

РЕГУЛЯЦИЯ ТРАНСКРИПЦИИ БЕЛОККОДИРУЮЩИХ ГЕНОВ МАКРОНУКЛЕУСА ИНFUЗОРИЙ-СПИРОТРИХ

© А. Ю. Пименов,¹ И. Н. Сквородкин, И. Б. Райхель, С. О. Скарлато

Институт цитологии РАН, Санкт-Петербург;

¹ электронный адрес: iamshura@yahoo.com

Представления о механизмах регуляции транскрипции и строении основного промотора белоккодирующих генов эукариотических организмов формировались на основе данных, полученных с использованием двух подходов: 1) использовании различных методов направленного мутагенеза (экспериментальный подход) и 2) сопоставлении нуклеотидных последовательностей промоторных областей исследованных генов с нуклеотидными последовательностями хромосом, не подвергавшихся экспериментальному изучению (сравнительно-молекулярный подход). Изучение представителей типа Ciliophora привело исследователей к выводам об отсутствии в генах класса II этих организмов классических элементов регуляции, характерных для большинства эукариот. Причина этого заключается в том, что основным методом большинства исследователей был сравнительно-молекулярный анализ нуклеотидных последовательностей. При этом экспериментальная проверка соответствующих областей хромосом отсутствовала. В настоящей работе на основе новых экспериментальных данных, полученных в результате исследования гена $\alpha 1$ -тубулина инфузории *Stylonychia lemnae*, произведен анализ функциональной роли регуляторных компонентов 5'-некодирующей области генов тубулина, катепсина и белка теплового шока 70 этой инфузории. Выполнен сравнительно-молекулярный анализ промоторных областей и впервые создана классификация механизмов регуляции транскрипции генов класса II инфузорий-спиротрих.

Ключевые слова: инфузории, макронуклеус, мини-хромосомы, регуляция транскрипции.

Принятые сокращения: п. н. — пара нуклеотидов, ТАФ — ТСБ-ассоциированные факторы, ТСБ — ТАТА-боксысвязывающий белок.

Специфичность связывания фермента РНК-полимеразы II с молекулой ДНК в процессе формирования транскрипционного комплекса определяется, с одной стороны, участками нуклеотидной последовательности ДНК, называемыми регуляторными, с другой — ферментом и различными транскрипционными факторами, имеющими белковую природу. Регуляторные последовательности ДНК, необходимые для инициации транскрипции, являются сайтами связывания определенных субъединиц фермента и белковых транскрипционных факторов (Ленинджер, 1974; Альбертс и др., 1994).

Представления о механизмах регуляции транскрипции и строении регуляторных нуклеотидных последовательностей ДНК белоккодирующих генов эукариотических организмов формировались на основе данных, полученных с использованием двух подходов: использовании различных методов направленного мутагенеза (экспериментальный подход) и сопоставлении нуклеотидных последовательностей промоторных областей исследованных генов с нуклеотидными последовательностями хромосом, не подвергавшихся экспериментальному изучению, с целью поиска известных элементов регуляции (сравнительно-молекулярный подход) (Hoffman et al., 1995; Пименов и др., 2006).

Изучение регуляции инициации транскрипции генов класса II эукариотических организмов привело к созда-

нию классической схемы строения промоторной области этих генов (Lee, Young, 2000). Так, для генов класса II большинства эукариот постулировано присутствие таких регуляторных нуклеотидных последовательностей, как СААТ-боксы и ТАТА-боксы в положениях 50—150 и 25—30 п. н. левее нуклеотида, с которого начинается транскрипция, соответственно и инициатора нуклеотидной последовательности, окружающей нуклеотид, с которого начинается синтез РНК.

Одновременно с этим изучение представителей типа Ciliophora привело исследователей к выводам об отсутствии в генах класса II этих организмов описанных классических элементов регуляции (Ghosh et al., 1994). Причина этого заключается в том, что основным методом большинства исследователей был теоретический анализ нуклеотидных последовательностей различных генов с целью поиска уже известных регуляторных нуклеотидных последовательностей в их классических положениях (уже упомянутый сравнительно-молекулярный подход) (Helftenbein, 1985; Conzelmann, Helftenbein, 1987; Helftenbein, Muller, 1988; Helftenbein et al., 1989; Steinbruck, 1990; Пименов и др., 2006). При этом экспериментальная проверка соответствующих областей хромосом отсутствовала, что ставит под сомнение выводы, сделанные исследователями.

Особый интерес для исследования механизмов регуляции транскрипции у ресничных простейших представ-

ляет группа брюхоносных инфузорий. Говоря о таких инфузориях, мы придерживаемся классификации типа Ciliophora, согласно которой этот тип включает в себя класс Spirotrichea, содержащий отряды Spirotrichida (род *Stylonychia* и род *Oxytricha* (*Sterkiella*)) и Hypotrichida (род *Euplotes*), класс Oligohymenophorea (род *Tetrahymena*) и класс Nassophorea (род *Paramecium*) (Lynn, Small, 1985). При этом представителей отрядов Spirotrichida и Hypotrichida исторически принято объединять в группу брюхоносных инфузорий, поэтому, говоря о родах *Stylonychia*, *Oxytricha* (*Sterkiella*) и *Euplotes*, мы будем использовать в равной степени термины «брюхоносные инфузории» и «инфузории-спиротрихи».

Ядерный гетероморфизм, характерный для всего типа Ciliophora, представлен у инфузорий-спиротрих наиболее ярко (Райков, 1978, 1989; Осипов, 1981; Klobutcher, Prescott, 1986). ДНК в транскрипционно активном макронуклеусе этих животных представлена короткими линейными молекулами — так называемыми мини-хромосомами, кодирующими, как правило, один-единственный ген. Таким образом, каждый макронуклеарный ген представляет собой автономную репликационную и транскрипционную единицу, а все регуляторные элементы молекулы ДНК, необходимые для корректной инициации транскрипции, сосредоточены в сравнительно короткой лидерной фланкирующей области мини-хромосом, что значительно сужает диапазон исследований (Осипов, 1981; Klobutcher, Prescott, 1986).

Появившееся в литературе большое количество противоречивых данных относительно состава основного промотора генов класса II представителей типа Ciliophora одновременно с удобством использования генов брюхоносных инфузорий в качестве экспериментального материала дало толчок к адаптации классических и разработок новых экспериментальных подходов в изучении рассматриваемого вопроса (Skovorodkin et al., 1999, 2001a, 2001b; Pimenov et al., 2002; Райхель, 2003; Пименов и др., 2006). Так, с использованием инфузории-спиротрихии *Stylonychia lemnae* в качестве модельного объекта был разработан метод котрансфекций инфузорий мутантными векторами, созданными на основе мини-хромосом этих организмов и содержащими делеции и нуклеотидные замены в области мини-хромосомы, расположенной между триплетом АТГ, с которого начинается кодирующая область гена, и теломерой. Эту область мини-хромосом в данной работе мы будем обозначать как 5'-некодирующую область гена (Skovorodkin et al., 1999). Такой подход позволяет с точностью до нуклеотида определять строение, местоположение и функции регуляторных последовательностей ДНК. В результате с использованием мини-хромосомы, несущей ген $\alpha 1$ -тубулина *S. lemnae*, был впервые полностью исследован промотор гена класса II инфузорий. Описаны такие классические эукариотические регуляторные нуклеотидные последовательности, как СААТ-бокс, ТАТА-бокс, инициатор, энхансер и репрессор (Pimenov et al., 2002, 2007; Skovorodkin et al., 2007). Впервые для генов класса II инфузорий-спиротрих в основном промоторе локализована нуклеотидная последовательность, контролирующая индукцию транскрипции гена $\alpha 2$ -тубулина *S. lemnae* (Skovorodkin et al., 2007). Экспериментально определены точные местоположения нуклеотидов, с которых начинается синтез РНК, в генах тубулина, катепсина и белка теплового шока *S. lemnae* (Pimenov et al., 2002).

Результаты экспериментальных исследований регуляции инициации транскрипции гена α -тубулина *S. lemnae*

предоставили большое количество материала для сравнительно-молекулярного анализа белоккодирующих генов инфузорий-спиротрих. В связи с этим в настоящем обзоре был проведен анализ функциональной роли регуляторных компонентов 5'-некодирующей области гена $\alpha 1$ -тубулина инфузории *S. lemnae*, выполнен сравнительно-молекулярный анализ промоторных областей и создана классификация механизмов регуляции транскрипции белоккодирующих генов инфузорий-спиротрих.

Регуляция транскрипции генов класса II эукариот

Транскрипция — один из ключевых процессов, обеспечивающих реализацию генетической информации организма и выражающийся в передаче информации о нуклеотидной последовательности ДНК на уровень РНК (Ленинджер, 1974). Транскрипция генов, кодирующих белки, осуществляется с помощью фермента РНК-полимеразы II, что и определяет их отнесение к группе генов класса II (Ленинджер, 1974; Альбертс и др., 1994).

Транскрипция включает в себя три основных этапа: инициацию, элонгацию цепи РНК и терминацию. Регуляция синтеза РНК, безусловно, происходит на всех этапах транскрипции, однако считается, что ключевую роль играют правильность и частота инициации этого процесса (Альбертс и др., 1994).

Для правильной инициации транскрипции необходимо связывание фермента РНК-полимеразы II с молекулой ДНК в нужном участке с образованием транскрипционного комплекса. Специфичность такого связывания определяется, с одной стороны, участками последовательности ДНК, называемыми регуляторными, с другой — ферментом и различными транскрипционными факторами, имеющими белковую природу. Транскрипция начинается только после сборки всех необходимых элементов в единый транскрипционный комплекс (Roberts, 2000) (рис. 1).

Нуклеотидные последовательности ДНК, необходимые для инициации транскрипции, можно условно подразделить на две группы (Альбертс и др., 1994; Pedersen et al., 1999; Lee, Young, 2000). Во-первых, это нуклеотидные последовательности, расположенные в непосредственной близости от нуклеотида, с которого начинается синтез РНК, — так называемый основной промотор (рис. 2). Эти элементы принимают прямое участие во взаимодействии с ферментом и формировании транскрипционного комплекса. Ко второй группе можно отнести области молекулы ДНК, участвующие в связывании дополнительных факторов регуляции транскрипции и располагающиеся на значительном расстоянии от нуклеотида, с которого начинается синтез РНК.

В состав основного промотора большинства генов класса II эукариот входит АТ-богатый участок, называемый ТАТА-боксом (Альбертс и др., 1994; Pedersen et al., 1999; Lee, Young, 2000) (рис. 2). У высших эукариот ТАТА-боксы располагаются в области -25 — -30 п. н., у дрожжей — в области -40 — -120 п. н. (здесь и далее приводится расстояние в числе пар нуклеотидов от нуклеотида, с которого начинается транскрипция, в направлении 5'-теломера) и служит сайтом для связывания так называемого ТАТА-боксысвязывающего белка (ТСБ), который вместе с ТСА-ассоциированными факторами (ТАФ) формирует один из главных факторов транскрипции — TFIID (рис. 1). Классическая нуклеотидная последовательность

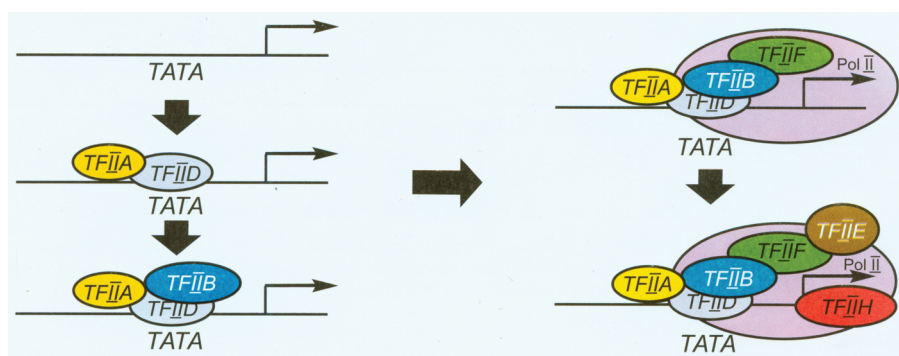


Рис. 1. Схема сборки транскрипционного комплекса (по: Roberts, 2000, модифицировано).

Молекула ДНК изображена в виде отрезка прямой; *изогнутая стрелка* — местоположение нуклеотида, с которого начинается транскрипция; *TATA* — TATA-боксы; *Pol II* — субъединицы фермента РНК-полимеразы II; *TFIIA*, *TFIIB*, *TFIID*, *TFIIE*, *TFIIF* и *TFIIH* — белковые транскрипционные факторы.

этого регуляторного элемента характеризуется алгоритмом 5'-ТАТА(А/Т)А(А/Т)-3' и, как правило, имеет нуклеотидный состав ТАТААА. Однако в некоторых случаях ТСБ может связываться и с другими последовательностями, что затрудняет точную идентификацию ТАТА-боксов (Pedersen et al., 1999; Lee, Young, 2000).

Еще одним элементом основного промотора является инициатор (Альбертс и др., 1994; Smale et al., 1998; Pedersen et al., 1999; Lee, Young, 2000) (рис. 2). Это нуклеотидная последовательность, окружающая нуклеотид, с которого начинается транскрипция, и также связывающаяся с регуляторными факторами. Нуклеотидный состав инициатора более вариабелен, нежели таковая ТАТА-боксы, и часто характеризуется алгоритмом 5'-YYANT/AYY-3', где Y — пиримидиновые основания, а N — любой нуклеотид (Smale et al., 1998). У высших эукариот инициатор богат пиримидинами, между тем у жгутиконосцев рода *Giardia* в нем преобладают тимин и аденин (Elmendorf et al., 2001).

По наличию ТАТА-боксов и инициатора промоторы генов класса II подразделяют на несколько групп. Так, для большинства высших эукариот характерны промоторы, содержащие только ТАТА-боксы. Напротив, у целого ряда генов этих организмов в промоторах присутствует только инициатор. Встречаются также гены, в регуляции транскрипции которых принимают участие оба рассматриваемых элемента. Интересно, что существуют промоторы, в которых отсутствуют и ТАТА-боксы, и инициатор (Pedersen et al., 1999; Lee, Young, 2000).

Отсутствие каких-либо консервативных регуляторных элементов не подразумевает полного отсутствия регуляции транскрипции. Во многих генах найдены так называемые ТАТА-боксы-подобные последовательности и целый ряд дополнительных регуляторных элементов, имеющих различную структуру и местоположение (Ghosh et al., 1994; Pedersen et al., 1999; Lee, Young, 2000). В случае отсутствия консервативной последовательности

ТАТА-боксов в районе его классического местоположения могут присутствовать другие регуляторные элементы, специфические для конкретных генов. Так, в гене β-тубулина у *Drosophila* в области между -26 и -32 п. н. присутствует регуляторная последовательность GAACAT, так же как и ТАТА-боксы, участвующая в образовании транскрипционного комплекса (Santel et al., 2000). Кроме этого, дополнительные регуляторные элементы были найдены и в генах, промоторы которых содержат ТАТА-боксы и инициатор (Lee, Young, 2000).

Говоря о дополнительных регуляторных элементах основного промотора, необходимо отметить, что из них наиболее часто встречаются так называемый GC-боксы и CAAT-боксы, характеризующиеся алгоритмами строения 5'-GGGCGGG-3' или 5'-CCCGCCC-3' и 5'-GG(T/C)CAATCT-3' соответственно. Как правило, эти элементы располагаются в области -50—-150 п. н. (Жимулев, 2002; Hoffman et al., 1995) (рис. 2).

Наконец, последнюю группу регуляторных элементов, содержащихся в основном промоторе, составляют нуклеотидные последовательности, располагающиеся правее нуклеотида, с которого начинается синтез РНК, в том числе и в кодирующей области гена (Альбертс и др., 1994; Santel et al., 2000) (рис. 2). Такое расположение регуляторных последовательностей, характерное главным образом для генов класса III, встречается и в некоторых генах класса II. Например, в гене β-тубулина у *Drosophila* правая регуляторная последовательность располагается в области от +49 до +72 п. н. (Santel et al., 2000). Многие авторы считают эти элементы компонентами основного промотора, поскольку они принимают участие в связывании фактора TFIID, но не через ТСБ, а через ТАФ, т. е. в промоторах, в которых отсутствует ТАТА-боксы, являются функциональными аналогами (Santel et al., 2000).

Разнообразие существующих регуляторных элементов, контролирующих инициацию транскрипции, не огра-

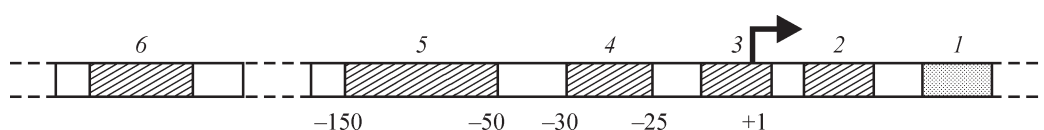


Рис. 2. Схема строения классического эукариотического промотора.

Представлена 5' — некодирующая область белкокодирующего гена эукариот. Цифрами под схемой указано число пар нуклеотидов от местоположения нуклеотида, с которого начинается транскрипция. *Изогнутая стрелка* — местоположение нуклеотида, с которого начинается транскрипция. 1 — кодирующая область гена, 2 — правые промоторные элементы, 3 — инициатор транскрипции, 4 — ТАТА-боксы, 5 — возможное положение дополнительных регуляторных элементов — СААТ-боксов и GC-боксов, 6 — возможное положение дополнительных регуляторных элементов — энхансеров и репрессоров.

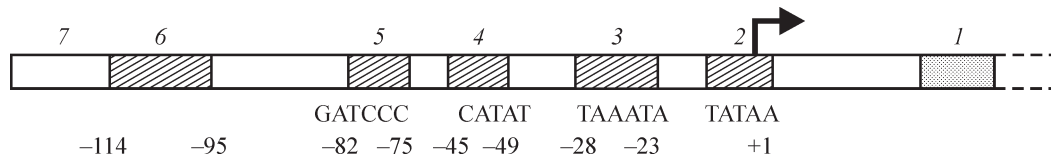


Рис. 3. Схема промотора гена $\alpha 1$ -тубулина *Styloynchia lemnae*.

Представлена 5'-некодирующая область гена $\alpha 1$ -тубулина *S. lemnae*. Цифрами под схемой указано число пар нуклеотидов от местоположения нуклеотида, с которого начинается транскрипция. Изогнутая стрелка — местоположение нуклеотида, с которого начинается транскрипция. 1 — кодирующая область гена, 2 — инициатор транскрипции, 3 — TATA-бокс, 4 — CAAT-бокс, 5 — GC-бокс, или энхансер, 6 — репрессор, 7 — теломера.

ничивается только последовательностями, принимающими участие в образовании транскрипционного комплекса. На процесс инициации транскрипции могут оказывать активирующее, ингибирующее или блокирующее влияние элементы, расположенные на значительном расстоянии от нуклеотида, с которого начинается синтез РНК (Альбертс и др., 1994; Lee, Young, 2000). Эти элементы также являются сайтами связывания различных транскрипционных факторов, но их влияние носит, как правило, опосредованный характер.

Несмотря на отсутствие четкой классификации данного типа регуляторных последовательностей, их можно разделить на несколько классов в соответствии с характером влияния на инициацию транскрипции (Lee, Young, 2000). К первой группе обычно относят активаторы, или энхансеры. Это кластеры сайтов для связывания белков, активирующих инициацию транскрипции. Важно, что действие энхансеров не зависит от их пространственной ориентации по отношению к кодирующей области гена.

Вторую группу рассматриваемых регуляторных последовательностей составляют репрессоры, роль которых сводится к ингибированию или блокаде инициации транскрипции. Действие репрессоров, так же как и активаторов, не зависит от их местоположения.

Характер действия белковых факторов транскрипции очень разнообразен. Один и тот же белок-активатор, связывающийся с энхансерами, может приводить к активации транскрипции в целом спектре генов. Вместе с тем транскрипция в каждом отдельном гене активируется многими активаторами. Этот факт говорит о том, что белки-активаторы, возможно, работают кооперативно. Действительно, в большинстве случаев наблюдается объединение этих элементов в агрегаты (энхансеромы) с помощью белок-белковых взаимодействий, что значительно повышает эффективность их работы (Lee, Young, 2000).

Белковые факторы, осуществляющие репрессорную функцию и имеющие, как правило, высококонсервативное строение, подразделяются на две группы — общие и геноспецифические (Lee, Young, 2000). Белки-репрессоры общего действия, как правило, связываются с комплексом ДНК—ТСБ и вызывают его диссоциацию, сопровождающуюся гидролизом АТФ. Большинство геноспецифических белков-репрессоров конкурирует с белками-активаторами за взаимодействие с сайтами связывания последних и тем самым подавляет инициацию транскрипции.

Работа активаторов и репрессоров во многом предопределяет не только частоту, но и подчас самую возможность инициации транскрипции. Однако точная регуляция этого процесса не может осуществляться без участия другой группы белков-регуляторов, а именно тех элементов, которые формируют транскрипционный комплекс и определяют точность связывания субъединиц фермента с молекулой ДНК (Burley, 1998; Smale et al., 1998). Необходи-

мо отметить, что общая номенклатура этих белковых факторов инициации транскрипции не отражает их функций и особенностей строения. Так, в буквенных обозначениях известных транскрипционных факторов — TFIIA, TFIIB, TFIIC, TFIID, TFIIIE, TFIIF и TFIIN — первые две буквы обозначают «транскрипционный фактор», цифра «II» свидетельствует о том, что этот фактор работает во взаимодействии с ферментом РНК-полимеразой II, а последняя буква указывает на то, каким по порядку этот фактор был открыт (Альбертс и др., 1994).

Наконец, необходимо сказать о роли самого фермента в определении точного местоположения нуклеотида, с которого начинается транскрипция. РНК-полимераза II, состоящая из нескольких субъединиц, способна к прочному, неспецифическому и обратимому связыванию с молекулой ДНК (Альбертс и др., 1994). В момент взаимодействия фермента с необходимым сигнальным участком ДНК происходит распознавание нуклеотидной последовательности одной из субъединиц фермента с последующим ее отделением. Оставшийся холофермент способен к прочному связыванию ДНК и остается на месте до начала элонгации цепи.

Функциональная роль регуляторных компонентов, влияющих на инициацию транскрипции, локализованных в 5'-некодирующей области гена $\alpha 1$ -тубулина инфузории *S. lemnae*

В результате экспериментального исследования 5'-некодирующего участка гена $\alpha 1$ -тубулина инфузории *S. lemnae*. У этой инфузории были обнаружены такие классические эукариотические регуляторные элементы, влияющие на инициацию транскрипции, как инициатор, TATA-бокс, CAAT-бокс и две дополнительные нуклеотидные последовательности, также влияющие на синтез РНК (Pimenov et al., 2002; Skovorodkin et al., 2007; Pimenov et al., 2007) (рис. 3). Различные нарушения в процессе инициации транскрипции, возникающие вследствие изменения нуклеотидного состава регуляторных элементов ДНК методами направленного мутагенеза, позволяют понять схему их взаимодействия.

Процесс сборки транскрипционного комплекса гена $\alpha 1$ -тубулина начинается с того, что ТАФ связывается со специфической последовательностью ТАА инициатора, а ТСБ — с областью TAAATA TATA-бокса, расположенной между позициями -23—-28 п. н., распознавая регуляторную последовательность. Эффективность взаимодействия ТАФ и инициатора повышается при участии фактора TFIIA, который, по-видимому, вызывает конформационные изменения в комплексе TFIIA—ДНК. Далее следует присоединение всех остальных факторов и фер-

мента с образованием единого комплекса посредством белок-белковых и белок-ДНК взаимодействий. При этом ТСБ, распознавая необходимую последовательность, связывается с молекулой ДНК точно в области ТАААТА ТА-ТА-бокса. Это взаимодействие обуславливает место прикрепления к молекуле ДНК ТАФ, определяя тем самым местоположение нуклеотида, с которого начинается транскрипция. Присутствие в инициаторе специфической нуклеотидной последовательности ТАА увеличивает скорость и точность взаимодействия ТАФ с молекулой ДНК, уменьшая время сборки транскрипционного комплекса. Такая функциональная роль инициатора объясняет, почему уровень транскрипции генов, в промоторе которых инициатор отсутствует, низок (Pimenov et al., 2002, 2007; Skovorodkin et al., 2007).

Большое функциональное значение имеет расстояние между ТАТА-боксом и инициатором. В случае нарушения нуклеотидной последовательности ТАТА-бокса место присоединения ТСБ к молекуле ДНК может быть несколько смещено, что приводит к неточному присоединению ТАФ. В результате местоположения нуклеотида, с которого начинается транскрипция, определяется только ТАФ, которые взаимодействуют с подходящими нуклеотидными последовательностями ТАА, присутствующими в необходимой области мини-хромосомы. Возникающая таким образом вариабельность местоположения сайта присоединения ТАФ к молекуле ДНК определяет возможность появления дополнительных абберрантных местоположений нуклеотида, с которого начинается транскрипция. Нуклеотидных последовательностей, имеющих нуклеотидный состав инициатора ТАА в области нуклеотида, с которого начинается транскрипция гена $\alpha 1$ -тубулина в норме, три, что соответствует трем типам РНК, образующимся в случае нарушения нуклеотидной последовательности ТАТА-бокса (Pimenov et al., 2002, 2007; Skovorodkin et al., 2007).

СААТ-бокс, располагающийся в 5'-некодирующей области гена $\alpha 1$ -тубулина между -45 и -49 н. п. и имеющий строение САТАТ, является типичным для эукариот элементом кор-промотора (Жимулев, 2002). При этом его функция заключается в контроле первоначального связывания РНК-полимеразы II с промотором, что в свою очередь обеспечивает правильное позиционирование прединициационного комплекса транскрипции на соответствующем сегменте молекулы ДНК. Известно, что если промотор содержит только инициатор транскрипции, а другие основные транскрипционные факторы отсутствуют, очищенная РНК-полимераза II инициирует транскрипцию малоэффективно (Cargano et al., 1991). Некоторые другие данные, например снижение эффективности транскрипции в результате сайт-специфических мутаций последовательностей ДНК, расположенных в определенных областях промотора, также подтверждают предположение о том, что СААТ-бокс модулирует эффективность транскрипции у эукариот (McKnight, Kingsbury, 1982; Graves et al., 1986; Barberis et al., 1987). По-видимому, у *S. lemnae* СААТ-бокс в гене, кодирующем $\alpha 1$ -тубулин, также играет роль вспомогательного элемента кор-промотора, усиливающего интенсивность транскрипции этого гена.

Наряду с регуляторными нуклеотидными последовательностями, входящими в состав основного промотора, 5'-некодирующая область гена $\alpha 1$ -тубулина содержит и две дополнительные нуклеотидные последовательности ДНК, оказывающие антагонистическое влияние на синтез РНК (Pimenov et al., 2007; Skovorodkin et al., 2007).

Так GC-бокс, выполняющий функцию активатора, расположен внутри нуклеотидной последовательности ТТААГАТССССТТАТСААА гена $\alpha 1$ -тубулина между -65 и -86 п. н. Допуская возможность некоторых отклонений от классической нуклеотидной последовательности GC-бокса, которая представлена нуклеотидами 5'-GGGCGGG-3' или 5'-CCCGCC-3', можно сделать предположение о том, что в гене $\alpha 1$ -тубулина *S. lemnae* GC-бокс представлен нуклеотидами GATССС, расположенными между -75 и -82 п. н.

Противоположное, ингибирующее влияние на синтез РНК оказывает участок гена $\alpha 1$ -тубулина, расположенный между 5'-теломером и позицией -95 п. н. Основываясь на литературных данных о механизмах сборки транскрипционного комплекса и особенностях ее регуляции, можно предположить, что в области между теломером и позицией -95 п. н. расположен один из следующих сайтов: либо 1) сайт связывания некоего транскрипционного фактора, участие которого в процессе инициации замедляет сборку транскрипционного комплекса и тем самым снижает частоту актов синтеза цепей РНК; либо 2) сайт связывания белка, оказывающего опосредованное влияние на процесс инициации транскрипции с участием известных реакций ацетилирования—деацетилирования, фосфорилирования—дефосфорилирования хромосомы; либо 3) сайт связывания факторов, активаторов и репрессоров, включающихся в транскрипционный комплекс на конкурентной основе. Во всех представленных вариантах роль рассматриваемой нуклеотидной последовательности может быть охарактеризована как репрессорная, т. е. снижающая количество синтезируемой РНК. При этом, принимая во внимание особенности строения мини-хромосом инфузорий-спиротрих, наиболее вероятным представляется механизм, основанный на конкуренции факторов транскрипции, активаторов и репрессоров, входящих в состав транскрипционного комплекса. При наличии в комплексе инициации транскрипции белкового фактора, выполняющего репрессорную функцию, увеличивается время сборки комплекса, это приводит к снижению частоты актов синтеза цепей РНК, что характерно для гена $\alpha 1$ -тубулина *S. lemnae* в норме. Однако можно предположить ситуацию, в которой в транскрипционном комплексе гена $\alpha 1$ -тубулина данный белок-репрессор отсутствует или заменяется на белок-активатор, что приводит к увеличению количества синтезируемой РНК.

Строение регуляторных компонентов, влияющих на инициацию транскрипции, локализованных в 5'-некодирующей области генов $\alpha 2$ -, $\beta 1$ - и $\beta 2$ -тубулина, катепсина и белка теплового шока 70 *S. lemnae*

Сравнение нуклеотидной последовательности промотора гена $\alpha 1$ -тубулина *S. lemnae*, исследованной экспериментально, с нуклеотидными последовательностями 5'-некодирующей области генов $\alpha 2$ -, $\beta 1$ - и $\beta 2$ -тубулина, катепсина и белка теплового шока 70 этой инфузории, в которых местоположения нуклеотидов, с которых начинается транскрипция, были определены ранее (Pimenov et al., 2002), показало, что ТАТА-бокс в гене $\alpha 1$ -, $\alpha 2$ - и $\beta 2$ -тубулина представлен нуклеотидами ТАААТА, расположенными в случае генов $\alpha 1$ - и $\alpha 2$ -тубулина в позиции -23 п. н., а в случае $\beta 2$ -тубулина — в позиции -24 п. н. (рис. 4). ТАТА-бокс в промоторе гена $\beta 1$ -тубулина иной,

ATCTTAAATA CTAAASTTA ACTAGTAATTAT **A** ATTATTT CATTTTTTCA AACAAAATCA ACTCTTCATC ATG
 ↓
 α1-турбулин ATCTTAAATA CTAAASTTA ACTAGTAATTAT **A** ATTATTT CATTTTTTCA AACAAAATCA ACTCTTCATC ATG
 α2-турбулин CCTATAAATA GCGAAATAAA GTGGGAAATTAT **A** ACTTTAT TTTATATTA AACAAAATCA AATCTTCATC ATG
 β1-турбулин ATTCAGAGAA TGAATATAAG GATACCAATT TTAAGAAGTATT **A** ATTATTC TTAATTTTCA AAAACAAACC ATG
 β2-турбулин TGTATAAAT ATGCGCCCTC TCTAACAAT ATAT **A** ATTCA AATAATTTT AAAACAAAAC TCAAGTCACA ATG
 Катепсин TAA TTATAGTCAC TAGGATCCAA ATTAATTAT **A** AATAATTATA ATG
 Белок теплового шока 70 TCATTAT TTAAGAGCGC TTAGTAAAT CAAGATT **A** AT TGTAAATAA AGATACCSTA ACAAACACAA ATG

Рис. 4. 5' — некодирующие участки генов α1-, α2-, β1- и β2-тубулина, катепсина и белка теплового шока 70 *Stylonychia lemnae*.

Жирным шрифтом выделены триплеты ATG, с которых начинается кодирующая область гена; ТАТА-боксы и инициаторы подчеркнуты двойной линией; нуклеотиды, с которых начинается транскрипция, указаны стрелкой.

Stylonychia

Актин-1 TAAGTAATA GGGGATATCTGAAATTAATAAATAATTAT **A** AGTAGATTTTATCAAGAACAT ATG
 Кальмодулин AAGAAATCATATTTAAGAGCGCTTAGTTAAATCAATAT **A** ATTGTTAAATAAAGATACCCCTAACAAACACAA ATG
 eRF1 ACAAATATAATTAAATTTATAAATAATTAATTCAT **A** AATAATAAACCCCAAAGGTTACTCTATATATACTATAAAA ATG
 H4 (клон H4g) CTAGTCTATATAAGGATCCACCTCTGTAAATTAAT **A** ATTTCATAAATAAATAAATCAA ATG
 H4 (клон H4k) TGTGAGACTATTTAAGCGACTGCATTTTAAAGTCT **A** AATAAGAAATTAATAA ATG
 Rob4 ATATATTTGGCCGGCCACCGGTTATTTAATAATT **A** ATATAAATTATAACTAATAT ATG

Oxytricha

eRF1 AATTAATAATATAAAGAGTCTTAATCATTATAAATAAT **A** AACACAAAAGGCTACCTATATATACTATAAAA ATG
 eRF2 AATTAATAATATAAAGAGTCTTAATCATTATAAATAAT **A** AACACAAAAGGCTACCTATATATACTATAAAA ATG
 eRF3 CCCTAGGAGTAATTGTAATAATTTATAATTAATAAATAAT **A** AATAAGTCTTAGAGTAGCCTTTATCAAAAAACAAA ATG
 Шаперон GAGCATACTTAATAAATTTAAGGAGCGAAGTCT **A** ATTATTTAAATTTCAATTATAAATCT ATG
 Убиквитин GTTTATAGTATAAAGCCAGTGAATTCTATTAATAAATAAT **A** ATTATAAATAAATACTTTTACACACAACAAAACAACAAC ATG
 TERT GATGAGTTTTATTGGCGTAATAATAAATAAATAATTT **A** ATATATCTTAAGACTATTAAGTGATCTGAA ATG
 PGK TTAAGTTTGAATGATGACT **A** ATTATATCTTTAACTTCACTCAAAATCATT ATG
 Ев-фосфофара AGAGGATATTAGAGTGGATATAAATAATATCAGGAATA **A** ATAGTATTAATAAATAATAGATTAAAGGATTA ATG
 α-тубулин TGTACCCCTATATAATTTAAGCGGATACGCCGATAGT **A** ATATAATTTTTTTCTTATTTTCAAAACAACCTTCTCAACATC ATG

Рис. 5. 5' — некодирующие участки белоккодирующих генов инфузорий родов *Stylonychia* и *Sterkiella* (*Oxytricha*).

Жирным шрифтом выделены триплеты ATG, с которых начинается кодирующая область гена; ТАТА-боксы и инициаторы подчеркнуты двойной линией; нуклеотиды, с которых начинается транскрипция, указаны стрелкой.

Euplotes focardii AAAAAACCAAAA ATTAAATAA TTCTTATACA TCTTTATATC TAATTAATAA ACAAACATG ATG
 α-турбулин Euplotes octocarinatus CCCCCAAAACC CTATAGGGA GCCCAAAATT GAAATTTATA TAATAATAAA CTACCAAG ATG
 β2-турбулин Euplotes nobilii TAAATACAT TTAAGTAAAA TTCTGGGAGA STAATTAATT AATAAATTT AATAAATAAT ATG
 Белок теплового шока 70 Euplotes focardii ATAGGTATTT AAAAATCCGA AAATAGTCTT CATTAATTAA AAATTAATAA CACTAACAAA ATG
 Белок теплового шока 70 Sterkiella histriomuscorum TTAAGATAGT TAATATAATT ATTTTCATTT TTTTTCATA CATACATACA ACTCTTCATC ATG
 α-турбулин Sterkiella histriomuscorum TATAAATGGG ATAGACCTAG TAATTGTAAT ACTTATCTT AATTAACAAC CACACACCCC ATG
 β-турбулин Sterkiella histriomuscorum TTTAGGACTT ATAAATTGAT GGATACCTTT TAATTTATA ATAATATCTA AATTTTTATA ATG
 Катепсин Sterkiella histriomuscorum TATAAAGCCT GAGATGGATC GAGAATTATA ATAATAATTT AACAAATCAC TTACATAAAT ATG
 Катепсин Oxytricha nova TTTACTATATATAAGCC TTTGCTGCCT TAATTTATAA AGTTATTATA CTAACACAA CAACACTCTC AAATATACAA ATG
 Белок теплового шока 70

Рис. 6. 5' — некодирующие участки генов α1-, α2-, β1- и β2-тубулина, катепсина и белка теплового шока 70 инфузорий родов *Euplotes* и *Sterkiella* (*Oxytricha*).

Жирным шрифтом выделены триплеты ATG, с которых начинается кодирующая область гена; предполагаемые ТАТА-боксы и инициаторы подчеркнуты двойной линией.

он представлен нуклеотидами ТАТАА в позиции –24 п. н. (рис. 4).

В рассматриваемых генах также обнаруживаются сходные нуклеотидные последовательности в области нуклеотида, с которого начинается синтез РНК (рис. 4). Так, в генах катепсина, $\alpha 1$ -, $\alpha 2$ - и $\beta 2$ -тубулина эта нуклеотидная последовательность имеет состав ТАТАА, в генах $\beta 1$ -тубулина и белка теплового шока 70 — АТТАА. Однако во всех генах присутствует последовательность ТАА, где первый аденозин является нуклеотидом, с которого начинается транскрипция (рис. 4).

Рассмотренные результаты сравнительно-молекулярного анализа промоторных областей генов $\alpha 1$ -, $\alpha 2$ -, $\beta 1$ - и $\beta 2$ -тубулина, катепсина и белка теплового шока 70 *S. lemnae* позволяют говорить о большом сходстве нуклеотидного состава ТАТА-боксов и инициатора у исследованных генов (Pimenov et al., 2002; Skovorodkin et al., 2007) (рис. 4).

Сравнительно-молекулярный анализ 5'-некодирующих областей генов класса II инфузорий-спиротрих

При сравнении нуклеотидных последовательностей 5'-некодирующих областей генов инфузорий родов *Stylonychia* и *Sterkiella* (*Oxytricha*), в которых местоположение нуклеотидов, с которых начинается транскрипция, было определено экспериментально разными авторами, последовательность инициатора ТАТАА была найдена во всех 6 исследованных генах инфузорий рода *Stylonychia* и в 6 из 9 генах инфузорий рода *Sterkiella* (рис. 5) (Пименов, 2007). В 3 генах инфузорий рода *Sterkiella* было обнаружено смещение последовательности инициатора на 3 п. н. ТАТА-бокс, имеющий строение ТААТА, был обнаружен только в 3 из 6 генов инфузорий рода *Stylonychia* и в 5 из 9 генов инфузорий рода *Sterkiella* (рис. 5) (Пименов, 2007).

При сравнении нуклеотидных последовательностей промоторных областей генов $\alpha 1$ -, $\alpha 2$ -, $\beta 1$ - и $\beta 2$ -тубулина, а также катепсина и белка теплового шока 70 у представителей родов *Euplotes*, *Sterkiella* и *Tetrahymena* оказалось, что в указанных генах местоположение нуклеотидов, с которых начинается транскрипция, экспериментально не определено ни одним исследователем (рис. 6). В соответствии с этим на основании имеющихся нуклеотидных последовательностей нами были сделаны предположения относительно нуклеотидного состава ТАТА-боксов и инициатора, а также местоположения нуклеотидов, с которых начинается синтез РНК. В результате нуклеотидные последовательности, которые могут соответствовать инициатору, были обнаружены во всех 9 генах (рис. 6). В соответствии с предполагаемым положением нуклеотидов, с которых начинается транскрипция, ТАТА-бокс-подобные последовательности были также обнаружены во всех 9 генах (рис. 6).

Сравнительный анализ 5'-некодирующих областей генов $\alpha 1$ -, $\alpha 2$ -, $\beta 1$ - и $\beta 2$ -тубулина, катепсина и белка теплового шока 70 у инфузорий родов *Euplotes*, *Sterkiella* (*Oxytricha*) и *Tetrahymena* показал, что элементы регуляции транскрипции могут быть обнаружены с помощью сравнительно-молекулярных методов только тогда, когда местоположение нуклеотида, с которого начинается синтез РНК, определено экспериментально.

Сопоставление строения ТАТА-боксов, инициатора, СААТ-боксов и GC-боксов в генах $\alpha 1$ -, $\alpha 2$ -, $\beta 1$ - и $\beta 2$ -тубули-

на, катепсина и белка теплового шока 70 инфузории *S. lemnae* с классическими эукариотическими алгоритмами строения этих элементов свидетельствует о том, что лишь ТАТА-бокс является в полной мере консервативным элементом и его нуклеотидный состав одинаков как для генов класса II спиротрих, так и для большинства других эукариот. Строение других элементов регуляции транскрипции значительно варьирует, поэтому можно говорить об их консервативности лишь в пределах определенной группы генов или организмов.

Заключение

Подводя итог сравнительно-молекулярным исследованиям генов класса II инфузорий-спиротрих, можно прийти к выводам о том, что регуляция транскрипции генов класса II у этих организмов осуществляется при участии промоторов разного строения. Промоторы генов класса II инфузорий-спиротрих могут быть подразделены на три основные группы: а) промоторы, построенные по классической схеме, характерной для многих эукариот, и содержащие только ТАТА-бокс и инициатор; б) промоторы, также характерные для многих эукариот и содержащие наряду с ТАТА-боксом и инициатором транскрипции ряд дополнительных регуляторных элементов, таких как СААТ-бокс, GC-бокс, различные энхансеры и репрессоры; в) промоторы, содержащие специфические регуляторные нуклеотидные последовательности, характерные для некоторых генов и групп брюхохоресничных инфузорий.

Учитывая вариабельность строения нуклеотидных последовательностей ДНК, необходимых для инициации транскрипции генов класса II как инфузорий-спиротрих, так и эукариотических организмов в целом, можно прийти к заключению о том, что сравнительно-молекулярный анализ можно использовать только в том случае, когда местоположение нуклеотидов, с которых начинается синтез РНК, определено экспериментально. В этом случае при нахождении классических последовательностей ТАТА-боксов и инициатора можно говорить о механизмах регуляции транскрипции в исследуемых генах. Определение местоположения и нуклеотидного состава дополнительных элементов регуляции транскрипции, таких как СААТ-бокс, GC-бокс, различные энхансеры и репрессоры, ввиду их сильной вариабельности возможно только экспериментально.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 07-04-00662).

Список литературы

- Альбертс Б., Брей Д., Льюис Дж., Рэфф М., Робертс К., Уотсон Дж. 1994. Молекулярная биология клетки. М.: Мир. 2: 540 с.
- Жимулев И. Ф. 2002. Общая и молекулярная генетика. Новосибирск: Изд-во Новосибирск. ун-та. 459 с.
- Ленинджер А. 1974. Биохимия. М.: Мир. 953 с.
- Осипов Д. В. 1981. Проблемы гетероморфизма ядер у одноклеточных организмов. Л.: Наука. 167 с.
- Пименов А. Ю., Райхель И. Б., Сквородкин И. Н., Подлипаева Ю. И., Скарлато С. О. 2006. *Stylonychia lemnae* — модельный объект для изучения мини-хромосом макронуклеуса инфузорий. Цитология. 48 (8): 619—635.
- Райков И. Б. 1978. Ядро простейших. Морфология и эволюция. Л.: Наука. 328 с.

- Райков И. Б. 1989. Ядерный геном простейших. В кн.: Организация генома. М.: Наука. 110—154.
- Райхель И. Б. 2003. Регуляторные участки ДНК в минихромосомах инфузории *Stylonychia lemnae*, несущих гены $\alpha 1$ - и $\beta 2$ -тубулина: Автореф. канд. дис. СПб.: СПб. гос. политех. ун-т. 26 с.
- Barberis A., Superti-Furga G., Busslinger M. 1987. Mutually exclusive interaction of the CCAAT-binding factor and of a displacement protein with overlapping sequences of a histone gene promoter. *Cell*. 50 : 347—359.
- Burley S. K., Roeder R. G. 1996. Biochemistry and structural biology of transcription factor IID (TEIID). *Annu. Rev. Biochem.* 65 : 769—799.
- Carcamo J., Buckbinder L., Reinberg D. 1991. The initiator directs the assembly of a transcription factor IID-dependent transcription complex. *Proc. Nat. Acad. Sci. USA*. 88 : 8052—8056.
- Conzelmann K. K., Helftenbein E. 1987. Nucleotide sequence and expression of two beta-tubulin genes in *Stylonychia lemnae*. *J. Mol. Biol.* 198 : 643—653.
- Elmendorf H. G., Singer S. M., Pierce J., Cowan J., Nash T. E. 2001. Initiator and upstream elements in the tubulin promoter of *Giardia lamblia*. *Mol. Biochem. Parasitol.* 113 : 157—169.
- Ghosh S., Jaraczewski J. W., Klobutcher L. A., Jahn C. L. 1994. Characterization transcription initiation, translation initiation, and poly(A) addition sites in the gene-sized macronuclear DNA molecules of *Euplotes*. *Nucl. Acids Res.* 22 : 214—221.
- Graves B. J., Johnson P. F., McKnight S. L. 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. *Cell*. 44 : 565—576.
- Helftenbein E. 1985. Nucleotide sequence of a macronuclear DNA molecule coding for alpha-tubulin from the ciliate *Stylonychia lemnae*. Special codon usage: TAA is not a translation termination codon. *Nucl. Acids Res.* 13 : 415—433.
- Helftenbein E., Conzelmann K. K., Becker K. F., Fritzenschaf H. 1989. Regulatory structure of gene expression, DNA-replication and DNA-rearrangement in macronuclear genes of *Stylonychia lemnae*, a Hypotrichous Ciliate. *Eur. J. Protistol.* 25 : 158—167.
- Helftenbein E., Muller E. 1988. Both alpha-tubulin genes are transcriptionally active in *Stylonychia lemnae*. *Curr. Genet.* 13 : 425—432.
- Hoffman D. C., Anderson R. C., DuBois M. L., Prescott D. M. 1995. Macronuclear gene-sized molecules of hypotrichs. *Nucl. Acids Res.* 23 : 1279—1283.
- Klobutcher L. A., Prescott D. M. 1986. The special case of the hypotrichs. In: The molecular biology of ciliated protozoa. New York: Acad. Press. 111—154.
- Lee T. H., Young R. A. 2000. Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* 34 : 77—137.
- Lynn D. H., Small E. B. 2002. Phylum Ciliophora. In: An illustrated guide to the Protozoa. Lawrence, Kansas: Society of Protozoologists. 371—656.
- McKnight S. L., Kingsbury R. 1982. Transcription control signals of a eukaryotic protein-coding gene. *Science*. 217 : 316—324.
- Pedersen A. G., Baldi P., Chauvin Y., Brunak S. 1999. The biology of eukaryotic promoter prediction — a review. *Computers and Chemistry*. 23 : 191—207.
- Pimenov A., Skovorodkin Y., Raykhel I., Schimanski B., Ammermann D., Günzl A. 2007. Components of Alpha-tubulin minichromosome promoters in the stichotrichous ciliates *Stylonychia lemnae* involved in transcription regulation. *Protistology*. 5 (1) : 62.
- Pimenov A., Zassoukhina I. B., Ammermann D., Skovorodkin I. N. 2002. Characterization of transcription initiation in alpha and beta tubulin genes of *Stylonychia lemnae*. In: 21 Jahrestagung der deutschen Gesellschaft fuer Protozoologie. Konstanz. 35.
- Roberts S. 2000. Mechanisms of action of transcription activation and repression domains. *Cell Mol. Life Sci.* 57 : 1149—1160.
- Santel A., Kaufmann J., Hyland R., Renkawitz-Pohl R. 2000. The initiator element of the *Drosophila* $\beta 2$ tubulin gene core promoter contributes to gene expression *in vivo* but is not required for male germ-cell specific expression. *Nucl. Acids Res.* 6 : 1439—1446.
- Skovorodkin I. N., Bollgonn S., Ammermann D., Günzl A. 1999. Stable transfection of the hypotrichous ciliate *Stylonychia lemnae* with tagged alpha 1 tubulin minichromosomes. *Eur. J. Protistol.* 35 : 70—80.
- Skovorodkin I., Pimenov A., Raykhel I., Schimanski B., Ammermann D., Günzl A. 2007. Analysis of α -tubulin minichromosome promoters in the stichotrichous ciliate *Stylonychia lemnae*. *Eukaryot. Cell*. 6 : 28—36.
- Skovorodkin I. N., Zassoukhina I., Ammermann D., Günzl A. 2001a. Characterization of the promoter region in alpha-1 tubulin gene of the hypotrichous ciliate *Stylonychia lemnae*. In: XI Intern. Congr. Protozool. Salzburg. 13.
- Skovorodkin I. N., Zassoukhina I. B., Hojak S., Ammermann D., Günzl A. 2001b. Minichromosomal DNA replication in the macronucleus of the hypotrichous ciliate *Stylonychia lemnae* is independent of chromosome-internal sequences. *Chromosoma*. 110 : 352—359.
- Smale S. T., Jain A., Kaufmann J., Emami K. H., Garraway I. P. 1998. The initiator element: a paradigm for core promoter heterogeneity within Metazoan protein-coding genes. *Cold Spring Harbor Lab. Press*. 63 : 21—31.
- Steinbrueck G. 1990. Recent advances in the study of ciliate genes. *Eur. J. Protistol.* 26 : 2—14.

Поступила 7 IV 2008

TRANSCRIPTIONAL REGULATION MACRONUCLEAR PROTEIN-CODING GENES IN STICHOTRICHIOUS CILIATES

A. Yu. Pimenov,¹ L. N. Skovorodkin, I. B. Raikhel, S. O. Skarlato

Institute of Cytology RAS, St. Petersburg; ¹ e-mail: iamshura@yahoo.com

The concept of transcriptional regulation and base promoter structure of genes in eukaryotic organisms rests on two approaches — the experimental (using different methods of directional mutagenesis) and the comparative molecular (comparison of nucleotide sequences in promoter regions of different genes). Investigation of ciliates has led researchers to the conclusion that the protein-coding genes of these organisms lack the classical eukaryotic regulatory elements in the promoter region. This conclusion is based mainly on the usage of the comparative approach, while experimental investigations of such genes are practically absent so far. In the present paper, the comparative molecular analysis of the promoter regions of genes and analysis of the functional role of tubulin, cathepsin and Hsp 70 5'-uncoding areas was performed using new experimental data obtained during investigation of *Stylonychia lemnae* α -tubulin gene. We suggest a new classification of mechanism of transcriptional regulation of protein-coding genes in stichotrichous ciliates.

Key words: ciliates, macronucleus, minichromosomes, transcriptional regulation.