

ФУНКЦИОНАЛЬНАЯ ГЕНОМИКА ЦИАНОБАКТЕРИЙ

С.В. Шестаков

Московский государственный университет им. М.В. Ломоносова
e-mail: shestakov@1.gen.bio.msu.ru

В задачи функциональной геномики входят идентификация функции генов и генетических элементов, регулирующих экспрессию генов, расшифровка механизмов координации работы генов в клетке. Основными методами функциональной геномики являются: 1) предсказание функции гена на основе изучения гомологии нуклеотидных и аминокислотных последовательностей по базам данных об идентифицированных генах других организмов; 2) ген-направленный мутагенез: введение в синтезированный и клонированный ген (или его фрагмент) инсерции/делеции или сайт-специфической мутации с последующим замещением гена дикого типа мутантным аллелем (метод так называемой "обратной генетики"); 3) изучение экспрессии гена с помощью ДНК-РНК гибридизации, репортерных систем (например, *lac*- или *luxAB* генов), бессмысловых РНК-конструкций, иммунохимического анализа, методов определения белок-белковых взаимодействий; 4) анализ промоторных, энхансерных и других регуляторных элементов, определяющих особенности транскрипции анализируемых генов, их взаимодействие с другими генами; 5) исследование координированной регуляции экспрессии генов в полногеномном масштабе с применением *microarray*-технологии (микрочипов с адресным расположением генных зондов). Эта технология позволяет установить, какие гены в целом геноме экспрессируются, включаются или выключаются в различных условиях (в частности, при стрессовых воздействиях) при мутационном блокировании функции других генов. С помощью этого метода можно выявить участие новых генов в том или ином биологическом процессе, т.е. идентифицировать ген, о клеточной функции которого не было сведений.

На основе использования методов функциональной геномики в последние годы достигнут существенный прогресс в понимании структурно-функциональной организации ге-

номов многих организмов, в том числе цианобактерий, которые являются модельными объектами изучения окислительного фотосинтеза и азотфиксации, биогенеза мембран и клеточной дифференцировки, сигнальных систем и механизмов адаптации, различных процессов, присущих фотосинтезирующим клеткам.

В таблице 1 приведены сведения о сиквенированных к 2004 г. геномах цианобактерий. На кафедре генетики МГУ были разработаны системы генетической трансформации у *Synechococcus* 7942 (Shestakov, Nguen Than Khyen, 1970) и *Synechocystis* 6803 (Grigorieva, Shestakov, 1982), которые сейчас широко используются во многих лабораториях мира. В 1966 г. японские авторы (Kaneko *et al.*, 1996) определили полную нуклеотидную последовательность генома *Synechocystis* 6803 – это был первый сиквенированный геном фотосинтезирующего организма.

Исследования в области молекулярной генетики *Synechocystis* 6803 внесли решающий вклад в расшифровку систем окислительного фотосинтеза (Shestakov, 2002). Дефектные по фотосинтезу мутанты этой фотогетеротрофной цианобактерии способны расти в темноте на среде с глюкозой. К ней применимы все основные методы бактериальной генетики и геномики, что позволяет изучать функции любых генов. В геноме *Synechocystis* 6803 85 % приходится на последовательности, кодирующие белки; 30 % из них гомологичны генам с известными функциями, 15 % имеют сходство с функционально идентифицированными генами и 10 % с гипотетическими генами других организмов, для 45 % генов нет аннотированных гомологов в геномах других объектов (Kotani, Tabata, 1998). Таким образом, главной задачей геномики этой цианобактерии является определение клеточных функций каждого гена с учетом того, что многие из них входят в большие семейства (например, около 50 генов кодируют разные пептидазы, более 100 – белки регуляторы транскрипции и т. д.)

Таблица 1

Геномы цианобактерий

Цианобактерия, вид, штамм	Биологические особенности	ГЦ % состав	Размер генома, Мб	Число ORF ¹
<i>Synechococcus</i> sp. PCC ² 6301 ³	одноклеточная, облигатно фотоавтотрофная	55,1	2,69	2842
<i>Synechococcus</i> sp. PCC 7002	короткие филаменты, фотогетеротрофная	49,1	3,12	2910
<i>Synechocystis</i> sp. PCC 6803	одноклеточная, фотогетеротрофная	47,7	3,57	3168
<i>Gloeobacter violaceus</i> PCC 7421	отсутствуют тилакоиды	62,1	4,65	4430
<i>Anabaena</i> sp. PCC 7120	азотфиксирующая, филаментозная	42,5	6,42	5368
<i>Thermosynechococcus elongatus</i> BP1	одноклеточная, термофильная		2,59	2475
<i>Prochlorococcus marinus</i> S120	морская пикоцианобактерия	36	1,75	1884

Примечание. 1 – ORF – открытые рамки считывания; 2 – PCC – номер в коллекции Пастеровского института, Париж; 3 – штамм, эквивалентный PCC 7942 (*Anacystis nidulans* R2).

В 1998–2003 гг. опубликовано более 700 работ, посвященных изучению функции генов *Synechocystis* 6803 с помощью методологии "обратной генетики" (рис.). Следует отметить, что при замещении в результате рекомбинации "нормального" гена в штамме дикого типа геном, инактивированным инсерцией (нокаут-мутация), далеко не во всех случаях можно получить полностью сегрегированный мутант с видимым фенотипическим проявлением.

Клетки *Synechocystis* 6803 содержат 8–12 идентичных хромосомных копий; если полной сегрегации мутации (т.е. замещения во всех копиях) достичь не удастся, то это означает, что ген отвечает за жизненно необходимую функцию, без которой клетки не выживают. Для изучения таких генов можно использовать генно-инженерные конструкции, обеспечивающие регулируемую экспрессию дополнительной копии анализируемого гена или регуляцию экспрессии с помощью антисмысловой РНК. Если же получен полностью сегрегированный мутант (что проверяется Саузерн-гибридизацией или ПЦР-методом) без видимых фенотипических изменений, то это может означать, что: 1) открытая рамка считывания (ORF) не экспрессируется ни при каких условиях (псевдоген, "реликтовая" последовательность и т. п.); 2) ген экспрессируется только при определенных условиях, например, при адаптивном ответе на стрессовые воздействия; 3) функция инактивированного гена компенсируется другим геном (например, кросс-взаимодействие протеинкиназ и регуляторных белков в двухкомпонентных системах сигнальной трансдукции); 4) блокирование функции гена не проявляется на уровне видимых морфо-физиологических характеристик; 5) возможно сохранение фенотипа в результате супрессорной мутации, локализованной в ином гене. Таким

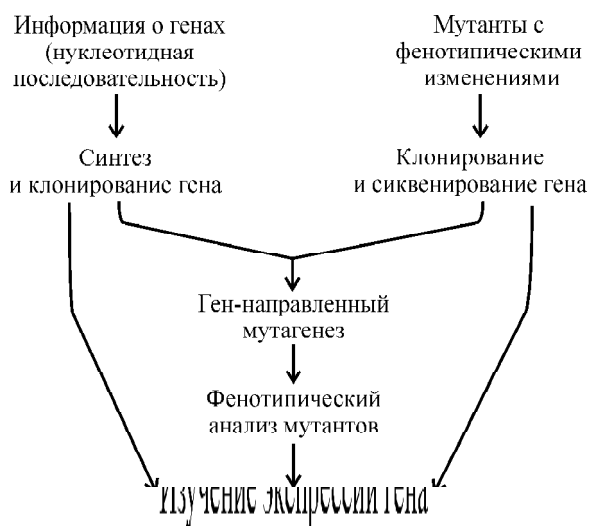


Рис. Функциональная геномика.

Левая часть схемы – методология «обратной генетики».

Таблица 2

Функциональный анализ некоторых генов *Synechocystis* 6803

Семейства генов, кодирующих	Количество генов, клеточные функции		Гены, функции которых исследуются в МГУ		
	идентифицированы	не установлены	инсерционный мутагенез	сегрегированные мутанты	определен фенотип
пептидазы	8	36	27	17	6
серин/треониновые протеинкиназы	2	14	12	11	2
response-белки 2-компонентных систем регуляции	6	33	24	22	2
Ycf- белки, гомологи белков хлоропластов	6	26	3	2	2

образом, для идентификации функции инактивируемого гена в большинстве случаев требуется детальный молекулярно-генетический и физиолого-биохимический анализ.

В таблице 2 представлена информация о некоторых направлениях исследований в МГУ по функциональной геномике *Synechocystis* 6803. Для 10 из 27 инактивированных генов, кодирующих пептидазы различного типа (Sokolenko *et al.*, 2002), не удалось получить полностью сегрегированных мутантов, что указывает на жизненно важные функции этих генов, к числу которых относятся 3 гена *clpP* семейства, а также гены *clpC*, *clpX*, *clpB1*, *htrA*, *ctpC*, *lepB2*, *ftsH1*. Изучение "гомозиготных" мутантов с фенотипическими изменениями позволило впервые установить конкретные физиологические функции ряда пептидаз. Мутанты по генам *ctpA*, *clpP2*, *lepB1* не способны к фотоавтотрофному росту. Ген *ctpA* кодирует С-терминальную пептидазу, осуществляющую процессинг белка D1 фотосистемы II (Shestakov *et al.*, 1994), пептидаза ClpP2 участвует в защите клеток от фотоинактивации (Паничкин и др., 2001), лидерная пептидаза LepB1 необходима для транслокации белков в тилакоидной мем-бране, белок ClpB2 с функцией шаперона вовлечен в регуляцию терморезистентности, пептидазы HhoA и HhoB являются белками теплового шока (Sokolenko *et al.*, 2002), пептидаза SppA1 отвечает за протеолиз линкерных белков фикобилисом, обес-

печивая адаптивную редукцию фотосинтетической антенны при высокой интенсивности света.

Инсерционная инактивация большинства генов серин/треониновых протеинкиназ эукариотического типа не привела к видимым фенотипическим изменениям (Галкин и др., 2003). Нокаут-мутация гена *spkA* блокирует подвижность клеток, у мутанта по гену *pknE* повышается резистентность к метиламину. Изучение экспрессии протеинкиназных генов с помощью мiсroаgау-анализа в различных стрессовых условиях может дать важную информацию о клеточных функциях этих генов. На основе такого подхода изучаются функции белков двухкомпонентных систем регуляции. В лаборатории Н. Мураты (Япония) осуществлен интерпозиновый мутагенез всех 40 *hik*-генов, кодирующих гистидиновые протеинкиназы, а в лабораториях В. Зинченко (МГУ) и Д. Лося (Институт физиологии растений РАН) получены инсерционные мутанты практически по всем *rre*-генам, кодирующим response-белки регуляторы двухкомпонентных систем. Наличие мутантов дает возможность идентифицировать взаимодействующие пары сенсорной гистидинкиназы и белка-регулятора и выяснить физиологические функции двухкомпонентных систем регуляции в адаптивных ответах клеток. Таким способом была расшифрована роль системы сигнальной трансдукции с участием гистидиновых протеинкиназ Hik33 и Hik19 в адаптации к холодному воздействию. Инактивация гена *hik33* влияет на регуляцию транскрипции более 60 ге-

нов (Suzuki *et al.*, 2001). При этом выявлено более 30 генов, контролируемых протеинкиназой Hik33, о клеточных функциях которых вообще не было информации.

Выявлены двухкомпонентные системы, ответственные за адаптивные ответы клеток *Synechocystis* 6803 при действии факторов осмотического солевого, светового, теплового или окислительного стресса. Наглядным примером исследования двухкомпонентных систем с использованием мутантов и микроаггау-анализа является работа по расшифровке механизма регуляции транспорта ионов марганца (Yamaguchi *et al.*, 2002). При высокой концентрации ионы марганца токсичны, при дефиците Mn^{2+} в клетках не собирается комплекс фотосистемы II. Поэтому внутриклеточная концентрация Mn^{2+} поддерживается на оптимальном уровне за счет регуляции экспрессии генов *mntABC*-оперона, кодирующих белки транспортной системы. По данным микроаггау-анализа в клетках штамма дикого типа при нехватке ионов марганца усиливается транскрипция *mnt*-генов, при избытке Mn^{2+} работа этих генов репрессируется. Регуляция экспрессии *mntABC*-генов осуществляется двухкомпонентной системой, включающей сенсорную гистидинпротеинкиназу MntS и регуляторный белок MntR. В клетках инсерционных мутантов по этим генам регуляция транскрипции *mntABC* оперона нарушается. При достаточной концентрации ионов марганца происходит их связывание с белком MntS, который активируется и переводит белок MntR в форму активного репрессора, блокирующего экспрессию *mntABC*-оперона. При дефиците ионов марганца происходит дерепрессия этого оперона, так как MntS находится в неактивной форме и не может инициировать образование репрессора.

В таблицу 2 включены также три гена, имеющие высокую степень гомологии с генами хлоропластной ДНК высших растений. Ранее был разработан метод функционального анализа таких *ycf*-генов растений на основе инсерционного мутагенеза гомологичных генов *Synechocystis* 6803 (Wilde *et al.*, 1995). При использовании этого подхода нам в сотрудничестве с Х. Пакраси (США) удалось установить, что белок Ycf4 участвует в защите фото-

системы I от фотоинактивации, а белок Ycf3 необходим для сборки и стабилизации реакционного центра этой фотосистемы. Анализ "нокаут-мутанта" по гену *psbJ* позволил доказать участие продукта этого гена в регуляции потока электронов в комплексе фотосистемы II (Regel *et al.*, 2001). На основе изучения мутантов с инактивированными *gap*-генами (кодирующими глицеральдегид-3-фосфат дегидрогеназу) было установлено, что ген *gap1* отвечает за гликолитическое расщепление глюкозы, а фермент Gap2, подобный белку хлоропластов, участвует в фотосинтетическом цикле (Koksharova *et al.*, 1998). Эта работа демонстрирует высокую эффективность метода "обратной генетики" в изучении физиологических функций дивергентных генов, относящихся к одному семейству.

Детально исследован открытый на кафедре генетики МГУ новый ген, *prqR*, кодирующий белок регулятор транскрипции генов антиоксидантных ферментов и антипортерных систем, участвующих в защите клеток от действия параквата (Бабыкин и др., 2003). Идентифицированы домены белка PrqR, ответственные за репрессорные функции, и расшифрован механизм авторегуляции этого гена, играющего важную роль в адаптивном ответе клеток цианобактерий при окислительном стрессе. В геноме *Synechocystis* 6803 обнаружено много генов, предположительно кодирующих белки регуляторы транскрипции. Нами совместно с другими лабораториями ведется работа по инактивации каждого из этих генов для последующего анализа генных сетей с помощью микроаггау-технологии.

Особое внимание, естественно, уделяется полногеномному изучению глобальной координации работы генов при изменениях светового режима (в зависимости от спектра и интенсивности освещения). Исследование дифференциальной экспрессии генов показало, что в процессе световой адаптации в клетках цианобактерии происходят глубокие перестройки метаболизма, связанные с регуляцией активности большого числа генов (Hihara *et al.*, 2001; Huang *et al.*, 2002). В фоторецепции участвуют различные фитохромы и светособирающие комплексы, которые передают сигналы через каскады регуляторных систем

Таблица 3

Регуляция экспрессии генов *Synechocystis* 6803 при адаптации клеток к свету высокой интенсивности (по данным microarray-анализа (Hihara *et al.*, 2001; Huang *et al.*, 2002)

Гены	Количество генов (экспрессия)			
	активация		подавление	
	15 ¹	120 ²	15 ¹	120 ²
с известной функцией,	> 30	> 30	> 40	> 20
включая ³	rbc, ccm ndh, sodB, glp clpB, hsp	rpl, rps hem, chl aps, psb	psb, psa hem, chl apc, cpc	ndh clpB, hsp
с неизвестной функцией	18	8	28	12

Примечание. 1 – в первые 15 мин; 2 – после 60–180 мин инкубации; 3 – гены, контролируемые: rbc, ccm – CO₂ метаболизм; ndh – дыхание; sodB, glp – антиоксидантные ферменты; hsp, clpB – шапероны; rpl, rps – белки рибосом; psa, psb – белки фотосистем I и II; hem, chl – ферменты биосинтеза хлорофилла; apc, cpc – ферменты биосинтеза фикобилинов.

до уровня транскрипции генов, контролирующего циркадный ритм, фототаксис, синтез белков аппарата фотосинтеза, антиоксидантных ферментов и других светозащитных и репарационных систем (Mullineaux, 2001). В таблице 3 суммированы обобщенные сведения о динамике изменений в профилях экспрессии генов при адаптации клеток *Synechocystis* 6803 к высокой интенсивности света (200–300 мЕ/м²сек⁻¹). По результатам microarray-анализа можно выделить две основные фазы адаптационных сдвигов. В первой кратковременной фазе (15 мин) наблюдается репрессия генов, ответственных за синтез компонентов фотосистем II и I (гены *psb* и *psa*); происходит редукция светособирающих антенн, что приводит к снижению поглощения энергии и, соответственно, образования токсичных радикалов, повреждающих клеточные структуры. Одновременно индуцируется синтез большой группы белков, ответственных за углеродный метаболизм и дыхание, увеличивается экспрессия генов, кодирующих антиоксидантные ферменты, синтез каротиноидов, белков теплового шока, некоторых шаперонов и других защитных систем. В начале второй фазы (> 60 мин) уровень синтеза ряда защитных белков (ClpB, Hsp) снижается, но повышается экспрессия генов рибосомного аппарата (гены *rps*, *rpl* и ряд других). Затем постепенно начинается активизация генов фотосинтеза, увеличивается соотношение комплексов фотосисте-

мы II/фотосистемы I. Среди генов, индуцируемых в первой фазе адаптации, выявлено более 20 генов, клеточные функции которых еще не известны. Не идентифицированы функции 30 других генов, экспрессия которых подавляется в условиях высокой освещенности. Таким образом, предстоит решить два важных вопроса: 1) какую физиологическую роль играют эти неидентифицированные гены в жизнедеятельности цианобактерий; 2) как осуществляется согласованная регуляция экспрессии генов в интегральной геномной сети при световой адаптации. Ответы на эти вопросы позволят понять сущность механизмов глобальной координации клеточных процессов, специфичных для фотосинтезирующих организмов. Высокие темпы исследований в области функциональной геномики *Synechocystis* 6803 позволяют надеяться, что клеточные функции практически всех генов этой цианобактерии будут установлены в ближайшие годы.

Прогресс в разработке проблем геномики цианобактерий открывает новые возможности и для обсуждения актуальных вопросов эволюционной биологии. В субгеномах хлоропластов растений выявлено около 80 генов, обладающих высокой степенью гомологии с генами *Synechocystis* 6803, что является еще одним убедительным свидетельством в пользу эндосимбиотического происхождения хлоропластов. Однако многие гены цианобактерии имеют большую гомологию

и с ядерными генами растений. Анализ 386 филогенетических деревьев показал, что почти 9 % ядерных генов *Arabidopsis thaliana* могут иметь цианобактериальное происхождение (Rujan, Martin, 2001). Идентификация клеточных функций генов цианобактерий позволит внести вклад в геномику растений, понимание закономерностей ядерно-цитоплазматических взаимодействий. Изучение этих проблем может пролить свет на сущность процессов редуccionной эволюции субгеномов органелл и природу генетических механизмов перемещения генов из хлоропластов в ядро на ранних этапах биологической эволюции.

Литература

- Бабыкин М.М., Сидорук К.В., Зинченко В.В. и др. Об участии регуляторного гена *prqR* в развитии устойчивости к метилвиологену у цианобактерии *Synechocystis* sp. PCC 6803 // Генетика. 2003. Т. 39. С. 25–32.
- Галкин А.Н., Михеева Л.Е., Шестаков С.В. Инсерционная инактивация генов, кодирующих серин/треониновые протеинкиназы эукариотического типа у цианобактерии *Synechocystis* sp. PCC 6803 // Микробиология. 2003. Т. 72. С. 64–69.
- Паничкин В.Б., Глазер В.М., Зинченко В.В. и др. Ген *clp2*, кодирующий пептидазу у цианобактерии *Synechocystis* sp. PCC 6803, контролирует чувствительность клеток к фотингибированию // Изв. АН. Сер. биол. 2001. № 3. С. 312–317.
- Grigorieva G.A., Shestakov S.V. Transformation in the cyanobacterium *Synechocystis* sp. 6803 // FEMS Microbiol. Lett. 1982. V. 13. P. 367–370.
- Hihara Y., Kamei A., Kanehisa M. *et al.* DNA microarray analysis of cyanobacterial gene expression during acclimation to high light // Plant Cell. 2001. V. 13. P. 393–806.
- Huang L., McCluskey M.P., Ni H., LaRossa R.A. Global gene expression profiles to the cyanobacterium *Synechocystis* strain sp. PCC 6803 in response to irradiation with UV-B and white light // J. Bacteriol. 2002. V. 184. P. 6845–6858.
- Kaneko T., Sato S., Kotani H. *et al.* Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803 // DNA Res. 1996. V. 3. P. 109–136.
- Koksharova O.A., Shubert M., Shestakov S.V., Cerff R. Genetic and biochemical evidence for distinct key functions of two highly divergent GAPDH genes in catabolic and anabolic carbon flow of the cyanobacterium *Synechocystis* sp. PCC 6803 // Plant Molec. Biol. 1998. V. 36. P. 183–184.
- Kotani T., Tabata S. Lessons from sequencing of the genome of a unicellular cyanobacterium *Synechocystis* sp. PCC 6803 // Annu. Rev. Plant Physiol. Plant Molec. Biol. 1998. V. 49. P. 151–171.
- Mullineaux C.W. How do cyanobacteria sense and respond to light? // Mol. Microbiol. 2001. V. 41 (5). P. 965–971.
- Regel R.E., Ivleva N.B., Zer H. *et al.* Dereglulation of electron flow within photosystem II in the absence of the PsbJ protein // J. Biol. Chem. 2001. V. 276. P. 41473–41478.
- Rujan T., Martin W. How many genes in *Arabidopsis* come from cyanobacteria? An estimate from 386 protein phylogeny // Trends Genet. 2001. V. 17, N 3. P. 113–120.
- Shestakov S.V. Gene-targeted and site-directed mutagenesis of photosynthesis genes in cyanobacteria // Photosynth. Res. 2002. V. 73. P. 279–284.
- Shestakov S.V., Anbudurai P.R., Stanbekova G.E. *et al.* Molecular cloning and characterization of the *ctpA* gene encoding a carboxyl-terminal processing protease // J. Biol. Chem. 1994. V. 269. P. 19354–19359.
- Shestakov S.V., Nguen Than Khyen. Evidence for genetic transformation in blue-green alga *Anacystis nidulans* // Mol. Gen. Genetics. 1970. V. 107. P. 372–375.
- Sokolenko A., Pojidaeva E., Zinchenko V. *et al.* The gene complement for proteolysis in the cyanobacterium *Synechocystis* sp. PCC 6803 and *Arabidopsis thaliana* chloroplasts // Current Genet. 2002. V. 41. P. 291–310.
- Suzuki I., Kanesaki Y., Mikami K. *et al.* Cold-regulated genes under control of the cold sensor Hik33 in *Synechocystis* // Mol. Microbiol. 2001. V. 40 (1). P. 235–244.
- Wilde A., Hartel H., Hubschmann T. *et al.* Inactivation of a *Synechocystis* sp. strain PCC 6803 gene with homology to conserved chloroplast open reading frame 184 increases the photosystem II to photosystem I ratio // Plant Cell. 1995. V. 7. P. 649–658.
- Yamaguchi K., Suzuki I., Yamamoto H. *et al.* A two-component Mn²⁺-sensing system negatively regulates expression of the *mmtCAB* operon in *Synechocystis* // Plant Cell. 2002. V. 14. P. 2901–2913.