

В.В. Суслов, К.В. Гунбин,
Н.А. Колчанов
Институт цитологии и генетики
СО РАН, г. Новосибирск

ГЕНЕТИЧЕСКИЕ МЕХАНИЗМЫ КОДИРОВАНИЯ БИОЛОГИЧЕСКОЙ СЛОЖНОСТИ

✿ Рост сложности организмов — глобальный тренд эволюции. Качественно более высокая сложность эукариот по сравнению с прокариотами отражается в особенностях организации их геномов и механизмах реализации генетических программ. Рассмотрены генетические механизмы кодирования биологической сложности у про- и эукариот: надтриплетные коды, комбинаторика генетических блоков и блоков генных сетей и их иерархическое взаимодействие.

✿ **Ключевые слова:** биологическая сложность, прокариоты, эукариоты, генетические коды, генные сети, регуляция.

ВВЕДЕНИЕ

Рост сложности организмов — глобальный тренд эволюции и признак эволюционного прогресса [87]. Можно выделить ряд характеристик, связанных с ростом сложности биологической организации: (i) увеличение количества элементов; (ii) связей между ними; (iii) уровней иерархии; (iv) количества элементов и связей, работающих в единицу времени и/или в единице объема; (v) разнообразия режимов поведения. Трудности изучения биологической сложности обусловлены тем, что ее классические определения слишком формальны. Так, по Колмогорову [7] сложность генетического текста есть оценка наименьшего числа генерирующих его операций (дубликаций, делеций, замен символов). По Кауфману [65] сложность определяется количеством конфликтующих параметров системы (например, конкурирующих транскрипционных факторов, регулирующих экспрессию гена). Хотя эти определения применяют в изучении ряда аспектов биологической сложности, в целом они не охватывают всей

Таблица 1

Сравнительная характеристика геномов про- и эукариот
(по [35, 114] с изменениями)

Таксон	Вид	Гаплоидный геном млн п.н.	Число генов в геноме
Прокариоты			
Микоплазмы	<i>Mycoplasma genitalium</i>	0,58	470
	<i>Mycoplasma pneumoniae</i>	0,82	~670
Риккетсии	<i>Rickettsia prowazekii</i>	1,1	834
Археобактерии	<i>Archaeoglobus fulgidus</i>	2,18	2436
	<i>Methanopyrus kandleri</i>	1,69	1738
Цианобактерии	<i>Synechocystis</i> sp.	3,57	3168
Эубактерии	<i>Escherichia coli</i>	4,6–5,5	4288
	<i>Campylobacter jejuni</i>	1,64	1654
	<i>Aquifex aeolicus</i>	1,55	1512
	<i>Neisseria meningitidis</i>	2,27	2121
	<i>Bacillus subtilis</i>	4,2	4100
Низшие эукариоты			
Грибы	<i>Saccharomyces cerevisiae</i>	11,4	6241
	<i>Schizosaccharomyces pombe</i>	13,8	4824
	<i>Aspergillus nidulans</i>	31	
Протисты	<i>Amoeba dubia</i>	670 000	
	<i>Entamoeba histolytica</i>	20	
	<i>Dictyostelium discoideum</i> ¹	32	11 000
Высшие эукариоты			
Высшие растения	<i>Lilium longiflorum</i>	90 000	
	<i>Arabidopsis thaliana</i>	115,7	27 540
	<i>Oryza sativa</i>	466	46 022–55 615
Первичноротые	<i>Caenorhabditis elegans</i>	97	19 049
	<i>Drosophila melanogaster</i>	120	13 600
Вторичноротые	<i>Protopterus aethiopicus</i>	139 000	
	<i>Fugu rubriceps</i>	365–400	30–40 тыс.
	<i>Homo sapiens</i>	3000	30 000
	<i>Mus musculus</i>	2500	37 000

¹ Протист, имеющий многоклеточную стадию жизненного цикла — плодовое тело [22].

проблемы. Так как фенотипические признаки организмов кодируются их геномами, ожидалось, что в разных таксонах геномы сильно различаются по числу генов. Но расшифровка геномов выявила что: (i) сложность прокариот в целом коррелирует с размерами геномов и числом генов; (ii) наблюдаются рост размера геномов и числа генов при переходе от прокариот к эукариотам и от одноклеточных к многоклеточным; (iii) однако у эукариот отсутствует связь между биологической сложностью, размерами геномов и числом генов (табл. 1) [34, 114].

В статье показано, что глобальный тренд усложнения биологической организации связан с качественным усложнением механизмов регуляции экспрессии генов.

1. ПРОГЕНОТЫ: РАННИЕ ЭТАПЫ ОБРАЗОВАНИЯ ГЕНЕТИЧЕСКИХ МЕХАНИЗМОВ КОДИРОВАНИЯ БИОЛОГИЧЕСКОЙ СЛОЖНОСТИ

Жизнь на Земле существует в форме генетических самовоспроизводящихся систем (ГСВС). Матричный синтез генетических макромолекул (ДНК, РНК, белков) — неотъемлемое свойство ГСВС. Предположительно существовавшие на древней Земле автокаталитические гиперциклы¹, способные к самовоспроизведению и эволюции за счет мутаций и дарвиновского отбора [48], не имели универсальных механизмов воспроизведения кодирующих матриц, что накладывало жесткое ограничение на число компонент гиперцикла и на число взаимодействий между ними. С вовлечением фосфат-ионов и азотистых оснований в гиперциклы сахаров могли возникать АТФ и РНК [3, 14]. Открытие природных РНК-ферментов — рибозимов [46] позволило сформулировать гипотезу предкового «Мира РНК» [63, 84]. Селекс-техникой («эволюция в пробирке») из пула случайных РНК удалось получить спектр рибозимов, замыкавших цикл самовоспроизведения: от синтеза рибонуклеотидов до синтеза РНК по РНК-матрице [62, 112, 113]. На этой стадии могли возникнуть сайзеры² — простейшие ГСВС [16, 17], имевшие эволюционное преимущество широкого профиля — универсальные процессы матричного синтеза, что обеспечило более высокую по сравнению с гипер-

циклом мутационную и параметрическую устойчивость сайзеров, и возможность наращивания их сложности путем объединения кодирующих матриц.

Следующим глобальным ароморфозом можно считать возникновение триплетного генетического кода — основы матричного биосинтеза еще одного класса генетических макромолекул — белков, образуемых из 20 типов аминокислот, обеспечивавших потенциально огромное разнообразие структур и активностей³.

Первые клеткоподобные организмы — прогены появились согласно палеонтологическим данным ~3,8 млрд лет назад [109] возможно путем объединения липидных [13] или протеиноидных микросфер [113] и сайзеров. Включение в геном информации о синтезе компонентов микросферы, вероятно, шло параллельно с усложнением сайзеров [17] или подобных им СВС [35].

Процессы комбинаторного объединения матриц, которые привели к возникновению геномов, вероятно начались на ранних этапах биологической эволюции и были связаны с возникновением сложных сайзеров, имевших большое количество кодирующих единиц в пределах одной самореплицирующейся системы, что позволяло им выполнять более сложные генетические программы. В возникших таким образом геномах на порядок выросла надежность хранения информации, что обеспечило дальнейшее наращивание ее количества за счет роста длины геномов и появления сложной системы надтриплетных кодов. Появление рекомбинации — основы блочно-модульной эволюции⁴ — позволило примитивным ГСВС наращивать сложность не только за счет фиксации мутаций, но и качественно новым способом — комбинаторикой геномных блоков, функция каждого из которых «проверена» эволюцией [16, 17].

Любая популяция ГСВС имеет верхнюю границу темпов мутирования. Гаплоидные популяции достигают ее, когда за один цикл репликации возникает одна летальная мутация на геном [48]. Следовательно, чем выше частота мутаций, тем ниже предельный размер генома⁵. Поэтому усложнение ГСВС, требующее роста генома, невозможно без роста надежности хранения и копирования генетической информации. Это ограничение было особенно важно на ранних этапах эволюции, послужив мо-

¹ Нелинейные автокаталитические цепи, в которых ферменты и кодирующие их матрицы кооперируются: матрица M_1 кодирует фермент E_1 ; ферменты циклически катализируют репликацию матриц, например, E_1 способствует репликации M_2 , E_2 способствует репликации M_3 , ..., E_n способствует репликации M_1 [17].

² Самовоспроизводящиеся системы, содержащие матрицу M , кодирующую ген фермента репликации E_1 , ген фермента трансляции E_2 и гены других белков, необходимых для жизнедеятельности. В отличие от гиперцикла, в сайзере репликация матрицы идет только с помощью фермента E_1 , а все белки (в том числе и E_2) транскрибируются со своих матриц одним и тем же ферментом E_2 [17].

³ Для ознакомления с проблемой возникновения и эволюции триплетного генетического кода, выходящей за рамки данной статьи см. обзор [119].

⁴ В возникновении рекомбинации возможно сыграли большую роль ГСВС-паразиты (примитивные вирусы, плазмиды транспозоны) [27].

⁵ Эйген показал, что средняя вероятность мутирования на позицию $(1-q)$, длина оптимальной матрицы (V_m) и параметры ее отбора σ_m связаны неравенством $V_m < V_{max} = (\ln \sigma_m) / (1-q)$. Оптимальной считается матрица (геном), имеющая наибольший параметр отбора σ_m , поэтому допустимые длины последовательностей ограничены сверху величиной V_{max} . Нарушение же этого неравенства является катастрофой мутационных ошибок [17, 48]. Для гаплоидных геномов это такая ситуация, когда в каждом цикле репликации геном получает как минимум одну летальную мутацию [17].

тивацией для возникновения и совершенствования высочайших систем репликации и репарации ДНК.

2. ПРОКАРИОТЫ: ГЕНЕТИЧЕСКИЕ МЕХАНИЗМЫ КОДИРОВАНИЯ БИОЛОГИЧЕСКОЙ СЛОЖНОСТИ

Особенности организации прокариотической клетки и генома

Суммируя структурно-функциональные особенности организации прокариотической клетки [35, 37], можно определить характерные черты организации, отражающие уровень биологической сложности прокариот и возможности их эволюционного прогресса. Это гаплоидность (i), пассивный механизм сегрегации хромосом, ассоциированный с мембраной (ii), отсутствие хроматина (iii), малые размеры клеток (iv), узкоспециализированный тип питания — пиноцитоз, связанный с наличием жесткого экзоскелета (v), отсутствие компартментов и, как следствие, невозможность разделения метаболических процессов в пространстве (vi), отсутствие активного внутриклеточного транспорта, роль которого играет диффузия (vii).

В бактериальных геномах элементарной единицей регуляции транскрипции является оперон¹. Число белковых регуляторов у бактерий связано с общим числом оперонов (групп одновременно транскрибируемых генов) аллометрически — степенной функцией, то есть растет очень быстро в зависимости от их количества [41]. Поэтому количество независимо транскрибируемых групп генов у бактерий должно быть ограничено сверху. Преимущество оперонной транскрипции генов — возможность появления необходимых белковых продуктов в стехиометрических соотношениях, что критически важно для оптимизации метаболизма². Недостаток — отсутствие гибкости в реализации генетических программ.

Очевидно, бактерии подошли к границе мутационной катастрофы ошибок [38], что наложило ограничение на размер гаплоидного генома. Интересно, что теоретические оценки предельного размера генома, вычисленные исходя из мутационной катастрофы ошибок (~ 6 млн п.о. [17]), близки к реальным данным (см. табл. 1). Примечательно, что в бактериальной хромосоме в ходе эволюции минимизировалось количество некодирующей ДНК [104], а также проходил горизонтальный перенос генов³.

Таким образом, особенности организации генома и структурно-функциональной организации клетки прокариот определили вектор эволюции бактерий как ус-

ложнение не морфологии, а метаболизма (в том числе и путем формирования бактериальных сообществ). В итоге, благодаря разнообразию их метаболизмов, бактерии замкнули биогеохимические циклы, что впоследствии стало основой существования биосферы.

Бактериальные сообщества

Бактерии в естественной среде, как правило, присутствуют в виде сообществ [4, 40], в которых существует общий пул метаболитов и генов (метагеном), взаимодействующих друг с другом на основе гибридных генных сетей⁴, сформированных бактериями разных видов или разных штаммов одного и того же вида⁵ [40]. Информационная емкость метагенома потенциально неограничена и может возрастать без увеличения размеров отдельного генома. Кроме того, в таких сообществах решается проблема компартментализации — роль компартментов играют бактерии с разным метаболизмом. Пик усложнения прокариот — сообщества на твердых субстратах — естественных (маты) [4] или секретируемых бактериями (био-пленки) [40, 75]. Метаболические пути на твердом субстрате могут быть разнесены в пространстве (аналог морфологии) [5], а внеклеточный матрикс биопленок даже формирует аналог проводящей системы [40].

Прокариоты: межклеточные коммуникации

Межклеточные коммуникации у бактерий ограничены экзоскелетом и осуществляются либо небольшими молекулами (например, «кворум-чувствительные» сигналы, обеспечивающие внутривидовые коммуникации) [88], либо с помощью неуниверсальных механизмов (пили миксобактерий [47], белковые комплексы полярной септы *Bacillus subtilis* [72] и др.).

Специальные системы коммуникаций (фаги и плазмиды) обеспечивают обмен генетическим материалом и, как следствие, возможность реализации дополнительных генетических программ, генерацию генетического разнообразия и увеличение информационно-генетической интеграции бактериальных сообществ. В частности, в задачи таких систем входит обеспечение устойчивости к некоторым неблагоприятным факторам среды, недопущение повторной фаговой инфекции в уже инфицированной популяции бактерий (абортивная инфекция) и оптимизация числа плазмид в бактерии [15].

Такие мобильные модули могут встраиваться в бактериальные геномы, обеспечивая горизонтальный перенос генетической информации и формирование

¹ В свою очередь опероны образуют видоспецифичные консервативные надоперонные ансамбли, оптимизирующие транскрипционную регуляцию путей метаболизма [115].

² Поэтому оперонная организация генов распространена у про- и эукариот среди генов, продукты которых тесно (часто физически) связаны [116].

³ Для ознакомления с этой проблемой, выходящей за рамки данной статьи, см. обзоры [59, 61, 97].

⁴ Генных сетей, состоящих из компонентов (генов, белков, небелковых веществ), относящихся к различным организмам.

⁵ Метаболические связи в таких сообществах порой столь тесны, что развитие одних бактерий невозможно, если продукты их метаболизма не утилизируют другие бактерии [4].

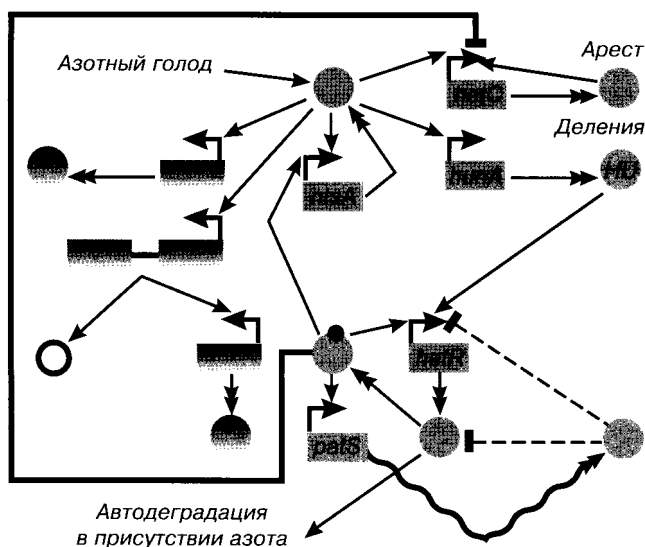


Рис. 1. Генная сеть дифференцировки гетероцисты.

NtsA активирует гены азотфиксации и с участием гена *hupA* включает ген *hetR*. При азотном голоде белок HetR фосфорилируется, что предотвращает его автодеградацию. Азотный голод активирует белок NtsA, стимулирующий синтез белков HetC, HetR и самого себя. HetC связан с остановкой деления. HetR подавляет *hetC*, разрешая синтез ДНК и тем самым транскрипцию оперонов азотфиксации, генов *patS* и *ntcA*. Взаимодействие положительной (*hetR*>*ntcA*>*hetC*) и отрицательной (*hetR*>*hetC*) обратных связей обеспечивает устойчивость процесса дифференцировки гетероцисты. Белок PatS, благодаря малым размерам [88], диффундирует в клетки, соседствующие с гетероцистой, подавляя *hetR* (аналог градиента морфогенов эукариот) [49]

качественно новых признаков, дающих селективное преимущество бактериальной популяции. Например, в геноме *Escherichia coli* описан комплекс генов *mazF*, *mazE*, *relA*, кодирующих стойкий токсин, нестойкий антитоксин (подавляющий активность токсина или разрушающий его) и регуляторный белок соответственно [50]. В голодных клетках *E. coli* синтез *mazE* и *mazF* подавлен, антитоксин быстро разрушается, что приводит к их гибели, уменьшению конкуренции бактерий за субстрат, а также использованию автолизата погибших бактерий в качестве источника питания [50]. Возможно, эти гены имели плазмидное происхождение [25]. Действительно, известны малокопийные плазмиды, механизм фиксации которых в бактерии связан с так называемым «модулем привыкания» — наработкой внутри клетки стойкого токсина и нестойкого антитоксина. С потерей плазмиды антитоксин быстро разрушается и токсин лизирует бактерию [50]¹.

¹ Такая комбинаторика разных ГСВС могла идти уже в Мире РНК, на что указывают РНК-антитоксины. Например, РНК гена *Sok*, являясь бессмысловой к участку мРНК белка-токсина *Nok*, блокирует его экспрессию [53].

Бактериальная многоклеточность

Замечательно, что у ряда бактерий (цианобактерии, спорулирующие бактерии, миксобактерии) существует более консолидированная форма организации, сравнимая с многоклеточностью. Например, у цианобактерий² азотный голод индуцирует дифференцировку азотфиксирующих клеток — гетероцист из обычных клеток [88] (рис. 1). Показано, что у бактерий с многоклеточными стадиями геном превосходит 4 млн п.н. (см. табл. 1). В образовании и функционировании гетероцист прямо задействовано ~140 [122], а косвенно — до 1000 [82] генов. За споруляцию *B. subtilis* (формирование двуклеточного спорангия) прямо отвечают 164 гена [90], косвенно — много больше [15]. За формирование плодового тела миксобактерий (2–4 типа клеток) прямо отвечают не менее 300 генов, ~200 генов ответственны за таксисы и еще примерно столько же связаны с этими процессами косвенно [120]. Таким образом, многоклеточность у бактерий обеспечивает минимум тысяч генов.

Что же помешало появлению на основе прокариот мира многоклеточных, сравнимого с эукариотами? Мы полагаем, что отсутствие в метагеноме централизованного, иерархически высокого уровня регуляции препятствовало образованию многоклеточной жизни на основе бактериальных сообществ. К факторам, затруднившим образование многоклеточности на основе отдельных видов бактерий относились: (i) жесткая организация регуляции экспрессии генов (оперонная организация генома, короткий размер регуляторных районов, контролирующих транскрипцию оперонов), что ограничивало разнообразие генетических программ, закодированных в бактериальных геномах; (ii) ограниченное количество генов и, как следствие, малая информационная емкость геномов, которой не хватало для кодирования сложных процессов онтогенеза, дифференцировки клеток, межклеточных коммуникаций, активного транспорта, сложных поведенческих реакций, движений многоклеточных агрегатов; (iii) экзоскелет, препятствовавший межклеточным взаимодействиям, а также ограничивавший разнообразие морфотипов клеток [35, 37, 121].

3. ЭУКАРИОТЫ: ГЕНЕТИЧЕСКИЕ МЕХАНИЗМЫ КОДИРОВАНИЯ БИОЛОГИЧЕСКОЙ СЛОЖНОСТИ

Качественно более высокая биохимическая, физиологическая, морфологическая, поведенческая сложность у эукариот по сравнению с прокариотами находит отражение как в особенностях их геномной организации, так и механизмах реализации генетических

² Которые существуют, как правило, в виде многоклеточных бактериальных тяжей и обладают способностью к передвижению [15].

программ, контролирующих фенотипические характеристики эукариотических организмов.

Характерная особенность геномов эукариот: высокая насыщенность повторами

Характерная особенность организации геномов эукариот — исключительно низкая плотность белок-кодирующих последовательностей. Если у прокариот такие последовательности занимают до 95% геномной ДНК, то у эукариот — около 5% [114].

Известно, что при репликации в результате неправильного спаривания комплементарных нитей ДНК по повторам, возможны делеции и дупликации, причем частота делеций превышает частоту дупликаций [38]. Решением проблемы неправильного спаривания ДНК по повторам у эукариот было появление хроматина [38], обеспечившего несколько уровней укладки и плотной упаковки ДНК. По мере репликации ДНК с ней прочно связываются белки хроматина [24]. Поэтому длина одонитевой ДНК в репликативной вилке у эукариот (100–200 п.н.) в 10 раз меньше, чем у бактерий (1000–2000 п.н.) [24], что существенно снижает частоту и размеры делеций и дупликаций по повторам [38]. С появлением хроматина стала возможна упорядоченная сегрегация хромосом при клеточных делениях [36], что создало

предпосылки для появления (i) сегментированных геномов и (ii) диплоидности. Диплоидность отодвинула границу мутационной катастрофы ошибок [17] — длина геномов эукариот на 3–6 порядков больше прокариотических (см. табл. 1). Диплоидность обеспечила возникновение кроссинговера — блочной перетасовки фрагментов гомологичных хромосом [36]. Аберрации же кроссинговера, неравный кроссинговер (НК), идущий по участкам tandemных протяженных повторов, обеспечили возникновение дупликаций. Постоянство среднего уровня насыщенности повторами в популяции обусловлено появлением при НК геномов как с увеличенным, так и уменьшенным числом повторов [2]. Таким образом, устойчивость геномов эукариот по отношению к повторам привела к появлению принципиально новых свойств, отсутствовавших у прокариот: мультигенности, наличию кластеров изофункциональных генов.

Кодирование генетической сложности: надтриплетные генетические коды

Первоначально некодирующую ДНК эукариот рассматривали как нефункциональную, а геномы эукариот как неоптимально организованные, так как доля белок-кодирующей ДНК в них очень мала. Однако постепенно

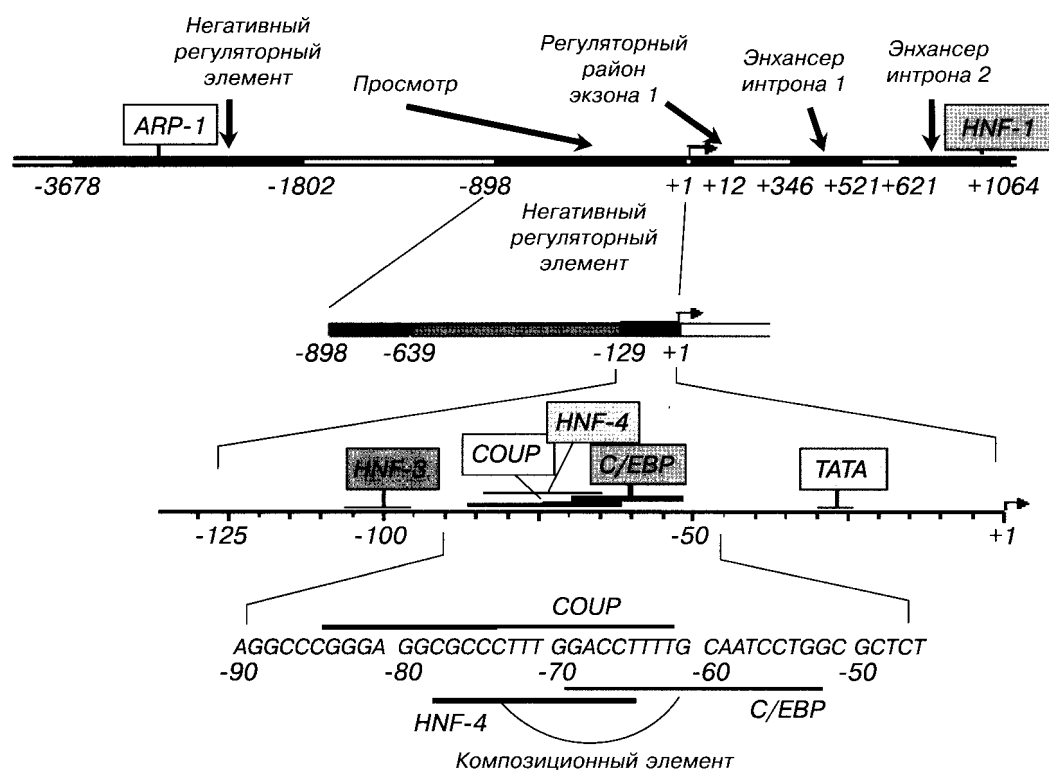


Рис. 2. Фрагмент иерархически организованного РР гена апо-липопротеина В [68].

Ген апо-липопротеина В содержит множество регуляторных элементов, которые могут находиться на большом расстоянии от старта транскрипции, а также в интронах и 3'-фланкирующем районе гена

стала формироваться иная точка зрения [20]: «Преобладающие в геномах эукариот некодирующие последовательности, вероятно, кодируют нечто иное, что не требует привлечения традиционного триплетного кода. Иными словами, кроме триплетного кода и трансляционной машины, в клетке имеются другие коды и средства их чтения. При этом под кодом понимается любой тип нуклеотидного контекста, значимый для выполнения определенной биомолекулярной функции». Рассмотрим некоторые из этих кодов, называя их надтриплетными, взяв за основу концепцию Э. Трифонова [20], расширенную с учетом современных данных.

Коды регуляции транскрипции

Количественная величина активности любых функциональных сайтов (ФС) ДНК, в том числе и регуляторных районов (РР) генов, как у эукариот, так и у прокариот определяется спецификой их взаимодействий с регуляторными белками, зависящей от их конформационных/физико-химических свойств [99]. Так кодируются самые разнообразные свойства ФС. Например, средство к регуляторным белкам сайтов связывания транскрипционных факторов (ССТФ), время жизни ДНК-белковых комплексов, а также кинетические характеристики их формирования и т. п. [98]. Интересно, что даже одиночные нуклеотидные замены способны существенно менять величину активности ФС — от полной ее потери до выраженного увеличения, а также приводить к появлению активных ФС в ранее нефункциональной ДНК [98].

РР генов прокариот имеют простую организацию: небольшие размеры (до 60–100 п.о.) и ограниченное количество сайтов связывания регуляторных белков [64]. РР генов эукариот имеют качественно иной уровень сложности (рис. 2): (i) большую длину — от тысяч до десятков тысяч п.н.; (ii) большое количество (до многих десятков) ССТФ в пределах РР и (iii) их сложную иерархическую организацию (ССТФ, композиционные элементы (КЭ), формируемые парами сближенных ССТФ, энхансеры и сайленсеры, образованные комбинациями ССТФ и КЭ). Размер регуляторного района гена может быть на порядки больше размера его кодирующей части [68, 124].

Существенно также, что ядро любой клетки многоклеточного организма имеет в зависимости от ее функционального состояния (определяемого типом клетки, стадией клеточного цикла, типом ткани, стадией развития организма, внешней средой, действием индукторов и т. д.) определенный набор транскрипционных факторов (ТФ) [68]. Взаимодействуя с ССТФ, белками базального комплекса и между собой, ТФ формируют генспецифический транскрипционный комплекс, определяющий уровень транскрипции конкретного гена в клетке, находящейся на определенной стадии клеточного цикла и дифференцировки в многоклеточном организме под дей-

ствием определенного внешнего стимула [68]. Структура транскрипционного комплекса определяется регуляторными элементами, их расположением относительно старта транскрипции, набором факторов, присутствующих в ядре и последовательностью его формирования во времени [39, 68]. Она также определяет не только интенсивность транскрипции конкретного гена, но и начало старта транскрипции: для многих генов выявлены альтернативные и множественные старты транскрипции, обеспечивающие разнообразие вариантов гЯРНК, отличающихся структурой первого интрона [68, 124].

В рамках простейшей бинарной модели (сайт связывания взаимодействует с транскрипционным фактором или свободен от него) емкость кода регуляции транскрипции W оценивается как 2^N . Например, при количестве ССТФ $N = 30$, $W = 2^{30} \sim 10^9$. То есть даже простой комбинаторики ССТФ, расположенных в РР данного гена, и ТФ достаточно для кодирования огромного разнообразия вариантов транскрипции гена в множестве клеток, тканей на различных этапах развития многоклеточного организма при его различных функциональных состояниях [9]. Более того, по разным оценкам, доля белков, связанных с транскрипцией, составляет от 6 до 15% от протеома разных эукариот [58, 94, 118]. Также важно отметить, что скорость эволюции белков транскрипционной машины относительно высока; например, скорость эволюции белков аппарата транскрипции *Arabidopsis thaliana* превосходит скорости эволюции других групп белков *A. thaliana* [117]. Эволюция же белков аппарата транскрипции, как правило, адаптивна, например, это подтверждается исследованием режимов эволюции различных групп генов эволюционно близких к *Drosophila melanogaster* видов [85, 103].

По причине важной роли дупликаций в ходе эволюции эукариот [105, 123] рост числа различных вариантов транскрипции генов может быть связан с широким распространением кластеров изофункциональных генов. Такие кластеры содержат гомологичные гены с частично различающейся структурой и функцией. Экспрессия генов, входящих в состав кластеров, осуществляется в зависимости от стадии индивидуального развития, функционального состояния организма и т. п. Порядок и интенсивность экспрессии генов может определяться специальным классом регуляторных элементов иерархически высокого (надгенного) уровня — LCR (локус-контролирующими районами) [33, 68, 76]. Каждый LCR образован специфической группой ССТФ и располагается иногда на очень большом (до десятков тысяч п.о.) расстоянии от контролируемой кассеты генов [33, 76].

Коды формирования нуклеосом и хроматин

У эукариот по сравнению с прокариотами конформационные/физико-химические свойства ДНК играют существенно большую роль, определяя особенности ну-

клеосомной организации ДНК. Нуклеосомы формируются предпочтительно в определенных участках геномной ДНК с характерным динуклеотидным контекстом [20], и их расположение в геномах неслучайно. Элементами кода нуклеосомной упаковки ДНК являются короткие олигонуклеотиды, распределенные определенным образом вдоль нуклеосомного сайта и определяющие конформационные свойства ДНК [107], оптимальные для взаимодействия с гистоновым октамером.

Нуклеосомная организация ДНК соответствует базовому уровню организации хроматина, который отсутствует у прокариот и является характерной особенностью эукариот [70]. Хроматин не только обеспечивает высокую степень компактизации геномной ДНК в ядре, что критически значимо при больших размерах геномов эукариот. Он является также важнейшим фактором регуляции транскрипции генов эукариот, обеспечивая «разметку» транскрипционно-активных участков геномов [74]. Регуляция транскрипции генов требует особого расположения нуклеосом, например, для обеспечения доступности функциональных сайтов ДНК регуляторным белкам или наоборот — для экранирования таких сайтов. Например, промоторам генов домашнего хозяйства свойственна ослабленная нуклеосомная упаковка ДНК или ее полное отсутствие [74]. По-видимому, это обеспечивает легкий доступ к промоторам белков базального транскрипционного комплекса, что необходимо для эффективной транскрипции. Напротив, промоторы тканеспецифических генов, как правило, характеризуются сильной нуклеосомной упаковкой ДНК, разрушение которой при транскрипции требует перестройки хроматина [74, 92].

Замечательно, что у эукариот существуют специальные механизмы перестройки структуры хроматина (*remodeling*), меняющие нуклеосомную «разметку» ДНК в зависимости от дифференцировки и функционального состояния клетки [39, 92]. Механизм такой перестройки структуры хроматина высоко консервативен, но его запуск определяется генными сетями функционального состояния клетки и клеточной дифференцировки, варьирующими у разных эукариот [39]. Ремоделинг хроматина — не только важнейший фактор регуляции экспрессии генов, но и механизм кодирования эпигенетических эффектов. Например, спецификация Нох-генами тканей сегментов тела *D. melanogaster* эпигенетически наследуема и зависит от белков семейств *Polycomb*, *Polyhomeotic* и *tritorax* [83]. Белки этих семейств ответственны за модификацию структуры хроматина, поддерживающих транскрипцию Нох-генов в потомках единожды детерминированной клетки [31, 83].

Коды блочной организации кодирующих районов генов эукариот

Экзон-интронная структура свойственна большинству расшифрованных генов эукариот [56, 106, 108]. Ее

возникновение — один из важнейших ароморфозов, создавших предпосылки для компактного кодирования генетической информации на основе комбинаторики экзонов и интронов. В результате альтернативного сплайсинга на основе одного и того же гена возможно кодирование большого разнообразия вариантов мРНК, отличающихся наборами экзонов, и как следствие — большого разнообразия вариантов белков. Например, ген *Dscam* дрозофилы путем альтернативного сплайсинга кодирует тысячи вариантов белка, контролирующего рост и формирование аксонов [29], тем самым тонко регулируя формирование нервной системы. Экзон-интронная структура и альтернативный сплайсинг эукариот являются новым способом кодирования генетической информации [41, 57], позволяющим наращивать сложность генетических программ регуляции экспрессии генов без существенного увеличения размеров геномов.

Замечательно также, что интроны не являются генетически инертной ДНК [52, 86], так как содержат большое разнообразие регуляторных элементов, влияющих на экспрессию генов — сайты связывания транскрипционных факторов, энхансеры, альтернативные промоторы и т. д. [68]. Показано [45, 74], что интроны обладают более высоким потенциалом формирования нуклеосом по сравнению с функционально более нагруженными экзонами. Предполагается, что само по себе возникновение мозаичной структуры генов связано с нуклеосомной упаковкой ДНК [18, 42]. Нуклеосомная упаковка геномной ДНК требует формирования нуклеосом в среднем на расстоянии 150–250 п.о. Однако в силу ограничений со стороны кодирующих последовательностей, гены не всегда могут иметь нуклеотидный контекст, обеспечивающий формирование стабильных нуклеосом [74]. Кодирующие последовательности, не удовлетворяющие этим условиям, могли быть разделены интронами.

Для многих белков установлено соответствие между экзонами генов и доменами кодируемых ими белков [44]. Это создает молекулярную основу для блочной эволюции за счет перекомбинирования экзонов — образованию новых белков из функциональных частей их предшественников [55, 78]. У прокариот проблема мутационного поиска новых вариантов генов не имеет большой остроты из-за гигантских численностей популяций и широкой распространенности горизонтального переноса генов [73]. Однако она существенна для большинства эукариот из-за, как правило, малой численности их популяций. Один из путей решения проблемы состоял в широком распространении интронов и возникновения на их основе механизмов изменчивости на основе комбинаторики блоков генов и белков [81]. Таким образом, эукариоты смогли приобретать новые варианты молекулярных структур существенно быстрее, чем прокариоты, без необходимости увеличения частоты точковых мутаций.

Пожалуй, наибольшего совершенства комбинаторика блоков генов и белков достигает в иммунной системе. Например, у человека V- и С-области иммуноглобулиновых генов кодируются отдельными локусами на одной и той же хромосоме. V-области легких цепей формируется путем случайного объединения сегментов ДНК из 2-х наборов — {V} и {J}, а V-области тяжелой цепи — из трех наборов: {V}, {D} и {J}. Генерируемое при этом разнообразие вариантов иммуноглобулинов достигает $\sim 10^7$ – 10^8 молекул [21]. Кодирование такого количества белков без использования комбинаторики потребовало бы увеличения размеров генома до 10^{10} – 10^{11} п.н., то есть в 10–100 раз.

Коды модуляции функций геномов

Одним важнейшим для эукариот механизмом кодирования биологической сложности являются коды модуляции [20]. Один из элементов этого кода — тандемные повторы. Их модулирующее влияние на функции генов проявляется при изменении числа копий в кластере [91]. Изменение числа повторов в кластерах — наиболее частый тип геномного полиморфизма [77]. Пространственный и временной паттерны экспрессии генов, находящи-

ся под влиянием повторов, могут изменяться после изменения копийности повторов [91]. В этом случае поиск оптимальных вариантов модуляторов осуществляется существенно быстрее, чем при точковых мутациях. Элементами кода модуляции геномов являются также диспергированные повторы и МГЭ, инсерции которых в различные районы геномов могут приводить к изменению паттерна экспрессии близкорасположенных генов [32]. Более того, МГЭ способствуют адаптивной пластичности эукариот, увеличивая мутабельность конкретных локусов в определенные периоды времени [6, 66].

ГЕННЫЕ СЕТИ: КОДЫ ФУНКЦИОНИРОВАНИЯ

Генная сеть (ГС) — группа координированно функционирующих генов, контролирующая формирование определенного фенотипического признака организма (молекулярно-генетического, биохимического, физиологического, морфологического, поведенческого и т. д.). В основе функционирования любой генной сети лежит определенный набор элементарных структур (генов, их РНК, белков, а также сигнальных молекул, метаболитов и т. д.) и элементарных событий — взаимодействий меж-

Таблица 2

Роль генов *Drosophila melanogaster* в различных биологических процессах [118]

Ген	Роль гена	
	В индивидуальном развитии	В иных процессах
run	Ранний эмбриогенез, развитие центральной нервной системы, развитие периферической нервной системы	Определение пола
dl	Определение дорзовентральной оси тела, развитие эктодермы, миграция клеток зародышевого пути, развитие эмбрионального сердца, развитие мезодермы, развитие центральной нервной системы	Иммунный ответ, функционирование жирового тела личинки
Tl	Определение дорзовентральной оси тела, развитие мускулатуры	Иммунный ответ
Kr	Ранний эмбриогенез, развитие Мальпигиевых трубок, развитие эмбриональной мышечной системы, развитие центральной нервной системы, развитие амниосерозы, развитие глаза	
sgg	Детерминация сегментной полярности, морфогенез щетинок, развитие сердца, морфогенез крыла	Циркадный ритм
ci	Ранний эмбриогенез, развитие имагинальных дисков, развитие аксонов сетчатки, развитие фолликулярных клеток	Клеточный цикл
Amph	Развитие глаза	Эндоцитоз синаптических пузырьков, регуляция мышечных сокращений
dsh	Развитие женских гонад, овогенез, развитие крыла	Движение клеток

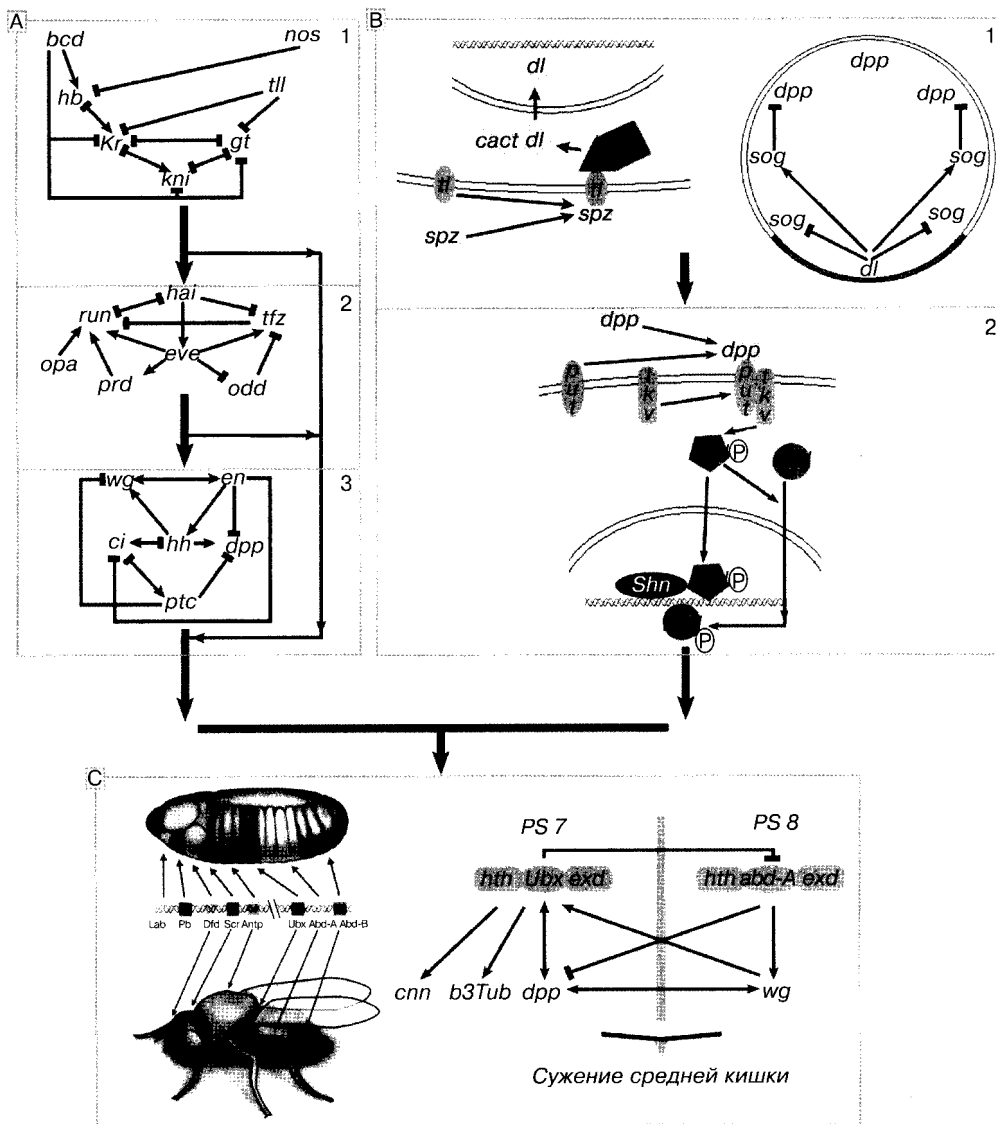


Рис. 3. Общая схема генной сети раннего эмбриогенеза *Drosophila melanogaster* (см. текст).
 А. Подсеть установления антерио-постериорной оси тела и ее спецификации. В. Подсеть установления и спецификации дорзовентральной оси. С. Подсеть морфофункциональной спецификации сегментов тела регулируемая компонентами вышеназванных подсетей (см. А. и В.). На рис. представлена схема генной сети спецификации висцеральной мезодермы 7-го и 8-го парасегментов (PS)

ду элементарными структурами (биохимических реакций, регуляторных, транспортных событий и т. д.) [8, 69].

Каждая эукариотическая генная сеть имеет [111]: 1) определенный набор путей передачи сигналов от клеточной мембраны до конкретных генов; 2) центральный регулятор — ТФ, обеспечивающий каскадную активацию большой группы генов, одновременная экспрессия которых необходима для выполнения функции генной сети; 3) определенный набор регуляторных механизмов, положительных или отрицательных обратных связей (ПОС и ООС соответственно), обеспечивающих поддержание определенного показателя ГС около некоторого оптимального уровня, или, напротив, эффективно отклоняющих контролируемый показатель ГС от начального значения.

Выделяют 4 основных класса ГС [111]: (i) ГС гомеостаза (в них преобладают ООС); (ii) ГС циклических процессов (имеется баланс между ПОС и ООС); (iii) ГС стрессового ответа (ПОС и ряд других активирующих механизмов уводят ГС от исходного состояния, а ООС и различные репрессирующие механизмы возвращают ее к состоянию покоя); (iv) ГС морфогенеза (в них важнейшую роль играют положительные обратные связи).

Режим функционирования ГС в организме (выполняемая ею программа) определяется [12, 67, 111]: 1) набором элементарных структур (генов, мРНК, белков, метаболитов и т. д.) — переменных этой ГС; 2) совокупностью элементарных процессов — взаимодействий между элементарными структурами; 3) константами элементарных про-

цессов; 4) графом ГС, вершины которого соответствуют элементарным структурам, а ребра — элементарным процессам; 5) начальным состоянием переменных этой ГС; 6) состоянием критических параметров внешней и внутренней среды организма, в которой функционирует ГС.

У многоклеточных эукариот существуют многочисленные механизмы комбинаторного кодирования сложности генетических программ, выполняемых генными сетями. Например, один и тот же ген (или группа генов) могут входить в состав нескольких ГС, работающих одновременно или в различных функциональных состояниях организма (табл. 2). Одна и та же ГС в зависимости от начальных данных функционирования (концентраций мРНК, белков, метаболитов и т. д.), определяемых ранее функционировавшими ГС, может иметь качественно отличающиеся режимы функционирования, например циклический и стационарный. Появление или исчезновение лишь одной регуляторной связи в ГС может привести к качественному изменению режима ее функционирования [12]. Подобного рода изменения, по-видимому, широко распространены в природе, затрагивая в основном регуляторные районы генов, и могут играть большую роль в морфологической эволюции [80, 101]. Возможно, мутации, изменяющие динамику ГС и/или комбинаторику различных ГС редуцировали число жилок в крыле двукрылых. Например, известны мутанты *D. melanogaster*, жилкование крыла которых сходно с жилкованием древних насекомых, то есть ГС предкового жилкования сохраняется. Значит, жилкообразование у насекомых, имеющих, в отличие предков, меньшее число жилок, должно ингибироваться в определенных местах крыла [28, 54]¹.

Локальные ГС эукариот могут объединяться в более сложные генные ансамбли. Наиболее ярким примером этого является ансамбль ГС эмбриогенеза. Например, ГС раннего эмбриогенеза *D. melanogaster* подразделяется на 2 большие по числу компонентов взаимосвязанных подсети (рис. 3): (А) подсеть установления антерио-постериорной оси тела и ее спецификации и (В) подсеть установления и спецификации дорзовентральной оси. Первая подсеть (рис. 3А) включает три ГС. ГС установления антерио-постериорной оси тела и начала сегментации (А1) обеспечивает разделение тела вдоль антерио-постериорной оси на большие участки. Эта ГС гибридна, образована гар-генами зиготы и генами материнского эффекта *bcd* и *nos*, экспрессирующимися в организме самки, продукты которых активно транспортируются в ооцит. ГС начальных этапов сегментации (А2) обеспечивает разделе-

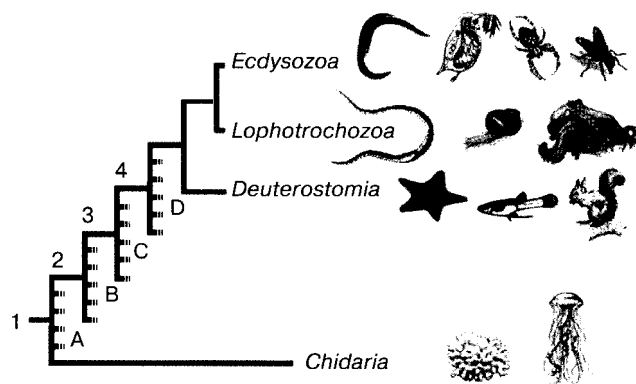


Рис. 4. Эволюция механизмов эмбриогенеза.

Цифрами обозначены основные ароморфозы (см. текст), буквами — основные адаптивные радиации (идеоадаптации), следовавшие за ароморфозами: А) радиация молекулярных механизмов клеточной дифференцировки; В) радиация молекулярных механизмов эмбриогенеза; С) радиация механизмов взаимодействия генных сетей дифференцировки и регуляции клеточного цикла; D) радиация механизмов детерминации структур взрослой формы. Эволюционные взаимоотношения между группами Bilateria реконструированы на основе 18S рРНК [96]

ние тела на «четные» и «нечетные» сегменты генами парного правила. Начальные условия функционирования этой ГС задаются предыдущей геной сетью — диффузией продуктов гар-генов. ГС конечных этапов сегментации (А3) включает в себя генами сегментной полярности и отвечает за образование истинных сегментов тела. Гены сегментации активируют или ингибируют экспрессию Нох-генов, морфофункционально специфицирующих сегменты тела (рис. 3С) [93]. Вторая подсеть (рис. 3В) включает две ГС. ГС определения вентральной стороны яйца (В1) обеспечивается dl-каскадом сигналов; dl-каскад запускается из перивителлинового пространства. Он ингибирует *dpp*-каскад, функционирующий на более поздних стадиях развития [102]. ГС дифференцировки тканей дорзальной стороны эмбриона (В2) определяется *dpp*-каскадом. Морфоген *dpp* отвечает за активацию или ингибирование как генов, определяющих дифференцировку клеток по дорсальному типу, так и Нох-генов в отдельных тканях и органах (рис. 3С) [110]. Вся ГС раннего эмбриогенеза *D. melanogaster* имеет каскадный характер (обеспечивающийся за счет ПОС) с небольшим числом компонентов и связей на входе, и множеством взаимосвязанных компонентов, организованных в процессы, запускающие морфогенез конкретных структур — на выходе [30].

В зависимости от функционального состояния организма одна и та же локальная ГС может входить в состав нескольких ГС иерархически более высокого уровня, что обеспечивает дополнительные возможности для комбинаторного кодирования сложности. Например, геновая сеть апоптоза функционирует в составе ГС онтогенеза [89], ГС сети иммунного [71] и противоопухолевого ответа [79] и

¹ Другой яркий пример — крылатые и бескрылые расы у муравьев. Генетические исследования показали, что в разных филумах муравьев прерывание работы ГС формирования крыла возникло независимо и происходит на разных этапах развития. Замечательно, что муравьи «научились» в ходе эволюции управлять своей ГС, регулируя соотношение крылатых и бескрылых форм в зависимости от ситуации в муравейнике [23].

множестве других. Объединение локальных ГС в более сложные функциональные ансамбли осуществляется с участием особого класса ГС — интеграторов [19, 111], которые в зависимости от стадии развития, ткани и функционального состояния организма могут подключать к выполнению программы те или иные наборы локальных ГС. Например, Нох-гены являются центральными регуляторами многих ГС-интеграторов морфогенеза, работа которых контролируется также компонентами сигнальных каскадов и тканеспецифичными ТФ. В итоге формируется сложный ансамбль ГС, контролируемый обратными связями, разделенными во времени и пространстве [60]¹.

Наиболее высокому уровню интеграции соответствует глобальная генная сеть (ГГС) многоклеточного организма, имеющая иерархически-сетевую организацию и образованная большим количеством взаимосвязанных функциональных ансамблей ГС [1, 111]. Взаимодействия между ГС и их ансамблями могут реализовываться через регуляторные контуры с положительными и отрицательными обратными связями на основе систем внутриклеточных и межклеточных коммуникаций и иметь как активирующий, так и репрессирующий характер. В пределах каждого такого ансамбля могут реализовываться сложные сценарии соподчинения локальных ГС. Например, ГС базового клеточного метаболизма зависят от функции ГС клеточного цикла, которые, в свою очередь, могут контролироваться сложными программами клеточной дифференцировки, а те — ГС онтогенеза и т. д. [1, 111]. ГС иерархически подчиненного уровня могут, в свою очередь, оказывать управляющие воздействия на функцию ГС иерархически высокого уровня. Например, в эмбриональном развитии ГС клеточного цикла взаимодействуют с ГС клеточной дифференцировки. Более того, рассогласование этих ГС во времени и пространстве в ходе развития организма может привести к радикальным изменениям морфологии организма, времени наступления половой зрелости (неотении) и т. д. [11].

Сложность ГГС увеличивалась в ходе эволюции. Особенно заметно увеличение сложности ГГС в эволюции эукариот. Онтогенез, существовавший у протистов [22], дополняется у многоклеточных эмбриональным развитием (эмбриогенезом) — сжатым во времени этапом онтогенеза, отмеченным резкими изменениями в организации тела: от одной клетки зиготы до тела личинки или взрослой особи [10]. Эволюцию эмбриогенеза можно представить как ряд ароморфозов, приводившим к качественно новым особенностям эмбрионального развития, за которыми следовали идиоадаптации, приводившие к разнообразию вариантов онтогенеза на основе появившегося свойства. Главными ароморфозами на пути от примитивных многоклеточных эукариот до современных сложнейших билатеральных животных (*Bilateria*) являются [43, 51, 95] (рис. 4): 1) возникновение ГС клеточной дифференцировки; 2) возникновение ГС эмбриогенеза с

детерминацией бластомеров сразу после оплодотворения (тип 1 эмбриогенеза, характерный для личинок примитивных первично- и вторичноротых); 3) интеграция ГС пролиферации и дифференцировки, и, как следствие, появление клеток, способных к дифференциации в ткани и органы после завершения примитивного эмбриогенеза; 4) появление ГС-интеграторов морфогенеза — молекулярных механизмов специализации таких клеток (например, Нох-гены). Таким образом, каждый из этих ароморфозов сопровождался появлением нового иерархического уровня организации ГС онтогенеза. С появлением эмбриогенеза, сходного с таковым у современных *Bilateria* (~700 млн лет назад [26]), темпы морфологической эволюции резко возросли [51] вероятно вследствие мощнейшего роста возможностей комбинаторики ГС эмбриогенеза и появления многих новых иерархических уровней ГС.

ЗАКЛЮЧЕНИЕ

Хотя у прокариот существуют механизмы наращивания комбинаторной сложности режимов функционирования генных сетей, они ограничены их более примитивной организацией: 1) оперонной структурой геномов, резко снижающей возможности независимой экспрессии индивидуальных генов; 2) малым размером регуляторных районов оперонов (генов), снижающим потенциал комбинаторной регуляции транскрипции; 3) примитивным строением обратных связей, замыкаемых, как правило, при участии конечных продуктов функционирования оперонов на уровне внутри- и межклеточных путей метаболизма; 4) примитивностью межклеточных коммуникаций.

Что касается эукариот, то у них комбинаторные механизмы кодирования генетических программ имеют качественно более высокий уровень сложности, обусловленный: 1) существенно большим количеством генов; 2) сложным строением их регуляторных районов; 3) возможностью кодирования комбинаторики генетических программ на уровне транскрипции; 4) наличием сплайсинга — качественно нового генетического механизма кодирования сложности; 5) распределенностью компонент генных сетей у эукариот по множествам компарментов, в том числе (в случае многоклеточных) принадлежащих различным клеткам, тканям и органам; 6) качественно более сложными и разнообразными механизмами межклеточных коммуникаций; 7) огромным разнообразием механизмов перепрограммирования генных сетей и их интеграции в ансамбли более высокого уровня сложности. Генные сети эукариотических организмов организованы таким образом, что позволяют реализовать огромное множество разнообразных генетических программ формирования фенотипических признаков организмов.

¹ Ярким примером такого комплексного взаимодействия является молекулярный механизм дифференцировки сердечной трубки *D. melanogaster* [100].

Неудивительно, что корреляция между числом генов и морфологической сложностью у многоклеточных эукариот не соблюдается, так как имеющегося разнообразия генных сетей, режимов их функционирования, взаимодействия и интеграции вполне достаточно для выполнения генетических программ, контролирующих сколь угодно сложные фенотипические признаки организмов.

Выражаем благодарность за ценную помощь в обсуждении статьи Левицкому В.Г., Подколотной О.А., Степаненко И.Л., Омелянчуку Л.В., Орлову Ю.Л.

Работа поддержана следующими грантами: РФФИ 03-04-48506-а, РФФИ 03-01-00328, интеграционным проектом СО РАН № 119, интеграционным проектом СО РАН № 145, интеграционным проектом СО РАН № 148, проектом «Описание и анализ биоразнообразия динамики экосистем Сибири с использованием информационных технологий», программы РАН по биоразнообразию (12.4) и проектом «Компьютерное моделирование и экспериментальное конструирование генных сетей», программы РАН по физико-химической биологии (10.4).

Литература

1. *Ананько Е.А., Игнатъева Е.В., Недосекина Е.А.* и др. Классификация генных сетей на основе информации базы данных GeneNet // Конференция, посвященная 90-летию со дня рождения А. А. Ляпунова: Тез. докл. — Новосибирск, Академгородок, 2001. <http://www.ict.nsc.ru/ws/Lyap2001/2528/>
2. *Бердников В.А., Родин С.Н., Жарких А.А.* Эволюция величины генома на основе неравного кроссинговера // Докл. АН СССР. — 1982. — Т. 263, № 2. — С. 464–467.
3. *Галимов Э.М.* Феномен жизни. — М.: Изд-во УРСС, 2001.
4. *Заварзин Г.А.* Проблемы доантропогенной эволюции биосферы // Развитие микробных сообществ в истории Земли. — М.: Наука. — 1993. — С. 212–222.
5. *Заварзин Г.А.* Экосистемные перестройки и эволюция биосферы // Реликтовые прокариотные сообщества гипергалинных водослов морского происхождения. — Вып. 1. — М.: Недра, — 1994. — С. 318–325.
6. *Захаров И.К., Иванников А.В., Юрченко Н.Н.* Динамика мутационного процесса и генофонд природных популяций *Drosophila melanogaster* // Современные концепции эволюционной генетики / Ред.: В.К. Шумный, А.Л. Маркель. — Новосибирск: ИЦиГ СО РАН. — 2000. — С. 151–159.
7. *Колмогоров А.Н.* Избранные труды. Математика и механика. — М.: Наука, 1985.
8. *Колчанов Н.А., Ананько Е.А., Колпаков Ф.А.* и др. Генные сети // Мол. биология. — 2000. — Т. 34, № 4. — С. 617–629.
9. *Колчанов Н.А., Сулов В.В., Шумный В. К.* Молекулярная эволюция генетических систем // Палеонтологический журнал. — 2003. — Т. 37, № 6. — С. 617–629.
10. *Короткова Г.П.* Происхождение и эволюция онтогенеза. — Л.: Издательство ЛГУ, 1979.
11. *Корочкин Л.И.* Введение в генетику развития. — М.: Наука, 1999.
12. *Лихошвай В.А., Матушкин Ю.Г., Фадеев С.И.* О связи графа генной сети с качественными режимами ее функционирования // Мол. биология. — 2001. — Т. 35, № 6. — С. 1080–1087.
13. *Опарин А.И.* Жизнь, ее природа, происхождение и развитие. — М.: Наука, 1968.
14. *Пармон В.Н.* Физико-химические движущие силы и направление естественного отбора и эволюции пребиотических автокаталитических систем // Журн. физ. химии. — 2002. — Т. 76, № 1. — С. 149–158.
15. *Прозоров А.А.* Альтруизм в мире бактерий? // Усп. совр. биол. — 2002. — Т. 122, № 5. — С. 403–413.
16. *Ратнер В.А.* Генетика. Молекулярная кибернетика. Личности и проблемы. — Новосибирск: Наука, 2002. — С. 104–121.
17. *Ратнер В.А., Жарких А.А., Колчанов Н.А.* и др. Проблемы теории молекулярной эволюции. — Новосибирск: Наука, 1985.
18. *Соловьев В.В., Колчанов Н.А.* Экзон-интронная структура генов эукариот может быть обусловлена нуклеосомной организацией хроматина и связанными с ней особенностями регуляции экспрессии генов // Докл. АН СССР. — 1985. — Т. 284, № 1. — С. 232–237.
19. *Степаненко И.Л.* Регуляция генных сетей стрессового ответа активными формами кислорода // Экологическая генетика. — 2004. (в печати)
20. *Трифонов Э.Н.* Генетическое содержание последовательностей ДНК определяется суперпозицией многих кодов // Мол. Биология. — 1997. — Т. 31, № 4. — Р. 759–767.
21. *Фогель Ф., Мотульски А.* Генетика человека. — М.: Мир, 1990.
22. *Хаусман К.* Протозология. — М.: Мир, 1988.
23. *Abouheif E., Wray G.A.* Evolution of the gene network underlying wing polyphenism in ants // Science. — 2002. — Vol. 297, N 5579. — P. 249–252.
24. *Alberts B., Johnson A., Lewis J. et al.* Molecular Biology of the Cell, Fourth Edition. New York: Garland Science, 2002.
25. *Ameisen J.C.* On the origin, evolution, and nature of programmed cell death: a timeline of four billion years // Cell Death Differ. — 2002. — Vol. 9, N 4. — P. 367–393.
26. *Ayala F.J., Rzhetsky A., Ayala F.J.* Origin of the metazoan phyla: molecular clocks confirm paleontological estimates // Proc. Natl. Acad. Sci. U.S.A. — 1998. — Vol. 95, N 2. — P. 606–611.
27. *Bell G.* The sexual nature of the eukaryote genome // Journal of Heredity. — 1993. — Vol. 84, N 5. — P. 351–359.
28. *Biehs B., Sturtevant M.A., Bier E.* Boundaries in the *Drosophila* wing imaginal disc organize vein-specific genetic programs // Development. — 1998. — Vol. 125, N 21. — P. 4245–4257.
29. *Black D.L.* Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology // Cell. — 2000. — Vol. 103, N 3. — P. 367–370.
30. *Bodnar J.W.* Programming the *Drosophila* embryo // J. Theor. Biol. — 1997. — Vol. 188, N 4. — P. 391–445.
31. *Bonifer C.* Long-distance chromatin mechanisms controlling tissue-specific gene locus activation // Gene. — 1999. — Vol. 238, N 2. — P. 277–289.
32. *Brosius J.* Genomes were forged by massive bombardments with retroelements and retrosequences // Genetica. — 1999. — Vol. 107, N 1–3. — P. 209–238.
33. *Bulger M., Groudine M.* Looping versus linking: toward a model for long-distance gene activation // Genes & Development. — 1999. — Vol. 13, N 19. — P. 2465–2477.
34. *Carroll S.B.* Chance and necessity; the evolution of morphological complexity and diversity // Nature. — 2001. — Vol. 409, N 6823. — P. 1102–1109.
35. *Cavalier-Smith T.* Obccells as proto-organisms: membrane heredity, lithophosphorylation, and the origins of the genetic code, the first cells, and photosynthesis // J. Mol. Evol. — 2001. — Vol. 53, N 4–5. — P. 555–595.
36. *Cavalier-Smith T.* Origins of the machinery of recombination and sex // Heredity. — 2002. — Vol. 88, N 2. — P. 125–141.
37. *Cavalier-Smith T.* The neomuran origin of archaebacteria, the eubacterial root of the universal tree and bacterial mega-classification // Int. J. Syst. Evol. Microbiol. — 2002. — Vol. 52. — Pt. 1. — P. 7–76.
38. *Computer Analysis of Genetic Macromolecules: structure, function and evolution / Eds. N.A. Kolchanov, H.A. Lim.* Singapore, etc.: World Scientific, 1994.
39. *Cosma M.P.* Ordered recruitment: gene-specific mechanism of transcription activation // Molecular Cell. — 2002. — Vol. 10, N 2. — P. 227–236.

40. *Crespi B.J.* The evolution of social behavior in microorganisms // *Trends Ecol. Evol.* — 2001. — Vol. 16, N 4. — P. 178–183.
41. *Croft L.G., Lercher M.J., Gagen M.J., Mattick J.S.* Is prokaryotic complexity limited by accelerated growth in regulatory overhead? // *Genome Biol.* — 2003. — Vol. 5, N 1. — P. 2. — Epub 2003 Dec 15.
42. *Csordas A.* A proposal for a possible role of nucleosome positioning in the evolutionary adjustment of introns // *Int. J. Biochem.* — 1989. — Vol. 21, N 5. — P. 455–461.
43. *Davidson E.H., Peterson K.J., Cameron R.A.* Origin of adult bilaterian body plans: Evolution of developmental regulatory mechanisms // *Science.* — 1995. — Vol. 270, N 5240. — P. 1319–1325.
44. *De Souza S. J., Long M., Schoenbach L. et al.* Intron positions correlate with module boundaries in ancient proteins // *Proc. Natl. Acad. Sci. USA.* — 1996. — Vol. 93, N 25. — P. 14632–14636.
45. *Denisov D.A., Shpigelman E.S., Trifonov E.N.* Protective nucleosome centering at splice sites as suggested by sequence-directed mapping of the nucleosomes. *Genes.* — 1997. — Vol. 205, N 1–2. — P. 145–149.
46. *Doudna J.A., Cech T.R.* The chemical repertoire of natural ribozymes // *Nature.* — 2002. — Vol. 418. — № 6894. — P. 222–228.
47. *Dworkin M.* Recent advances in the social and developmental biology of the myxobacteria // *Microbiol. Rev.* — 1996. — Vol. 60, N 1. — P. 70–102.
48. *Eigen M.* Selforganization of matter and the evolution of biological macromolecules // *Naturwissenschaften.* — 1971. — B. 58. — N 10. — P. 465–523.
49. *El-Shehawey R., Kleiner D.* The mystique of irreversibility in cyanobacterial heterocyst formation: parallels to differentiation and senescence in eukaryotic cells // *Physiol. Plantarum.* — 2003. — Vol. 119, N 1. — P. 49–55.
50. *Engelberg-Kulka H., Glaser G.* Addiction modules and programmed cell death and antideath in bacterial cultures // *Annu. Rev. Microbiol.* — 1999. — Vol. 53. — P. 43–70.
51. *Erwin D.H., Davidson E.H.* The last common bilaterian ancestor // *Development.* — 2002. — Vol. 129, N 13. — P. 3021–3032.
52. *Fedorova L., Fedorov A.* Introns in gene evolution // *Genetica.* — 2003. — Vol. 118, N 2. — P. 123–131.
53. *Franch T., Gerdes K.* Programmed cell death in bacteria: translational repression by mRNA end-pairing // *Mol. Microbiol.* — 1996. — Vol. 21, N 5. — P. 1049–1060.
54. *Garcia-Bellido A., de Celis J.F.* Developmental genetics of the venation pattern of *Drosophila* // *Annu. Rev. Genet.* — 1992. — Vol. 26. — P. 277–304.
55. *Gilbert W., De Souza S.J., Long M.* Origin of Genes // *Proc. Natl. Acad. Sci. USA.* — 1997. — Vol. 94, N 15. — P. 7698–7703.
56. *Gopalan V., Tan T.W., Lee B.T.K., Ranganathan S.* Xpro: database of eukaryotic protein-encoding genes // *Nucleic Acids Res.* — 2004. — Vol. 32. — Database issue. — P. D59–D63.
57. *Graveley B.R.* Alternative splicing: increasing diversity in the proteomic world // *Trends Genet.* — 2001. — Vol. 17, N 2. — P. 100–106.
58. *Hermoso A., Aguilar D., Aviles F.X., Querol E.* TrSDB: a proteome database of transcription factors // *Nucleic Acids Res.* — 2004. — Vol. 32. — Database issue. — P. D171–D173.
59. *Holmes A.J., Gillings M.R., Nield B.S. et al.* The gene cassette metagenome is a basic resource for bacterial genome evolution // *Environ. Microbiol.* — 2003. — Vol. 5, N 5. — P. 383–394.
60. *Hombria J. C.-G., Lovegrove B.* Beyond homeosis — Hox function in morphogenesis and organogenesis // *Differentiation.* — 2003. — Vol. 71, N 8. — P. 461–476.
61. *Jain R., Rivera M.C., Moore J.E. et al.* Horizontal gene transfer in microbial genome evolution // *Theor. Popul. Biol.* — 2002. — Vol. 61, N 4. — P. 489–495.
62. *Johnston W.K., Unrau P.J., Lawrence M.S., et al.* RNA-catalyzed RNA polymerization: Accurate and general RNA-templated primer extension // *Science.* — 2001. — Vol. 292, N 5520. — P. 1319–1325.
63. *Joyce G.F.* The antiquity of RNA-based evolution // *Nature.* — 2002. — Vol. 418, N 6894. — P. 214–221.
64. *Karp P.D., Arnaud M., Collado-Vides J. et al.* The E. coli EcoCyc database: No longer just a metabolic pathway database // *American Society for Microbiology News.* — 2004. — Vol. 70, N 1. — P. 25–30.
65. *Kauffman S.A.* The Origins of Order. New York: Oxford University Press, 1993.
66. *Kidwell M.G., Lisch D.R.* Transposable elements and host genome evolution // *Trends in Ecology and Evolution.* — 2000. — Vol. 15, N 3. — P. 95–99.
67. *Kolchanov N.A., Ananko E.A., Likhoshvai V.A. et al.* Gene networks description and modelling in the GeneNet system // Gene regulation and metabolism: post-genomic computational approaches / Eds.: J. Collado-Vides, R. Hofstad. Cambridge, etc.: MIT Press. — 2002. — P. 149–179.
68. *Kolchanov N.A., Ignatieva E.V., Ananko E.A. et al.* Transcription Regulatory Regions Database (TRRD): its status in 2002 // *Nucleic Acids Research.* — 2002. — Vol. 30, N 1. — P. 312–317.
69. *Kolchanov N. A., Nedosekina E. A., Ananko E. A. et al.* GeneNet database: description and modeling of gene networks. // *In Silico Biol.* — 2002. — Vol. 2, N 2. — P. 97–110.
70. *Kornberg R. D., Lorch Y.* Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome // *Cell.* — 1999. — Vol. 98, N 3. — P. 285–294.
71. *Krammer P. H.* CD95's deadly mission in the immune system // *Nature.* — 2000. — Vol. 407, N 6805. — P. 789–795.
72. *Kroos L., Zhang B., Ichikawa H., Yu Y.T.* Control of sigma factor activity during *Bacillus subtilis* sporulation // *Mol. Microbiol.* — 1999. — Vol. 31, N 5. — P. 1285–1294.
73. *Levin B.R., Bergstrom C.T.* Bacteria are different: Observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes // *Proc. Natl. Acad. Sci. USA.* — 2000. — Vol. 97, N 13. — P. 6981–6985.
74. *Levitsky V.G., Podkolodnaya O.A., Kolchanov N.A., Podkolodny N. L.* Nucleosome formation potential of eukaryotic DNA: tools for calculation and promoters analysis // *Bioinformatics.* — 2001. — Vol. 17, N 11. — P. 998–1010.
75. *Lewis K.* Programmed death in bacteria // *Microbiol. Mol. Biol. Rev.* — 2000. — Vol. 64, N 3. — P. 503–514.
76. *Li Q., Peterson K.R., Fang X., Stamatoyanopoulos G.* Locus control regions // *Blood.* — 2002. — Vol. 100, N 9. — P. 3077–3086.
77. *Li Y. C., Korol A. B., Fahima T. et al.* Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. // *Molecular Ecology.* — 2002. — Vol. 11, N 12. — P. 2453–2465.
78. *Long M., Deutsch M., Wang W. et al.* Origin of new genes: evidence from experimental and computational analyses // *Genetica.* — 2003. — Vol. 118, N 2–3. — P. 171–182.
79. *Lowe S.W., Lin A.W.* Apoptosis in cancer // *Carcinogenesis.* — 2000. — Vol. 21, N 3. — P. 485–495.
80. *Ludwig M.Z., Bergman C., Patel N.H., Kreitman M.* Evidence for stabilizing selection in a eukaryotic enhancer element // *Nature.* — 2000. — Vol. 403, N 6769. — P. 564–567.
81. *Lynch M., Conery J.S.* The origins of genome complexity // *Science.* — 2003. — Vol. 302, N 5649. — P. 1401–1404.
82. *Lynn M.E., Bantle J.A., Ownby J.D.* Estimation of gene expression in heterocysts of *Anabaena variabilis* by using DNA-RNA hybridization. // *J. Bacteriol.* — 1986. — Vol. 167, N 3. — P. 940–946.
83. *Mahmoudi T., Verrijzer C.P.* Chromatin silencing and activation by Polycomb and trithorax group proteins // *Oncogene.* — 2001. — Vol. 20, N 24. — P. 3055–3066.
84. *Martin W., Russell M.J.* On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells // *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* — 2003. — Vol. 358, N 1429. — P. 59–85.
85. *Martinez-Castilla L.P., Alvarez-Buylla E.R.* Adaptive evolution in the Arabidopsis MADS-box gene family inferred from its complete resolved phylogeny // *Proc. Natl. Acad. Sci. USA.* — 2003. — Vol. 100, N 23. — P. 13407–13412.
86. *Mattick J.S., Gagen M.J.* The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the

- development of complex organisms // *Mol. Biol. Evol.* — 2001. — Vol. 18, N 9. — P. 1611–1630.
87. *McShea D.W.* The minor transitions in hierarchical evolution and the question of a directional bias // *Journal of Evolutionary Biology.* — 2001. — Vol. 14, N 3. — P. 502–518.
 88. *Meeks J.C., Elhai J., Potts M. et al.* An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium // *Photosyn. Res.* — 2001. — Vol. 70, N 1. — P. 85–106.
 89. *Meier P., Finch A., Evan G.* Apoptosis in development // *Nature.* — 2000. — Vol. 407, N 6805. — P. 796–801.
 90. *Moszer I., Jones L. M., Moreira S. et al.* SubtiList: the reference database for the *Bacillus subtilis* genome // *Nucleic Acids Res.* — 2002. — Vol. 30, N 1. — P. 62–65.
 91. *Nakamura Y., Koyama K., Matsushima M.* VNTR (variable number of tandem repeat) sequences as transcriptional, translational, or functional regulators // *Journal of Human Genetics.* — 1998. — Vol. 43, N 3. — P. 149–152.
 92. *Narlikar G.J., Fan H.-Y., Kingston R.E.* Cooperation between complexes that regulate chromatin structure and transcription // *Cell.* — 2002. — Vol. 108, N 4. — P. 475–487.
 93. *Nasiadka A., Dietrich B.H., Krause H.M.* Anterior-posterior patterning in the *Drosophila* embryo // *Advances in developmental biology and biochemistry*, Vol. 12 / Ed. M. DePamphilis. New York: Elsevier, 2002. — P. 155–204.
 94. *Peri S., Navarro J.D., Amanchy R. et al.* Development of human protein reference database as an initial platform for approaching systems biology in humans // *Genome Research.* — 2003. — Vol. 13, N 10. — P. 2363–2371.
 95. *Peterson K.J., Cameron R.A., Davidson E.H.* Bilateral Origins: significance of new experimental observations // *Dev. Biol.* — 2000. — Vol. 219, N 1. — P. 1–17.
 96. *Peterson K. J., Eernisse D. J.* Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences // *Evol. Dev.* — 2001. — Vol. 3, N 3. — P. 170–205.
 97. *Philippe H., Douady C.J.* Horizontal gene transfer and phylogenetics // *Current Opinion in Microbiology.* — 2003. — Vol. 6, N 5. — P. 498–505.
 98. *Ponomarenko J.V., Furman D.P., Frolov A.S. et al.* ACTIVITY: a database on DNA/RNA sites activity adapted to apply sequence-activity relationships from one system to another // *Nucleic Acids Res.* — 2001. — Vol. 29, N 1. — P. 284–287.
 99. *Ponomarenko J.V., Ponomarenko M.P., Frolov A.S. et al.* Conformational and physicochemical DNA features specific for transcription factor binding sites // *Bioinformatics.* — 1999. — Vol. 15, N 7–8. — P. 654–668.
 100. *Ponzielli R., Astier M., Chartier A. et al.* Heart tube patterning in *Drosophila* requires integration of axial and segmental information provided by the Bithorax Complex genes and hedgehog signaling // *Development.* — 2002. — Vol. 129, N 19. — P. 4509–4521.
 101. *Purugganan M.D.* The molecular population genetics of regulatory genes // *Mol. Ecol.* — 2000. — Vol. 9, N 10. — P. 1451–1461.
 102. *Riechmann V., Ephrussi A.* Axis formation during *Drosophila* oogenesis // *Current Opinion in Genetics & Development.* — 2001. — Vol. 11, N 4. — P. 374–383.
 103. *Rifkin S.A., Kim J., White K.P.* Evolution of gene expression in the *Drosophila melanogaster* subgroup // *Nat. Genet.* — 2003. — Vol. 33, N 2. — P. 138–144.
 104. *Rogozin I.B., Makarova K.S., Natale D.A. et al.* Congruent evolution of different classes of non-coding DNA in prokaryotic genomes // *Nucleic Acids Res.* — 2002. — Vol. 30, N 19. — P. 4264–4271.
 105. *Rubin G.M., Yandell M.D., Wortman J.R. et al.* Comparative genomics of the eukaryotes // *Science.* — 2000. — Vol. 287, N 5461. — P. 2204–2215.
 106. *Sakharkar M., Passetti F., de Souza J.E. et al.* ExInt: an exon intron database // *Nucleic Acids Res.* — 2002. — Vol. 30, N 1. — P. 191–194.
 107. *Satchwell S.C., Drew H.R., Travers A.A.* Sequence periodicities in chicken nucleosome core DNA // *J. Mol. Biol.* — 1986. — Vol. 191, N 4. — P. 659–675.
 108. *Saxonov S., Daizadeh I., Fedorov A., Gilbert W.* EID: the exon-intron database—an exhaustive database of protein-coding intron-containing genes // *Nucleic Acids Res.* — 2000. — Vol. 28, N 1. — P. 185–190.
 109. *Schidlowski M.* A 3,800-million-year isotopic record of life from carbon in sedimentary rocks // *Nature.* — 1988. — Vol. 333, N 6171. — P. 313–318.
 110. *Stathopoulos A., Levine M.* Dorsal gradient networks in the *Drosophila* embryo // *Developmental Biology.* — 2002. — Vol. 246, N 1. — P. 57–67.
 111. *Stepanenko I.L., Podkolodnaya O.A., Kolchanov N.A.* Gene networks: principles of organization and mechanisms of operation and integration // *Proceedings of the third international conference on Bioinformatics of Genome Regulation and Structure.* — 2002. — Vol. 2. — P. 111–115.
 112. *Strobel S.A.* Biological catalysis: Repopulating the RNA World // *Nature.* — 2001. — Vol. 411, N 6841. — P. 1003–1004.
 113. *Szostak J.W., Bartel D.P., Luisi P.L.* Synthesizing life // *Nature.* — 2001. — Vol. 409, N 6818. — P. 387–390.
 114. *Taft R.J., Mattick J.S.* Increasing biological complexity is positively correlated with the relative genome-wide expansion of non-protein-coding DNA sequences // *Genome Biology.* — 2003. — Vol. 5, N 1. P1. Epub 2003 Dec 01
 115. *Tamames J.* Evolution of gene order conservation in prokaryotes // *Genome Biol.* — 2001. — Vol. 2, N 6. Research0020. Epub 2001 Jun 01
 116. *Teichmann S.A., Babu M.M.* Conservation of gene co-regulation in prokaryotes and eukaryotes // *Trends Biotechnol.* — 2002. — Vol. 20, N 10. — P. 407–410.
 117. The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana* // *Nature.* — 2000. — Vol. 408, N 6814. — P. 796–815.
 118. The Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource // *Nucleic Acids Res.* — 2004. — Vol. 32. Database issue. — P. D258–D261.
 119. *Trifonov E.N.* Consensus temporal order of amino acids and evolution of the triplet code // *Gene.* — 2000. — Vol. 261, N 1. — P. 139–151.
 120. *Velicer G.J., Lenski R.E., Kroos L.* Rescue of social motility lost during evolution of *Myxococcus xanthus* in an asocial environment. // *J. Bacteriol.* — 2002. — Vol. 184, N 10. — P. 2719–2727.
 121. *Vellai T., Vida G.* The origin of eukaryotes: the difference between prokaryotic and eukaryotic cells // *Proc. R. Soc. Lond., B, Biol. Sci.* — Vol. 266, N 1428. — P. 1571–1577.
 122. *Wolk C.P.* Prokaryotic Development // *Heterocyst formation in Anabaena* / Eds.: Y. Braun, L. J. Shimkets. Washington: American Society of Microbiology, 2000. — P. 83–103.
 123. *Zhang J.* Evolution by gene duplication: an update // *Trends in Ecology and Evolution.* — 2003. — Vol. 18, N 6. — P. 292–298.
 124. *Zuckermandl E.* Why so many noncoding nucleotides? The eukaryotic genome as an epigenetic machine // *Genetica.* — 2002. — Vol. 115, N 1. — P. 105–129.

V.V. Suslov, K.V. Gunbin, N.A. Kolchanov
Institute of cytology and genetics CO RAS; Novosibirsk State University

✿ THE SUMMARY: Increase in organism complexity is a global trend in evolution. Qualitatively extended complexity in eukaryotes in comparison to prokaryotes is provided by genome organization and genetic program realization. Genetic mechanisms of encoding biological complexity in pro- and eukaryotes are considered: above-triplet codes, combinatorial analysis of genetical blocks and gene network blocks, and their hierarchical interaction.

✿ KEY WORDS: biological complexity, prokaryotes, sequence codes, gene networks, regulatory mechanisms.