

ОБЗОРЫ

УДК 575.113

© Коллектив авторов

СИНТЕЗ ИСКУССТВЕННОГО ГЕНОМА КАК ОСНОВА СИНТЕТИЧЕСКОЙ БИОЛОГИИ

С.П. Радько, А.П. Ильина, Н.В. Бодоев, А.И. Арчаков*

Государственное учреждение Научно-исследовательский институт биомедицинской химии им. В.Н.Ореховича РАМН, 119121 Москва, Погодинская ул., 10;
тел. (495) 246-16-41; факс (495) 246-37-71; эл. почта: radko@ibmc.msk.ru

Достижения последних лет в секвенировании геномов, особенно вирусных и бактериальных, в совокупности с развитием методов биоинформатики и молекулярной биологии, создали предпосылки для перехода от синтеза генов к сборке целых геномов из химически синтезированных блоков – олигонуклеотидов. Создание искусственных геномов и, в перспективе, искусственных клеток, без сомнения, окажут огромное влияние на углубление знаний о механизмах функционирования живых систем на клеточном уровне, на пути зарождения и эволюции жизни, а также на биотехнологию будущего, и сформируют предпосылки для дальнейшего развития синтетической биологии и нанобиотехнологии.

Ключевые слова: синтез генома, синтетическая биология, искусственная клетка, геннобиотехнология.

ВВЕДЕНИЕ. Интерес к созданию синтетических геномов тесно связан с концепцией “минимального генома”, то есть минимального набора генов, необходимых (микро)организму для жизнедеятельности и размножения в наиболее благоприятных (лабораторных) условиях, а именно в присутствии всех веществ, необходимых для жизнеобеспечения и в отсутствии “стрессовых нагрузок” со стороны окружающей среды [1]. Наименьшим бактериальным геномом, известным на сегодня, является геном *Mycoplasma genitalium*, чей размер составляет 580070 пар оснований (п.о.) и содержит около 470 генов [2]. Наименьший известный геном был обнаружен у *Buchnera* - внутриклеточного симбиота афидов – его размер составляет 450 тыс. п.о. [3]. Предполагая, что нижний предел минимального генома задается генами, ответственными за системы трансляции, транскрипции, репликации, поддержание примитивного метаболизма (в основном ограниченного гликолизом) и минимальной транспортной системы, их количество оценивается как 100-150 [1]. При этом “минимальный геном” размером ~150 генов включает 25 генов, отвечающих за полный синтез ДНК/РНК, 120 генов, кодирующих белки (включая синтез РНК и 55 рибосомных белков) и 4 гена, участвующих в липидном метаболизме [1]. Принимая во внимание утверждение, что первичные рибосомы состоят из рРНК, ассоциированной с основными пептидами [4], и исключая часть генов, кодирующих рибосомальные белки и некоторые другие ферменты, можно предположить, что минимальное число генов составляет около 110 [5]. Это число

* - адресат для переписки

может быть ещё немного уменьшено, если принять предположение [6-9], что, как и в случае упрощенной версии белкового синтеза, возможно существование более простого ферментативного набора репликации. В частности, установлено, что одна единственная полимеразы может играть мультиплетную роль: как ДНК-полимераза, транскриптаза и праймаза [10]. Однако вопрос о том, сможет ли такой “минималистский организм” поддерживать себя и размножаться с какой-то разумной скоростью даже в самых благоприятных условиях, остается открытым.

Полное секвенирование геномов ряда микроорганизмов открыло возможность оценить размер минимального генома, используя методы сравнительной биоинформатики [1, 11-15]. Впервые данный анализ был выполнен на полностью секвенированных геномах *Haemophilus influenzae* и *M. genitalium* и выявил, что размер минимального набора лежит в пределах от 265 до 350 генов [11]. Однако проведенное позднее более тщательное сравнительное исследование геномов, чьи полные последовательности известны (всего в работе были использованы геномы 21 бактерии, археобактерий и эукариотических организмов), показало, что только около 80 общих генов присутствовало во всех исследованных организмах – число, которое ниже даже самых смелых оценок [1]. Возможно, дивергенция достигает такой степени, что делает гены неузнаваемыми для используемых алгоритмов сравнения, или что многие гены возникли независимо друг от друга в процессе эволюции [16]. В любом случае, полученные результаты демонстрируют ограниченность подхода, основанного только на сравнительной геномике, для определения минимального набора генов.

Первые экспериментальные подходы к оценке размера минимального генома основывались на мутагенезе ограниченного числа хромосомных локусов микроорганизма с последующей экстраполяцией частоты незаменимых генов на весь геном [17]. Используя такой подход и *B. subtilis* как модельный организм, размер минимального генома был оценён как 560 тыс. п.о. Hutchinson и соавторы [18] предложили использовать глобальный транспозонный мутагенез для определения числа генов “минимального генома”. По их оценке, из 480 генов генома *M. genitalium*, кодирующих белки, только 265-350 являются существенными для поддержания жизни и размножения бактерии в лабораторных условиях. Однако их оценка является скорее “оценкой сверху”, так как может отражать число генов, необходимых для саморазмножения бактерии с определенной скоростью, которая задается экспериментатором субъективно [16, 18]. Сравнительный анализ геномов микроорганизмов в совокупности с экспериментальными данными о генах, существенных для клеточного жизнеподдержания и размножения, дал оценку размера минимального генома как состоящего из 206 генов [14].

Создание синтетических геномов также связано с идеей конструирования “искусственных клеток”. Сам термин “искусственная клетка” трактуется многими исследователями различно, часто в зависимости от тех задач, на решение которых они нацелены [5, 19-25]. Диапазон интересов, которые стимулируют исследования в области конструирования искусственных клеток, простирается от фундаментального, заключающегося в попытке понять, как зародилась жизнь и как она эволюционировала в сверхсложные системы, каковыми являются сегодняшние живые организмы, до сугубо утилитарных [26]. Практический аспект работ по конструированию искусственных клеток состоит, по сути, в создании наноразмерных биореакторов, производящих требуемые биосубстанции в нанокolicествах [25, 27]. Такие “нанореакторы” должны быть способны к самоподдержанию, как минимум, и к самовоспроизводству, как максимум. Их способность к саморазмножению может быть использована как для простого масштабирования производства, так и для решения локальных задач, например, адресная доставка терапевтических средств в очаг патологического процесса [28-31].

Большинство проблем, стоящих на пути к созданию искусственной клетки, далеки сегодня даже от своего теоретического решения. Тем не менее, в этой области проводятся достаточно интенсивные исследования. В США часть таких

исследований ведется в рамках National Nanotechnology Initiative – специальной межведомственной программы с бюджетом в сотни миллионов долларов (www.nano.gov). Европейский Союз в декабре 2004 года учредил специальный институт The European Centre for Living Technology (бюджет 9 млн. евро) для финансовой поддержки и координации работ в области создания искусственной клетки в рамках европейского проекта Programmable Artificial Cell Evolution (<http://134.147.93.66/bmcmyp/Data/PACE/Public>).

Конструирование искусственной клетки требует, как минимум: 1) создания оболочки, разграничивающей внутриклеточный объем и внешнюю среду и позволяющей осуществлять направленный и контролируемый транспорт веществ внутрь клетки и наружу; 2) создания молекулярного аппарата для транскрипции и трансляции генетической информации в белковые продукты и для выполнения корректного фолдинга синтезируемых белков; 3) создания синтетических геномов.

По составу к мембранам клеток наиболее близки оболочки липидных бислойных везикул (липосом). Впервые липосомы (мультиламелярные липосомы), состоящие из концентрированных липидных бислоев, были получены в шестидесятых годах прошлого века [32]. Разработаны стандартные методики изготовления мультиламелярных липосом размером от 100 нм до 1 мкм [33, 34]. При обработке липосом ультразвуком образуются более однородные униламелярные липосомы (от 25 до 100 нм). Поскольку диаметр клетки составляет порядка 0,2-0,5 мкм, то легко получаемые и стабильные большие мультиламелярные липосомы могут вмещать достаточное число макромолекул и быть пригодными для формирования искусственной клеточной системы. Но липосомы сами по себе являются просто оболочкой, поэтому степень пропускной способности липидных бислоев (диффузия маленьких, нейтральных молекул через стенки липосом и барьер для заряженных цитерионов и макромолекул) регулируется путем манипуляции молекулярного веса и состава бислоя [35, 36].

Вторая проблема – создание молекулярного аппарата для транскрипции и трансляции генетической информации. Существует большое число исследований относительно проведения биохимических процессов в липосомах [37-44]. Работа по ферментативной полимеризации ADP в поли(А) с помощью полинуклеотид-фосфорилазы в липосомах [38] способствовала проведению исследований относительно репликации в липосомах матрицы РНК с помощью Q β репликазы [41] и синтезу мРНК используя матрицу ДНК и T7 РНК полимеразу [44]. Другой важный процесс в клетках это белковый синтез, где большинство реакций протекает на рибосомах. В 1999 г. появилось сообщение о рибосомальном синтезе в липосомах поли(Phe) с помощью поли(U), используемого в качестве РНК-мессенжера [44]. Таким образом, комбинируя синтез мРНК и белков, можно сконструировать ген-экспрессирующую систему. Кроме того, было продемонстрировано, что с помощью сопряженных бесклеточных транскрипционной и трансляционной систем в липосомах может быть синтезирован GFP – green fluorescent protein (зелёный флуоресцирующий белок) [45, 46]. Следующим шагом стало конструирование в липосомах функциональной бесклеточной генетической сети [47], создание транскрипционного активационного каскада, в котором белковый продукт первого шага необходим для запуска второго шага. В исследованиях Ishikawa и сотрудников был выбран T7 РНК полимеразный ген для первого шага и, для второго шага, GFP ген, в котором экспрессия белка контролировалась T7 промотором [48].

Создание синтетических геномов является, пожалуй, наиболее близко стоящей к успешному решению проблемой, связанной с конструированием самовоспроизводящихся искусственных клеток. По крайней мере, современное состояние биоинформатики в совокупности с методами молекулярной биологии и автоматизированного химического синтеза олигонуклеотидов, могут в принципе позволить собрать геном, сравнимый с размером бактериального генома. Единственный вопрос – цена и длительность работы. Например, Institute for

Biological Energy Alternatives, учрежденный и руководимый J. Craig Venter (один из столпов программы “Геном человека”) и Hamilton O. Smith (Нобелевский лауреат 1978 г. в области медицины), получил от Министерства энергетики США 3-х годичный грант в 12 млн. долларов для реализации проекта, посвященного созданию “минимального организма” (www.microbialgenome.org/announcement/doereward.shtml). В рамках проекта предлагается удалить весь генетический материал из модельного микроорганизма, *M. genitalium* синтезировать искусственную хромосому, содержащую минимальный набор генов, необходимый для поддержания жизни данной бактерии, вставить ее в клетку и проверить способность такой клетки к выживанию и репродукции.

1. Синтез искусственного генома.

Синтез генома берет свое начало в химическом синтезе генов, история которого насчитывает уже более 30 лет. Khorana и сотрудники впервые предложили план синтеза гена тРНК^{Ala} дрожжей протяженностью 77 п.о. путём сборки дуплекса ДНК за счет перекрывающихся комплементарных участков коротких (5-20 нуклеотидов) олигонуклеотидов с последующим энзиматическим лигированием [49]. Последовавшая за этой публикацией серия статей была посвящена экспериментальному синтезу ряда генов [50], что представляло по тем временам титаническую работу. В дальнейшем химическому синтезу генов были посвящены многочисленные публикации различных научных коллективов [51]. Внедрение в практику синтеза олигонуклеотидов фосфорамидитов и тетразола как катализатора [52] с последующим появлением и быстрым распространением автоматических синтезаторов ДНК к началу 90-х годов позволило упростить синтез олигонуклеотидов. Одновременно происходило развитие методов сборки олигонуклеотидов в протяженные молекулы ДНК. Была продемонстрирована возможность использования химического лигирования для сборки олигонуклеотидов [53]. Дополнительно было предложено собирать гены из протяженных фрагментов ДНК, используя “серийное клонирование” (serial cloning). Кольцевая форма плазмиды, линеаризованной с помощью *BsmI*, восстанавливается *in vivo* с помощью “олигонуклеотид-зависимой” репарации двунитевых разрывов. Синтетический олигонуклеотид, вставляемый в плазмиду после линеаризации, несет участок гена протяженностью 70-100 нуклеотидов (нт) и сайт *BsmI* на 3'-конце, что позволяет восстановить сайт рестрикции и продолжить присоединение следующих участков гена. Процесс повторяется до тех пор, пока ген не будет полностью собран [54]. Следующим шагом на пути совершенствования сборки генов было появление полимеразной циклической сборки (ПЦС) (polymerase cycling assembly, PCA), впервые использованной для синтеза гена HIV-2 Rev длиной 303 п.о. [55]. Развитие методологической базы позволило преодолеть символический барьер в 1 тыс.п.о. В 1990 Mandeck и соавт. собрали плазмиду протяженностью 2,1 тыс.п.о. из 30-ти химически синтезированных олигонуклеотидов с помощью “серийного клонирования” [56]. Позднее Stemmer и соавторы, используя ПЦС, синтезировали фрагмент ДНК протяженностью 1100 п.о., содержащий ген TEM-1 бета-лактомазы, из 56 олигонуклеотидов и плазмиду размером 2,7 тыс. нуклеотидов из 136 олигонуклеотидов [57]. К настоящему времени синтетическая ДНК наибольшего размера (~32 тыс. нуклеотидов) содержащая кластер генов, кодирующих мегафермент поликетидсинтазу *E. coli*, была получена Kodumal и соавторами [58]. Авторы использовали олигонуклеотиды длиной 40 нт, которые собирались в дуплексы ДНК (“синтоны”) протяженностью 500-800 п.о. с помощью ПЦС. Сегменты ДНК размером ~ 5000 п.о. собирались из синтонов клонированием с использованием оригинальной селекции по комбинации устойчивости к нескольким антибиотикам. Кластер генов поликетидсинтазы был собран из сегментов с использованием традиционных методов клонирования. Функциональность полученных генов была подтверждена их способностью обеспечить синтез мегафермента и его продуктов в бактерии.

В 2002 г. Cello и соавт. [59] впервые сообщили о синтезе полноразмерной кДНК генома полиовируса (~ 7500 п.о.). Синтез включал следующие основные этапы: 1) химический синтез и очистку олигонуклеотидов “положительной” и “отрицательной” полярности (в соответствии с направлением от 5’- к 3’-концам и наоборот); 2) сборку сегментов ДНК длиной 400-600 п.о. за счет перекрывающихся комплементарных участков олигонуклеотидов с последующим лигированием или использованием ПЦС; 3) сборку методами клонирования трех фрагментов генома размерами от ~2 000 до ~3 000 п.о. из сегментов с последующей сборкой генома из данных фрагментов. Синтетическая кДНК транскрибировалась в вирусную РНК с помощью РНК-полимеразы фага Т7. Полученная РНК вызывала синтез вирусных частиц при помещении в бесклеточные экстракты. Полученные *de novo* вирусы обладали инфекционной активностью и имели биохимические и патогенетические характеристики полиовируса [59].

Для получения полноразмерной геномной ДНК второго полноразмерного синтезированного генома - генома фага Х174 размером 5386 п.о. Smith и соавторы [60] использовали только лигирование и ПЦС. Как сообщают авторы, весь процесс занял около 2 недель. Все 259 химически синтезированных олигонуклеотидов размером 42 нуклеотида каждый, очищенные с помощью гель-электрофореза и полностью покрывающие последовательности обеих нитей ДНК, были смешаны в эквимоллярных количествах и лигированы. Полученные фрагменты различной протяженности (лигирование не давало полноразмерных ДНК из-за присутствия некоторого количества дефектных олигонуклеотидов, с одной стороны, и того, что эффективность “сборки” олигомеров в протяженные дуплексы ДНК не бывает 100-процентной) были достроены до “геном-размерных” с помощью ПЦС и амплифицированы. Линейная полноразмерная геномная ДНК была циркуляризована энзиматическим лигированием с использованием заранее предусмотренных сайтов рестрикции для получения липких концов. Трансфекция клеток циркулярным синтетическим геномом приводила к появлению инфекционно-активных фагов [60].

Поскольку задача синтеза отдельных генов и целых геномов бактериальных и вирусных организмов, а также создание на их основе модельных систем, содержащих “минимальный геном”, уже решается, в настоящее время огромный интерес и значение приобретает проблема синтеза искусственных генов, а в перспективе и геномов, эукариотических организмов, в частности, человека. Перспектива создания искусственного генома человека имеет глобальное значение как для дальнейшего развития всей фундаментальной науки, так и для развития медицины, а именно диагностики и лечения различных форм наследственных заболеваний, механизмов их профилактики и лечения. Очевидно, что имеющиеся на сегодня технологии, используемые при синтезе генов и геномов вирусов и бактерий, не отвечают тому объему работы и сложности задач, которые связаны с синтезом генома эукариотического организма. Однако наблюдающийся бурный прогресс в области нанотехнологий может привести к появлению подходов к синтезу и сборке геномов, которые сегодня не известны. Как ни фантастично звучит идея синтеза генома человека сегодня, идея химического синтеза вирусных геномов звучала не менее фантастично не многим более двух десятилетий назад.

2. Методические подходы и технологические решения.

Синтез искусственного генома можно разбить на четыре основных этапа: конструирование набора олигонуклеотидов, химический синтез и очистка олигонуклеотидов, сборка участков генома из олигонуклеотидов, сбор всего генома с применением методов молекулярного клонирования. Каждый из этапов характеризуется собственными подходами и требует решения ряда задач с целью снижения количества нежелательных мутаций в “конечном продукте” – синтетическом геноме.

Первый шаг синтеза генома заключается в создании набора (коротких в 40-50 оснований) олигонуклеотидов, являющихся фрагментами геномной ДНК с оптимизированным триплетным составом. Вторым шагом получения искусственного генома является синтез 40-50-членных олигонуклеотидов, где 5'- и 3'-концы одного олигонуклеотида комплементарны, соответственно, 3'- и 5'-концам двух других и так далее. На третьем этапе работы проводится сборка участков генома из имеющихся олигонуклеотидов с помощью ПЦС. Ясно, однако, что геном-размерную ДНК собрать таким образом невозможно из-за возрастающего количества перекрывающихся участков с высокой степенью гомологии, вплоть до идентичности. Так как спаривание комплементарных участков является температурно-зависимым динамическим процессом, степень дискриминации образования дуплексов между полностью комплементарными последовательностями и последовательностями, образующими одну (или более) неправильную пару оснований, будет определяться тем, насколько близка температура проведения гибридизации, T_H , к температуре плавления данного дуплекса, T_M . Однако различия в T_M дуплексов, образующихся при разбиении условного генома на участки, очевидно, будут существенными. Необходимо также отметить возможность повторов, как прямых, так и обратных, в условном геноме, а также на возможное наличие самокомплементарных участков. Всё это накладывает ограничения на длину фрагмента ДНК, который может быть правильно собран из синтезированных олигонуклеотидов. Последующий сбор генома из фрагментов методами клонирования требует, в свою очередь, введения в них сайтов рестрикции в предусмотренных местах. Следовательно, необходимо, чтобы идентичные натуральные сайты рестрикции отсутствовали в условном геноме. Причем как введение, так и удаление сайтов не должно приводить к изменению в последовательности кодируемых белков. Все или большинство этих проблем можно решить путем оптимизации триплетного состава на первом этапе работы.

Перекодировка генома за счет вырожденности генетического кода позволяет, в принципе, решить указанные выше проблемы как полностью (сайты рестрикции, повторы, самокомплементарные участки), так и частично (унификация T_M). Сама по себе перекодировка генов является процедурой, рутинно используемой для оптимизации продукции рекомбинантных белков через выбор кодонов, соответствующих тРНК, наиболее представленной в клетке-хозяине [61]. Однако перекодировка даже одного гена представляет достаточно сложную и трудоемкую задачу, для решения которой используются компьютеризованные подходы [62, 63] (компьютерная программа Gene2Oligo свободно доступна через <http://www.berry.engin.umich.edu/gene2oligo>). Ясно, что перекодировка геном-размерной ДНК невозможна без дальнейшего развития соответствующих компьютеризованных подходов из-за возрастающего объема работы и сложности задачи. Например, при замене одного кодона на другой, кодирующий ту же аминокислоту, должна быть учтена вышеупомянутая регуляция белкового синтеза через наличие и концентрацию тРНК, соответствующей данному кодону, если условный синтетический геном предназначен для последующего помещения в определенную клетку-хозяина. Так, конструирование набора олигонуклеотидов, использованного для сборки фрагмента ДНК размером более 32 тыс. п.о., потребовало создания собственной компьютерной программы [58].

Химический синтез олигонуклеотидов является многошаговым процессом, где каждый шаг включает депротекцию (т.е. отсоединение защитной, диметокситритильной группы с образованием свободной 5'-гидроксильной группы), присоединение (присоединение фосфорамидитного нуклеозида к растущей цепи через взаимодействие с гидроксильной группой) и кэпирование (деактивацию не прореагировавших 5'-гидроксильных групп ацетилированием). Казалось бы, что такой процесс будет требовать много времени и затрат. Но с решением этой проблемы успешно справляются автоматические синтезаторы олигонуклеотидов. Однако ни одна из реакций не протекает со 100-процентной

эффективностью; это приводит к тому, что суммарная эффективность каждого шага составляет от 98% до 99,5%. В результате, выход полноразмерных олигонуклеотидов при синтезе, например, 50-мера составит 60%, а 500-мера – 0,7% при эффективности синтеза 99%. Хотя образование делеций является наиболее часто встречающейся ошибкой при химическом синтезе олигонуклеотидов, могут также происходить вставки и нуклеотидные замены, например, дезаминирование основания в процессе синтеза может привести к превращению *C* в *U*.

Присутствие дефектных олигонуклеотидов является главным источником ошибок при последующей сборке генов и геномов, приводящим как к низкому выходу полноразмерных фрагментов ДНК, так и к высокой частоте мутаций. Выявление и коррекция мутаций требуют клонирования полученных полноразмерных фрагментов, определения их нуклеотидной последовательности и последующей коррекции мутаций с помощью, например, сайт-направленного мутагенеза, что делает синтез геном-размерной ДНК неприемлемым по времени и трудозатратам. Таким образом, выделение и эффективная очистка олигонуклеотидов становится одной из центральной задач, требующих решения на пути к синтетическому геному.

В то время как последовательности с однонуклеотидной делецией составляют наибольшую фракцию дефектных олигонуклеотидов, рутинный способ очистки олигонуклеотидов с помощью гель-электрофореза является довольно трудоемким и плохо поддающимся автоматизации. Хотя использование достаточно протяженных гелей и позволяет разделять олигонуклеотиды длиной до 100 нт, эффективность разделения *N*- и (*N*-1)-меров прогрессивно падает при *N* около 50-ти и выше. Присутствие дефектных олигонуклеотидов в количестве даже нескольких процентов может резко повысить частоту мутаций. Гель-электрофорез не позволяет также дифференцировать и однонуклеотидные замены. В качестве подхода к проблеме очистки олигонуклеотидов предложено использовать селекцию, основанную на кинетике гибридизации полностью комплементарных последовательностей и последовательностей с одним или более неспаренными основаниями [64]. Для каждой олигонуклеотидной последовательности синтезируются два олигомера, комплементарные к “левой” и “правой” части последовательности и полностью ее перекрывающие. Длина олигомера подбирается в соответствии с набором составляющих его нуклеотидов таким образом, что бы дуплексы, образуемые ими с последовательностями, имели практически одинаковую T_M . Набор “левых” и “правых” олигонуклеотидов иммобилизуется на твердом носителе (например, на магнитных микросферах) и смесь олигонуклеотидных последовательностей проводят через несколько циклов “инкубация-отмывка”, попеременно с каждым из наборов иммобилизованных олигомеров (при температуре, равной T_M). Такой способ очистки, с одной стороны, может быть относительно легко автоматизирован, а с другой, - позволяет эффективно устранить дефектные олигонуклеотиды, снижая частоту ошибок в последовательностях синтезированных генов до 1/1394 п.о., по сравнению с 1/160 п.о. и 1/455 п.о. при использовании соответственно неочищенных олигонуклеотидов и очищенных гель-электрофорезом [64].

Другой проблемой, требующей своего решения в рамках синтеза генома, является создание высокопроизводительных автоматизированных подходов к синтезу олигонуклеотидов. Так, сборка генома размером в 150 тыс.п.о. потребует синтеза 6000 50-мерных олигонуклеотидных последовательностей. Например, при использовании 8-колоночного автоматического ДНК-синтезатора, только синтез такого количества олигонуклеотидов потребует от года до полутора лет. Несомненно, использование одного или нескольких более производительных синтезаторов, как, например, ABI 3900 НТ, позволяющего вести одновременный синтез 48 последовательностей, может частично или полностью решить данную проблему. В принципе, число колонок может быть и больше [65]. Существуют также коммерческие синтезаторы ДНК/РНК, которые могут синтезировать

до 384 олигонуклеотидов одновременно – синтез ведётся в 96-луночных микропланшетах в аргоновой камере (например, синтезатор MerMade-384 [www.bioautomation.com]). Недавно был предложен альтернативный подход, где синтез олигонуклеотидов проводили, используя программируемый микрофлюидный “пикореактор” - упорядоченно расположенные на микрочипе 3698 реакционных камер объемом менее кубического пикометра, связанных системой микроканалов [64, 66]. Это позволяет одновременно синтезировать несколько тысяч последовательностей, используя фотоиндуцированный олигонуклеотидный синтез, хотя и в гораздо меньшем весовом количестве. Для того, что бы получить олигонуклеотиды в количествах, достаточных для сборки генома, каждый нуклеотид несет универсальный праймер с сайтом рестрикции, что позволяет проводить их амплификацию до необходимого количества и последующее удаление праймера.

Третий этап синтеза генома - сборка полученных олигонуклеотидов. ПЦС как сам по себе, так и в комбинации с лигированием, несомненно, является наиболее популярным и перспективным методом сборки олигонуклеотидов в протяженные днДНК (двунитевые ДНК), включая полноразмерные вирусные геномы и кластеры генов. ПЦС представляет, по сути, приложение метода, известного как “DNA shuffling” [67], к смеси химически синтезированных олигонуклеотидов. Как уже упоминалось, олигонуклеотиды суммарно представляют полную последовательность обеих нитей синтезируемой ДНК и имеют взаимно перекрывающиеся участки. После отжига они образуют разнообразные дуплексы, размер которых может простирается от двух олигонуклеотидов с перекрывающимися участками до, возможно, полноразмерной ДНК, которую предстоит синтезировать. Каждый такой дуплекс имеет выступающие одонитевые концы, которые позволяют достроить дуплекс с помощью полимеразной реакции, тем самым, увеличивая длину олигонуклеотидов. В следующем цикле происходит денатурация дуплексов, образование новых комбинаций при отжиге и увеличение размера олигомеров при достройке дуплексов. В результате многочисленных циклов (как правило, около 30-ти) появляется незначительное количество полноразмерных одонитевых ДНК, которые амплифицируют путем добавления праймеров к фланкирующим последовательностям. Проведению ПЦС может предшествовать лигирование олигонуклеотидов, собранных в дуплексы различной длины [60]. Также, “фланкирующие” праймеры могут быть добавлены к смеси олигонуклеотидов с самого начала [64].

На последнем этапе синтеза генома использование ПЦС позволяет получать достаточно протяженные фрагменты ДНК размерами от сотен до тысяч п.о. Такие фрагменты, в свою очередь, выступают как “блоки” для сборки традиционными методами клонирования сегментов генома, содержащих кластеры генов, и имеющих размеры в десятки тыс.п.о. [58]. Как правило, для такой сборки во фланкирующие последовательности фрагментов вводятся сайты рестрикции для создания “липких концов” и последующего лигирования. Предполагается, что такие фрагменты могут быть использованы для сборки синтетических геномов с размерами ~150 или более тыс.п.о. [60].

Хотя общая стратегия сборки генома представляется на сегодня ясной, оптимальный размер фрагментов ДНК, синтезированных с использованием ПЦС, остается темой для дискуссий. Причина этого состоит в том, что пока не ясно, какой подход окажется наиболее эффективным для снижения количества нежелательных мутаций в конечном синтетическом геноме. Так, количество мутаций в условном фрагменте генома размером 5 тыс.п.о. из-за дефектов олигонуклеотидов, очищенных с помощью гель-электрофореза, составит около 10 (исходя из частоты 1/500 нт). Частота ошибок *Taq*-полимеразы лежит, по разным данным, в пределах 10^{-4} – 10^{-5} , а наиболее надёжная из известных полимераз – *Pfu* - имеет частоту ошибок $\sim 10^{-6}$. Если принять чистоту ошибок для коммерческих

смесей полимераз менее 10^{-5} , то количество привнесенных мутаций в процессе сборки фрагмента из олигомеров может быть оценено как не более трех при условии проведения 30 циклов амплификации. Таким образом, дефекты олигонуклеотидов определяют конечное число мутаций, которое в таком случае может быть оценено как порядка 10 на фрагмент размером 5 тыс.п.о. Как следствие, фракция фрагментов без единой мутации не превысит 3×10^{-5} . Найти такую ДНК методами клонирования без специфического отбора не представляется возможным. Следует отметить, что при синтезе генома фага ФХ174 (размер $\sim 5,4$ тыс.п.о.), авторы фактически проводили отбор “жизнеспособных” вирусных геномов, трансфицируя клетки синтетической вирусной ДНК. Согласно их оценкам, доля таких геномов (включая несущие молчащие мутации и мутации, совместимые с размножением фагов) составила 5×10^{-5} [60].

Наиболее рациональным на сегодня представляется подход, при котором размер фрагментов ДНК, собираемых из олигонуклеотидов, составляет от 500 до 1000 п.о. В этом случае количество мутаций, привнесенных как полимеразами, так и дефектными олигонуклеотидами, не превысит уровень, при котором фрагмент с последовательностью, идентичной заданной, может быть легко найден клонированием и последующим секвенированием. Сборка таких фрагментов в сегменты генома может быть осуществлена с помощью последовательных ПЦС за счет того, что они фланкируются уникальными перекрывающимися (“липкими”) линкерами. С помощью таких 30-мерных перекрывающихся линкеров, предусмотренных на концах, 21 рибосомальный ген, синтезированный с использованием ПЦС, был собран с помощью последовательных ПЦС-реакций в оперон размером 14,6 тыс. п.о., ответственный за синтез белков малой субчастицы рибосом *E. coli* [65]. В другом подходе [58], авторы сконструировали плазмиды, каждая из которых несет гены устойчивости к антибиотикам. Сайты рестрикции и комбинация генов устойчивости к антибиотикам предусмотрены таким образом, что после клонирования фрагментов в данные векторы, объединение фрагментов возможно через традиционные рекомбинантные методы, причем правильное соединение фрагментов приводит к уникальной комбинации устойчивости к антибиотикам, позволяя проводить эффективный отбор. Пошаговая сборка позволила получить крупные фрагменты днДНК размером около 5 тыс. п.о. [58].

ЗАКЛЮЧЕНИЕ. Принимая во внимание уровень сегодняшнего развития молекулярно-биологических методов, технологий химического синтеза и очистки олигонуклеотидов, методов биоинформатики и компьютерного анализа, синтез минимального генома саморазвивающейся живой клетки выглядит как весьма вероятное событие ближайших пяти лет. Это даст толчок дальнейшему развитию синтетической биологии и особенно исследованиям, связанным с созданием искусственной клетки. Прогресс в изучении протеомов и функциональной геномики бактерий, несомненно, будет определяющим для формирования технологических подходов к решению задач создания искусственной клетки. В отличие от синтеза саморазвивающегося генома сложность задачи создания синтетической клетки делает весьма затруднительной, если вообще возможной, оценку сегодня временных рамок её решения. Однако, очевидно, что исследования в области создания искусственной клетки будут, тем не менее, продолжаться с увеличивающейся интенсивностью в ближайшее десятилетие, так как перспективы, которые открывает создание контролируемых и саморазмножающихся нанобиореакторов для индустрии, энергетики и медицины, весьма заманчивы.

ЛИТЕРАТУРА

1. Koonin E.V. (2000) *Annu. Rev. Genomics Hum. Genet.*, **1**, 99-116.
2. Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M., Fritchman R.D., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J., Nguyen D., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M., Tomb J.F., Dougherty B.A., Bott K.F., Hu P.C., Lucier T.S., Peterson S.N., Smith H.O., Hutchison C.A. 3rd, Venter J.C. (1995) *Science*, **270**, 397-403.
3. Gil R., Sabater-Munoz B., Latorre A., Silva F.J., Moya A. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 4454-4458.
4. Weiner A.M., Maizels N. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 7383-7387.
5. Luisi P.L., Ferri F., Stano S. (2006) *Naturwissenschaften*, **93**, 1-13.
6. Frick D.N., Richardson C.C. (2001) *Annu. Rev. Biochem.*, **70**, 39-80.
7. Lazcano A., Guerrero R., Margulis L., Oro J. (1988) *J. Mol. Evol.*, **27**, 283-290.
8. Lazcano A., Valverde V., Hernandez G., Gariglio P., Fox G.E., Oro J. (1992) *J. Mol. Evol.*, **35**, 524-536.
9. Suttle D.P., Ravel J.M. (1974) *Biochem. Biophys. Res. Commun.*, **57**, 386-393.
10. Luisi P.L., Oberholzer T., Lazcano A. (2002) *Helv. Chim. Acta*, **85**, 1759-1777.
11. Mushegian A.R., Koonin E.V. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 10268-10273.
12. Mushegian A.R. (1999) *Curr. Opin. Gen. Dev.*, **9**, 709-714.
13. Koonin E.V. (2003) *Nat. Rev. Microbiol.*, **1**, 127-136.
14. Gil R., Silva F.J., Pereto J., Moya A. (2004) *Microbiol. Mol. Biol. Rev.*, **68**, 518-537.
15. Islas S., Becerra A., Luisi P.L., Lazcano A. (2004) *Orig. Life Evol. Biosph.*, **34**, 243-256.
16. Peterson S.N., Fraser C.M. (2001) *Genome Biol.*, **2**, 1-8.
17. Itaya M. (1995) *FEBS Lett.*, **362**, 257-260.
18. Hutchison C.A., Peterson S.N., Gill S.R., Cline R.T., White O., Fraser C.M., Smith H.O., Venter J.C. (1999) *Science*, **286**, 2165-2169.
19. Morowitz H.J. (1967) *Prog. Theor. Biol.*, **1**, 35-58.
20. Dyson F.J. (1982) *J. Mol. Evol.*, **18**, 344-350.
21. Walde P., Wick R., Fresta M., Mangone A., Luisi P.L. (1994) *J. Am. Chem. Soc.*, **116**, 11649-11654.
22. Jay D., Gilbert W. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1978-1980.
23. Luisi P.L., Oberholzer T., Lazcano A. (2002) *Helv. Chim. Acta*, **85**, 1759-1777.
24. Oberholzer T., Luisi P.L. (2002) *J. Biol. Phys.*, **28**, 733-744.
25. Pohorille A., Deamer D. (2002) *Trends Biotechnol.*, **3**, 123-128.
26. Szostak J.W., Bartel D.P., Luisi P.L. (2001) *Nature*, **409**, 387-390.
27. Noireaux V., Bar-Ziv R., Godefroy J., Salman H., Libchaber A. (2005) *Phys. Biol.*, **2**, 1-8.
28. Fielding R.M., Lasic D.D. (1999) *Expert. Opin. Therapeut. Patents*, **9**, 1679-1688.
29. Zignani M., Drummond D.C., Meyer O., Hong K., Leroux J.C. (2000) *Biochim. Biophys. Acta*, **1463**, 383-394.
30. Chang T.M.S. (1999) *Trends Biotechnol.*, **17**, 61-67.
31. Rudolph A.S., Philips W.T. (1998) in: *Blood Substitutes: Principles, Methods, Products and Clinical Trials (Vol. II)* (T.M.S. Chang, ed.) Karger, Basel, Switzerland, pp. 197-205.
32. Bangham A.D., Standish M.M., Watkins J.C. (1965) *J. Mol. Biol.*, **13**, 238-252.
33. Deamer D., Bangham A.D. (1976) *Biochim. Biophys. Acta*, **443**, 629-634.
34. Hope M.J., Bally M.B., Webb G., Cullis P.R. (1985) *Biochim. Biophys. Acta*, **812**, 55-65.
35. Paula S., Volkov A.G., Van Hoek A.N., Haines T.H., Deamer D.W. (1996) *Biophys. J.*, **70**, 339-348.
36. Monnard P.A., Deamer D.W. (2001) *Orig. Life Evol. Biosphere*, **31**, 147-155.

37. Chakrabarti A.C., Breaker R.R., Joyce G.F., Deamer D.W. (1994) *J. Mol. Evol.*, **39**, 555-559.
38. Walde P., Goto A., Monnard P.A., Wessicken M., Luisi P.L. (1994) *J. Am. Chem. Soc.*, **116**, 7541-7547.
39. Walde P., Wick R., Fresta M., Mangone A., Luisi P.L. (1994) *J. Am. Chem. Soc.*, **116**, 11649-11654.
40. Oberholzer T., Albrizio M., Luisi P.L. (1995) *Chem. Biol.*, **2**, 677-682.
41. Oberholzer T., Wick R., Luisi P.L., Biebricher C.K. (1995) *Biochem. Biophys. Res. Commun.*, **207**, 250-257.
42. Oberholzer T., Nierhaus K.H., Luisi P.L. (1999) *Biochem. Biophys. Res. Commun.*, **261**, 238-241.
43. Tsumoto K., Nomura S.M., Nakatani Y., Yoshikawa K. (2001) *Langmuir*, **17**, 7225-7228.
44. Fischer A., Franco A., Oberholzer T. (2002) *Chembiochem*, **3**, 409-417.
45. Yu W., Sato K., Wakabayashi M., Nakaishi T., Ko-Mitamura E.P., Shima Y., Urabe I., Yomo T. (2001) *J. Biosci. Bioeng.*, **92**, 590-593.
46. Nomura S.M., Tsumoto K., Hamada T., Akiyoshi K., Nakatani Y., Yoshikawa K. (2003) *Chembiochem*, **4**, 1172-1175.
47. Noireaux V., Bar-Ziv R., Libchaber A. (2003) *Proc. Natl. Acad. Sci. USA*, **100**, 12672-12677.
48. Ishikawa K., Sato K., Shima Y., Urabe I., Yomo T. (2004) *FEBS Lett.*, **576**, 387-390.
49. Khorana H.G., Agarwal K.L., Buchi H., Caruthers M.H., Gupta N.K., Kleppe K., Kumar A., Otsuka E., RajBhandary U.L., Van de Sande J.H., Sgaramella V., Terao T., Weber H., Yamada T. (1972) *J. Mol. Biol.*, **72**, 209-217.
50. Khorana H.G. (1979) *Science*, **203**, 614-625.
51. Engels J., Uhlmann E. (1988) *Adv. Biochem. Eng. Biotechnol.*, **37**, 73-127.
52. McBride L.J., Caruthers M.H. (1983) *Tetrahedron Lett.*, **24**, 245-248.
53. Shabarova Z.A., Merenkova I.N., Oretskaya T.S., Sokolova N.I., Skripkin E.A., Alexeyeva E.V., Balakin A.G., Bogdanov A.A. (1991) *Nucleic Acids Res.*, **19**, 4247-4251.
54. Hayden M.A., Mandecki W. (1988) *DNA*, **8**, 571-577.
55. Dillon P.J., Rosen C.A. (1990) *Biotechniques*, **298**, 300.
56. Mandecki W., Hayden M.A., Shallcross M.A., Stotland E. (1990) *Gene*, **94**, 103-107.
57. Stemmer W.P., Crameri A., Ha K.D., Brennan T.M., Heyneker H.L. (1995) *Gene*, **164**, 49-53.
58. Kodumal S.J., Patel K.G., Reid R., Menzella H.G., Welch M., Santi D.V. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 15573-15578.
59. Cello J., Paul A.V., Wimmer E. (2002) *Science*, **297**, 1016-1018.
60. Smith H.O., Hutchison C.A. 3rd, Pfannkoch C., Venter J.C. (2003) *Proc. Natl. Acad. Sci. USA*, **100**, 15440-15445.
61. Gustafsson C., Govindarajan S., Minshull J. (2004) *Trends Biotechnol.*, **22**, 346-353.
62. Hoover D.M., Lubkowski J. (2002) *Nucleic Acids Res.*, **30**, e43.
63. Rouillard J.M., Lee W., Truan G., Gao X., Zhou X., Gulari E. (2004) *Nucleic Acids Res.*, **32** (Web Server issue), W176-180.
64. Tian J., Gong H., Sheng N., Zhou X., Gulari E., Gao X., Church G. (2004) *Nature*, **432**, 1050-1054.
65. Lashkari D.A., Hunicke-Smith S.P., Norgren R.M., Davis R.W., Brennan T. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 7912-7915.
66. Zhou X., Cai S., Hong A., You Q., Yu P., Sheng N., Srivannavit O., Muranjan S., Rouillard J.M., Xia Y., Zhang X., Xiang Q., Ganesh R., Zhu Q., Matejko A., Gulari E., Gao X. (2004) *Nucleic Acids Res.*, **32**, 5409-5417.
67. Stemmer W.P. (1994) *Nature*, **370**: 389-391.

Поступила: 20. 10. 2006.

SYNTHESIS OF ARTIFICIAL GENOME AS THE BASIS OF SYNTHETIC BIOLOGY

S.P. Rad'ko, A.P. Il'ina, N.V. Bodoev, A.I. Archakov

Orehovich Institute of Biomedical chemistry RAMS, Pogodinskaya ul., 10, Moscow, 119121 Russia;
tel.: (495)246-16-41; fax: (495)2463771; e-mail: radko@ibmc.msk.ru

Recent achievements in the whole-genome sequencing especially viral and bacterial ones, together with the development of methods of bioinformatics and molecular biology, have created preconditions for transition from synthesis of genes to assembly of the whole genomes from chemically synthesized blocks, oligonucleotides. The creation of artificial genomes and artificial cells, will undoubtedly render huge influence on a deepening of knowledge of mechanisms of functioning of living systems at a cellular level, on a way of origin and evolution of life, and also on biotechnology of the future, and will generate preconditions for the further development of synthetic biology and nanobiotechnology.

Key words: synthensis of genome, synthetic biology, nanobiotechnology.