

УДК 577.241

## ГЕНЫ МАЛЫХ ЯДРЫШКОВЫХ РНК

© 2007 г. Ю. А. Макарова, Д. А. Крамеров

Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва 119991;  
факс: (495) 135-14-05; e-mail: kramerov@eimb.ru

Поступила в редакцию 03.05.2006 г.

Малые ядрышковые РНК (мякРНК) представляют собой одну из наиболее многочисленных и хорошо изученных групп, не кодирующих белок РНК. В комплексе с белками мякРНК осуществляют две самые распространенные модификации нуклеотидов рРНК: метилирование рибозы по 2'-ОН и псевдоуридилрование. Несмотря на то, что механизмы модификаций и строение мякРНК весьма консервативны, гены мякРНК проявляют удивительное разнообразие организации. Помимо независимо транскрибирующихся генов встречаются гены, локализованные в интронах других генов, гены, формирующие кластеры, которые транскрибируются с общего промотора, а также кластеры генов, локализованные в интронах. Интересно, что у разных таксонов встречается преимущественно один из способов организации. У позвоночных мякРНК образуются из интронов генов белков, однако небольшая группа мякРНК кодируется интронами генов некодирующих РНК.

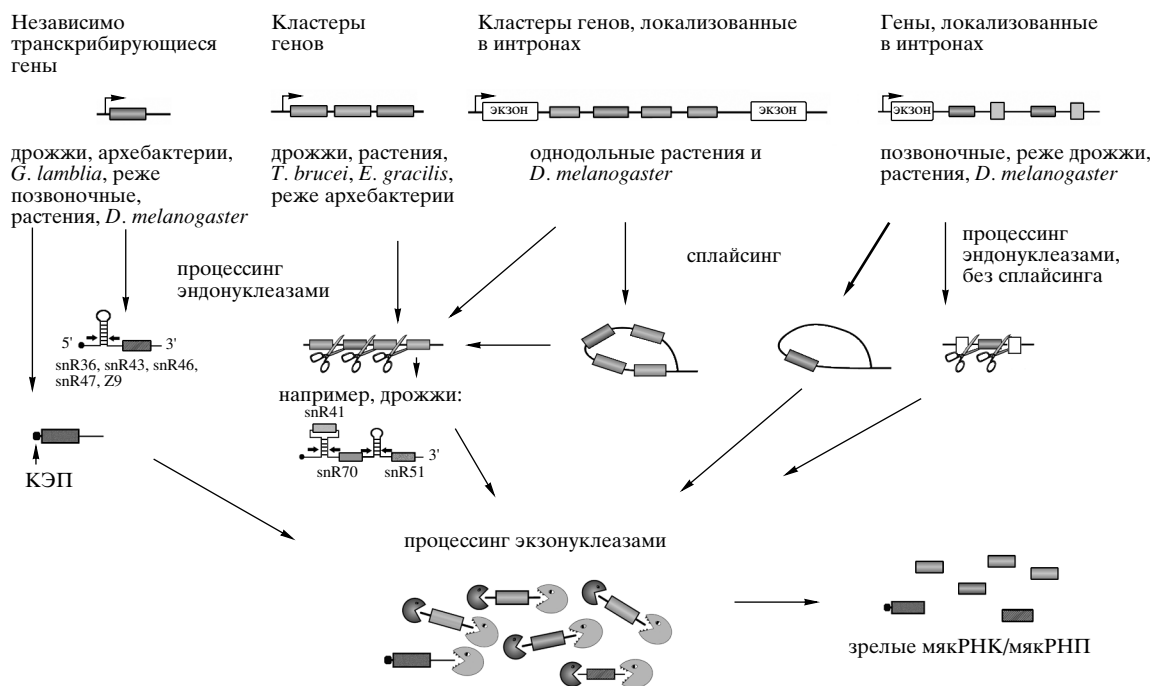
Значительная часть транскрибируемых последовательностей генома не кодирует белок. В последние несколько лет интерес к таким последовательностям необычайно возрос благодаря обнаружению новых классов некодирующих РНК. Оказалось, что эти РНК выполняют в клетке удивительно широкий спектр функций, используя не известные ранее механизмы. Они вовлечены в регуляцию транскрипции и трансляции, репрессию мобильных генетических элементов и другие регуляторные пути – процессы, участие в которых считалось ранее прерогативой белков. Описанные на сегодняшний день некодирующие РНК по самым разным оценкам представляют собой только вершину айсберга (см., например, [1, 2]).

Одним из классов некодирующих РНК являются малые ядрышковые РНК (мякРНК, англ. – small nucleolar RNAs, snoRNAs). Несколько видов мякРНК необходимо для разрезания пре-рРНК, тогда как функция большинства состоит в определении нуклеотидов рРНК, малых ядерных РНК и мРНК, которые будут модифицированы одним из двух способов: метилированием 2'-ОН-группы рибозы или превращением уридина в псевдоуридин [3]. Это две самые распространенные модификации, число которых только в рРНК превышает сотню. В соответствии с видом модификации и наличием консервативных элементов в нуклеотидной последовательности мякРНК делят на два семейства: С/Д (box С/Д snoRNAs) и Н/АСА (box Н/АСА snoRNAs). Некоторые мякРНК семейства С/Д участвуют в разрезании пре-рРНК, тогда как прочие направляют 2'-О-метилирование. В состав нуклеотидной последовательности всех С/Д мякРНК входят консервативные элементы С (UGAUGA), D (CUGA) и их часто

вырожденные копии С' и D'. Помимо этого, в состав мякРНК семейства входит так называемый антисенс-элемент – последовательность длиной 10–20 нуклеотидов, полностью комплементарная соответствующей последовательности модифицируемой РНК и способная взаимодействовать с ней. 2'-О-метилированию подвергается остаток, входящий в образующуюся РНК/РНК спираль и отделенный четырьмя нуклеотидами от элемента D и/или D' [4]. Метилирование осуществляют белки С/Д мякРНК.

Почти все мякРНК семейства Н/АСА определяют нуклеотид РНК, который будет подвергнут псевдоуридилрованию, и только небольшое число РНК этого семейства участвует в разрезании пре-рРНК. Все Н/АСА мякРНК содержат консервативные последовательности Н (ANANNA), АСА (ACA) и, подобно С/Д мякРНК, один или реже два антисенс-элемента. Выбор уридилового остатка, который будет модифицирован, как и в случае С/Д мякРНК, определяется благодаря комплементарным взаимодействиям антисенс-элемента с нуклеотидной последовательностью модифицируемой РНК, причем собственно реакцию псевдоуридилрования осуществляет один из белков, связанных с Н/АСА мякРНК [5].

К мякРНК также относят РНК, входящую в состав РНКазы MRP, и иногда похожую на нее РНК РНКазы Р. РНКазы MRP осуществляет разрезание пре-рРНК в сайте А3 первого внутреннего транскрибируемого спейсера, давая зрелую форму 5.8 S рРНК [6], хотя ее функции этим не исчерпываются. РНКазы Р вовлечена в удаление 5'-концевой лидерной последовательности у пре-рРНК [6].



**Рис. 1.** Способы организации генов и виды процессинга мякРНК у разных систематических групп. Сайты разрезания эндорибонуклеазы Rnt1p отмечены маленькими стрелками (см. раздел 4) (по [8, 9], с изменениями).

Более подробно строение и функции мякРНК обсуждаются в другой статье авторов [7]. Основной темой данного обзора служат способы организации генов мякРНК у разных таксономических групп.

## 1. ГЕНЫ мякРНК

Гены мякРНК проявляют удивительное разнообразие организации: некоторые гены представляют собой самостоятельные транскрипционные единицы, тогда как другие расположены в интронах белковых генов и процессируются при сплайсинге. Кроме того, гены мякРНК могут образовывать кластеры, с которых синтезируется полицистронный транскрипт. Иногда такой кластер локализован в интроне другого гена. Встречаются и более экзотические варианты. Интересно, что в различных таксонах встречается преимущественно один из способов организации. Так, у позвоночных почти все мякРНК кодируются однокопийными генами, расположенными в интронах; у дрожжей большинство генов одиночные, а у растений и простейших они собраны в кластеры (рис. 1).

### 1.1. Гены мякРНК позвоночных

Почти все мякРНК позвоночных кодируются очень необычным способом: их гены расположе-

ны в интронах генов белков, причем в одном интроне присутствует только один ген мякРНК (рис. 1). Только гены U3, U8, U13, а также гены РНК РНКаз Р и МРР транскрибируются соответственно РНК-полимеразами II и III со своих собственных промоторов [10]. Кроме них, со своих промоторов транскрибируются РНК-полимеразой II две необычные С/D мякРНК, содержащие дополнительные копии элементов С и D [11]. Большинство генов мякРНК однокопийные. К немногочисленным исключениям относится представленный множеством копий жизненно важный ген U3 РНК [10].

### 1.2. Гены мякРНК дрожжей

У дрожжей, как и у животных, почти все гены мякРНК однокопийные, но, в отличие от животных, большинство генов мякРНК дрожжей представляет собой независимые транскрипционные единицы, хотя известно семь генов, локализованных в интронах, и пять полицистронных кластеров (два дицистронных, два трицистронных и один гептацистронный [12]). Интересный случай представляет ген U86 РНК, закодированный внутри открытой рамки считывания [13]. Образование U86 РНК и мРНК с этой рамки – альтернативные процессы. У позвоночных U86 РНК кодируется интроном.

### 1.3. Гены мякРНК растений

У растений большинство генов С/D и Н/АСА мякРНК образуют кластеры (см. рис. 1), которые могут быть гомогенными, т.е. состоять из двух-четырёх копий одного гена мякРНК, или гетерогенными, когда в состав кластера входят несколько разных генов мякРНК. В последнем случае один или несколько генов гетерогенного кластера тоже могут быть представлены несколькими копиями. Кластеры состоят обычно из 2–5 генов и рассеяны по всем хромосомам [14], причем некоторые представлены несколькими (2–5) копиями [15]. У двудольных почти все кластеры генов мякРНК расположены в промежутках между другими генами и транскрибируются в виде полицистронных РНК со своих собственных промоторов. Например, у *Arabidopsis thaliana* 133 из 175 известных генов мякРНК организованы в 49 кластеров (включая несколько кластеров, представленных двумя и тремя копиями), из которых только четыре расположены в интронах белковых генов. У однодольных, хотя кластерная организация генов остается предпочтительной, картина несколько иная: более половины кластеров генов мякРНК расположены в интронах генов белков [15]. Этот уникальный способ организации не встречается у позвоночных и дрожжей (см. рис. 1). У обоих классов имеются сравнительно немногочисленные одиночные гены, закодированные, подобно большинству генов позвоночных, в интронах [8].

Несколько генов мякРНК (около 10 [14]) транскрибируются со своих собственных промоторов. Так, гены U3 РНК и РНКазы MRP представляют собой независимые транскрипционные единицы у всех эукариот. Однако у растений, в отличие от животных и дрожжей, ген U3 РНК транскрибируется не РНК-полимеразой II, а РНК-полимеразой III. Промотор, подобно промоторам малых ядерных РНК растений, содержит так называемый USE (англ. *upstream sequence element*) и ТАТА-боксы. Вообще следует отметить некоторое структурное сходство промоторов РНК-полимеразы II и III у растений [16].

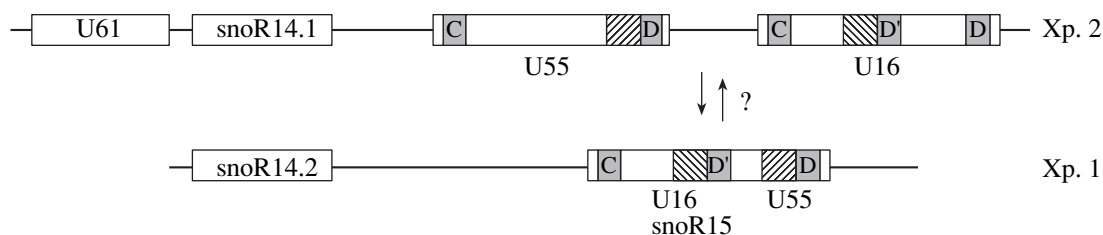
У растений был обнаружен еще один необычный способ кодирования мякРНК: и у двудольных (*A. thaliana*), и у однодольных (*Oryza sativa*) найдены несколько дицистронных генов, образованных генами тРНК и мякРНК (ген мякРНК расположен после гена тРНК). Для их транскрипции используется промотор тРНК [17].

В отличие от позвоночных и дрожжей, около половины мякРНК растений кодируются несколькими генами: например, у *A. thaliana* около 65% генов представлено более чем одной копией [18]. Это отчасти связано с большим числом дупликаций фрагментов генома: гены, кодирующие одну и ту же мякРНК, нередко входят в состав

дублированного кластера. Кроме того, ген мякРНК может претерпевать тандемную дупликацию внутри одного кластера. Уровень сходства между такими дублированными генами сильно варьирует. Вероятно, он пропорционален времени, прошедшему с момента дупликации генов. Часто разные копии отличаются друг от друга лишь несколькими заменами и/или вставками/делециями. Как правило, изменения не затрагивают консервативные последовательности (элементы С, D, С', D' и антисенс-элемент), однако иногда происходят и в них. Впрочем, замены обычно столь незначительны, что вряд ли приводят к утрате геном мякРНК своей функции. Но иногда такие незначительные изменения могут приводить к появлению новых сайтов модификации рРНК. Например, в число изменений, отличающих одну копию гена snoR20 мякРНК *A. thaliana* от другой, входит вставка одного нуклеотида между боксом D и антисенс-элементом. В результате появляется новый сайт модификации рРНК, и, таким образом, два варианта гена snoR20 мякРНК определяют модификацию двух разных (соседних) нуклеотидов в рРНК [14]. Копии одного и того же гена мякРНК могут различаться и весьма существенно, иногда одна из копий может быть представлена всего лишь фрагментом псевдогена [14].

Гены некоторых мякРНК имеют ортологов и у позвоночных, и у дрожжей, и у растений. При этом, как правило, сходство сохраняется только между консервативными последовательностями и антисенс-элементом. Назвать такие гены ортологами можно только достаточно условно. Иногда ситуация осложняется тем, что помимо ортолога существует еще один ген, содержащий еще один антисенс-элемент помимо имеющегося у ортолога. Например, U33 РНК присутствует и у позвоночных, и у растений. Однако у растений есть еще одна РНК, snoR34, содержащая антисенс-элемент U33 РНК и еще один антисенс-элемент [14]. Иногда внутри одного вида две копии гена имеют один общий антисенс-элемент и различаются вторым. Например, мякРНК snoR16.1 и snoR16.2 направляют метилирование U48 в 25S рРНК, но при этом snoR16.1 направляет метилирование еще и U2445, а snoR16.2 – метилирование U36 в 25S рРНК [14].

Можно привести еще один интересный пример: на хромосоме 2 *A. thaliana* расположен кластер, состоящий из четырех генов мякРНК, в том числе из генов U55 и U16 РНК (рис. 2). И U55, и U16 содержат по одному антисенс-элементу. На хромосоме 1 имеется кластер, состоящий из двух генов, один из которых, snoR15, содержит антисенс-элементы и U55, и U16 РНК. Кластеры очевидно родственны, поскольку имеют общий ген (snoR14). Авторы, описавшие эти кластеры [8, 14], полагают, что U55 и U16 РНК образовались в результате дупликации гена РНК snoR15, располо-



**Рис. 2.** Строение двух кластеров генов мьякРНК *A. thaliana*. Антисенс-элементы обозначены штриховками (по [8, 14], с изменениями).

женного на хромосоме 1. Альтернативное объяснение заключается в том, что кластер на хромосоме 1 сначала содержал и U55, и U16 РНК. Затем в гене U55 РНК образовался антисенс-элемент, характерный для U16 РНК: для этого должны были произойти только три замены. Ген U16 РНК был затем утерян [14].

Таким образом, гены мьякРНК растений открывают широкое поле для построения гипотез о возможных путях эволюции генов.

#### 1.4. Гены мьякРНК простейших

Гены мьякРНК трипаносом (класс Mastigophora, подкласс Zoomastigina, отряд Kinetoplastidae), подобно генам растений, образуют кластеры, содержащие мьякРНК обоих семейств и обычно повторенные в геноме несколько раз [19] (см. рис. 1). Все кластеры располагаются в промежутках между генами, полицистронный транскрипт синтезируется РНК-полимеразой II. Таким же образом организованы гены мьякРНК другого представителя того же класса, но относящегося к другому подклассу – Phytomastigina – *Euglena gracilis* [20]. Однако у более близкого трипаносомам таксона, отряда Diplomonadida, относящегося, как и трипаносомы, к подклассу Zoomastigina, а именно у *Giardia lamblia*, гены 20 изученных мьякРНК одиночные, однокопийные, расположены между генами белков и транскрибируются РНК-полимеразой II со своих собственных промоторов [21]. Это интересное наблюдение, поскольку все предыдущие данные позволяли предполагать, что чем более сходны таксоны, тем более похожи предпочтительные способы организации генов мьякРНК.

#### 1.5. Гены мьякРНК дрозофилы

В последнее время выясняется, что и у животных способы организации генов мьякРНК более разнообразны, чем предполагалось ранее: у *Drosophila melanogaster* помимо характерных для животных генов мьякРНК, закодированных в интронах (один интрон – один ген) [22], и некоторого количества независимых генов недавно обнаружены кластеры генов мьякРНК, кодируемые

интронами [23] (рис. 1). В каждом кластере находится до девяти генов. Как правило, кластер состоит из двух-трех генов, представленных несколькими копиями. Ранее такой тип организации был описан только у растений, преимущественно у однодольных. У дрозофил обнаружен еще один необычный способ кодирования мьякРНК: по меньшей мере 5 генов локализовано в экзонах генов белков, так что образование мьякРНК и мРНК – это альтернативные процессы. До сих пор единственный пример такой организации среди эукариот найден у дрожжей (см. раздел 1.2). Кроме того, этот способ используется архебактериями. Еще в нескольких случаях ген мьякРНК расположен в экзоне, подверженном альтернативному сплайсингу, так что иногда мьякРНК оказывается включенной в экзон, иногда – в интрон. Наконец, одна мьякРНК локализована на границе экзона и интрона [24].

#### 1.6. Гены мьякРНК архебактерий

МьякРНК найдены даже у такого далекого от эукариот таксона, как архебактерии, хотя совершенно отсутствуют в бактериях. У архебактерий гены мьякРНК однокопийные и одиночные, хотя известно несколько примеров расположенных рядом генов, транскрибирующихся, вероятно, в виде единого предшественника. Гены мьякРНК локализованы между генами белков и, видимо, имеют собственные промоторы. 10–20% генов мьякРНК расположено необычно: их 5'-концевая часть перекрывается с 3'-концом открытой рамки считывания. Таким образом, последовательность, кодирующая несколько С-концевых аминокислот, одновременно кодирует 5'-концевую часть гена мьякРНК. Наконец, известен пример гена мьякРНК, кодируемого интроном тРНК [25].

#### 1.7. Разнообразие способов организации генов мьякРНК

Обычно у крупных таксонов, например у позвоночных и растений, одна и та же мьякРНК кодируется разными способами. Однако в пределах такого таксона мьякРНК, как правило, кодируется одним способом. Например, мьякРНК U23 кодиру-

ется 12-м интроном гена нуклеолина у всех классов позвоночных. Тем не менее нередки случаи, когда мякРНК меняет генов-хозяев. Так, мякРНК U15 у человека кодируется интронами гена рибосомного белка S3, а у *Xenopus laevis* – интронами гена рибосомного белка S1. Кроме того, мякРНК, подобно U14 мякРНК *X. laevis*, может иметь разных генов-хозяев у одного вида [10]. Такие данные позволили предположить, что мякРНК произошли от мобильных элементов. Однако у генов мякРНК отсутствуют предсказываемые таким сценарием концевые повторы [10].

Возникает еще один вопрос: почему в пределах одного организма разные мякРНК кодируются различными способами. Ответ пока кажется очевидным только для независимо экспрессирующихся генов: так, у позвоночных все независимые гены кодируют мякРНК, участвующие в разрезании рРНК и потому жизненно необходимые. Хотя даже в этом случае все не так очевидно: у позвоночных две обнаруженные недавно C/D мякРНК, кодируемые независимыми генами, направляют 2'-О-метилирование малых ядерных РНК [11], а у дрожжей большинство генов мякРНК независимые. Неясно также, в чем заключается биологический смысл (если он существует) того, что многие ортологичные мякРНК кодируются в разных крупных таксонах по-разному. Самым простым было бы предположение о том, что выбор способа кодирования случаен, однако оно не согласуется с наблюдаемым для мякРНК позвоночных феноменом: большинство генов-хозяев так или иначе вовлечены в трансляцию (см. раздел 2). Непонятно, наконец, с чем вообще связано такое разнообразие организации генов мякРНК – среди других групп генов на сегодняшний день неизвестно ничего похожего.

Известно, что полицистронная организация генов позволяет координированно регулировать их экспрессию. У растений мякРНК организованы в кластеры, а у позвоночных локализованы в интронах генов, кодирующих в основном белки трансляционного аппарата. Эти гены, таким образом, тоже кодируют несколько продуктов, и поэтому тоже в некотором роде являются полицистронными. Похоже, что проблема координированной экспрессии функционально связанных генов у этих таксонов решается по-разному. Следует отметить, что для успешного процессинга почти всех генов мякРНК позвоночных необходим сплайсинг, тогда как у растений процессинг мякРНК не зависит от сплайсинга, что позволяет разнообразить организацию генов мякРНК. В условиях, когда сплайсинг угнетен (например, при тепловом шоке [26]), мякРНК растений продолжают образовываться. Возможно, такая повышенная по сравнению с позвоночными устойчивость продуцирования мякРНК связана с тем, что

растения в гораздо большей степени, чем животные, подвержены экстремальным нагрузкам.

## 2. ГЕНЫ-ХОЗЯЕВА мякРНК ПОЗВОНОЧНЫХ

Большинство генов, из интронов которых образуются мякРНК, кодируют белки, необходимые для трансляции (факторы инициации и элонгации, белки рибосом и т.д.) или вовлеченные в функционирование ядрышка. Такое расположение генов мякРНК представляется неслучайным и необходимо, вероятно, для координации экспрессии транскриптов РНК-полимераз I и II: интенсивность образования рРНК и в конечном итоге образование рибосом оказываются связанными с уровнем транскрипции мРНК белков трансляционного аппарата [10].

Почти все гены-хозяева мякРНК позвоночных принадлежат к так называемому 5'-ТОР семейству (5'-terminal oligouridine) генов домашнего хозяйства [27]. К этому семейству относятся гены рибосомных белков, некоторых факторов трансляции и ряд других генов. Первым нуклеотидом в их мРНК является цитозин, за которым следует короткий (4–13 нуклеотидов) тракт пиримидиновых остатков. Такое строение в сочетании со специфическими последовательностями, входящими в 5'-нетранслируемую область мРНК, позволяет координированно экспрессировать гены семейства в ответ на внешние сигналы (например, ростовые факторы). В результате происходит увеличение интенсивности трансляции [28].

Интересно, что обнаружено несколько относящихся к 5'-ТОР-семейству генов-хозяев мякРНК (UNG, U17HG, U19H, gas5, U50HG и U87HG), экзоны которых содержат на всем протяжении множественные стоп-кодоны и поэтому не способны кодировать белок [27, 29–33]. В ходе транскрипции из интронов образуются мякРНК, а некодирующий транскрипт гена-хозяина полиаденилируется и либо остается в ядре (U19H РНК), либо выходит в цитоплазму. Для трех транскриптов (UNG, gas5 и U87HG) показана ассоциация с рибосомами, причем “мРНК” UNG и gas5 подвергаются быстрой деградации, зависимой от активности трансляционного аппарата. РНК U87HG, хотя также ассоциирована с рибосомами, отличается гораздо большей стабильностью [34]. Следует отметить, что транскрипты некодирующих генов-хозяев отличаются очень низкой консервативностью, сопоставимой с уровнем сходства нейтральных последовательностей [31]. Это породило предположение о том, что единственная функция таких генов-хозяев заключается в продуцировании мякРНК, а экзоны не несут самостоятельной функциональной нагрузки [27, 29, 30, 32]. Единственным исключением является транскрипт гена U87HG: уровень его консервативности выше, чем у дру-

гих некодирующих генов-хозяев, и соответствует уровню сходства нетранслируемых областей мРНК (во всех случаях речь идет о сходстве в пределах класса млекопитающих). Это позволяет предполагать наличие у транскрипта гена U87HG самостоятельной функции в клетке [34], тем более что ряд некодирующих РНК с установленными функциями имеет аналогичный уровень консервативности [35, 36]. По поводу остальных некодирующих генов-хозяев нужно отметить следующее. В последнее время все чаще оказывается, что и при отсутствии консервативности некодирующие РНК могут быть функциональными [35]. Такие РНК стали открывать в последнее время, во многом потому, что раньше никто не пытался искать гены среди неконсервативных последовательностей. Можно ожидать, что в ближайшее время список таких РНК будет стремительно расти. Поэтому утвердившееся представление о том, что экзоны некодирующих генов-хозяев не несут самостоятельной нагрузки, следует признать по меньшей мере не окончательным.

Известно еще несколько случаев, когда гены мякРНК локализованы в интронах некодирующих генов, однако ни один из этих генов-хозяев не относится к 5'-TOP-семейству. Так, мякРНК RVII-36 крысы, экспрессирующаяся только в мозге и лишенная участков, комплементарных рРНК или малым ядерным РНК, кодируется интроном гена некодирующей РНК *Bsr*, экспрессирующейся только в нервной системе [37]. Гены еще трех обнаруженных недавно C/D мякРНК локализованы в интронах подверженного импринтингу некодирующего гена *Ihm*. Интересно, что у человека гомологи этих мякРНК не найдены [38]. Другим примером служит подверженный импринтингу ген *SNURF-SNURPN*, экспрессирующийся с отцовского аллеля во всех тканях [39]. Его длина составляет около 450 тпн, а количество экзонов превышает 150. В дицистронной мРНК последовательно закодированы белок *SNURF* (англ. *SNRPN upstream reading frame*; экзоны 1–3) и сплайсосомный белок *SmN* (экзоны 4–10). Все оставшиеся экзоны некодирующие, а в их интронах расположены гены мякРНК, по одному гену на интрон: интроны 21–52 содержат 27 копий мякРНК *НВП-85*, а интроны 63–142 содержат 47 копий мякРНК *НВП-52*. Еще пять мякРНК, представленные меньшим числом копий, присутствуют в других интронах. Все эти мякРНК лишены участков, комплементарных рРНК или малым ядерным РНК, и экспрессируются преимущественно или исключительно в головном мозге. 3'-концевая некодирующая часть гигантского транскрипта служит антисенс-РНК для транскрибирующегося в противоположном направлении белкового гена *UBE3A* [40]. Делеция гена *SNURF-SNURPN* вместе со всеми некодирующими экзонами и мякРНК служит причиной большинства случаев

так называемого синдрома Прадера–Вилли (*Prader-Willi syndrome, PWS*) [41] – тяжелого наследственного заболевания, характеризующегося умственной отсталостью, ожирением, гипотонией мышц и рядом других признаков. В последнее время получен ряд данных о том, что делеция мякРНК, в частности *НВП-85*, играет существенную, если не ведущую, роль в патогенезе этого заболевания (см., например, [42, 43]).

Еще одна кодируемая этим геном мякРНК, *НВП-52/МВП-52* (даны названия гомологов у человека/мышь), содержит протяженный участок (18 нуклеотидов), полностью комплементарный фрагменту экспрессирующейся только в головном мозге мРНК гена серотонинового рецептора 5-НТ2С [44]. Интересно, что гены и *МВП-52*, и 5-НТ2С экспрессируются в мозге, но совместная их экспрессия наблюдается только в некоторых отделах. Потенциальной мишенью 2'-О-метилирования, которое могла бы направлять мякРНК *НВП-52/МВП-52*, служит адениловый остаток, который обычно подвергается редактированию и превращается в инозин (I). Совсем недавно [45] для мыши получены убедительные свидетельства того, что *МВП-52* действительно осуществляет такое метилирование и что при этом существенно снижается эффективность дезаминирования А-I. Иными словами, 2'-О-метилирование снижает эффективность редактирования. При трансляции инозин прочитывается как гуанозин. Вероятно, таким образом осуществляется тонкая регуляция взаимодействия серотонинового рецептора с G-белком. Интересно, что мыши, нокаутные по гену 5-НТ2С, имеют дефекты, похожие на наблюдаемые при *PWS*.

В последнее время обнаружено следующее. Известно, что в результате альтернативного сплайсинга образуются две изоформы мРНК 5-НТ2С. Они различаются размером пятого экзона. Только длинная форма мРНК способна продуцировать функциональный рецептор. Во второй части пятого экзона (5b, рис. 3) расположена последовательность (сайленсер), которая препятствует включению 5b в состав мРНК, и, соответственно, происходит образование короткой формы мРНК 5-НТ2С. В составе 5b расположены 5 нуклеотидов, которые подвергаются редактированию. Редактирование обеспечивает включение 5b в состав мРНК, и в результате образуется полноразмерная форма. Однако редактирование снижает чувствительность рецептора 5-НТ2С к серотонину в 10–100 раз. Обнаружено, что мякРНК *НВП-52/МВП-52*, антисенс-элемент которой комплементарен фрагменту 5b, обеспечивает включение 5b и соответственно образование длинной формы мРНК *НВП-52/МВП-52* независимо от редактирования [46]. Продукт такой нередактированной мРНК обладает высокой чувствительностью к серотонину. В норме клеточная популяция мРНК



**Рис. 3.** мякРНК HVI-52 и мРНК серотонинового рецептора 5-HT2C способны к комплементарным взаимодействиям. *a* – фрагмент мРНК рецептора 5-HT2C человека. Экзоны обозначены прямоугольниками и пронумерованы. Звездочкой и стрелкой отмечены соответственно проксимальный и дистальный сайты сплайсинга; *б* – комплементарные взаимодействия между мякРНК HVI-52 и мРНК 5-HT2CR. Пять нуклеотидов, подверженных редактированию (происходит дезаминирование A с образованием I), выделены рамками. Проксимальный сайт сплайсинга обозначен звездочкой и подчеркнут. Бокс D выделен прямоугольником (по [46], с изменениями).

5-HT2C содержит как редактированные, так и неотредактированные формы. Очевидно, это необходимо для регуляции чувствительности нервных клеток к серотонину. В мозге людей, больных PWS, не только отсутствует мякРНК HVI-52, но и значительно снижено количество мРНК 5-HT2C с неотредактированной последовательностью 5b [46]. Последнее обстоятельство, очевидно, вносит вклад в патогенез PWS. Благодаря проведенным исследованиям открывается путь для поиска стратегии лечения этой тяжелой болезни.

### 3. ТРАНСКРИПЦИЯ ГЕНОВ мякРНК

Транскрипцию самостоятельных генов осуществляют РНК-полимераза II (например, U3, U8 и U13РНК) и РНК-полимераза III (РНК РНКазы MRP, U3 РНК растений). 5'-концы транскриптов имеют 2,2,7-триметилгуанозиновый кэп или  $\gamma$ -монотимилфосфат соответственно. мякРНК, кодирующиеся в интронах или экспрессирующиеся в составе полицистронного транскрипта, имеют немодифицированные 5'-концы. Такие мякРНК вырезаются из предшественника, после чего их концы формируются с участием экзонуклеаз.

3'-конец всех известных мякРНК не модифицирован [10].

### 4. СОЗРЕВАНИЕ И ТРАНСПОРТ мякРНК

Существует несколько путей созревания мякРНК. Это связано прежде всего с различной организацией их генов, транскрибирующихся независимо в виде моно- и полицистронных транскриптов или локализованных в интронах. В любом случае в предшественнике мякРНК после действия специфических эндонуклеаз или после разрезания последовательности интрона в точке

ветвления образуются свободные 5'- и 3'-концы. Они подвергаются атаке 5'-3' и 3'-5' экзонуклеаз, в результате активности которых формируются концы зрелых мякРНК. Полной экзонуклеолитической деградации мякРНК препятствует связывание с ними коровых белков.

Начальные стадии процессинга независимо транскрибирующихся генов осуществляют эндонуклеазы. У дрожжей большинство генов мякРНК именно такие и входят в состав моно- и полицистронных транскриптов. Большинство этих транскриптов содержит сайты узнавания эндорибонуклеазы Rnt1p [47] (см. рис. 1). Rnt1p обладает эндонуклеазной активностью по отношению к двуцепочечной РНК и по этому признаку относится к семейству, объединяющему белки с общим названием РНКазы III [48]. РНКазы этого семейства присутствуют у всех живых организмов и участвуют в процессинге рРНК, тРНК, мРНК, микроРНК, а также могут инициировать деградацию мРНК [49]. Помимо мякРНК Rnt1p осуществляет процессинг рРНК и некоторых малых ядерных РНК (U5 и U2) [50, 51]. В ходе процессинга мякРНК Rnt1p нарезает как полицистронные транскрипты [51], так и моноцистронные транскрипты [52] (см. рис. 1). После ее действия остаются свободные концы незрелой мякРНК, которые укорачиваются экзонуклеазами: 5'-коцевая область подвергается атаке 5'-3' экзонуклеаз Rat1p и Xrn1p [53], а 3'-концевая часть процессируется экзосомой, в частности ее компонентом Rgrbr [54]. Экзосомы представляют собой комплексы большого числа (не менее 10) клеточных 3'-5' экзонуклеаз; они присутствуют как в ядре, так и в цитоплазме [55] и осуществляют формирование 3'-концов многих клеточных РНК, деградацию мРНК в цитоплазме и деграда-

цию межцистронных последовательностей пре-рРНК [56].

Некоторые мякРНК дрожжей не содержат последовательностей, узнаваемых Rnt1p. Формирование свободных 5'- и 3'-концов у предшественников таких РНК осуществляют белки, которые вносят разрыв в РНК при разрезании/полиаденировании мРНК [9].

Процессинг мякРНК, кодируемых интронами, осуществляется двумя способами. Для созревания большей части мякРНК необходим сплайсинг и последующий гидролиз 2'-5' фосфодиэфирной связи в точке ветвления. Образовавшийся продукт (мякРНК с прилегающими к ней последовательностями) подвергается атаке экзонуклеаз, и в результате формируется зрелая мякРНК [57] (см. рис. 1). Процессинг небольшого числа интронных мякРНК, в том числе у позвоночных [58], не зависит от сплайсинга: сайты для экзонуклеаз образуются в результате действия на несплайсированную пре-мРНК эндонуклеаз (рис. 1). Например, у дрожжей в этом процессе участвует Rnt1p, которая вносит разрывы в короткий двуцепочечный стебель, формируемый интронными последовательностями, фланкирующими ген мякРНК, а экзонуклеолитическая деградация осуществляется, как и в случае независимых генов, Rat1, Xrn1 [53] и экзосомой [54]. В этом случае процессы образования мякРНК и мРНК являются альтернативными путями, и возникает вопрос, каким образом осуществляется выбор между ними. Для дрожжей, где почти все немногочисленные интронные мякРНК процессируются таким образом, показано, что 5 из 6 генов этих мякРНК фланкированы последовательностями, которые не способны образовывать двуспиральный участок, являющийся субстратом Rnt1p. Тем не менее эти последовательности способны к комплементарному взаимодействию, однако в результате образуется неканонический субстрат для Rnt1p, так что узнавать его она может только в присутствии корового белка мякРНК Nop1p. Таким образом, процессинг мякРНК происходит только при наличии белков мякРНК, в противном случае происходит сплайсинг и образуется мРНК гена-хозяина [59]. Другие эксперименты также свидетельствуют о том, что коровые белки мякРНК могут принимать активное участие в процессинге [60].

Созревание интронных мякРНК происходит рядом с их генами [61], тогда как процессинг полицистронных транскриптов может происходить в так называемых тельцах Кахала и в ядрышке [62]. Тельца Кахала (ТК, англ. – Cajal bodies) часто ассоциированы с ядрышком, имеют диаметр 0.3–0.5 мкм и состоят из спиральных фибриллярных цепей, образованных главным образом белком коилином (p80 coilin) [63]. Имеющиеся данные позволяют предполагать, что ТК служат ме-

стом созревания и сборки малых ядерных РНК и транскрипционных комплексов, которые функционируют затем в других отделах ядра [63]. В тельцах Кахала продолжается сборка и созревание некоторых мякРНК, например U3 [64], и некоторых мякРНК растений [60]. Кроме того, в ТК удерживаются дефектные мякРНК [65]. Наконец, собранные мякРНК транспортируются в ядрышко, хотя некоторые этапы созревания могут проходить и в нем [60].

В заключение можно отметить, что С/D и Н/АСА мякРНК имеются у всех эукариот, связаны с одинаковым набором гомологичных белков и могут быть уложены в общие вторичные структуры. Таким образом, строение мякРНК и осуществляемые ими процессы модификации клеточных РНК весьма консервативны. С такой консервативностью удивительно контрастирует разнообразие организации генов мякРНК. Можно надеяться, что дальнейшие исследования позволят объяснить это интересное явление.

Работа выполнена при финансовой поддержке РФФИ (проект № 05-04-49553).

#### СПИСОК ЛИТЕРАТУРЫ

1. Claverie J.M. Fewer genes, more noncoding RNA // Science. 2005. V. 309. P. 1529–1530.
2. Vaughn M.W., Martienssen R. It's a small RNA world, after all // Science. 2005. V. 309. P. 1525–1526.
3. Kiss T. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs // EMBO J. 2001. V. 20. P. 3617–3622.
4. Tollervey D. Small nucleolar RNAs guide ribosomal RNA methylation // Science. 1996. V. 273. P. 1056–1057.
5. Ganot P., Bortolin M.L., Kiss T. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs // Cell. 1997. V. 89. P. 799–809.
6. Xiao S., Scott F., Fierke C.A. et al. Eukaryotic ribonuclease P: a plurality of ribonucleoprotein enzymes // Annu. Rev. Biochem. 2002. V. 71. P. 165–189.
7. Макарова Ю.А., Крамеров Д.А. Малые ядрышковые РНК // Молекулярная биология. 2007. Т. 41 (в печати).
8. Brown J.W., Echeverria M., Qu L.H. Plant snoRNAs: functional evolution and new modes of gene expression // Trends Plant. Sci. 2003. V. 8. P. 42–49.
9. Fatica A., Morlando M., Bozzoni I. Yeast snoRNA accumulation relies on a cleavage-dependent/polyadenylation-independent 3'-processing apparatus // EMBO J. 2000. V. 19. P. 6218–6229.
10. Maxwell E.S., Fournier M.J. The small nucleolar RNAs // Annu. Rev. Biochem. 1995. V. 64. P. 897–934.
11. Tycowski K.T., Aab A., Steitz J.A. Guide RNAs with 5' caps and novel box C/D snoRNA-like domains for modification of snRNAs in metazoa // Curr. Biol. 2004. V. 14. P. 1985–1995.



12. Lowe T.M., Eddy S.R. A computational screen for methylation guide snoRNAs in yeast // *Science*. 1999. V. 283. P. 1168–1171.
13. Filippini D., Renzi F., Bozzoni I. et al. U86, a novel snoRNA with an unprecedented gene organization in yeast // *Biochem. Biophys. Res. Commun.* 2001. V. 288. P. 16–21.
14. Brown J.W., Clark G.P., Leader D.J. et al. Multiple snoRNA gene clusters from Arabidopsis // *RNA*. 2001. V. 7. P. 1817–1832.
15. Chen C.L., Liang D., Zhou H. et al. The high diversity of snoRNAs in plants: identification and comparative study of 120 snoRNA genes from *Oryza sativa* // *Nucl. Acids Res.* 2003. V. 31. P. 2601–2613.
16. Kiss T., Marshallsay C., Filipowicz W. Alteration of the RNA polymerase specificity of U3 snRNA genes during evolution and *in vitro* // *Cell*. 1991. V. 65. P. 517–526.
17. Kruszka K., Barneche F., Guyot R. et al. Plant dicistronic tRNA-snoRNA genes: a new mode of expression of the small nucleolar RNAs processed by RNase Z // *EMBO J.* 2003. V. 22. P. 621–632.
18. Vision T.J., Brown D.G., Tanksley S.D. The origins of genomic duplications in Arabidopsis // *Science*. 2000. V. 290. P. 2114–2117.
19. Uliel S., Liang X.H., Unger R. et al. Small nucleolar RNAs that guide modification in trypanosomatids: repertoire, targets, genome organization, and unique functions // *Int. J. Parasitol.* 2004. V. 34. P. 445–454.
20. Russell A.G., Schnare M.N., Gray M.W. Pseudouridine-guide RNAs and other Cbf5p-associated RNAs in *Euglena gracilis* // *RNA*. 2004. V. 10. P. 1034–1046.
21. Yang C.Y., Zhou H., Luo J. et al. Identification of 20 snoRNA-like RNAs from the primitive eukaryote, *Giardia lamblia* // *Biochem. Biophys. Res. Commun.* 2005. V. 328. P. 1224–1231.
22. Tycowski K.T., Steitz J.A. Non-coding snoRNA host genes in *Drosophila*: expression strategies for modification guide snoRNAs // *Eur. J. Cell. Biol.* 2001. V. 80. P. 119–125.
23. Huang Z.P., Zhou H., Liang D. et al. Different expression strategy: multiple intronic gene clusters of box H/ACA snoRNA in *Drosophila melanogaster* // *J. Mol. Biol.* 2004. V. 341. P. 669–683.
24. Accardo M.C., Giordano E., Riccardo S. et al. A computational search for box C/D snoRNA genes in the *Drosophila melanogaster* genome // *Bioinformatics*. 2004. V. 20. P. 3293–3301.
25. Dennis P.P., Omer A., Lowe T. A guided tour: small RNA function in Archaea // *Mol. Microbiol.* 2001. V. 40. P. 509–519.
26. Bond U. Heat shock but not other stress inducers leads to the disruption of a sub-set of snRNPs and inhibition of *in vitro* splicing in HeLa cells // *EMBO J.* 1988. V. 7. P. 3509–3518.
27. Pelczar P., Filipowicz W. The host gene for intronic U17 small nucleolar RNAs in mammals has no protein-coding potential and is a member of the 5'-terminal oligopyrimidine gene family // *Mol. Cell. Biol.* 1998. V. 18. P. 4509–4518.
28. Amaldi F., Pierandrei-Amaldi P. TOP genes: a translationally controlled class of genes including those coding for ribosomal proteins // *Prog. Mol. Subcell. Biol.* 1997. V. 18. P. 1–17.
29. Tycowski K.T., Shu M.D., Steitz J.A. A mammalian gene with introns instead of exons generating stable RNA products // *Nature*. 1996. V. 379. P. 464–466.
30. Bortolin M.L., Kiss T. Human U19 intron-encoded snoRNA is processed from a long primary transcript that possesses little potential for protein coding // *RNA*. 1998. V. 4. P. 445–454.
31. Smith C.M., Steitz J.A. Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes // *Mol. Cell. Biol.* 1998. V. 18. P. 6897–6909.
32. Tanaka R., Satoh H., Moriyama M. et al. Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma // *Genes Cells*. 2000. V. 5. P. 277–287.
33. Makarova Yu.A., Kramerov D.A. Mammalian 87 small nucleolar RNA and its host gene // *Mol. Biol. (Rus.)*. 2005. V. 39. P. 655–663.
34. Makarova J.A., Kramerov D.A. Noncoding RNA of U87 host gene is associated with ribosomes and is relatively resistant to nonsense-mediated decay // *Gene*. 2005. V. 363. P. 51–60.
35. Mattick J.S. The functional genomics of noncoding RNA // *Science*. 2005. V. 309. P. 1527–1528.
36. Tam W. Identification and characterization of human BIC, a gene on chromosome 21 that encodes a noncoding RNA // *Gene*. 2001. V. 274. P. 157–167.
37. Cavaille J., Vitali P., Basyuk E. et al. A novel brain-specific box C/D small nucleolar RNA processed from tandemly repeated introns of a noncoding RNA gene in rats // *J. Biol. Chem.* 2001. V. 276. P. 26374–26383.
38. Xiao Y., Zhou H., Qu L.H. Characterization of three novel imprinted snoRNAs from mouse Irm gene // *Biochem. Biophys. Res. Commun.* 2006. V. 340. P. 1217–1223.
39. Gray T.A., Saitoh S., Nicholls R.D. An imprinted, mammalian bicistronic transcript encodes two independent proteins // *Proc. Natl Acad. Sci. USA*. 1999. V. 96. P. 5616–5621.
40. Runte M., Huttenhofer A., Gross S. et al. The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A // *Hum. Mol. Genet.* 2001. V. 10. P. 2687–2700.
41. Ohta T., Gray T.A., Rogan P.K. et al. Imprinting-mutation mechanisms in Prader-Willi syndrome // *Am. J. Hum. Genet.* 1999. V. 64. P. 397–413.
42. Ding F., Prints Y., Dhar M.S. et al. Lack of Pwcr1/MBII-85 snoRNA is critical for neonatal lethality in Prader-Willi syndrome mouse models // *Mamm. Genome*. 2005. V. 16. P. 424–431.
43. Schule B., Albalwi M., Northrop E. et al. Molecular breakpoint cloning and gene expression studies of a novel translocation t(4;15)(q27;q11.2) associated with Prader-Willi syndrome // *BMC Med. Genet.* 2005. V. 6. P. 18.
44. Cavaille J., Buiting K., Kieffmann M. et al. Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization //

- Proc. Natl Acad. Sci. USA. 2000. V. 97. P. 14311–14316.
45. Vitali P., Basyuk E., Le Meur E. et al. ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs // *J. Cell. Biol.* 2005. V. 169. P. 745–753.
  46. Kishore S., Stamm S. The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2 C // *Science*. 2006. V. 311. P. 230–232.
  47. Chanfreau G., Legrain P., Jacquier A. Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism // *J. Mol. Biol.* 1998. V. 284. P. 975–988.
  48. Elela S.A., Igel H., Ares M., Jr. RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site // *Cell*. 1996. V. 85. P. 115–124.
  49. Carmell M.A., Hannon G.J. RNase III enzymes and the initiation of gene silencing // *Nat. Struct. Mol. Biol.* 2004. V. 11. P. 214–218.
  50. Abou Elela S., Ares M., Jr. Depletion of yeast RNase III blocks correct U2 3' end formation and results in polyadenylated but functional U2 snRNA // *EMBO J.* 1998. V. 17. P. 3738–3746.
  51. Chanfreau G., Rotondo G., Legrain P. et al. Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1 // *EMBO J.* 1998. V. 17. P. 3726–3737.
  52. Kufel J., Allmang C., Chanfreau G. et al. Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding // *Mol. Cell. Biol.* 2000. V. 20. P. 5415–5424.
  53. Petfalski E., Dandekar T., Henry Y. et al. Processing of the precursors to small nucleolar RNAs and rRNAs requires common components // *Mol. Cell. Biol.* 1998. V. 18. P. 1181–1189.
  54. Allmang C., Kufel J., Chanfreau G. et al. Functions of the exosome in rRNA, snoRNA and snRNA synthesis // *EMBO J.* 1999. V. 18. P. 5399–5410.
  55. van Hoof A., Parker R. The exosome: a proteasome for RNA? // *Cell*. 1999. V. 99. P. 347–350.
  56. Butler J.S. The yin and yang of the exosome // *Trends Cell. Biol.* 2002. V. 12. P. 90–96.
  57. Kiss T., Filipowicz W. Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns // *Genes Dev.* 1995. V. 9. P. 1411–1424.
  58. Caffarelli E., Ares M., Santoro B. et al. *In vitro* study of processing of the intron-encoded U16 small nucleolar RNA in *Xenopus laevis* // *Mol. Cell. Biol.* 1994. V. 14. P. 2966–2974.
  59. Giorgi C., Fatica A., Nagel R. et al. Release of U18 snoRNA from its host intron requires interaction of Nop1p with the Rnt1p endonuclease // *EMBO J.* 2001. V. 20. P. 6856–6865.
  60. Filipowicz W., Pogacic V. Biogenesis of small nucleolar ribonucleoproteins // *Curr. Opin. Cell. Biol.* 2002. V. 14. P. 319–327.
  61. Samarsky D.A., Fournier M.J., Singer R.H. et al. The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization // *EMBO J.* 1998. V. 17. P. 3747–3757.
  62. Shaw P.J., Beven A.F., Leader D.J. et al. Localization and processing from a polycistronic precursor of novel snoRNAs in maize // *J. Cell. Sci.* 1998. V. 111. Pt 15. P. 2121–2128.
  63. Gall J.G. Cajal bodies: the first 100 years // *Annu. Rev. Cell. Dev. Biol.* 2000. V. 16. P. 273–300.
  64. Gerbi S.A., Borovjagin A.V., Lange T.S. The nucleolus: a site of ribonucleoprotein maturation // *Curr. Opin. Cell. Biol.* 2003. V. 15. P. 318–325.
  65. Narayanan A., Speckmann W., Terns R. et al. Role of the box C/D motif in localization of small nucleolar RNAs to coiled bodies and nucleoli // *Mol. Biol. Cell.* 1999. V. 10. P. 2131–2147.

## Small Nucleolar RNA Genes

Yu. A. Makarova and D. A. Kramerov

*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia;*  
*fax: (495)135-14-05; e-mail: kramerov@eimb.ru*

Small nucleolar RNAs (snoRNAs) are one of the most numerous and well-studied groups of non-protein-coding RNAs. In complex with proteins, snoRNAs perform the two most common nucleotide modifications in rRNA: 2'-O-methylation of ribose and pseudouridylation. Although the modification mechanisms and snoRNA structures are highly conserved, the snoRNA genes are surprisingly diverse in organization. In addition to genes transcribed independently, there are genes that are in introns of other genes, form clusters transcribed from a common promoter, or cluster in introns. Interestingly, one type of gene organization usually prevails in different taxa. Vertebrate snoRNAs mostly originate from introns of protein-coding genes; a small group of snoRNAs are encoded by introns of genes for noncoding RNAs.