



УДК 577.214

## НЕОДНОЗНАЧНОСТЬ ГРАНИЦ ТРАНСКРИПЦИИ ЭУКАРИОТИЧЕСКИХ ГЕНОВ

В. М. Кавсан

На основании данных анализа структуры ряда вирусных и эукариотических генов были точно определены места, соответствующие сайтам кэпирования и полиаденилирования зрелых мРНК. Эти исследования привели к широко принятому сейчас заключению о том, что инициация транскрипции начинается недалеко от последовательности гена, соответствующей сайту кэпирования матричной РНК. Сайт, терминирующий транскрипцию, менее определен, однако долгое время считалось, что он должен соответствовать нуклеотидной последовательности, кодирующей сайт полиаденилирования зрелой мРНК.

Действительно, в ряде работ описано определение предшественников мРНК с размерами, не превышающими длину соответствующих генов. Например, путем гибридизации с кДНК  $\beta$ -глобина такие молекулы с коэффициентом седиментации 15—17S были обнаружены в ядерной РНК эритроидных клеток различных животных [1]. 15S-предшественник и зрелая мРНК  $\beta$ -глобина имеют одну и ту же 5'-концевую последовательность, одну и ту же структуру кэпа. 5'- и 3'-концы 15S пре-мРНК кодируются в тех же местах хромосомной ДНК, что и соответствующие концы зрелой мРНК. Аналогичные результаты были получены и для некоторых других пре-мРНК  $\beta$ -глобинового семейства. В подобных экспериментах кДНК  $\alpha$ -глобина гибридизовалась только с фракцией 11S РНК, что коррелирует с размерами генов для  $\alpha$ -глобина, и не гибридизовалась с фракцией 15S ядерной РНК [1, 2].

**Обнаружение гигантских ядерных РНК, намного превышающих по размерам соответствующие гены.** Наряду с 15—17S и 11S-предшественниками информационных РНК  $\beta$ - и  $\alpha$ -глобинов соответственно многие исследователи обнаружили глобинспецифические последовательности в молекулах ядерных РНК с размерами, значительно превышающими эти значения. Так, при гибридизации с радиоактивно меченой кДНК глобина наличие комплементарных последовательностей было показано в ядерной РНК с молекулярной массой более  $1,7 \cdot 10^6$  [3—10].

Однако, несмотря на специально поставленные для достижения той же цели эксперименты, другим авторам, как отмечалось выше, не удавалось определить глобинспецифические последовательности во фракциях ядерных РНК с коэффициентом седиментации более 17S. В то же время Стрэр и др. [11] нашли, что меченная тритием по уридиону глобинспецифическая ядерная РНК незрелых клеток крови уток седиментирует в денатурирующих условиях со скоростью, равной или несколько большей, чем 28S-рибосомная РНК. Через 15 или 60 мин после введения метки молекулы, содержащие последовательности мРНК глобина, седиментировали со значительно меньшей скоростью (16,5S). При исследовании немеченых ядерных РНК около 2% общей глобиновой РНК ядра седиментировало примерно с той же скоростью, что и 28S-рРНК; 16,5S- и 28S-глобиновые РНК при повторном центрифугировании показывали те же значения коэффициентов седиментации. Авторы предположили, что отрицательные результаты при определении

радиоактивно меченных 28S-глобиновых РНК в цитированных выше работах могут объясняться тем, что эти молекулы не накапливаются в достаточной мере по отношению к общему пулу меченой глобиновой РНК за 5 мин после введения метки (их определение затруднено ввиду 0,01 %-ного фона гибридизации). 28S-глобиновые РНК, возможно, имеют очень короткое время жизни и потому не могут быть определены через 15 мин после введения метки.

В экспериментах, показавших существование гигантских пре-мРНК, чаще всего использовали метод гибридизации немеченой клеточной РНК с радиоактивно меченым кДНК-зондом [3—6]. Этот метод значительно более чувствителен, чем гибридизация импульсно меченой клеточной РНК с немеченой кДНК, применяя которую авторы не сумели показать наличия гигантских глобинспецифических ядерных РНК [1, 12]. Тем не менее Шаул и др. [13], используя технику импульсной метки и последующего денатурирующего электрофореза, подтвердили известные ранее данные о существовании гигантских молекул с последовательностями мРНК глобина в ядрах клеток селезенки и культуре клеток крови мышей с эритролейкемией. Бастос и Авив [14] через 5 мин после введения импульсной метки в клетки Фрейнда и клетки селезенки мыши обнаружили до 2/3 всей радиоактивности во фракции ядерной РНК с коэффициентом седиментации 27S, которая аффинно связывалась на колонке с кДНК глобина, присоединенной к целлюлозе. Однако Хейнес и др. [15] не смогли обнаружить гигантских молекул в таких же клетках через 10 мин после введения метки при использовании такой же аффинной колонки.

При гибридизации пре-мРНК с кДНК, иммобилизованной на нерастворимом носителе, имеет место двухфазное распределение молекул. Поскольку иммобилизованную ДНК нельзя считать полностью находящейся в растворе, очень сложные последовательности при малой концентрации (такие, как пре-мРНК) могут быть недоступными для гибридизации и, следовательно, не быть обнаруженными. Этим можно объяснить то, что глобиновые последовательности величиной более 5000 нуклеотидов определены только по гибридизационной кинетике с кДНК, но не выделением молекул. В лаборатории Шеррера [16] был предложен метод получения подобных молекул гибридизацией в водной фазе с последующим выделением гибридов аффинной хроматографией. Благодаря такому подходу были получены молекулы РНК, содержащие глобиновые последовательности величиной 1500, 4000 и до 10 000 нуклеотидов.

Длительную дискуссию о ядерных глобинспецифических РНК, значительно более длинных, чем соответствующий ген между сайтами экзонирования и полиаденилирования зрелых мРНК, можно было считать оконченной признанием их существования. Однако в 1984 году при помощи денатурирующего электрофореза с последующим перенесением разделенной РНК на бумагу и гибридизацией с кДНК было показано, что определение гигантских глобинспецифических молекул РНК может быть артефактом в результате поперечной сшивки рРНК и мРНК, когда в качестве денатурирующего агента применяются альдегиды [17]. Использование двух независимых методов электрофореза с денатурирующими агентами, различающимися по своим механизмам действия на РНК, перенесение разделенных РНК на различные твердые подложки, получение высокоочищенной поли(А)-содержащей ядерной РНК, а также применение различных ДНК-зондов, в том числе специфичных только для интронов пре-мРНК, позволили доказать, что  $\alpha$ - и  $\beta$ -глобинспецифические транскрипты дискретной длины, намного превышающей размеры соответствующих генов, действительно существуют в ядерной РНК эритроидных клеток [18—21].

Обнаружение ядерных РНК гигантских размеров свидетельствовало против предполагавшейся ранее многими исследователями модели транскриптона, в котором транскрипция начинается и кончается в сайтах, соответствующих зрелой мРНК. Эти данные указывали на то, что

у эукариот имеет место более сложный механизм транскрипции и процессинга предшественников матричных РНК.

**Образование гигантских ядерных РНК в результате транскрипции последовательности ДНК за сайтом полиаденилирования.** Возможность транскрипции через сайт полиаденилирования впервые была показана у ДНК-содержащих вирусов, способных интегрировать с клеточной ДНК. Так, в случае мажорного позднего транскрипта [22] и двух ранних транскриптов районов 2 и 4 аденовируса 2 [23], как и в случае некоторых транскриптов SV40 [24], РНК-полимераза проходит несколько тысяч нуклеотидных пар за нуклеотид, соответствующий месту присоединения поли(А) в зрелой мРНК. Первичные неполиаденилированные РНК-транскрипты, гетерогенные по 3'-концу, затем подравниваются специальным механизмом для наращивания поли(А). Таким образом, здесь не обнаруживается специфической транскрипционной единицы, в противоположность тому, что было найдено для неинтегрирующих с ядерной ДНК вирусных геномов.

Прочитывание через сайт полиаденилирования описано сейчас и для ряда клеточных генов. Так, гибридизацией с фрагментами клонированного гена  $\beta^{\text{maj}}$ -глобина мыши показано, что до 80 % молекул ядерных глобинспецифических РНК содержат с 3'-конца по меньшей мере 1000 нуклеотидов, переписанных с участка, лежащего за сайтом полиаденилирования [25]. Транскрипция заканчивается в районе между 700 и 2000 нуклеотидов за поли(А)-сайтом, но не в одном месте [26]. Для  $\beta_1$ -глобинового гена кролика терминация одного из таких транскриптов была картирована в области около 1300 нуклеотидов за каноническим поли(А)-сайтом [27]. Терминация транскрипции  $\alpha$ -амилазного гена мыши на 90 % проходит в районе между 2500 и 4000 нуклеотидных пар за поли(А)-сайтом [28]. Транскрипция через сайт полиаденилирования имеет место для генов X и Y курицы [29], для глобинового гена в SV40-глобиновом рекомбинанте [30].

Чрезвычайно интересный пример транскрипции через сайт полиаденилирования представляет синтез мРНК для тяжелой цепи ( $\mu$ ) связывающейся с мембраной формы IgM [31]. Транскрипция через сайт полиаденилирования, обычно использующийся при синтезе мРНК для секреторируемой формы белка ( $\mu_s$ ), и полиаденилирование у сайта, находящегося дальше в 3'-направлении, продуцируют пре-мРНК, которая включает последовательность еще одного сегмента, кодирующего гидрофобный С-конец связывающейся с мембраной формы ( $\mu_m$ ). Матричная РНК для этой формы белка образуется путем вырезания из пре-мРНК части последовательности, кодирующей 19 последних аминокислотных остатков  $\mu_s$  вместе с терминирующим кодоном и далее до начала, характерного для формы  $\mu_m$  экзона, с последующим сплайсингом остающихся фрагментов. Поскольку IgM<sub>s</sub> и IgM<sub>m</sub> продуцируются на различных стадиях созревания В-клеток, было предположено, что на каждой стадии происходит специфическое узнавание различных сайтов полиаденилирования [31]. Возможно, нечто похожее имеет место и при образовании множественных мРНК для  $\delta$ -цепей IgD [32], а также при образовании мРНК для  $k$ -цепей [33].

Более одного сайта полиаденилирования показано и для других клеточных генов, например для мРНК  $\alpha$ -амилазы мышей, у которой более чем на 200 нуклеотидов отличается длина 3'-нетранслируемых районов в зависимости от того, где она транскрибируется — в клетках печени или клетках слюнных желез [28, 34]. Для мРНК пролактина также показаны вариации места полиаденилирования [35]. Несколько типов мРНК дигидрофолатредуктазы мыши, отличающихся длиной 3'-нетранслируемых районов (от 80 до 930 нуклеотидов) — также результат полиаденилирования в различных сайтах транскрипта одного и того же гена [36]. Такое же явление, возможно, лежит в образовании необычно длинных мРНК, кодирующих  $\alpha$ - и  $\beta$ -интерфероны [37]. Для гена овомукоида показано, что полиаденилирование транскрибируемых с него мРНК может происходить в трех разных местах [38]. Так же

объясняется образование двух различающихся по длине мРНК фибробластного интерферона человека [39], двух отличающихся по размерам 3'-нетранслируемого района мРНК, переписанных с одного гена виментина курицы [40], трех мРНК кальмодулина угря [41], по меньшей мере двух типов мРНК с различными размерами 3'-конца для  $\beta_2$ -микроглобулина тератокарциномы [42], двух мРНК для кальцитонина и кальцитонинподобного нейропептида, переписываемых с одного гена [43], двух мРНК для про- $\alpha$ (I)-коллагена человека [44]; про- $\alpha$ (III)-коллагеновый ген также транскрибируется в виде множественных мРНК, различающихся по длине 3'-нетранслируемого района [45].

Частичное переписывание через обычно использующийся сайт полиаденилирования в поиске расположенных далее по направлению транскрипции потенциальных экзонов, которые могут быть включены в транскрипционную единицу, имеет очевидное эволюционное значение.

**Терминация транскрипции генов, кодирующих белок, у ядерных организмов.** В образовании 3'-конца зрелой мРНК у эукариот принимают участие два различных механизма: терминация транскрипции и процессинг. У дрожжей терминация происходит, очевидно, недалеко от сайта полиаденилирования. Например, 3'-конец мРНК цитохрома *c* у дрожжей образуется объединенным процессом терминации — полиаденилирования, который происходит в АТ-богатой последовательности, обладающей двойной симметрией [46].

Однако для большинства генов у высших эукариот не обнаружено четких сигналов терминации транскрипции [47]. Гибридизацией меченых ядерных транскриптов с клонированными фрагментами  $\beta_1$ -глобинового гена кролика было показано, что транскрипция доходит до 2447 нуклеотидов за канонический поли(А)-сайт. Эффективность транскрипции начинает снижаться где-то после 600 нуклеотидов за поли(А)-сайтом, постепенно затухая на протяжении следующих 568 нуклеотидов. При этом никаких дискретных сайтов для терминации транскрипции РНК-полимеразой II не было обнаружено, так же как не было обнаружено их и для  $\beta^{\text{maj}}$ -глобинового гена мыши [47]. В обоих случаях наблюдалась чрезвычайно подобная картина: уровень транскрипции резко снижался в районе за 600-м нуклеотидом после поли(А)-сайта, играющем роль аттенюаторного участка. Остановка транскрипции в этом участке может осуществляться за счет наличия в нем двух специфических последовательностей. Одна из них представляет инвертированный повтор, способный образовывать шпильчатую структуру, которая обычно встречается как терминатор транскрипции у прокариот [49]. Аналогичные структуры, как было предположено, вызывают аттенюацию транскрипции генов SV40 — стабильность этой вторичной структуры влияла на эффективность остановок в месте аттенюации.

Вторая короткая последовательность в районе остановки транскрипции  $\beta$ -глобиновых генов может транскрибироваться с противоположной цепи и таким образом интерферировать с транскрипцией  $\beta_1$ -глобинового гена, возможно, диссоциацией молекул РНК-полимеразы из транскрипционного комплекса. Далее в 3'-направлении следует последовательность, принадлежащая к семейству D-повторов и представляющая, по-видимому, самостоятельную транскрипционную единицу для РНК-полимеразы III, а еще далее — АТ-богатый участок. Хотя только высокое содержание АТ не является достаточным для окончания транскрипции, наличие таких участков вместе с другими факторами может способствовать терминации.

Таким образом, по-видимому, в терминации транскрипции эукариотических генов могут участвовать различные сигналы. Кроме перечисленных, такими сигналами могут служить последовательности средних диспергированных повторов, как это было предположено для гена дигидрофолатредуктазы мыши. Транскрипция этого гена идет через семь сайтов полиаденилирования и заканчивается около такого повтора [47]. Нельзя исключить, что в терминации транскрипции генов, кодирующих

белки у эукариот, принимают участие и трансдействующие факторы, подобные тем солерастворимым факторам, участие которых предполагается в терминации транскрипции рРНК, гистоновых мРНК, РНК SV40. В принципе короткий транскрипт, синтезирующийся в противоположном направлении от РНК-транскрипта  $\beta$ -глобинового гена, может образовывать с последним двухцепочечную структуру, составляя потенциальный сигнал для регуляторного фактора [48].

**Сигналы, определяющие 3'-конец мРНК.** У высших эукариот созревание 3'-конца мРНК включает процессинг первичного транскрипта эндонуклеотическим и/или экзонуклеотическим расщеплением. В формировании сигнала для этого этапа принимает участие каноническая последовательность AAUAAA, впервые описанная для глобиновых и иммуноглобулиновых генов [50], а затем обнаруженная у большинства генов эукариот. Расположение этой последовательности или ее вариантов стандартно: в среднем 10—20 нуклеотидов от сайта полиаденилирования [2, 51]. Консервативность этой последовательности чрезвычайно высока. Компьютерный анализ 134 мРНК позвоночных, за исключением мРНК гистонов, позволил вывести следующую общую формулу:  $A_{98}A_{91}U_{100}A_{99}A_{99}A_{98}$ . Варианты: AUUAAA (12%), AAUAAA (2%), AAUACA (2%), AAUAAU (2%), AAUAAC (1%) и CAUAAA (1%) также обнаружены в природных мРНК [52].

Было показано, что реакция расщепления первичного транскрипта предшествует полиаденилированию поздних мРНК SV40. У них обнаруживаются идентичные 3'-концы, содержащие AAUAAA на расстоянии 12 нуклеотидов от сайта полиаденилирования. Удаление этой последовательности препятствовало полиаденилированию, а мутанты с небольшими делециями между AAUAAA и нормальным сайтом полиаденилирования образовывали мРНК, которые были полиаденилированы в новых местах, расположенных дистально [53].

Единственная трансверсия  $U \rightarrow G$  в канонической последовательности AAUAAA гена IA раннего района аденовируса 2 снижала эффективность расщепления РНК-транскрипта, но не уменьшала эффективность полиаденилирования для тех молекул, у которых уже произошло отщепление «избыточного» 3'-района. По-видимому, последовательность AAUAAA необходима скорее для эффективного расщепления, чем для полиаденилирования [54].

Природная мутация AATAAA  $\rightarrow$  AACAAA также уменьшала эффективность образования нормального 3'-конца глобиновой мРНК: вместо него у пациентов с  $\beta$ -талассемией обнаруживались глобинспецифические мРНК на 900 нуклеотидов длиннее, чем в норме [55]. Аналогичный результат был получен при исследовании в клетках *HeLa* транскрипта  $\alpha$ -глобинового гена, взятого у больного с  $\alpha$ -талассемией и содержавшего мутацию AATAAA  $\rightarrow$  AATAAG [56]. При этом, возможно, используются криптические сайты полиаденилирования, как, например, гексануклеотид AATAAA, отстоящий примерно на 1300 нуклеотидов от 3'-сайта полиаденилирования зрелой мРНК  $\beta_1$ -глобина кролика и используемый для полиаденилирования глобинспецифических транскриптов длиной около 2600 нуклеотидов [20].

Ген тропомиозина I дрозофилы имеет шесть или даже семь поли(A)-сигналов, которые используются *in vivo*. Три из них — обычные AATAAA, два — AATAAA и еще два — AACAAA. В пределах размеров транскрипта AATAAA встречается трижды и трижды используется в качестве сигнала полиаденилирования, AATAAA используется только в двух случаях из пяти. Возможно, объяснение этого обстоятельства состоит в том, что обычные эффективные сигналы AATAAA расположены вблизи от неактивных AATAAA. Последовательность AACAAA также функциональна только в двух случаях из пяти, обнаруженных в 3'-концевых участках гена [57]. Необычный сигнал полиаденилирования — ATTAAA — имеется также в гене цыпленка [58], гене  $\alpha$ -амилазы мыши [34]. В последнем также обнаружена еще одна функциональная сигнальная последовательность — AATATA [28]. В  $\beta_1$ -глобиновом гене

лягушки минорный сайт полиаденилирования расположен на 46 нуклеотидов в 3'-направлении дальше от мажорного и используется только у 1% молекул мРНК. Если «мажорный» сайт делегирован, «минорный» используется более чем в 90% случаев. При изменении ААТААА на ААТАСА до 35% мРНК еще полиаденилировано в мажорном сайте, что указывает на определенную степень гибкости в обозначении сигнала полиаденилирования [59].

Хотя последовательность ААUААА необходима для расщепления предшественника мРНК, по-видимому, она не всегда достаточна для образования полного сигнала. Эта последовательность встречается часто в кодирующих областях мРНК, где она не участвует в образовании 3'-конца [53]. Некоторые другие консервативные последовательности, кроме ААUААА, имеют значение для формирования 3'-конца мРНК. Так, у большинства генов обнаружена консервативная последовательность YGTGTTYU, расположенная чаще всего на расстоянии 24—30 нуклеотидов в 3'-направлении от ААТААА. Удаление этого участка в генах вируса герпеса приводило к укорачиванию 3'-концевой области. Примерно треть исследованных генов не имеет с 3'-конца мотива последовательности YGTGTTYU. Ее отсутствие, возможно, компенсируется другими сигналами. Например, отмечено, что 3'-концевые районы содержат Т-богатые участки [60].

Пока не ясно, где образует последовательность YGTGTTYU сигнал — в ДНК или РНК, и участвует ли этот сигнал в терминации транскрипции или процессинге. Некоторые косвенные данные свидетельствуют в пользу его роли в терминации транскрипции: анализ распределения последовательностей YGTGTTYU у большинства генов вируса герпеса, где различные транскрипционные единицы четко локализованы в вирусном геноме, и в отличие от клеточных генов их экспрессия нарушается, если не будет достаточно четкой остановки транскрипции, указывает на отсутствие корреляции между наличием этого сигнала с селекцией 3'-конца при альтернативном процессинге. Его удаление не оказывает влияния на правильную инициацию транскрипции, но ингибирует образование 3'-конца. Образующиеся при этом РНК не полиаденилированы и, хотя транспортируются в цитоплазму, не транслируются там [60].

Как показано *in vitro*, при использовании лизатов клеток *HeLa* каноническая последовательность ААТААА необходима все же и для полиаденилирования. Полиаденилирующая активность узнает сигнал и в отсутствие эндонуклеазной активности, локализуя 3'-конец РНК на расстоянии даже более 400 нуклеотидов от ближайшего ААUААА и используя этот конец как праймер для поли(А)-синтеза. Интересно, что длина поли(А)-хвоста, наращенного к более длинному 3'-концевому фрагменту (расстояние от ААUААА до 3'-конца), обратно пропорциональна этому расстоянию. Создается впечатление, что поли(А)-полимераза может каким-то образом считать количество добавленных нуклеотидов и отсюда количество нуклеотидов при поиске 3'-конца. Значит, число добавленных аденозинов будет обратно пропорционально числу нуклеотидов, разделяющих сигнал ААUААА и 3'-конец РНК [61]. Из данных, полученных *in vitro*, следует, что не всегда полиаденилирование может происходить через 10—30 нуклеотидов после сигнала ААUААА. *In vivo* также отмечено, что в РНК вирусов Т-клеточной лейкемии человека I и II, а также вируса бычьего лейкоза, расстояние между ААUААА и поли(А)-сайтом составляет более 200 нуклеотидов [61].

**Формирование 3'-конца гистоновых мРНК.** До сих пор не сообщалось об идентификации предшественников гистоновых мРНК, однако из результатов экспериментов с инъекцией генов в ооциты лягушки можно заключить, что транскрипция пре-мРНК гистонов заканчивается во многих точках на протяжении примерно 100 нуклеотидов в 3'-направлении от нуклеотида, соответствующего 3'-концу мРНК [62].

Гистоновые мРНК в подавляющем большинстве случаев не поли-

аденилированы, однако, как и у поли(А)-содержащих мРНК, 3'-конец мРНК гистонов формируется процессингом более длинного предшественника. Процессинг 3'-концевого участка не зависит от 5'-концевых последовательностей, однако 5'-эпирование необходимо для достижения максимальной скорости процессинга 3'-конца. Общим свойством почти всех гистоновых мРНК является наличие палиндромной структуры у 3'-конца. Эта структура имеется у всех исследованных видов животных от морского ежа до человека; она чрезвычайно консервативна, но не узнается в качестве терминирующего сигнала РНК-полимеразой II в бесклеточной системе транскрипции, а скорее необходима для указания места созревания 3'-концевой последовательности [63].

Как показано в серии работ Бирнстила и др., краткий обзор которых дан в статьях [64, 65], процесс созревания 3'-конца гистоновых мРНК осуществляется с помощью трансдействующих факторов, представляющих собой частицы няРНК. Так, было показано, что одновременное введение вместе с геном гистона H3 морского ежа фракции, содержащей частицы 12S РНК, было необходимым для образования правильных 3'-концов в РНК, транскрибированной с этого гена в ооцитах лягушки. Фракция няРНК с коэффициентом 12S в основном содержит два типа няРНК — около 100 и 60 нуклеотидов. Очищенная из этих частиц няРНК (по классификации няРНК U7) длиной 55—68 нуклеотидов полностью обеспечивала образование правильных 3'-концов мРНК гистона H3 в описанных условиях.

Основываясь на комплементарности последовательностей, полученных с помощью секвенирования гистоновых генов и няРНК U7, была предложена модель [63], где 5'-конец няРНК U7 гибридизуется с двумя консервативными районами 3'-конца гистоновой пре-мРНК. Согласно этой модели, няРНК превращает шпилькообразную структуру 3'-конца пре-мРНК в линейную форму спариванием с 13 из 16 нуклеотидов шпильки. Одновременно с этим няРНК U7 образует шесть комплементарных нуклеотидных пар с последовательностью CAAGAAAGA, находящейся через шесть нуклеотидов после места, соответствующего 3'-концу мРНК. Выпячивание участка гистоновой пре-мРНК между этими консервативными последовательностями в виде небольшой петли, по-видимому, способствует эндонуклеотическому расщеплению, продуцирующему 3'-конец зрелой мРНК. Последовательность CAAGAAAGA законсервирована точно в таком виде только у морских ежей; у высших организмов она несколько изменена, что, возможно, объясняет необходимость в няРНК U7 при правильном процессинге мРНК гистона H3 морского ежа в ооцитах лягушки.

**Образование 5'-конца мРНК у эукариот.** Одной из характерных особенностей мРНК и их предшественников у эукариот является наличие на 5'-конце особой «кэп»-структуры,  $m^7G(5')ppp(5')N$ . Она обнаружена на 5'-конце большинства вирусных и клеточных пре-мРНК. Кэп ядерных транскриптов сохраняется во время созревания пре-мРНК в мРНК. Таким образом, кэп представляет собой уникальный маркер для 5'-концов мРНК и их предшественников [51].

Впервые обнаруженная в позднем районе вируса полиомы гетерогенность сайта эпирования [66] сейчас показана для многих транскрипционных единиц ДНК-содержащих вирусов [67]. Подобная микрогетерогенность имеет место и для клеточных мРНК: изо-1-цитохрома дрожжей, овальбумина, лизоцима,  $\beta$ -глобина курицы [68]. Кроме такой микрогетерогенности сайтов эпирования описаны также примеры гетерогенности сайтов начала транскрипции на значительно более протяженных участках гена. Так, в районе для поздних мРНК генома вируса полиомы были локализованы по меньшей мере 15 различных сайтов эпирования на протяжении 94 нуклеотидных пар [69].

Поскольку структура кэпа образуется из трифосфорилированных концов только что синтезированного транскрипта [70], множественность сайтов эпирования подразумевает варибельность сайтов инициации. В ряде случаев она заключается только в небольшой неточности выбо-

ра соответствующего иницирующего нуклеотида. Транскрипционные единицы с такой микрогетерогенностью сайтов инициации обычно, хотя и не без исключения, имеют ТАТА-бокс (так называемый бокс Хогнесса), расположенный на расстоянии примерно 30 нуклеотидных пар перед районом инициации. Как считают Бэйкер и Зифф [67], РНК-полимераза II может отмерять определенное расстояние от ТАТА-бокса и затем сканировать участок из 2—7 нуклеотидных пар для одного или более подходящих мест начала транскрипции, чаще всего выбирая для этой цели пурины. Согласно гипотезе, на микрогетерогенность может влиять структура либо длина района, включающего ТАТА-последовательность, которые могут вносить вариабельность в отмеряемое расстояние, и/или природа последовательностей около потенциальных сайтов инициации.

Отсутствие ТАТА-последовательности в районе промоторов некоторых генов аденовирусов и паповавирусов, а также появление новых стартов транскрипции после удаления ТАТА из района промотора генов SV40, гистона H2A,  $\beta$ -глобинового гена кролика свидетельствуют, по-видимому, о том, что для РНК-полимеразы II существуют и другие типы промоторов. В клетках яйцевода, например, вместе с транскриптами гена овомукоиды, в инициации синтеза которых принимал участие бокс Хогнесса, имелись также продукты, транскрипция которых иницировалась с других промоторов, расположенных на расстоянии 80 и 29 нуклеотидных пар от основного [71].

Гетерогенность мРНК плацентарного лактогена человека обусловлена множественными сайтами инициации транскрипции. В системе *in vitro* наряду с основным сайтом инициации, отстоящим на 29 нуклеотидных пар от ТАТА-бокса, 8 % транскриптов начинались с участка, отстоящего на 53 нуклеотидные пары в 5'-направлении. По меньшей мере еще три сайта, отстоящих от него на 15, 23 и 39 нуклеотидных пар в 3'-направлении, использовались *in vivo* [72]. Наряду с мажорной гомогенной по 5'-концу фракцией мРНК ренина 1 и ренина 2, которые начинаются с одного места, обнаруживалась и минорная фракция транскриптов, начинающихся в 5'-направлении от основного места инициации [73].

До последнего времени результаты большинства работ свидетельствовали о том, что инициация транскрипции  $\beta$ -глобиновых генов происходит исключительно в каноническом сайте экпирования, как указывалось в начале статьи. Некоторые авторы сохраняют это убеждение и сейчас [48]. Однако в ряде работ показано, что это положение не совсем справедливо. Так, для  $\epsilon$ -глобинового гена человека были обнаружены по меньшей мере девять сайтов инициации транскрипции, которые располагаются в четырех районах: —65 до —250, —900, —1480 и —4500 нуклеотидных пар от кодона инициации трансляции; во всех случаях транскрипты были экпированы [74].

Транскрипция части мРНК  $\gamma$ - и  $\beta$ -глобинов человека, похоже, начинается на 200 нуклеотидов раньше подавляющего большинства молекул остальной гомологичной мРНК [75]. Эти сайты инициации транскрипции, так же как и сайт, лежащий на 1400 нуклеотидных пар впереди  $\epsilon$ -глобинового гена [74], размещены около последовательности, которая может служить промотором РНК-полимеразы III. Синтез  $\beta$ -глобинспецифических РНК человека *in vitro*, начинающихся между —235-м нуклеотидом и каноническим сайтом экпирования, не ингибируется низкими дозами  $\alpha$ -аманитина (2 мкг/мл); для его ингибирования необходимы значительно большие дозы (100 мкг/мл) [75].

Небольшая часть РНК ретикулоцитов мыши содержит последовательность  $\beta^{\text{maj}}$ -глобина, присоединенную к последовательностям, транскрибированным с 5'-пограничного района  $\beta^{\text{maj}}$ -глобинового гена. Эти мРНК также полиаденилированы, содержат 700—800 нуклеотидов и имеют гетерогенные 5'-концы. Подобные молекулы РНК транскрибировались *in vitro*, когда в качестве матрицы использовалась кольцевая гибридная плазида, включавшая ген  $\beta^{\text{maj}}$ -глобина. Синтез большей



части этой РНК также угнетался низкими дозами  $\alpha$ -аманитина, указывая на то, что она транскрибируется другими ферментами, отличными от РНК-полимеразы II [77]. Было показано, что транскрипты, синтезированные РНК-полимеразой III *in vitro*, могут быть значительной протяженности и захватывать зоны, обычно транскрибируемые РНК-полимеразой II [78].

Места инициации транскрипции вблизи глобинового гена изучали *in vitro* в изолированных клетках эритролейкемии Фрейнда в присутствии радиоактивно меченных  $\gamma$ -тиотрифосфатов. Синтез  $\beta$ -глобинспецифических транскриптов иницировался  $\gamma$ -тиоАТР в сайте экпирования и угнетался  $\alpha$ -аманитином в концентрации 1 мкг/мл. Еще несколько сайтов инициации отмечались в 5'-направлении от сайта экпирования; синтез в них начинался  $\gamma$ -тиоАТР, но не  $\gamma$ -тиоГТР. Наряду с этим выявлялись дополнительные сайты инициации транскрипции, не угнетавшиеся  $\alpha$ -аманитином в той же концентрации [79].

Ген соматомаммотропина и ген соматотропина человека имеют по два сайта инициации транскрипции. Около 95 % транскриптов начинается через 30 нуклеотидных пар после ТАТААА-последовательности, остальные же начинаются через 30 нуклеотидных пар после САТААА-последовательности, размещенной на 55 нуклеотидных пар впереди ТАТА-бокса [80]. Единственный ген, кодирующий большую субъединицу рибулозбифосфаткарбоксилазы у кукурузы, также имеет два сайта инициации транскрипции, расстояние между которыми более 200 нуклеотидов [81]. Множественные мРНК 3-окси-3-метилглутарилкоэнзим А редуктазы образовались как за счет гетерогенности 3'-концевых последовательностей, так и множественности сайтов инициации транскрипции [82]. Ген субъединицы II цитохромоксидазы гороха образует две мРНК с различающимися 5'-концами [83]; мРНК дигидрофолатредуктазы человека и мыши гетерогенны не только по 3'-концу [35, 36], но и 5'-концы этих мРНК также гетерогенны. Такая гетерогенность отчасти коррелирует с присутствием в гене двух последовательностей промоторных элементов. Для чего нужны два промотора гену дигидрофолатредуктазы, пока неясно. Такое же наблюдение сделано для гена  $\alpha$ -амилазы мыши, алкогольдегидрогеназы дрозофилы, миозина скелетной мышцы курицы, инвертазы дрожжей [84, 85], соматомаммотропина и соматотропина человека [80].

**Заключение.** У бактерий информационные РНК обычно представляют собой первичный транскрипт — точную реплику последовательностей ДНК в геноме. У эукариот первичный транскрипт преобразуется в мРНК только после серии модификаций: экпирования, расщепление с образованием 3'-конца, полиаденилирование, метилирование внутренних оснований, сплайсинг. В принципе каждый из этих этапов представляет возможность для регуляции образования мРНК.

Похоже, что контроль дифференцированной экспрессии генов путем регуляции сайта терминации транскрипции и/или полиаденилирования широко осуществляется у эукариот. Не исключено, что различия в 3'-нетранслируемых районах в какой-то степени влияют на субклеточную компарментализацию мРНК и, следовательно, на судьбу транслированного белка. Например, один белок, транслированный с одного образца мРНК, может секретироваться, в то время как другой может ассоциировать еще с каким-либо из белков [42].

Сайты полиаденилирования в мРНК кальцитонина и родственного кальцитонину пептида, транскрибирующиеся с одного и того же гена, расположены в 4-м и 6-м экзонах соответственно. Транскрипция всего гена не связана с преимущественным производством того или иного пептида. Однако, хотя *in vitro* используются оба сайта полиаденилирования, проксимальный (кальцитониновый) сайт используется предпочтительно в ядрах из тканей, продуцирующих мРНК кальцитонина. По-видимому, выбор дистального сайта дает транскрипт, для которого более кинетически благоприятен сплайсинг, приводящий к образованию экзона кальцитонинподобного пептида [43]. Тканеспецифическая экспрессия

была отмечена для двух типов мРНК, транскрибирующихся с одного гена виментина курицы. Оба образца мРНК присутствуют в мышечных клетках, фибробластах, спинном мозге и хрусталике, но в эритроидных клетках 10—15-дневных эмбрионов в основном доминирует мРНК меньших размеров. Различие этих мРНК лежит в различной длине их 3'-нетранслируемых районов [40]. Множественные промоторы также могут играть роль в дифференцированной экспрессии генов в различных типах клеток. Например, продукт реакции, катализируемой дигидрофолатредуктазой (тетрагидрофолат), необходим для синтеза *de novo* глицина, пуринов и тимидилата. Глицин и пурины необходимы для синтеза белка и РНК во всех клетках, в то время как тимидилат нужен только во время синтеза ДНК. Два промотора могут работать таким образом, что один из них функционирует в неделящихся клетках, другой — обеспечивает значительно большую необходимость в тетрагидрофолате при делении клеток [84, 85].

Наличие более одного сайта инициации транскрипции одного и того же гена может сказаться на различных способах сплайсинга, в результате приводя к образованию различных мРНК. С таким явлением мы встречаемся в случае гена  $\alpha$ -амилазы, однако здесь не происходит изменения структуры кодируемого полипептида [86]. РНК-транскрипты генов легких цепей миозина курицы, начинающиеся с различных сайтов, по-разному сплайсируются, что приводит к образованию двух различных мРНК и отсюда двух различных полипептидов [87].

Возможно, существование нескольких промоторов, выполняющих различные регуляторные функции (например, гормональная индукция, тканевая специфичность, определение времени появления в развитии), является общим свойством многих генов эукариот.

#### AMBIGUITY OF THE TRANSCRIPTION BOUNDARIES OF EUKARYOTIC GENES

V. M. Kavsan

Institute of Molecular Biology and Genetics,  
Academy of Sciences of the Ukrainian SSR, Kiev

##### Summary

A great number of experimental data testify to a more complicated mechanism of the eucaryotic mRNA generation as compared to that of procaryotic mRNAs. The latter have a transcription process which begins and terminates in the sites corresponding to 5'- and 3'-ends of functional mRNAs. The nuclear cells appear to have special control mechanisms of differential gene expression by selecting sites for initiation and termination of the transcription and/or for polyadenylation of mRNAs.

1. Coutelle C. The precursor to animal cell mRNA // *Biochem. J.*—1981.—197, N 1.— P. 1—6.
2. Кавсан В. М. Строение семейства глобиновых генов как модель строения генов эукариот // *Молекуляр. биология.*—1983.—17, № 1.— С. 6—32.
3. Williamson R., Drewienkiewicz C., Paul J. Presence of globin messenger sequences in high molecular weight RNA from mouse liver // *Nature.*—1973.—241, N 107.— P. 66—68.
4. Imaizumi T., Diggelmann A., Scherrer K. Demonstration of globin messenger in giant nuclear precursors of messenger RNA of avian erythroblasts // *Proc. Nat. Acad. Sci. USA.*—1973.—70, N 4.— P. 1122—1126.
5. Spohr G., Imaizumi T., Scherrer K. Synthesis and processing of nuclear precursor-messenger RNA in avian erythroblasts and *HeLa* cells // *Ibid.*—1974.—71, N 12.— P. 5009—5013.
6. Messing J. Globin messenger precursor RNA in duck immature red blood cells // *Eur. J. Biochem.*—1978.—91, N 2.— P. 587—609.
7. Hayashi Y., Mikami E. Precursor of globin mRNA in erythroid-enriched bone marrow cells of mouse // *J. Biochem.*—1978.—83, N 3.— P. 699—711.
8. Reynaud C.-A., Imaizumi-Scherrer M. T., Scherrer K. Size determination of the transcriptional units of the avian globin genes defined at the pre-messenger RNA level // *Formation of messenger RNA in eucaryotic cells: Abstr. IV Arolla Workshop.*—Arolla, 1980.— P. 91.
9. Therwath A., Mengod G., Scherrer K. Transcription of globin genes in AEV transformed erythroblasts // *Ibid.*— P. 107.

10. Reynaud C.-A., Imaizumi-Scherrer M.-T., Scherrer K. The size of the transcriptional units of the avian globin genes defined at the pre-messenger RNA level // *J. Mol. Biol.*—1980.—140, N 4.—P. 481—505.
11. Strair R. K., Skoultchi A. J., Shatritz D. A. A characterization of globin mRNA sequences in the nucleus of duct immature red blood cells // *Cell.*—1977.—12, N 1.—P. 133—141.
12. Ross J. Purification and structural properties of the precursors of the globin mRNAs // *J. Mol. Biol.*—1978.—119, N 1.—P. 21—37.
13. Shaul Y., Kaminchik T., Aviv H. Large globin RNA molecules and their processing // *Eur. J. Biochem.*—1981.—116, N 3.—P. 461—466.
14. Bastos R. N., Aviv H. Globin RNA precursor molecules: biosynthesis and processing in erythroid cells // *Cell.*—1977.—11, N 3.—P. 641—650.
15. The absence of a precursor larger than 16S to globin mRNA / J. R. Haynes, V. F. Kolb, P. Rastock, A. Longrel // *FEBS Lett.*—1978.—91, N 2.—P. 173—177.
16. Kessler-Icekson G., Moreau J., Scherrer K. Isolation of globin pre-messenger RNA on thiol-agarose by terminally mercurocurated complementary DNA // *Mol. Biol. Repts.*—1981.—7, N 1—3.—P. 83—93.
17. Rohrbaugh M., Hardison R. Analysis of rabbit  $\beta$ -like globin gene transcripts during development // *J. Mol. Biol.*—1983.—164, N 3.—P. 395—419.
18. Кавсан В. М. Синтез и строение генов глобинов и инсулинов: Автореф. дис. ... д-ра биол. наук.—М., 1985.—43 с.
19. Золотухин С. Б., Кавсан В. М. Глобинспецифические ядерные РНК эритроидных клеток костного мозга кролика // *Биополимеры и клетка.*—1985.—1, № 4.—С. 208—213.
20. Транскрипция генов глобинов кролика: множественность сайтов терминации / С. Б. Золотухин, В. И. Кашуба, А. П. Коваль, В. М. Кавсан // V Всесоюз. биохим. съезд: Тез. стенод. сообщ.—Киев, 1985.—Т. 2.—С. 355.
21. Solutuchin S. B., Kavsan V. M. Ambiguity of transcription boundaries of eucaryotic genes // *Macromolecules in the functioning cell*: Abstr. 4-th symp. USSR—Italy (Kiev, 1984).—Kiev: Naukova dumka, 1984.—P. 117.
22. The major late adenovirus type-2 transcription is downstream from the last poly(A) site / N. W. Fraser, J. R. Nevins, E. Ziff, J. E. Darnell // *J. Mol. Biol.*—1979.—129, N 4.—P. 643—666.
23. Nevins J. R., Blanchard J. M., Darnell J. E. Transcription units of adenovirus type-2, termination of transcription beyond the poly(A) addition site in early regions 2 and 4 // *Ibid.*—1980.—144, N 3.—P. 377—386.
24. Ford J. P., Hsu M.-T. Transcription pattern of *in vivo* labelled late simian virus 40 RNA: equimolar transcription beyond the mRNA 3' terminus // *J. Virol.*—1978.—28, N 3.—P. 759—801.
25. Hofer E., Darnell J. E. The primary transcription unit of the mouse  $\beta$ -major globin gene // *Cell.*—1981.—23, N 2.—P. 585—595.
26. Transcription termination occurs within a 1000 base pair region downstream from the poly(A) site of the mouse  $\beta$ -globin (major) gene / B. Cötron, E. Falck-Pedersen, M. Saldiff-Georgieff, J. E. Darnell // *Nucl. Acids Res.*—1984.—12, N 22.—P. 8723—8732.
27. Solutuchin S. B., Koval A. P., Kavsan V. M. Bidirectional *in vivo* transcription of DNA segment downstream of the rabbit  $\beta$ -globin gene // 9th Nucle(ol)ar Work Shop (Cracow, 13—17 september, 1985).—Cracow, 1985.—P. 47.
28. Hagenbüchle O., Bovey R., Young R. A. Tissue-specific expression of mouse  $\alpha$ -amylase genes: nucleotide sequence of isoenzyme mRNAs from pancreas and salivary gland // *Cell.*—1980.—21, N 1.—P. 179—189.
29. The ovalbumin gene family: hormonal control of X and Y gene transcription and mRNA accumulation / M. Le Meur, N. Glanville, I. L. Mandel et al. // *Ibid.*—1981.—23, N 2.—P. 561—573.
30. Hamer D. H., Leder P. Expression of the chromosomal mouse  $\beta^{\text{maj}}$ -globin gene cloned in SV40 // *Nature.*—1979.—281, N 5826.—P. 35—40.
31. Two mRNAs can be produced from a single immunoglobulin  $\mu$  gene by alternative RNA processing pathways / P. Early, J. Rogers, M. Davis et al. // *Cell.*—1980.—20, N 2.—P. 313—321.
32. Mouse spleen and IgD-secreting plasmacytomas contain multiple IgD  $\delta$  chain RNAs / L. Fitzmaurice, J. Owens, F. R. Blattner et al. // *Nature.*—1982.—296, N 5856.—P. 459—462.
33. Transcription of mouse  $\kappa$  chain genes: implications for allelic exclusion / R. P. Perry, D. E. Kelley, C. Coleclough et al. // *Proc. Nat. Acad. Sci. USA.*—1980.—77, N 4.—P. 1937—1941.
34. Multiple polyadenylation sites in a mouse  $\alpha$ -amylase gene / M. Tosi, R. A. Young, O. Hagenbüchle, U. Schibler // *Nucl. Acids Res.*—1981.—9, N 10.—P. 2313—2323.
35. Variation in the polyadenylation site of bovine prolactin mRNA / N. L. Sasavage, M. Smith, S. Gillam et al. // *Proc. Nat. Acad. Sci. USA.*—1982.—79, N 2.—P. 223—227.
36. Size heterogeneity in the 3' end of dihydrofolate reductase m-RNAs in mouse cells / D. R. Setzer, M. M. Grogan, J. H. Nunberg, R. T. Schimke // *Cell.*—1980.—22, N 2.—P. 361—370.
37. Sehgal P. B. How many human interferons are there // *Interferon.*—1982.—4.—P. 1—23.

38. *Multiple initiation and polyadenylation sites for the chicken ovomucoid transcription unit* / P. Gerlinger, A. Krust, M. Lemeur et al. // *J. Mol. Biol.*—1982.—162, N 2.—P. 345—364.
39. *Human fibroblast interferon RNA transcripts of different sizes in poly(I)·poly(C) induced cells* / J. Content, L. DeWit, J. Tavernier, W. Fiers // *Nucl. Acids Res.*—1983.—11, N 9.—P. 2627—2637.
40. *Tissue-specific expression of two mRNA species transcribed from a single vimentin gene* / Y. G. Capetanaki, J. Ngai, C. N. Flytzanis, E. Lazarides // *Cell.*—1983.—35, N 2.—P. 411—420.
41. *Identification of multiple species of calmodulin messenger RNA using a full length complementary DNA* / L. Lagace, T. Chandra, S. L. C. Woo, A. K. Means // *J. Biol. Chem.*—1983.—258, N 3.—P. 1684—1689.
42. *Parnes J. R., Robinson R. R., Seidman J. C.* Multiple mRNA species with distinct 3' termini are transcribed from the  $\beta_2$ -microglobulin gene // *Nature.*—1983.—302, N 5907.—P. 449—452.
43. *Amara S. G., Evans R. M., Rosenfeld M. G.* Calcitonin / calcitonin gene-related peptide transcription unit: tissue-specific expression involves selective use of alternative polyadenylation sites // *Mol. and Cell. Biochem.*—1984.—4, N 10.—P. 2151—2160.
44. *Fine structural analysis of the human pro- $\alpha_1$ (I) collagen gene. Promoter structure, AluI repeats, and polymorphic transcripts* / M.-L. Chu, W. Wet, M. Bernard, F. Ramirez // *J. Biol. Chem.*—1985.—260, N 4.—P. 2315—2320.
45. *Isolation of cDNA and genomic clones encoding human pro- $\alpha_1$ (III) collagen. Partial characterization of the 3' end region of the gene* / M.-L. Chu, D. Weil, W. Wet et al. // *Ibid.*—N 7.—P. 4357—4364.
46. *Zaret K. S., Sherman F.* DNA sequence required for efficient transcription termination in yeast // *Cell.*—1982.—28, N 2.—P. 563—573.
47. *Salditt-Georgieff M., Darnell J. E., Jr.* A precise termination site in the mouse  $\beta$ -major globin transcription unit // *Proc. Nat. Acad. Sci. USA.*—1984.—81, N 7.—P. 2274.
48. *Transcription unit of the rabbit  $\beta_1$ -globin gene* / M. Rohrbaugh, J. Johnson, M. James, B. Hardison // *Mol. and Cell. Biochem.*—1985.—5, N 1.—P. 147—160.
49. *Transcription of the mouse dihydrofolate reductase gene proceeds unabated through seven polyadenylation sites and terminates near a region of repeated DNA* / E. Frayne, E. Leys, G. Crause et al. // *Ibid.*—1984.—4, N 12.—P. 2921—2924.
50. *Proudfoot N. J., Brownlee G. G.* Sequence at the 3' end of globin mRNA shows homology with immunoglobulin light chain mRNA // *Nature.*—1974.—252, N 5482.—P. 359—362.
51. *Газарян К. Г., Тарантул В. З.* Геном эукариот.—М.: Изд-во Моск. ун-та, 1983.—268 с.
52. *Wickens M., Stephenson P.* Role of the conserved AAUAAA sequence four AAUAAA point mutants prevent messenger RNA 3' end formation // *Science.*—1984.—226, N 4678.—P. 1045—1052.
53. *Fitzgerald M., Shenk T.* The sequence 5'-AAUAAA-3' forms part of the recognition site for polyadenylation of late SV40 mRNAs // *Cell.*—1981.—24, N 1.—P. 251—261.
54. *Inhibition of RNA cleavage but not polyadenylation by a point mutation in mRNA 3' consensus sequence AAUAAA* / C. Montell, E. F. Fisher, M. H. Caruthers, A. I. Berk // *Nature.*—1983.—305, N 5935.—P. 600—605.
55. *Thalassemia due to a mutation in the cleavage-polyadenylation signal of the human  $\beta$ -globulin gene* / S. H. Orkin, T. Cheng, C. Antonarakis, H. H. Kazazian // *EMBO J.*—1985.—4, N 2.—P. 453—456.
56.  *$\alpha$ -Thalassemia caused by a polyadenylation signal mutation* / D. R. Higgs, S. E. Goodbourn, J. Lamb et al. // *Nature.*—1983.—306, N 5941.—P. 398—400.
57. *Boardman M., Basi G., Sterti R.* Multiple polyadenylation sites in a *Drosophila* tropomyosin gene are used to generate functional mRNAs // *Nucl. Acids Res.*—1985.—13, N 5.—P. 1763—1777.
58. *Exons encode functional and structural units of chicken lysozyme* / A. Jung, A. E. Sippel, M. Grez, G. Schutz // *Proc. Nat. Acad. Sci. USA.*—1980.—77, N 10.—P. 5759—5764.
59. *Polyadenylation of the Xenopus  $\beta_1$ -globin mRNA at a downstream minor site in the absence of the major site and utilization of an AAUACA polyadenylation signal* / P. J. Mason, M. B. Jones, J. A. Elkington, J. C. Williams // *EMBO J.*—1985.—4, N 1.—P. 205—213.
60. *The consensus sequence YGTGTTY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini* / J. McLauchlan, D. Yaffing, J. L. Whitton, J. B. Clements // *Nucl. Acids Res.*—1985.—13, N 4.—P. 1347—1368.
61. *Manley J. L., Yu H., Ryner L.* RNA sequence containing hexanucleotide AAUAAA directs efficient mRNA polyadenylation *in vitro* // *Mol. and Cell. Biochem.*—1985.—5, N 2.—P. 373—379.
62. *3'-Editing of mRNAs: sequence requirements and involvement of a 60-nucleotide RNA in maturation of histone mRNA precursors* / C. Birchmeier, D. Schumperli, G. Sconzo, M. L. Birnstiel // *Proc. Nat. Acad. Sci. USA.*—1984.—81, N 4.—P. 1057—1061.
63. *Stunnenberg H., Birnstiel M. L.* Bioassay for components regulating eucaryotic gene expression: a chromosomal factor involved in the generation of histone mRNA 3'-termini // *Ibid.*—1982.—79, N 20.—P. 6201—6205.
64. *Mattaj J. W.* snRNAs from gene architecture to RNA processing // *Trends Biochem. Sci.*—1984.—9, N 10.—P. 435—438.

65. *Turner Ph.* Gene regulation: controlling roles for snurps // *Nature*.— 1985.—316, N 6024.— P. 105—107.
66. *Multiple 5'-terminal cap structures in late polyoma virus RNA / A. J. Flavell, A. Cowre, L. Legon, R. Kamen // Cell*.— 1979.—16, N 2.— P. 357—371.
67. *Baker C. C., Ziff E. B.* Promoters and heterogeneous 5'-termini of the mRNAs of adenovirus serotype 2 // *J. Mol. Biol.*— 1981.—149, N 2.— P. 189—223.
68. *Dolan M., Dodgson J. B., Engel J. D.* Analysis of the adult chicken  $\beta$ -globin gene. Nucleotide sequence of the locus, microheterogeneity at the 5' end of  $\beta$ -globin mRNA, and aberrant nuclear RNA species // *J. Biol. Chem.*— 1983.—258, N 6.— P. 3983—3991.
69. *Cowie A., Tyndall C., Kamen R.* Sequences at the capped 5' ends of polyoma virus late region mRNAs: an example of extreme terminal heterogeneity // *Nucl. Acids Res.*— 1981.—9, N 23.— P. 6305—6323.
70. *Shatkin A. J.* Introduction. Elucidating mechanism of eukaryotic genetic expression by studying animal viruses // *Current topics in microbiology and immunology* // Eds. W. Henle et al.— New York etc: Springer-Verlag, 1981.— V. 93.— P. 1—4.
71. *Heterogeneous initiation regions for transcription of the chicken ovomucoid gene / E. C. Lai, D. R. Roop, M.-J. Tsai et al. // Nucl. Acids Res.*— 1982.—10, N 18.— P. 5553—5568.
72. *Multiple origins of transcription for the human placental lactogenes / Ch. Selvanayagam, S. Tsai, M. Tsai et al. // J. Biol. Chem.*— 1984.—259, N 23.— P. 14642—14647.
73. *Expression of tissue-specific ren1 and ren2 genes of mice: comparative analysis of 5'-proximal flanking regions / L. J. Field, W. M. Philbrick, P. N. Howles et al. // Mol. and Cell. Biochem.*— 1984.—4, N 11.— P. 2321—2331.
74. *Allan M., Lanyon W. G., Paul J.* Multiple origins of transcription in the 4.5 Kb upstream of the globin gene // *Cell*.— 1983.—35, N 1.— P. 187—197.
75. *Alternative sites of transcription initiation upstream of the canonical cap site in human  $\gamma$ -globin and  $\beta$ -globin genes / G. J. Grindlay, W. G. Lanyon, M. Allan, J. Paul // Nucl. Acids Res.*— 1984.—12, N 4.— P. 1811—1820.
76. *Carlson D. P., Ross J.* Human  $\beta$ -globin promoter and coding sequences transcribed by RNA polymerase III // *Cell*.— 1983.—34, N 3.— P. 857—864.
77. *Carlson P. P., Ross R.*  $\alpha$ -Amanitin-insensitive transcription of mouse  $\beta^{\text{major}}$ -globin 5'-flanking and structural gene sequences correlates with mRNA expression // *Proc. Nat. Acad. Sci. USA*.— 1984.—81, N 24.— P. 7782—7786.
78. *Manley J. L., Colozzo M. T.* Synthesis *in vitro* of an exceptionally long RNA transcript promoted by an Alu1 sequence // *Nature*.— 1982.—300, N 5890.— P. 376—379.
79. *Winitcov I., Weidner D.* Transcription initiation in  $\beta$ -globin region from induced friend cell nuclei // 13th Int. Congr. Biochem.: Abstr. (Friday, August 30).— Amsterdam, 1985.— P. 656.
80. *Analysis of a major human chorionic somatomammotropin gene / M. J. Selby, A. Barla, J. D. Baxter et al. // J. Biol. Chem.*— 1984.—259, N 21.— P. 13131—13138.
81. *Crossland L. D., Rodermel S. R., Bogorad L.* Single gene for the large subunit of ribulosebiphosphate carboxylase in maize yields of two differentially regulated mRNAs // *Proc. Nat. Acad. Sci. USA*.— 1984.—81, N 13.— P. 4060—4064.
82. *Reynolds G. A., Goldstein J. L., Brown M. S.* Multiple mRNAs for 3-hydroxy-3-methylglutaryl coenzyme A reductase by multiple transcription initiation sites and intron splicing sites in the 5'-untranslated region // *J. Biol. Chem.*— 1985.—260, N 18.— P. 10369—10377.
83. *Moon E., Kao T., Wu R.* Pea cytochrome oxidase subunit II gene has no intron and generates two mRNA transcripts with different 5'-termini // *Nucl. Acids Res.*— 1985.—13, N 9.— P. 3195—3213.
84. *Heterogeneity at the 5'-termini of mouse dihydrofolate reductase mRNAs. Evidence for multiple promoter regions / M. McGrogan, C. C. Simonsen, D. T. Smouse et al. // J. Biol. Chem.*— 1985.—260, N 4.— P. 2307—2314.
85. *Masters J. N., Attardi G.* Discrete human dihydrofolate reductase gene transcripts present in polysomal RNA map with their 5'-ends several hundred nucleotides upstream of the main mRNA start site // *Mol. and Cell. Biochem.*— 1985.—5, N 3.— P. 493—500.
86. *Two promoters of different strengths control the transcription of the mouse alpha-amylase gene amyla in the parotid gland and the liver / U. Schibler, O. Hagenbüchle, P. K. Wellauer, A. C. Pittet // Cell*.— 1983.—33, N 2.— P. 501—508.
87. *Alternative transcription and two modes of splicing result in two myosin light chains from one gene / Y. Nabeshima, Y. Fujii-Kuriyama, M. Muramatsu, K. Ogata // Nature*.— 1984.—308, N 5997.— P. 333—338.