

УДК 577.21

Регуляторные элементы в векторах для эффективного получения клеточных линий-продуцентов целевых белков

О. Г. Максименко, Н. Б. Гасанов, П. Г. Георгиев*

Институт биологии гена РАН, 119334, Москва, ул. Вавилова, 34/5

*E-mail: georgiev_p@mail.ru

Поступила в редакцию 14.04.2015

РЕФЕРАТ В настоящее время все больше лекарственных препаратов производят в культурах клеток млекопитающих. С целью повышения уровня и стабильности экспрессии целевых рекомбинантных белков в клеточных культурах используют различные регуляторные элементы, механизмы действия которых пока недостаточно изучены. В обзоре охарактеризованы и обсуждены потенциальные механизмы действия таких регуляторных элементов.

КЛЮЧЕВЫЕ СЛОВА инсуляторы, рекомбинантные белки, продукция белков в клеточных линиях, UCOE, S/MAR, STAR.

СПИСОК СОКРАЩЕНИЙ РБ – рекомбинантный белок; СНО – культура клеток яичников китайского хомячка; НТО – нетранслируемая область гена; S/MAR – последовательности ДНК, ассоциированные с белками ядерного матрикса; инсуляторы – регуляторные элементы, которые блокируют взаимодействие между энхансером и промотором; UCOE – регуляторные элементы, содержащие сильные промоторы генов «домашнего хозяйства»; STAR – регуляторные элементы, которые защищают от HP1-зависимой репрессии; MSX – L-метионинсульфоксид; MTX – метотрексат.

ВВЕДЕНИЕ

Терапевтические белки – это основные структурные и регуляторные молекулы, необходимые для нормальной жизнедеятельности человеческого организма. Часть рекомбинантных белков небольших размеров и не требующих дополнительных модификаций производится в наиболее экономичном биореакторе – клетках кишечной палочки. Однако производство белков в бактериях связано с рядом ограничений: у некоторых белков нарушен фолдинг, не вносятся правильные модификации, невозможно производство крупных молекул [1, 2]. Часть этих ограничений снимается при использовании клеток дрожжей, в которых можно получить качественные рекомбинантные белки с достаточно низкой себестоимостью. Однако многие рекомбинантные белки человека должны подвергнуться определенным модификациям, возможным только при их экспрессии в клетках высших эукариот [3]. Поэтому в настоящее время все больше лекарственных препаратов получают в культурах клеток млекопитающих, выращенных в биореакторах, прежде всего, в клетках яичника китайского хомячка (СНО) [4, 5]. Впервые клетки СНО были выделены в 1957 году [6].

Вскоре стало понятно, что эти клетки оптимальны для масштабирования биомассы в процессе наработки рекомбинантных белков на уровне биореакторов, так как они нетребовательны к условиям культивирования. На основе первичного клона клеток СНО получены несколько линий, среди которых наибольшее распространение получила линия СНО-К1 [7]. С использованием данной линии были подобраны оптимальные условия для обеспечения высокой плотности роста клеток СНО в биореакторах, в результате чего удалось добиться значительного увеличения выхода продукта – целевого белка, при достаточно низкой вероятности передачи вирусов человека [1, 4, 5]. Однако основную проблему при производстве рекомбинантных белков в культурах клеток представляет крайне высокая себестоимость продукта, поэтому постоянно пытаются снизить затраты при получении высокопродуктивных клеточных культур и увеличить выход белкового продукта за счет повышения уровня экспрессии целевого белка и плотности клеточной культуры, а также снижения гибели клеток. Одно из таких направлений – усовершенствование векторов для получения трансгенов, что может значительно снизить себестоимость получения

клеточных линий-продуцентов. В настоящем обзоре рассмотрены регуляторные элементы, которые используются в векторных конструкциях, предназначенных для получения трансгенов.

ВЕКТОРНЫЕ КОНСТРУКЦИИ, ИСПОЛЬЗУЕМЫЕ ДЛЯ ПОЛУЧЕНИЯ КЛЕТОЧНЫХ ЛИНИЙ, ЭКСПРЕССИРУЮЩИХ ЦЕЛЕВЫЕ БЕЛКИ

Наибольшее распространение в промышленной биотехнологии получил метод трансфекции клеток линейаризованной плазмидной ДНК [3, 8], который позволяет получить клеточные линии, содержащие много копий экспрессионного вектора, обычно встроенных в один, реже – в несколько геномных сайтов. Механизм интеграции векторных конструкций в геном изучен не до конца. При попадании линейной ДНК в клеточные ядра происходит индукция систем репарации, которые обеспечивают сшивку концов линейных молекул ДНК по двум основным механизмам: гомологичной рекомбинации и соединения нехомологичных концов ДНК [9–11]. В результате формируются длинные линейные ДНК, содержащие несколько копий векторной конструкции, которые с некоторой вероятностью могут встраиваться в геном. Линейная ДНК с помощью механизма гомологичной рекомбинации может встраиваться в участок генома, в который ранее уже встроились другие копии такой же ДНК, что приводит к увеличению числа копий конструкции, интегрированных в одно место генома [12]. Таким образом, в один геномный сайт может встроиться до нескольких сотен копий векторной конструкции.

С целью повышения продуктивности клонов в промышленной биотехнологии используется селективное увеличение числа копий конструкции, которая обеспечивает экспрессию целевого белка [8]. Это достигается путем снижения функциональной активности репортерного гена, кодирующего фермент дигидрофолатредуктазу (DHFR), которая катализирует превращение дигидрофолата в тетрагидрофолат и необходима для биосинтеза глицина, пуринов и тимидиновой кислоты [13, 14]. Клеточные линии, ауксотрофные по дигидрофолатредуктазе, способны расти на среде, которая обязательно содержит глицин, источник пуринов (гипоксантин), и тимидин. Достаточно давно с помощью случайного мутагеназа линии CHO-K1 была получена производная линия CHO-DG44 [15, 16], в которой мутированы оба аллеля гена, кодирующего дигидрофолатредуктазу. В результате появилась возможность использовать ген дигидрофолатредуктазы в качестве селекционного гена для получения клеточных линий-продуцентов. При этом для селективного увеличения количества копий экспрессионного вектора, обычно коррелирующего с повышением уровня экспрессии

целевого белка, применяют метотрексат (МТХ), способный избирательно ингибировать превращение дигидрофолата в тетрагидрофолат [2]. Обработка трансфицированных линий клеток метотрексом приводит к тому, что выживают только клетки, в которых значительно повышен уровень дигидрофолатредуктазы. В большинстве случаев это происходит в результате увеличения числа копий гена *dhfr*, входящего в состав конструкции и, как следствие, гена, кодирующего целевой белок. Другой часто используемый селективный маркер – ген, кодирующий глутаминсинтазу (GS). При работе с этим маркером используют клеточные линии CHO-K1, содержащие мутантный аллель гена *gs*. В данном случае в качестве селективного агента, улучшающего отбор наиболее эффективных клонов, используется L-метионинсульфоксимин (MSX) [2].

Селекцию клонов можно проводить с использованием и других репортерных генов. Наиболее привлекательны варианты, в которых не надо выводить мутантную клеточную линию. В качестве такой системы можно рассматривать технологию с флуоресцентными белками [17, 18], которая позволяет по уровню свечения клеточного клона на определенной длине волны отбирать клетки, экспрессирующие целевой ген на максимальном уровне. С помощью этого подхода можно культивировать стабильные клеточные линии, которые, в отличие от мутантных производных, дефектных по генам *dhfr* и *gs*, имеют гораздо более высокий потенциал деления и жизнеспособность. К недостаткам применения в качестве селективных маркеров генов, кодирующих флуоресцентные белки, относится невозможность амплификации числа копий конструкции, что во многих случаях способствует увеличению продуктивности клеточного клона. В последнее время приступили к широкому использованию специальных роботов, способных отбирать отдельные клеточные клоны, наиболее эффективно экспрессирующие целевой белок, идентифицируемый с помощью антител [4]. Наряду с другими преимуществами, эта технология позволяет не использовать маркеры, экспрессия которых не всегда коррелирует с уровнем синтеза целевого белка.

Векторные конструкции разработаны также на основе вирусов, мобильных элементов, системы антифаговой защиты бактерий и систем рекомбинации фагов и дрожжей [3, 19]. Использование этих векторов позволяет вводить единичные инсерции целевого гена, в некоторых случаях – в заданный район генома, что широко применяют в генной терапии, при создании трансгенных клеточных линий и животных с целью получения модельных систем для изучения процессов регуляции экспрессии генов [20].

Основная проблема при получении клеточных линий-продуцентов, содержащих много копий экспрессионного вектора, – гетерохроматизация повторяющихся последовательностей векторных конструкций, которая часто усиливается в ряду клеточных делений. Основную роль в гетерохроматизации повторяющихся последовательностей ДНК играют РНК-интерференция и некодирующие РНК, которые могут стимулировать процесс формирования на промоторах зон репрессивного хроматина, метилирование CpG-сайтов в области промоторов, что снижает эффективность связывания транскрипционных факторов с данными участками [21–23]. В результате, через некоторое время после получения высокопродуктивных клеточных линий уровень экспрессии целевого белка в них может значительно снизиться. Одновременно формирование на повторяющихся копиях векторной конструкции зоны гетерохроматина может негативно влиять на активность клеточных генов, расположенных рядом, что часто приводит к снижению жизнеспособности клеточных линий-продуцентов. Репрессия транскрипции в интегрированных повторах векторной конструкции обусловлена защитной реакцией клетки на введение в геном чужеродной информации, экспрессию которой необходимо подавить. Таким образом, крайне актуальным представляется использование регуляторных элементов, способных поддерживать эффективную работу гена целевого белка и одновременно изолировать регуляторные элементы векторной конструкции от регуляторных элементов генома.

ПРОМОТОРЫ И ЭНХАНСЕРЫ, ИСПОЛЬЗУЕМЫЕ ДЛЯ ПОЛУЧЕНИЯ КЛЕТОЧНЫХ ЛИНИЙ-ПРОДУЦЕНТОВ ЦЕЛЕВЫХ БЕЛКОВ

Для экспрессии трансгена обычно используют сильные вирусные промоторы – цитомегаловирусный промотор (CMV) и ранний промотор вируса SV40, а также сильные клеточные промоторы генов «домашнего хозяйства», таких, как гены β -актина и фактора EF1 α [24].

Сильные вирусные и клеточные промоторы содержат минимальный промотор размером примерно 100 п.н., с которого начинается транскрипция, и сильный энхансер, расположенный в непосредственной близости от промотора. Так, наиболее часто используемый CMV-промотор состоит из коровой части, расположенной между позициями –62...–1 п.н. относительно начала транскрипции, и энхансера (от –544 до –63 п.н.) [25]. В области минимального промотора находятся несколько мотивов, которые определяют связывание с данной областью различных компонентов основного комплекса транскрипции (TFIID): TATA [26], INR [27], DPE [28], BRE [29],

DCE [30] и MTE [31]. Однако эти участки не относятся к обязательным, так как многие сильные промоторы не содержат этих элементов либо принадлежат к CG-богатым промоторам. Можно предположить, что существует практически неизученная группа так называемых архитектурных белков, определяющих способность корового промотора рекрутировать TFIID-комплекс [32]. К сожалению, архитектура промоторов на настоящий момент описана не в той мере, чтобы по последовательности определять способность минимального промотора эффективно связывать основные факторы транскрипции, необходимые для функционирования сильного промотора.

Один из подходов использования эффективных промоторов в биотехнологии – идентификация сильных промоторов непосредственно в тех клеточных культурах, на основе которых планируется получать продуценты целевых белков. Таким образом, был проведен полногеномный скрининг наиболее сильных промоторов в клетках СНО, наиболее часто используемых в качестве системы экспрессии в клетках млекопитающих [33–35]. Как и ожидалось, эффективными оказались промоторы генов «домашнего хозяйства», в том числе некоторых рибосомных генов. Однако подводным камнем такого подхода является довольно высокая вероятность того, что многие промоторы, идентифицированные с помощью полногеномного анализа, эффективно работают только в определенном участке генома («эффект положения») или содержат сложную регуляторную область, что сильно снижает привлекательность таких промоторов для использования в экспрессионных системах с целью получения требуемых белков в трансгенных линиях клеток СНО.

Одно из решений проблемы «эффекта положения» состоит в использовании длинных регуляторных последовательностей активно транскрибирующихся генов «домашнего хозяйства», расположенных с обеих сторон от кодирующей части. Так, высокий уровень экспрессии целевых белков (в 6–35 раз превышающий уровень, который обеспечивает стандартный промотор CMV) достигнут с помощью вектора, содержащего 12 т.п.н. регуляторной области гена EF1 α китайского хомячка и 4 т.п.н. 3'-области этого же гена [36]. Одной из проблем использования длинных регуляторных последовательностей с целью экспрессии целевого белка являются нестабильность больших векторов и снижение эффективности получения многокопийных линий, содержащих преимущественно полные конструкции, способные экспрессировать целевой белок. Перспективным представляется использование конструкции, содержащей длинные регуляторные участки ДНК из концевой повтора вируса Эпштейна–Барр, который на порядок усиливает

эффективность получения стабильно трансфицированных клеток [37].

Еще один интересный подход – искусственная модификация промоторов с целью усиления их активности. Например, показано, что сильный CMV-промотор находится под негативным контролем, который приводит к метилированию GC-участков в сайтах связывания транскрипционных факторов (ТФ), входящих в состав промоторов и, как следствие, к ингибированию рекрутирования ТФ на промотор. В результате активность CMV-промотора значительно снижается. Такой негативный эффект можно блокировать, встраивая между энхансером и коровыми участками CMV-промотора регуляторную последовательность, с которой связываются транскрипционные факторы, блокирующие процессы метилирования ДНК [38]. На основе двух CMV-промоторов создан эффективный промотор, состоящий из пары разнонаправленных коровых элементов, между которыми встроен один CMV-энхансер [39]. Такой двусторонний промотор способен с примерно одинаковой эффективностью экспрессировать два противоположно направленных гена, что имеет большое значение при продукции белков, состоящих из двух разных субъединиц (например, моноклональных антител).

Оригинальным способом повышения уровня экспрессии целевого белка является искусственное привлечение на промотор комплексов, ассоциированных с эффективной транскрипцией [40, 41]. Например, гистон-ацетилтрансфераза р300 связывается с активными энхансерами и промоторами и участвует в стимуляции транскрипции [42]. Рекрутирование р300 на промоторы значительно увеличивает эффективность получения стабильных клеточных клонов с высоким уровнем экспрессии репортерного гена [43]. Интересно, что противоположные результаты получены в аналогичных опытах с ремодулирующим комплексом Brahma, обеспечивающим повышение подвижности нуклеосом и способным позитивно/негативно регулировать транскрипцию в зависимости от конкретного гена.

В некоторых случаях с целью усиления транскрипции репортерного гена и снижения зависимости экспрессии трансгена от окружающего хроматина используются сильные клеточные энхансеры [44, 45]. Один из наиболее часто применяемых энхансеров – LCR (Locus Control Region), который контролирует экспрессию генов β -глобинового локуса человека [46]. Основным недостатком использования энхансеров для усиления уровня экспрессии трансгена обусловлен их специфичностью, т.е. функционированием только в определенных клеточных линиях, что накладывает определенные ограничения на их исполь-

зование в качестве универсальных регуляторных элементов. В этом направлении перспективен поиск энхансеров, наиболее эффективных в клеточных линиях, применяемых для наработки белков в промышленных масштабах [47].

ПЕРСПЕКТИВЫ ИНСУЛЯТОРОВ В ПОВЫШЕНИИ ЭФФЕКТИВНОСТИ ЭКСПРЕССИИ ЦЕЛЕВОГО ГЕНА В КЛЕТОЧНЫХ ЛИНИЯХ-ПРОДУЦЕНТАХ

С целью увеличения эффективности и стабильности экспрессии целевого белка используют известные инсуляторы [41, 48–51]. Инсуляторами называются регуляторные элементы, которые блокируют взаимодействие между энхансером и промотором, если находятся между ними [52, 53]. При этом инсуляторы не влияют непосредственно на активность энхансера и промотора, т.е. промотор может активироваться любым другим энхансером, а энхансер способен активировать любой другой промотор. Кроме того, часть инсуляторов может служить границей между транскрипционно активным хроматином и гетерохроматином. Наиболее хорошо изучены инсуляторы плодовой мушки дрозофилы и позвоночных. Исходно предполагали, что инсуляторы определяют границы транскрипционных доменов, внутри которых экспрессия генов не зависит от негативных эффектов окружающего генома [54, 55]. Однако в дальнейшем было показано, что инсуляторные белки значительно более гибко интегрированы в регуляторные системы генов [52, 53].

Недавно было обнаружено, что некоторые инсуляторные белки участвуют в организации специфичных дистанционных взаимодействий между удаленными участками хроматина [56–59]. Инсуляторные белки могут поддерживать взаимодействия между энхансерами и промоторами, границами транскрипционных доменов, которые часто располагаются на расстояниях, достигающих сотен тысяч пар нуклеотидов [34, 60–62]. Полученные результаты дали основания отнести инсуляторные белки к классу архитектурных белков хроматина [32, 53].

К настоящему моменту наиболее хорошо описаны инсуляторные/архитектурные белки (ИАБ) дрозофилы, что во многом связано с простотой получения трансгенных линий этих мух. Исследование свойств инсуляторов в трансгенных линиях дрозофилы показало, что с каждым инсулятором связывается несколько ИАБ, которые и определяют специфичность дистанционных взаимодействий [32, 53]. В результате два идентичных инсулятора способны достаточно специфично обеспечивать сверхдальние взаимодействия между регуляторными элементами в трансгенных линиях дрозофилы [63, 64], что позволило предложить модель, согласно которой ИАБ, специфичные

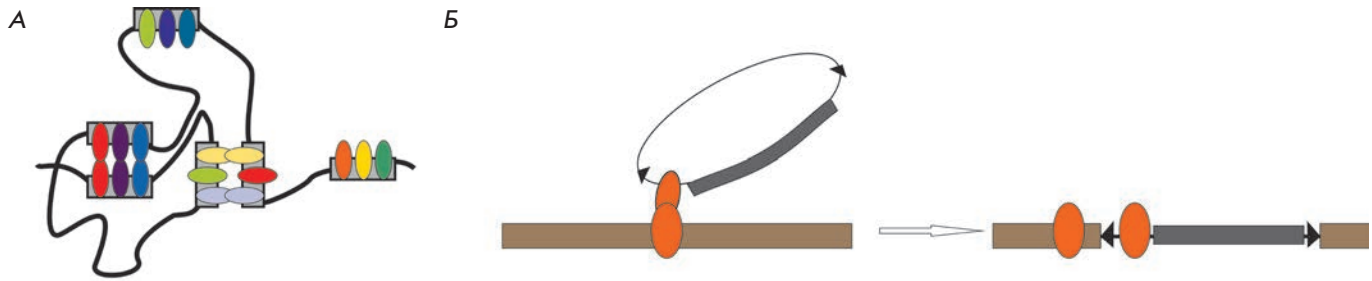


Рис. 1. А – модель установления специфичных дистанционных взаимодействий между регуляторными элементами. С каждым элементом связывается несколько ИАБ (инсуляторные архитектурные белки). В результате два одинаковых элемента способны высокоспецифично и эффективно обеспечивать сверхдальние взаимодействия между регуляторными элементами, при частичном совпадении сайтов связывания ИАБ в составе регуляторных элементов формируются более слабые контакты. Серые прямоугольники – регуляторные элементы, с которыми связываются ИАБ; разноцветные овалы – комбинации ИАБ, специфично связывающиеся со своими сайтами и взаимодействующие друг с другом. Б – специфичная интеграция трансгенной конструкции в определенное место генома. Трансгенная конструкция представлена в виде кольца. Черные треугольники – концы Р-транспозона, серый прямоугольник – элементы в составе трансгенной конструкции. Оранжевые овалы – регуляторный элемент, с которым связывается ИАБ. Коричневый прямоугольник – участок генома

сайты связывания которых входят в состав регуляторных элементов, создают код, определяющий, насколько эффективным будет установленное между данными регуляторными элементами дистанционное взаимодействие [32] (рис. 1А).

Обычно трансгенные линии дрозофилы получают путем инъекции вектора, содержащего концы Р-транспозона, которые фланкируют трансгенную конструкцию, вместе с геном, кодирующим транспозазу, необходимую для интеграции конструкции в геном [65]. Описаны случаи высокоспецифичной интеграции Р-транспозона, в состав которого входит инсулятор или сильный промотор, в определенный район генома, содержащий аналогичный эндогенный регуляторный элемент [66, 67]. Подобную высокоспецифичную интеграцию Р-транспозона, названную «возвращением домой» (homing), можно объяснить тем, что после введения конструкции в эмбрион с регуляторным элементом в составе Р-транспозона связывается ИАБ. В результате обеспечивается специфичное взаимодействие Р-транспозона с аналогичным эндогенным регуляторным элементом, и происходит интеграция транспозона в определенный район хромосомы (рис. 1Б).

Полногеномные исследования распределения сайтов связывания ИАБ однозначно показали, что инсуляторы не являются стационарными границами между транскрипционными доменами [62, 68]. В трансгенных линиях дрозофилы, клеточных линиях и в системе *in vitro* описаны два механизма, с помощью которых инсуляторы способны блокиро-

вать активность энхансеров [53]. Первый механизм основан на наведении топологических препятствий, блокирующих взаимодействия между энхансером и промотором. В результате образования стабильных контактов между инсуляторами один из взаимодействующих регуляторных элементов оказывается изолированным в независимой хроматиновой петле. Такой механизм действия инсуляторов обнаружили с помощью искусственно созданных инсуляторов, состоящих из сайтов связывания белков, способных эффективно взаимодействовать друг с другом и образовывать стабильные хроматиновые петли [69, 70], и в трансгенных линиях дрозофилы [71]. Этот механизм эффективно проявляется только в том случае, когда инсуляторы непосредственно соседствуют с блокируемыми элементами (энхансерами и промоторами). При увеличении размеров хроматиновой петли, формируемой инсуляторами, блокирование энхансера по такому механизму не происходит [72]. Второй механизм, более общий, основан на установлении прямых контактов между белками, связанными с инсулятором, и энхансер-промоторным комплексом. Например, показано, что в трансгенных линиях дрозофилы инсуляторы могут прямо взаимодействовать с промоторами и энхансерами [71, 73]. В данной ситуации, когда инсулятор находится между энхансером и промотором, инсуляторные белки мешают установлению правильных контактов между транскрипционными комплексами, собранными на энхансере и промоторе, что приводит к частичной или полной неспособности энхансера стимулировать транскрипцию с промотора.

Механизмы, определяющие барьерную функцию инсуляторов, описаны достаточно подробно у дрозофилы и млекопитающих. В частности, установлено, что ИАБ помогают рекрутировать на инсуляторы белковые комплексы, которые отвечают за ремоделирование и модификации нуклеосом [74–78], что в результате приводит к созданию зон открытого хроматина. Одновременно некоторые инсуляторы могут привлекать белковые комплексы, непосредственно принимающие участие в стимуляции транскрипции [75]. За счет создания свободных от нуклеосом участков ДНК и рекрутирования транскрипционных комплексов инсуляторы блокируют распространение репрессивного хроматина, что, тем не менее, не исключает возможности прямого взаимодействия инсуляторов с сайленсерами, иницирующими гетерохроматинизацию участков генома.

К настоящему времени описан только один инсуляторный ДНК-связывающий белок позвоночных, CTCF [79, 80], что, вероятно, связано с отсутствием удобных модельных систем для изучения инсуляторов. CTCF способен поддерживать дистанционные взаимодействия между удаленными участками хроматина [60, 79–81]. Таким образом, CTCF – первый охарактеризованный архитектурный белок в геноме млекопитающих [32, 80].

Несмотря на ключевую роль в формировании архитектуры хроматина, домены белка CTCF остаются слабо изученными, и механизм функционирования CTCF в процессе поддержания дистанционных взаимодействий до сих пор не охарактеризован. В центральной части белка расположены 11 цинковых пальцев (ZF) типа C2H2, из них только четыре (с четвертого по седьмой) необходимы для узнавания корового мотива ДНК [82]. Предполагается, что остальные цинковые пальцы также способны узнавать определенные нуклеотидные последовательности, стабилизируя тем самым связывание CTCF с ДНК. Наиболее логично предположить, что белок, поддерживающий дистанционные взаимодействия, способен к эффективной ди- или мультимеризации. Действительно, показано, что CTCF способен гомодимеризоваться, однако домен, ответственный за эту активность, не был локализован [83]. Также получены данные, согласно которым С-концевой домен белка CTCF прямо взаимодействует со своими цинковыми пальцами [84]. Однако такое взаимодействие не может быть высокоспецифичным, так как цинковые пальцы CTCF связываются со многими факторами транскрипции – CHD8, Sin3A, УВ-1 [85–87].

Предполагается, что существенную роль в организации дистанционных взаимодействий играет когезиновый комплекс [60, 80, 88, 89], взаимодействующий непосредственно с CTCF [90]. Когезиновый

комплекс рекрутируется на хроматин с помощью CTCF и способствует установлению дистанционных взаимодействий между геномными CTCF-сайтами. Эта модель согласуется с данными полногеномных исследований, показывающих высокую степень колокализации CTCF и субъединиц когезинов [91, 92]. Однако в экспериментах по инактивации CTCF наблюдается совсем незначительное снижение связывания когезинов с хроматином, что предполагает участие и других факторов транскрипции в рекрутировании когезинового комплекса на хроматин [92–94]. Более того, инактивация CTCF и когезинов приводит к разным нарушениям архитектуры хроматина [95, 96], что можно объяснить независимым функционированием этих белков.

В биотехнологии используют наиболее хорошо изученный инсулятор позвоночных, HS4, состоящий из 1200 п.н. и расположенный на 5'-конце β-глобинового локуса кур [97, 98] (рис. 2). В этом инсуляторе выявлен коровый участок из 250 п.н., обладающий полной инсуляторной активностью и содержащий пять фрагментов (FI, FII, FIII, FIV, FV), каждый из которых имеет функциональное значение. В участке FII идентифицирован сайт связывания белка CTCF, необходимого и достаточного для проявления энхансерблокирующей активности HS4 [99]. За способность создавать границу между активным хроматином и гетерохроматином отвечают белки USF1 и USF2, которые связываются в виде гетеродимера с районом FIV [100]. Показано, что USF рекрутирует белковые комплексы, которые отвечают за модификации гистонов, ассоциированных со стимуляцией транскрипции [100, 101]. С другими районами инсулятора HS4 (FI, FIII, FV) связывается белок BGP1/Vezf1, который имеет ДНК-связывающий домен, состоящий из цинковых пальцев [102]. Белок BGP1/Vezf1 защищает GC-богатые участки инсулятора от метилирования, которое нарушает связывание инсуляторных белков с ДНК и, как следствие, приводит к инактивации инсулятора. Согласно существующей модели BGP1/Vezf1 терминирует слабую транскрипцию из гетерохроматинового района, что может играть важную роль в защите β-глобинового локуса от распространения неактивного хроматина [103].

Инсулятор HS4 с момента его открытия активно использовали для защиты экспрессии трансгена в культурах клеток млекопитающих [98]. Две полные копии HS4 встроили в вектор для получения трансгенных животных, в молоке которых экспрессируется требуемый белок [104]. Показано, что инсулятор значительно увеличивал экспрессию целевых белков, но не влиял существенно на специфичность экспрессии трансгена только в молочной железе

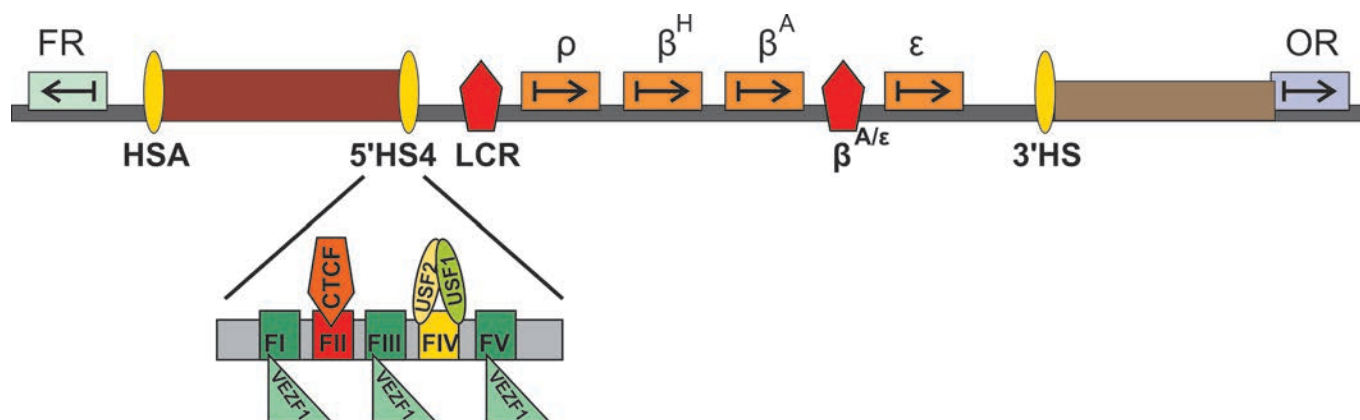


Рис. 2. Схема β-глобинового локуса и прилегающих последовательностей. ρ, β^H, β^A, ε соответствуют генам β-глобинового локуса; FR – ген фолатного рецептора; OR – ген обонятельного рецептора. Стрелками показано направление транскрипции генов. HSA, 5'HS4, 3'HS – инсультаторы; LCR, β^{A/ε} – энхансеры β-глобинового локуса. Детально представлена схема инсультатора HS4. FI, FIII, FV – участки связывания белка Vezf1/BGP1. FII и FIV – сайты связывания белка CTCF и гетеродимера USF1/2 соответственно

и не поддерживал прямую корреляцию между количеством копий конструкции и уровнем наработки целевого белка [105].

Наиболее эффективно инсультатор HS4 можно применять в векторах, размер которых по каким-либо причинам ограничен. Так, при использовании полного инсультатора размером 1.2 т.п.н. значительно снижается эффективность трансформации клеток лентивирусными векторами (возможно, из-за ограничений, накладываемых на размер вирусной частицы). Поэтому в векторах этого класса используют дублированный коровый элемент HS4 размером 250 п.н., который содержит сайты связывания всех идентифицированных факторов транскрипции, необходимых для проявления активности инсультатора [98, 106]. Инсультаторы также с успехом использовали для защиты экспрессии репортерного гена в векторах, созданных на основе мобильных элементов [51] и ретровирусов [107].

Несмотря на примеры успешного использования 1.2 т.п.н. инсультатора HS4 или его корового участка [98, 108], получены многочисленные данные, согласно которым HS4 не оказывал положительного воздействия на экспрессию целевого гена. Это можно объяснить тем, что используемые клеточные культуры значительно различаются по набору факторов транскрипции, способных связываться с инсультатором HS4.

В заключение можно постулировать, что основными механизмами защиты экспрессии трансгена с помощью инсультатора HS4 являются: 1) создание свободного от нуклеосом участка ДНК, который может прерывать линейное распространение гетерохроматина; 2) рекрутирование белковых комплексов, которые увеличивают подвижность нуклеосом, мо-

дифицируют гистоны, стимулируют транскрипцию, защищают от метилирования CpG-участки и терминируют слабую транскрипцию. Не установлено, способен ли инсультатор HS4 направлять интеграцию конструкции в зоны транскрипционно активного хроматина и непосредственно взаимодействовать с промотором целевого гена для последующей стимуляции транскрипции. Очевидно, что основным недостатком HS4 и других инсультаторов является зависимость их активности от набора конкретных факторов транскрипции, которые экспрессируются в данной клеточной линии.

ПОВЫШЕНИЕ УРОВНЯ ЭКСПРЕССИИ ТРАНСГЕНА С ПОМОЩЬЮ А/Т-БОГАТЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ, АССОЦИИРОВАННЫХ С БЕЛКАМИ ЯДЕРНОГО МАТРИКСА (S/MAR)

С целью увеличения эффективности трансформации и повышения стабильности экспрессии трансгена [108–110] с начала 1990-х годов широко используются последовательности (300–5000 п.н.), обычно А/Т-богатые, которые, как показано в опытах *in vitro*, взаимодействуют с фракцией ядерного матрикса (S/MAR, scaffold matrix attachment region). S/MAR обладают рядом характерных свойств: это А/Т-богатые, чувствительные к ДНКазе I участки, потенциально склонные к формированию левозакрученной спирали и триплексных структур [111, 112]. Предполагается, что именно А/Т-богатый состав данных элементов приводит к дестабилизации двойной спирали и способности MAR формировать участки, насыщенные различными вторичными структурами [113, 114].

Исходя из особенностей вторичной структуры, в геноме человека было предсказано порядка 50000

элементов, предположительно обладающих свойствами S/MAR [115, 116]. Из этого пула участков отобрали 1500, максимально соответствующих особенностям S/MAR. Оказалось, что только несколько из них обладают высокой степенью гомологии с мышинными ортологами, т.е. нуклеотидная последовательность в участках S/MAR не обладает выраженными консервативными элементами.

Структура S/MAR-элементов предполагает, что они могут быть горячими точками рекомбинации. Действительно, показано, что места разрывов при инверсиях, связанных с болезнями человека, часто находятся в области S/MAR-элементов [117, 118], а интеграция ретровирусов в геном с высокой частотой происходит в непосредственной близости от S/MAR [119, 120]. Согласно некоторым данным S/MAR участвуют в регуляции репликации ДНК [121–123]. Установлено, что S/MAR-элементы увеличивают экспрессию трансгена и уменьшают вариабельность экспрессии в процессе получения стабильных клеточных линий [41, 124]. Экспериментально показано, что экспрессия гена, окруженного S/MAR-элементами, пропорциональна числу копий гена [125]. Предположили, что функционально S/MAR-элементы можно рассматривать как инсульты, защищающие экспрессию трансгена от позитивных/негативных эффектов окружающего хроматина.

Исходно считалось, что основным компонентом белков ядерного матрикса являются ламинины [126]. Впоследствии в составе ядерного матрикса обнаружили много дополнительных белков, в том числе факторов транскрипции [127]. В S/MAR часто находятся сайты связывания транскрипционных факторов SATB1, Fast1, СЕВР, SAF-A и SAF-B (белки, которые предпочтительно связываются с А/Т-богатыми участками), NMP4 (белок матрикса), CTCF и белков Нох-семейства [83, 116, 128–131]. Топоизомераза II также преимущественно связывается с А/Т-богатыми участками S/MAR [132–134]. Снижение плотности распределения нуклеосом на S/MAR-элементах и повышение концентрации комплексов, ацетилирующих гистоны, объясняется связыванием многочисленных транскрипционных факторов и способностью этих элементов формировать вторичные структуры.

SATB1 – наиболее хорошо изученный белок матрикса, вовлеченный во многие биологические процессы, такие, как дифференцировка Т-клеток и эпидермиса [135–137]. SATB1 можно отнести к классу архитектурных белков, способных поддерживать специфичные дистанционные взаимодействия [135]. SATB1 образует гомодимеры и связывается с А/Т-богатыми последовательностями с помощью двух CUT-доменов и С-концевого гомеобокса. Помимо

участия в формировании хроматиновых доменов, SATB1 рекрутирует ASF1 (АТФ-зависимый фактор, участвующий в организации хроматина) и комплекс ISWI (повышает подвижность нуклеосом) [128, 138].

Другой белок матрикса SAF-A содержит ДНК-связывающий (SAF) и РНК-связывающий (RGG) домены. Интересно, что РНК Xist, регулирующая дозовую компенсацию у млекопитающих, связывается с RGG-доменом. Это взаимодействие определяет локализацию РНК Xist на X-хромосоме [139]. Согласно существующей модели [140] SAF-A рекрутирует РНК Xist на S/MAR-элемент, расположенный в районе участка инициации формирования гетерохроматина на X-хромосоме. В результате взаимодействий белков SAF-A и SATB1 в дальнейшем формируются петли между соседними S/MAR-комплексами, что в итоге приводит к распространению РНК Xist по X-хромосоме и ее последующей инактивации.

Согласно наиболее распространенной модели S/MAR-элементы взаимодействуют с белками ядерного скелета (матриксными белками), что приводит к формированию хроматиновых петель, основанием которых являются S/MAR [141]. Предполагается, что гены, находящиеся внутри хроматиновой петли, формируемой S/MAR, защищены от негативного эффекта окружающего хроматина [109]. Однако до настоящего момента непонятна структура самого ядерного матрикса и роль S/MAR в организации архитектуры хромосом. Согласно современным представлениям матрикс представляет собой лабильные конгломераты белков, с которыми транзитно взаимодействуют белковые комплексы S/MAR, входящие в состав хромосом [141].

Несмотря на отсутствие понимания механизмов действия S/MAR, получены многочисленные экспериментальные данные, показавшие эффективность использования этих элементов для увеличения уровня экспрессии целевых белков в культуре клеток млекопитающих [142]. Например, S/MAR из гена лизоцима куриного яйца в 5–10 раз увеличивают уровень экспрессии моноклональных антител в клетках CHO [109, 143]. В дальнейшем были охарактеризованы более эффективные S/MAR млекопитающих [144]. S/MAR успешно использовали для увеличения уровня экспрессии эритропоэтина, человеческого рогового фактора, рецептора II типа TGF-β [108]. При этом оказалось, что S/MAR функционируют как в вирусных векторах [145], так и векторах, созданных на основе мобильных элементов [146]. S/MAR эффективно защищают экспрессию трансгена от репрессии (барьерная активность), а также поддерживают более высокий уровень транскрипции с промоторов в составе экспрессирующих векторов (стимулирующая активность) [144].

Некоторые S/MAR способны повышать эффективность интеграции векторной конструкции в хромосомы [147, 148]. Более того, S/MAR могут обеспечивать интеграцию новых дополнительных копий векторных конструкций в район, в котором уже находятся встроенные копии конструкции. Наиболее вероятным объяснением этого свойства S/MAR может быть то, что взаимодействие между архитектурными белками (такими, как SAF-A и SATB1) обеспечивает контакт между копиями S/MAR, которые находятся в геноме и плазмиде (аналог описанного выше феномена homing-эффекта у дрозофилы). Вероятно, что повышенная рекомбинационная активность, обеспечиваемая S/MAR-элементами, усиливает эффективность интеграции трансгена в определенные участки генома. Таким образом, позитивное влияние S/MAR на экспрессию трансгена может в значительной степени определяться направленной интеграцией векторной конструкции в S/MAR-содержащие районы, обычно транскрипционно активные. Такой вывод согласуется с наблюдением, согласно которому многие исследованные S/MAR позитивно влияют на экспрессию генов только при интеграции конструкции в геном [125, 144]. Для увеличения степени амплификации конструкции в клеточных клонах S/MAR объединили с районом инициации репликации у млекопитающих [149, 150]. Обработка первичных трансфектантов MTX позволила добиться масштабной амплификации конструкций в клеточных клонах, что приводило к многократному стабильному увеличению продукции целевого белка [150].

В заключение можно сделать вывод, что S/MAR – это регуляторные элементы, менее изученные по сравнению с инсуляторами. Наиболее вероятными механизмами действия S/MAR в процессе повышения уровня экспрессии трансгена являются: 1) направленная интеграция S/MAR-содержащих конструкций в районы транскрипционно активного хроматина и увеличение числа копий интегрированной конструкции; 2) связывание с элементами комплексов S/MAR, которые наводят зоны транскрипционно активного хроматина и тем самым блокируют распространение гетерохроматина; 3) прямая активация промотора транскрипционными факторами, которые непосредственно связываются с S/MAR.

ПОВЫШЕНИЕ УРОВНЯ ЭКСПРЕССИИ ТРАНСГЕНА В КУЛЬТУРАХ КЛЕТОК С ПОМОЩЬЮ РЕГУЛЯТОРНЫХ ЭЛЕМЕНТОВ, СОДЕРЖАЩИХ СИЛЬНЫЕ ПРОМОТОРЫ ГЕНОВ «ДОМАШНЕГО ХОЗЯЙСТВА»

В 2000–2002 годах небольшая компания Cobra Therapeutics разработала технологическую платформу для создания эффективных клеточных линий-продуцентов рекомбинантных белков на основе

регуляторных элементов, выделенных из промоторов генов «домашнего хозяйства», которые активно работают на всех стадиях развития и во всех клетках организма [45]. Такие регуляторные элементы названы «элементами с открытым хроматином» (ubiquitous chromatin opening elements, UCOE), так как для промоторных областей активно транскрибирующихся генов характерна низкая плотность нуклеосом, связанная непосредственно с присутствием ДНК-связывающих ТФ, стимулирующих транскрипцию. К наиболее хорошо охарактеризованному UCOE относятся области ДНК, содержащие разнонаправленные пары промоторов генов *HNRPA2B1* и *CBX3* или *TBP* и *PSNB1*, которые интенсивно транскрибируются во всех клетках организма [151]. В первых экспериментах использовали большие регуляторные элементы размером 12–16 т.п.н., которые значительно увеличивали процент трансфицированных клеток и обеспечивали высокую и стабильную экспрессию трансгена на протяжении длительного времени культивирования клеточных клонов [151, 152]. Так, UCOE в 16 раз усиливают эффективность работы промотора CMV, часто подверженного ингибированию путем РНК-интерференции и метилирования CpG-районов [152, 153]. Показано, что UCOE могут поддерживать высокий уровень экспрессии трансгена, встроенного в прицентромерный гетерохроматин. UCOE эффективны в составе лентивирусных векторов [154–159]. Можно предположить, что с UCOE связываются транскрипционные факторы, привлекающие комплексы, которые предотвращают метилирование CpG-повторов и создают участок хроматина с пониженной плотностью нуклеосом в области промоторов в составе лентивирусных векторов [155, 160].

Нужно отметить, что промоторы генов «домашнего хозяйства» способны эффективно функционировать в различных клеточных линиях, в отличие от других регуляторных элементов, таких, как LCR и энхансеры, которые проявляют выраженную клеточную специфичность. В результате экспериментов по уменьшению размеров UCOE, необходимых для оценки возможности использования UCOE в экспрессионных векторах, получены более компактные варианты UCOE размером 1.5–3 т.п.н. Такие укороченные элементы полностью сохраняют свою активность в процессе получения высокопродуктивных клеточных линий [152].

UCOE – активный участник процесса регуляции транскрипции, что подразумевает существование прямых взаимодействий между регуляторными элементами промотора, ответственного за экспрессию репортерного гена в составе вектора, и транскрипционных факторов, связанных с UCOE. Вследствие этого UCOE может эффективно работать только с опре-

деленными промоторами, и их функциональная активность не распространяется на все клеточные линии [161]. В некоторых работах показано, что сами UCOE могут использоваться в качестве промоторов для обеспечения стабильной экспрессии репортерного гена [162]. Однако непонятна роль и неоднозначен вклад в экспрессию трансгена транскрипции, индуцируемой с промоторов, входящих в состав UCOE. В частности, существуют экспериментальные данные, согласно которым UCOE не всегда эффективно усиливают уровень экспрессии целевого белка в клетках CHO [162, 163]. Негативные результаты, получаемые при использовании UCOE, можно объяснить тем, что сильные промоторы, входящие в состав UCOE, индуцируют транскрипцию, способную в ряде случаев инициировать РНК-интерференцию и/или рекрутирование репрессивных комплексов на промотор, с которого транскрибируется репортерный ген.

Показано также, что совмещение двух сильных промоторов может в некоторых случаях способствовать получению стабильных клеточных линий и увеличивать уровень экспрессии трансгена [164]. Анализ различных комбинаций двух промоторов: CMV, SV40, RPL32, EF1 α и β -актина, показал, что только промотор гена RPL32, встроенный перед любым из проверяемых промоторов, способен значительно увеличить эффективность отбора стабильных клеточных клонов. Интересно, что направление промотора RPL32 должно совпадать с направлением промотора, ответственного за экспрессию репортерного гена, и коровые элементы промотора RPL32 в данном случае являются необходимыми компонентами системы для обеспечения стимулирующего эффекта.

В целом, комбинация сильных промоторов – перспективный способ увеличения эффективности получения клеточных клонов-продуцентов целевых белков. Сильные промоторы, состоящие из комбинации энхансера и корового промотора, рекрутируют белковые комплексы, которые, в свою очередь, поддерживают транскрипционно активное состояние хроматина. Согласно данным полногеномных исследований промоторы служат эффективными границами, способными обеспечивать защиту районов генома от распространения зон репрессивного хроматина [165, 166]. Транскрипционные факторы, которые связываются с правильно подобранными парами промоторов, могут взаимно усиливать активность друг друга. Вероятно, использование некоторых промоторов может приводить к преимущественной интеграции конструкции в определенные места хромосом, в которых уровень транскрипции достигает наиболее высоких уровней. Более полное понимание механизмов активации транскрипции позволит в дальнейшем

модифицировать промоторы с целью увеличения их эффективности при работе с экспрессионными системами.

ПОВЫШЕНИЕ УРОВНЯ ЭКСПРЕССИИ ТРАНСГЕНА В КУЛЬТУРАХ КЛЕТОК С ПОМОЩЬЮ РЕГУЛЯТОРНЫХ ЭЛЕМЕНТОВ, ЗАЩИЩАЮЩИХ ОТ HP1-ЗАВИСИМОЙ РЕПРЕССИИ

Компания Chromagenic разработала технологическую платформу, основанную на тест-системе, которая позволяет эффективно идентифицировать регуляторные элементы, обладающие способностью блокировать распространение зон гетерохроматина [45]. Тест-система основана на рекрутировании белка HP1, ответственного за формирование гетерохроматина, на плазмиду с помощью ДНК-связывающего домена белка Lex [167]. Химерный белок HP1-Lex связывается с lex-сайтами на плазмиде и привлекает другие компоненты гетерохроматинового комплекса, что и приводит к запуску инактивации расположенного рядом промотора. В результате происходит репрессия гена *zeo^R*, ответственного за устойчивость к антибиотику зеоцину (Zeocin), и гибель трансфицированных клеток при культивировании на селективной среде с добавлением зеоцина. Скрининг, направленный на поиск фрагментов ДНК, встраивание которых между сайтами связывания Lex и промотором гена *zeo^R* обеспечивает защиту промотора от HP1-зависимой репрессии, позволил найти серию регуляторных элементов размером от 500 до 2000 п.н., названных антирепрессорами (STAR). Другие известные регуляторные элементы, инсулятор HS4, MAR и UCOE, не обладают такой способностью. Сравнительный анализ различных регуляторных элементов [168] показал, что STAR-элементы наиболее эффективны при создании высокопродуктивных клеточных линий CHO. Однако механизм действия STAR-элементов остается неизученным. До сих пор не установлено, какие транскрипционные факторы связываются с элементами данного класса и обеспечивают их функциональную активность.

ЗАКЛЮЧЕНИЕ

К настоящему времени не найдены универсальные регуляторные элементы с понятным механизмом действия, которые можно эффективно использовать в составе векторных конструкций всех типов, предназначенных для получения высокопродуктивных клеточных линий-продуцентов различных белков. Во многом это связано со сложностью механизмов, регулирующих активность промоторов, а также с отсутствием фактических подтверждений первоначальных представлений о строгой организации генов с одинаковым профилем экспрессии в транс-

крипционные домены, ограниченные особыми регуляторными элементами класса инсультаторов или S/MAR. Очевидно, что даже для самых сильных промоторов должны существовать механизмы супрессии «чрезмерной» транскрипции. Один из таких механизмов – РНК-интерференция. Возможно, детальное понимание механизмов активации и супрессии транскрипции позволит создать искусственные про-

моторы, с помощью которых реальным станет получение стабильно высоких уровней экспрессии целевого гена в трансгенных системах. ●

Работа поддержана Российским научным фондом (грант № 14-24-00166).

СПИСОК ЛИТЕРАТУРЫ

1. Demain A.L., Vaishnav P. // *Biotech. Advances*. 2009. V. 27. P. 297–306.
2. Durocher Y., Butler M. // *Curr. Opin. Biotech.* 2009. V. 20. P. 700–707.
3. Khan K.H. // *Adv. Pharm. Bull.* 2013. V. 3. P. 257–263.
4. Kim J.Y., Kim Y.G., Lee G.M. // *Appl. Microbiol. Biotechnol.* 2012. V. 93. P. 917–930.
5. Hacker D.L., De Jesus M., Wurm F.M. // *Biotechnol. Adv.* 2009. P. 1023–1027.
6. Tjio J.H., Puck T.T. // *J. Exp. Med.* 1958. V. 108. P. 259–271.
7. Kao F.T., Puck T.T. // *Proc. Natl. Acad. Sci. USA*. 1968. V. 60. P. 1275–1281.
8. Browne S.M., Al-Rubeai M. // *Trends Biotech.* 2007. V. 25. P. 425–432.
9. Folger K.R., Wong E.A., Wahl G., Capecchi M.R. // *Mol. Cell Biol.* 1982. V. 2. P. 1372–1387.
10. Folger K.R., Thomas K., Capecchi M.R. // *Mol. Cell Biol.* 1985. V. 5. P. 59–69.
11. Robins D.M., Ripley S., Henderson A.S., Axel R. // *Cell*. 1981. V. 23. P. 29–39.
12. Thomas K.R., Folger K.R., Capecchi M.R. // *Cell*. 1986. V. 44. P. 419–428.
13. Hayward B.E., Hussain A., Wilson R.H., Lyons A., Woodcock V., McIntosh B., Harris T.J. // *Nucl. Acids Res.* 1986. V. 14. P. 999–1008.
14. Wuest D.M., Harcum S.W., Lee K.H. // *Biotechnol. Adv.* 2012. V. 30. P. 629–638.
15. Urlaub G., Chasin L.A. // *Proc. Natl. Acad. Sci. USA*. 1980. V. 77. P. 4216–4220.
16. Urlaub G., Kas E., Carothers A.M., Chasin L.A. // *Cell*. 1983. V. 33. P. 405–412.
17. Inouye S., Tsuji F.I. // *FEBS Lett.* 1994. V. 341. P. 277–280.
18. Stepanenko O.V., Verkhusha V.V., Kuznetsova I.M., Uversky V.N., Turoverov K.K. // *Curr. Protein Pept. Sci.* 2008. V. 9. P. 338–369.
19. Niu J., Zhang B., Chen H. // *Mol. Biotechnol.* 2014. V. 56. P. 681–688.
20. Wijshake T., Baker D.J., van de Sluis B. // *Biochim. Biophys. Acta*. 2014. V. 1842. P. 1942–1950.
21. Castel S.E., Martienssen R.A. // *Nat. Rev. Genet.* 2013. V. 14. P. 100–112.
22. Fatica A., Bozzoni I. // *Nat. Rev. Genet.* 2014. V. 15. P. 7–21.
23. Morris K.V., Mattick J.S. // *Nat. Rev. Genet.* 2014. V. 15. P. 423–437.
24. Lai T., Yang Y., Ng S.K. // *Pharmaceuticals (Basel)*. 2013. V. 6. P. 579–603.
25. Bradley A.J., Lurain N.S., Ghazal P., Trivedi U., Cunningham C., Baluchova K., et al. // *J. Gen. Virol.* 2009. V. 90. P. 2375–2380.
26. Goldberg M.L. *Sequence Analysis of Drosophila Histone Genes*: PhD thesis. Stanford University, 1979.
27. Corden J., Wasylyk B., Buchwalder A., Sassone-Corsi P., Kedinger C., Chambon P. // *Science*. 1980. V. 209. P. 1406–1414.
28. Burke T.W., Kadonaga J.T. // *Genes Dev.* 1996. V. 10. P. 711–724.
29. Lagrange T., Kapanidis A.N., Tang H., Reinberg D., Ebricht R.H. // *Genes Dev.* 1998. V. 12. P. 34–44.
30. Lewis B.A., Kim T.K., Orkin S.H. // *Proc. Natl. Acad. Sci. USA*. 2000. V. 97. P. 7172–7177.
31. Lim C.Y., Santoso B., Boulay T., Dong E., Ohler U., Kadonaga J.T. // *Genes Dev.* 2004. V. 18. P. 1606–1617.
32. Maksimenko O., Georgiev P. // *Front. Genet.* 2014. V. 5. P. 28.
33. Xu X., Nagarajan H., Lewis N.E., Pan S., Cai Z., Liu X., Chen W., Xie M., Wang W., Hammond S., et al. // *Nat. Biotechnol.* 2011. V. 29. P. 735–741.
34. Becker J., Hackl M., Rupp O., Jakobi T., Schneider J., Szczepanowski R., Bekel T., Borth N., Goesmann A., Grillari J., et al. // *J. Biotechnol.* 2011. V. 156. P. 227–235.
35. Jakobi T., Brinkrolf K., Tauch A., Noll T., Stoye J., Pühler A., Goesmann A. // *J. Biotechnol.* 2014. V. 190. P. 64–75.
36. Running Deer J., Allison D.S. // *Biotechnol. Prog.* 2004. V. 20. P. 880–889.
37. Orlova N.A., Kovnir S.V., Hodak J.A., Vorobiev I.I., Gabibov A.G., Skryabin K.G. // *BMC Biotechnol.* 2014. V. 14. P. 56.
38. Mariati, Koh E.Y., Yeo J.H., Ho S.C., Yang Y. // *Bioengineered*. 2014. V. 5. P. 340–345.
39. Andersen C.R., Nielsen L.S., Baer A., Tolstrup A.B., Weilguny D. // *Mol. Biotechnol.* 2011. V. 48. P. 128–137.
40. de Groote M.L., Verschure P.J., Rots M.G. // *Nucl. Acids Res.* 2012. V. 40. P. 10596–10613.
41. Kwaks T.H., Otte A.P. // *Trends Biotechnol.* 2006. V. 24. P. 137–142.
42. Holmqvist P.H., Mannervik M. // *Transcription*. 2013. V. 4. P. 18–23.
43. Kwaks T.H., Sewalt R.G., van Blokland R., Siersma T.J., Kasiem M., Kelder A., Otte A.P. // *J. Biotechnol.* 2005. V. 115. P. 35–46.
44. Moltó E., Fernández A., Montoliu L. // *Brief Funct. Genomic Proteomic*. 2009. V. 8. P. 283–296.
45. Palazzoli F., Bire S., Bigot Y., Bonnin-Rouleux F. // *Nat. Biotechnol.* 2011. V. 29. P. 593–597.
46. Kim A., Dean A. // *Mol. Cells*. 2012. V. 34. P. 1–5.
47. Vishwanathan N., Le H., Le T., Hu W.S. // *Curr. Opin. Biotechnol.* 2014. V. 30. P. 113–119.
48. Recillas-Targa F., Valadez-Graham V., Farrell C.M. // *BioEssays*. 2004. V. 26. P. 796–807.
49. Emery D.W. // *Hum. Gene Ther.* 2011. V. 22. P. 761–774.
50. Furlan-Magaril M., Rebollar E., Guerrero G., Fernández A., Moltó E., González-Buendía E., Cantero M., Montoliu L.,

- Recillas-Targa F. // *Nucl. Acids Res.* 2011. V. 39. P. 89–103.
51. Bire S., Ley D., Casteret S., Mermod N., Bigot Y., Rouleux-Bonnin F. // *PLoS One.* 2013. V. 8. e82559.
52. Chetverina D., Aoki T., Erokhin M., Georgiev P., Schedl P. // *BioEssays.* 2014. V. 36. P. 163–172.
53. Kyrchanova O., Georgiev P. // *FEBS Lett.* 2014. V. 588. P. 8–14.
54. Labrador M., Corces V.G. // *Cell.* 2002. V. 111. P. 151–154.
55. Bell A.C., West A.G., Felsenfeld G. // *Science.* 2001. V. 291. P. 447–450.
56. Handoko L., Xu H., Li G., Ngan C.Y., Chew E., Schnapp M., Lee C.W., Ye C., Ping J.L., Mulawadi F., et al. // *Nat. Genet.* 2011. V. 43. P. 630–638.
57. Dixon J.R., Selvaraj S., Yue F., Kim A., Li Y., Shen Y., Hu M., Liu J.S., Ren B. // *Nature.* 2012. V. 485. P. 376–380.
58. Sanyal A., Lajoie B.R., Jain G., Dekker J. // *Nature.* 2012. V. 489. P. 109–113.
59. Sexton T., Yaffe E., Kenigsberg E., Bantignies F., Leblanc B., Hoichman M., Parrinello H., Tanay A., Cavalli G., et al. // *Cell.* 2012. V. 148. P. 458–472.
60. Holwerda S., de Laat W. // *Front. Genet.* 2012. V. 3. P. 217.
61. Nora E.P., Dekker J., Heard E. // *BioEssays.* 2013. V. 35. P. 818–828.
62. Gibcus J.H., Dekker J. // *Mol. Cell.* 2013. V. 49. P. 773–782.
63. Kravchenko E., Savitskaya E., Kravchuk O., Parshikov A., Georgiev P., Savitsky M. // *Mol. Cell Biol.* 2005. V. 25. P. 9283–9291.
64. Li H.B., Müller M., Bahechar I.A., Kyrchanova O., Ohno K., Georgiev P., Pirrotta V. // *Mol. Cell Biol.* 2011. V. 31. P. 616–625.
65. Venken K.J., Bellen H.J. // *Meth. Mol. Biol.* 2012. V. 859. P. 203–228.
66. Fujioka M., Wu X., Jaynes J.B. // *Development.* 2009. V. 136. P. 3077–3087.
67. Fujioka M., Sun G., Jaynes J.B. // *PLoS Genet.* 2013. V. 9. e1003883.
68. Tanay A., Cavalli G. // *Curr. Opin. Genet. Dev.* 2013. V. 23. P. 197–203.
69. Ameres S.L., Drupeppel L., Pfliederer K., Schmidt A., Hillen W., Berens C. // *EMBO J.* 2005. V. 24. P. 358–367.
70. Bondarenko V.A., Jiang Y.I., Studitsky V.M. // *EMBO J.* 2003. V. 22. P. 4728–4737.
71. Kyrchanova O., Maksimenko O., Stakhov V., Ivlieva T., Parshikov A., Studitsky V.M., Georgiev P. // *PLoS Genet.* 2013. V. 9. e1003606.
72. Savitskaya E., Melnikova L., Kostuchenko M., Kravchenko E., Pomerantseva E., Boikova T., Chetverina D., Parshikov A., Zobacheva P., Gracheva E., et al. // *Mol. Cell Biol.* 2006. V. 26. P. 754–761.
73. Erokhin M., Davydova A., Kyrchanova O., Parshikov A., Georgiev P., Chetverina D. // *Development.* 2011. V. 138. P. 4097–4106.
74. Nakayama T., Shimojima T., Hirose S. // *Development.* 2012. V. 139. P. 4582–4590.
75. Ghirlando R., Giles K., Gowher H., Xiao T., Xu Z., Yao H., Felsenfeld G. // *Biochim. Biophys. Acta.* 2012. V. 1819. P. 644–651.
76. Vorobyeva N.E., Mazina M.U., Golovnin A.K., Kopytova D.V., Gurskiy D.Y., Nabirochkina E.N., Georgieva S.G., Georgiev P.G., Krasnov A.N. // *Nucl. Acids Res.* 2013. V. 41. P. 5717–5730.
77. Yajima M., Fairbrother W.G., Wessel G.M. // *Development.* 2012. V. 139. P. 3613–3622.
78. Li M., Belozherov V.E., Cai H.N. // *Mol. Cell Biol.* 2010. V. 30. P. 1067–1076.
79. Herold M., Bartkuhn M., Renkawitz R. // *Development.* 2012. V. 139. P. 1045–1057.
80. Merckenschlager M., Odom D.T. // *Cell.* 2013. V. 152. P. 1285–1297.
81. Chaumeil J., Skok J.A. // *Curr. Opin. Immunol.* 2012. V. 24. P. 153–159.
82. Nakahashi H., Kwon K.R., Resch W., Vian L., Dose M., Stavreva D., Hakim O., Pruett N., Nelson S., Yamane A., et al. // *Cell Rep.* 2013. V. 3. P. 1678–1689.
83. Yusufzai T.M., Felsenfeld G. // *Proc. Natl. Acad. Sci. USA.* 2004. V. 101. P. 8620–8624.
84. Pant V., Kurukuti S., Pugacheva E., Shamsuddin S., Mariano P., Renkawitz R., Klenova E., Lobanekov V., Ohlsson R. // *Mol. Cell Biol.* 2004. V. 24. P. 3497–3504.
85. Chernukhin I.V., Shamsuddin S., Robinson A.F., Carne A.F., Paul A., El-Kady A.I., Lobanekov V.V., Klenova E.M. // *J. Biol. Chem.* 2000. V. 275. P. 29915–29921.
86. Lutz M., Burke L.J., Barreto G., Goeman F., Greb H., Arnold R., Schultheiss H., Brehm A., Kouzarides T., Lobanekov V., et al. // *Nucl. Acids Res.* 2000. V. 28. P. 1707–1713.
87. Ishihara K., Oshimura M., Nakao M. // *Mol. Cell.* 2006. V. 23. P. 733–742.
88. Mehta G.D., Kumar R., Srivastava S., Ghosh S.K. // *FEBS Lett.* 2013. V. 587. P. 2299–2312.
89. Lee B.K., Iyer V.R. // *J. Biol. Chem.* 2012. V. 287. P. 30906–30913.
90. Xiao T., Wallace J., Felsenfeld G. // *Mol. Cell Biol.* 2011. V. 31. P. 2174–2183.
91. Parelho V., Hadjur S., Spivakov M., Leleu M., Sauer S., Gregson H.C., Jarmuz A., Canzonetta C., Webster Z., Nesterova T., et al. // *Cell.* 2008. V. 132. P. 422–433.
92. Wendt K.S., Yoshida K., Itoh T., Bando M., Koch B., Schirghuber E., Tsutsumi S., Nagae G., Ishihara K., Mishiro T., et al. // *Nature.* 2008. V. 451. P. 796–801.
93. Hadjur S., Williams L.M., Ryan N.K., Cobb B.S., Sexton T., Fraser P., Fisher A.G., Merckenschlager M. // *Nature.* 2009. V. 460. P. 410–413.
94. Nativio R., Wendt K.S., Ito Y., Huddleston J.E., Uribe-Lewis S., Woodfine K., Krueger C., Reik W., Peters J.M., Murrell A. // *PLoS Genet.* 2009. V. 5. e1000739.
95. Zuin J., Dixon J.R., van der Reijden M.I., Ye Z., Kolovos P., Brouwer R.W., van de Corput M.P., van de Werken H.J., Knoch T.A., van IJcken W.F., et al. // *Proc. Natl. Acad. Sci. USA.* 2014. V. 111. P. 996–1001.
96. Gosalia N., Neems D., Kerschner J.L., Kosak S.T., Harris A. // *Nucl. Acids Res.* 2014. V. 42. P. 9612–9622.
97. Chung J.H., Bell A.C., Felsenfeld G. // *Proc. Natl. Acad. Sci. USA.* 1997. V. 94. P. 575–580.
98. Emery D.W., Yannaki E., Tubb J., Stamatoyannopoulos G. // *Proc. Natl. Acad. Sci. USA.* 2000. V. 97. P. 9150–9155.
99. Bell A.C., West A.G., Felsenfeld G. // *Cell.* 1999. V. 98. P. 387–396.
100. West A.G., Huang S., Gaszner M., Litt M.D., Felsenfeld G. // *Mol. Cell.* 2004. V. 16. P. 453–463.
101. Huang S., Li X., Yusufzai T.M., Qiu Y., Felsenfeld G. // *Mol. Cell Biol.* 2007. V. 27. P. 7991–8002.
102. Dickson J., Gowher H., Strogantsev R., Gaszner M., Hair A., Felsenfeld G., West A.G. // *PLoS Genet.* 2010. V. 6. e1000804.
103. Giles K.E., Gowher H., Ghirlando R., Jin C., Felsenfeld G. // *Cold Spring Harbor Symp. Quant. Biol.* 2010. V. 75. P. 1–7.
104. Максименко О.Г., Дейкин А.В., Ходарович Ю.М., Георгиев П.Г. // *Acta Naturae.* 2013. T. 5. № 1(16). С. 33–47.
105. Rival-Gervier S., Pantano T., Viglietta C., Maeder C., Prince S., Attal J., Jolivet G., Houdebine L.M. // *Transgenic Res.* 2003. V. 12. P. 723–730.
106. Hanawa H., Yamamoto M., Zhao H., Shimada T., Persons D.A. // *Mol. Therapy.* 2009. V. 17. P. 667–674.

107. Aker M., Tubb J., Groth A.C., Bukovsky A.A., Bell A.C., Felsenfeld G., Kiem H.P., Stamatoyannopoulos G., Emery D.W. // *Hum. Gene Ther.* 2007. V. 18. P. 333–343.
108. Kim J.M., Kim J.S., Park D.H., Kang H.S., Yoon J., Baek K., Yoon Y. // *J. Biotech.* 2004. V. 107. P. 95–105.
109. Girod P.A., Zahn-Zabal M., Mermod N. // *Biotech. Bioeng.* 2005. V. 91. P. 1–11.
110. Harraghy N., Gaussin A., Mermod N. // *Curr. Gene Ther.* 2008. V. 8. P. 353–366.
111. Wang T.Y., Han Z.M., Chai Y.R., Zhang J.H. // *Mol. Biol. Rep.* 2010. V. 37. P. 3553–3560.
112. Bode J., Schlake T., Rios-Ramirez M., Mielke C., Stengert M., Kay V., Klehr-Wirth D. // *Int. Rev. Cytol.* 1995. V. 162. P. 389–454.
113. Platts A.E., Quayle A.K., Krawetz S.A. // *Cell Mol. Biol. Lett.* 2006. V. 11. P. 191–213.
114. Evans K., Ott S., Hansen A., Koentges G., Wernisch L. // *BMC Bioinformatics.* 2007. V. 8. P. 71.
115. Liebich I., Bode J., Frisch M., Wingender E. // *Nucl. Acids Res.* 2002. V. 30. P. 372–374.
116. Girod P.-A., Nguyen D.-Q., Calabrese D., Puttini S., Grandjean M., Martinet D., Regamey A., Saugy D., Beckmann J.S., Bucher P., et al. // *Nat. Methods.* 2007. V. 4. P. 747–753.
117. Welsh P.L., King M.C. // *Hum. Mol. Genet.* 2001. V. 10. P. 705–713.
118. Iarovaia O.V., Shkumatov P., Razin S.V. // *J. Cell Sci.* 2004. V. 117. P. 4583–4590.
119. Kulkarni A., Pavithra L., Rampalli S., Mogare D., Babu K., Shiekh G., Ghosh S., Chattopadhyay S. // *Biochem. Biophys. Res. Commun.* 2004. V. 322. P. 672–677.
120. Johnson C.N., Levy L.S. // *Virology.* 2005. V. 2. P. 68.
121. Koina E., Piper A. // *J. Cell Biochem.* 2005. V. 95. P. 391–402.
122. Mearini G., Chichiarelli S., Zampieri M., Masciarelli S., D'Erme M., Ferraro A., Mattia E. // *FEBS Lett.* 2003. V. 547. P. 119–124.
123. Piechaczek C., Fetzter C., Baiker A., Bode J., Lipps H.J. // *Nucl. Acids Res.* 1999. V. 27. P. 426–428.
124. Galbete J.L., Buceta M., Mermod N. // *Mol. Biosyst.* 2009. V. 5. P. 143–150.
125. Bode J., Benham C., Knopp A., Mielke C. // *Crit. Rev. Eukaryot. Gene Expr.* 2000. V. 10. P. 73–90.
126. Berezney R., Coffey D.S. // *Biochem. Biophys. Res. Commun.* 1974. V. 60. P. 1410–1417.
127. Albrethsen J., Knol J.C., Jimenez C.R. // *J. Proteomics.* 2009. V. 72. P. 71–81.
128. Cai S., Han H.J., Kohwi-Shigematsu T. // *Nat. Genet.* 2003. V. 34. P. 42–51.
129. Renz A., Fackelmayer F.O. // *Nucl. Acids Res.* 1996. V. 24. P. 843–849.
130. Agrelo R., Souabni A., Novatchkova M., Haslinger C., Leeb M., Komnenovic V., Kishimoto H., Gresh L., Kohwi-Shigematsu T., Kenner L., et al. // *Dev. Cell.* 2009. V. 16. P. 507–516.
131. Arope S., Harraghy N., Pjanic M., Mermod N. // *PLoS One.* 2013. V. 8. e79262.
132. Berrios M., Osheroff N., Fisher P.A. // *Proc. Natl. Acad. Sci. USA.* 1985. V. 82. P. 4142–4146.
133. Feister H.A., Onyia J.E., Miles R.R., Yang X., Galvin R., Hock J.M., Bidwell J.P. // *Bone.* 2000. V. 26. P. 227–234.
134. Adachi Y., Kas E., Laemmli U.K. // *EMBO J.* 1989. V. 8. P. 3997–4006.
135. Kohwi-Shigematsu T., Kohwi Y., Takahashi K., Richards H.W., Ayers S.D., Han H.J., Cai S. // *Methods.* 2012. V. 58. P. 243–254.
136. Burute M., Gottimukkala K., Galande S. // *Immunol. Cell Biol.* 2012. V. 90. P. 852–859.
137. Naito T., Tanaka H., Naoe Y., Taniuchi I. // *Int. Immunol.* 2011. V. 23. P. 661–668.
138. Yasui D., Miyano M., Cai S., Varga-Weisz P., Kohwi-Shigematsu T. // *Nature.* 2002. V. 419. P. 641–645.
139. Hasegawa Y., Brockdorff V.N., Kawano S., Tsutsui K., Nakagawa S. // *Dev. Cell.* 2010. V. 19. P. 469–476.
140. Tattermusch A., Brockdorff N. // *Hum. Genet.* 2011. V. 130. P. 247–253.
141. Razin S.V., Iarovaia O.V., Vassetzky Y.S. // *Chromosoma.* 2014. V. 123. P. 217–224.
142. Harraghy N., Buceta M., Regamey A., Girod P.A., Mermod N. // *Methods Mol. Biol.* 2012. V. 801. P. 93–110.
143. Zahn-Zabal M., Kobr M., Girod P.A., Imhof M., Chatellard P., de Jesus M. // *J. Biotech.* 2001. V. 87. P. 29–42.
144. Harraghy N., Regamey A., Girod P.-A., Mermod N. // *J. Biotechnol.* 2011. V. 154. P. 11–20.
145. Buceta M., Galbete J.L., Kostic C., Arsenijevic Y., Mermod N. // *Gene Therapy.* 2011. V. 18. P. 7–13.
146. Ley D., Harraghy N., Le Fourn V., Bire S., Girod P.A., Regamey A., Rouleux-Bonnin F., Bigot Y., Mermod N. // *PLoS One.* 2013. V. 8. e62784.
147. Grandjean M., Girod P.A., Calabrese D., Kostyrko K., Wicht M., Yerly F., Mazza C., Beckmann J.S., Martinet D., Mermod N. // *Nucl. Acids Res.* 2011. V. 39. e104.
148. Puttini S., van Zwieten R.W., Saugy D., Lekka M., Hogger F., Ley D., Kulik A.J., Mermod N. // *BMC Mol. Biol.* 2013. V. 14. P. 26.
149. Shimizu N., Miura Y., Sakamoto Y., Tsutsui K. // *Cancer Res.* 2001. V. 61. P. 6987–6990.
150. Noguchi C., Araki Y., Miki D., Shimizu N. // *PLoS One.* 2012. V. 7. e52990.
151. Antoniou M., Harland L., Mustoe T., Williams S., Holdstock J., Yague E., Mulcahy T., Griffiths M., Edwards S., Ioannou P.A., et al. // *Genomics.* 2003. V. 82. P. 269–279.
152. Williams S., Mustoe T., Mulcahy T., Griffiths M., Simpson D., Antoniou M., Irvine A., Mountain A., Crombie R. // *BMC Biotech.* 2005. V. 5. P. 17.
153. Benton T., Chen T., McEntee M., Fox B., King D., Crombie R., Thomas TC, Bebbington C. // *Cytotechnology.* 2002. V. 38. P. 43–46.
154. Zhang F., Thornhill S.I., Howe S.J., Ulaganathan M., Schambach A., Sinclair J., Kinnon C., Gaspar H.B., Antoniou M., Thrasher A.J. // *Blood.* 2007. V. 110. P. 1448–1457.
155. Zhang F., Frost A.R., Blundell M.P., Bales O., Antoniou M.N., Thrasher A.J. // *Mol. Ther.* 2010. V. 18. P. 1640–1649.
156. Dighe N., Khoury M., Mattar C., Chong M., Choolani M., Chen J., Antoniou M.N., Chan J.K. // *PLoS One.* 2014. V. 9. e104805.
157. Pfaff N., Lachmann N., Ackermann M., Kohlscheen S., Brendel C., Maetzig T., Niemann H., Antoniou M.N., Grez M., Schambach A., et al. // *Stem Cells.* 2013. V. 31. P. 488–499.
158. Brendel C., Müller-Kuller U., Schultze-Strasser S., Stein S., Chen-Wichmann L., Krattenmacher A., Kunkel H., Dillmann A., Antoniou M.N., Grez M. // *Gene Ther.* 2012. V. 19. P. 1018–1029.
159. Boscolo S., Mion F., Licciulli M., Macor P., De Maso L., Brce M., Antoniou M.N., Marzari R., Santoro C., Sblattero D. // *Nat. Biotechnol.* 2012. V. 29. P. 477–484.
160. Lindahl A.M., Antoniou M. // *Epigenetics.* 2007. V. 2. P. 227–236.
161. Nair A.R., Jinger X., Hermiston T.W. // *BMC Res. Notes.* 2011. V. 4. P. 178.
162. Jonuschies J., Antoniou M., Waddington S., Boldrin L., Muntoni F., Thrasher A., Morgan J. // *Curr. Gene Ther.* 2014. V. 14. P. 276–288.

163. Otte A.P., Kwaks T.H., van Blokland R.J., Sewalt R.G., Verhees J., Klaren V.N., Siersma T.K., Korse H.W., Teunissen N.C., Botschuijver S., et al. // *Biotechnol. Prog.* 2007. V. 23. P. 801–807.
164. Hoeksema F., Hamer K., Siep M., Verhees J.A., Otte A.P. // *Biotechnol. Res. Int.* 2011. 492875.
165. Wang J., Lawry S.T., Cohen A.L., Jia S. // *Cell Mol. Life Sci.* 2014. V. 71. P. 4841–4852.
166. Kwaks T.H.J., Barnett P., Hemrika W., Siersma T., Sewalt R.G., Satijn D.P., Brons J.F., van Blokland R., Kwakman P., Kruckeberg A.L., et al. // *Nat. Biotechnology.* 2003. V. 21. P. 553–558.
167. Schwartz Y.B., Linder-Basso D., Kharchenko P.V., Tolstorukov M.Y., Kim M., Li H.B., Gorchakov A.A., Minoda A., Shanower G., Alekseyenko A.A., et al. // *Genome Res.* 2012. V. 22. P. 2188–2198.