

UDC 577.214.6

Фактор транскрипции CTCF и организация генома млекопитающих

Е. С. Котова, С. Б. Акопов, Е. Д. Свердлов, Л. Г. Николаев

Институт биоорганической химии им. академиков М. М. Шемякина и Ю. А. Овчинникова Российской академии наук
Ул. Миклухо-Маклая, 16/10, Москва, В-437, 117997 ГСП

lev@ibch.ru

Фактор транскрипции CTCF считается одним из основных участников различных сетей регуляции генов, включая активацию и репрессию транскрипции, образование независимо функционирующих доменов хроматина, регуляцию импринтинга и т. д. Секвенирование геномов человека и других организмов позволяет выявить геномное распределение сайтов связывания CTCF и идентифицировать CTCF-зависимые регуляторные элементы, включая инсуляторы. В обзоре суммированы новые данные по функционированию CTCF в рамках гипотезы участия петельных доменов хроматина в крупномасштабной регуляции активности генома. Фундаментальные свойства CTCF позволяют ему действовать как регулятор транскрипции, инсуляторный белок, а также как распределенный по геному пограничный элемент, способный привлекать различные факторы, появляющиеся в ответ на разнообразные внешние и внутренние воздействия и таким образом осуществлять свои сигнал-специфичные функции.

Ключевые слова: фактор транскрипции CTCF, хроматин, регуляция транскрипции.

Введение. CTCF (CCCTC binding factor) – эволюционно консервативный транскрипционный фактор позвоночных. Он связывается с различными функциональными элементами генома и выполняет разнообразные регуляторные функции [1–3].

Транскрипционный фактор CTCF впервые идентифицирован как белок, узнающий последовательность трех прямых повторов CCCTC в регуляторной области гена *MYC* кур [4, 5]. Одновременно был описан белок NeP1, связывающийся с F1-элементом сайленсера гена лизоцима кур [6]. Как оказалось, NeP1 и CTCF представляют собой один и тот же белок [7].

Первичная структура CTCF высококонсервативна; ортологи CTCF выявлены у различных позвоночных: человека, собаки, мыши, кур, лягушек и др. [8–10]. Аминокислотные последовательности белка CTCF человека и кур идентичны на 93 % [11].

Полипептидная цепь CTCF человека состоит из 727 аминокислотных остатков (а. о.). Центральный ДНК-связывающий домен содержит 11 цинковых пальцев, фланкированных богатыми лизином и ар-

гинином положительно заряженными N- и C-концевыми доменами, содержащими соответственно 267 и 150 а. о. [12]. После ДНК-связывающего домена располагается глицин-богатый мотив, характерный для АТР- и GTP-связывающих белков. Ближе к C-концу полипептида находится сигнал ядерной локализации, за ним располагаются сайты фосфорилирования [9, 11].

Разнообразие выполняемых белком CTCF функций предполагает существование процесса регуляции активности самого CTCF, осуществляющегося в том числе и посредством его посттрансляционных модификаций, среди которых следует отметить фосфорилирование киназой 2 [13], ADP-рибозилирование [14], сумоилирование [15–17].

Эволюционно CTCF появляется у билатеральных многоклеточных животных и отсутствует у растений и простейших [18]. CTCF экспрессируется конститутивно и необходим для функционирования многих типов клеток позвоночных в многоклеточном организме. Для жизни клеток млекопитающих в культуре наличие CTCF, по всей видимости, не является обязательным [2]. Тем не менее, на большое значение белка CTCF для развития позвоноч-

ных указывает тот факт, что мышинные эмбрионы, гомозиготные по поврежденному гену белка CTCF, гибли еще до имплантации [19, 20]. Повреждение гена, кодирующего CTCF в ооцитах мышей, препятствует нормальному развитию бластоцисты после оплодотворения [21].

Функциональные свойства CTCF. Сайт связывания CTCF. С развитием технологий ChIP-chip (Chromatin Immunoprecipitation on chip) и ChIP-seq (Chromatin Immunoprecipitation-Sequencing) [22–24] получен ряд консенсусных последовательностей ДНК, с которыми предпочтительно взаимодействует CTCF [22–23, 25]. Оказалось, что не только аминокислотная последовательность белка CTCF, но и нуклеотидная последовательность его сайта связывания высококонсервативны для различных видов позвоночных.

Позднее процедура ChIP-seq была дополнена обработкой экзонуклеазой (ChIP-exo) [26], что позволило более точно локализовать сайты связывания CTCF в геноме. Показано также, что сайт связывания CTCF можно подразделить на четыре блока – каждый со своей консенсусной последовательностью. Около половины сайтов связывания CTCF содержат лишь два центральных блока. Различные сайты связывания CTCF включают или все четыре блока, или комбинацию двух или трех блоков, либо же только один блок.

Дальнейшие исследования привели к понятию коровой области связывания CTCF [27], лежащей в основе подавляющего большинства сайтов связывания (рис. 1).

Среди CTCF-связывающих последовательностей всегда присутствует небольшое число тех, для которых нельзя выявить какого-либо консенсуса [26, 27]. Предполагается, что эти последовательности связывают CTCF не напрямую, а через промежуточные белки [27]. Экспериментально обнаруженные CTCF-связывающие последовательности сведены в базу данных CTCF Binding Site Database [28].

Взаимодействие CTCF с ДНК. Методом торможения в геле исследовали взаимодействие фрагментов домена цинковых пальцев CTCF с некоторыми известными CTCF-связывающими последовательностями, в частности, сайтом из бета-глобинового локуса кур [8], сайтами локуса *Igf2/H19* мышей [29, 30] и сайтом ARVbeta, расположенным в промоторе предшественника бета-амилоида человека [31].

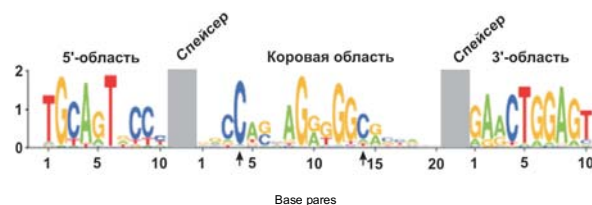


Рис. 1. Общая нуклеотидная структура сайта связывания CTCF [27]. Стрелками обозначены потенциальные сайты метилирования CpG

Поскольку цинковые пальцы – характерная ДНК-связывающая структура многих ядерных белков [32], было сделано предположение о том, что именно содержащий их домен отвечает за взаимодействие белка CTCF с ДНК. Показано, что четыре цинковых пальца (с 4-го по 7-й) являются минимальным набором, необходимым для специфического взаимодействия CTCF с его сайтом связывания *in vitro*. При дальнейшем уменьшении количества цинковых пальцев ДНК-связывающая способность фрагмента CTCF резко уменьшалась [30].

CTCF может использовать для связывания с ДНК различные комбинации цинковых пальцев. В частности, связывание с регуляторным сайтом гена *MYC* кур осуществляют цинковые пальцы со 2-го по 7-й, а с сайтом, расположенным рядом с промотором P2 гена *MYC* человека, CTCF связывается при помощи цинковых пальцев с 3-го по 11-й [11]. Для связывания с фрагментом F1 сайленсера гена лизоцима кур *in vitro* необходимы цинковые пальцы с 5-го по 8-й [7].

Позднее эти данные были подтверждены и расширены с использованием массивированного секвенирования [27]. Показано, что 8-й цинковый палец не обязателен для специфического узнавания CTCF его сайта связывания, но стабилизирует комплекс CTCF–ДНК за счет неспецифического взаимодействия. Цинковые пальцы 9–11 отвечают за взаимодействие с небольшим участком CTCF-связывающей последовательности, отделенным от корового участка спейсером. Этот фрагмент назван U-элементом (upstream element), он соответствует блоку 1, идентифицированному в работе [26]. Интересно, что при удалении цинковых пальцев 8–11 взаимодействие CTCF с сайтом связывания, содержащим U-элемент, полностью исчезает [27]. Возможно, что в том случае, когда коровая последовательность далека от консенсусной, ДНК-белковый комплекс стабилен только при дополнительном связывании CTCF с U-элементом. Цинковый палец 3 вовлечен во взаимодействие с ДНК в отсутствие U-элемента, а цинко-



Рис. 2. Гипотетическая модель образования петли при помощи CTCF: *A* – связывание CTCF со своим сайтом (обозначен стрелкой) за счет части цинковых пальцев; *B* – с помощью основной части цинковых пальцев белок связывается со своим сайтом, остальные цинковые пальцы остаются свободными для образования петли (при связывании CTCF ДНК изгибается, что позволяет свободным цинковым пальцам замкнуть петлю [33]); *B* – образование петли вследствие связывания двух молекул CTCF с двумя сайтами, расположенными на границах функционального домена (возможно, с участием или без него промежуточного белка (зеленый овал)); свободные цинковые пальцы могут принимать участие в белково-белковых взаимодействиях [33]; заключенные в квадрат области соответствуют инсуляторам или пограничным элементам домена)

вые пальцы 1 и 2, по-видимому, не участвуют в специфическом связывании с ДНК, но отвечают за общую стабилизацию ДНК-белкового комплекса.

Предложены гипотетические механизмы образования петли (рис. 2), посредством которых CTCF может участвовать в формировании доменов хроматина и функционирует как белковый компонент инсуляторов [33, 34].

Влияние метилирования CpG ДНК на связывание CTCF. Показано, что метилирование CpG в сайтах связывания CTCF приводит к подавлению их взаимодействия с белком CTCF *in vitro* [30, 35–37]. Кроме того, CTCF не связывается с метилированной областью ICR локуса отцовского аллеля *in vivo* [38, 39], а связывание CTCF с ICR локуса *Igf2/H19* материнской хромосомы препятствует ее метилированию при развитии организма [39].

Ванг и соавт. [40] исследовали дифференциальное метилирование *in vivo* CTCF-связывающих последовательностей на 19 типах клеток в культуре и тканей человека методами ChIP-seq и массивированного бисульфитного секвенирования. 36 % сайтов связывания CTCF присутствовали во всех 19 типах клеток. Для 67 % дифференциально метилированных сайтов связывания CTCF показана обратная зависимость между степенью их метилирования и способностью связывать CTCF. Таким образом, метилирование ДНК является важным фактором, определяющим, будет ли белок CTCF связываться с данной нуклеотидной последовательностью. Также установлено, что около 29 % CTCF-связывающих последовательностей в геноме содержат CpG хотя бы в одном из двух положений: –1 и 11, что со-

ответствует положениям 4 и 14 коровой последовательности сайта связывания CTCF (рис. 1, [27]) и положениям –5 и 5 [26].

Взаимодействие CTCF с белками. При иммунопреципитации из клеточного лизата CTCF выделяется вместе с белком ядрышка нуклеофосмином. С другой стороны, методом иммунопреципитации хроматина с антителами к нуклеофосмину продемонстрировано, что нуклеофосмин взаимодействует с CTCF-зависимыми инсуляторами *in vivo* [41]. Содержащая хромодомен хеликаза CHD8 также связывается с CTCF и CTCF-зависимыми инсуляторами. В отсутствие CHD8 область ICR локуса *Igf2/H19* перестает выполнять инсуляторную функцию, хотя связывание CTCF с этой областью сохраняется [42]. В белковых комплексах CTCF-зависимых инсуляторов выявлено присутствие большой субъединицы РНК-полимеразы II с фосфорилированным и дефосфорилированным С-концевым доменом [43].

Коиммунопреципитацией установлено, что белок CTCF взаимодействует с ДНК-связывающим белком YY1. По-видимому, во взаимодействии участвует N-концевой домен CTCF. Связывание CTCF с YY1 необходимо во время инактивации одной X-хромосомы для поддержания второй X-хромосомы в активном состоянии [44].

К настоящему времени различными методами обнаружено еще несколько белков, способных взаимодействовать с CTCF. Это YB1 (Y-box-binding protein 1) – мультифункциональный ДНК и РНК-связывающий белок, участвующий в регуляции процессов репликации и репарации ДНК, транскрипции, процессинга РНК, также способный взаимо-

действовать с YY1 [45], и Kaiso – транскрипционный фактор, обладающий способностью связывать участки ДНК с повышенным содержанием метилированных CpG-сайтов. По-видимому, он может замещать CTCF в случае метилирования сайта связывания последнего [46]. С белком CTCF также взаимодействуют транскрипционный корепрессор Sin3A [47], гистон H2A.Z [22, 41], PARP [48], p68 (DDX5) [49] и другие белки [2, 50].

Еще одним важным белком, взаимодействующим с CTCF, является когезин. Когезин отвечает за удерживание вместе сестринских хроматид, необходимое для успешного протекания митоза и мейоза [51–53] и представляет собой белковый комплекс, состоящий из четырех субъединиц: Smc1, Smc3, Scc1 (Rad21) и Scc3 (SA1 или SA2). Четыре субъединицы образуют структуру в форме кольца, внутри которого, предположительно, оказываются удерживаемые вместе хроматиды [54]. Известно, что когезин отвечает не только за взаимодействие между собой сестринских хроматид, необходимое для успешного протекания митоза и мейоза, но и вовлечен в регуляцию экспрессии генов [55–57]. По результатам ChIP-seq анализа обнаружено, что примерно половина сайтов связывания когезина перекрываются с сайтами связывания CTCF [58–60]. Когезин связывается с участком в С-концевой области CTCF через свою субъединицу SA2 [61].

Методом 3С выявлена пространственная сближенность участков интерфазного хроматина, с которыми взаимодействуют одновременно когезин и CTCF. При этом участок хромосомы между сайтами связывания когезина образует петлю [56].

Роль транскрипционного фактора CTCF в регуляции ДНК-зависимых процессов. Связывание CTCF с ДНК может влиять на экспрессию генов различным образом: в одних случаях CTCF выполняет роль активатора транскрипции, в других – репрессора или же обеспечивает активность инсулятора.

Инсуляторами называют фрагменты ДНК, препятствующие взаимодействию регуляторных элементов, между которыми они располагаются. В частности, находясь между промотором и энхансером, инсулятор блокирует активирующее действие последнего, а также, будучи помещенным по краям генетической конструкции в составе геномной ДНК эукариотической клетки, препятствует эффекту положения [62]. Все найденные на данный момент ин-

суляторы позвоночных, за редкими исключениями [63], связывают транскрипционный фактор CTCF.

Кроме того, белок CTCF участвует в инактивации X-хромосомы, импринтинге генетической информации и, по-видимому, регулирует процесс сплайсинга РНК. Разработка новых методов изучения взаимодействий удаленных друг от друга областей эукариотического генома позволила получить множество свидетельств того, что CTCF играет важнейшую роль в образовании трехмерной структуры эукариотического генома [34, 64–68]. Понимание механизмов, лежащих в основе столь многочисленных и различных функций белка CTCF, должно прояснить его роль в регуляции ДНК-зависимых процессов.

Гипотеза функциональных доменов хроматина. Инсуляторы. В конце 80-х годов XX века была выдвинута гипотеза, в соответствии с которой весь хроматин эукариотической клетки разделен на структурно-функциональные домены. В соответствии с этой гипотезой хроматиновый домен представляет собой петлю, содержащую один или несколько генов, концы которой закреплены в ядерном матриксе. Для отдельных петель характерна независимая сверхспирализация ДНК. Хроматин одного домена может независимо от хроматина других доменов переходить в открытую (транскрипционно активную) или закрытую (неактивную) конформацию [69, 70].

Ко времени формулировки гипотезы хроматиновых доменов было получено множество данных, подтверждающих, дополняющих и вносящих изменения в эту гипотезу. В настоящий момент установлено, что основным фактором, контролирующим декомпактизацию хроматиновой фибриллы, и, следовательно, возможность начала транскрипции на данном участке хромосомы, является ацетилирование гистонов [71, 72]. Выявлены так называемые LCR (locus control regions) – участки ДНК, определяющие транскрипционный статус домена [73–75].

Изучение бета-глобинового локуса кур показало, что в клетках, экспрессирующих бета-глобины, хроматин локуса находится в декомпактизованном состоянии, уровень ацетилирования гистонов повышен. При этом за пределами локуса хроматин находится в конденсированном состоянии [76–78]. Концевые участки бета-глобинового локуса взаимодействуют с ядерным матриксом и сближены между собой, при этом локус образует петлевую структуру [79–83]. Показано также, что петли могут фор-

мироваться не только при сближении границ функциональных доменов, но и при сближении отдельных участков ДНК, например, энхансеров и промоторов внутри доменов [81, 84, 85].

В недавних работах установлено, что транскрипционная активность гена зависит не только от регуляторных элементов внутри домена, но и от того, в какой части ядра данный ген располагается [64–68]. При этом ген может подвергаться воздействию регуляторных элементов, локализованных в других доменах и даже на других хромосомах [86–88].

Таким образом, дополненная и уточненная гипотеза структурно-функциональных хроматиновых доменов остается актуальной.

Если существуют хроматиновые домены, то должны быть и функциональные элементы, препятствующие влиянию на гены одного домена регуляторных элементов, принадлежащих другим доменам. Такими функциональными элементами генома являются инсуляторы [89, 90].

Инсуляторы предотвращают нежелательную активацию или репрессию генов под влиянием окружения. Нежелательная активация гена энхансером подавляется блокированием его действия на промотор только в том случае, когда инсулятор располагается между ними. Нежелательную репрессию гена инсулятор предотвращает, ограничивая распространение конденсированного хроматина вдоль хроматиновой фибриллы [89, 91]. Некоторые авторы называют инсуляторы пограничными элементами (*border elements*) в связи с тем, что они часто находятся на границах доменов (рис. 2) [92].

Подавляющее большинство инсуляторов, обнаруженных в геномах позвоночных, обладают способностью связывать фактор транскрипции CTCF [8, 63, 89]. CTCF играет важнейшую роль в образовании хроматиновых петель. Показано, что в бета-глобиновых локусах позвоночных сайты связывания CTCF на границах локусов контактируют друг с другом [80, 81, 93]. При этом в экспрессирующих бета-глобина клетках не происходит активации энхансерами локуса промоторов соседних доменов. Аналогично регуляторные элементы соседних доменов не влияют на экспрессию бета-глобиновых генов в эритроидных клетках [76–78, 94, 95].

Анализ генома человека показал, что распределение CTCF-связывающих сайтов коррелирует с плотностью распределения генов, но слабо коррели-

рует с длиной хромосом [24]. Несмотря на зависимость количества сайтов связывания CTCF от количества генов, большая часть этих сайтов (46 %) располагается очень далеко от промоторов – на расстоянии в среднем 48 тыс. п. н. [24, 96]. Это отличает распределение сайтов связывания CTCF от такового большинства других транскрипционных факторов и согласуется с инсуляторной функцией CTCF.

В участки генома, обедненные сайтами связывания CTCF, входят в основном семейства генов с общей регуляцией транскрипции, тогда как обогащенные сайтами связывания CTCF домены содержат гены с альтернативно регулируемым промоторами. Эти наблюдения согласуются также с инсуляторной функцией CTCF [24].

Блокирование энхансера. Первые данные об энхансер-блокирующих свойствах сайтов связывания CTCF получены с использованием конструкций, содержащих находящийся под контролем промотора репортерный ген и энхансер. Между энхансером и промотором вводили CTCF-связывающую последовательность [8, 97]. Анализ эффективности экспрессии репортерного гена при наличии и отсутствии сайта связывания CTCF позволяет судить об энхансер-блокирующей активности этого фрагмента ДНК. Большая часть CTCF-связывающих последовательностей проявляет энхансер-блокирующую активность в составе таких конструкций как при транзientной трансфекции, так и при интеграции в геном [98–100].

Метод оценки экспрессии репортерного гена при транзientной трансфекции имеет свои недостатки. Сайты связывания CTCF в норме функционируют в составе хроматина, а плазмидный вектор не имитирует полностью хроматинового окружения. Кроме того, во многих случаях плазмидный вектор представляет собой гетерологичную систему, в которой энхансер, промотор, сайт связывания CTCF и репортерный ген могут происходить из различных организмов.

В работе [101] при изучении механизма функционирования CTCF-зависимого инсулятора использовали искусственную минихромосому, содержащую энхансер из LCR локуса бета-глобина человека и ген эпсилон-глобина с собственным промотором. CTCF-зависимый инсулятор 5'HS4 кур клонировали между энхансером и промотором гена эпсилон-глобина и для контроля – рядом с энхансером, с противоположной стороны от промотора.

При расположении инсулятора между энхансером и промотором продемонстрирован CTCF-зависимый энхансер-блокирующий эффект, а также CTCF-зависимое снижение количества РНК-полимеразы II, связанной с промотором, и увеличение количества полимеразы, взаимодействующей с энхансером и CTCF-связывающей последовательностью. Этот эффект можно объяснить тем, что CTCF-содержащий ДНК-белковый комплекс мешает продвижению РНК-полимеразы от энхансера к промотору. Отмечено также CTCF-зависимое падение уровня ацетилирования гистонов H3 и H4 между 5'HS4 инсулятором и геном, включая область промотора, в том случае, когда инсулятор 5'HS4 расположен между энхансером и промотором [101].

Введение нескольких коровых последовательностей куриного инсулятора 5'HS4 в состав плазмидной конструкции между энхансером и промотором приводило к более осязаемому энхансер-блокирующему эффекту, нежели встраивание одной такой последовательности [102]. Подобный эффект наблюдается и при помещении между энхансером и промотором нескольких CTCF-зависимых инсуляторов HS1 или HS2 из ICR локуса *Igf2/H19* мыши [37, 38]. Таким образом, несколько CTCF-содержащих ДНК-белковых комплексов затрудняют перемещение РНК-полимеразы II от энхансера к промотору эффективней, чем один.

Однако передачу сигнала вдоль хроматиновой фибриллы вряд ли можно рассматривать как единственный механизм функционирования энхансеров. Многие энхансеры отделены от своих промоторов-мишеней миллионами нуклеотидных пар [103]. И энхансер, и промотор зачастую оказываются разделенными несколькими хромосомными локусами с различной структурой хроматина. Кроме того, описана активация промотора энхансером, расположенным на другой хромосоме [86, 88].

Некоторые авторы считают, что этот механизм вообще не реализуется в случае настоящих эукариотических энхансеров, а только для энхансер-подобных элементов прокариотов [104].

В последнее время с помощью технологии 3C (chromosome conformation capture) и ее производных [105] показано непосредственное взаимодействие энхансера и промотора, активируемого энхансером [106, 107], хотя механизм, по которому формирование хроматиновой петли и сближение энхансера и промотора ведет к активации промотора, точ-

но неизвестен [108]. Возможно, за счет приближения промотора к энхансеру происходит привлечение промотора в тот компартмент ядра, где есть условия для активной транскрипции [108]. Способность стабилизировать хроматиновые петли была продемонстрирована методом ChIA-PET для различных белков, в том числе для CTCF [34, 87].

В соответствии с моделью ловушки [62, 109] CTCF-связывающие инсуляторы могут конкурировать с промоторами за взаимодействие с энхансерами, то есть перехватывать энхансер, не позволяя ему связаться с промотором. В работе [34] показано, что значительная часть CTCF-связывающих последовательностей, участвующих в удаленных взаимодействиях, связана с последовательностями, обладающими свойствами энхансеров. С другой стороны, если инсулятор функционирует по модели ловушки, то между энхансером и промотором должна быть передача сигнала вдоль хроматиновой фибриллы, иначе ловушка будет действовать одинаково независимо от того, с какой стороны от энхансера она расположена, что противоречит определению инсулятора. В бета-глобиновых локусах позвоночных методом 3C продемонстрировано, что энхансеры и промоторы физически сближены между собой [80], что свидетельствует в пользу функционирования энхансера по комбинированному механизму tracking-looping.

Инсулятор как пограничный элемент. При введении в геном клетки конструкций, содержащих трансген, экспрессия трансгена сначала происходит приблизительно на одном уровне во всех клетках, но через какое-то время большая часть встроившихся в геном конструкций становится неактивной из-за конденсации хроматина. Если у дрозофилы и дрожжей это ингибирование экспрессии обратимо, то у позвоночных становится необратимым из-за метилирования ДНК [110]. Инсулятор 5'HS4, расположенный на 5'-конце бета-глобинового локуса, препятствует распространению находящегося рядом участка конденсированного хроматина на весь локус.

В работе [111] методом ChIP-seq показано, что часть сайтов связывания CTCF локализуется в местах перехода компактизованного хроматина в декомпактизованный. Такие места определяли по изменению уровня триметилирования остатка лизина 27 гистона H3 (H3K27me3) и ацетилирования остатка лизина 5 гистона H2A (H2AK5ac). Модифи-

кация H3K27me3 характерна для компактного хроматина, а H2AK5ac – для декомпактизованного. Доля CTCF-связывающих последовательностей, расположенных в местах перехода компактизованного хроматина в декомпактизованный, невелика, но с высокой значимостью показан неслучайный характер расположения этих последовательностей, что указывает на возможную роль сайтов связывания CTCF в разделении хромосомных доменов *in vivo*.

Авторы работы [112] продемонстрировали значительное обогащение сайтами связывания CTCF пограничных областей так называемых топологических доменов. Они предположили, что пограничные области топологических доменов могут представлять собой инсуляторы, блокирующие распространение гетерохроматина. Действительно, в дифференцированных клетках пограничным элементам топологических доменов соответствуют области с пониженным уровнем содержания H3K9me, характерные для перехода гетерохроматина в менее компактный хроматин. По-видимому, при формировании окончательной структуры геномных доменов стабилизация состояния хроматина происходит после образования петлевой структуры [112].

Возможно, что CTCF-связывающие последовательности не участвуют напрямую в блокировании распространения гетерохроматина. Однако четкое разделение в геноме участков компактизованного и декомпактизованного хроматина может быть следствием образования CTCF-содержащими пограничными элементами петлевых структур.

CTCF как регулятор транскрипции. CTCF способен как активировать, так и ингибировать транскрипцию в зависимости от целевого гена. Механизмы активации при этом могут быть различными. Во-первых, CTCF может выступать как белок – регулятор транскрипции, способствуя или препятствуя образованию инициаторного комплекса. Во-вторых, связывание CTCF может приводить к формированию/разрушению хроматинового домена с соответствующим изменением структуры хроматина, что, в свою очередь, приводит к изменению уровня экспрессии находящихся в данном домене генов. Нужно отметить, что четкую границу между этими механизмами можно провести далеко не всегда. По первому механизму, скорее всего, осуществляется репрессия гена *MYC* [5], по второму – репрессия гена *PUMA* [113]. Некоторые примеры регуляции с участием CTCF приведены ниже.

При связывании CTCF с 5'-регуляторной областью гена *APP* (amyloid β protein precursor) наблюдалась активация промотора этого гена при транскрипции *in vitro* в ядерном экстракте клеток HeLa [31, 114]. При добавлении к ядерному экстракту олигонуклеотидов, конкурирующих за связывание CTCF, активность промотора снижалась. Тот же эффект наблюдался при удалении CTCF из ядерного экстракта методом иммунопреципитации, однако при внесении в обедненный ядерный экстракт белка CTCF активность промотора восстанавливалась. За активацию ответствен участок белка CTCF между аминокислотными остатками 1 и 248 [114].

Показано, что гиперчувствительный к ДНКазе I участок ДНК HS5-1, расположенный в кластере протокадгеринных генов, обладает свойствами энхансера [115] и, по данным иммунопреципитации хроматина, взаимодействует с CTCF в клетках головного мозга мышей. При делеции этой CTCF-связывающей последовательности из генома трансгенных мышей активность промотора снижается [116]. Методом иммунопреципитации хроматина с последующим масштабным секвенированием продемонстрировано, что CTCF и когезин взаимодействуют с промотором альтернативных форм альфа-протокадгерина. Связывание прямо коррелирует с экспрессией альтернативных изоформ [117]. Методом 3C показано пространственное сближение промоторов альфа-протокадгеринных 4, 8 и 12 с потенциальными энхансерами HS5-1 и HS7 [84].

Сайленсер гена лизоцима кур состоит из двух модулей F1 и F2. Последний связывает рецептор тиреоидных гормонов, тогда как F1 содержит сайт связывания CTCF, при этом оба модуля могут подавлять транскрипцию независимо друг от друга. Когда каждый из них связывает свой белковый фактор, их действие суммируется [6]. Возможно, в механизмы супрессии–активации гена лизоцима кур вовлечен косупрессор Sin3A и комплексы гистоновых ацетилаз и деацетилаз [47, 91, 118].

Взаимодействие CTCF с двумя CpG-островками в первом интроне гена *Bcl6* не позволяет этому гену активно экспрессироваться. При некоторых разновидностях лимфомы эти CpG-островки метилируются, что препятствует связыванию CTCF с ДНК и приводит к увеличению количества мРНК *Bcl6* в клетках [119].

Связывание CTCF с участком в первом экзоне гена теломеразы человека *hTERT* подавляет транс-

крипцию этого гена. Метилирование CTCF-связывающей последовательности, по-видимому, играет важную роль в регуляции экспрессии этого гена. Первый экзон гена теломеразы в составе плазмидной конструкции оказывает репрессивное действие на ряд промоторов как в клетках, экспрессирующих теломеразу, так и в клетках, ее не экспрессирующих. При этом в клетках, экспрессирующих теломеразу, наблюдается метилирование CTCF-связывающей последовательности в первом экзоне этого гена и отсутствие взаимодействия CTCF с этой последовательностью [120].

После обработки 5-аза-2'-дезоксцитидином, вызывающей глобальное деметилирование ДНК, в клеточных линиях, экспрессирующих *hTERT*, наблюдали связывание CTCF с фрагментом 1-го экзона этого гена и подавление образования мРНК *hTERT*. Ингибирование экспрессии CTCF с помощью коротких шпилечных РНК приводило к увеличению количества мРНК *hTERT* в клеточных линиях [120, 121]. Установлено также участие CTCF в регуляции транскрипции генов *Bax* [123], опухолевого репрессора p16 [123] и генов рибосомных РНК [124].

Участие белка CTCF в импринтинге генетической информации. Под генным импринтингом понимают способ наследования, при котором у потомства экспрессируется только один из родительских аллелей гена. Какой из аллелей будет экспрессирован, определяется его наследованием от отцовского или материнского организма. CTCF взаимодействует с областями, контролирующими импринтинг (ICR – imprinting control region). Наиболее известным примером участия CTCF в импринтинге является регуляция генов локуса *Igf2/H19* [37, 125]. Ген *Igf2* (Insulin-like growth factor 2) кодирует эмбриональный митоген [37], с гена *H19* транскрибируется некодирующая РНК, замедляющая рост плода [62, 126, 127]. Экспрессия *Igf2* происходит только с отцовской, а *H19* – только с материнской хромосомы, что достигается при помощи расположенного между этими генами ICR. ICR на материнской хромосоме мышей содержит два сайта, способных связывать CTCF и обладающих свойствами инсуляторов. На отцовской хромосоме ICR метилирован независимо от ткани и стадии развития и не способен связывать CTCF. Таким образом, на отцовской хромосоме CTCF-зависимые инсуляторы неактивны, промотор свободно взаимодействует с энхансерами и *Igf2* экспрессируется [128].

Взаимодействуя с ICR материнской хромосомы, CTCF предположительно защищает от метилирования близлежащие области локуса, в том числе промотор и внутренние области гена *H19*. Если материнский ICR мутирован и не связывает CTCF, метилирование распространяется на промотор и внутригенные участки *H19*, и уровень экспрессии этого гена снижается [129, 139]. Таким образом, метилирование отцовского ICR, подавляя связывание CTCF, приводит к метилированию промотора гена *H19* и подавлению экспрессии отцовского аллеля [129–131].

CTCF и инактивация X-хромосомы. CTCF участвует в процессе инактивации X-хромосомы млекопитающих [132]. В состав X-хромосомы входит центр инактивации (Xic, X-chromosome inactivation center). Этот элемент контролирует инактивацию одной из X-хромосом в каждой клетке в ходе эмбриогенеза и дальнейшее поддержание ее в неактивном состоянии. В состав Xic входят гены *Xist* (X-inactive specific transcript), *Tsix* (название образовано от *Xist* перестановкой букв в обратном порядке, что отражает обратную ориентацию гена *Tsix* по отношению к *Xist*) и *Xite* (X-inactivation intergenic transcription element), с которых транскрибируются некодирующие РНК. X-хромосома, на которой расположен активно экспрессирующийся аллель гена *Xist*, инактивируется. Транскрипция генов *Tsix* и *Xite* происходит с активной X-хромосомы. Транскрипт гена *Tsix* инактивирует экспрессию гена *Xist*, расположенного на той же хромосоме, а экспрессия гена *Xite* активирует экспрессию *Tsix*, что приводит к подавлению экспрессии гена *Xist*, в результате чего X-хромосома, на которой находятся транскрипционно активные аллели генов *Tsix* и *Xite*, остается активной [50, 133].

Инактивация X-хромосомы происходит в несколько стадий [133–136]. Это «подсчет» количества X-хромосом в клетке, а, точнее, определение отношения количества X-хромосом к количеству аутосом [133, 134]. На следующей стадии осуществляется гомологичное спаривание двух X-хромосом и выбор той, которая будет инактивирована. X-хромосомы взаимодействуют в районе Xic, причем для их слияния необходим и достаточен фрагмент Xic, включающий гены *Xite* и *Tsix* [137, 138]. Эта область изобилует сайтами связывания CTCF, а также белка YY1 [44, 132, 139]. Показано, что при отсутствии в клетке CTCF спаривание X-хромосом подавляется. Белок YY1, способный взаимодейст-

вовать с CTCF, не является необходимым для протекания этой стадии [139]. Существует модель, в соответствии с которой в результате спаривания X-хромосом происходит необратимый перенос белковых факторов, таких как CTCF и Oct-4, с одной X-хромосомы (в будущем неактивной) на другую (в будущем активную) [44, 138–142]. Если эта модель верна, то белок CTCF причастен и к выбору X-хромосомы, которая будет инактивирована. Последняя стадия – это сам процесс инактивации X-хромосомы. При этом происходит активная экспрессия гена *Xist*, транскрипт гена *Xist* покрывает ту хромосому, на которой он был синтезирован и хроматин этой X-хромосомы компактизируется [134–136].

При инактивации X-хромосомы отмечено взаимодействие репрессорного комплекса PRC2 с 5'-областью гена *Xist* и обогащение хроматина этой области гистонам H3, триметилированным по остатку лизина 27 [143, 144]. Для инактивации X-хромосомы важна транскрипция небольшого участка гена *Xist*, в результате чего образуется РНК, которую называют RepA [145, 146]. По-видимому, RepA взаимодействует с комплексом PRC2 [144]. Показано, что X-хромосома, содержащая в промоторе гена *Xist* сайт связывания CTCF с повышенным родством к этому белку, чаще подвергается инактивации, чем X-хромосома, на которой такой фрагмент ДНК обладает пониженным родством к CTCF [147]. Вместе с YY1 CTCF, по-видимому, участвует в активации экспрессии гена *Tsix* [44]. CTCF также обеспечивает функционирование инсуляторных элементов, отделяющих расположенные на инактивированной X-хромосоме активные гены от транскрипционно неактивного окружения [148].

Котранскрипционная регуляция альтернативного сплайсинга. Недавно стали появляться свидетельства в пользу того, что CTCF может быть вовлечен в котранскрипционную регуляцию альтернативного сплайсинга. Продемонстрировано, что взаимодействие CTCF с фрагментом ДНК, расположенным в пятом экзоне гена *CD45*, приводит к более частому включению этого экзона в состав зрелой мРНК. Уменьшение количества CTCF в клетке с использованием РНК-интерференции или метилирование сайта связывания CTCF в пятом экзоне гена *CD45* подавляет образование сплайс-формы, содержащей пятый экзон. Установлено, что связывание CTCF с этим сайтом приводит к тому, что РНК-

полимераза II, достигая участка ДНК, взаимодействующего с CTCF, делает паузу. Авторы работ [149, 150] сравнили полногеномные данные, полученные методом иммунопреципитации хроматина с антителами к CTCF и РНК-полимеразе II, и данные анализа транскриптома. В результате получено подтверждение того, что связанный с ДНК белок CTCF задерживает транскрибирующую РНК-полимеразу II [149, 150].

На основе этих данных предложена гипотеза о роли белка CTCF в регуляции альтернативного сплайсинга. Начальные этапы сплайсинга РНК проходят во время транскрипции, при этом преимущественное образование той или иной сплайс-формы может зависеть от скорости элонгации. Если в определенный промежуток времени (после появления в составе транскрипта относительно слабого сплайс-сайта, но до появления более сильного) РНК-полимераза II будет задержана связанным с ДНК белком CTCF, то более слабый сплайс-сайт получит преимущество. Таким образом, белок CTCF может влиять на скорость элонгации и регулировать альтернативный сплайсинг [149, 151].

Выводы. Основными функциями CTCF, известными к настоящему времени, являются:

1. Непосредственная регуляция транскрипции некоторых генов;
2. Организация доменной структуры хроматина;
3. Организация инсуляторов (энхансер-блокирующих и пограничных элементов);
4. Участие в геномном импринтинге и инактивации X-хромосомы;
5. Регуляция альтернативного сплайсинга.

Точный механизм осуществления этих функций в большинстве случаев исследован недостаточно.

Большое количество и разнообразие сайтов связывания CTCF, множество выполняемых им функций, а также тот факт, что CTCF появляется лишь у относительно развитых многоклеточных животных и отсутствует у растений и простейших, позволяет предположить, что его основная задача состоит в регуляции развития организма и организации генома клетки таким образом, чтобы она могла быть составной частью многоклеточного организма.

Финансирование. Работа финансировалась программой поддержки ведущих научных школ России (проект НШ_1674.2012.4), программой Президиума РАН «Молекулярная и клеточная биология» и РФФИ (проекты 10-04-01365, 10-04-01472).

Transcription factor CTCF and mammalian genome organization

E. S. Kotova, S. B. Akopov, E. D. Sverdlov, L. G. Nikolaev

Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry,
Russian Academy of Sciences

16/10, Miklukho-Maklaya Str., Moscow, 117997, Russian Federation

Summary

The CTCF transcription factor is thought to be one of the main participants in various gene regulatory networks including transcription activation and repression, formation of independently functioning chromatin domains, regulation of imprinting etc. Sequencing of human and other genomes opened up a possibility to ascertain the genomic distribution of CTCF binding sites and to identify CTCF-dependent cis-regulatory elements, including insulators. In the review, we summarized recent data on CTCF functioning within a framework of the chromatin loop domain hypothesis of large-scale regulation of the genome activity. Its fundamental properties allow CTCF to serve as a transcription factor, an insulator protein and a dispersed genome-wide demarcation tool able to recruit various factors that emerge in response to diverse external and internal signals, and thus to exert its signal-specific function(s).

Keywords: transcription factor CTCF, chromatin, transcription regulation.

Фактор транскрипції CTCF і організація геному ссавців

О. С. Котова, С. Б. Акопов, Е. Д. Сverdlov, Л. Г. Ніколаєв

Резюме

Фактор транскрипції CTCF вважають одним з основних учасників різних мереж регуляції генів, з-поміж яких активація і репресія транскрипції, утворення незалежно функціонуючих доменів хроматину, регуляція імпринтингу тощо. Секвенування геномів людини та інших організмів дозволяє виявляти геномний розподіл сайтів зв'язування CTCF та ідентифікувати CTCF-залежні регуляторні елементи, до яких належать інсулятори. В огляді підсумовано нові дані з функціонування CTCF у рамках гіпотези участі петельних доменів хроматину у великомасштабній регуляції активності геному. Фундаментальні властивості CTCF дозволяють йому діяти як регулятор транскрипції, інсуляторний білок, а також як розподілений по геному прикордонний елемент, здатний залучати різні фактори, які з'являються у відповідь на різноманітні зовнішні та внутрішні чинники і таким чином виконувати свої сигнал-специфічні функції.

Ключові слова: фактор транскрипції CTCF, хроматин, регуляція транскрипції.

REFERENCES

- Holwerda SJ, de Laat W. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philos Trans R Soc Lond B Biol Sci*. 2013;**368**(1620):20120369.
- Nikolaev LG, Akopov SB, Didych DA, Sverdlov ED. Vertebrate protein CTCF and its multiple roles in a large-scale regulation of genome activity. *Curr Genomics*. 2009;**10**(5):294–302.
- Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell*. 2009;**137**(7):1194–211.
- Lobanenkov VV, Gudvin GG. CCCTC-binding protein: a new nuclear protein factor which interaction with 5'-flanking sequence of chicken c-myc oncogene correlates with repression of the gene. *Dokl Akad Nauk SSSR*. 1989;**309**(3):741–5.
- Lobanenkov VV, Nicolas RH, Adler VV, et al. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene*. 1990;**5**(12):1743–53.
- Baniahmad A, Steiner C, Kohne AC, Renkawitz R. Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. *Cell*. 1990;**61**(3):505–14.
- Burcin M, Arnold R, Lutz M, et al. Negative protein 1, which is required for function of the chicken lysozyme gene silencer in conjunction with hormone receptors, is identical to the multivalent zinc finger repressor CTCF. *Mol Cell Biol*. 1997;**17**(3):1281–8.
- Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*. 1999;**98**(3):387–96.
- Klenova EM, Nicolas RH, Paterson HF, et al. CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Mol Cell Biol*. 1993;**13**(12):7612–24.
- Robinett CC, O'Connor A, Dunaway M. The repeat organizer, a specialized insulator element within the intergenic spacer of the *Xenopus* rRNA genes. *Mol Cell Biol*. 1997;**17**(5):2866–75.
- Filippova GN, Fagerlie S, Klenova EM, et al. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol Cell Biol*. 1996;**16**(6):2802–13.
- Ohlsson R, Renkawitz R, Lobanenkov V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet*. 2001;**17**(9):520–7.
- Klenova EM, Chernukhin IV, El-Kady A, et al. Functional phosphorylation sites in the C-terminal region of the multivalent multifunctional transcriptional factor CTCF. *Mol Cell Biol*. 2001;**21**(6):2221–34.
- Klenova E, Ohlsson R. Poly(ADP-ribosylation) and epigenetics. Is CTCF PART of the plot? *Cell Cycle*. 2005;**4**(1):96–101.
- Kitchen NS, Schoenherr CJ. Sumoylation modulates a domain in CTCF that activates transcription and decondenses chromatin. *J Cell Biochem*. 2010;**111**(3):665–75.
- MacPherson MJ, Beatty LG, Zhou W, Du M, Sadowski PD. The CTCF insulator protein is posttranslationally modified by SUMO. *Mol Cell Biol*. 2009;**29**(3):714–25.
- Wang J, Wang Y, Lu L. De-SUMOylation of CCCTC binding factor (CTCF) in hypoxic stress-induced human corneal epithelial cells. *J Biol Chem*. 2012;**287**(15):12469–79.
- Heger P, Marin B, Bartkuhn M, Schierenberg E, Wiehe T. The chromatin insulator CTCF and the emergence of metazoan diversity. *Proc Natl Acad Sci USA*. 2012;**109**(43):17507–12.
- Heath H, Ribeiro de Almeida C, Sleutels F, et al. CTCF regulates cell cycle progression of alphabeta T cells in the thymus. *EMBO J*. 2008;**27**(21):2839–50.
- Splinter E, Heath H, Kooren J, et al. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev*. 2006;**20**(17):2349–54.
- Fedoriw AM, Stein P, Svoboda P, Schultz RM, Bartolomei MS. Transgenic RNAi reveals essential function for CTCF in *H19* gene imprinting. *Science*. 2004;**303**(5655):238–40.
- Barski A, Cuddapah S, Cui K, et al. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007;**129**(4):823–37.
- Chen X, Xu H, Yuan P, et al. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*. 2008;**133**(6):1106–17.
- Kim TH, Abdullaev ZK, Smith AD, et al. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell*. 2007;**128**(6):1231–45.
- Whittington T, Perkins AC, Bailey TL. High-throughput chromatin information enables accurate tissue-specific prediction of transcription factor binding sites. *Nucleic Acids Res*. 2009;**37**(1):14–25.

26. Rhee HS, Pugh BF. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell*. 2011; **147**(6):1408–19.
27. Nakahashi H, Kwon KR, Resch W, et al. A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell Rep*. 2013; **3**(5):1678–89.
28. Ziebarth JD, Bhattacharya A, Cui Y. CTCFBSDB 2.0: a database for CTCF-binding sites and genome organization. *Nucleic Acids Res*. 2013; **41**(Database issue):D188–94.
29. Engel N, West AG, Felsenfeld G, Bartolomei MS. Antagonism between DNA hypermethylation and enhancer-blocking activity at the H19 DMD is uncovered by CpG mutations. *Nat Genet*. 2004; **36**(8):883–8.
30. Renda M, Baglivo I, Burgess-Beusse B, et al. Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci. *J Biol Chem*. 2007; **282**(46):33336–45.
31. Vostrov AA, Quitschke WW. The zinc finger protein CTCF binds to the APBbeta domain of the amyloid beta-protein precursor promoter. Evidence for a role in transcriptional activation. *J Biol Chem*. 1997; **272**(52):33353–9.
32. Klug A. The discovery of zinc fingers and their development for practical applications in gene regulation and genome manipulation. *Q Rev Biophys*. 2010; **43**(1):1–21.
33. MacPherson MJ, Sadowski PD. The CTCF insulator protein forms an unusual DNA structure. *BMC Mol Biol*. 2010; **11**:101.
34. Handoko L, Xu H, Li G, et al. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet*. 2011; **43**(7): 630–8.
35. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*. 2000; **405**(6785):482–5.
36. Filippova GN, Thiens CP, Penn BH, et al. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat Genet*. 2001; **28**(4):335–43.
37. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*. 2000; **405**(6785):486–9.
38. Kanduri C, Pant V, Loukinov D, et al. Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr Biol*. 2000; **10**(14):853–6.
39. Pant V, Mariano P, Kanduri C, et al. The nucleotides responsible for the direct physical contact between the chromatin insulator protein CTCF and the H19 imprinting control region manifest parent of origin-specific long-distance insulation and methylation-free domains. *Genes Dev*. 2003; **17**(5):586–90.
40. Wang H, Maurano MT, Qu H, et al. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res*. 2012; **22**(9):1680–8.
41. Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell*. 2004; **13**(2):291–8.
42. Ishihara K, Oshimura M, Nakao M. CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol Cell*. 2006; **23**(5):733–42.
43. Chernukhin I, Shamsuddin S, Kang SY, et al. CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Mol Cell Biol*. 2007; **27**(5):1631–48.
44. Donohoe ME, Zhang LF, Xu N, Shi Y, Lee JT. Identification of a Ctf cofactor, Yy1, for the X chromosome binary switch. *Mol Cell*. 2007; **25**(1):43–56.
45. Chernukhin IV, Shamsuddin S, Robinson AF, et al. Physical and functional interaction between two pluripotent proteins, the Y-box DNA/RNA-binding factor, YB-1, and the multivalent zinc finger factor, CTCF. *J Biol Chem*. 2000; **275**(38):29915–21.
46. De La Rosa-Velazquez IA, Rincon-Arango H, Benitez-Bribiesca L, Recillas-Targa F. Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. *Cancer Res*. 2007; **67**(6):2577–85.
47. Lutz M, Burke LJ, Barreto G, et al. Transcriptional repression by the insulator protein CTCF involves histone deacetylases. *Nucleic Acids Res*. 2000; **28**(8):1707–13.
48. Guastafierro T, Cecchinelli B, Zampieri M, et al. CCCTC-binding factor activates PARP-1 affecting DNA methylation machinery. *J Biol Chem*. 2008; **283**(32):21873–80.
49. Yao H, Brick K, Evrard Y, Xiao T, Camerini-Otero RD, Felsenfeld G. Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev*. 2010; **24**(22):2543–55.
50. Zlatanova J, Caijafa P. CCCTC-binding factor: to loop or to bridge. *Cell Mol Life Sci*. 2009; **66**(10):1647–60.
51. Hirano T. At the heart of the chromosome: SMC proteins in action. *Nat Rev Mol Cell Biol*. 2006; **7**(5):311–22.
52. Ishiguro K, Watanabe Y. Chromosome cohesion in mitosis and meiosis. *J Cell Sci*. 2007; **120**(Pt 3):367–9.
53. Peters JM, Tedeschi A, Schmitz J. The cohesin complex and its roles in chromosome biology. *Genes Dev*. 2008; **22**(22):3089–114.
54. Haering CH, Lowe J, Hochwagen A, Nasmyth K. Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol Cell*. 2002; **9**(4):773–88.
55. Degner SC, Verma-Gaur J, Wong TP, et al. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proc Natl Acad Sci U S A*. 2011; **108**(23):9566–71.
56. Hadjur S, Williams LM, Ryan NK, et al. Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature*. 2009; **460**(7253):410–3.
57. Kagey MH, Newman JJ, Bilodeau S, et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature*. 2010; **467**(7314):430–5.
58. Parelho V, Hadjur S, Spivakov M, et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*. 2008; **132**(3):422–33.
59. Rubio ED, Reiss DJ, Welcsh PL, et al. CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci USA*. 2008; **105**(24): 8309–14.
60. Wendt KS, Yoshida K, Itoh T, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*. 2008; **451**(7180):796–801.
61. Xiao T, Wallace J, Felsenfeld G. Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol Cell Biol*. 2011; **31**(11):2174–83.
62. Herold M, Bartkuhn M, Renkawitz R. CTCF: insights into insulator function during development. *Development*. 2012; **139**(6): 1045–57.
63. Magdiner F, Yusufzai TM, Felsenfeld G. Both CTCF-dependent and -independent insulators are found between the mouse T cell receptor alpha and Dad1 genes. *J Biol Chem*. 2004; **279**(24): 25381–9.
64. Chambeyron S, Bickmore WA. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev*. 2004; **18**(10):1119–30.
65. Ferrai C, Xie SQ, Luraghi P, et al. Poised transcription factories prime silent uPA gene prior to activation. *PLoS Biol*. 2010; **8**(1): e1000270.
66. Mahy NL, Perry PE, Bickmore WA. Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *J Cell Biol*. 2002; **159**(5): 753–63.
67. Osborne CS, Chakalova L, Brown KE, et al. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet*. 2004; **36**(10):1065–71.

68. Wei Z, Huang D, Gao F, et al. Biological implications and regulatory mechanisms of long-range chromosomal interactions. *J Biol Chem*. 2013;**288**(31):22369–77.
69. Bodnar JW. A domain model for eukaryotic DNA organization: a molecular basis for cell differentiation and chromosome evolution. *J Theor Biol*. 1988;**132**(4):479–507.
70. Goldman MA. The chromatin domain as a unit of gene regulation. *Bioessays*. 1988;**9**(2–3):50–5.
71. Angueta E, Johnson CA, Wood WG, Turner BM, Higgs DR. Identification of a conserved erythroid specific domain of histone acetylation across the alpha-globin gene cluster. *Proc Natl Acad Sci U S A*. 2001;**98**(21):12114–9.
72. Krajewski WA, Becker PB. Reconstitution of hyperacetylated, DNase I-sensitive chromatin characterized by high conformational flexibility of nucleosomal DNA. *Proc Natl Acad Sci USA*. 1998;**95**(4):1540–5.
73. Forrester WC, Epner E, Driscoll MC, et al. A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. *Genes Dev*. 1990;**4**(10):1637–49.
74. Grosveld F, van Assendelft GB, Greaves DR, Kollias G. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell*. 1987;**51**(6):975–85.
75. Razin SV, Farrell CM, Recillas-Targa F. Genomic domains and regulatory elements operating at the domain level. *Int Rev Cytol*. 2003;**226**:63–125.
76. Felsenfeld G. Chromatin structure and the expression of globin-encoding genes. *Gene*. 1993;**135**(1–2):119–24.
77. Hebbes TR, Clayton AL, Thorne AW, Crane-Robinson C. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *EMBO J*. 1994;**13**(8):1823–30.
78. Litt MD, Simpson M, Recillas-Targa F, Prioleau MN, Felsenfeld G. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J*. 2001;**20**(9):2224–35.
79. Kurukuti S, Tiwari VK, Tavoosidana G, et al. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc Natl Acad Sci U S A*. 2006;**103**(28):10684–9.
80. Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de Laat W. The beta-globin nuclear compartment in development and erythroid differentiation. *Nat Genet*. 2003;**35**(2):190–4.
81. Ulianov SV, Gavrilov AA, Razin SV. Spatial organization of the chicken beta-globin gene domain in erythroid cells of embryonic and adult lineages. *Epigenetics Chromatin*. 2012;**5**(1):16.
82. Wang L, Di LJ, Lv X, et al. Inter-MAR association contributes to transcriptionally active looping events in human beta-globin gene cluster. *PLoS One*. 2009;**4**(2):e4629.
83. Yusufzai TM, Felsenfeld G. The 5'-HS4 chicken beta-globin insulator is a CTCF-dependent nuclear matrix-associated element. *Proc Natl Acad Sci USA*. 2004;**101**(23):8620–4.
84. Guo Y, Monahan K, Wu H, et al. CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter choice. *Proc Natl Acad Sci U S A*. 2012;**109**(51):21081–6.
85. Mitchell JA, Fraser P. Transcription factories are nuclear sub-compartments that remain in the absence of transcription. *Genes Dev*. 2008;**22**(1):20–5.
86. Hu Q, Kwon YS, Nunez E, et al. Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. *Proc Natl Acad Sci U S A*. 2008;**105**(49):19199–204.
87. Li G, Ruan X, Auerbach RK, et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell*. 2012;**148**(1–2):84–98.
88. Patel B, Kang Y, Cui K, et al. Aberrant TAL1 activation is mediated by an interchromosomal interaction in human T-cell acute lymphoblastic leukemia. *Leukemia*. 2014;**28**(2):349–61.
89. West AG, Gaszner M, Felsenfeld G. Insulators: many functions, many mechanisms. *Genes Dev*. 2002;**16**(3):271–88.
90. Wallace JA, Felsenfeld G. We gather together: insulators and genome organization. *Curr Opin Genet Dev*. 2007;**17**(5):400–7.
91. Dunn KL, Zhao H, Davie JR. The insulator binding protein CTCF associates with the nuclear matrix. *Exp Cell Res*. 2003;**288**(1):218–23.
92. Gohl D, Aoki T, Blanton J, Shanower G, Kappes G, Schedl P. Mechanism of chromosomal boundary action: roadblock, sink, or loop? *Genetics*. 2011;**187**(3):731–48.
93. Chien R, Zeng W, Kawauchi S, et al. Cohesin mediates chromatin interactions that regulate mammalian β -globin expression. *J Biol Chem*. 2011;**286**(20):17870–8.
94. Bulger M, Bender MA, van Doorninck JH, et al. Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters. *Proc Natl Acad Sci U S A*. 2000;**97**(26):14560–5.
95. Schubeler D, Francastel C, Cimbora DM, Reik A, Martin DI, Groudine M. Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. *Genes Dev*. 2000;**14**(8):940–50.
96. Xie X, Mikkelsen TS, Gnirke A, Lindblad-Toh K, Kellis M, Lander ES. Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proc Natl Acad Sci USA*. 2007;**104**(17):7145–50.
97. Recillas-Targa F, Bell AC, Felsenfeld G. Positional enhancer-blocking activity of the chicken beta-globin insulator in transiently transfected cells. *Proc Natl Acad Sci USA*. 1999;**96**(25):14354–9.
98. Akopov SB, Ruda VM, Batrak VV, et al. Identification, genome mapping, and CTCF binding of potential insulators within the FXVD5-COX7A1 locus of human chromosome 19q13.12. *Mamm Genome*. 2006;**17**(10):1042–9.
99. Didych DA, Kotova ES, Akopov SB, Nikolaev LG, Sverdlov ED. DNA fragments binding CTCF *in vitro* and *in vivo* are capable of blocking enhancer activity. *BMC Res Notes*. 2012;**5**:178.
100. Vetchinova AS, Akopov SB, Chernov IP, Nikolaev LG, Sverdlov ED. Two-dimensional electrophoretic mobility shift assay: identification and mapping of transcription factor CTCF target sequences within an FXVD5-COX7A1 region of human chromosome 19. *Anal Biochem*. 2006;**354**(1):85–93.
101. Zhao H, Dean A. An insulator blocks spreading of histone acetylation and interferes with RNA polymerase II transfer between an enhancer and gene. *Nucleic Acids Res*. 2004;**32**(16):4903–19.
102. Chung JH, Bell AC, Felsenfeld G. Characterization of the chicken beta-globin insulator. *Proc Natl Acad Sci USA*. 1997;**94**(2):575–80.
103. Harmston N, Lenhard B. Chromatin and epigenetic features of long-range gene regulation. *Nucleic Acids Res*. 2013;**41**(15):7185–99.
104. Kulaeva OI, Nizovtseva EV, Polikanov YS, Ulianov SV, Studitsky VM. Distant activation of transcription: mechanisms of enhancer action. *Mol Cell Biol*. 2012;**32**(24):4892–7.
105. Simonis M, Kooren J, de Laat W. An evaluation of 3C-based methods to capture DNA interactions. *Nat Methods*. 2007;**4**(11):895–901.
106. Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell*. 2002;**10**(6):1453–65.
107. Tsytsykova AV, Rajsbaum R, Falvo JV, Ligeiro F, Neely SR, Goldfeld AE. Activation-dependent intrachromosomal interactions formed by the TNF gene promoter and two distal enhancers. *Proc Natl Acad Sci USA*. 2007;**104**(43):16850–5.
108. Krivega I, Dean A. Enhancer and promoter interactions-long distance calls. *Curr Opin Genet Dev*. 2012;**22**(2):79–85.
109. Geyer PK. The role of insulator elements in defining domains of gene expression. *Curr Opin Genet Dev*. 1997;**7**(2):242–8.

110. Pikaart MJ, Recillas-Targa F, Felsenfeld G. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.* 1998;**12**(18):2852–62.
111. Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res.* 2009;**19**(1):24–32.
112. Dixon JR, Selvaraj S, Yue F, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* 2012;**485**(7398):376–80.
113. Gomes NP, Espinosa JM. Gene-specific repression of the p53 target gene PUMA via intragenic CTCF-Cohesin binding. *Genes Dev.* 2010;**24**(10):1022–34.
114. Vostrov AA, Taheny MJ, Quitschke WW. A region to the N-terminal side of the CTCF zinc finger domain is essential for activating transcription from the amyloid precursor protein promoter. *J Biol Chem.* 2002;**277**(2):1619–27.
115. Ribich S, Tasic B, Maniatis T. Identification of long-range regulatory elements in the protocadherin-alpha gene cluster. *Proc Natl Acad Sci USA.* 2006;**103**(52):19719–24.
116. Kehayova P, Monahan K, Chen W, Maniatis T. Regulatory elements required for the activation and repression of the protocadherin-alpha gene cluster. *Proc Natl Acad Sci USA.* 2011;**108**(41):17195–200.
117. Monahan K, Rudnick ND, Kehayova PD, et al. Role of CCCTC binding factor (CTCF) and cohesin in the generation of single-cell diversity of protocadherin- α gene expression. *Proc Natl Acad Sci USA.* 2012;**109**(23):9125–30.
118. Lutz M, Burke LJ, LeFevre P, et al. Thyroid hormone-regulated enhancer blocking: cooperation of CTCF and thyroid hormone receptor. *EMBO J.* 2003;**22**(7):1579–87.
119. Lai AY, Fatemi M, Dhasarathy A, et al. DNA methylation prevents CTCF-mediated silencing of the oncogene BCL6 in B cell lymphomas. *J Exp Med.* 2010;**207**(9):1939–50.
120. Renaud S, Loukinov D, Bosman FT, Lobanekov V, Benhattar J. CTCF binds the proximal exonic region of hTERT and inhibits its transcription. *Nucleic Acids Res.* 2005;**33**(21):6850–60.
121. Renaud S, Loukinov D, Abdullaev Z, et al. Dual role of DNA methylation inside and outside of CTCF-binding regions in the transcriptional regulation of the telomerase hTERT gene. *Nucleic Acids Res.* 2007;**35**(4):1245–56.
122. Mendez-Catala CF, Gretton S, Vostrov A, et al. A novel mechanism for CTCF in the epigenetic regulation of Bax in breast cancer cells. *Neoplasia.* 2013;**15**(8):898–912.
123. Rayess H, Wang MB, Srivatsan ES. Cellular senescence and tumor suppressor gene p16. *Int J Cancer.* 2012;**130**(8):1715–25.
124. Huang K, Jia J, Wu C, et al. Ribosomal RNA gene transcription mediated by the master genome regulator protein CCCTC-binding factor (CTCF) is negatively regulated by the condensin complex. *J Biol Chem.* 2013;**288**(36):26067–77.
125. Kanduri C, Holmgren C, Pilartz M, et al. The 5' flank of mouse H19 in an unusual chromatin conformation unidirectionally blocks enhancer-promoter communication. *Curr Biol.* 2000;**10**(8):449–57.
126. Gabory A, Jammes H, Dandolo L. The H19 locus: role of an imprinted non-coding RNA in growth and development. *Bioessays.* 2010;**32**(6):473–80.
127. Leighton PA, Saam JR, Ingram RS, Stewart CL, Tilghman SM. An enhancer deletion affects both H19 and Igf2 expression. *Genes Dev.* 1995;**9**(17):2079–89.
128. Lewis A, Murrell A. Genomic imprinting: CTCF protects the boundaries. *Curr Biol.* 2004;**14**(7):R284–6.
129. Schoenherr CJ, Levorse JM, Tilghman SM. CTCF maintains differential methylation at the Igf2/H19 locus. *Nat Genet.* 2003;**33**(1):66–9.
130. Szabo PE, Tang SH, Silva FJ, Tsark WM, Mann JR. Role of CTCF binding sites in the Igf2/H19 imprinting control region. *Mol Cell Biol.* 2004;**24**(11):4791–800.
131. Gao ZH, Suppola S, Liu J, Heikkila P, Janne J, Voutilainen R. Association of H19 promoter methylation with the expression of H19 and IGF-II genes in adrenocortical tumors. *J Clin Endocrinol Metab.* 2002;**87**(3):1170–6.
132. Chao W, Huynh KD, Spencer RJ, Davidow LS, Lee JT. CTCF, a candidate trans-acting factor for X-inactivation choice. *Science.* 2002;**295**(5553):345–7.
133. Lee JT. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes Dev.* 2009;**23**(16):1831–42.
134. Avner P, Heard E. X-chromosome inactivation: counting, choice and initiation. *Nat Rev Genet.* 2001;**2**(1):59–67.
135. Clerc P, Avner P. Random X-chromosome inactivation: skewing lessons for mice and men. *Curr Opin Genet Dev.* 2006;**16**(3):246–53.
136. Erwin JA, Lee JT. New twists in X-chromosome inactivation. *Curr Opin Cell Biol.* 2008;**20**(3):349–55.
137. Bacher CP, Guggiari M, Brors B, et al. Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. *Nat Cell Biol.* 2006;**8**(3):293–9.
138. Xu N, Tsai CL, Lee JT. Transient homologous chromosome pairing marks the onset of X inactivation. *Science.* 2006;**311**(5764):1149–52.
139. Xu N, Donohoe ME, Silva SS, Lee JT. Evidence that homologous X-chromosome pairing requires transcription and CTCF protein. *Nat Genet.* 2007;**39**(11):1390–6.
140. Nicodemi M, Panning B, Prisco A. A thermodynamic switch for chromosome colocalization. *Genetics.* 2008;**179**(1):717–21.
141. Nicodemi M, Prisco A. Self-assembly and DNA binding of the blocking factor in x chromosome inactivation. *PLoS Comput Biol.* 2007;**3**(11):e210.
142. Nicodemi M, Prisco A. Symmetry-breaking model for X-chromosome inactivation. *Phys Rev Lett.* 2007;**98**(10):108104.
143. Sun BK, Deaton AM, Lee JT. A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Mol Cell.* 2006;**21**(5):617–28.
144. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science.* 2008;**322**(5902):750–6.
145. Hoki Y, Kimura N, Kanbayashi M, et al. A proximal conserved repeat in the Xist gene is essential as a genomic element for X-inactivation in mouse. *Development.* 2009;**136**(1):139–46.
146. Wutz A, Rasmussen TP, Jaenisch R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet.* 2002;**30**(2):167–74.
147. Pugacheva EM, Tiwari VK, Abdullaev Z, et al. Familial cases of point mutations in the XIST promoter reveal a correlation between CTCF binding and pre-emptive choices of X chromosome inactivation. *Hum Mol Genet.* 2005;**14**(7):953–65.
148. Filippova GN, Cheng MK, Moore JM, et al. Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Dev Cell.* 2005;**8**(1):31–42.
149. Shukla S, Kavak E, Gregory M, et al. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature.* 2011;**479**(7371):74–9.
150. Wada Y, Ohta Y, Xu M, et al. A wave of nascent transcription on activated human genes. *Proc Natl Acad Sci USA.* 2009;**106**(43):18357–61.
151. Kornblihtt AR. CTCF: from insulators to alternative splicing regulation. *Cell Res.* 2012;**22**(3):450–2.

Received 15.03.14