

УДК 575.83

## ЭВОЛЮЦИЯ ЗАПАСНЫХ ГЛОБУЛИНОВ СЕМЯН И СУПЕРСЕМЕЙСТВО КУПИНОВ

© 2011 г. А. Д. Шутов\*, И. А. Каховская

*Лаборатория биохимии растений Государственного университета Молдовы, MD-2009, Кишинев, Молдова*

Поступила в редакцию 25.10.2010 г.

Принята к печати 08.11.2010 г.

Обширное суперсемейство купинов (клан cl09118) в настоящее время объединяет тысячи функционально и структурно разнородных белков прокариот и эукариот, содержащих  $\beta$ -баррель, образованный антипараллельными  $\beta$ -стрендами (купиновый модуль). Возможные пути формирования суперсемейства купинов описаны на основе сопоставления первичных и третичных структур белков из ряда консервативных семейств купинов, в том числе, запасных глобулинов семян и оксалатоксидаз растений (джерминов), бактериальных оксалатдекарбоксилаз, гентизатдиоксигеназ и эпимераз. Получены свидетельства о происхождении двудоменных структур запасных глобулинов семян из двудоменных оксалатдекарбоксилаз цианобактерий. Прослежен эволюционный путь однодоменных джерминов, ранее считавшихся прямыми предшественниками запасных глобулинов. Общими эволюционными корнями джерминов и оксалатдекарбоксилаз являются ныне существующие белки бактерий и археобактерий, примитивная структура которых целиком состоит из купинового модуля. Эти корневые белки отражают гипотетическую структуру прокупина, давшего начало, по крайней мере, части из общего многообразия членов суперсемейства купинов (например, купинового модулю гентизатдиоксигеназ). Главная проблема описания суперсемейства купинов – разграничение между дивергентной эволюцией и конвергенцией. Примером структурной конвергенции может быть формирование  $\beta$ -барреля в структурах чрезвычайно консервативного семейства эпимераз археобактерий и бактерий.

**Ключевые слова:** запасные глобулины семян, джермины, оксалатдекарбоксилазы, суперсемейство купинов, молекулярная эволюция.

EVOLUTION OF SEED STORAGE GLOBULINS AND CUPIN SUPERFAMILY, by A. D. Shutov, I. A. Kakhovskaya (Laboratory of Plant Biochemistry, State University of Moldova, MD-2009, Chişinău, Moldova; \*e-mail: shutovandrei@yahoo.com). An extensive superfamily of cupins (clan cl09118) currently combines thousands of functionally and structurally diverse prokaryote and eukaryote proteins, which contain a  $\beta$ -barrel of antiparallel  $\beta$ -strands (cupin module). Possible ways of the formation of the cupin superfamily were suggested based on the comparison of primary and tertiary structures of proteins from several conserved families of cupins including seed storage globulins and plant oxalate oxidases (germins), and bacterial oxalate decarboxylases, gentisate dioxygenases and epimerases. The origin of two-domain structure of seed storage globulins from cyanobacterial two-domain oxalate decarboxylases has been deduced. The evolutionary pathway of single-domain germins previously suggested to be immediate progenitors of storage globulins was traced back. Common evolutionary roots of germins and oxalate decarboxylases descend from recent bacterial and archaeobacterial proteins whose primitive structure is restricted to the cupin module. These root proteins reflect the hypothetical structure of a pro-cupin that probably gave rise to at least a part of the total diversity of members of the cupin superfamily (for instance, to the cupin module of gentisate dioxygenases). The major dilemma for the description of the cupin superfamily is distinguishing evolutionary divergence from convergence. The structural convergence can be exemplified by formation of a  $\beta$ -barrel inside extremely conserved structures of the otherwise unrelated epimerases from *Archaea* and bacteria.

**Keywords:** seed storage globulins, germins, oxalate decarboxylases, cupin superfamily, molecular evolution.

Термины легумин и вицилин введены Т.Б. Осборном в конце XIX века для обозначения двух типов запасных глобулинов в семенах четырех бобовых растений. С тех пор стало известно, что оба

запасных глобулина характерны для семян всех сперматофитов (см. [1]). При сопоставлении первичных [2, 3], а затем и третичных [4] структур выяснилось, что легумин и вицилин происходят от об-

Принятые сокращения: ОД – оксалатдекарбоксилаза; ГД – гентизат-1,2-диоксигеназа (гентизатдиоксигеназы); ЭП – дегидро-рамнозо-3,5-эпимераза (эпимеразы).

\* Эл. почта: shutovandrei@yahoo.com

щего эволюционного предшественника. В этой роли может выступать специфичный для спор папоротника вицилиноподобный белок саа91187, аминокислотная последовательность которого несет признаки, общие для этих двух белков [5–7].

Субъединица запасного глобулина содержит два структурно эквивалентных домена, каждый из которых образован  $\beta$ -баррелем (из восьми основных антипараллельных  $\beta$ -стрендов BCDE-FGHI), соединенным с группой из двух или трех  $\alpha$ -спиралей [2, 4]. Аминокислотные последовательности доменов гомологичны и, очевидно, происходят из общего однодоменного предшественника, в ходе эволюции претерпевшего дубликацию [2]. Таким предшественником мог бы быть однодоменный белок джермин (растительная оксалатоксидаза), гомологичный запасным глобулинам [8] (гипотеза I). Эта гипотеза стимулировала дальнейший поиск эволюционных предшественников запасных глобулинов и джермина, в результате которого было обнаружено обширное суперсемейство так называемых купинов [9, 10]. В основе поиска использованы в качестве ключевых два узловых мотива аминокислотных последовательностей в структуре джермина [11], соответствующие двум укороченным половинкам  $\beta$ -барреля:  $Gx_5HxHx_{3,4}Ex_6G$  ( $\beta$ -стренд CDE) и  $Gx_5PxGx_2Hx_3N$  ( $\beta$ -стренд FGH).

С точки зрения эволюции легумина и вицилина, из других купинов наибольший интерес представляет семейство бактериальных оксалатдекарбоксилаз (ОД), структура которых образована двумя доменами, гомологичными доменам запасных глобулинов. Высказано предположение, что бактериальные ОД – прямые и единственные предшественники запасных глобулинов [12] (гипотеза II). Не исключено, что ген, кодирующий ОД цианобактерий, унаследованный растениями в результате симбиогенеза, далее был перенесен в ядро путем эндосимбиоза [13], где впоследствии эволюционировал с образованием ныне существующих запасных глобулинов.

Современные банки данных содержат многие тысячи аминокислотных последовательностей, отнесенных к суперсемейству купинов (клан cl09118). Реальное присутствие предполагаемого для всех купинов  $\beta$ -барреля показано в третичных структурах более сотни белков архебактерий, бактерий и эукариот. Как полагают, из общего эволюционного предшественника произошли, по меньшей мере, 18 подклассов купинов [10]. Однако гомологичность коротких отрезков аминокислотных последовательностей купинов между упомянутыми подклассами далеко не всегда очевидна. Сходство первичных структур в некоторых случаях может иметь конвергентный характер в связи с формированием сходных третичных структур.

В свете двух упомянутых выше альтернативных гипотез происхождения запасных глобулинов семян [8, 12] в настоящем обзоре прослежен эволю-

ционный путь легумина и вицилина и сопоставлены первичные и высшие структуры легумина, джермина и бактериальных ОД. Предпринята попытка оценить вероятность дивергентного и/или конвергентного характера эволюции некоторых других белков суперсемейства купинов.

## МЕТОДЫ И СТРАТЕГИЯ АНАЛИЗА

В работе использованы следующие программы: BLAST (<http://www.ncbi.nih.gov/>) – для поиска гомологичных аминокислотных последовательностей белков, ClustalW2 – для их выравнивания и TREECON [14] – для их эволюционного анализа, а также программа DeepView/Swiss-Pdb Viewer – для анализа третичных структур белков, построения ленточных диаграмм и структурного выравнивания.

Для получения статистически достоверной последовательности формирования кластеров на эволюционном древе запасных глобулинов и некоторых других купинов мы руководствовались следующей стратегией.

1. Анализируемый участок последовательностей должен быть, насколько это возможно, максимально протяженным. Выравнивание последовательностей из предварительного широкого набора купинов (в том числе, с известной третичной структурой) показал, что этот участок может охватывать, как минимум, всю последовательность  $\beta$ -барреля (все восемь  $\beta$ -стрендов).

2. При создании коллекции аминокислотных последовательностей для анализа (наиболее критичный этап) из их широкого предварительного набора должны быть избраны только те, что занимают позицию, наиболее близкую к главному стволу эволюционного древа. Этот выбор может опираться на оценку статистической достоверности формирования кластеров и отдельных ветвей при предварительном эволюционном анализе широкого набора последовательностей. С другой стороны, в пределах каждого из кластеров должно быть представлено все таксономическое разнообразие видов, которым принадлежат анализируемые белки.

3. Последовательности, пригодные к использованию в качестве корней (внешней группы) эволюционного древа (outgroup sequences), должны принадлежать анализируемому семейству, но при этом быть наименее сходными со всеми другими последовательностями.

4. Особую сложность при сопоставительном анализе последовательностей  $\beta$ -баррелей купинов создает участок между  $\beta$ -стрендами E и F, во многих случаях образующий протяженную переменную петлю, лишенную вторичных структур [7, 10]. Для установления границ этой петли мы использовали структурное выравнивание, основанное на наложении  $C^\alpha$  атомов в скелете третичных структур ку-

пинов в их парных сочетаниях. О достоверности выравнивания, использованного для построения эволюционного древа, судили по уровню его совпадения со структурным выравниванием.

5. В выбранной коллекции купинов однодоменные белки с известной третичной структурой объединены в гомодимеры, структурно эквивалентные гомологичным доменам двудоменных белков (бикупинов). Это позволяет использовать для выравнивания и эволюционного анализа удвоенную информацию: сумму последовательностей N- и C-концевых  $\beta$ -баррелей бикупинов и удвоенные последовательности  $\beta$ -баррелей однодоменных белков. Правомочность такого приема подтверждается идентичностью топологии эволюционных деревьев, построенных на основе независимого анализа последовательностей как N-концевых, так и C-концевых  $\beta$ -баррелей бикупинов, и  $\beta$ -баррелей однодоменных белков.

### ЭВОЛЮЦИОННЫЙ ПУТЬ ЗАПАСНЫХ ГЛОБУЛИНОВ СЕМЯН

Применяя описанную выше стратегию исследования, мы обнаружили пять кластеров, образующих статистически достоверные эволюционные ветви, каждая из которых может быть поддержана белками с известной третичной структурой (рис. 1а). Общим корнем эволюционного древа, по-видимому, является группа однодоменных белков, структура субъединиц которых почти полностью состоит только из  $\beta$ -барреля (рис. 1а,б). Во всех продуктах дальнейшей эволюции (как однодоменных, так и двудоменных) последовательность доменов содержит C-концевое удлинение, которое в белках с известной структурой образует группу  $\alpha$ -спиралей. Структура двудоменных ОД образовалась, вероятно, в результате дубликации их однодоменного предшественника, примером которого может быть однодоменная ОД *Mycobacterium abscessus* cam64834 (рис. 1).

Мы нашли гомологию практически полных последовательностей однодоменной ОД *M. abscessus* и однодоменного джерминоподобного сферулина *Physarum polycephalum* aaa29979 (рис. 1в). Это пока единственное прямое свидетельство предполагавшегося ранее [15] происхождения джерминов из однодоменных бактериальных белков путем классической дивергентной эволюции. Отметим, что присутствие генов сферулина, являющегося безусловным предшественником джерминов [10], в геноме *P. polycephalum* (*Amoebozoa*) исключает участие симбиогенеза в эволюции джерминов [15].

Двудоменная структура запасных глобулинов образовалась либо в результате независимой дубликации однодоменного растительного джерминоподобного предшественника (гипотеза I [8, 15]), либо была напрямую унаследована от двудоменных ОД (гипотеза II [12]). Согласно топологии эволюционного древа (рис. 1а), джермины и ОД,

образующие объединенный кластер, равновероятно могли быть предшественниками запасных глобулинов. Выбор между этими двумя вариантами может опираться на сопоставление полных первичных и высших структур ОД, джерминов и запасных глобулинов. В пользу гипотезы II свидетельствует тот факт, что, в отличие от джерминов, структура доменов ОД и запасных глобулинов стабилизирована дополнительными  $\beta$ -стрендами A' и J', а междоменные взаимодействия — дополнительным  $\beta$ -стрендом Z. Необычайно высокий уровень сходства первичных, третичных и четвертичных структур ОД цианобактерий и растительных запасных глобулинов может быть результатом симбиогенеза.

Итак, прямыми эволюционными предшественниками двудоменных запасных глобулинов семян, вероятнее всего, были двудоменные ОД, произошедшие из однодоменных ОД. В свою очередь, вероятно, что последние могли быть предшественниками джерминоподобных белков (сферулинов *Amoebozoa*) и джерминов растений (рис. 2а).

### СУПЕРСЕМЕЙСТВО КУПИНОВ

Общим признаком, позволившем объединить в купинное суперсемейство белки, отличающиеся по функциям, аминокислотным последовательностям и доменной организации, послужило присутствие в их третичных структурах (установленных или предполагаемых)  $\beta$ -барреля, содержащего узловые мотивы (в той или иной мере вырожденные), характерные для джерминов [10]. Задача установления достоверных филогенетических взаимоотношений между членами суперсемейства купинов вряд ли вообще разрешима, если опираться лишь на сопоставленные во множественном выравнивании короткие отрезки последовательностей, содержащие узловые мотивы (52 аминокислотных остатка,  $\beta$ -стренды CDE-FGH, см. рис. 1б).

Суперсемейство купинов образовано многочисленными семействами функционально родственных белков, полные аминокислотные последовательности которых в пределах семейств безусловно гомологичны. Поэтому задача эволюционного анализа каждого из отдельных семейств легко разрешима. Намного более сложной является задача установления филогенетических взаимоотношений между семействами, если сходство между ними ограничивается областью купинового модуля. Феномен присутствия в структурах функционально разнородных белков купинового  $\beta$ -барреля можно объяснить четырьмя причинами, которые рассматриваются ниже.

1. Эволюция двух или нескольких семейств купинов развивается в соответствии с канонами классической дивергентной эволюции. Исходным их предшественником может быть прокупин, структу-



**Рис. 1.** Эволюция запасных глобулинов семян. *а* – Эволюционное древо. Проанализированная область последовательностей соответствует сумме двух купиновых модулей, имеющихся в двудоменных белках и в гомодимерах однодоменных белков (156 позиций выравнивания). Цифры над ветвями отражают статистическую поддержку кластеров (% из 1000 репликаций). Звездочками отмечены последовательности однодоменной оксалатдекарбоксилазы (ОД) и сферулина. Справа показаны ленточные диаграммы третичных структур субъединиц легумина *Glycine max* Iod5 [16] и ОД *Synechocystis* sp. 2vqa [17] (С-концевые домены) и джермина *Hordeum vulgare* Ifi2 [11]. *б* – Структурное выравнивание последовательностей, представляющих кластеры легумина, джермина и ОД, а также корневой кластер (купин 2dct *Thermus thermophilus*). Строчными буквами даны аминокислотные остатки, не совпадающие по положению ни в одном из парных сочетаний структур. Символом > обозначены С-концевые аминокислоты белков. Отметим практически полное совпадение в области купинового модуля (стренды BCDE-FGHI) структурного выравнивания с выравниванием, построенным по первичной структуре белков. Звездочками отмечены позиции, содержащие характерные для купинов консервативные аминокислотные остатки [10]. *в* – Гомология практически полных аминокислотных последовательностей однодоменной бактериальной ОД и сферулина. Область купинового модуля подчеркнута.

ра которого целиком состоит из  $\beta$ -барреля. Последующая эволюция прокупина приводит к образованию усложненной структуры общего вторичного предшественника семейств. Таким образом, аминокислотные последовательности семейств сохраняют признаки гомологии далеко за пределами купинового модуля. Эволюция запасных глобулинов семян служит примером (пока единственным) описанной выше дивергентной эволюции купиновых семейств (рис. 2а).

2. Аналогичным образом, прокупин является общим предшественником двух или нескольких семейств купинов, однако последующая его эволюция может развиваться по двум или нескольким независимым сценариям, образуя различные вторичные предшественники двух или нескольких семейств. В этом случае семейства сохраняют признаки гомологии исключительно в области купинового  $\beta$ -барреля. Примером такого рода могут быть эволюционные взаимоотношения между корневым кластером ветви запасных глобулинов и семейством гентизатдиоксигеназ (ГД) (рис. 2б,в).

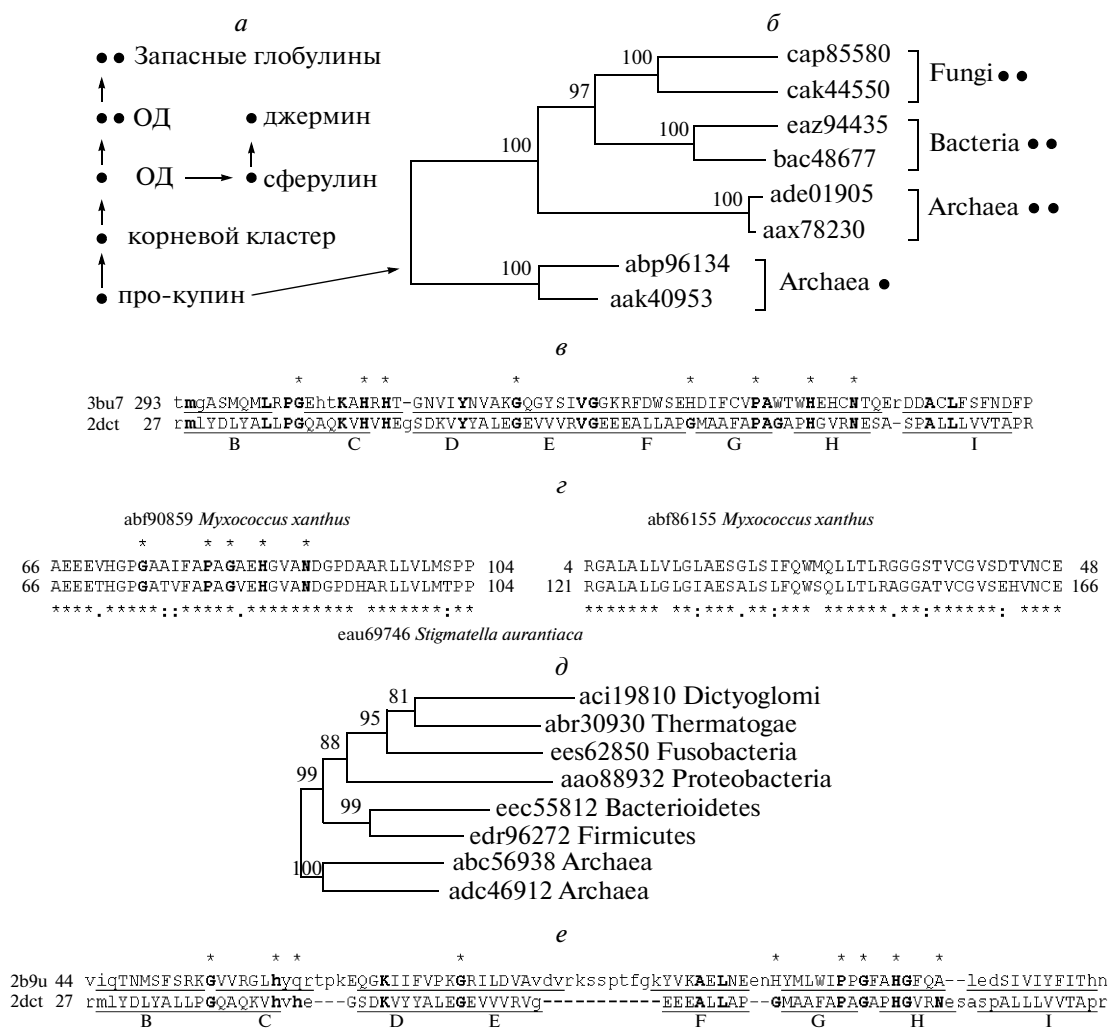
3. Для многих семейств купиновый  $\beta$ -баррель может быть примером эволюционно мобильного модуля. Такой модуль может оказаться настолько успешным для структурной и функциональной адаптации, что появляется во множестве белков в весьма различном структурном обрамлении [18]. Будучи внедренным в последовательность того или иного белка, купиновый мобильный модуль в этом новом обрамлении претерпевает модификацию, что затрудняет установление источника исходного модуля. Между тем, именно такое установление и является единственной задачей, решение которой позволит описать отнюдь не филогенетические взаимоотношения ныне существующих семейств купинов, а историю возникновения купинового модуля в каждом из них. Модификация купинового  $\beta$ -барреля, внедренного в чужеродный белок, развивается во времени, и факт недавнего события внедрения может быть легко обнаружен (рис. 2г). Отметим, что термин “мобильный модуль” мы используем здесь в широком смысле вне зависимости от генетических механизмов, ответственных за его мобильность.

4. До настоящего времени нет достоверных свидетельств предполагаемого [10] происхождения  $\beta$ -барреля у всех членов суперсемейства купинов из единого эволюционного предшественника. Весьма вероятно, что признаки сходства между купиновыми модулями двух или нескольких семейств купинов – следствие конвергентной эволюции, обусловленной формированием сходных структур  $\beta$ -барреля. Возможная независимая эволюция двух таких семейств рассматривается на примере сопоставления купинов корневого кластера ветви запасных глобулинов и эпимераз (ЭП) (рис. 2д,е).

Четыре предполагаемых варианта развития структуры  $\beta$ -барреля в функционально разнородных семействах купинов иллюстрируются приведенными ниже примерами.

**Эволюционная ветвь запасных глобулинов.** Купины относительно консервативного корневого кластера отражают примитивную структуру  $\beta$ -барреля гипотетического прокупина – общего предка двух или большего числа купиновых семейств. Однодоменные ОД, произошедшие из корневого кластера, приобрели более сложную структуру, которая далее развивалась по двум независимым сценариям: с образованием однодоменных сферулинов и затем джерминов (а), с образованием двудоменных ОД и запасных глобулинов (б) (рис. 2а).

**Семейство гентизатдиоксигеназ (ГД)** содержит белки архебактерий, бактерий и грибов, образованные двумя гомологичными структурно эквивалентными доменами. Полные последовательности доменов ГД, далеко выходящие за пределы купинового  $\beta$ -барреля, гомологичны полным последовательностям однодоменных ГД архебактерий (рис. 2б). Поскольку область гомологии ГД и ОД ограничивается  $\beta$ -баррелем, очевидно, что они образовались в результате двух независимых событий дубликации различных однодоменных предшественников (но не единого события, как это предполагалось ранее [10]). В области  $\beta$ -барреля структурное и обычное (построенное по первичной структуре белков) выравнивание ГД протеобактерии *Ruegeria pomeroyi* 3bu7 [19] и купина *T. thermophilus* (рис. 1) практически совпадают (рис. 2в). Сходство как высших, так и первичных структур купинового модуля ГД и белков корневого кластера ветви ОД позволяет предпо-



**Рис. 2.** Возможные пути эволюции структуры  $\beta$ -барреля в функционально разнородных семействах купинов. *a* – Эволюционные ветви запасных глобулинов и джерминов. Символами ● и ●● обозначены соответственно однодоменные и двудоменные белки. *б* – Двудоменные гентизатдиоксигеназы (ГД) образовались, вероятно, в результате дупликации однодоменных ГД (корневой кластер). Проанализированная область последовательностей соответствует сумме полных доменов, присутствующих в двудоменных белках и в гомодимерах однодоменных белков (249 позиций выравнивания). *в* – Структурное выравнивание ГД протеобактерии *Ruegeria pomeroyi* 3bu7 и купина *T. thermophilus* 2dct. *г* – Последовательность витамин-К-эпоксид–редуктазы *S. aurantiaca*, содержащая купиновый модуль (показана его С-концевая часть), гомологична сумме независимых последовательностей купина и витамин-К-эпоксид–редуктазы *M. xanthus*. *д* – Филогенетические взаимоотношения между эпимеразами (ЭП) бактерий и археобактерий. Проанализированы полные последовательности белков (178 позиций выравнивания). *е* – Структурное выравнивание ЭП археобактерии *Sulfolobus tokodaii* 2b9u и купина *T. thermophilus* 2dct. Звездочками отмечены позиции, содержащие характерные для купинов консервативные аминокислотные остатки [10]. Строчными буквами показаны аминокислотные остатки, не совпадающие по положению в пространственно совмещенных структурах белков.

лагать, что они происходят из общего предшественника (прокупина, существовавшего на ранних этапах возникновения жизни).

**Купиновый  $\beta$ -баррель как эволюционно мобильный модуль.** В качестве примера приведем последовательность eau69746 протеобактерии *Stigmatella aurantiaca*, купиновый модуль которой успешно использовали в корневом кластере эволюционного древа запасных глобулинов (рис. 1а). Последовательность *S. aurantiaca* состоит из двух частей: купинового  $\beta$ -барреля и витамин-К-эпоксид–редукта-

зы. Обе части существуют отдельно в качестве независимых белков многих бактерий. В частности, сумма таких последовательностей белков abf90859 и abf86155 протеобактерии *Mycobacterium xanthus* в высокой степени гомологична (лишь 7% неконсервативных замен) полной последовательности *S. aurantiaca* (рис. 2г). Поэтому представляется достаточно вероятным, что гены, кодирующие соответствующие независимые белки, присутствуют в геноме не только *M. xanthus*, но и *S. aurantiaca*. По-видимому, формирование мультидоменного белка eau69746 *S. aurantiaca*

са — следствие копирования и встраивания гена, кодирующего короткую купиновую последовательность, в 5'-область гена витамин-К-эпоксид-редуктазы. Множественность событий такого копирования (дубликаций) нуклеотидных последовательностей в эволюции геномов эукариот [20] позволяет предполагать, что дубликация купинового модуля является одним из основных механизмов формирования  $\beta$ -барреля в разнообразных структурах белков суперсемейства купинов.

**Семейство эпимераз (ЭП).** В состав обширного кластера однодоменных ЭП входят белки архебактерий и разнообразных бактерий (рис. 2*д*), а также животных. Собственно купиновый  $\beta$ -баррель составляет приблизительно треть последовательности ЭП, которые необычайно консервативны. Структурное (рис. 2*е*) и обычное выравнивание в области  $\beta$ -барреля ЭП архебактерии *Sulfolobus tokodaii* и купина *T. thermophilus* далеки от совпадения (как в парном, так и множественном вариантах обычного выравнивания). Поэтому вряд ли можно с уверенностью предположить, что ОД и ЭП происходят из общего древнего предшественника. Так, совпадение в структурном выравнивании трех остатков Gly и остатка Pro (рис. 2*д*) может быть следствием конвергенции в связи с образованием сходных структур  $\beta$ -баррелей, в которых эти остатки ответственны за формирование поворотов полипептидной цепи между  $\beta$ -стрендами В/С, D/E и G/H.

В настоящем обзоре мы попытались показать, насколько сложна задача описания филогенетических взаимоотношений между семействами белков, объединяемых в суперсемейство купинов. По-видимому, рациональный путь решения такой задачи заключается в идентификации эволюционных предшественников для каждого из семейств и в сопоставлении первичных и третичных структур этих предшественников. Можно ожидать, что это позволит различить дивергентную и конвергентную эволюцию как источник сходства между купиновыми  $\beta$ -баррелями различных семейств и, возможно, проследить путь перемещения купинового модуля как эволюционно мобильного структурного элемента.

## СПИСОК ЛИТЕРАТУРЫ

- Shewry P.R., Casey R. 1999. *Seed Proteins*. Dordrecht: Cluwer.
- Lawrence M.C., Izard T., Beuchat M., Blagrove R.J., Colman P.M. 1994. Structure of phaseolin in 2.2 Å resolution. *J. Mol. Biol.* **238**, 748–776.
- Shutov A.D., Kakhovskaya I.A., Braun H., Bäumlein H., Müntz K. 1995. Legumin and vicilin-like seed storage proteins: evidence for a common single-domain ancestral gene. *J. Mol. Evol.* **41**, 1057–1069.
- Adachi M., Takenaka Y., Gidamis A.B., Mikami B., Utsumi S. 2001. Crystal structure of soybean proglycinin A1aB1b homotrimer. *J. Mol. Biol.* **305**, 291–305.
- Shutov A.D., Braun H., Chesnokov Yu.V., Bäumlein H. 1998. A gene encoding a vicilin-like protein is specifically expressed in fern spores. Evolutionary pathway of seed storage globulins. *Eur. J. Biochem.* **252**, 79–89.
- Shutov A.D., Bäumlein H. 1999. Origin and evolution of seed storage globulins. In: *Seed Proteins*. Eds Casey R., Shewry P. Kluwer: Dordrecht, pp. 543–561.
- Shutov A.D., Blattner F.R., Bäumlein H., Müntz K. 2003. Storage and mobilization as antagonistic functional constraints of seed storage globulin evolution. *J. Exp. Bot.* **54**, 1645–1654.
- Bäumlein H., Braun H., Kakhovskaya I.A., Shutov A.D. 1995. Seed storage proteins of spermatophytes share a common ancestor with desiccation proteins of fungi. *J. Mol. Evol.* **41**, 1070–1075.
- Dunwell J.M., Khuri S., Gane P.J. 2000. Microbial relatives of the seed storage proteins of higher plants: conservation of structure and diversification of function during evolution of the cupin superfamily. *Microbiol. Mol. Biol. Rev.* **64**, 153–179.
- Dunwell J.M., Culham A., Carter C.E., Sosa-Aguirre C.R., Goodenough P.W. 2001. Evolution of functional diversity in the cupin superfamily. *Trends Biochem. Sci.* **26**, 740–746.
- Woo E.J., Dunwell J.M., Goodenough P.W., Marvier A.C., Pikersgill R.W. 2000. Germin is a manganese containing homo-hexamer with oxalate oxidase and superoxide dismutase activities. *Nature Struct. Biol.* **7**, 1036–1040.
- Dunwell J.M., Gane P.J. 1998. Microbial relatives of seed storage proteins; conservation of motifs in a functionally diverse superfamily of enzymes. *J. Mol. Evol.* **46**, 147–154.
- Martin W., Herrmann R.G. 1998. Gene transfer from organelles to the nucleus: how much, what happens, and why. *Plant Physiol.* **118**, 9–17.
- Van de Peer Y., De Wachter R. 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput. Appl. Biosci.* **10**, 569–570.
- Shutov A.D., Blattner F.R., Bäumlein H. 1999. Evolution of structurally conserved protein module from *Archaea* to plants. *Trends Genet.* **15**, 348–349.
- Adachi M., Kanamori J., Masuda T., Yagasaki K., Kitamura K., Mikami B., Utsumi S. 2003. Crystal structure of soybean 11S globulin: glycinin A3B4 homo-hexamer. *Proc. Natl. Acad. Sci. USA.* **100**, 7395–7400.
- Tottey S., Waldron K.J., Firbank S.J., Reale B., Bessant C., Sato K., Cheek T.R., Gray J., Banfield M.J., Dennison C., Robinson N.J. 2008. Protein-folding location can regulate manganese-binding versus copper- or zinc-binding. *Nature.* **455**, 1138–1142.
- Doolittle R.F. 1995. The multiplicity of domains in proteins. *Ann. Rev. Biochem.* **64**, 287–314.
- Chen J., Li W., Wang M., Zhu G., Liu D., Sun F., Hao N., Li X., Rao Z., Zhang X.C. 2008. Crystal structure and mutagenic analysis of GDOsp, a gentisate 1,2-dioxygenase from *Silicibacter pomeroyi*. *Protein Sci.* **17**, 1362–1373.
- Thomas E.E., Srebro N., Sebat J., Navin N., Healy J., Mishra B., Wigler M. 2004. Distribution of short paired duplications in mammalian genomes. *Proc. Natl. Acad. Sci. USA.* **101**, 10349–10354.