

УДК 577.21

## ПОВТОРЫ ДНК КАК ИНСТРУМЕНТ БИОЛОГИЧЕСКОГО РАЗНООБРАЗИЯ

© 2011 г. В. В. Гречко\*

*Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва 119991*

Поступила в редакцию 25.11.2010 г.

Принята к печати 29.12.2010 г.

В свете современных представлений о существовании объемных регуляторных сетей как основы жизнеобеспечения обсуждается роль некодирующих повторяющихся элементов ДНК в клетке, а также в эволюции таксонов. В органичной связи с этой проблемой рассматривается ведущая роль принципа удвоения и умножения любых отрезков ДНК — от нуклеотида до сколь угодно длинных фрагментов в сегментных дупликациях — в возникновении механизмов организации и реконструкций генома, в его взаимодействии с белками на уровне хроматина. Это обеспечивает разнообразие архитектоники генома и хромосом как основных участников морфогенеза. В большой степени пластичность генома и его регуляция определяются структурой и эволюцией тандемных и диспергированных повторов ДНК, их тесным сцеплением с историей развития живых организмов, их участием в соматических и наследуемых процессах, включая эпигенетическую регуляцию за счет самих повторов, их модификаций и их транскриптома. Высказывается мнение, имеющее поддержку в работах других авторов, о бессодержательности и бесполезности использования — для оценки смысла существования и функционирования не кодирующих белки последовательностей — их метафорических определений как “эгоистических”, “мусорных” или “паразитических”. Сформулирована гипотеза, согласно которой эту часть геномной ДНК, напротив, можно рассматривать как основную, “хозяйскую”, действующую субстанцию, некий каркас, на котором, во взаимодействии с которым и для поддержания которого функционирует белково-лигандный аппарат. Собственная молекулярная эволюция каркаса определяет паттерн и эволюцию структурных белков, обеспечивающих разнообразие морфогенеза. Консервативному общему “стержню” процессов жизнедеятельности всех организмов — метаболических, энергетических и обеспечиваемых генами “домашнего хозяйства” — отводится роль в образовании защитной оболочки наследственного вещества, а также в обеспечении ферментативного пути репликации и транскрипции и в эпигенетической регуляции.

**Ключевые слова:** сателлитные повторы, диспергированные повторы, “эгоистические, мусорные ДНК”, принцип повторемости в геноме, морфогенез, биоразнообразие, молекулярная эволюция.

REPEATED DNA SEQUENCES AS AN ENGINE OF BIOLOGICAL DIVERSIFICATION, *by V. V. Grechko\** (Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow 119991 Russia; \*e-mail: grechko@eimb.ru). Several aspects of the functional role of non-protein-coding DNA elements in the cell life and taxa evolution were discussed in connection with modern views on three-dimensional regulatory network as a basis for life support and evolution. In accordance with this problem, the leading role of duplication and multiplication of any DNA fragments (from single nucleotides to segmental duplications) in origination of genome structure and reshaping were considered in relationships with chromatin proteins. By this way the diversities in genome and chromosome architectonics leads to diversity of morphogenesis. The regulation and plasticity of the genome is determined by the structure, plasticity and evolution of genomic satellite and dispersed repetitive elements, which, being tightly bound with life phylogeny, result in somatic and inherited changes. Repeated sequences take part and perform some epigenetic regulation via repeats themselves, their modifications, or via their RNA transcripts. There is a growing evidence that the usage of metaphoric designations of protein-non-coding sequences as “egoistic, junk or parasitic” are senseless and useless. To the contrary, the hypothesis is formulated that the repetitive non-coding DNA part could be considered as a main substrate of life, something like a “carcass” serving as a basis for protein-coding sequences. The own molecular evolution of this carcass defines the pattern of DNA transcription and then evolution of structural proteins and morphological differentiation. Genes involved in main energetic, “house-keeping” and metabolic processes are forming some kind of the protecting envelope for the hereditary material, and also provides the enzymatic pathway of replication, transcription and epigenetic regulation.

**Keywords:** tandem (satellite) repeats, dispersed repeats, “egoistic, junk” DNA, DNA duplication and multiplication, morphogenesis, biodiversity, molecular evolution.

\* Эл. почта: grechko@eimb.ru

*Without denying the obvious existence of an island/sea contrast between coding and noncoding sequences, it will be shown that available evidence does not refute, and increasingly support the view that the sea itself is to a significant extent permeated by function.*

E. Zuckerkandl [1]

## ВВЕДЕНИЕ

Со времени открытия сателлитных и диспергированных повторов в ДНК, когда эти избыточные, как полагали, нефункциональные регионы рассматривались как не необходимые для жизнедеятельности клетки или даже паразитически (или эгоистически) истощающие ресурсы организма для своего размножения, прошло более 40 лет. Вначале геном как бы подразделяли на две ипостаси, одна из которых синтезировала белки, была источником формообразования и поддержания жизни, а вторая, самодостаточная, только отнимала у клетки часть ресурсов для самоподдержания, как истинный вредитель функционирующей части генома. И эта позиция рассматривалась как своего рода парадигма [2, 3].

С течением времени — и анализу этой точки зрения в историческом аспекте посвящен данный обзор — самые смелые теоретические и логические предсказания авторов, которые высказывали противоположную точку зрения, оказались гораздо более продуктивными в оценке возможностей повторов как носителей структурной, кодирующей, регуляторной, эволюционной функции [1, 4–9]. В последние годы эта точка зрения, казалось бы, стала общепринятой, однако даже сейчас ее поддерживают некоторые авторы, по-прежнему использующие в своих докладах и статьях метафорическое описание повторов как мусорных и эгоистических ([10, 11] и другие). И это совсем не так безопасно, как можно полагать, поскольку известно, как идеологические предрассудки могут сдерживать развитие научного знания. Чего стоит, например, следующее высказывание автора обзора данных, представленного на годичном собрании Международного общества молекулярных эволюционистов (ISME) (A. Eyre-Walker [12]). Цитируя работу Карла Шмида, открывшего, что в ответ на стрессовые обстоятельства начинается транскрипция диспергированного повтора типа SINE, рецензент все так же — неспящая — отмечает: *“Эта поразительная картина пытается заставить нас думать, что SINE имеют функцию, однако требуется доказать, что иметь SINE — это преимущество. Мне, лично, трудно вообразить, что SINE и LINE, которые очевидно являются эгоистическими и паразитическими участками ДНК, могут обладать какими-то иными свойствами, чем эти”*. Автор не принимает во внимание и то, что с этимологией понятия “паразитизм” или “мусор” вряд ли согласуется то, что нередко доля не кодиру-

ющих белки регионов в геноме достигает десятков процентов.

Точка зрения, согласно которой не следовало бы увлекаться произвольными метафорическими терминами в научном обиходе, близка автору обзора. Термины сайленсинг, сателлит, орфон, мусор, эгоизм, псевдоген и другие, хотя и способствуют творческому воображению, но отвлекают от попыток строгого, логического и доказательного описания новых явлений или их характеристик. Недавний пример — оценка некодирующей (в основном, повторяющейся части ДНК), как “пустыни” (desert) (Venter et al., 2001, см. ссылку в [13]), оказалась также бессодержательной, если не сказать дезориентирующей, после открытия важнейшей регуляторной роли некодирующей части ее транскриптома (см. обзоры [13–15]).

Благодаря множеству публикаций (не все из которых, конечно, могут быть упомянуты в пределах одного обзора), подтверждающих теснейшее и взаимопроницающее взаимодействие кодирующей и не кодирующей белки частей ДНК как единой, совместно эволюционирующей, функционирующей и обеспечивающей жизнедеятельность системы, представляется возможным и необходимым еще раз рассмотреть некоторые из, на наш взгляд, наиболее убедительных.

\* \* \*

Коснемся эволюции взглядов на роль некодирующей ДНК. Приведенная в эпиграфе философски-биологическая метафора проблемы, которой посвящен обзор, сформулирована несколько осторожно, будучи написанной в 1992 г. Е. Цукеркандль (Zuckerkandl) [1] рассматривает в качестве “некодирующей” только как таковую огромную, часто подавляющую, часть генома эукариот, которая состоит из повторов любого типа. Он не привлекает еще явления “псевдогенизации” структурных генов; он не рассматривает тот факт, что так называемая некодирующая часть ДНК подвергается транскрипции с образованием различных типов РНК; он не учитывает того, что малые РНК участвуют в специфичной регуляции и в эпигенетических процессах, что, само по себе, представляет некий код влияния генома на жизнь клетки. И что поэтому относить повторяющиеся районы ДНК к разряду вообще ничего “не кодирующих” вряд ли разумно, как это и рассматривал Трифонов (Trifonov) [5, 6]. Автор эпиграфа имел в виду не кодирующую БЕЛКИ часть генома. Однако метафорический образ, высказанный Цукеркандлем, очень выразителен, используется другими авторами [15] и полностью применим с позиций, изложенных в данном обзоре.

Однако была четко сформулирована другая точка зрения, в которой введена метафора, используемая и поныне (Дулиттл (Doolittle) и Сапиенца (Sapi-

enza) [3]). Авторы полагают, что естественный отбор внутри генома неизбежно приводит к появлению ДНК, не выражающихся фенотипически, чья единственная функция есть переживание внутри генома. Прокариотические транспозоны и эукариотические среднеповторяющиеся последовательности могут рассматриваться как такие – паразитические или мусорные, или эгоистические – ДНК, и никаких фенотипических или эволюционных функций можно им не приписывать. Этой точке зрения, как известно, соответствуют взгляды Оргелла (Orgell) и Крика (Crick) [2]. Конечно, вряд ли кто-нибудь из современных исследователей, использующих до сих пор эти броские, но по существу бессмысленные и антибиологические метафоры, помнит высказывание дословно и подпишет сейчас под утверждением авторов. И, тем не менее, даже и сейчас, зная огромные достижения последних 15–20 лет в изучении молекулярной биологии повторов и других разделов неструктурной части генома (см., например, [16–18] и другие, рассмотренные далее в тексте), некоторые авторы продолжают использовать эти характеристики, несмотря на соображения, никоим образом не поддерживающие их.

Наиболее систематически эта проблема рассмотрена в обзорах Маттика (Mattick) и соавт. [13–15]. Согласно им, долгое время считавшаяся ортодоксальной точка зрения, приписывающая не только первично структурную, но и главную функциональную и регуляторную роль только белкам (и их комплексам) живой клетки, неверна. Мнение о том, что для осуществления специфической и организующей программы биологического существования необходимы и достаточны только белки, пересматривается и не соответствует действительности [19, 20]. При этом отмечается, что еще в 1975 г. Кинг (King) и Уилсон (Wilson) (см. [19]) сформулировали мнение, что лишь малая часть эволюционных изменений в молекулярных основах жизни обеспечивается кодирующими белки участками генома. Главные эволюционные изменения основаны на механизмах, контролирующих экспрессию генов, и на некодирующих последовательностях. (Эта точка зрения вынесена в заголовок следующего раздела обзора). Эволюция не может определяться только все более усложняющимся репертуаром белков, но должна зависеть в большой степени от набора инструкций от некодирующей части генома, подавляющая часть которой состоит из повторов разного типа.

Возникло представление о регуляторной сети, связывающей воедино миллионы обменных и структурных взаимосвязей живого организма и опосредуемой, в основном, не кодирующими белки (далее, для краткости, – “небелковые гены”) повторами ДНК [6, 21]. Именно регуляторная сеть, ее природа, архитектоника и вариативность могут обеспечивать невероятное разнообразие индивидуальных и таксонных различий. Она в той или иной

степени присутствует у всех эукариот [22] и прокариот (см. [23] и ссылки там). Представления о возможной регулирующей роли повторов ДНК в индивидуальном развитии клетки сформулированы А. Оловниковым уже в 1996 г. в детально разработанной гипотезе [24], в которой повторяющейся части ДНК отводится исключительная по значению роль “локационной”, направляющей перевод линейной информации ДНК в трехмерную форму организма.

Более умеренная позиция сформулирована Свердловым [25], который полагает, что современная парадигма эволюции должна включать примерно равноважные составляющие: биохимическую эволюцию (структура и функция белков), регуляторную компоненту (эволюция регуляторных сетей и, в том числе, повторов) и эволюцию типа “less-is-more” [26] (возникновение новых функций у паралогов белков, о чем будет сказано далее). Автору обзора более близка точка зрения, согласно которой главным компонентом эволюции жизнедеятельности являются регуляторные сети, основное действующее начало которых – некодирующие белки регионов, а также общий (и, видимо, всеобщий) принцип дуплицируемости любого участка ДНК.

Цель настоящего обзора – попытаться из поистине безбрежного моря исследований показать наиболее характерные, подтверждающие точку зрения, которая предполагает возможность рассматривать небелковые повторы как основную силу эволюции и морфообразования.

### **ВСЕ ФОРМЫ ЖИЗНИ ОБЪЕДИНЯЮТСЯ НА ОСНОВЕ ПРИНЦИПА КОНСЕРВАТИВНОСТИ СТРУКТУРНЫХ ГЕНОВ БЕЛКОВ ОСНОВНОГО ОБМЕНА И ДИВЕРСИФИЦИРУЮТСЯ НА ОСНОВЕ ПРИНЦИПА ВАРИАбельНОСТИ НЕКОДИРУЮЩЕЙ ЧАСТИ ДНК**

На каком же основании могла возникнуть и долгое время преобладать упомянутая вначале парадигма? На наш взгляд, на априорном и казавшемся ранее логическим представлении о том, что исключительную роль в морфообразовании играют белки и их комплексы и, соответственно, гены, кодирующие эти белки. Эволюция живого рассматривалась – и не только в стране, почитавшей Энгельса, – лишь как форма жизни белков. То простое и логическое соображение, что основные законы энергетического и синтетического обмена и законы функционирования аппарата синтеза белков у всех прокариот и эукариот практически одинаковы в качестве единого принципа жизнедеятельности, а поэтому менее вероятны в роли субстрата для возникновения био-разнообразия, не рассматривалось и начинает завоевывать позиции только сейчас. Совокупность ортологов всех важнейших для обмена структурных генов (метагенов по терминологии Дж. Стюарта

(Stuart) [27]) “нанизывает” весь живой мир на некий общий стержень, для которого фантастическое разнообразие форм, в котором он существует, вообще говоря, не обязательно. Автор полагает, что действуют силы отбора общего соотношения, дизайна генетических путей, что необходимо для поддержания работы всего механизма за счет высококонсервативных генов [27]. Это — закон для всего живого, в каких бы морфологических формах оно ни существовало.

Отсюда следует вывод, что возникновение разнообразия живого вещества вряд ли связано единственно со структурными генами (хотя их значение неоспоримо, см. обзор [28]). Во всех отделах живого мира (человек, дрожофила, нематода, дрожжи, растения и бактерии) коэкспрессируются 22 163 структурных ортологичных гена [27]. Они “продержались” в достаточно — принципиально — сходном виде всю историю про- и эукариот, что свидетельствует и об их общем корне, получившем селективное преимущество, и о том, что не они являются главными “ответчиками” за неисчислимо морфологическое разнообразие живого. Анализируя базу данных профилей экспрессии шести основных таксонов, Бергман (Bergman) и соавт. [29] подчеркивают, что коэкспрессия генов со сходными функциями у всех групп про- и эукариот консервативна, особенно это относится к генам метаболических путей. Некоторые коэкспрессионные модули одинаковы у всех названных представителей, в других случаях они тесно коррелируют (например, процессинг рибосомной РНК), хотя имеются и уникальные свойства, формирующиеся за счет случайных мутаций. Тем не менее, в целом генные сети структурных белков выглядят сходно, несмотря на различия в поведении отдельных генных сетей [30].

Точка зрения на эволюцию живого как на взаимодействие сил, сохраняющих основные консервативные свойства, с пластичностью архитектуры генома, его склонностью к перестройкам и с разнообразием эпигенетических механизмов регуляции [19, 20, 31] не вызывает уже серьезных возражений.

Что же определяет пластичность хромосом и генома?

### *Поиски причин (основ) биологического разнообразия*

В этих поисках обратили внимание на фракцию повторяющихся ДНК, поскольку с самого начала и до наших дней находят тесное сцепление между разнообразием повторов и таксономическим и эволюционным положением таксонов (см. обзоры [32–34]). Как полагали вначале, явление повторяемости присуще только некоторым участкам ДНК, не кодирующим белки, и отражает их неконтролируемое и ненужное для клетки размножение. И только изредка, как бы исключительно и в неболь-

шом числе копий, оно функционально оправдано и наблюдается у структурных (типа иммуноглобулинов, гистонов или глобинов) и рибосомных генов в случае непреодолимой потребности организма в увеличении дозы генов для конкретных функций организма. Однако выяснилось и сейчас общепринято, что повторяемость, с образованием множества одинаковых или принципиально сходных элементов, есть общий принцип эволюции генома. Это касается в равной степени и структурных генов: ~5% генома человека участвует в сегментных дупликациях, около 35% белковых локусов амплифицировано (см. обзоры [25, 35, 36]).

Имеются и весьма активно функционируют универсальные ферментативные механизмы, осуществляющие эти удвоения и амплификации участков генома, любых — в широчайшем диапазоне длин умножающихся участков во всех отделах живого мира — от 1–2 н. до сколь угодно длинных фрагментов ДНК (сегментные дупликации) и целых геномов ([25], см. также [37]). Тандемизации могут подвергаться и участки ДНК, содержащие повторы типа SINE, как это наблюдается у некоторых саламандр Европы (но не у американских видов), тритонов и ксенопуса [38].

Постепенно стало понятно, что появление сегментных дупликаций также определяется сложной “игрой” различных механизмов, зависит от неких неравномерных событий эволюции, что привело к нескольким волнам этих дупликаций (у приматов, по крайней мере [37]). Наблюдается очевидная взаимосвязь между дупликациями, геномной нестабильностью и крупными хромосомными перестройками. Сегментные дупликации не только способны приводить к появлению новых генных семейств, что может влиять на современные генные и фенотипические вариации приматов, но и удельный вес таких изменений так значителен, как и не представляли до сих пор [37]. Это в равной степени относится к геному любых эукариот [39].

Эти механизмы дупликации, по-видимому, не отдают “предпочтения” каким-либо фрагментам с разной степенью функциональной значимости, и сохранение или уничтожение дубликатов предоставляется силам отбора. Умножающийся фрагмент скорее всего проходит стадию удвоения, если используются предполагаемые ныне механизмы удвоения ДНК (гомологической и эктопической рекомбинации, генной конверсии или другие [40–42]).

Амплификация может происходить с использованием механизма катящегося кольца (найлены кольцевые формы сателлитной ДНК [43]). При использовании механизма обратной транскрипции и обратного встраивания кДНК в геном возможна взрывная амплификация псевдогена (в широком смысле этого слова — как варианта гена).

Поскольку такие механизмы работают на сохранение и расширение спектра этого явления от про-

кариот к эукариотам, то, придерживаясь парадигмы “мусорной” ДНК, следовало бы признать, что вся эволюция генома за счет амплификации направлена на выработку и умножение так называемой “паразитической” ДНК. В эту категорию пришлось бы зачислить вполне функционально важные не только структурные, но и повторяющиеся регуляторные участки ДНК, например, амплифицированные сайты связывания факторов транскрипции в регуляторных модулях, *цис*-расположенных, и часто на огромных расстояниях, по отношению к структурному гену [44, 45]. Эти сайты факторов транскрипции оказались к тому же очень консервативными: они относительно мало изменились за последние 450 млн. лет (около 75% сходства), но тандемизация наблюдается только у позвоночных (т.е. эволюционно более молодых организмов). Следы дубликации прослеживаются более 500 млн. лет.

Стало очевидным, что эволюционное значение явления повторяемости (избыточности) должно иметь большое значение [39]. В последние годы предпринято теоретическое изучение различий в структурных паралогах на филогенетическом уровне в попытках осмысления (путем математического моделирования) последствий возникновения паралогов и избыточности генов вообще (см. обзор [46] и ссылки в нем). Основное заключение: естественный отбор не может просто поддерживать некий уровень “избыточности” участков, степень избыточности должна увеличиваться.

Сравнение сходства последовательностей паралогов, представленных в банке данных, между собой и с ортологом того же гена показало, что в ряде случаев ортолог и его паралог в другом таксоне более сходны, чем другие паралоги этого таксона между собой. По-видимому, имеется линия сохранения исходной функции ортолога в одном из них в эволюции и псевдогенизация (т.е. потеря или перестройка функции) – в других [47, 48]. Причем удвоимость (или большие степени амплификации) генов с тенденцией сохранять функцию дубликата увеличивается по мере усложнения организма [49].

Структурно обоснована гипотеза, постулирующая, что между ранними хордовыми и позвоночными произошли две полноразмерные дубликации генома, в результате чего, естественно, увеличилось число наборов генов. Дубликации генов характерны для перехода от про- к эукариотам. Примеры можно было бы умножить. Следы дубликации, характерной для рыб, прослеживаются и у человека на фоне других перестроек генома [35].

Появление дубликатов-паралогов генов далеко не всегда и не неизбежно приводит к полностью нефункциональному состоянию истинной псевдогенизации (“испорченный” ген). Многие “псевдогенизированные” гены, например, некоторые паралоги так называемых “кассетных генов” (в том числе гены белков с “цинковыми пальцами” –

*ZNF*-гены и другие) сохраняют и диверсифицируют свои функциональные свойства [42]. Показано, что они экспрессируются тканеспецифично [50, 51]. Паралоги одного, функционирующего в настоящее время, гена могут увеличивать его потенциал (как в случае амплификации генов гистонов или иммуноглобулинов) или приобретать дополнительную или вовсе новую функцию. Само это явление может иметь равновероятно положительное или отрицательное значение (вероятность зарегистрировать второе – больше [47]). Даже потерявшие или модифицировавшие свою функцию псевдогены могут начать работать на “другом поле”.

Например, у человека имеется большое число непроцессированных генов, так называемых псевдогенов, исходные функции которых были утрачены после разделения клады “человек–шимпанзе” [52]. Установлено, что псевдогенизация в некоторых случаях имеет адаптивное значение. Так, нулевой аллель гена *CASP12* (цистеиниласпартат-протеиназы), в норме участвующий в противовоспалительных реакциях и иммунном ответе и потерявший эти функции, сцеплен с ответом организма на сепсис. Имеются и другие примеры (см. ссылки в [52]), из которых упомянем псевдогенизацию гена *H*-цепи миозина 16 (*MYH16*), которая привела, как полагают, к уменьшению объема жевательных мышц и к последовавшему за этим увеличению размера головного мозга у человека

Эти и другие данные соответствуют высказанной ранее Олсоном (Olson) [26] гипотезе (“less-is-more hypothesis”), постулирующей, что уменьшение одной функции может прямо способствовать появлению другой, более успешной на определенном этапе. Гипотеза стимулировалась данными о том, что нокаут даже большинства генов у дрожжей не влияет на их фенотип и выживание в лабораторных условиях, иногда даже улучшая ситуацию (см. ссылки в [26]). У человека примером может быть потеря группы крови Duffy, которая сопровождается появлением устойчивости к возбудителю малярии; устойчивость к вирусу HIV связана с потерей функции гена *CCR5*. В этих случаях теряются некоторые рецепторы хемокинов, необходимые паразиту для инвазии. Известны и другие примеры адаптивной компенсации потери функции гена.

Изменение функций паралогов, возникающих в результате сегментных дубликаций участка локализации ортолога, зависит от свойств и числа копий так называемых внутренних тандемных повторов (ITS), которые входят в состав непроцессированного “белкового” гена. Они могут изменить или их кодирующие части, или паттерн сплайсинга, или характер тканевой экспрессии [26]. Более того, в настоящее время известно, что многие белки с практически одинаковыми высшими структурами представляют собой, по существу, паралоги и выполняют в разных тканях (т.е. в разном окружении)

разные функции. Не исключено, что они являются продуктом экспрессии генов-паралогов (см. [25]). Возник даже термин “moonlighting proteins”, т.е. “белки, работающие по совместительству” [53]).

Само умножение есть принцип дополнительной подстраховки любой функции — так полагали Вагнер (Wagner) и соавт. (1994; см. ссылку в [46]). Возможно, эта антропосоциальная метафора все-таки здесь неуместна и заданной цели “подстраховки” какой-либо функции быть не может (ламаркизм). Но с тем, что случайно развившиеся механизмы удвоения и амплификации любого участка могли бы оказаться эволюционно полезными, очевидно, не приходится спорить [7]. Почему? Потому что феномен проявляется в эволюции, по-видимому, в нарастающей — от низших к высшим — степени с расширением спектра амплифицируемых участков ДНК и с усложнением их структур. Само по себе увеличение числа однородных элементов могло бы стать (в совокупности с обычным мутационным процессом) предметом изучения как одной из причин возникновения биологического разнообразия [54].

Явление мультипликации локусов, характерное для геномов всего сообщества эукариот (и проявляющееся уже у низших форм), открывает возможность возникновения несходства, разнообразия на основе огромного числа вариантов амплификации, главным образом, неструктурных генов, но также и структурных, если они будут подвергаться “псевдогенизации” ([46]; см. также ссылки в [25]). Все эти события приводят к геномной и хромосомной реорганизации. Таким образом, один из основных признаков современной эволюции высших — сегментные дубликации генома, вызывающие хромосомную нестабильность, индивидуальный полиморфизм, болезни — “драматическую пластичность генома” [36]. Этот феномен, и это очевидно, сохраняется в эволюции.

Одна из попыток математического моделирования современной эволюции, предпринятая Болдогкой (Boldogkoi) [55], приводит к такому же заключению. Первоначально модель была построена на допущении, что эволюция определяется постоянным изменением частот возникновения вариантов (аллелей), различающихся по признаку приспособленности организма. И здесь автор [55] сталкивается с интуитивно очевидной ситуацией, рассмотренной выше, когда собственно структурные гены и белки не дают достаточно сведений о молекулярных основах разнообразия. Такие сведения можно получить только из рассмотрения регуляторных регионов ДНК, а также всех, включая сателлитные, диспергированные и мобильные элементы, регионов генома, не связанных прямо с синтезом белка. Это привело автора к необходимости нового обобщения, а именно, к модели, названной “гипотеза сети эгоистических генов”, в которой рассматривают генные сети, внутривидовой регуляторный поли-

морфизм, а не аллельный полиморфизм индивидуальных структурных генов [55]. Полиморфизм генерируется вариациями регуляторных компонентов генных сетей, а не вариациями их кодирующих последовательностей. Эволюция происходит через непрерывное реструктурирование состава сетей, а не за счет фиксации специфических аллелей или какого-либо одного варианта сети.

Эти выводы коррелируют с представлениями многих исследователей о существовании неслучайных пространственных сетей генов и их регуляторного обеспечения (network theories) (см. ссылки в [48, 55], а также в более ранних работах [4–6, 18, 24]).

Обратимся собственно к повторам, значение которых как присущих функционирующему геному не подвергается сомнению, но их участие в процессе жизнедеятельности до недавних пор рассматривалось некоторыми авторами только с позиций их потенциальной разрушительной деятельности (или, по крайней мере, деятельности, не способствующей адаптации организма) [56, 57]. Впрочем, как уже отмечалось, противоположная точка зрения высказывалась неоднократно, начиная с 70-х годов [4]. Доля некодирующих повторов во всех царствах живого мира может варьировать от нескольких процентов (у дрожжей) до десятков (у высших эукариот) [25]. Разнообразие этой части повторов невероятно велико — в пределах нескольких основных способов организации их последовательностей.

Среди них, во-первых, группа сателлитных повторов, образующих непрерывные ряды единиц, однотипных в популяции и виде, но специфических и отличающихся в каждом таксоне. Число семейств и копий, их размеры и последовательности исключительно разнообразны, хотя и совершенно неслучайны по своему происхождению и эволюции [32–34, 58, 59]. Некоторые сходные по структуре, родственные в одном крупном таксоне, семейства сателлитов характерны для более низких таксономических единиц по той или иной части входящих в их состав специфических и консервативных фрагментов. Это свидетельствует об их общем происхождении и о подверженности положительному отбору, т.е. о том, что они сохраняются в истории развития таксона.

Во-вторых, это разбросанные (диспергированные) вдоль генома повторяющиеся таксоноспецифичные участки. Размеры некоторых из них лежат в пределах 100–500 н. (повторы типа SINE), другие существенно больше — порядка тысяч и десятков тысяч нуклеотидов (повторы типа LINE, ДНК-транспозоны, ретротранспозоны, ретровирусы). Повторяемость их чрезвычайно велика и разнообразна. Некоторые из них не содержат открытых рамок считывания для собственных белков, другие кодируют обратную транскриптазу, нуклеазу и некоторые другие ферменты, используемые, в частно-

сти, для их размножения и встраивания обратно в геном [60, 61].

В-третьих, это гены транспортных, больших и малых рибосомных РНК. Как можно полагать, идея о “паразитической” или “мусорной” природе этого класса не пришла никому в голову, видимо, только потому, что функция рибосом стала очевидной до того, как эти повторы были обнаружены в ипостаси самостоятельных существ генома, а не как просто диспергированные или тандемные повторы.

И наконец, это псевдогены, совсем недавно рассматриваемые как отработанный материал эволюции, как настоящий “мусор”, но роль которых в эволюции и их “жизнь после смерти” в качестве гена некой другой функции только теперь начинает проясняться и вызывает огромный интерес молекулярных генетиков [25, 53].

В данной статье основное внимание будет уделено значению первой из упомянутых групп повторов, как наиболее контрастной, функциональный смысл существования которой подвергался наибольшему сомнению и наиболее труден для экспериментального моделирования.

### ФУНКЦИИ САТЕЛЛИТНЫХ (ТАНДЕМНЫХ) ДНК В ОБЕСПЕЧЕНИИ ЖИЗНЕДЕЯТЕЛЬНОСТИ

С самого момента открытия сателлитов вопрос об их функциональной роли вызывал огромный интерес в научных кругах (см. обзоры [32, 34]), но только в настоящее время оформилось мнение о важности этой роли [36, 62]. Напомним, что длины рядов повторяющихся тандемных повторов лежат в огромном диапазоне — от сотен до миллионов копий, которые следуют друг за другом по принципу “голова-к-хвосту”. Разнообразны и размеры повторяющихся единиц в разных семействах повторов — от единиц (микросателлиты) и десятков (минисателлиты), через сотни (порядка 200–400 п.н. в одной копии) до порядка  $10^3$  (мегасателлиты). Например, около 2% длины плеча Y-хромосомы человека занимают 2000 копий длиной 2450 п.н. каждая. У китовых имеются сателлиты с длиной единиц примерно 1500 и 1800 п.н. [34, 63]. Видоспецифичным может быть и число мономеров в ряду. Так, например, у трех видов дрозофилы сильно и видоспецифично различается содержание гомологичных тандемных повторов (21.5, 5 и 0.4% соответственно) [32, 33, 56].

Круг возможностей для функционирования сателлитов очерчен Бриттенем (Britten) и Дэвидсоном (Davidson) [4]. Они полагали, что повторы поддерживаются естественным отбором и сохраняются в течение миллионов лет, поскольку оказались, по-видимому, структурными компонентами — например, спейсерами для жизненно важных регионов, которые неустойчивы при эволюционных измене-

ниях. Авторы впервые предположили, что повторы могут участвовать в регуляции транскрипции, а перестройка генома за счет числа и организации повторов в эволюции может вести к образованию новых регуляторных сетей. *“Разнообразие повторяющихся последовательностей, очевидно, существует сотни миллионов лет и является источником для образования новых регуляторных взаимоотношений. Кроме того, ... плодотворные перестройки ранее образовавшихся регуляторных сетей, по-видимому, продолжают и в настоящее время”* [4].

Эти взгляды прямо переплетаются с современными представлениями, имеющими уже некоторую экспериментальную поддержку ([55], см. также обзор [64]). Сходные идеи были развиты и поддержаны в других работах (см. обзоры [1, 58, 63, 65]).

Особенно важным представляется факт увеличения доли повторяющихся последовательностей в эволюционном движении таксонов [66], и их возможное участие как регуляторов перестройки генома на стадии онтогенеза и гаметогенеза (см. обзоры [20, 67]), что прямо связывается с популяционной дифференциацией и репродуктивной изоляцией. Перестройкам повторов в геномах придается первостепенное значение. Так была создана концептуальная схема, отводящая большую, если не определяющую, роль некодирующих повторов в регуляции и перестройке генома в качестве субстрата эволюционного движения (см. ссылки во Введении). В настоящее время она стала наполняться экспериментальным содержанием. Особой опорой для этой схемы стали многочисленные данные о разнообразии сателлитов в разных таксонах живого мира.

### Феноменология роли сателлитных ДНК в филогенезе

Далее приведены некоторые общие сведения об этой взаимосвязи “с высоты птичьего полета”, поскольку сколько-нибудь детально рассмотреть в одной статье многие сотни публикаций по сателлитам, даже опираясь на упомянутые выше обзоры, нет никакой возможности.

Таксономическое разнообразие тандемных повторов по последовательностям, размеру повторяющейся единицы, числу единиц, организации рядов (к примеру, возможность внутренней повторяемости небольших участков внутри копии или самих копий и другие особенности) невероятно велико. Иногда участки разного типа “чередуются” внутри ряда, хотя это зачастую трудно определить вследствие “размывания” границ между ними и другими типами повторов вследствие случайных мутаций (см. [33, 34, 59]). Некоторые повторы типа SINE могут составлять “чистые” тандемные ряды, в других случаях в ряды сателлитов могут внедряться также другие участки; тандемные повторы могут состоять

из фрагментов диспергированных повторов [34], или некоторые фрагменты могут располагаться тандемно (у одних таксонов) или рассеяно (у других) [34]. Каждый таксон может содержать несколько совершенно разных семейств сателлита.

Филогенетический анализ множества таксонов показал, что индивидуальное сходство копий каждого семейства сателлитных повторов высоко (иногда они практически полностью идентичны благодаря “концертной” эволюции) на уровне как особи, так и популяции. Уровень дивергенции между популяциями несколько выше и увеличивается по мере расхождения популяций и видов (таксоноспецифичность). Обычно это коррелирует с увеличением систематической (по морфологии) дистанции родства между таксонами, но, если это не так, такие противоречия рассматриваются как предлог для пересмотра или уточнения таксономических границ (см. обзор [33]).

Эти заключения можно обосновать результатами работ на дрозофиле [68], жуках [69, 70], пауках (цит. по [66]), рыбах (см. [32], а также [71]), нематоде [72] и множестве других организмов [73]. Например, у рептилий сем. Lacertidae каждый из исследованных родов обладает своей, обнаруживаемой во всех видах каждого рода, гомологичной группой сателлита, но каждая из них может сильно отличаться — как по структуре, так и по особенностям организации — в разных родах семейства [74–76]. Близкие виды внутри рода сохраняют очевидную гомологию сателлита CLsat, и чем менее родственны виды, тем более различаются их сателлиты. Степень этих различий отражает степень генетического родства популяций. Чаще всего это родство соответствует представлению зоологов о морфологической таксономии этих рептилий, но в некоторых случаях позволяет более точно определять таксономическое родство и систематическое положение.

То же обнаружено у амфибий: высокосходный по структуре сателлит найден во всех хромосомах *Rana catesbiana* и еще у четырех видов этого рода (Wu et al., 1986, см. ссылку в [33]), но отсутствует в ДНК родов *Xenopus*, *Bombina*, *Acris*, а также у географически удаленных китайских видов рода *Rana*. Минисателлитные повторы у нематоды (и некоторых насекомых и моллюсков) наиболее гомогенны и малоизменчивы в эволюции [73]. Коралловые моллюски, не изменившиеся со времен девона, содержат небольшое число повторов, мало отличающихся в разных таксонах. Отсутствие морфологических признаков видообразования у географически очень отдаленных популяций нематоды согласуется с чрезвычайной мономорфностью их сателлитов [25].

Как правило, на уровне выше родового сателлитные семейства уже практически несходны, хотя в некоторых случаях, например у жуков *Tribolium*, *Palorus* и *Pimelia* (сем. Tenebrionidae), сателлитный кластер остается сходным после дивергенции родов

и семейств, насчитывающих 60 млн. лет своей истории [69]. То же относится к трем родам сем. Acipenseridae (осетровые), PstI-сателлит которых объединяет их на протяжении 100 млн. лет [77].

При любой степени сходства и отличий в последовательностях сателлита дивергенция их отнюдь не стохастична, подвержена отбору и не является результатом случайного мутирования [78]: в копиях имеются более консервативные участки, проявляющие значимую ортологию даже в наименее сходных популяциях рода, и вариабельные участки, в которых наряду со случайными мутациями есть и специфические, по которым можно проследить возможный путь их эволюции в таксоне более низкого ранга [33, 34]. Сателлиты ряда насекомых служат примером наибольшей консервативности свойств на уровне семейства (несмотря на их частичную вариабельность) как по общей длине повторяющегося ряда, так и мономера; сходны мотивы их вторичной и третичной структур (см. далее раздел о физической структуре сателлитов) (см. обзор [63]).

Во всех работах [17, 69] снова и снова обращается внимание на привязку определенных специфических сателлитов к разным таксонам и присутствие в них потенциально функциональных элементов. Не могут быть случайными специфические паттерны вариабельности и сохранение некоторых консервативных участков таксоноспецифичных сателлитов. Например, установлено, что все 12 видов единственного рода *Stenomys* (туко-туко, семейство грызунов Америки) содержат специфический сателлитный кластер. Его нет в других обширных семействах (Muridae и Cricetidae), но он обладает весьма слабым сходством с сателлитным кластером представителей семейства Octodontidae, с которым туко-туко разошелся около 10 млн. лет назад [79].

Довером (Dover) [80] давно показано, что различающиеся по плавучей плотности фракции сателлитных повторов приматов составляют группы, различающиеся по нуклеотидной последовательности. SatIII и SatIV человека имеются также у шимпанзе, гориллы и орангутана; SatII — только у человека;  $\alpha$ -сателлит — только у высших приматов и человека. Выводы этой работы подтверждены в наше время. Компьютерная обработка и описание  $\alpha$ -сателлитов из фрагментов шот-гановской суммы ДНК (набор фрагментов ДНК, расщепленной по случайному признаку) приматов подтверждает фундаментальные различия в распределении и эволюции сателлита у обезьян Старого Света и высших обезьян [81]. В наше время установлено, что две группы, из которых состоит SatIII акроцентрических хромосом, возникли в разное время. Скорость их эволюции видоспецифична (внутри приматов) и неодинакова в разных хромосомах [82].

Эволюция  $\beta$ -сателлитной ДНК приматов по структуре участков четко указывает на их сопряженность с эволюционно разными группами [83]. В



геноме орангутана основная масса этого повтора представлена очень короткими сцепками — от 100 до 175 п.н., погруженными в дублицированные сегменты ДНК длиной 60–80 т.п.н. Число копий этого дуплика с  $\beta$ -сателлитом колеблется от 7 до 100 на гаплоидный геном. В ДНК макака и гиббона этот дубликат маркируется в одном хромосомном регионе на границе с рДНК маркерной хромосомы. Однако только у гиббона дубликат включает  $\beta$ -сателлит длиной 100 п.н. Таким образом, предковая форма дубликата появилась у обезьян Старого Света 25–30 млн. лет назад, в то время как прототип  $\beta$ -сателлита появился у предшественника гиббона после отделения клады горилла + шимпанзе от обезьян Старого Света (около 35 млн. лет назад). Детально изучена собственная эволюция  $\alpha$ -сателлитной центромерной ДНК [84]. Поиски более молодых и более древних копий в огромном массиве доступных сейчас данных по мономерам и филогенетическое сравнение их у человека и приматов привело к заключению, что ряды этого сателлита проходили через периоды гипермутабильности. И каждая волна экспансии коррелирует по времени с возникновением нового таксона приматов [85].

Филогения приматов прослеживается на другом сателлите (28 п.н.), лежащем в энхансерном районе гена *TSER* (тимидилатсинтаза), слева от него [86]. Этот район не только полиморфен по числу копий (что отражается на экспрессии гена), но и структурно подразделяется на три независимо эволюционировавших части (R1, R2, R3). Это следует из того факта, что степень их сходства в разных видах приматов выражена больше, чем между тремя частями внутри одного вида. По-видимому, первичная дивергенция между ними произошла до расхождения обезьян Старого Света и гоминид и далее продолжалась независимо в каждом из таксонов. У приматов степень полиморфизма увеличивается на пути к человеку, наиболее мономорфны примитивные таксоны.

Работ такого типа множество, и они продолжают вплоть до сегодняшнего времени (например, [87]). Все изложенные факты и подобные им свидетельствуют о неразрывной взаимосвязи между формированием биоразнообразия и специфичностью сателлитных ДНК. И этой точки зрения придерживаются многие исследователи, прямо рассматривающие высокую вероятность видообразовательной функции сателлитов [17, 33, 69, 78]. Особенно показательна взаимосвязь в процессе индивидуального развития между содержанием и качеством сателлитных наборов, филогеографией и некоторыми морфофизиологическими признаками (см. ссылки в [5, 6]). Это, например, вариации в числе тандемных повторов, что может влиять как на фенотип особей на разных этапах индивидуального развития, так и путем наследования — на их потомков. Неожиданную поддержку этой точки зрения можно найти в современных данных, которые свидетель-

ствуют о полиморфизме числа копий структурных генов (CNP) в геноме человека. Так любые два индивида различаются по CNP примерно 105 генов; то же наблюдается при сравнении других видов [88].

В упомянутой выше работе [5] Трифонов полагает, что прямое экспериментальное изучение модуляции повторов и их связи с фенотипом при плейотропности фенотипического признака, зависящего от многих генов, вряд ли возможно, однако число и конгруэнтность частичных данных может служить признаком верности посылок. Приведем некоторые из множества упомянутых примеров (см. ссылки в [5]).

Содержание гетерохроматина (состоящего, как известно, в основном из сателлитов и других повторов) в ДНК австралийской саранчи *Antractomorpha similis* повышается клинально от севера к югу, что соответствует географическому продвижению и эволюции вида. В 12 видах кенгуровой крысы (*Dipodomys*) содержание сателлита колеблется от 36 до 74%, и разные величины соответствуют разным способам локомоции зверька (передвижение с помощью или двух, или четырех конечностей), что коррелирует с определенными остеологическими изменениями.

В работах лаборатории Нево показано, что у растений *Festuca ovina* и *Agrostis canina*, растущих на цинк-содержащих почвах, вырабатывается устойчивость к металлу, что коррелирует с содержанием сателлита. Другие авторы обнаружили изменения в морфологии растущего на “перекормленных” почвах льна, что сопряжено с увеличением числа тандемных повторов нескольких типов. При дифференцировке клеток моркови, культивируемых на твердой среде с фитогормонами, некоторые фракции сателлитной ДНК начинают реплицироваться раньше, чем основная часть ДНК. Амплификация гетерохроматина наблюдается в культуре клеток при воздействии лекарственных средств (колхицина, метотрексата и других) (см. ссылки в обзоре [5]).

У растений табака при образовании каллуса (дифференцировка) образуется временная фракция сателлитной ДНК. Содержание сателлита в слюнных железах личинки дрозофилы по мере развития падает от 60% до следовых количеств. В модели дифференцировки клеток дрожжей активация репортерного гена *CYC7* (изо-2-цитохром C) зависит от повтора из 57 п.н. на его 5'-конце и от типа клеток (см. ссылки в обзоре [5]). Количество сателлита RAE180 отличается в ДНК разных видов щавеля (*Rumex*). В кладе с системой половых хромосом XX/XY сателлит локализован только в малом аутосомном локусе; в кладе XX/XY(I) $\gamma$ 2 имеется дополнительный массивный локус в Y-хромосоме. Первый блок гомогенен, но видоспецифичен, второй — менее гомогенен, его внутривидовая дивергенция высока и сопоставима с межвидовой [89].

Таким образом, предположение, высказанное еще в 1971 г., вскоре после открытия сателлитных повторов, о том, что их, такой быстрый и неодинаковый в разных линиях, мутационный процесс выражается, вероятно, в видовой специфичности (в биоразнообразии), поддерживается многими данными [58]. Однако потребовалось более 30 лет, чтобы пошатнуть отношение многих исследователей к повторам как к ненужному компартменту ДНК. К настоящему времени накопилось достаточно данных и размышлений, которые исключают представления об эгоистической или, тем более, мусорной природе сателлитов.

### *Тандемные повторы в хроматине и центромере*

Основная доля тандемных повторов клетки содержится в гетерохроматине и центромере, которая представляет собой особую его структуру. Поэтому проблема возможной роли сателлитов как главной составляющей части хромосомы и как участников функционирования центромеры и кинетохора привлекала особое внимание (см. обзор по организации этих функциональных регионов и ссылки в нем [90–93]). Центромера по существу представляет собой комплекс специфических доменов (или компартментов), состоящих из разного типа сателлитных повторов, представленных в большем или меньшем количестве, со специфическими белками, который осуществляет процесс агрегации и расхождения хромосом в митозе и мейозе.

Из замечательных работ последних лет особо отметим работу Гвенатри (Guenatri) и соавт. [91], в которой изучали роль различных фракций сателлитных повторов в центромере и прицентромерном гетерохроматине мышей. Использовали двухцветный FISH-анализ живых ядер и выделенных нуклеосом в сочетании с иммуноблоттингом белков и определением времен начала репликации в ядре. На основе этих данных разработаны объемные модели структуры сателлитных доменов в их взаимодействии с белками в течение клеточного цикла. Оказалось, что собственно центромерные и прицентромерные повторы обладают разными характеристиками. Основная (мажорная) фракция образует кластеры с гетерохроматиновым белком 1 $\alpha$ , а минорная представляет собой отдельную сущность, взаимодействующую с другими — центромерными — белками. Оба комплекса содержат метилированный гистон H3, но по-разному чувствительны к нуклеазе микрококков. Каждая фракция образует повторяющуюся динуклеосомную единицу. Оба домена реплицируются асинхронно, причем сцепление хроматид поддерживается мажорным сателлитом дольше, чем минорным, что каким-то образом связано с работой гистоновой метилтрансферазы. Такие пространственно-временные параметры характерны, очевидно, для разных этапов сцепления и расхож-

дения центромер и расхождения хромосом с участием тандемных повторов.

Некоторые аспекты роли сателлитов в центромере, кроме предполагаемой опорно-структурной, суммированы в обзоре Плоля (Plohl) и соавт. [94]. Здесь оцениваются (1) особые и важные свойства структурной организации сателлитов и их специфическая локализация в центромере; (2) закономерности эволюционной гомогенизации тандемных рядов, поскольку, наряду с существенной дивергенцией в целом, имеются весьма консервативные в эволюции элементы копий или целых семейств сателлитов; (3) способность центромерных и прицентромерных сателлитных повторов — как таковых и в их мозаичном взаимодействии с другими элементами генома — взаимодействовать с рядом специфических белков; (4) несомненное участие сателлита в функционировании центромеры и, следовательно, в делении клетки. Анализируются (5) данные о роли сателлитов в эпигенетической регуляции транскрипции посредством участия в метилировании ДНК и через образование и действие малых интерферирующих РНК как продукта транскрипции самих сателлитов; а также (6) их возможная роль в перестройках хромосом как одного из механизмов видообразования и/или репродуктивной изоляции (и некоторые другие вопросы). Подчеркнем, что уже 20 лет назад было высказано мнение, основанное на детальном изучении паттернов гетерохроматина хромосом и свойств сателлитов у тритона, согласно которому даже малые изменения во временах отдельных этапов клеточного цикла вследствие изменений в количестве гетерохроматина могут иметь многократно умноженный эффект на рост организма и быть адаптивно важными [95].

Об очевидной вспышке интереса к этим проблемам в наше время свидетельствует и то, что из 140 работ, цитированных только в одном обзоре [74], половина опубликована в период с 2002 по 2007 годы, а в реальности таких работ гораздо больше. Некоторые из этих, основанных на логике и косвенных данных, теоретических положений получают в настоящее время экспериментальные подтверждения.

### *Генноинженерные и другие экспериментальные подходы к изучению роли сателлитов*

Один из таких подходов — экспериментальный трансгенез с использованием кассет, содержащих ген-мишень и различные виды сателлитов в разных ориентациях, с разным числом единиц и из разных источников [96]. Например, в клетки эритролейкоза мыши вводили ген флуоресцирующего белка (eGFP) в составе кассеты rYB, содержащей фланкирующие ген eGFP последовательности вектора YAC/BAC. Известно, что векторные участки кассеты подавляют транскрипцию гена eGFP (эффект угнетения, сайленсинга) [97]. Если в этих условиях вставить в кассету ряд определенных сателлитов

(искусственно полученных, с известной структурой), варьируя в большом диапазоне число их копий, то ген реактивируется. По состоянию метилированности (ди- или триметилирование остатка Lys9) центромерного варианта гистона H3 можно заключить, что трансген находится в деконденсированном состоянии, и это состояние не передается на расположенные рядом конденсированные участки вектора. Это явление ярко выражено при использовании  $\gamma$ -сателлита приматов, слабо выражено — у мажорного сателлита мышей и вообще не проявляется при введении  $\alpha$ -сателлита человека. Антиингибиторное (инсуляторное) действие  $\gamma$ -сателлита не есть следствие его прямого промоторного или энхансерного действия как сайта ДНК, оно коррелирует с образованием специфической структуры хроматина, которая и становится пермиссивно-транскрипционной. Киму (Kim) и соавт. [96] удалось показать, что ранее обнаруженные в  $\gamma$ -сателлите сайты связывания так называемого белка Ikaros (определенной его изоформы), которые представляют собой связку двух пентамеров GGGAA, разделенных несколькими случайными нуклеотидами, нарушаются при определенных мутациях. Это приводит к подавлению инсуляторных возможностей  $\gamma$ -сателлита.

Эти данные свидетельствуют в пользу представлений о том, что  $\gamma$ -сателлит определяющим образом (эпигенетически) влияет на транскрипцию генов, используя как механизмы метилирования гистонов (и тем самым влияя на компактизацию хроматина), так и взаимодействуя с одним из ключевых белков клеточной дифференцировки и репрограммирования транскрипции (см. ссылки в [96]).

Ранее установили, что  $\gamma$ -сателлит (по крайней мере, он) может выступать в роли так называемого “хроматинового барьера”, который разделяет функциональные состояния центромерного и прицентромерного гетерохроматина, не допускает размывания границы, что приводило бы к нарушению сегрегации хромосом в мейозе [98]. Авторы предлагают для функции  $\gamma$ -сателлита понятие “арест гетерохроматина” (“heterochromatin-arresting repeat”). О роли этого сателлита в нативном состоянии свидетельствует тот факт, что он консервативен у разных приматов и содержит консенсус сайта связывания определенной изоформы белка Ikaros. Примечательно, что фактор транскрипции YY1 связывается с прицентромерным  $\gamma$ -сателлитом только при вхождении клетки в цикл деления [98].

Еще в конце восьмидесятых обнаружили, что введение фрагмента  $\alpha$ -сателлитной ДНК человека в клетки хомячка в составе экспрессионного вектора приводит к резкой дестабилизации хромосом, выражающейся в искажении как структуры (абберации), образование дицентромерных или кольцевых форм), так и числа хромосом. Вдвое увеличивается

частота обмена хроматид, что сопровождается амплификацией введенного фрагмента сателлита [99].

В обзоре Брахмачари (Brachmachari) и соавт. [100] проанализированы примеры того, что простые последовательности типа (TG/CA)<sub>n</sub>, CTG<sub>n</sub>/CAG<sub>n</sub> и TTAGGG (теломерные повторы) и некоторые другие могут определять степень спирализации ДНК во время транскрипции, участвовать в рекомбинации; формировать новые типы структур хроматина для обеспечения спаривания хромосом; или нарушать регуляцию генной экспрессии *in vivo* при некоторых генетических болезнях человека. Приводятся некоторые экспериментальные доказательства этих положений. Показано, например, что, если в плазмиду перед репортерным геном *lacZ* встроить гексамерные Alu1-повторы сателлитной ДНК, то экспрессия гена изменится в зависимости от дозы и ориентации повторов [101].

В фундаментальной работе Айхлера (Eichler) и соавт. [51] исследовали организацию, архитектуру и эволюцию большого кластера (около 4 млн. п.н.) семейства генов цинк-содержащих регуляторных белков KRAB-ZNF, локализованных в определенной цитогенетической полосе. Эти гены образуют tandemные кластеры в разных местах хромосом. Функционирующие (как репрессоры транскрипции) гены KRAB-ZNF заключены внутри блока  $\beta$ -сателлитных повторов, которые располагаются между кассетами “белковых” генов. Этот кластер возник примерно 50 млн. лет назад в ранней эволюции приматов. В модели, предложенной авторами,  $\beta$ -сателлит использовался как промотор транскрипции и “движитель” при быстрой экспансии большого семейства генов за короткий промежуток эволюционного времени.

О работе Гвенатри и соавт. [91] упоминалось ранее. Напомним, что в геноме мыши имеются два пространственно разделенных домена сателлита, располагающихся в прицентромерной и центромерной областях хромосом (мажорный и минорный). Они специфичны и различаются по времени репликации и сегрегации в хромосомах, участвуя в центромерном сцеплении и расхождении (см. ссылки в [91]).

Эти современные, далеко не единичные, данные базируются на более ранних и перекликаются с более общими данными, показавшими, что сателлиты обладают потенциалами как вариабельности, так и консервативности, т.е. разными скоростями мутирования различных их участков (см. обзор [17]). Консервативные участки дают возможность взаимодействовать с эволюционно важными консервативными белками, а вариабельные — приспосабливаться к более изменчивым в процессе адаптации белкам типа центромерного гистона CENH3, гораздо менее консервативного, чем обычный нуклеосомный гистон H3 [102].

В упомянутом обзоре [17] приведена схема, суммирующая представления о том, как сателлитные ряды могут участвовать, главным образом, в регуляции транскрипции, определяя транскрипционно активные или неактивные структуры гетерохроматина (взаимодействуя с белками хроматина), и за счет продуктов своей собственной транскрипции (siРНК и другие [103, 104]), и за счет рибозимных свойств сателлитных РНК и способности связывать факторы сплайсинга (см. схемы в [17] и в обзоре [104]). Особенно важен тот факт, что некоторые кодирующие мРНК человека и эмбрионов курицы содержат  $\alpha$ -сателлит-подобные повторы в составе 5'- и 3'-нетранслируемых участков, входящих в состав мРНК. Экспрессия мРНК служит, по-видимому, мишенью для siРНК, произошедших от  $\alpha$ -сателлит-подобных повторов [105].

Сателлит, расположенный на 3'-фланке гена *HRAS1*, связывает четыре белка, которые на самом деле сходны с факторами транскрипции группы rel/NF- $\kappa$ B [106]. Интересно, что энхансер вируса SV40 можно заменить tandemным рядом, состоящим из любых коротких последовательностей, свойственных нативному энхансеру, и активность этого ряда прямо пропорциональна числу копий.

Предпринимались попытки изучения функции сателлитов при их трангенезе *in vivo* и в культурах клеток. Внедрение длинных рядов типа, например сателлита IV крупного рогатого скота, влияет на фенотипическое проявление ДНК мыши, а сами повторы наследуются [107, 108]. Этот сателлит (примерно 3.8 т.п.н.) насыщен прямыми, инвертированными и симметричными повторами, есть районы микросателлитных повторов с потенциальными функциями горячих точек рекомбинации, с сайтами узнавания топоизомеразы I, сигналами переключения классов иммуноглобулинов. Определены места вставок сателлита трансгенных животных, и эти вставки могли передаваться по наследству в гемизиготном состоянии, но сателлит неустойчив, некоторые его участки утрачиваются. При этом возникают аномалии развития трансгенных эмбрионов, которые можно связать с потерей определенных участков сателлита. Например, альфоидные сателлиты нестабильны при атипичном липоматозном раке, что приводит к видимым нарушениям структуры хромосом (образование дополнительных мостиков и хроматиновых нитей) и образованию дополнительных кинетохора в анафазе [109].

Эти изменения прослеживаются у мышей как *in vivo*, так и в случае митотической нестабильности трансфицированных сателлитом клонов эмбриональной тератокарциномы мыши. Работы такого типа (см. ссылки в [108–111]) свидетельствуют о регуляторном потенциале больших сателлитов.

В самом общем виде коснемся особой роли сателлитов во взаимодействии хромосом (и генома) с элементами структурного каркаса ядра в клеточном

цикле ([112, 113] и ссылки в них). Обогащение ядерного матрикса (NM) сателлитами давно установлено Разиным (Razin) [114], который и предположил их важную роль в транскрипции и репликации. В NM имеются сайты связывания с промоторами и интронами (показано для ряда генов). Предполагают, что они локализованы в основании хроматиновых петель. Эти сайты называются MAR (Matrix Attachment Regions, около 1% генома), в них определенные белки NM (TopoII, SAF-B, ARBB/MeCP2, CENP-B, ламинины и некоторые другие) взаимодействуют с прицентромерной сателлитной ДНК (у человека это  $\alpha$ -сателлит, у мыши – SatMi), причем этот контакт, как правило, не связан с нуклеотидной последовательностью сателлита (за исключением слабо выраженного консенсуса из 15 п.н., специфичной по отношению к CENP-B) [112]. Сателлиты включены в образование ДНК-белковых нитей, физически связывающих хромосомы в митозе. Например, все четыре известных типа центромерных сателлитов мыши участвуют в образовании нитей, причем теломерные сателлиты с этим не связаны [115].

Однако, как показано авторами цитируемой статьи [115] и обосновано в более ранних [114, 116, 117] и других работах (см. ссылки в [112]), белки распознают особенности высшей структуры сателлитной ДНК, обогащенной AT-участками, в частности, ее изгибаемость (кривизна, bent mode), т.е. это взаимодействие не специфично к последовательности (за исключением белка CENP-B). Такое состояние фрагментов ДНК (определенной длины и AT-состава) показано в многочисленных опытах, проведенных на разных ДНК разными авторами, и необходимо, например, в процессе фазирования нуклеосом. Взаимодействие белков NM с сателлитной ДНК зависит от степени ее метилирования. В контрольных опытах показано, что аффинность MAR к  $\alpha$ 1-сателлитной ДНК в 25–50 раз больше, чем к эухроматину. Таким образом, MAR-связывающие белки являются одновременно сателлит-связывающими.

#### **Филогенетический подход к анализу структурной организации**

Это подход основан на сравнении нуклеотидных последовательностей и структурной организации родственных сателлитов в разных таксонах. О том, что любые перестройки и транзиции содержащих сателлиты регионов служат инструментом и маркером эволюционных событий, свидетельствует работа Хорвата (Horvath) и соавт. (см. [118] и ссылки в ней). На основании филигранных опытов получена существенная поддержка общих представлений работ, в которых показана взаимосвязь между свойствами сателлитов и эволюционным многообразием и специфичностью таксонов. Внутри прицентромерного района одной из хромосом человека

включены (в определенный момент эволюции) дублированные участки эухроматиновых районов (так называемое явление “pericentromeric seedings”). Это происходило после расхождения линий высших и низших приматов (примерно 23 млн. лет назад), но перед началом расхождения человека и высших обезьян (10–20 млн. лет назад).

После этого разделения уже не наблюдалось перехода эухроматического материала в прицентромерный район, хотя в последующие 4–8 млн. лет происходила дубликация внутри него (так называемый “pericentromeric swapping”). Конечно, воспроизвести в пробирке какие-либо этапы эволюции приматов с использованием масштабных перестроек гетерохроматина невозможно, но нельзя не принимать во внимание такие очевидные корреляции между морфологическими свойствами живого организма и поведением сателлит-содержащих районов ДНК. Во многих экспериментальных работах отмечается, что гетерохроматинизация эухроматина приводит к существенным изменениям в схемах транскрипционной активности (Suzuki et al., 2006, см. ссылку [96]).

Как упоминалось во Введении, явление сегментной дубликации крупных участков ДНК может иметь важное значение в перестройках регуляторных сетей и в образовании новых, поскольку оно может затрагивать и “перетасовывать” различные по функциональной значимости регионы независимо от того, включают ли они повторы или нет и в какой форме их мозаичности. Размеры дублируемых районов могут достигать более 150 т.п.н, а их сходство остается при этом на уровне 98% и более. Локализуются они главным образом в прицентромерных и прителомерных областях хромосом. В геноме человека такие мозаики в одной сегментной дубликации включают разные типы [119, 120] тандемных и диспергированных повторов (а также ДНК-транспозоны).

Распределение дублированных сегментов между хромосомами также различается [121]. Если представить, что такого рода перестройки с захватом как гетеро-, так и эухроматического материала происходили в эволюции неоднократно [96], то роль этого процесса в ходе морфогенеза не может не быть наиважнейшей.

Об этом свидетельствуют не только данные Хорвата (Horvath) и соавт. [118], но и ряд других, в которых показана радикальная разница в структуре центромер (читай – сателлитов) у различных приматов: у низших обезьян эта организация наиболее проста, у высших – неизмеримо сложнее и разнообразнее [83, 122, 123]. Как уже упоминалось, при компьютерной обработке всех известных на настоящий момент “шот-гановских” фрагментов  $\alpha$ -сателлитов также обнаружилось фундаментальное различие в распределении и эволюции центромерных сателли-

тов у высших и низших приматов ([81], см. также обзор [34]).

На более ранних стадиях жизни, например, в геноме гидромицета *Botrytis cinerea*, сателлит появляется только в одном локусе – в интроне гена АТФ-синтазы, что можно связать с его строгой родоспецифичностью: ни один из родов этого таксона, включая наиболее близкий род *Sclerotinia*, не имеет этого сателлита в этом месте [73]. У жуков-чернотелок Канарского архипелага эволюционная дивергенция сателлитов увеличивается по мере освоения насекомыми новых островов и связана с таксономическим разнообразием [66].

Расположенные тандемно гены, кодирующие U2 нмЯРНК человека, возникли в таком виде после отделения бабуинов от высших приматов (около 35 млн. лет назад) и сохраняют высокую гомологию и цитогенетическую локализацию (локус RNU2) [124] при том, что место их присоединения во фрагменте ДНК (размером от 1 до примерно 10 т.п.н.) подвергалось перестройке; другими словами, гомогенность этого локуса в эволюции важна и поддерживается отбором. Интересно, что у многих млекопитающих и низших обезьян имеются эти гены, но они рассеяны по геному. Таким образом, тандемизация повторов прямо связана с переходом на более высокую ступень эволюции.

Один из наиболее интригующих вопросов касается работы центромеры – каким образом выполняется ее весьма консервативная эволюционная роль на фоне разнообразных в филогенетическом ряду тандемных повторов? В попытке понять это видимое противоречие рассмотрим возможности, которые открываются при изучении высших структур разных последовательностей и при функционировании тандемных и других повторов.

#### **“Структурные аспекты” вероятной роли сателлитных ДНК**

Именно эти аспекты выходят в последние годы на первый план, и именно с этой точки зрения рассматриваются действующие начала механизма специфического взаимодействия регуляторных сетей (в том числе – с участием сателлитов) в экспрессии ДНК (см. обзор [125]). Архитектоника различных регионов ДНК определяется особенностями первичной структуры этих регионов – в их зависимости от ионных и лигандных условий. В образование высших структур могут вовлекаться различные сочетания нуклеотидов, включая их гомогенные ряды или микросателлиты. Биофизические работы, начиная с 80-х годов и по сию пору, открывают все новые (помимо канонической правозакрученной двойной спирали ДНК в В-форме) вторичные и третичные структуры. Это левозакрученная Z-форма, крестообразные структуры, шпильки, стабильная нескрученная ДНК, тройные цепи, G-квартеты

и G-квадруплексы, i-форма олиго(C)-фрагментов. Каждая из них образуется в специфических участках ДНК, они иммуногенны и обнаруживаются *in vivo*. Кроме всего прочего, информативным оказалось такое свойство двухцепочечной ДНК, как кривизна (bending) фрагментов ее молекулы и возможность распознавания вариантов этого свойства белками и другими комплексообразователями [117, 126].

Эта проблема интенсивно изучается. Яркий пример — взаимодействие ДНК-топоизомеразы I, не обладающей строгой специфичностью к нуклеотидной последовательности. Она распознает минимальные ключевые участки (9 и 5 н.) на расщепляемой и нерасщепляемой цепи ДНК соответственно. Эти “ключи” обеспечивают необходимую пространственную изогнутость участка молекулы ДНК, узнаваемую ферментом. Скорость реакции зависит от того, насколько затруднено прохождение стадий конформационной “подгонки”, что можно смоделировать и изучать на известных пространственных структурах фермента и лиганда [127].

Давно показано, что открытая левозакрученная Z-форма, перемежающаяся с B-формой в составе рекомбинантной плазмиды, существенно влияет на степень суперскрученности плазмидной ДНК (Singleton et al., 1982, см. ссылку в [125]). Исходя из физических предпосылок, авторы утверждают, что даже ничтожные изменения в конформации локальных участков (например, в результате мутирования) могут глубоко влиять на степень суперскрученности хромосомной ДНК, которая, в свою очередь, оказывает влияние на транскрипционные свойства разных генов. Это позволило рассматривать конформационные переходы во вторичной структуре ДНК как фактор генетической регуляции [128].

Эти данные имеют непосредственное отношение к тандемным повторам. Ранее Трифоновым (см. обзор [129] и ссылки в нем) было отмечено: имеются все предпосылки того, что естественная кривизна молекулы ДНК могла бы стать важным фактором в распознавании ДНК и белка. В сателлитной ДНК эта кривизна двойной спирали может зависеть от числа и соотношения нуклеотидов в слегка варьирующих последовательностях и может влиять на взаимодействие копий в каждом специфичном сателлите с белками. Это обосновано экспериментально (см. обзоры [129, 130]). Если делеция двух нуклеотидов в мономере практически не влияет на общую форму отдельной копии, то она модифицирует относительную позицию других копий и влияет на суперспирализацию и периодичность ряда. Можно полагать, что паттерны вариантов кривизны как-то накапливаются и трансформируются при переходе к длинному ряду сцепленных повторов. Если соотнести число возможных структурных вариантов сателлитов только по этому признаку с относительной “склонностью” сателлитов к мутированию, то становится очевиден

огромный потенциал изменчивости этих повторов на уровне их третичной структуры, что согласуется с рассуждениями Фогта (Vogt) [131].

При помощи различных экспериментальных подходов, электронной микроскопии и компьютерного моделирования на сателлите ксенопуса [132] и жуков-чернотелок [133] показано, что в двухцепочечном состоянии эти сателлиты обладают естественной кривизной (bending, curvature, способностью изгибаться в трехмерном пространстве), рисунок которой консервативен у всех членов сателлита, но обладает слабой специфичностью. Консервативность сателлитов необходима для позиционирования нуклеосом [134]. Эти данные поддерживают результаты работы Радика и соавт. (Radice et al. 1987, см. ссылку в [135]), в которой впервые обнаружили, что сателлитная ДНК мыши обладает свойствами, характерными для молекул с определенной кривизной. В модельных условиях конденсация центромерного хроматина уменьшается при обработке дистамицином, снимающим эту кривизну ДНК. Возможность изгибания ДНК обнаружена в опытах на *E. coli*, отдельные белки которой взаимодействуют только с сайтами, образующими кривизну, и не взаимодействуют с ними в присутствии дистамицина. Другой известный пример: связывающий белок теряет способность распознавать изогнутую структуру сателлита AluI (*у Artemia salina*) в присутствии дистамицина. Эти и многие другие работы, устанавливающие и анализирующие значение высших структур ДНК в разнообразных взаимодействиях с лигандами, суммированы в фундаментальном сборнике обзоров ведущих в этой области авторов [135], к которому отсылаем желающих более глубоко ознакомиться с проблемой.

Таким образом, современные представления о высших структурах ДНК предполагают, что варианты таких структур могут играть функциональную роль, включая не только лево- или правозакрученность спирали или образование крестообразных структур в палиндромных последовательностях, но и возможность структурирования участков, которые содержат только гуаниловые или только цитидиловые остатки, как в пределах одной цепи в участках, образующих внутренние дуплексы, так и между разными цепями (Weizmann et al., 1997, ссылки см. обзор [136]) с образованием так называемых квадруплексов. В этом случае очевидна возможность возникновения топологических препятствий для продвижения полимераз, но возникают и некие реперные участки для взаимодействия со структурными и другими факторами. Они могут определять и лабильность отрезков ДНК, где они находятся, и взаимодействие с белками, вовлеченными в транскрипционную и репликативную активность генома, и образование хроматина.

Очевидно, что повторяющиеся последовательности, включая сателлитные ДНК, должны быть

подвержены такому структурообразованию и, возможно, их высшие структуры могли бы иметь дополнительную периодичность. Это было показано вначале на микросателлитных повторах (см. обзор [100]). Однако по мере того как обнаруживаются все новые варианты, способные образовывать такого (или иного) типа высшие структуры на фоне фантастического разнообразия всех сателлитов, можно полагать, что их функции также разнообразны и многообразны.

Изучение микросателлитов (до 10 п.н.), роль которых в этиологии множества генетических заболеваний человека велика (см. обзор [137]), показало, что хромосомные теломеры, состоящие из коротких повторов типа  $d(\text{TTAGG})_n$  (у позвоночных) или  $(\text{TTTAGGG})$  (у остальных эукариот), определяют время жизни клетки. В теломере человека имеется примерно 1000 повторов, заканчивающихся одноцепочечным “хвостом” (1–200 н.), обогащенных гуанином (для посадки РНК-затравки при репликации). Теломерные повторы спонтанно складываются за счет копланарного взаимодействия четырех гуанинов из разных положений цепи (G-квартеты) с последующим образованием стопок таких тетрад — G-квадруплексов. Теломерный конец выглядит как серия “бусин”-квадруплексов, нанизанных на “шнур” неструктурированных нуклеотидов. Это подтверждено иммунологически, электронно-микроскопически и с помощью специфических к квадруплексам лигандов (см. обзор [136]).

Вероятно, такие структуры микросателлитных повторов блокируют возможность удлинения теломера теломеразой, и этот процесс нарушается при неконтролируемом росте клетки при злокачественных перерождениях. Оказалось, что теломерные повторы могут образовывать различные модификации таких структур, находящихся в равновесии друг с другом [125].

Кроме этого, микросателлитные повторы, по-видимому, участвуют в регуляции транскрипции, используя, возможно, тот же механизм образования четвертичных структур: квадруплексы присутствуют в районе промоторов примерно 15% всех генов, располагаясь на расстоянии около 1 т.п.н. перед стартом инициации транскрипции (TSS) [138]. Район  $d(\text{GGGGaGGGtGGGGaGGGtGGGGaaGG})$  чувствителен к нуклеазам, т.е. свободен от гистонов и связывает некий специфический для G-квадруплекса лиганд, что влияет на экспрессию гена *ctmic* (у авторов). Компьютерный анализ показывает, что примерно половина всех генов содержит структуры, способные образовывать квадруплексы и именно в области, близкой к TSS (см. [136]). Интересно, что эта доля еще выше в области генов, включающихся при злокачественном перерождении (до 70%).

По-видимому, такие структуры очень важны как участники одного из механизмов регуляции тран-

скрипции на основе стерических возможностей повторов. Однако не исключено, что G-квадруплексы могут служить и уникальными сайтами распознавания белков — их открыто уже около 30 (см. обзор [139]) — или, например, для сдвига транскрипции в пользу одной из цепей. Это явление характерно для  $\gamma$ -сателлита в эмбриогенезе мыши (см. обзор [136]). Комплементарные ряды, обогащенные цитидином, также могут образовывать локальные высшие структуры, значение которых пока не изучено. Теоретически возможна и обсуждается проблема образования пентад и гептад другими сочетаниями повторов (например, GGA) (см. обзор [140]). Олигопиримидиновые участки типа AAGAGAG в сателлитных последовательностях дрозофилы в физиологических условиях образуют тройную спираль, которая стабилизируется ионами меди [141]. По-видимому, “регуляторная игра” на основе различных высших структур повторов с участием разных ионов может оказаться очень важным фактором.

По этой проблеме опубликованы сотни статей, интерес к ней подогревается и поисками лекарственных средств, стимулирующих образование такого рода структур или угнетение теломеразы, а также интересом к действию этих препаратов на работу генома в целом. Не исключено, что высшие структуры могут формироваться и в сателлитных повторах большей длины, часто содержащих обогащенные гуанинами участки, как у сателлитной ДНК дрозофилы (см. [141]). О возможности таких параллелей свидетельствует тот факт, что среди виновников генетических заболеваний числятся и сателлиты большой длины (см. обзоры [5, 117]).

Возможно, топологические взаимодействия могут послужить ключом к разрешению парадокса сателлитов центромер “одна функция — разные последовательности” [142]. Весьма вероятно, что различающиеся последовательности могут образовывать сходные пространственные структуры, в которых функциональны, в строгом смысле слова, лишь определенные детали. По крайней мере известно, что топоизомераза II, локализуемая в центромере, распознает и разрезает в петле шпильчатые структуры  $\alpha$ -сателлита. Мутационный анализ показал, какие свойства последовательности необходимы для этого. Не исключено, что ДНК-топоизомераза II может взаимодействовать с двумя шпильками, выступая в качестве медиатора сцепления хромосом [143].

Подводя краткий итог этому разделу, можно констатировать, что вероятное (и в некоторых случаях экспериментально обоснованное *in vitro*) влияние разнообразных высших структур сателлитных ДНК на протекание важнейших жизненных процессов в клетке открыло новую область функционирования повторов. Вряд ли окажется возможным исследовать эту проблему *in vivo* на целом организме, но нет и принципиальных преград для осмысле-

ния имеющегося материала и теоретических построений; не исключено, что экспериментальные подходы *in vivo* будут найдены.

Одно из первых продуктивных мнений о возможных конкретных и общебиологических функциях сателлитов было высказано Трифоновым [6]. Он полагает, что сателлитные повторы – во всем их многообразии и разнообразии структур – могут рассматриваться как настройщики (“tuning”) различных процессов. Многообразие и тонкая настройка этих функций опосредуются различными факторами. Среди них – нуклеотидная последовательность структурообразующих сайтов и их фланков; состояние локальной структуры хроматина и транскрипционная активность региона в данный период; скорости репликации сателлитных рядов с разным составом; пул нуклеотидов в клетке; в какой из цепей – ведущей или отстающей – образуются эти структуры во время репликации. Сюда же можно присовокупить вариации в ионных условиях (специфичность и комбинации ионов металлов и рН), равно как и специфичность и дозу белков, взаимодействующих с ДНК.

При этом очевидно, что все эти состояния и “перестроенные” процессы первично характерны для ДНК, т.е. причинно-следственная зависимость “настройка ДНК → функционирование клетки → разнообразие морфогенеза” выглядит именно так. Один из конкретных примеров: изменение структурной конформации белка HMG-1 зависит от стехиометрии его взаимодействия с сателлитным повтором и степени кривизны повтора, т.е. структура белка индуцируется конформацией ДНК, а не наоборот [144]. Авторы распространяют этот вектор влияния на формирование структуры хроматина и генной регуляции. Более высокие структуры в составе хромосом, их видоспецифичная трехмерная организация сцеплены с количеством сателлитной ДНК, с механизмом прикрепления политеменных хромосом к ядерному матриксу (см. обзоры [145, 146]).

Неожиданным аспектом проблемы адаптации белков центромеры к структуре сателлита мог бы стать пример неоцентромеры, образующейся из белков центромеры в отсутствие сателлита [147]. В свое время возможность паллиативного существования неоцентромеры рассматривалась как доказательство отсутствия влияния сателлита на формирование полной центромеры. Однако можно высказать предположение, что образование и эволюция современных белков центромеры (и их комплексов) шло по пути приспособления их пространственной конфигурации к высшей структуре сателлита. Такой “белковый слепок” мог бы частично имитировать функции центромеры в отсутствие сателлита. Напомним, что эпизодическое появление неоцентромеры указывает на “ненормальное” протекание процесса, и вряд ли отсутствие сателлитов в неоцентромере может служить отрицанием их

функции при нормальном делении клетки; скорее, это свидетельствует об обратном.

### *Другие аспекты вероятной функциональной роли повторов*

В обзоре можно лишь слегка затронуть ряд важнейших и интенсивно исследуемых сейчас проблем. Например, это процессы метилирования ДНК (см. обзоры [148, 149] и не прямое влияние сателлитов через их РНК-продукты (см. обзор [63]), которые регулируют экспрессию на транскрипционном и посттранскрипционном уровнях. Сами сателлиты и их транскрипция могут контролироваться РНК-интерференцией с других участков ДНК, как это показано на яйчниках дрозофилы [150]. Большая часть генома человека транскрибируется, и более 90% всех транскриптов синтезируются с некодирующих участков. Среди них копии интронов, регуляторных районов, повторов гетерохроматина. И, что крайне важно, паттерны экспрессии тканеспецифичны, зависят от пола и стадии развития организма [13]. Особенно важно, что экспрессия сателлитов обладает именно такими свойствами: в клетках насекомых она зависит от стадии развития, пола и касты особи; в эмбриогенезе мыши наблюдается дифференциальная транскрипция одной из цепей  $\gamma$ -сателлита [151], как и при клеточной дифференцировке [64]. Транскрипция диспергированных повторов типа SINE приводит к образованию РНК-транскриптов, потенциально значимых для некой функции. Это отражается и на тканеспецифичности паттернов экспрессии повторов, и в увеличении их числа при клеточных инсультах, и в регуляции собственной экспрессии, например, *Alu* на разных уровнях регуляции, и в воздействии на синтез белка с помощью фактора инициации трансляции, с которым взаимодействуют РНК-транскрипты *Alu*-элементов [152].

Транскрипты сателлитов каким-то образом участвуют в защите клетки от стресса [153]. При тепловом шоке в культуре клеток транскрипция сателлита III (*SatIII*), связанного с образующимися при этом в ядре “стрессовыми тельцами”, активируется на четыре порядка по сравнению с нормой. Транскрипция сильно асимметрична, большинство транскриптов содержат копии обогащенной гуанином цепи сателлита, которая остается связанной с тельцами. Как уже упоминалось, осмотический и окислительный шок, ауксины, состав и увлажненность почвы, тяжелые металлы и облучение ультрафиолетовым светом приводят к перестройке сателлитов или появлению их новых форм (см. ссылки в обзоре [5]). При осмотическом и тепловом шоке считывание зависит от различных факторов транскрипции. Ранее установлено, что транскрипция сателлитной ДНК у насекомых происходит с обеих цепей, причем у самок гораздо более активно, чем у самцов [103].



Упомянем способность некоторых транскриптов сателлитной ДНК приобретать особую структуру (“hammer-head”), обладающую “ферментативной” (рибозимы) способностью к саморасщеплению и расщеплению других РНК [154–156]. Предполагается, что эти свойства сателлитов задействованы в процессах регуляции.

Еще одна важная проблема – взаимосвязь между сателлитными повторами типа Ste и Su/Ste, локализованными в X- и Y-хромосомах дрозофилы. Эти повторы представляют собой тандемно амплифицированные гены регуляторной  $\beta$ -субъединицы протеинкиназы СК2 [157]. В X-хромосоме со всех образовавшихся копий синтезируется  $\beta$ -субъединица СК2 в исходном ее состоянии, а в Y-хромосоме этот ген при амплификации псевдогенизируется, утрачивает кодирующую способность, но приобретает функцию супрессии гена в X-хромосоме за счет ингибирующего действия дцРНК-супрессора [158]. Сверхэкспрессия повторов Ste X-хромосомы – при полном или частичном отсутствии локуса супрессии Su/Ste в Y-хромосоме – приводит к частичной стерильности самцов (нарушение мейоза). Эти повторы супрессируют плодовитость, что является важным фактором популяционной эволюции. Повидимому, особи, у которых функция гена сохранилась при его амплификации одновременно в обеих хромосомах, были отброшены отбором, и сохранились только те особи, у которых произошла истинная псевдогенизация (потеря функции), что помогло сдерживать экспрессию гена, не нужного в таком количестве. Другими словами, можно предположить, что “игра” на присутствие/отсутствие определенного вида сателлита является элементом морфообразования. Вряд ли это явление может быть названо “паразитическим” или “эгоистическим”, так как оно поддерживается отбором и сцеплено с очевидным видообразующим сдвигом, изменяя баланс полов в популяции.

То же относится к элементу Responder, функционирование которого всегда сцеплено с большим фрагментом сателлитных повторов. Приспособляемость дрозофил уменьшается при делеции сателлита Rsp [159]. Этот сателлит всегда делетируется или транслоцируется в случае аберрантной конденсации хромосом в сперматогенезе. Число повторов в ряду во всех хромосомах хорошо коррелирует с их чувствительностью к нарушениям сегрегации (distortion) [160]. Содержание сателлитных и сателлитоподобных последовательностей определенных типов в крупных регионах (порядка 200 т.п.н.), фланкирующих гены, связано (у дрозофилы) со временами репликации более чем 4000 генов, и вероятность этой ассоциации выходит за пределы случайной [161].

Кодирование “смысловых” мРНК в тандемных повторах заложено в мегасателлите RS477 генома человека. Они содержат и промоторные участки, и

поли(А), и открытую рамку считывания для новой формы деубиквитирующего фермента, уникальной тем, что уровень экспрессии гена в головном мозге может модулироваться антисмысловым транскриптом комплементарного участка ДНК [162].

В этом обзоре удастся лишь упомянуть такую неисчерпаемую тему, как участие микросателлитных повторов в возникновении множества генетических заболеваний. Число триплетов в экзонных частях генов прямо влияет на структуру и функцию белка за счет умножения в нем числа аминокислот, кодируемых данным триплетом-кодоном [137]. Однако помимо этого вредоносного действия, микросателлиты могут выступать как инструмент морфо- и видообразования. Например, экспансия или сокращение числа микросателлитных повторов в регуляторных областях генов, определяющих эволюцию собак разных пород, количественно связаны с морфологическими характеристиками конечностей и черепа [163]. Высказывается мнение, что связанные с генами тандемные повторы функционируют как ускорители (промоторы) эволюции, давая мощные, крупные вариации генома и допуская экстремально быструю эволюцию новых форм.

### ДИСПЕРГИРОВАННЫЕ ПОВТОРЫ, ТРАНСПОЗОНЫ

Как отмечено во Введении, можно полагать, что выводы из более подробного рассмотрения функционального значения тандемных (сателлитных) повторов, особенно часто рассматривавшихся как “мусорная или паразитическая” фракция генома, применимы и к оценке гораздо более сложной и разнообразной (и структурно, и функционально) группы повторов – диспергированных мобильных элементов (TE, Transposable Element), и обосновывают противоположную точку зрения. Под этим термином объединяется широкий спектр рассеянных по геному повторов, распадающихся на две основные группы. Одна из них, ДНК-транспозоны, обладают системой вырезания, переноса и встраивания копии в другие участки ДНК как внутри собственного организма, так и между геномами разных таксонов. Число копий этих элементов невелико, составляя несколько десятков на геном, но их значение в возникновении биоразнообразия чрезвычайно велико. Вторая группа, ретротранспозоны, размножается и переносится внутри одного генома путем образования РНК-копий с последующей их обратной транскрипцией и встраиванием в новом месте. Ретропозоны образуют несколько структурно различающихся классов (LTR-содержащие и не содержащие, LINE и SINE), число копий которых в геноме может достигать многих сотен тысяч. Поскольку в обзоре и в данной главе обсуждаются только общее значение всей группы повторов, связанное с их перемещением и встраиванием в новые участки генома, мы коснемся лишь этого вопроса,

не рассматривая системно каждый из упомянутых классов и выделяя наиболее важные аспекты их способности к перемещениям, необходимым, как можно полагать, для формирования биоразнообразия.

Из этих повторов собственно “паразитически”, в этимологическом значении слова, могли бы быть — и до сих пор таковыми часто считаются — вирусоподобные частицы типа ДНК-транспозонов, способные самостоятельно перемещаться, “заражать” новые популяции, приносить очевидный вред геному с потерей адаптивности организма. Такая точка зрения неоднократно высказывалась (см., например, [164]), несмотря на признание и описание теми же авторами множества данных, свидетельствующих о важной интегрирующей и регулирующей роли разнообразной по строению и способам перемещения в ДНК группы мобильных элементов в жизнедеятельности.

Противоположная точка зрения предполагает наиболее важную и чуть ли не решающую роль всех групп перемещающихся элементов в эволюции генома и, тем самым, в эволюции биологического разнообразия [7, 164–166]. Приведем несколько примеров.

Вклад ТЕ в физиологию практически всех эукариот сильно зависит от их положения при транспозиции и от частоты таких событий [167]. У дрожжей *Schizosaccharomyces pombe* доля транспозона (Tf1) в геноме которых невелика (около 1%), они располагаются, главным образом, вблизи промоторов PolIII. Анализ 73250 участков включения ТЕ подтвердил, что эти промоторы (примерно их треть) служат мишенями для включения мобильных элементов в 75% межгенных районов, содержащих эти промоторы. Зависимости между числом сайтов и уровнем транскрипции генов не наблюдается, но отчетливо проявляется приверженность ТЕ к генам, связанным с возникновением стресса, что свидетельствует об их участии в процессе адаптации. В другом случае, регуляторные сайты и сайты связывания факторов транскрипции, необходимые для экспрессии гена одного из интерферонов (INF- $\lambda$ 1), располагаются в мобильном элементе и в Alu-подобном участке, что прямо свидетельствует об участии ТЕ в регуляции синтеза интерферона [168].

Проблемы, связанные с внедрением транспозонов и проявляющиеся в ядре и цитоплазме в регуляции соотношения полов, во влиянии ТЕ на экспрессию генов, в индукции точечных мутаций, в предмейотическом метилировании и интерференции РНК, а также участие ТЕ в “дизайне” генома и, наконец, их внутривидовая (индивидуальная) изменчивость, рассмотрены в ряде работ ([146, 165, 166, 169–171]). Все данные свидетельствуют в пользу исключительно важного значения диспергированных мобильных элементов в регуляции жизнедеятельности и в видообразовании [170].

Точка зрения о вирусоподобной болезнетворной природе ТЕ при горизонтальном переносе также не может считаться убедительной. В этом случае следовало бы признать, что в ту же категорию паразитических попадают все события горизонтальных переносов, даже потенциально дружественных и адаптивных. Наиболее яркий пример — горизонтальный перенос комплекса генов обмена каротиноидов (не синтезирующихся у животных) из ДНК грибов в ДНК насекомых типа Aphididae и их успешное функционирование в новом хозяине. Это подтверждается и появлением оранжевой окраски в обычно зеленой популяции, и увеличением адаптивности этой популяции [172]. В настоящее время все чаще высказывается точка зрения о том, что роль повторов вообще и мобильных элементов, в частности, в видообразовании и эволюции до сих пор недооценивалась (см. обзоры [146, 170, 173]).

Велика роль повторов типа LINE млекопитающих [174]. Например, если полноразмерный L1 млекопитающих располагается в интроне одного гена в антисмысловой ориентации, то он может как бы “разрывать” ген. Среди продуктов транскрипции обнаружены предсказываемые структуры, которые могут образоваться при этом: они содержат последовательности правого или левого экзонов и фрагменты LINE. После анализа 15 генов появились основания полагать, что белки, кодируемые “разорванными генами”, способны функционировать или во взаимодействии друг с другом, или отдельно.

Повторы типа LINE участвуют, по-видимому, в выборе того, будут ли два аллеля гена экспрессироваться по кодоминантному или моноаллельному типу [175]. Контекст в хромосомах вокруг моноаллельно экспрессирующегося гена (плотность и число LINE, их сохранность или молодость, степень метилирования) существенно отличается от контекста при кодоминантном наследовании. Моноаллельные гены фланкированы более “молодыми” и менее дивергировавшими LINE1, меньшим числом CpG-островков и меньшим числом элементов SINE, чем биаллельно экспрессирующиеся.

Особенно важно участие повторов в функционировании теломер хромосом дрозифилы. Теломеры образуются в результате транспозиции в определенном порядке ретротранспозонов трех видов LINE, не содержащих LTR. Этот процесс строго регулируется, поскольку добавление полного ретротранспозона увеличивает длину хромосомы на несколько тысяч пар нуклеотидов, в то время как теломера укорачивается всего на 70–75 п.н. за одно поколение. В этом процессе происходит гетерогенизация концевых повторов у разных особей и даже в разных типах клеток у одной особи; при этом активность теломеразы особенно важна в герминальных тканях. Это кардинально отличает дрозифилу от других организмов, теломера которых состоит из коротких

повторов. Таким образом, “собственная” жизнь ТЕ оказывается существенной для наиболее важного процесса программируемой летальности. Так же важны взаимодействия специфичных теломеразных белков с теломерой в процессах позиционирования их внутри ядра, что связано со спариванием и расхождением хромосом [176].

Удивительно, что мобильные элементы задействованы в межвидовой и межродовой гибридизации, которая может быть источником видообразования. При гибридизации кенгуру двух родов (*Macropus eugenii* и *Wallabia bicolor*) гибрид получает оба набора родительских хромосом, в которых присутствуют хромосомные перестройки, в том числе, амплификация ретроэлементов в их неметилированном состоянии. Центромера аутосом, полученных от матери, сильно увеличена в размерах. При этом гибриды фенотипически нормальны и вполне жизнеспособны [177].

Общебиологические примеры рассмотрены в обзоре Фонтдевила (Fontdevil) [178]. Автор суммирует, что в большом числе случаев приспособляемость гибридов оказывается или лучше, или остается без изменения. Это характерно и для растений, и для животных, таких как рыбы (*Gambusia*, *Geospiza*), амфибии (*Hyla*), некоторые насекомые (*Drosophila*), млекопитающие (*Sorex*, *Felis*, *Canis* и ряд других). У ириса F1-гибриды образуются редко, но эти особи обладают лучшей приспособляемостью к определенным новым условиям. Известно, что у рыб и рептилий гибриды также активно оккупируют новые места обитания.

В этой связи Фонтдевил [178] напоминает высказывание МакКлинток (1980 г.) о том, что новые типы геномных перестроек, индуцируемых мобильными элементами, открывают новые пределы существования таксона в эволюции.

Можно обобщить, что с открытием явления горизонтального переноса понятие “геномного паразитизма” для любого фрагмента ДНК, подвергнувшегося переносу, становится бессодержательным, поскольку каждый такой фрагмент, даже если он содержит вполне легально функционирующие гены, мог бы считаться паразитом. То же самое можно сказать по поводу важного явления дубликации и мультипликации в принципе любых участков и регионов ДНК внутри одного генома, что совсем не предполагает их обязательного повреждающего эффекта, а представляет собой естественный процесс разнообразных перестроек, подвергающихся затем естественному же отбору. Это явление широко распространено и вовсе не требует, и не предполагает необходимости оценки его как отражения отношений “паразит–хозяин”, и существует как атрибут взаимодействия и эволюции живого на генетическом уровне (см. обзор [179]).

Следует подчеркнуть, что, поскольку повреждающий эффект любых перестроек нагляден и прояв-

ляется или фенотипически, или в заболевании, оценка его значения может быть преувеличена по сравнению с положительными эффектами, которые могут не проявляться немедленно, труднее фиксируемы, и для их выявления требуются специальные усилия. Тем не менее, эволюционное значение переноса транспозонов как внутри таксонов, так и между ними — даже на самом высоком таксономическом уровне, как в случае переносов между позвоночными и беспозвоночными, велико [180] и подтверждается во многих работах (см. обзоры [181–183]). Так перенос трех типов ТЕ из благополучной популяции дрозофилы в угнетенную приводит в некоторых случаях к резкому и очевидному увеличению ее адаптивности (по определенным маркерам и по состоянию популяции) [184], но в других — к ухудшению адаптивности [185]. В геномах дрозофил перераспределение в результате скрещиваний (или горизонтальных переносов) нескольких типов ТЕ вызывает их “массовую мобилизацию”. Своего рода локальная “хромосомная революция”, приводящая к гибриднему дисгенезу, может рассматриваться не только как отрицательный признак, но и как источник хромосомного полиморфизма и возникновения репродуктивной изоляции между популяциями, что может быть причиной начала видообразования [170, 186].

Что касается “несамостоятельных” диспергированных повторов типа SINE (ретропозоны), то их эволюционное значение как наиболее адекватных маркеров эволюции сейчас не подвергается сомнению [60, 61]. Как и сателлитные повторы, ретропозоны обладают ярко выраженной таксонной специфичностью, свидетельствующей о них, если не как о первичной движущей причине в эволюции геномов (хотя это не исключено), то как об абсолютно сопряженных с эволюцией таксонов. Изучено около сотни семейств SINE, размеры и возраст которых, нуклеотидные последовательности копий, их число и организация невероятно разнообразны и специфичны для разных групп таксонов [60, 61]. Имеется, по крайней мере одна, наиболее древняя и вырожденная группа повторов (MIR), лежащая в основании древа эукариот (см. ссылки в [60]). Известны более молодые группы повторов, характерные для организмов, составляющих крупные разделы систематики (различные отряды млекопитающих, приматы, отряды рыб, амфибий [60, 61, 187], рептилий [188]) и еще более молодые подгруппы, специфичные для уровня семейств и родов. Все эти свойства и признаки коротких ретропозонов, а также то обстоятельство, что их собственная эволюция направлена только в сторону увеличения числа, делают их наиболее удобными и практически не подверженными гомоплазии молекулярными маркерами филогении [187]. Очевидно, что разнообразные функциональные проявления ретропозонов, о которых будет кратко сказано далее, делают их участие в не-

посредственном формировании биоразнообразия, их таксонообразующую роль высоко вероятной.

Мнение одного из лидеров этого направления, Окада (Okada) [187], об этом выражено четко: ретропозоны, которые вначале рассматривались как геномные паразиты или мусорные ДНК, во многих случаях приобретают новые функции, будь то новый поли(А)-сигнал, сигнал энхансера или обновленный экзон. Допускается участие транскриптов SINE в промотировании трансляции при стрессах (Chu et al., 1998, цит. по [187]). Не исключается также возможность их пассивного переноса за счет захвата при переносе мобильными элементами или при горизонтальном переносе [189].

Найдены многочисленные примеры прямого участия диспергированных повторов (см. обзоры [60, 61, 152, 190]) в функционировании ДНК, содержащих промоторные, энхансерные, супрессорные и другие элементы типа сайтов связывания рецепторов гормонов и других белков (см. обзор [191]). Они могут быть источником новых экзонов, как Alu-повторы приматов [192], могут включать ретропозоны на границе интронов и экзонов, что приводит к возникновению нового фенотипа, как в случае гена *SILV*, отвечающего за пигментацию шерсти и нарушения зрения у собак. Определенная делеция в участке ретропозона, обогащенном олиго(dA), восстанавливает нормальное состояние собак [193]. В геноме собак примерно половина структурных локусов находится в биаллельном состоянии — один аллель содержит специфический для собак тип SINE, а другой — не содержит, причем разные породы различаются по паттернам появления таких биаллельных локусов [194] (другие примеры см. в обзорах [43, 44]).

Примеры косвенного, регуляторного, эпигенетического влияния диспергированных повторов еще более многочисленны и впечатляющи (см. обзоры [195, 196]). Давно открытое метилирование цитозина связано, очевидно, с тканеспецифичным и гендерным распределением метилирования Alu-повторов у приматов. Количество транскриптов Alu-повторов существенно меняется при клеточных стрессах, приводя к перераспределению интенсивности синтеза определенных ферментов.

Сама организация диспергированных повторов, например их спорадическая тандемизация в одном локусе с последующей транскрипцией, дает различные паттерны РНК, что функционально значимо [197]. siРНК угнетает функцию ретропозонов (сайленсинг) в зародышевых клетках дрозофилы [198]. В состав Alu-повторов входят кластеры сигналов транскрипции, направляемой РНК-полимеразой II [199]. Очевидна и важна роль Alu-повторов в посттранскрипционной регуляции на уровне сплайсинга и в специфическом участии в считывании мРНК.

Таким образом, участие SINE в механизмах мутагенеза обусловлено их включением в различные

области генома — экзоны, интроны, сайты связывания факторов транскрипции, репликации, рецепторов гормонов, что может приводить как к разрушающему, так и созидающему воздействию на геном. SINE участвует в механизмах структурирования и реструктурирования генома в связи с появлением новых генов; в механизмах регуляции через образование РНК-транскриптов, тканеспецифически опосредующих посттранскрипционный уровень формирования признаков; в механизмах взаимодействия ядерного матрикса с белками и сателлитами; в механизмах эпигенетической регуляции за счет метилирования—деметиличивания и в изменении паттерна этой модификации (см. обзор [60]). Роль транскриптов LINE как движущей силы эволюции обсуждается и высоко оценивается другими авторами [200].

Особая роль диспергированных повторов связана со структурой и функцией центромеры, где они “работают” наряду с сателлитными повторами, как, например, в случае транспозонов, предоставляющих сайты для определенных, специфичных для центромеры, белков (см. обзор [201]). Некоторые из работ последних лет отмечу особо. Важно, что последовательности TE, обычно обнаруживаемые в нетранслируемых районах мРНК, более часто встречаются в мРНК более молодых генов (типа вовлеченных в иммунитет) и гораздо более редко — в транскриптах высококонсервативных генов с функциями основного обмена (см. обзор [190]). Около двух третей всех известных генов содержат Alu в интронах или в 3'-нетранслируемых участках. Не исключено, что они могут быть сигналами нестандартных транскриптов [96]. Alu-повторы, как отмечалось, содержатся (примерно в 5% случаев) и в экзонах некоторых генов, причем даже единичные замены в них приводят к образованию либо альтернативных, либо новых форм белка и к возникновению болезней (см. обзоры [202, 203]).

Последовательности ретропозонов могут служить субстратами для гомологичных и негомологичных рекомбинаций, их взаимная ориентация и расположение в геноме прямо отражаются на стабильности хромосом (см. обзор [204]).

## ЗАКЛЮЧЕНИЕ

В заключение нельзя не коснуться упомянутой в начале обзора теоретической работы, посвященной осмыслению сложнейших взаимосвязей между молекулярными участниками процесса морфогенеза [24]. Оловников предполагает, что в этой объемной сети, названной им “стереогеномом”, существуют уровни и подуровни взаимодействий, включающих большее или меньшее число функционирующих единиц. Наиболее логичные из них — “морфоны” — по существу представляют собой различные домены некодирующей части ДНК, которую автор называет “локационной”, имея в виду управление по-

рядком “выхода на сцену” той или иной единицы в морфогенезе. Морфоны, по идее автора, — это носители любых, в основном, повторяющихся последовательностей, кроме генов “домашнего хозяйства”, т.е. они локализованы в гетерохроматине. *“Хорошо известное из исследований по молекулярной эволюции явление филогенетических изменений качества и количества различных повторов, вероятно, как раз и является ни чем иным, как проявлением работы по созданию и модификации морфонов”*, — пишет автор. Высказывается убеждение, что если *“... филогенез, в ходе которого базовые системы, вроде генетического кода или биоэнергетики, меняются мало, основан на игре морфологий, то роль всего, что связано с чтением позиционной информации, приобретает для эволюции особое значение .... Бывшая “эгоистическая ДНК” ...получает новую функцию, — она становится “локационной” при оформлении и работе морфонов”* [24].

В настоящем обзоре высказывается точка зрения, наиболее идейно близкая взглядам Оловникова и, как надеется автор данного обзора, она свидетельствует о важнейшей и первенствующей роли повторов в жизнедеятельности. Разделяется мнение многих авторов (и автор обзора лишь только несколько заостряет ее) о том, что сателлитные и диспергированные повторы во всем их разнообразии играют главную роль в структурировании и функционировании сложнейшей регуляторной сети клетки и в непосредственном взаимодействии с другими компонентами клетки. В настоящее время полностью осознано, что амплификация “белковых” генов и вообще любых локусов (сегментная дупликация) во всем разнообразии ее проявления также служит важнейшим фактором перестроек генома и возникновения вариантов биологического разнообразия.

Эта точка зрения, в том или ином (более мягком) виде сформулированная и другими авторами, тем не менее, не является общепринятой. Попытки рассматривать процессы жизнедеятельности с позиций их якобы паразитической природы, их “ненужности” или “эгоистичности” встречаются до сих пор, как будто авторы этих высказываний действительно знают, что “нужно” эволюции, а что — нет для поддержания и возникновения бесчисленных форм живого вещества ....

Еще более обобщая, — принцип умножения любых регионов ДНК как кодирующих, так и не кодирующих белки, обеспечивает и поддерживает непрерывную модификацию генетического и регуляторного материала, является прямым источником биологического разнообразия форм живого. В этом процессе преобразования и взаимодействия всех регионов ДНК возникают варианты, поддерживаемые отбором, или варианты, которые несут несомнимые с основным стержнем биохимического и

энергетического обмена признаки, и поэтому отбрасываются.

Поскольку проблемы такого рода не могут быть смоделированы экспериментально во всей их сложнейшей совокупности, приходится опираться на частные модельные опыты, на сравнительное изучение разных филогенетических уровней и на знание первичных и высших структур повторов во всем их разнообразии и особенностях взаимодействия с другими компонентами клетки. Невозможность прямого моделирования эволюции должна компенсироваться, хотя бы частично, конгруэнтностью результатов и выводов, которые поддерживают теоретические предпосылки об определяющей роли повторов и, в целом, о принципах дупликации/амплификации любой части генома как движущей силы в возникновении разнообразных биологических форм и, следовательно, эволюции. В данном случае совокупность упомянутых (а еще большего числа, к сожалению, не упомянутых) результатов свидетельствует о важнейшей роли повторяющихся регионов, о динамике их собственной жизнедеятельности и эволюции, о роли их первичных и высших структур в формировании и определении огромного числа вариантов регуляторных и структурных потенций, которые во всех своих проявлениях могут быть причиной индивидуального и таксонного разнообразия живых существ.

Если совсем заострить ситуацию, то можно предположить, что вся не кодирующая белки повторяющаяся часть генома, которая составляет, как известно, большую (а у высших эукариот — большую) часть генетического материала, является главным признаком и основой живого вещества. Можно высказать гипотезу, что эта часть геномной ДНК представляет собой основную, “хозяйскую” действующую субстанцию, некий “каркас”, на котором, во взаимодействии с которым и для поддержания которого функционирует белково-лигандный аппарат. Собственная молекулярная эволюция каркаса определяет паттерн и эволюцию структурных белков, обеспечивающих разнообразие морфогенеза, но не являющихся его причиной. Консервативному общему “стержню” процессов жизнедеятельности всех организмов — генам метаболических и энергетических процессов, а также генам “домашнего хозяйства” — отводится роль в образовании защитной оболочки наследственного вещества, а также в обеспечении ферментативного пути репликации, транскрипции и эпигенетической регуляции. Наше представление согласуется, в большой степени, с компьютерной модельной системой “регуляторных сетей эгоистических генов” Болдогкой (Boldogkoi) [55], упомянутой в тексте, и с представлениями Трифонова и Оловникова, оказавшими большое влияние на автора обзора.

Еще раз подчеркнем то обстоятельство, что, как выясняется в наши дни, феномен дуплика-

ции/мультипликации генетического материала касается практически любых регионов ДНК любой длины, в том числе, кодирующих ферменты и другие белки. Этот феномен выражен тем сильнее, чем сложнее и выше организован таксон, но проявляется уже на простейших формах жизни. Это свидетельствует о положительной эволюционной роли явления любой мультипликации. Оно создает дополнительные и новые возможности для вариативности (условия и пути эволюции) как через механизмы умножения числа сохраняющих при этом функцию белков, так и за счет истинной псевдогенизации (полочки с потерей функции). Но и псевдогенизация не всегда (и чаще всего не) является отрицательным фактором, поскольку позволяет белкам приобретать другие (или дополнительные) функции, проявляющиеся на другом поле. Примеров такого рода накапливается все больше (гипотеза “less-is-more”, [35]). Поэтому очевидно, что собственно белковые тела и их комплексы также включены в процесс эволюции, конечный результат которой – их суммарное взаимодействие.

Парадоксальным образом наша точка зрения смыкается, сделав круг, с точкой зрения адептов “мусорной” ДНК, предполагавших, что эта часть ДНК использует возможности транскрибируемой части генома для собственных нужд, тем самым паразитируя на нем. В нашей концепции – обратной – именно эта якобы “ненужная” часть ДНК представляется, напротив, главным компонентом и “хозяйном” жизнедеятельности и эволюции, именно она первично видоизменяется, именно она формирует и регулирует структурную часть генома и направляет ее в сторону постоянно работающего принципа образования все новых и новых биологических форм-вариантов – для “лап” естественного отбора.

Сходные представления высказаны Бартом (Burt) [205], который рассматривает моменты, неопровержимо свидетельствующие о присутствии некодирующим повторам свойства важнейших инструментов и участников эволюции естественных популяций, хотя и в этой работе они почему-то продолжают называться как “selfish” и “junk”.

Автор благодарит Т.Д. Машкову, Н.Ю. Опарину, А.М. Оловникова, Е.Д. Свердлова, а также Е.Ю. Дмитриеву за обсуждение статьи на стадии подготовки, за ценные замечания и советы, учтенные автором в конечном варианте обзора.

Работа получила финансовую поддержку Российского фонда фундаментальных исследований (10-04-000325) и Программы фундаментальных исследований Российской академии наук “Генофонды и генетическое разнообразие” (2010 г.).

## СПИСОК ЛИТЕРАТУРЫ

- Zuckerklund E. 1992. Revisiting junk DNA. *J. Mol. Evol.* **14**, 259–271.
- Orgell C.E., Crick F.H. 1980. Selfish DNA: the ultimate parasite. *Nature*. **284**, 604–607.
- Doolittle W.E., Sapienza C. 1980. Selfish genes: the phenotype paradigm and genome evolution. *Nature*. **284**, 601–693.
- Britten R.J., Davidson E.N. 1971. Repetitive and non-repetitive DNA sequences and a speculation on the origin of evolutionary novelty. *Quart. Rev. Biol.* **46**, 111–133.
- Trifonov E.N. 1999. Elucidating sequence codes: three codes for evolution. *Ann. N.Y. Acad. Sci.* **870**, 330–338.
- Trifonov E.N. 1999. Tuning function of tandemly repeating sequences: a molecular for fast adaptation. (Available from the author by request).
- Kazazian H.H., Jr. 2004. Mobile elements: drivers of genome evolution. *Science*. **303**, 1626–1632.
- Feschotte C. 2008. The contribution of transposable elements to the evolution of regulatory networks. *Nat. Rev. Genet.* **9**, 397–405.
- Pritham E.J. 2009. Transposable elements and factors influencing their success in eukaryotes. *J. Hered.* **100**, 648–655.
- Faulkner C.J., Carrininci P. 2009. Altruistic function for selfish DNA. *Cell Cycle*. **8**, 2895–2900.
- Malik Y.S., Henikoff S. 2002. Conflict begets complexity: the evolution of centromere. *Curr. Opin. Genet. Dev.* **12**, 711–718.
- Eyre-Walker A. 1999. Evolutionary genomics. *Trends Ecol. Evol.* **14**, 176–177.
- Mattick J.S. 2003. Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. *BioEssays*. **25**, 930–939.
- Mattick J.S., Makonin I.V. 2005. Small regulatory RNA in mammals. *Hum. Mol. Genet.* **14**, R121–R132.
- Mattick J.S., Taft R.S., Faulkner G.J. 2010. A global view of genetic information moving beyond the gene and the master regulation. *Trends Genet.* **26**, 21–28.
- Novak R. 1994. Mining treasures from “junk DNA”. *Science*. **263**, 608–610.
- Ugarkovic D. 2005. Functional elements residing within satellite DNA. *EMBO J.* **6**, 1035–1039.
- Castillo-Davis C.J. 2005. The evolution of noncoding DNA: how much junk, how much func? *Trends Genet.* **21**, 533–536.
- Carroll S.B. 2005. Evolution at the two levels: on genes and form. *PloS Biol.* **3**, e245.
- Carroll S.B. 2008. Evo-devo and expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*. **134**, 25–36.
- Kunarso G., Chia N.-Y., Jeyakani J., Hwang C., Lu X., Chan Y.-S., Ng H.-H., Bourque G. Transposable elements have rewired the core-regulatory network of human embryonic stem cells. *Nat. Genet.* **42**, 631–634.
- Richard G.-F., Kerrest A., Dujon B. 2008. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol. Mol. Biol. Rev.* **72**, 686–727.

23. Rocha E.P.C. 2003. DNA repeats lead to the accelerated loss of gene order in bacteria. *Trends Genet.* **19**, 600–603.
24. Оловников А.М. 1996. Молекулярные механизмы мофогенеза: теория локационной ДНК. *Биохимия.* **61**, 1948–1970.
25. Свердлов Е.Д. 2009. Взгляд на жизнь через окно генома. Т. 1. М.: Наука.
26. Olson M.V. 1999. Molecular evolution, 99. When less is more: gene loss as an engine of evolutionary change. *Am. J. Hum. Genet.* **64**, 18–23.
27. Stuart J.M., Segal E., Koller D., Kim S.K. 2003. A gene-coexpression network for global discovery of conserved genetic modules. *Science.* **302**, 249–255.
28. Свердлов Е.Д. 1999. Микрокосм генома. *Молекуляр. биология.* **33**, 917–939.
29. Bergman S., Ihmels J., Barkai N. 2004. Similarities and differences in genome-wide expression data of six organisms. *PLoS Biol.* **2**, E9.
30. Barabashi A.L., Oltvai Z.N. 2004. Network biology: understanding the cell's functional organization. *Nat. New Genet.* **5**, 101–113.
31. Koonin E.V., Wolf Y.I. 2010. Constraints and plasticity in genome and molecular-phenome evolution. *Nat. Rev. Genet.* **11**, 487–498.
32. Elder J.F., Turner B.J. 1995. Concerted evolution of repetitive DNA sequences in eukaryotes. *Quart. Rev. Biol.* **70**, 297–320.
33. Гречко В.В. 2002. Молекулярные маркеры в филогении и систематике. *Генетика.* **38**, 1013–1033.
34. Хамлебен В., Беридзе Т.Г., Бахман К., Коварик Я., Торрес Р. 2003. Сателлитные ДНК. В сб: *Успехи биологической химии.* М.: ВИНТИ, **43**, 267–306.
35. Stein L.D. 2004. End of the beginning. *Nature.* **431**, 915–916.
36. Margu-Bonet T., Girirajan S., Eichler E.S. 2009. The origin and impact of primate segmental duplications. *Trends Genet.* **25**, 443–454.
37. Bailey J.A., Eichler E.E. 2006. Primate segmental duplications: crucibles of evolution diversity and diseases. *Nat. Rev. Genet.* **7**, 5542–5564.
38. Batistoni R., Pesole G., Marracci S., Nardi J. 1995. A tandemly repeated DNA family originated from SINE-related elements in European Pleistocene salamanders (Amphibia, Urodela). *J. Mol. Evol.* **40**, 608–615.
39. Koszul R., Fischer G. 2009. A prominent role for segmental duplications in modelling Eukaryotic genomes. *C. R. Biologies.* **332**, 254–266.
40. Feliciello I., Picariello O., Chinali G. 2006. Intraspecific variability and unusual organization of repetitive units in a satellite DNA from *Rana dalmatina*: molecular evidence of a new mechanism of DNA repair acting on satellite DNA. *Gene.* **383**, 81–92.
41. Ugarkovic D. 2009. Centromere-completed DNA: structure and evolution. *Progr. Mol. Subcell. Biol.* **48**, 53–76.
42. Nowick K., Hamilton A.T., Zhang H., Stubbs L. 2010. Rapid sequence and expression divergence suggest selection for novel function in primate-specific FRAB-ZNF genes. *Mol. Biol. Evol.* **27**, 2606–2617.
43. Navratilova H., Koblikova A., Macas J. 2008. Survey of extrachromosomal circular DNA derived from plant satellite repeats. *BMC Plant Biol.* **8**, 90–95.
44. Cameron R.A., Show S.H., Berney F., Chin T.-Y., Ydan Q.-A., Kromer A., Helguero A., Ransick A., Yoh M., Davidson E.N. 2005. An evolutionary constraint: strongly disfavored class of change in DNA sequences during divergence of *cis*-regulatory modules. *Proc. Natl. Acad. Sci. USA.* **102**, 11769–11774.
45. Wagner G.P., Lynch V.J. 2010. Evolutionary novelties. *Curr. Biol.* **20**, R48–R52.
46. Wagner A. 1999. Redundant gene functions and natural selection. *J. Evol. Biol.* **12**, 1–16.
47. Ciccarelli F.D., von Mering C., Suyama M., Harrington E.D., Izauralde E., Bork P. 2005. Complex genomic rearrangements lead to novel primate gene function. *Genome Res.* **15**, 343–351.
48. Prince V.E., Pickett F.B. 2002. Splitting pairs: the diverging fates of duplicate genes. *Nat. New Genet.* **3**, 65–72.
49. Yang J., Lusk R., Li W.-H. 2003. Organismal complexity? Protein complexity and gene duplication. *Proc. Natl. Acad. Sci. USA.* **100**, 15661–15665.
50. Becker K., Nagle J., Canning R., Baddison K., Ozato K., Drew P. 1995. Rapid isolation and characterization of 118 novel C2H2-type finger cDNAs expressed in human brain. *Hum. Mol. Genet.* **4**, 689–691.
51. Eichler E.E., Hoffman S.M., Adamson A.A., Gordon L.F., McCreedy P., Lamerdin J.E., Wohrenweiser H. 1998. Complex beta-satellite repeat structures and the expansion of the zinc finger gene cluster in 19p12. *Genome Res.* **8**, 798–808.
52. Wang X., Grus W.E., Zhang J. 2006. Gene loss during human origin. *PLoS Biol.* **4**, e52.
53. Huberts D.H.E.W., van der Klei I.J. 2010. Moonlighting proteins: an intriguing mode of multitasking. *Biochim. Biophys. Acta.* **1803**, 520–525.
54. Lynch M., Connery J.S. 2000. The evolutionary fate and consequence of duplicate genes. *Science.* **240**, 1151–1155.
55. Boldogkoi Z. 2008. Gene network polymorphism is the raw material of natural selection: the selfish gene network hypothesis. *J. Mol. Evol.* **59**, 340–357.
56. Dawe R.K., Henikoff S. 2006. Centromeres put epigenetics in the driver seat. *Trends Biol. Soc.* **31**, 662–669.
57. Vinogradov A.E. 2003. Selfish DNA is maladaptive: evidence from the plant Red List. *Trends Genet.* **19**, 609–614.
58. Yahisa J.J., Yasmini W.G. 1971. Heterochromatin, satellite DNA, and cell function. Structural DNA of eukaryotes may support and protect genes and aid in speciation. *Science.* **174**, 1200–1209.
59. Micloush G.L.G. 1985. Localized highly repetitive DNA sequences in vertebrate and invertebrate genome. In: *Molecular evolutionary genetics.* Ed. MacIntyre R.G. N.Y.-London: Plenum Press, 241–321.
60. Kramerov D.A., Vasetskii N.S. 2005. Short retroposons in eukaryotic genomes. *Int. Rev. Cytol.* **247**, 165–220.

61. Крамеров Д.А., Васецкий Н.С. 2009. Короткие ре-тропозоны и их использование в филогенетических исследованиях. *Молекуляр. биология*. **43**, 795–806.
62. Adegá F., Guedes-Pinto H., Chaves R. 2009. Satellite DNA in the karyotype evolution of domestic animals – clinical considerations. *Cytogenet. Genome Res.* **126**, 12–20.
63. Palomec T., Lorite P. 2008. Satellite DNA in insects: a review. *Heredity*. **100**, 564–573.
64. Kosak S.T., Groudine M. 2004. Form follows function: the genomic organization of cellular differentiation. *Genes Dev.* **18**, 1371–1384.
65. Zuckerkandl E., Cavalli G. 2007. Combinatorial epigenetics, junk DNA, and the evolution of complex organisms. *Gene*. **390**, 232–242.
66. Pons J., Bruvo B., Petitpierre E., Plohl M., Ugarkovic D., Juan C. 2004. Complex structural features of satellite DNA sequences in the genus *Pimelia* (Coleoptera: Tenebrionidae): random differential amplification from a common “satellite DNA library”. *Heredity*. **92**, 418–427.
67. Корочкин Л.И. 2002. Связь онто- и филогенеза в генетическом освещении. Проблема макромутаций (морфологический и молекулярный аспекты). *Генетика*. **38**, 727–738.
68. Lohe A.R., Brutlag D.C. 1987. Adjacent satellite DNA sequences in *Drosophila*, structure and function. *J. Mol. Biol.* **194**, 171–179.
69. Mravinac B., Plohl M., Ugarkovic D. 2005. Preservation of high sequence conservation of selfish DNAs suggests functional constraints. *J. Mol. Evol.* **61**, 542–550.
70. Pons J., Gillespie R.G. 2004. Solution of satellite DNAs in a radiation of endemic Hawaiian spiders: does concerted evolution of highly repetitive sequences reflect evolutionary history? *J. Mol. Evol.* **59**, 632–641.
71. Saito Y., Edpalina R.R., Abe S. 2007. Isolation and characterization of salmonid telomeric and centromeric satellite DNA sequences. *Genetica*. **131**, 157–166.
72. Sivasundar A., Hey J. 2003. Population genetics of *Caenorhabditis elegans*: the paradox of low polymorphism in a wide-spread species. *Genetics*. **163**, 147–157.
73. Giraud T., Fortuni D., Lecis C., Brygoo Y. 1998. The minisatellite MSB1 in the fungus *Botrytis cinerea*, probably mutates by slippage. *Mol. Biol. Evol.* **15**, 1524–1531.
74. Capriglione T., DeSanto M.G., Odierna G., Olmo E. 1998. An alphoid-like satellite DNA sequence is present in the genome of Lacertidae lizards. *J. Mol. Biol.* **46**, 240–244.
75. Ciobanu D., Grechko V.V., Darevsky I.S., Kramerov D.A. 2004. New satellite DNA in *Lacerta s. str.* lizards (Sauria, Lacertidae): evolutionary pathways and phylogenetic impact. *J. Exp. Zool. Part B: Mol. Develop. Evol.* **302B**, 505–516.
76. Grechko V.V., Ciobanu D.G., Darevsky I.S., Kosushkin S.A., Kramerov D.A. 2006. Molecular evolution of satellite DNA repeats and speciation of lizards of the genus *Darevskia* (Sauria: Lacertidae). *Genome*. **49**, 1297–1307.
77. Robles F., Herran de la R., Ludwig A., Rejon C.R., Rejon M.R., Garrido-Ramos M.A. 2004. Evolution of ancient satellite DNAs in sturgeon genome. *Gene*. **338**, 133–142.
78. Csink A.R., Henikoff S. 1998. Something from nothing: the evolution and utility of satellite repeats. *Trends Genet.* **14**, 2000–2004.
79. Slamovits C.H., Cook J.A., Lessa E.P., Rossi S.M. 2001. Recurrent amplifications and deletions of satellite DNA accompanied chromosomal diversification in South American tuko-tukos (genus *Ctenomys*, Rodentia: Octodontidae): a phylogenetic approach. *Mol. Biol. Evol.* **18**, 1708–1719.
80. Dover G. 1982. Molecular drive: a cohesive mode of species evolution. *Nature*. **299**, 111–117.
81. Alkan C., Ventura M., Arenidiakono N., Rocchi M., Sahinalo S.C., Eichler E.E. 2007. Organization and evolution of primate centromeric DNA from whole-genome shotgun sequence data. *PLoS Comput. Biol.* **3**, 1807–1818.
82. Jarmuz M., Grotzbach C.D., Baily K.D., Bandyopadhyay R., Shaffer L.G. 2007. The evolution of satellite III DNA subfamilies among primates. *Am. J. Hum. Genet.* **80**, 495–501.
83. Cardone M.F., Ballarati L., Ventura M., Rocchi M., Marozzi A., Ginelli E., Meneveri R. 2004. Evolution of beta satellite DNA sequences: evidence for duplication-mediated repeat amplification and spreading. *Mol. Biol. Evol.* **81**, 1792–1799.
84. Alexandrov T., Kazakov A., Tumaneva I., Shepelev V., Yurov Y. 2001. Alpha-satellite DNA of primates: old and new families. *Chromosoma*. **110**, 253–266.
85. Shepelev V.A., Alexandrov A.A., Yurov Y.B., Alexandrov I.A. 2009. The evolutionary origin of man can be traced in the layers of defunct ancestral alpha satellite flanking the active centromeres of human chromosomes. *PLoS Genet.* **5**, e1000641.
86. Zhang Yu., Luo H., Ruder A., Zhang Ya. 2004. Evolution of the tandem repeats in thymidilate synthase enhancer region (TSER) in primates. *Gene*. **338**, 47–54.
87. Kuhn G.C.S., Sene F.M., Moteira-Filho O., Schwarzacher T. 2008. Sequence analysis? Chromosomal distribution and long-range organization show that rapid turnover of new and old satellite pBuM satellite DNA repeats leads to different patterns of variation in seven species of *Drosophila buzzatti* cluster. *Chromosome Res.* **16**, 307–324.
88. Schreider D.R., Hahn M.W. 2010. Gene copy-number polymorphism in nature. *Proc. R. Soc. B*. On-line, doi:10.1098./rspb.2010.1180.
89. Navajas-Perez R., Quesada del Bosque M.E., Garrido-Ramos M.A. 2009. Effect of location? Organization, and repeat-copy number in satellite evolution. *Mol. Genet. Genome*. **282**, 395–406.
90. Heslop-Harrison J.S., Brandes A., Schwarzacher T. 2003. Tandemly repeated DNA sequences and centromeric chromosomal regions of *Arabidopsis* species. *Chromosome Res.* **11**, 241–253.
91. Guenatry M., Bailly D., Almouzni G. 2004. Mouse centric and pericentric repeats form distinct functional heterochromatin. *J. Cell Biol.* **166**, 493–505.
92. Schueler M.G., Sullivan B. 2006. Structural and functional dynamics of human centromeric chromatin. *Annu. Rev. Genom. Hum. Genet.* **7**, 301–313.



93. Smith C.D., Shu S.Q., Mungdall C.J., Karpen G.H. 2007. Annotation of *Drosophila melanogaster* heterochromatin. *Science*. **316**, 1586–1591.
94. Plohl M., Luchetti A., Mestrovic N., Mantovani B. 2008. Satellite DNA between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromere heterochromatin. *Gene*. **409**, 72–82.
95. MacGregor H.C., Sessions S.K. 1986. The biological significance of variation in satellite DNA and heterochromatin in newt of the genus *Triturus*: the evolutionary perspective. *Phyl. Trans. Roy. Soc. B*. **312**, 243–259.
96. Kim J.-H., Ebersol T., Kouprina N., Noskov V.N., et al. 2009. Human gamma-satellite DNA maintains open chromatin structure and protects a transgene from epigenetic silencing. *Genome Res.* **19**, 533–544.
97. Suzuki M., Kasai K., Saeki Y. 2006. Plasmid DNA sequences present in conventional herpes simplex virus amplicon vectors cause transgene silencing by forming inactive chromatin. *J. Virol.* **80**, 3293–3300.
98. Shestakova E.A., Mansuroglu Z., Mokrani H., Ghinea N., Bonnefoy E. Transcription factor YY1 associates with pericentromeric Y-satellite DNA in cycling but not in quiescent (Go) cells. *Nucl. Acids Res.* **32**, 4390–4399.
99. Heartline M.W., Knoll J.H.M., Latt S.A. 1988. Chromosomal instability associated with human alphoid DNA transfected into the Chinese hamster genome. *Mol. Cell Biol.* **8**, 3611–3618.
100. Brahmachari S.K., Meera G., Sarkar P.S., Shaligram U., Pataskar S. 1995. Small repetitive sequences in the genome: structure and functional significance. *Electrophoresis*. **16**, 1705–1714.
101. Maiorano D., Cece R., Badaracco G. 1997. Satellite DNA from brine shrimp *Artemia* affects the expression on flanking gene in yeast. *Gene*. **189**, 13–18.
102. Henikoff S., Dalal Y. 2005. Centromeric heterochromatin: what makes it unique? *Curr. Opin. Genet.* **15**, 177–184.
103. Rouleux-Bennin F., Renault S., Bigot Y., Periquet G. 1996. Transcription of four satellites in *Diprion pini* (Hymenoptera, Symphyta, Diprionidae). *Eur. J. Biochem.* **238**, 752–759.
104. Гвоздев В.А. 2003. Мобильные элементы и явление РНК-интерференции. *Генетика*. **39**, 228–233.
105. Li Y.X., Kirby M.L. 2003. Coordinated and conserved expression of alphoid repeat-tagged coding sequences. *Dev. Dynamics*. **228**, 72–81.
106. Trepicchio W.L., Kroutiris T.G., 1992. Members of the *rel/NF-κB* family of transcriptional regulatory proteins bind the *HRAS1* minisatellite DNA sequences. *Nucl. Acids Res.* **20**, 2427–2434.
107. Попов А.В., Смирнов А.Ф., Сучкова И.О., Баранова Т.В., Сорокин Ф.В., Гайцкохи В.С., Паткин Е.Л. 2000. Моделирование хроматиновых районов у трансгенных животных. *Генетика*. **36**, 1119–1125.
108. Сучкова И.О., Сломинская Н.А., Кустова М.Е., Баранова Т.В., Голубков В.И., Сорокин А.В., Васильев В.Б., Паткин Е.А. 2004. Анализ нестабильности повторяющихся единиц чужеродной центромерной сателлитной ДНК у трансгенных мышей и трансфектных клетках. *Генетика*. **40**, 1034–1045.
109. Gisselson D., Hoglund M., Mertens F., Mandahl N. 1999. Variable stability of chromosomes containing amplified  $\alpha$ -satellite sequences in human mesenchymal tumors. *Chromosoma*. **108**, 271–277.
110. Harrington J.J., van Bokkelen G., Mays R.W., Gustashaw K., Willard H.F. 1997. Formation of *de novo* centromeres and construction of first generation human artificial microchromosomes. *Nat. Genet.* **15**, 345–355.
111. Lambert S., Saintigny Y., Delacote F., Amiot F., Chaput B., Lecomte M., Huck S., Bertrand P., Lopez B.S. 1999. Analysis of intrachromosomal homologous recombination in mammalian cell, using tandem repeat sequences. *Mutat. Res.* **433**, 159–168.
112. Lobov I.B., Tsutsui K., Mitchell., Podgornaya O.I. 2001. Specificity of SAF-A and lamin-B binding *in vitro* correlates with the satellite bending state. *J. Cell. Biochem.* **83**, 218–229.
113. Romanova L.Y., Deriagin G.V., Mashkova T.D., Tumeneva I.G., Mushegian A.R., Kisselev L.L., Alexandrov I.A. 1996. Evidence for selection of evolution of  $\alpha$ -satellite DNA: the central role of CENP-B/pJ $\alpha$  binding region. *J. Mol. Biol.* **261**, 334–340.
114. Razin S.V. 1987. DNA interaction with nuclear matrix and spatial organization of replication and transcription. *BioEssays*. **8**, 27–37.
115. Kuznetsova I.S., Erukashvily N.I., Noniashvily E.M., Shatrova A.N., Aksenov N.D., Zenin V.V., Dyban A.P., Podgornaya O.I. 2007. Evidence for the existence of satellite DNA-containing connection between metaphase chromosomes. *J. Cell Biochem.* **101**, 1046–1061.
116. Ulanovsky L.E., Trifonov E.N. 1987. Estimation of wedge components in curved DNA. *Nature*. **326**, 720–722.
117. Martinez-Balbas A., Rodriguez-Campos A., Garcia-Ramirez M., Carrera P., Aymami J., Azorin F. 1990. Satellite DNAs constrain sequences that induced curvature. *Biochemistry*. **29**, 2342–2348.
118. Horvath J.E., Gnilden C.C., Valente R.U., Eichler M.Y., Ventura M., McPherson J.D., Graves T.A., Wilson R.K., Schwartz S., Rocchi M., Eichler E.E. 2005. Punctuated duplication seedings events during the evolution of human chromosome 2p11. *Genome Res.* **15**, 914–927.
119. Horvath J.E., Schwartz S., Eichler E.E., 2000. The mosaic structure of human pericentromeric DNA: a strategy for characterizing complex regions of the human genome. *Genome Res.* **10**, 839–852.
120. Опарина Н.Ю., Лакруа М.-Э., Рычков А.А., Машкова Т.Д. 2003. Мозаичная структура сегментных дупликаций в геноме человека образуется при участии тандемных и диспергированных повторов. *Молекуляр. биология*. **37**, 228–233.
121. Лакруа М.-Э., Опарина Н.Ю., Машкова Т.Д. 2003. Сегментная дупликация в геноме человека. *Молекуляр. биология*. **37**, 212–220.
122. Cellamare A., Catacchio C.R., Alkan C., Gianuzzi G., Antonacci F., Cardone M.F., Della Valle G., Malig M., Rocchi M., Eichler E.E., Ventura M. 2009. New insights into centromere organization and evolution

- from white-cheeked gibbon and marmoset. *Mol. Biol. Evol.* **26**, 1889–1890.
123. Rudd M.K., Wray G.A., Willard H.F. 2006. The evolutionary dynamics of  $\alpha$ -satellite. *Genome Res.* **16**, 88–96.
  124. Pavelitz T., Liao D., Wiener A.M. 1999. Concerted evolution of the tandem array encoding primate U2 snRNA (the RNU2 locus) is accompanied by dramatic remodeling of the junctions with flanking chromosomal sequences. *EMBO J.* **18**, 3783–3792.
  125. Mirkin S.M. 2008. Discovery of alternative DNA structures: a heroic decade. *Front Biosci.* **13**, 1064–1071.
  126. Bolshoy A., McNamara P., Harrington R.E., Trifonov E.N. 1991. Curved DNA without A-A: experimental estimation of all 16 DNA wedge angle. *Proc. Natl. Acad. Sci. USA.* **88**, 2312–2316.
  127. Бугреев Д.В., Бунева В.Н., Невинский Г.А. 2003. Механизмы специфического расщепления суперскрученной ДНК ДНК-топоизомеразой I человека: влияние структуры лиганда на каталитическую стадию реакции. *Молекуляр. биология.* **37**, 325–339.
  128. Herbert A., Rich A. 1996. The biology of left-handed Z-DNA. *J. Biol. Chem.* **271**, 11595–11598.
  129. Trifonov E.N. 1991. DNA in profile. *Trends Biochem. Soc.* **16**, 467–470
  130. Canapa A., Cerioni P.N., Barucca M., Olmo E., Caputo V. 2002. A centromeric satellite DNA may involved in heterochromatin compactness in gobiid fishes. *Chromosome Res.* **10**, 297–304.
  131. Vogt P. 1992. Code domains in tandem repetitive DNA sequence structures. *Chromosoma.* **101**, 585–589.
  132. Pasero P., Sjakste N., Blettry C., Got C., Marilley M. 1993. Long-range organization and sequence-directed curvature of *Xenopus laevis* satellite 1 DNA. *Nucl. Acids Res.* **21**, 4703–4711.
  133. Barcello T., Pons J., Petitpierre E., Barjan I., Portugal J. 1997. Polymorphic curvature of satellite DNA in three subspecies of the beetle *Pimelia sparsa*. *Eur. J. Biochem.* **244**, 318–324.
  134. Fitzgerald D.J., Dryden G.L., Bronson E.C., Williams J.S., Anderson J.N. 1994. Conserved patterns of bending in satellite and nucleosome positioning DNA. *J. Biol. Chem.* **269**, 21303–21314.
  135. *DNA Conformation and Transcription. Molecular Biology Intelligence Unit.* 2005. Ed. Takashi Ohyama, USA: Springer Science+Business Media, Inc., 211 P.
  136. Huppert J.L. 2008. Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chem. Soc. Rev.* **37**, 1375–1384.
  137. Mirkin S.M. 2007. Expandable DNA repeats and human disease. *Nature.* **447**, 932–940.
  138. Siddiqui-Jain A., Grand C.L., Bearss D.J., Hurley D.R. 2002. Direct evidence for a G-quadruplex. Is a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. USA.* **99**, 11593–11598.
  139. Kerwin S.M. 2000. G-quadruplex DNA as a target for drug design. *Curr. Pharm. Design.* **6**, 441–471.
  140. Burge S., Parkinson G.N., Hazel P., Todd A.K., Neidle S. 2006. Quadruplex DNA: sequence topology and structure. *Nucl. Acids Res.* **34**, 5402–5415.
  141. Paris C., Geinguenaud F., Gonyette C., Liqueur J., Lacoste J. 2007. Mechanism of copper mediated triple helix formation at neutral pH in *Drosophyla* satellite repeats. *Biophys. J.* **92**, 2498–2506.
  142. Henikoff S., Ahmad K., Malik H.S. 2001. The centromere paradox: stable inheritance with rapidly evolved DNA. *Science.* **293**, 1098–1102.
  143. Jonstrup A.T., Thomsen T., Wang Y., Knudsen B.P., Koch J., Andersen A.H. 2008. Hairpin structure formed by  $\alpha$ -satellite DNA of human centromeres are cleaved by human topoisomerase II  $\alpha$ . *Nucl. Acids Res.* **36**, 6165–6174.
  144. Slama-Schwok A., Zakrzewska K., Leger G., Leroux Y., Takahashi M., Kas E., Debey P. 2000. Structural changes induced by binding of the high-mobility group I protein to a mouse satellite DNA sequence. *Biophys. J.* **78**, 2543–2559.
  145. Стегний В.Н. 2006. Эволюционное значение архитектоники хромосом как формы эпигенетического контроля онто- и филогенеза эукариот. *Генетика.* **42**, 1215–1224.
  146. Hurst D.D., Werren J.H. 2001. The role of selfish genetic elements in eukaryotic evolution. *Nature Rev. Genet.* **2**, 597–606.
  147. Amor D., Bentley K., Ryan J., Perry J., Wong L., Slater H., Choo K.H.A. 2004. Human centromere repositioning “in progress”. *Proc. Natl. Acad. Sci. USA.* **101**, 6542–6547.
  148. Bird A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21.
  149. Low J.A., Jacobsen S.E. 2010. Establishing, maintaining and modifying DNA methylation patterns in plant and animals. *Nat. Rev. Genet.* **11**, 204–220
  150. Usakin L., Abad J., Vagin V.V., de Pablos B., Villasante A., Gvozdev V.A. 2007. Transcription of the 1.688 satellite DNA family is under the control of RNA interference machinery in *Drosophyla melanogaster* ovaries. *Genetics.* **176**, 1343–1356.
  151. Rudert F., Bronner S., Garnier J.-M., Dolle P. 1995. Transcription from opposite strands of  $\gamma$ -satellite DNA are differently expressed during mouse development. *Mamm. Genome.* **6**, 76–83.
  152. Schmid C.W. 1998. Does SINE evolution preclude Alu function? *Nucl. Acids Res.* **26**, 4541–4550.
  153. Valgarsdottir P., Choidi I., Giordano M., Poggi A., Bazzini S., Ghigna C., Riva S., Biamondi G. 2008. Transcription of SatIII non-coding RNA is a general stress response in human cells. *Nucl. Acids Res.* **36**, 423–434.
  154. Coats S.R., Zhang Y., Epstein L.M. 1994. Transcription of satellite 2 DNA from the newts driven by a snRNA type of promoter. *Nucl. Acids Res.* **22**, 4697–4704.
  155. Ferbeyre G., Smith J.M., Cedergren R. 1998. Shistosome satellite DNA encodes active hammerhead ribozyme. *Mol. Cell Biol.* **18**, 3889–3888.
  156. Luzi E., Eckstein F., Barsacchi G. 1997. The newt ribozyme is part of riboprotein complex. *Proc. Natl. Acad. Sci. USA.* **94**, 9711–9716.

157. Коган Г.Л., Гвоздев В.А. 2002. Молекулярная эволюция тандемных повторов гетерохроматина в связи с их функцией в геноме *Drosophila melanogaster*. *Генетика*. **38**, 710–718.
158. Aravin A.A., Naumova N.M., Tulin A.V., Vagin V.V., Rozovsky Y.M., Gvozdev V.A. 2001. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *Drosophila melanogaster* germline. *Curr. Biol.* **11**, 1017–1027.
159. Wu C.-I., True J.R., Johnson N. 1989. Fitness reduction associated with the deletion of a satellite. *Nature*. **341**, 248–251.
160. Wu C.-I., Little T.W., Wu M.-L., Lin G.-A. 1998. Association between a satellite DNA sequence and the Responder of Segregation Distorter in *Drosophila melanogaster*. *Cell*. **54**, 178–189.
161. Regelson M., Eiler C.D., Horvath S., Marahrens Y. 2006. A link between repetitive sequences and gene replication time. *Cytogenet. Genome Res.* **112**, 184–193.
162. Saitoh Y., Miyamoto N., Okada T., Gondo Y., Showguchi-Miyata J., Hadano S., Ikeda J.F. 2000. RS447 human megasatellite tandem repetitive sequence encodes a novel deubiquitinating enzyme with a functional promoter. *Genomics*. **67**, 291–300.
163. Fondon III J.W., Garner G.R. 2004. Molecular origins of rapid and continuous morphological evolution. *Proc. Natl. Acad. Sci. USA*. **101**, 18058–18063.
164. Kidwell M.G., Lisch D. 2001. Perspective: transposable elements: parasitic DNA and genomic evolution. *Evolution*. **55**, 1–24.
165. Kunarso G., Chia N.Y., Jeyakani I., Hwang C., Lu X., Chan Y.S., Ng H.H., Bourque G. 2010. Transposable elements have rewired the corw regulatory network of human embryonic stem cells. *Nat. Genet.* **42**, 631–634.
166. Oliver K.R., Greene W.K. 2009. Transposable elements: powerful facilitators of evolution. *BioEssays*. **31**, 703–714.
167. Guo Y., Levin H.L. 2010. High-throughput sequencing of retroposon integration provides a saturated profile of target activity in *Schizosaccharomyces pombe*. *Genome Res.* **20**, 239–248.
168. Thomson S.J.P., Goh F.G., Banks H., Krausgruber T., Kotenko S.V., Foxwell B.H., Udalova I.A. 2009. The role of transposable elements in the regulation of IFN- $\lambda$ 1 gene expression. *Proc. Natl. Acad. Sci. USA*. **106**, 11564–11569.
169. Walser J.-C., Chen B., Feder M.E. 2006. Heat-shock promoters: targets for evolution by P transposable elements in *Drosophila*. *Plos Genet.* **2**, 1541–1555.
170. Евгеньев М.Б., Мнджоян Е.И., Зеленцова Е.С., Шостак Н.Г., Лезин Г.Т., Великодворская В.В., Полуэктова Е.В. 1998. Мобильные элементы и видообразование. *Молекуляр. биология*. **32**, 184–192.
171. Xing J., Zhang Y., Han K., et al. 2009. Mobile elements create structural variation: analysis of a complete human genome. *Genome Res.* **19**, 1516–1526.
172. Moran N.A., Jarvik T. 2010. Lateral transfer of genes from fungi underlies carotenoid production in aphids. *Science*. **328**, 624–627.
173. Muotry A.R., Marchetto M.C.N., Coufal N.G., Gage F.H. 2007. The necessary junk: new function for transposable elements. *Hum. Mol. Genet.* **16**, K159–K167.
174. Wheelan S.J., Aizawa Y., Han J.S., Boeke J.D. Gene-breaking: a new paradigm for human retrotransposon-mediated gene evolution. 2005. *Genome Res.* **15**, 1073–1078.
175. Allen E., Horvath S., Tong F., Kraft P., Spiteri E., Riggs A.D., Marahrens Y. 2003. High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes. *Proc. Natl. Acad. Sci. USA*. **100**, 9940–9946.
176. Шпиз С.Г., Калмыкова А.И. 2007. Структура теломерного хроматина у *Drosophila*. *Биохимия*. **72**, 759–773.
177. O'Neill R.J.W., O'Neill M.J., Graves J.A.M. 1998. Undermethylation associated with retroelement activation and chromosome remodeling in an interspecific mammalian hybrid. *Nature*. **393**, 68–72.
178. Fontdevila A. 2005. Hybrid genome evolution by transposition. *Cytogenet. Genome Res.* **110**, 49–55.
179. Brown J.R. 2003. Ancient horizontal gene transfer. *Nature*. **4**, 121–132.
180. Gilbert C., Schaack S., Pace II J.K., Brindley P.J., Feschotte C. 2010. A role for host parasite interaction in the horizontal transfer of transposon across phyla. *Nature*. **464**, 1347–1350.
181. Ратнер В.А., Васильева Л.А. 1993. Мобильные генетические элементы и эволюция генома. В кн.: “Современные проблемы эволюции”. под ред. Л.П. Татаринова. М.: Наука, 43–60.
182. Любомирская Н.В., Ильин Ю.В. 1999. Мобильные генетические элементы эукариот: прошлое, настоящее, будущее. *Молекуляр. биология*. **33**, 958–968.
183. Henikoff S., Malik H.S. 2002. Centromere: selfish drivers. *Nature*. **417**, 227.
184. Беляева Е.С., Пасюкова Е.Г., Гвоздев В.А. 1994. “Адаптивные транспозиции” ретротранспозонов в геноме *Drosophila melanogaster*, сопровождающиеся увеличением приспособленности организма. *Генетика*. **30**, 725–730.
185. Pasyukova E.G., Nuzhdin S.V., Morozova T.V., Mackay T.F.C. 2004. Accumulation of transposable elements in the genome of *Drosophila melanogaster* is associated with a decrease in fitness. *J. Hered.* **95**, 284–290.
186. Petrov D.A., Schultzman J.I., Hartl D.L., Lozovskaya E.R. 1995. Diverse transposable elements are mobilized in hybrid dysgenesis in *Drosophila virilis*. *Proc. Natl. Acad. Sci. USA*. **95**, 8050–8054.
187. Ogiwara I., Miya M., Oshima K., Okada N. 2002. V-SINEs: a new superfamily of vertebrate SINEs that are widespread in vertebrate genomes and retain a strongly conserved segment within each repetitive unit. *Genome Res.* **12**, 316–324.
188. Grechko V.V., Kosushkin S.A., Borodulina O.R., Butaeva F.G., Darevsky I.S. 2011. Short interspersed elements (SINEs) of Squamate reptiles (Squam1 and Squam2): structure and phylogenetic significance. *J. Exp. Zool. (Mol. Dev. Evol.)* **316**, 212–226

189. Dewannieux M., Esnault C., Hedmann T. 2003. LINE-mediated retrotransposition of marked Alu-sequences. *Nat. Genet.* **35**, 41–48.
190. van den Lagemaat L.N., Landry J.-R., Mager D.L.M., Medstrand P. 2003. Transposable elements in mammalian promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.* **19**, 530–536.
191. Polak P., Domany E. 2006. Alu elements contain many binding sites for transcription factors and may play a role of developmental process. *BMC Genomics.* **7**, 133–135.
192. Lin L., Shen S., Iue A., Cau J.J., Jiang P., Davidson B.L., Xing Y. 2008. Diverse splicing patterns of exonized Alu elements in human tissues. *PloS Genet.* **4**, e1000225.
193. Clark L.A., Wahl J.M., Rees C.A., Murphy K.E. 2006. Retroposon insertion in *SYLV* is responsible for merle patterning of the domestic dog. *Proc. Natl. Acad. Sci USA.* **103**, 1376–1381.
194. Wang W., Kirkness E.F. 2005. Short interspersed elements (SINEs) are a major source of canine genomic diversity. *Genome Res.* **15**, 1798–1808.
195. Grewal S.I., Moared D. 2003. Heterochromatin and epigenetic controlled gene expression. *Science.* **301**, 798–802.
196. Tomilin N.V. 2008. Regulation of mammalian gene expression by retrovirus and non-coding tandem repeats. *BioEssays.* **30**, 338–348.
197. Li T.-H., Schmid K. 2004. Alu's dimeric consensus sequence destabilizes its transcripts. *Gene.* **324**, 191–200.
198. Klenov M.S., Lavrov S.A., Stolyarchenko A.L., Rysansky S.S., Aravin A.A., Gvozdev V.A. 2007. Repeat-associated siRNA cause chromatin silencing of retroposons in the *Drosophila melanogaster* germline. *Nucl. Acids Res.* **35**, 5430–5438.
199. Oei S.L., Babich V.S., Kazakov V.T., Usmanova N.M., Kropotkov A.V., Tomilin N.V. 2004. Clusters of regulatory signals for RNA polymerase II transcription associated with Alu family repeats and CpG islands in human promoters. *Genomics.* **83**, 873–882.
200. Smalheiser N.R., Torvik V.I. 2005. Mammalian microRNAs derived from genomic repeats. *Trends Genet.* **21**, 322–326.
201. Kipling D., Warburton P.E. 1997. Centromeres, CENP-B, and tigger too. *Trends Genet.* **13**, 141–145.
202. Хитринская И.Ю., Степанов В.А., Пузырев В.П. 2003. Alu-повторы в геноме человека. *Молекуляр. биология.* **37**, 382–391.
203. Krealing J., Graveley B.R. 2004. The origin and implications of *Alu* alternative splicing. *Trends Genet.* **20**, 1–4.
204. Deininger P.L., Moran J.V., Batzer M.A., Kazazian Jr. H.H. 2003. Mobile elements and mammalian genome evolution. *Curr. Opin Genet. Develop.* **13**, 651–658.
205. Burt A. 2003. Site specific selfish genes as tools for control and genetic engineering of natural populations. *Proc. R. Soc. Biol. Sci. (London).* **270**, 921–928.