.

Несмотря на большой объем доступной информации по последовательностям микроорганизмов, треть открытых рамок считывания у обоих организмов не нашла гомологий с известными функциональными последовательностями. Это может означать, что организмы выработали

некоторые специализированные белки для выполнения уникальных функций. Возможно и более простое объяснение – информация по другим организмам еще недостаточно, для того чтобы делать обоснованные заключения.

Уроки и прогнозы

Полученные результаты показывают, что мы реально вступаем в эпоху сравнительного анализа геномов путем их тотального секвенирования. Уже сейчас несколько бактериальных геномов близки к полной расшифровке. Использование полученной структурной информации в сочетании с развитыми технологиями направленного мутагенеза, по-видимому, позволит в ближайшее время идентифицировать большое число неизвестных до сегодняшнего дня генов. Такой анализ с фундаментальной точки зрения даст возможность оценить стратегии построения различных геномов в соответствии с условиями их существования и приблизит науку к пониманию принципов организации живой материи и ее эволюции.

В процессе полного секвенирования бактериальных геномов накапливается опыт и разрабатывается технологическая база для секвенирования сложных геномов, включая геном человека. Многие склонны считать, что использованный шот-ган-принцип практически без изменений может быть перенесен и на геномы, состоящие не из миллионов, как бактериальные, а из миллиардов нуклеотидных звеньев. Вряд ли такие надежды оправдаются, но то, что опыт бактериального секвенирования позволит более рационально организовать работу с человеческим и другими сложными геномами, несомненно.

Помимо решения фундаментальных проблем исследования по полному секвенированию геномов открывают и важнейшие практические перспективы. Возможности современного анализа позволяют осуществить полное секвенирование геномов 25 главных бактериальных и паразитических патогенов за 5 лет. За примерно 100000000 долларов появится возможность узнать последовательность каждой детерминанты, определяющей вирулентность, каждого белкового антигена и каждой мишени для лекарства.

Особенно важно идентифицировать гены, определяющие вирулентность патогенов. В сочетании с информацией о структурах их мишеней это позволит выработать рациональные пути модификации микроорганизмов для медицинских, сельскохозяйственных, промышленных и экологических целей. Такая информация чрезвычайно важна ввиду постоянно возникающих новых инфекционных заболеваний и возвращения старых, которые опрометчиво были объявлены побежденными. Сегодня, когда сильно возрастает волна индивидуального и государственного терроризма, следует видеть и еще одну важную сторону использования получаемой структурно-функциональной информации: она позволит разработать эффективные способы контроля за использованием и распространением опасных биологических агентов, их быструю идентификацию и обезвреживание.

Использованные источники

- Novak R. // Science. 1995. V. 269. P. 468–470.
 Fleishman R.D. et al. // Science. 1995. V. 269. P. 613.
 Smith H.O. et al. // Science. 1995. V. 269. P. 538–540.
 Fraser C.M. et al. // Science. 1995. V. 270. P. 397–403.
 Goffeau A. // Science. 1995. V. 270. P. 445–446.
 Bloom B.R. // Nature. 1995. V. 378. P. 236.

НОВЫЕ ГЕНЫ БОЛЕЗНИ АЛЬЦГЕЙМЕРА

Болезнь Альцгеймера (А.), связанная с потерей памяти у пожилых людей, поражает ежегодно до 20 миллионов человек в мире. Болезнь может возникать самопроизвольно или носить выраженный наследственный характер. В 1991 г. на хромосоме 21 (хр. 21) был идентифицирован ген, ответственный за наследственную болезнь А. примерно в 2–3% случаев. Он кодирует предшественник амилоидного белка. Недавно ген, ответственный за 80% случаев наследственной болезни А., был локализован на хр. 14. Появившиеся недавно в журнале "Science" статьи описывают идентификацию еще одного такого гена на хр. 1. Новый ген, возможно, является причиной большинства остальных случаев наследственной болезни А. По последовательности он гомологичен

Гены, вовлеченные в возникновение болезни А.

Хромосома	Тип гена	Возраст больных	% случаев		Белковый продукт
			наслед.	всех	
14	Аутосомный, доминантный	30–60	70–80	5–10	Мембранный S182
19	Генетический фактор риска	>60	–	40–50	АpoE4
21	Аутосомный, доминантный	45–65	2–3	<1	Предшественник амилоидного белка
1	Аутосомный, доминантный	40–70	20	2–3	Мембранный

ранее локализованному на хр. 14 гена. Новый ген идентифицирован в семьях, происходящих от двух семей – выходцев из колонии волжских немцев, и картирован на хр. 1 с помощью молекулярно-генетических маркеров. Интересно, что клонирование гена было осуществлено с помощью развитой в ходе программы “Геном человека” системы EST (expressed sequenced tags – секвенированных экспрессированных ярлыков), представляющих собой короткие секвенированные фрагменты ДНК, происходящие из экспрессирующихся генов. В банке таких последовательностей был обнаружен один EST, гомологичный фрагменту гена болезни А. на хр. 14. Он был локализован на хр. 1 в той области, которая совпала с ожидаемой локализацией гена болезни А. на этой хромосоме. Определение последовательности гена у многих членов семей волжских немцев, страдающих этим наследственным дефектом, обнаружило у всех одну и ту же мутацию. Сравнение последовательностей генов на хр. 1 и 14 показывает, что оба они кодируют гомологичные, вероятно интегральные, мембранные белки длиной около 450 а. о. Функции новых генов пока неизвестны, хотя показано, что клетки, полученные от больных с дефектом на хр. 14, синтезируют аномально высокие количества β -амилоида, предшественник которого кодируется геном на хр. 21.

Современная информация о генах, вовлеченных в возникновение болезни А., представлена в таблице.

Использованные источники

- Barinaga M. // Science. 1995. V. 268. P. 1845–1846.
 Barinaga M. // Science. 1995. V. 269. P. 917–918.
 Levy-Lahad E. et al. // Science. 1995. V. 269. P. 970–973.
 Levy-Lahad E. et al. // Science. 1995. V. 269. P. 973–977.

НОКАУТ ГЕНОВ И БОЛЕЗНЬ ХАНТИНГТОНА

Болезнь Хантингтона, которая проявляется во взрослом возрасте, представляет собой доминантную нейродегенеративную болезнь, вызываемую удлинением САG-повтора в участке гена, кодирующем N-концевую часть белка “хантингтина”, функция которого неизвестна. В норме длина этого повтора варьирует у разных индивидуумов от 11 до 34 триплетов. У больных его длина колеблется от 37 до более чем 100 кодонов. Этот дефект, по-видимому, каким-то образом проявляется через измененный белок. Одна из возможностей заключается в том, что удлинение полиглутаминового кластера в белке уменьшает его нормальную активность (механизм “потери функции”). Однако индивидуумы, у которых одна копия гена инактивирована за счет повреждения

структуры хромосомы в этом месте, а не за счет удлинения повтора, не проявляют признаков болезни Хантингтона, хотя и имеют пониженное содержание “хантингтина”. Это противоречит механизму “потери функции”, хотя остается возможной так называемой негативной доминантности, при которой молекулы белка с утраченной функцией ингибируют активность нормального белка. Другая возможность заключается в том, что белок с удлиненным полиглутаминовым участком приобретает новую, вредную функцию.

Чтобы сделать выбор между этими двумя возможными механизмами возникновения болезни вследствие мутации, мышинный гомолог гена был инактивирован путем ген-таргеттинга. Гетерозиготные мыши, содержащие одну нормальную и одну поврежденную аллель, были фенотипически нормальны, тогда как гомозиготные мыши погибали на стадии эмбриона. Если бы болезнь была связана с эффектом негативной доминантности, следовало бы ожидать, что гетерозиготная мышь с одной инактивированной аллелью будет фенотипически нормальной подобно людям с поврежденной хромосомой. Гомозиготная же по инактивированному гену мышь в этом случае должна была рождаться нормальной и по аналогии с проявлением болезни у людей проявлять патологию во взрослом возрасте. Наоборот, если болезнь вызывается доминантным приобретением функции вследствие удлинения повтора, ни гетерозигота, ни гомозигота не должны развивать симптомов болезни, а наблюдаемый эффект гомозиготности должен быть связан с нарушением нормальной функции белка. Таким образом, полученные результаты согласуются с последним вариантом, означающим, что нейропатология включает в себя механизм “приобретения функции” и что “хантингтин” критичен для раннего эмбрионального развития, перед появлением нервной системы.

Использованный источник

- Duyao M.P. et al. // Science. 1995. V. 269. P. 407–410.

ВОЗМОЖНАЯ РОЛЬ ЭЛОНГАЦИИ ТРАНСКРИПЦИИ В ОПУХОЛЕВОЙ ТРАНСФОРМАЦИИ

РНК-полимераза II, фермент, который осуществляет транскрипцию генов, кодирующих белки у эукариотических организмов, так же как и РНК-полимеразы прокариот, в процессе элонгации делает многочисленные паузы. Часть этих остановок происходит практически сразу после начала синтеза РНК. Возможно, что большая часть времени транскрипции затрачивается РНК-полимеразой на этих остановках. В прокариотических системах регуляция транскрипции на уровне

элонгации хорошо известна. Например, белки N и Q бактериофага λ функционируют как антитерминаторы транскрипции. В последнее время появляются многочисленные данные, что регуляция на уровне элонгации может осуществляться и у эукариот, где длительное время наиболее важную роль в регуляции экспрессии генов приписывали инициации транскрипции. Нарушения в этой системе регуляции, возможно, могут проявляться, в частности, в опухолевой трансформации клеток.

Три статьи в журнале "Science" [1–3], сопровождаемые краткой рецензией [4], приводят данные, что фактор элонгации транскрипции элонгин негативно регулируется продуктом антионкогена фон Хиппеля (von Hippel-Lindau tumor suppressor gene, *VHL*). Дефект в этом антионкогене приводит к наследственной болезни фон Хиппеля – предрасположенности к различным формам рака. Спорадические мутации в этом гене также наблюдаются при раковых перерождениях. Ген болезни фон Хиппеля был клонирован и секвенирован в 1989 г., но его структура не давала оснований для выводов о возможной функции белка.

Элонгин представляет собой гетеротример, состоящий из большой каталитической субъеди-

ницы А и двух меньших субъединиц, В и С, играющих роль позитивных регуляторов в элонгине. Сборка элонгина осуществляется таким образом, что сначала ассоциируют субъединицы В и С, которые затем связываются с субъединицей А. Нормальный продукт гена *VHL* также связывается с комплексом В/С и не взаимодействует с белком А. Связывание с В/С-комплексом обусловлено фрагментом белка *VHL*, который гомологичен субъединице А. Взаимодействие белка фон Хиппеля с этими субъединицами нарушает функцию элонгина. Мутанты же *VHL* с измененной областью гомологии с субъединицей А не образуют комплекса с элонгином. Это позволяет предположить, что взаимодействие белка фон Хиппеля с элонгином лежит в основе его супрессорной функции.

Использованные источники

1. Duan R.X. et al. // Science. 1995. V. 269. P. 1402–1406.
2. Aso T. et al. // Science. 1995. V. 269. P. 1439–1443.
3. Kibel A. et al. // Science. 1995. V. 269. P. 1444–1446.
4. Krumm A., Groudine M. // Science. 1995. V. 269. P. 1400–1401.

Recent Advances in Studies of Genes and Genomes

E. D. Sverdlov

*Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,
ul. Miklukho-Maklaya 16/10, Moscow GSP-7, 117871 Russia*

Сдано в набор 03.01.96 г.

Подписано к печати 06.03.96 г.

Формат бумаги 60 × 88¹/₈

Офсетная печать

Усл. печ. л. 10.0

Усл. кр.-отт. 4.4 тыс.

Уч.-изд. л. 10.6

Бум. л. 5.0

Тираж 421 экз.

Зак. 3964