

УДК 577.21

СТСФ как пример ДНК-связывающих транскрипционных факторов, содержащих кластеры цинковых пальцев С2Н2-типа

О. Г. Максименко^{1,2*}, Д. В. Фурсенко¹, Е. В. Белова^{1,2}, П. Г. Георгиев^{1*}¹Институт биологии гена РАН, Москва, 119334 Россия²Центр высокоточного редактирования и генетических технологий для биомедицины ИБГ РАН, Москва, 119334 Россия

*E-mail: maksog@mail.ru; georgiev_p@mail.ru

Поступила в редакцию 17.09.2020

Принята к печати 12.11.2020

DOI: 10.32607/actanaturae.11206

РЕФЕРАТ В клетках млекопитающих большая часть границ топологически ассоциированных доменов и всех хорошо изученных инсуляторов обогащена сайтами связывания белка СТСФ. Согласно экспериментальным данным, СТСФ является ключевым фактором в организации архитектуры хромосом млекопитающих. Присутствие в центральной части СТСФ кластера из 11 доменов цинковых пальцев С2Н2-типа, пять из которых специфически связываются с протяженной, консервативной у большей части животных последовательностью ДНК, определяет основную часть функциональных свойств СТСФ как архитектурного белка. Класс транскрипционных факторов, содержащих кластер цинковых пальцев С2Н2-типа из пяти и более доменов (С2Н2-белки), широко представлен во всех группах животных. В настоящее время функции подавляющей части С2Н2-белков остаются неизвестными. На примере СТСФ позвоночных и нескольких С2Н2-белков дрозофилы в обзоре рассмотрено строение и вероятные архитектурные функции этих белков. **КЛЮЧЕВЫЕ СЛОВА** цинковые пальцы С2Н2-типа, архитектурные белки, регуляция транскрипции, инсуляторы, ТАД, энхансеры, промоторы, СТСФ.

СПИСОК СОКРАЩЕНИЙ ТАД – топологически ассоциированный домен; СТСФ – СССТС-связывающий фактор; С2Н2 – домен, состоящий из двух остатков цистеина и двух остатков гистидина, координированных ионом цинка.

ВВЕДЕНИЕ

Значительное усложнение процессов регуляции экспрессии генов у высших эукариот стало следствием дифференцировки клеток в сложно устроенных организмах. Известно, что специализация клеток определяется репертуарами транскрипционных факторов, собирающихся на геномных регуляторных элементах. Гены, ответственные за дифференцировку клеток, обычно регулируются энхансерами, каждый из которых активирует промотор в течение заданного временного промежутка в определенной группе клеток [1–3]. В ряде случаев транскрипция генов развития регулируется десятками энхансеров, часть из которых может быть удалена от регулируемого промотора на расстояния, достигающие сотен тысяч пар нуклеотидов.

Способность энхансеров стимулировать промоторы на больших расстояниях привела к предположению о том, что должны существовать особые

транскрипционные домены, или компартменты, внутри которых устанавливаются взаимодействия между энхансерами и промоторами [4]. Считалось, что на границе таких транскрипционных доменов должны находиться специальные регуляторные элементы, способные блокировать взаимодействия между энхансерами и промоторами, расположенными в соседних доменах [5, 6]. Согласно наиболее распространенному мнению, границы доменов взаимодействуют либо между собой, либо с ядерными структурами, прикрепленными к ядерной оболочке. Действительно, регуляторные элементы, имеющие предсказанные свойства, нашли сначала у дрозофилы, а потом у млекопитающих, и назвали инсуляторами [7]. С использованием модельных систем в трансгенных линиях дрозофилы описаны два основных свойства инсуляторов: способность блокировать взаимодействие между энхансерами и промоторами и защищать экспрессию трансгена от репрессии

при его интеграции в гетерохроматиновые области генома [5, 6].

Появление полногеномных методов анализа, позволяющих идентифицировать всю совокупность взаимодействий между участками хроматина *in vivo*, и микроскопии высокого разрешения [8–11] вывело изучение пространственной организации генома на новый уровень. Оказалось, что хромосомы всех эукариот организованы в топологически ассоциированные домены (ТАДы), которые часто формируются за счет предпочтительного взаимодействия друг с другом концов, или границ, домена [12–15]. Контакты внутри ТАДа устанавливаются с более высокой частотой по сравнению с контактами между ТАДами.

Открытие ТАДов дало основание предполагать, что их границы соответствуют инсуляторам, ограничивающим независимые регуляторные домены [16–18]. Однако исследования, проведенные прежде всего на единичных клетках, показали, что границы ТАДов формируются как совокупность предпочтительных контактов и не представляют собой строгих физических барьеров, блокирующих любые транс-взаимодействия между регуляторными элементами, расположенными в разных ТАДах [12, 14, 19, 20]. Большая часть охарактеризованных инсуляторов, взаимодействующих энхансеров и промоторов содержится внутри одного ТАДа, что, вероятно, способствует формированию его пространственной структуры. Увеличение разрешения карт пространственных взаимодействий внутри ТАДов привело к обнаружению субдоменов, которые обычно соответствуют локальным контактам между регуляторными элементами [19].

СТСФ КАК НАИБОЛЕЕ ИЗУЧЕННЫЙ БЕЛОК С КЛАСТЕРОМ ЦИНКОВЫХ ПАЛЬЦЕВ С2Н2-ТИПА У МЛЕКОПИТАЮЩИХ

Белок СТСФ (ССТС binding factor) позвоночных, хорошо изученный у человека и мыши [21, 22], экспрессируется на всех стадиях развития во всех типах клеток и необходим в процессе эмбрионального развития. В зависимости от контекста СТСФ может выступать в роли активатора или репрессора транскрипции, он принимает участие в инактивации одной из X-хромосом у млекопитающих, регулирует альтернативный сплайсинг пре-мРНК некоторых генов, контролирует импринтинг, участвует в рекомбинации и репарации, определяет активность энхансеров, промоторов и инсуляторов. Однако наиболее детально описана ключевая роль белка СТСФ позвоночных в организации архитектуры хромосом [23–25]. В геноме млекопитающих присутствует от 40000 до 80000 сайтов связывания СТСФ, из которых более 5000

сайтов консервативны между различными видами и клеточными линиями [21, 26]. Приблизительно 50% сайтов связывания СТСФ находятся в межгенных областях, ~ 15% расположены вблизи промоторов, а ~ 35% внутригенные (30% в интронах и 5% в экзонах) [27].

СТСФ млекопитающих состоит из неструктурированных концевых областей и расположенного в центральной части кластера 11 цинковых пальцев, из которых первые 10 относятся к типу С2Н2, а последний – С2НС. Стоит отметить, что белки, содержащие один и, реже, несколько кластеров доменов С2Н2, составляют значительную часть от общего количества С2Н2-белков [28].

Классический С2Н2-домен имеет консенсусную последовательность $CX_{2-4}CX_{12}HX_{2-8}H$. В присутствии иона цинка эта последовательность складывается в $\beta\alpha$ -структуру, в которой цинк тетраэдрически координируется двумя цистеинами на одном конце β -листа и двумя гистидинами в С-концевой части α -спирали. Структура стабилизируется гидрофобными связями. В каноническом комплексе тандемные цинковые пальцы С2Н2 погружены в большую бороздку ДНК α -спиральными участками. Высокоаффинное и специфичное связывание опосредовано специфическими взаимодействиями с азотистыми основаниями и неспецифическими контактами с фосфатным остовом ДНК. К любому триплету ДНК можно подобрать специфически узнающий его С2Н2-домен, который будет содержать подходящие аминокислоты в ключевых позициях α -спирали [29–31]. Поэтому уже через несколько лет после описания первой структуры С2Н2-доменов с ДНК химерные белки, состоящие из кластера С2Н2-доменов и домена FokI, вносящего двухцепочечные разрывы в последовательность ДНК, стали активно использоваться в качестве сайт-специфичных эндонуклеаз для внесения целевых изменений в геном [32, 33].

В белках с кластером цинковых пальцев С2Н2-типа короткие пятиаминокислотные линкеры, располагающиеся между доменами, имеют консенсусную последовательность TGEKP и являются характерной особенностью ДНК-связывающих С2Н2-белков [34]. Линкеры оказывают критическое влияние на аффинность и специфичность связывания с ДНК, а мутации в них могут приводить к потере функции белка *in vivo* [35, 36]. Считается, что каждый аминокислотный остаток линкера играет свою роль при взаимодействии с ДНК. Белковая структура из нескольких С2Н2-доменов, гибкая в свободном состоянии, «защелкивается», как только встречает правильную последовательность ДНК. При этом ОН-группа первого треонина Т1 (или серина) образует водородную связь с амидной группой глутаминовой кислоты Е3,

а глицин G2 обеспечивает необходимую для защелкивания гибкость основной цепи. Глутаминовая кислота E3 может играть роль в стабилизации контактов между цинковыми пальцами. Лизин K4 (или аргинин) образует контакт с фосфатным остовом ДНК. Пролин P5, вероятно, укрепляет связь между линкером и последующим цинковым пальцем, при этом он фиксирует следующий за ним консервативный остаток фенилаланина или тирозина, ароматическое кольцо которых укладывается на N-конец α -спирали [37]. В белке CTCF человека ДНК-связывающие C2H2-домены также соединены TGEKP-подобными линкерами (рис. 1).

Конформационные изменения в структуре ДНК, вносимые C2H2-доменами при образовании комплекса, ограничивают число взаимодействующих с ДНК C2H2-доменов, соединенных короткими линкерами, и, соответственно, длину канонического сайта связывания [37]. Вероятно, поэтому в большей части белков только 4–5 C2H2-доменов участвуют во взаимодействии и специфичном узнавании сайта ДНК длиной 12–15 п.н. В исследованиях с искусственно созданными C2H2-кластерами показано, что специфичность связывания белка с ДНК повышается, когда несколько коротких ДНК-узнающих кластеров C2H2-доменов соединены более длинными неканоническими линкерами [28]. Поэтому можно предположить, что белки, содержащие большое число C2H2-доменов в кластере, могут специфично узнавать сайты ДНК с разной последовательностью.

В составе CTCF человека 3–7 C2H2-домены отвечают за специфичное связывание с консенсусной последовательностью длиной 15 п.н. (рис. 1) [38]. C2H2-домен 8, находясь вне большой бороздки, не участвует в узнавании азотистых оснований ДНК и поэтому может служить мостиком, соединяющим 3–7 C2H2-домены, узнающие основной мотив, с 9–11 C2H2-доменами, которые могут специфично связываться с дополнительным ДНК-мотивом, присутствующим примерно в 15% сайтов связывания CTCF [39, 40]. 1–2 C2H2-домены также могут связываться с неконсервативной последовательностью ДНК [39]. Таким образом, разные комбинации C2H2-доменов белка CTCF могут с разной эффективностью связываться с широким спектром мотивов [41, 42].

В экспериментах *in vitro* показано, что связывание CTCF с ДНК ингибируется при метилировании цитозина в положении 2 в консенсусном сайте, в то время как метилирование цитозина в положении 12 практически не имеет эффекта. В положении 2 цитозин узнается остатком аспарагиновой кислоты, предпочитающим немодифицированное основание, а в положении 12 – остатком глутаминовой кислоты, демонстрирующим незначительное увеличение

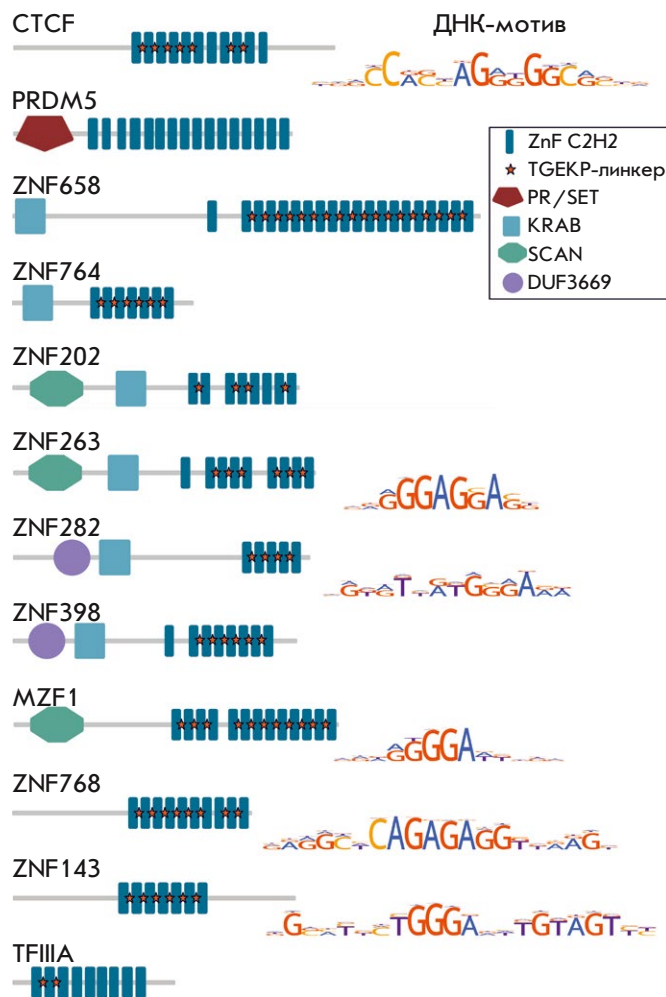


Рис. 1. C2H2-белки позвоночных с архитектурными функциями. Показаны доменная организация описанных белков и известные мотивы связывания

аффинности связывания с метилированным основанием [38]. Кроме того, важную роль в распознавании метильной группы играет остаток аргинина, образующий в комплексе с ДНК триаду 5-метилцитозин–аргинин–гуанин, характерную для всех комплексов C2H2-белков с метилированной ДНК [43, 44].

Метилирование цитозина в сайтах связывания может усилить, ослабить или полностью заблокировать связывание C2H2-белков с ДНК, т.е. является глобальным механизмом регуляции активности промоторов, энхансеров и инсультаторов [45]. Наиболее яркий пример роли метилирования сайтов связывания C2H2-белков – участие CTCF в геномном импринтинге, эпигенетическом механизме регуляции экспрессии аллелей одного гена, в зависимости от их происхождения – мужского или женского [46]. Импринтинг осуществляется с помощью специальных регулятор-

ных элементов, получивших название дифференциально метилированных областей (ДМО), которые часто содержат сайты связывания CTCF. Наиболее хорошо описан импринтинг на генах *Igf2* и *H19*, которые активируются группой расположенных рядом энхансеров. ДМО, выполняющий функцию инсультатора, находится между геном *Igf2* и энхансерами и состоит из четырех сайтов связывания CTCF, имеющих цитозин во втором положении. Метилирование ДМО поддерживается только в локусе *Igf2/H19* отцовского происхождения, что приводит к потере связывания CTCF и к активации гена *Igf2*. При этом в материнском локусе CTCF связывается со своими сайтами в ДМО, что приводит к блокированию взаимодействия между энхансерами и геном *Igf2*.

Метилирование сайтов связывания транскрипционных факторов и, в частности С2Н2-белков, также может участвовать в процессе инактивации транскрипции в одной из двух Х-хромосом млекопитающих [47].

Кластеры С2Н2-доменов могут участвовать в специфичном и неспецифичном взаимодействии с РНК [48, 49]. Наиболее хорошо исследовано специфичное взаимодействие белка TFIIIA с 5S РНК. Показано, что 1–3, 5 и 7–9 С2Н2-домены связываются с ДНК-мотивами в промоторной области гена 5S РНК, а 4, 5 и 6 С2Н2-домены взаимодействуют с 5S РНК. Таким образом, с одной стороны, С2Н2-домены 4 и 6 являются линкерами, расширяющими возможности связывания белка TFIIIA с ДНК, а с другой стороны, они, взаимодействуя с синтезируемой 5S РНК, стабилизируют ее при экспорте из ядра в цитоплазму до сборки рибосомы.

Два С2Н2-домена, 1 и 10, отвечают за неспецифичное взаимодействие CTCF с большим спектром РНК [50, 51]. При этом нарушение структуры С2Н2-доменов в результате мутации остатка гистидина не влияет на связывание РНК. Этот результат предполагает важную роль отдельных аминокислот С2Н2-доменов в связывании РНК, а не структуры цинкового пальца в целом. Существуют экспериментальные данные, показывающие, что взаимодействие CTCF с РНК может приводить к его мультимеризации, однако механизм этого процесса остается неизвестным [50, 52]. Так как достаточно большая часть сайтов CTCF находится в интронах генов, можно ожидать, что CTCF, неспецифично связываясь с РНК, участвует в регуляции процессов сплайсинга и терминации синтеза пре-мРНК, которые протекают сопряженно с этапами транскрипции. Например, CTCF способен замедлять движение РНК-полимеразы II, тем самым приводя к выбору альтернативного экзона при сплайсинге [53, 54] или альтернативного сигнала полиаденилирования

при терминации транскрипции [55]. В С-концевом домене белка CTCF картирован домен, способный взаимодействовать с РНК-полимеразой II, что может объяснить эффект ее торможения при движении через сайты связывания CTCF [56].

Большое количество экспериментальных данных показывает, что отдельные С2Н2-домены или их кластеры участвуют в белок-белковых взаимодействиях [34]. Однако детальные механизмы этих процессов и их специфичность почти неизучены. С2Н2-домены часто взаимодействуют с белковыми комплексами, ремоделирующими хроматин и вносящими модификации в гистоны. Согласно данным мутационного анализа, в таких взаимодействиях могут участвовать (в отличие от связывания с ДНК) любые аминокислоты в составе С2Н2-доменов и соединяющих их линкеров, поэтому предполагается, что С2Н2-домены в некоторых случаях могут, находясь в связанном с ДНК состоянии, привлекать на хроматин регуляторные комплексы.

Кластер С2Н2-доменов – единственная консервативная часть белка CTCF, имеющая высокий уровень гомологии у большей части позвоночных, насекомых и некоторых нематод [57–59]. Белок CTCF не найден у растений, дрожжей и круглых червей. Распределение сайтов связывания CTCF в геноме также характеризуется некоторой степенью консервативности, в частности, они найдены на границах регуляторных доменов гомеозисных генов млекопитающих, рыб и дрозофилы [60, 61], где CTCF выполняет инсультаторные функции, разграничивая зоны действия энхансеров, расположенных в соседних доменах [62–66]. Стоит отметить, что сайты связывания CTCF найдены в повторяющихся элементах генома млекопитающих, что, возможно, в процессе эволюции стало отправной точкой экспансии сайтов связывания CTCF в межгенных областях, в которых располагаются границы ТАДов [26, 67].

Несмотря на отсутствие участков гомологии, N-концевые домены белка CTCF у девяти видов животных разных классов представлены неструктурированными гомодимеризующимися доменами [68]. Делеция димеризующегося домена в составе CTCF дрозофилы приводит к значительному снижению функциональной активности мутантного CTCF [69]. На эмбриональных стволовых клетках мыши выявлено участие N-концевого домена в специфичном связывании CTCF со своими сайтами [70]. На N-конце CTCF человека между гомодимеризующимся доменом и С2Н2-кластером находится мотив YxF, необходимый для взаимодействия с когезиновым субкомплексом SA2–SCC1 [71]. Аналогичный мотив найден и у CTCF других видов животных. Таким образом, N-концевые домены белка CTCF разных ор-

ганизмов обладают общими характерными чертами строения, несмотря на отсутствие выраженной гомологии.

Ранее на С-конце белка CTCF *in vitro* картировали участок, который взаимодействует непосредственно с субъединицей SA2 когезинового комплекса [72], но в новом исследовании этот результат не был подтвержден [71].

Функции С2Н2-белков во многом определяются белками, с которыми они взаимодействуют. Достоверно выявлено более 90 белков-партнеров CTCF [73, 74], однако механизмы и специфичность таких взаимодействий остаются под вопросом. Большая часть белок-белковых взаимодействий локализована в кластере С2Н2-доменов и в неструктурированной С-концевой части белка CTCF. Вероятно, многие С2Н2-белки способны взаимодействовать с одинаковыми белковыми комплексами за счет С2Н2-доменов. Показано, что CTCF взаимодействует непосредственно с каталитической субъединицей BRG1 из комплекса ремоделирования хроматина SWI/SNF [74] и общим фактором транскрипции II-I (TFII-I) [75]. Таким образом, наиболее вероятная функция CTCF в области промоторов активно транскрибирующихся генов состоит в участии в наведении зоны открытого хроматина путем привлечения комплекса SWI/SNF, увеличивающего подвижность нуклеосом. Также CTCF может участвовать в стабилизации на промоторах TFIIID-комплекса, частью которого является TFII-I. При инактивации CTCF заметное падение уровня экспрессии детектируется только у тех генов, в промоторных областях которых расположены сайты связывания CTCF [76]. Таким образом, одна из основных функций CTCF – организация активных промоторов. Интересно, что CTCF, как и многие другие С2Н2-белки, содержит участки, обогащенные пролином и кислыми аминокислотами, что характерно для активаторов транскрипции, привлекающих на хроматин транскрипционные комплексы.

В С-концевой части CTCF идентифицирован домен, который взаимодействует с РНК-хеликазами DEAD [74, 77], что может быть связано с потенциальным участием CTCF в регуляции сплайсинга и терминации транскрипции. Для реализации функций CTCF важно его прямое взаимодействие с топоизомеразой II (Top2) [78], которая регулирует топологию хроматина, внося АТР-зависимые двойные разрывы в ДНК. Белок Top2 обнаружен примерно на половине сайтов связывания CTCF [78]. Активность Top2 наиболее часто реализуется в непосредственной близости от сайтов связывания CTCF [79]. Предполагается, что Top2 рекрутируется на открытые участки хроматина, формируемые на сайтах CTCF, что усиливается прямыми белок-белковыми контактами. Можно

предположить, что CTCF помогает рекрутировать Top2 в интроны и 3'-области генов, что может быть необходимым при прохождении этапов транскрипции генов.

Активность С2Н2-белков регулируется путем внесения различных посттрансляционных модификаций. Наиболее хорошо исследовано фосфорилирование С2Н2-белков по линкерам между С2Н2-доменами, происходящее в процессе митоза и приводящее к снижению эффективности связывания белков с хроматином [80–83]. С2Н2-белки могут подвергаться и другим модификациям, таким, как присоединение убиквитина, сумо, поли-ADP-рибозилирование [84]. Сайт рибозилирования находится на N-конце CTCF [85], и эта модификация может влиять на димеризацию белка и его связывание с когезиновым комплексом. Поли-ADP-рибозилирование влияет на локализацию белка CTCF в ядерных компартментах, связывание с хроматином и регуляцию транскрипции [85–87]. Интересно, что N-конец CTCF человека взаимодействует с С-концевым доменом нуклеофосмина 1 (NPM1), который может определять локализацию CTCF в клетке [88]. В концевых доменах белка CTCF найдены сайты ковалентного присоединения белка SUMO через лизин CTCF [89]. Белок Pc2, принадлежащий к группе репрессоров транскрипции Поликомб, идентифицирован в качестве SUMO E3-лигазы для CTCF. В ядрах клеток CTCF и Pc2 найдены в тельцах, обогащенных белками группы Поликомб. Предполагается, что SUMO, взаимодействуя с различными белками и формируя гомополимеры, служит катализатором формирования плотных белковых гранул, которые могут иметь много функций и, в том числе, быть источником запасных белков при формировании хроматина на вновь синтезируемой ДНК в процессе репликации [90, 91]. Сумоилирование CTCF на хроматине также может выступать регулятором привлечения транскрипционных комплексов на хроматин, меняя тем самым свойства CTCF при активации или репрессии транскрипции генов.

CTCF как типичный представитель С2Н2-белков обладает типичными для них структурными особенностями, такими, как кластер цинковых пальцев, обеспечивающий специфичное связывание с геномными мишенями и взаимодействие с РНК и белками, концевые домены, необходимые для установления дистанционных контактов и взаимодействия с различными регуляторными комплексами.

CTCF В ОРГАНИЗАЦИИ АРХИТЕКТУРЫ ХРОМОСОМ И ИНСУЛЯЦИИ У ПОЗВОНОЧНЫХ

Белок CTCF первоначально рассматривали как основной инсуляторный белок позвоночных [92].

Первый инсулятор позвоночных был описан на границе гетерохроматинового района и кластера β -глобиновых генов курицы [93, 94]. Инсулятор, ко- ровая часть которого состоит из 275 п.н., был кар- тирован в гиперчувствительном к ДНКазе I районе и поэтому назван HS4 [95]. В модельных клеточных трансгенных системах инсулятор HS4 способен в одной или нескольких копиях блокировать активность энхансеров и защищать экспрессию трансгена от ре- прессии окружающим хроматином. Кроме сайта связывания белка CTCF, в HS4-инсуляторе найде- ны сайты связывания белков USF1/USF2 [96] и три сайта связывания белка VEZF1 (Vascular endothelial zinc finger 1) [97]. Показано, что CTCF нужен для бло- кирования энхансеров и рекрутирования белков USF1/USF2, которые, в свою очередь, привлекают комплексы, ремоделирующие и модифицирующие нуклеосомы. В результате нуклеосомы в районе HS4-инсулятора, как и на активных промоторах, оказы- ваются обогащены гистоном H3, метилированным по лизину 4, и ацетилованными гистонами H3 и H4.

Белок VEZF1 содержит кластер, состоящий из ше- сти C2H2-доменов, и преимущественно связывается с активными промоторами [98]. Инактивация сайтов связывания белка VEZF1 в трансгенных клеточных линиях усиливает метилирование ДНК на HS4-инсуляторе и промоторе репортерного гена [97]. Предполагается, что VEZF1 рекрутирует комплекс, который деметилюет ДНК и тем самым поддержи- вает рекрутирование на HS4-инсулятор и соседние регуляторные элементы транскрипционных фак- торов, которые не могут эффективно связываться с метилированными сайтами. Таким образом, HS4-инсулятор представляет собой комбинацию сайтов связывания нескольких белков, которые связываются с хроматином и функционируют в тесной коопера- ции друг с другом.

Несмотря на многочисленные примеры ключе- вой роли сайтов CTCF в организации границ регу- ляторных доменов и инсуляции энхансеров [23], от- крытым остается вопрос о роли других неизвестных белков, связывание которых с конкретным регуля- торным элементом зависит от присутствия CTCF, как это показано для белков USF1/USF2 и HS4-инсулятора. Например, у млекопитающих иденти- фицировано большое количество CTCF-зависимых инсуляторов, которые блокируют распространение Поликомб-зависимого гетерохроматина, связанно- го с обогащением протяженных участков хромати- на модификацией H3K27me3. Однако инактивация CTCF не приводит к распространению модификации H3K27me3 в этих районах, что предполагает при- сутствие на границах других белков, блокирующих распространение репрессивного хроматина и тем

самым маскирующих отсутствие CTCF [76]. Таким образом, CTCF-зависимые инсуляторы, границы ре- гуляторных доменов и ТАДов состоят, скорее всего, из CTCF-сайтов в комбинации с сайтами связывания других транскрипционных факторов, включая еще неописанные C2H2-белки.

Согласно современным представлениям, подкре- пленным многочисленными экспериментальными результатами, CTCF млекопитающих в кооперации с когезиновым комплексом формирует границы хро- матиновых петель и определяет границы большей части ТАДов [19, 99]. Когезиновый комплекс участву- ет в процессах митоза, мейоза и регуляции экспрес- сии генов [100, 101]. Этот комплекс состоит из белков SMC1, SMC3 и SCC1 (Rad21), образующих кольцевую структуру и связывающихся через SCC1 с четвертой субъединицей, которая существует в виде двух изо- форм, STAG1 (SA1) и STAG2 (SA2). Предполагается, что SA1 и SA2 могут определять локализацию коге- зинового комплекса на разных хроматиновых сай- тах. Комплекс NIPBL/MAU2 и WAPL катализируют АТФ-зависимое связывание когезинового комплекса с хроматином и его последующую диссоциацию соот- ветственно [100].

В зависимости от исследуемых компонентов коге- зинового комплекса, используемых антител к CTCF и клеточных линий наблюдается 40–95% колокали- зация сайтов CTCF и когезина [102–104]. При инак- тивации CTCF наблюдается перераспределение ко- гезиновых комплексов с сайтов связывания CTCF на промоторы активных генов, что сопровождается частичным нарушением конфигурации ТАДов [76]. Инактивация субъединиц когезинового комплекса или белка NIPBL [105, 106], обеспечивающего ре- крутирование когезинового комплекса на хроматин, приводит к почти полному исчезновению ТАДов. Наоборот, при инактивации факторов, негативно влияющих на связывание когезина с хроматином, происходит стабилизация ТАДов и дистанционных взаимодействий в хроматине [106]. Наконец, мутации и делеции в CTCF, которые нарушают его взаимо- действие с когезиновым комплексом, также приводят к значительным нарушениям процессов формиро- вания дистанционных взаимодействий и ТАДов [71, 104]. Субъединицы Smc1 и Smc3 содержат АТФ-зависимые домены, и энергия расщепления АТФ нужна при по- садке и диссоциации когезинового комплекса [107, 108]. При этом мутации в субъединицах когезиново- го комплекса, нарушающие гидролиз АТФ, влияют на дистанционные взаимодействия и организацию хромосом в ТАДы [109].

Сайты CTCF на границе ТАДов обычно находят- ся в конвергентной ориентации [8, 110]. Показано, что взаимная ориентация мотивов CTCF определяет

пары сайтов CTCF, которые преимущественно стабилизируют петли ДНК [8, 110–112]. Для объяснения предпочтительного формирования хроматиновых петель между CTCF-сайтами, расположенными в конвергентной ориентации, предложена модель экструзии петель, согласно которой когезиновый комплекс после загрузки на хроматин запускает экструзию ДНК с формированием хроматиновой петли. CTCF может блокировать продвижение когезинового комплекса только в том случае, если его N-концевой домен, взаимодействующий с субкомплексом SA2–SCC1 [71], правильно экспонирован по отношению к движущемуся когезиновому комплексу.

Модель постулирует, что когезиновый комплекс может активно (с использованием энергии АТФ) или пассивно индуцировать экструзию хроматина с образованием петель. Действительно, *in vitro* показано, что в присутствии NIPBL и молекул АТФ когезиновый комплекс связывается с молекулой ДНК и перемещается по ней с образованием петли [113], даже если ДНК связана с нуклеосомами [114]. Когезин также может преодолевать небольшие белковые комплексы нуклеосомного размера, однако неодолимой преградой являются комплексы диаметром более 13 нм, которые при наличии моторной функции могут сами перемещать когезин [115]. Таким образом, конвергентные сайты CTCF ограничивают участки выпетливания хроматиновых петель, а само формирование петель обеспечивается молекулярными моторами.

Согласно полимерной модели процесс формирования ТАДов во многом зависит от физических свойств хроматина, у которого ярко выражено стремление к формированию доменов одного типа. Такая тенденция особенно заметна на модели дрозофилы, в которой показан механизм формирования ТАДов с помощью электростатических межнуклеосомных взаимодействий. В результате границы таких доменов представлены протяженными участками открытого хроматина, а ТАД оказывается более плотной структурой [13, 19, 116, 117]. При этом в рамках такой модели роль CTCF заключается в рекрутировании когезиновых комплексов, которые стабилизируют взаимодействия между уже сближенными участками хроматина. Однако такая модель не объясняет, почему хроматиновые петли у млекопитающих формируют преимущественно только конвергентно расположенные сайты связывания CTCF.

Согласно экспериментальным данным [107, 113, 118], размер хроматиновой петли зависит не от времени связывания когезина с ДНК, а от преград, ограничивающих его продвижение (подобных CTCF). CTCF динамически связывается с хроматином, что согласуется с гетерогенностью границ

ТАДов, наблюдаемой при исследовании единичных клеток [20]. На границах ТАДов сайты связывания CTCF обычно представлены кластерами, что, возможно, и обеспечивает связывание CTCF с геномными мишенями на протяжении более длительного времени [119].

Согласно экструзионной модели, когезиновые комплексы только временно блокируются на конкретном CTCF-сайте и могут продолжить протягивание хроматина либо в результате преодоления блока, созданного CTCF, либо в результате ухода CTCF с хроматина [20]. Инактивация белка WAPL приводит к стабилизации связывания когезиновых комплексов с хроматином, при этом наблюдается увеличение размера хроматиновых петель, что объясняется увеличением времени, которое когезиновый комплекс проводит на хроматине [106, 120, 121].

В процессе митоза происходит конденсация хромосом, сопряженная с масштабными изменениями в хроматине и потерей связывания с ДНК части транскрипционных факторов. Во время профазы митоза большая часть когезина уходит с хромосом за исключением когезина, ассоциированного с центромерами. Во время анафазы происходит диссоциация когезина под действием сепаразы, что способствует расхождению сестринских хроматид [101]. На компактных митотических хромосомах структура ТАДов практически теряется, но быстро восстанавливается к середине стадии G1 [122]. Данные о связывании CTCF со своими сайтами на митотических хромосомах противоречивы. По некоторым оценкам, 18.6% сайтов находятся в связанном с CTCF состоянии [122], но показано также, что связывание CTCF со своими сайтами в основном теряется при фосфорилировании линкеров между C2H2-доменами [123]. Вероятно, уход CTCF со своих сайтов помогает более эффективно удалить когезиновые комплексы с митотических хромосом. Однако после митоза происходит быстрое восстановление мест связывания CTCF, что может быть следствием ассоциации свободного CTCF с конденсированными хромосомами в процессе митоза [123]. Остается открытым вопрос, каким образом осуществляется эффективное восстановление связывания CTCF со своими сайтами после митоза. Наиболее вероятно, что другие транскрипционные факторы остаются на митотических хромосомах и поддерживают частично открытое состояние хроматина (выполняют роль меток), что облегчает связывание CTCF со своими сайтами после митоза. В результате после репликации ДНК наблюдается быстрое восстановление как профиля связывания CTCF, так и структуры ТАДов на дублированных хромосомах. Можно предположить, что избыточная часть CTCF

находится в специализированных ядерных компартментах, которые стабилизируются SUMO [89]. При репликации избыточное количество CTCF используется для взаимодействия с увеличивающимся числом сайтов связывания.

ДРУГИЕ C2H2-БЕЛКИ ПОЗВОНОЧНЫХ С АРХИТЕКТУРНЫМИ ФУНКЦИЯМИ

В работах по созданию искусственных цинковых пальцев C2H2-типа для обеспечения специфичного взаимодействия с определенной мишенью в геноме показано резкое повышение специфичности связывания с ДНК кластера из пяти правильно организованных цинковых пальцев. Поэтому в данной главе мы рассмотрим белки с такой структурной организацией (имеющие не менее пяти доменов C2H2-типа, разделенных типичным линкерным участком из 6 п.н.) как наиболее перспективные в качестве архитектурных белков.

Другие C2H2-белки пока сравнительно менее изучены, чем CTCF [124, 125]. Основные проблемы при исследовании белков этого класса обусловлены высокой степенью дублирования функций у разных C2H2-белков и отсутствием качественных специфичных антител к этим белкам, которые сделают возможным проведение полногеномных исследований для идентификации сайтов связывания C2H2-белков и их роли в поддержании дистанционных взаимодействий между регуляторными элементами и формировании архитектуры хромосом. В двух работах [126, 127] изучены сайты связывания 60 и 221 C2H2-белков, тагированных эпитопами GFP или HA, в клетках НЕК293Т. Оказалось, что сайты связывания одних и тех же C2H2-белков, изученных в обеих работах, перекрываются друг с другом совсем незначительно [128]. Нужно отметить, что в этих работах экспрессия тагированных C2H2-белков происходила на фоне эндогенного C2H2-белка, так что большая часть реальных сайтов связывания была закрыта нативным белком, в то время как тагированный белок связывался в основном неспецифично с районами в зонах открытого хроматина. Вероятно, использование CRISPR/Cas9-редактирования позволит в ближайшем будущем заменять эндогенные гены модифицированными, экспрессирующими тагированные варианты C2H2-белков, что упростит их детальное исследование.

Приблизительно половина генома млекопитающих насыщена разными вариантами повторяющихся последовательностей различной природы, включая мобильные элементы и ретровирусы [129]. Большая часть изученных C2H2-белков, и в том числе CTCF, имеют сайты связывания в мобильных элементах [130–133]. При этом такие повторяющиеся последо-

вательности становятся частью регуляторных систем генов и границ ТАДов [134], что значительно расширяет возможности тонкой адаптации экспрессии генов в процессе эволюции.

Примерно половина всех белков с C2H2-доменом содержит на N-конце еще один домен. Два эволюционно наиболее древних домена, которые встречаются у всех эукариот, это PR/SET-домен (например, белок PRDM5 (рис. 1)), обладающий, как правило, метилтрансферазной активностью [135], и VTB-домен, формирующий димеры или мультимеры и привлекающий регуляторы транскрипции к геномным мишеням [136]. Одна из самых многочисленных групп C2H2-белков млекопитающих имеет на N-конце KRAB-домен (например, белки ZNF658 и ZNF764 (рис. 1)). Полагают, что этот домен получил широкое распространение у млекопитающих благодаря своей репрессорной функции по отношению к мобильным элементам. Однако параллельно с эволюцией регуляторных систем генов, в которые интегрируются мобильные элементы, KRAB-C2H2-белки приобретают новые функции в регуляции экспрессии хозяйских генов [130, 131]. Часть таких C2H2-белков с KRAB-доменом содержит на N-конце дополнительный домен – SCAN (например, белки ZNF202 и ZNF263) или DUF3669 (например, белки ZNF282 и ZNF398) [137–139]. Некоторые C2H2-белки содержат только SCAN-домен (например, MZF1) и происходят из белков, потерявших KRAB-домен. Можно предположить, что часть функций этих белков связана со способностью SCAN и DUF3669 к формированию гомо- и гетеродимеров между SCAN-C2H2 и DUF3669-C2H2 соответственно [131, 137].

Наиболее хорошо описано участие C2H2-белков в создании зоны открытого хроматина на промоторах генов и привлечении транскрипционных комплексов, вовлеченных в активацию или репрессию транскрипции. Белок ZNF658 участвует в активации экспрессии генов рРНК, которые транскрибируются РНК-полимеразой I, и связывается с регуляторным элементом, расположенным рядом с 3525 промоторами [140, 141]. Белок ZNF764 экспрессируется повсеместно, он вовлечен в регуляцию активности глюкокортикоидных, андрогенных и тиреоидных гормонов [142]. Интересно, что сайты связывания преимущественно находятся в межгенных участках (60%) и интронах (31%), при этом в значительной степени они (37%) колокализуются с сайтами глюкокортикоидных рецепторов (GR) [143]. Экспериментально доказано прямое взаимодействие между KRAB-доменом ZNF764 и LBD-доменом GR, что предполагает кооперативное связывание этих белков с регуляторными районами.

Сайты связывания белков ZNF202 [126, 144], ZNF263 [145], MZF1 [146], ZNF768 [133], PRDM5 [147] преимущественно найдены в промоторных областях генов, что свидетельствует о возможной роли данных факторов в активации и репрессии транскрипции.

На N-конце белка ZNF768 (рис. 1) находятся 15 гептадных повторов, имеющих сходство с C-концевым доменом РНК-полимеразы II [133] и, предположительно, участвующих в рекрутировании на промотор комплекса элонгации транскрипции.

Показано, что, используя SCAN-домен, MZF1 может гетеродимеризоваться с другими SCAN-содержащими белками – ZNF24, ZNF174 и ZNF202 [148, 149]. Белки ZNF282 и ZNF398 гомо- и гетеродимеризуются через DUF3669-домен [150] и могут связываться с промоторами комбинаторным способом [126]. Белок PRDM5 содержит N-концевой PR/SET-домен, который утратил метилирующую активность и, вероятно, участвует в белок-белковых взаимодействиях [151, 152].

Наиболее хорошо охарактеризован белок ZNF143 (рис. 1), необходимый для эмбрионального развития млекопитающих [153]. В центральной части белка расположен кластер из семи C2H2-доменов. N-Концевой домен содержит три повтора из 15 аминокислот, между которыми находятся спейсеры длиной 10–12 аминокислот [154]. C-Концевой домен обогащен кислыми аминокислотами, что характерно для активаторов транскрипции. Сайты связывания ZNF143 локализованы в области примерно 2000 промоторов, которые регулируются РНК-полимеразами II и III [155–158]. Функциональная активность ZNF143 в области промоторов связана с формированием зон открытого хроматина и участия в привлечении комплексов, активирующих транскрипцию [159–161]. Белок ZNF143 имеет два частично перекрывающихся консенсусных сайта связывания с одинаковой коровой последовательностью CCCAGA [155], что можно объяснить разным вкладом отдельных C2H2-доменов в узнавание двух вариантов сайтов. Полногеномные исследования показали, что белок ZNF143 может участвовать в формировании хроматиновых петель между энхансерами и промоторами [155, 156, 162–164].

Достаточно большой процент сайтов связывания белков PRDM5 и ZNF143 колокализуются с CTCF [143, 152, 163]. Белок PRDM5 обнаружен в комплексе с когезином и CTCF [152]. На клетках HEK293T показано, что инактивация ZNF143 приводит к нарушению некоторых CTCF-зависимых хроматиновых петель [163]. Однако отсутствуют экспериментальные данные о том, что ZNF143 (в отличие от CTCF) может участвовать в локализации когезинового комплекса на хроматине.

Другой пример структурной функции C2H2-белков можно наблюдать при исследовании архитектуры хроматина, которая организуется TFIIIC-комплексом. Установлено, что сайты связывания TFIIIC-комплекса колокализуются с конденсинами и могут служить границами между активным хроматином и гетерохроматином, а также поддерживать дистанционные взаимодействия, т.е. активно участвовать в формировании архитектуры хромосом [165]. Интересно, что сайты связывания белков PRDM5, CTCF/когезина, ZNF143 находятся рядом или колокализуются с районами связывания TFIIIC [152, 155, 166], что предполагает кооперативное участие этих белков в организации TFIIIC-зависимых регуляторных элементов. Более того, PRDM5 выделен в комплексе с TFIIIC, что предполагает участие PRDM5 в рекрутировании TFIIIC-комплекса на хроматин [152].

В заключение следует упомянуть второй (после CTCF) хорошо описанный C2H2-белок, TFIIIA, который у всех эукариот связывается с Pol III-зависимыми промоторами генов, кодирующих 5S рРНК [167]. В отличие от CTCF у TFIIIA, который обычно состоит из девяти C2H2-доменов и C-концевого активационного домена, названного TAS (Transcription Activating Signal, сигнал активации транскрипции), консервативной является только общая организация. Белок связывается с регуляторным элементом, названным ICR, который находится в транскрибируемой части гена. Структурный анализ показал, что C2H2-домены 1–3 и 7–9 связываются с двумя участками (C- и A-боксы) в ICR-эlemente, при этом центральные C2H2-домены участвуют в специфичном связывании с 5S РНК [167]. Отсутствие гомологии в аминокислотной последовательности TFIIIA-белков из разных видов предполагает параллельную эволюцию последовательностей промотора, 5S РНК и C2H2-доменов, которые участвуют в специфичном связывании ДНК и РНК. Белок TFIIIA определяет открытый хроматин на промоторе, а TAS-домен участвует в рекрутировании и стабильном связывании с промотором TFIIIB-комплекса [168].

Таким образом, в число C2H2-белков входят, помимо CTCF, и другие белки, которые могут выполнять архитектурную функцию, но они мало изучаются. Многие белки участвуют в активации или репрессии транскрипции, определены сайты связывания в открытом хроматине, получены примеры организации локальных петель на уровне отдельных генов. Однако не хватает структурированных полногеномных исследований, направленных на детальное описание роли этих белков в организации архитектуры хромосом.

С2Н2-БЕЛКИ ДРОЗОФИЛЫ: РАЗНЫЕ СТРУКТУРЫ, СХОДНЫЕ СВОЙСТВА

В геноме дрозофилы найдено порядка 170 белков с кластерами, состоящими из пяти и более С2Н2-доменов. Однако к настоящему времени только для небольшой части этих белков получены данные по распределению сайтов их связывания в геноме и их функциональной роли в регуляции транскрипции генов и организации архитектуры хромосом (рис. 2). К наиболее хорошо изученным С2Н2-белкам относятся первый описанный у высших эукариот белок, обладающий инсуляторными свойствами, Su(Hw), и гомолог CTCF млекопитающих [22, 24, 169, 170]. Оба инсуляторных белка обладают сходным строением – содержат концевые неструктурированные домены и расположенный в центре кластер из 11 (dCTCF) или 12 (Su(Hw)) С2Н2-доменов. На N-конце белка dCTCF картирован неструктурированный глобулярный домен, способный формировать тетрамерные комплексы [68, 69], и потенциальный участок взаимодействия с когезиновым комплексом, имеющий гомологию с YxF-мотивом CTCF человека, взаимодействующим с комплексом SA2–SCC1 [71]. Интересная структурная особенность другого исследованного С2Н2-белка, Orbp [171], – наличие на N-конце атипичного цинкового пальца, способного к гомодимеризации (рис. 2). Orbp содержит кластер из пяти С2Н2-доменов, ответственных за специфичное связывание с ДНК, и дополнительных четырех С2Н2-доменов, которые могут участвовать во взаимодействиях с РНК и белками.

Остальные пять С2Н2-белков (M1BP, ZAF1, Pita, Zw5 и ZIPIC) относятся к большой группе ZAD-содержащих белков. ZAD (zinc-finger-associated domain)-домен найден у 98 белков дрозофилы, при этом около 70 из них содержат пять или более С2Н2-доменов [172, 173]. Обычно гены, кодирующие белки ZAD–С2Н2, расположены кластерами и, подобно белкам КРАВ–С2Н2 млекопитающих [174, 175], активно эволюционируют в результате множественных дупликаций исходных копий генов. Структура ZAD-доменов формируется двумя парами цистеинов, координируемых ионом цинка [176]. N-Концевая часть домена представляет собой глобулярную структуру, а С-концевой стебель сформирован длинной α-спиралью. Интересно, что ZAD-домены способны к гомодимеризации с формированием антипараллельного димера [176, 177].

Мутации в генах, кодирующих белки Pita и Zw5, вызывают эмбриональную летальность, что предполагает важную роль белков данного класса на ранних этапах развития дрозофилы [178, 179]. Инактивация белка Su(Hw) нарушает развитие гонад, что приводит к стерильности самок [180]. Как и у млекопи-

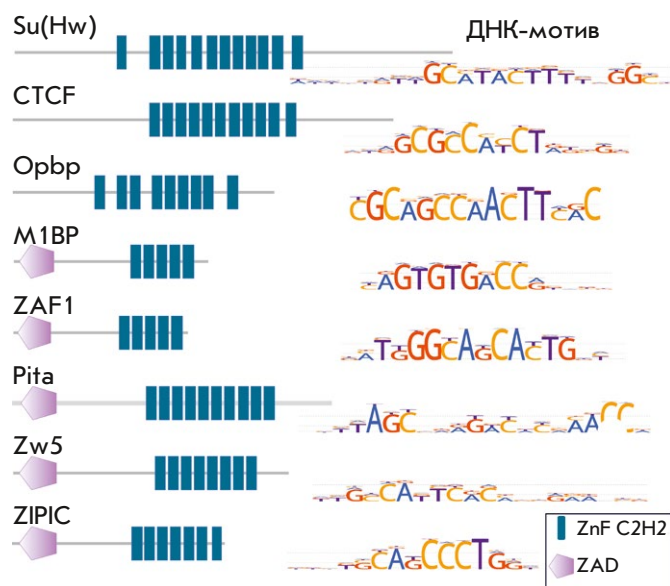


Рис. 2. С2Н2-белки дрозофилы с архитектурными функциями. Показаны доменная организация известных архитектурных белков дрозофилы и их мотивы связывания

тающих, белок CTCF дрозофилы участвует в регуляции экспрессии *hox*-генов [181, 182]. Несмотря на то что в геноме дрозофилы найдено только около 40 сайтов связывания белка Orbp, его инактивация приводит к летальности на стадии куколки [171].

Все исследованные С2Н2-белки используются для связывания с длинными (12–15 п.н.) ДНК-мотивами (рис. 2) четыре-пять С2Н2-доменов, образующих кластер [171, 183–186]. За исключением Su(Hw) сайты связывания С2Н2-белков локализованы преимущественно в области промоторов активных генов и интронов [177, 184–190]. Наиболее показательным примером белка данного класса является M1BP, который связывается с промоторами более 2000 генов [185] и, согласно экспериментальным данным [191], участвует в наведении зон открытого хроматина и привлечении основного промоторного комплекса. Белок Orbp также связывается исключительно с промоторами генов и примерно в половине из них колокализуется с M1BP [171]. В отличие от других С2Н2-белков, вероятно, участвующих в активации экспрессии, Su(Hw), связываясь с промоторами большой группы нейрональных генов, репрессирует их транскрипцию в гонадах самок и на ранних стадиях развития дрозофилы [192, 193].

Роль С2Н2-белков в установлении дистанционных взаимодействий и блокировании активности энхансера проанализирована в трансгенных линиях дрозофилы. *In vivo* С2Н2-белки эффективно взаимодействуют с искусственно синтезированными ДНК-

фрагментами, содержащими по четыре-пять сайтов связывания [177, 184, 188, 194]. В том случае, когда сайтами связывания С2Н2-белков окружен энхансер, происходит сильное блокирование его активности. Однако удаление любого из двух участков связывания С2Н2-белка восстанавливает активность энхансера, что доказывает важную роль взаимодействия между С2Н2-белками в процессе формирования петли хроматина, которая приводит к стерической изоляции энхансера. В трансгенной модельной системе сайты связывания С2Н2-белков способны сближать дрожжевой GAL4-активатор и промотор репортерного гена, что приводит к активации транскрипции [177, 184, 195]. При этом комбинации сайтов связывания разных С2Н2-белков не способны сближать GAL4-активатор с промотором [177, 195], что может объяснить значение предпочтительной гомодимеризации С2Н2-белков в обеспечении специфичных дистанционных взаимодействий между геномными элементами. Например, способность белков ZAF1 и ZIP1 поддерживать дистанционные взаимодействия определяется гомодимеризующимся ZAD-доменом [177, 184]. Таким образом, домены, обладающие способностью формировать гомодимеры, по всей видимости, играют важную роль в организации специфичных дистанционных взаимодействий между регуляторными элементами в хроматине.

Роль С2Н2-белков в организации границ регуляторных доменов наиболее наглядно можно продемонстрировать на примере *bithorax*-комплекса (ВХ-С), в состав которого входят три гомеозисных гена *Ubx*, *abd-A* и *Abd-B* [196, 197]. Регуляторная область ВХ-С делится на девять независимых доменов, каждый из которых активирует транскрипцию одного из трех гомеозисных генов в процессе развития. Несколько границ доменов детально охарактеризованы и картированы в виде минимальных фрагментов, способных функционировать как эффективные инсуляторы в трансгенных модельных системах [198–202]. Каждая охарактеризованная граница содержит в своем составе разные комбинации сайтов связывания белков Pita, dCTCF и Su(Hw), необходимых для ее функциональной активности [65, 66]. Границы могут быть заменены на мультиплицированные четыре-пять сайтов каждого из С2Н2-белков. Таким образом, несмотря на различия в структурной организации, белки Su(Hw), Pita или dCTCF имеют сходные функции и в процессе организации границ регуляторных доменов работают в кооперации друг с другом [66, 203, 204].

В отличие от млекопитающих, границы большинства ТАДов дрозофилы совпадают с кластерами генов «домашнего хозяйства» [205, 206]. Например, белок M1BP, сайты связывания которого локализованы

во многих промоторах генов «домашнего хозяйства», наиболее часто находят на границах ТАДов, тогда как сайты связывания других охарактеризованных С2Н2-белков обычно располагаются внутри ТАДов. В эмбрионах и эмбриональных клеточных линиях белок dCTCF, несмотря на когезинсвязывающий мотив, достаточно редко обнаруживается на границах ТАДов, при этом от 40 до 60% сайтов dCTCF колокализуются с когезиновыми комплексами на хроматине [205–207]. Связывание основной части когезина обнаруживается в зонах открытого хроматина активно транскрибируемых промоторов [208], поэтому нельзя исключить прямую или косвенную (наведение зон открытого хроматина) роль С2Н2-белков в привлечении когезиновых комплексов. Интересно, что в культуре клеток BG3, полученной из нервных тканей дрозофилы, большая часть границ ТАДов совпадает с сайтами связывания dCTCF [207]. Таким образом, границы ТАДов дрозофилы могут изменяться в процессе клеточной дифференцировки.

Наиболее вероятно, что границы ТАДов фиксируются благодаря взаимодействиям между белковыми комплексами, фланкирующими ТАД. Кроме того, на модели дрозофилы показано существование механизма формирования ТАДов за счет электростатических межнуклеосомных взаимодействий, в результате чего транскрипционно активные участки становятся своеобразными границами между ТАДами [116].

ЗАКЛЮЧЕНИЕ

В настоящее время С2Н2-белки высших эукариот остаются наименее изученным классом транскрипционных факторов. Хорошо исследованный белок CTCF млекопитающих дает общие представления о свойствах, партнерах и функциях транскрипционных факторов этого класса. CTCF, вероятно, является родоначальником всего класса С2Н2-белков, которые в процессе эволюции могли приобретать новые домены и связываться с новыми ДНК-последовательностями. В этом плане интересно то, что CTCF и у дрозофилы, и у млекопитающих участвует в организации границ транскрипционных доменов гомеозисных генов. На основании существующих результатов можно заключить, что С2Н2-белки у млекопитающих и дрозофилы часто участвуют в организации активных промоторов. Белки С2Н2, взаимодействуя с комплексами ремоделирования нуклеосом, могут формировать открытый хроматин и одновременно участвовать в привлечении основных транскрипционных факторов на промоторы. Многие хорошо исследованные регуляторные элементы, прежде всего промоторы и инсуляторы, содержат комбинации сайтов связывания С2Н2-белков, которые, взаимодействуя с хроматином, функциониру-

ют кооперативно. У части С2Н2-белков, в том числе СТСФ, идентифицированы N-концевые гомодимеризующиеся домены, которые могут быть вовлечены в организацию специфичных дистанционных взаимодействий. В настоящее время только в белке СТСФ найден мотив, с которым взаимодействует когезиновый комплекс. Однако С2Н2-белки, вероятно, способны взаимодействовать с другими поверхностями в когезиновых и конденсиновых комплексах, что согласуется с локализацией этих комплексов на активных промоторах.

Считается, что границы ТАДов и дистанционные взаимодействия у млекопитающих и дрозофилы формируются с помощью разных механизмов. Однако способен ли когезиновый комплекс млекопитающих активно выпетливать хроматин при формировании ТАДов и дистанционных взаимодействий между регуляторными элементами остается открытым вопросом. Также непонятно, почему аналогичный механизм не работает у других высших эукариот, несмотря на высокую степень консервативности когезинового комплекса. Интересно, что у рыбок данио на большей части границ ТАДов отсутствуют СТСФ и когезиновый комплекс [209], несмотря на то что СТСФ у данио и человека гомологичны на 86%. С другой стороны, СТСФ обнаружен на границах ТАДов в нервных клетках дрозофилы [207]. Можно предположить, что механизмы формирования ТАДов на самом деле намного более универсаль-

ны, чем кажется в настоящее время. Такие С2Н2-белки, как PRDM5 и ZNF143, могут стабилизировать связывание СТСФ с границами ТАДов млекопитающих и участвовать в дистанционных взаимодействиях. С2Н2-белки дрозофилы, связываясь в различных комбинациях с инсуляторами (например, в составе ВХ-С), позволяют двум идентичным копиям инсулятора поддерживать сверхдальние взаимодействия, т.е. по сути, формировать границы нового ТАДа. На границах ТАДов млекопитающих обычно находятся наиболее эволюционно консервативные кластеры СТСФ-сайтов [119]. Можно предположить, что на ранних этапах эволюции позвоночных размноженные копии одного или нескольких типов мобильных элементов, содержащих сайты связывания СТСФ в комбинации с сайтами других С2Н2-белков, организовывали дистанционные взаимодействия, часть из которых и привела к формированию ТАДов. Таким образом, несмотря на большой прогресс, достигнутый за последние годы в изучении пространственной организации генома и, в частности, архитектурной роли СТСФ, очень много вопросов остается без ответа из-за отсутствия данных о других участниках, необходимых для формирования архитектуры ядра. ●

*Работа выполнена при поддержке гранта
Российского научного фонда (проект 19-74-30026).*

СПИСОК ЛИТЕРАТУРЫ

- Spitz F, Furlong E.E. // Nature Reviews Genetics. 2012. V. 13. № 9. P. 613–626.
- Levine M., Cattoglio C., Tjian R. // Cell. 2014. V. 157. № 1. P. 13–25.
- Zabidi M.A., Stark A. // Trends Genet. 2016. V. 32. № 12. P. 801–814.
- Furlong E.E.M., Levine M. // Science. 2018. V. 361. № 6409. P. 1341–1345.
- Geyer P.K., Clark I. // Cell. Mol. Life Sci. 2002. V. 59. № 12. P. 2112–2127.
- West A.G., Gaszner M., Felsenfeld G. // Genes Dev. 2002. V. 16. № 3. P. 271–288.
- Gerasimova T.I., Corces V.G. // Annu. Rev. Genet. 2001. V. 35. P. 193–208.
- Rao S.S., Huntley M.H., Durand N.C., Stamenova E.K., Bochkov I.D., Robinson J.T., Sanborn A.L., Machol I., Omer A.D., Lander E.S., et al. // Cell. 2014. V. 159. № 7. P. 1665–1680.
- Boettiger A., Murphy S. // Trends Genet. 2020. V. 36. № 4. P. 273–287.
- Boettiger A.N., Bintu B., Moffitt J.R., Wang S., Beliveau B.J., Fudenberg G., Imakaev M., Mirny L.A., Wu C.T., Zhuang X. // Nature. 2016. V. 529. № 7586. P. 418–422.
- Dekker J., Misteli T. // Cold Spring Harbor Perspectives Biol. 2015. V. 7. № 10. P. a019356.
- Sikorska N., Sexton T. // J. Mol. Biol. 2020. V. 432. № 3. P. 653–664.
- Hansen A.S., Cattoglio C., Darzacq X., Tjian R. // Nucleus. 2018. V. 9. № 1. P. 20–32.
- Luppino J.M., Park D.S., Nguyen S.C., Lan Y., Xu Z., Yunker R., Joyce E.F. // Nat. Genet. 2020. V. 52. № 8. P. 840–848.
- Zheng H., Xie W. // Nat. Rev. Mol. Cell Biol. 2019. V. 20. № 9. P. 535–550.
- Sexton T., Yaffe E., Kenigsberg E., Bantignies F., Leblanc B., Hoichman M., Parrinello H., Tanay A., Cavalli G. // Cell. 2012. V. 148. № 3. P. 458–472.
- Dixon J.R., Selvaraj S., Yue F., Kim A., Li Y., Shen Y., Hu M., Liu J.S., Ren B. // Nature. 2012. V. 485. № 7398. P. 376–380.
- Nora E.P., Lajoie B.R., Schulz E.G., Giorgetti L., Okamoto I., Servant N., Pilot T., van Berkum N.L., Meisig J., Sedat J., et al. // Nature. 2012. V. 485. № 7398. P. 381–385.
- Szabo Q., Bantignies F., Cavalli G. // Science Advances. 2019. V. 5. № 4. P. eaaw1668.
- Chang L.H., Ghosh S., Noordermeer D. // J. Mol. Biol. 2020. V. 432. № 3. P. 643–652.
- Arzate-Mejia R.G., Recillas-Targa F., Corces V.G. // Development. 2018. V. 145. № 6. P. dev137729.
- Ali T., Renkawitz R., Bartkuhn M. // Curr. Opin. Genetics Dev. 2016. V. 37. P. 17–26.
- Braccioli L., de Wit E. // Essays Biochem. 2019. V. 63. № 1. P. 157–165.
- Chen D., Lei E.P. // Curr. Opin. Cell Biol. 2019. V. 58. P. 61–68.
- Merkenschlager M., Nora E.P. // Annu. Rev. Genomics Hum. Genet. 2016. V. 17. P. 17–43.
- Schmidt D., Schwalie P.C., Wilson M.D., Ballester B., Goncalves A., Kutter C., Brown G.D., Marshall A., Flicke P,

- Odom D.T. // *Cell*. 2012. V. 148. № 1–2. P. 335–348.
27. Chen H., Tian Y., Shu W., Bo X., Wang S. // *PLoS One*. 2012. V. 7. № 7. P. e41374.
28. Klug A. // *Quarterly Rev. Biophys.* 2010. V. 43. № 1. P. 1–21.
29. Persikov A.V., Wetzel J.L., Rowland E.F., Oakes B.L., Xu D.J., Singh M., Noyes M.B. // *Nucleic Acids Res.* 2015. V. 43. № 3. P. 1965–1984.
30. Garton M., Najafabadi H.S., Schmitges F.W., Radovani E., Hughes T.R., Kim P.M. // *Nucleic Acids Res.* 2015. V. 43. № 19. P. 9147–9157.
31. Persikov A.V., Singh M. // *Nucleic Acids Res.* 2014. V. 42. № 1. P. 97–108.
32. Durai S., Mani M., Kandavelou K., Wu J., Porteus M.H., Chandrasegaran S. // *Nucleic Acids Res.* 2005. V. 33. № 18. P. 5978–5990.
33. Kim Y.G., Cha J., Chandrasegaran S. // *Proc. Natl. Acad. Sci. USA*. 1996. V. 93. № 3. P. 1156–1160.
34. Brayer K.J., Segal D.J. // *Cell Biochem. Biophys.* 2008. V. 50. № 3. P. 111–131.
35. Ryan R.F., Darby M.K. // *Nucleic Acids Res.* 1998. V. 26. № 3. P. 703–709.
36. Crozatier M., Kongsuwan K., Ferrer P., Merriam J.R., Lengyel J.A., Vincent A. // *Genetics*. 1992. V. 131. № 4. P. 905–916.
37. Wolfe S.A., Nekludova L., Pabo C.O. // *Annu. Rev. Biophys. Biomol. Structure*. 2000. V. 29. P. 183–212.
38. Hashimoto H., Wang D., Horton J.R., Zhang X., Corces V.G., Cheng X. // *Mol. Cell*. 2017. V. 66. № 5. P. 711–720 e713.
39. Nakahashi H., Kwon K.R., Resch W., Vian L., Dose M., Stavreva D., Hakim O., Pruett N., Nelson S., Yamane A., et al. // *Cell Reports*. 2013. V. 3. № 5. P. 1678–1689.
40. Xiao T., Wongtrakongate P., Trainor C., Felsenfeld G. // *Cell Reports*. 2015. V. 12. № 10. P. 1704–1714.
41. Xu D., Ma R., Zhang J., Liu Z., Wu B., Peng J., Zhai Y., Gong Q., Shi Y., Wu J., et al. // *J. Phys. Chem. Lett.* 2018. V. 9. № 14. P. 4020–4028.
42. Yin M., Wang J., Wang M., Li X., Zhang M., Wu Q., Wang Y. // *Cell Research*. 2017. V. 27. № 11. P. 1365–1377.
43. Liu Y., Zhang X., Blumenthal R.M., Cheng X. // *Trends Biochem. Sci.* 2013. V. 38. № 4. P. 177–183.
44. Hudson N.O., Buck-Koehntop B.A. // *Molecules*. 2018. V. 23. № 10. P. 2555.
45. Ren G., Zhao K. // *Cell. Biosci.* 2019. V. 9. P. 83.
46. Noordermeer D., Feil R. // *Curr. Opin. Genet. Dev.* 2020. V. 61. P. 17–24.
47. Heard E., Distèche C.M. // *Genes Dev.* 2006. V. 20. № 14. P. 1848–1867.
48. Iuchi S. // *Cell. Mol. Life Sci.* 2001. V. 58. № 4. P. 625–635.
49. Hall T.M. // *Curr. Opin. Struct. Biol.* 2005. V. 15. № 3. P. 367–373.
50. Saldana-Meyer R., Gonzalez-Buendia E., Guerrero G., Narendra V., Bonasio R., Recillas-Targa F., Reinberg D. // *Genes Dev.* 2014. V. 28. № 7. P. 723–734.
51. Kung J.T., Kesner B., An J.Y., Ahn J.Y., Cifuentes-Rojas C., Colognori D., Jeon Y., Szanto A., del Rosario B.C., Pinter S.F., et al. // *Mol. Cell*. 2015. V. 57. № 2. P. 361–375.
52. Hansen A.S., Hsieh T.S., Cattoglio C., Pustova I., Saldana-Meyer R., Reinberg D., Darzacq X., Tjian R. // *Mol. Cell*. 2019. V. 76. № 3. P. 395–411 e313.
53. Shukla S., Kavak E., Gregory M., Imashimizu M., Shutinoski B., Kashlev M., Oberdoerffer P., Sandberg R., Oberdoerffer S. // *Nature*. 2011. V. 479. № 7371. P. 74–79.
54. Marina R.J., Sturgill D., Bailly M.A., Thenoz M., Varma G., Prigge M.F., Nanan K.K., Shukla S., Haque N., Oberdoerffer S. // *EMBO*. 2016. V. 35. № 3. P. 335–355.
55. Nanavaty V., Abrash E.W., Hong C., Park S., Fink E.E., Li Z., Sweet T.J., Bhasin J.M., Singuri S., Lee B.H., et al. // *Mol. Cell*. 2020. V. 78. № 4. P. 752–764 e756.
56. Chernukhin I., Shamsuddin S., Kang S.Y., Bergstrom R., Kwon Y.W., Yu W., Whitehead J., Mukhopadhyay R., Docquier F., Farrar D., et al. // *Mol. Cell Biol.* 2007. V. 27. № 5. P. 1631–1648.
57. Heger P., Marin B., Schierenberg E. // *BMC Mol. Biol.* 2009. V. 10. P. 84.
58. Heger P., Marin B., Bartkuhn M., Schierenberg E., Wiehe T. // *Proc. Natl. Acad. Sci. USA*. 2012. V. 109. № 43. P. 17507–17512.
59. Moon H., Filippova G., Loukinov D., Pugacheva E., Chen Q., Smith S.T., Munhall A., Grewe B., Bartkuhn M., Arnold R., et al. // *EMBO Reports*. 2005. V. 6. № 2. P. 165–170.
60. Holohan E.E., Kwong C., Adryan B., Bartkuhn M., Herold M., Renkawitz R., Russell S., White R. // *PLoS Genet.* 2007. V. 3. № 7. P. e112.
61. Kadota M., Hara Y., Tanaka K., Takagi W., Tanegashima C., Nishimura O., Kuraku S. // *Sci. Rep.* 2017. V. 7. № 1. P. 4957.
62. Narendra V., Rocha P.P., An D., Raviram R., Skok J.A., Mazzoni E.O., Reinberg D. // *Science*. 2015. V. 347. № 6225. P. 1017–1021.
63. Savitsky M., Kim M., Kravchuk O., Schwartz Y.B. // *Genetics*. 2016. V. 202. № 2. P. 601–617.
64. Luo H., Wang F., Zha J., Li H., Yan B., Du Q., Yang F., Sobh A., Vulpe C., Drusbosky L., et al. // *Blood*. 2018. V. 132. № 8. P. 837–848.
65. Kyrchanova O., Zolotarev N., Mogila V., Maksimenko O., Schedl P., Georgiev P. // *Development*. 2017. V. 144. № 14. P. 2663–2672.
66. Kyrchanova O., Maksimenko O., Ibragimov A., Sokolov V., Postika N., Lukyanova M., Schedl P., Georgiev P. // *Sci. Adv.* 2020. V. 6. № 13. P. eaaz3152.
67. Schwalie P.C., Ward M.C., Cain C.E., Faure A.J., Gilad Y., Odom D.T., Flicek P. // *Genome Biol.* 2013. V. 14. № 12. P. R148.
68. Bonchuk A., Kamalyan S., Mariasina S., Boyko K., Popov V., Maksimenko O., Georgiev P. // *Sci. Rep.* 2020. V. 10. № 1. P. 2677.
69. Bonchuk A., Maksimenko O., Kyrchanova O., Ivlieva T., Mogila V., Deshpande G., Wolle D., Schedl P., Georgiev P. // *BMC Biol.* 2015. V. 13. P. 63.
70. Nishana M., Ha C., Rodriguez-Hernaez J., Ranjbaran A., Chio E., Nora E.P., Badri S.B., Kloetgen A., Bruneau B.G., Tsigos A., et al. // *Genome Biol.* 2020. V. 21. № 1. P. 108.
71. Li Y., Haarhuis J.H.L., Sedenno Cacciatore A., Oldenkamp R., van Ruiten M.S., Willems L., Teunissen H., Muir K.W., de Wit E., Rowland B.D., et al. // *Nature*. 2020. V. 578. № 7795. P. 472–476.
72. Xiao T., Wallace J., Felsenfeld G. // *Mol. Cell Biol.* 2011. V. 31. № 11. P. 2174–2183.
73. Zlatanova J., Caiafa P. // *J. Cell Sci.* 2009. V. 122. № Pt 9. P. 1275–1284.
74. Marino M.M., Rega C., Russo R., Valletta M., Gentile M.T., Esposito S., Baglivo I., De Feis I., Angelini C., Xiao T., et al. // *J. Biol. Chem.* 2019. V. 294. № 3. P. 861–873.
75. Pena-Hernandez R., Marques M., Hilmi K., Zhao T., Saad A., Alaoui-Jamali M.A., del Rincon S.V., Ashworth T., Roy A.L., Emerson B.M., et al. // *Proc. Natl. Acad. Sci. USA*. 2015. V. 112. № 7. P. E677–686.
76. Nora E.P., Goloborodko A., Valton A.L., Gibcus J.H., Uebersohn A., Abdennur N., Dekker J., Mirny L.A., Bruneau B.G. // *Cell*. 2017. V. 169. № 5. P. 930–944 e922.
77. Yao H., Brick K., Evrard Y., Xiao T., Camerini-Otero R.D., Felsenfeld G. // *Genes Dev.* 2010. V. 24. № 22. P. 2543–2555.
78. Uuskula-Reimand L., Hou H., Samavarchi-Tehrani P.,

- Rudan M.V., Liang M., Medina-Rivera A., Mohammed H., Schmidt D., Schwalie P., Young E.J., et al. // *Genome Biol.* 2016. V. 17. № 1. P. 182.
79. Gittens W.H., Johnson D.J., Allison R.M., Cooper T.J., Thomas H., Neale M.J. // *Nat. Commun.* 2019. V. 10. № 1. P. 4846.
80. Jantz D., Berg J.M. // *Proc. Natl. Acad. Sci. USA.* 2004. V. 101. № 20. P. 7589–7593.
81. Dovat S., Ronni T., Russell D., Ferrini R., Cobb B.S., Smale S.T. // *Genes Dev.* 2002. V. 16. № 23. P. 2985–2990.
82. Rizkallah R., Alexander K.E., Hurt M.M. // *Cell Cycle.* 2011. V. 10. № 19. P. 3327–3336.
83. Luo H., Yu Q., Liu Y., Tang M., Liang M., Zhang D., Xiao T.S., Wu L., Tan M., Ruan Y., et al. // *Science Adv.* 2020. V. 6. № 8. P. eaaw4651.
84. Caiafa P., Zlatanova J. // *J. Cell. Physiol.* 2009. V. 219. № 2. P. 265–270.
85. Farrar D., Rai S., Chernukhin I., Jagodic M., Ito Y., Yammine S., Ohlsson R., Murrell A., Klenova E. // *Mol. Cell Biol.* 2010. V. 30. № 5. P. 1199–1216.
86. Pavlaki I., Docquier F., Chernukhin I., Kita G., Gretton S., Clarkon C.T., Teif V.B., Klenova E. // *Biochim. Biophys. Acta Gene Regul. Mech.* 2018. V. 1861. № 8. P. 718–730.
87. Torrano V., Navascues J., Docquier F., Zhang R., Burke L.J., Chernukhin I., Farrar D., Leon J., Berciano M.T., Renkawitz R., et al. // *J. Cell Sci.* 2006. V. 119. Pt 9. P. 1746–1759.
88. Wang A.J., Han Y., Jia N., Chen P., Minden M.D. // *Leukemia.* 2020. V. 34. № 5. P. 1278–1290.
89. MacPherson M.J., Beatty L.G., Zhou W., Du M., Sadowski P.D. // *Mol. Cell Biol.* 2009. V. 29. № 3. P. 714–725.
90. Golovnin A., Volkov I., Georgiev P. // *J. Cell Sci.* 2012. V. 125. № Pt 8. P. 2064–2074.
91. Rosonina E., Akhter A., Dou Y., Babu J., Sri Theivakadacham V.S. // *Transcription.* 2017. V. 8. № 4. P. 220–231.
92. Wallace J.A., Felsenfeld G. // *Curr. Opin. Genet. Dev.* 2007. V. 17. № 5. P. 400–407.
93. Barkess G., West A.G. // *Epigenomics.* 2012. V. 4. № 1. P. 67–80.
94. Ghirlando R., Felsenfeld G. // *Genes Dev.* 2016. V. 30. № 8. P. 881–891.
95. Farrell C.M., West A.G., Felsenfeld G. // *Mol. Cell Biol.* 2002. V. 22. № 11. P. 3820–3831.
96. West A.G., Huang S., Gaszner M., Litt M.D., Felsenfeld G. // *Mol. Cell.* 2004. V. 16. № 3. P. 453–463.
97. Dickson J., Gowher H., Strogantsev R., Gaszner M., Hair A., Felsenfeld G., West A.G. // *PLoS Genet.* 2010. V. 6. № 1. P. e1000804.
98. Gowher H., Brick K., Camerini-Otero R.D., Felsenfeld G. // *Proc. Natl. Acad. Sci. USA.* 2012. V. 109. № 7. P. 2370–2375.
99. Fudenberg G., Imakaev M., Lu C., Goloborodko A., Abdennur N., Mirny L.A. // *Cell Reports.* 2016. V. 15. № 9. P. 2038–2049.
100. Nishiyama T. // *Curr. Opin. Cell Biol.* 2019. V. 58. P. 8–14.
101. Morales C., Losada A. // *Curr. Opin. Cell Biol.* 2018. V. 52. P. 51–57.
102. 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. № 3. P. 422–433.
103. 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. № 7180. P. 796–801.
104. Pugacheva E.M., Kubo N., Loukinov D., Tajmul M., Kang S., Kovalchuk A.L., Strunnikov A.V., Zentner G.E., Ren B., Lobanenkov V.V. // *Proc. Natl. Acad. Sci. USA.* 2020. V. 117. № 4. P. 2020–2031.
105. Rao S.S.P., Huang S.C., Glenn St Hilaire B., Engreitz J.M., Perez E.M., Kieffer-Kwon K.R., Sanborn A.L., Johnstone S.E., Bascom G.D., Bochkov I.D., et al. // *Cell.* 2017. V. 171. № 2. P. 305–320 e324.
106. Wutz G., Varnai C., Nagasaka K., Cisneros D.A., Stocsits R.R., Tang W., Schoenfelder S., Jessberger G., Muhar M., Hossain M.J., et al. // *EMBO.* 2017. V. 36. № 24. P. 3573–3599.
107. Ladurner R., Bhaskara V., Huis in 't Veld P.J., Davidson I.F., Kreidl E., Petzold G., Peters J.M. // *Curr. Biol.* 2014. V. 24. № 19. P. 2228–2237.
108. Elbatsh A.M.O., Haarhuis J.H.I., Petela N., Chopard C., Fish A., Celie P.H., Stadnik M., Ristic D., Wyman C., Medema R.H., et al. // *Mol. Cell.* 2016. V. 61. № 4. P. 575–588.
109. Vian L., Pekowska A., Rao S.S.P., Kieffer-Kwon K.R., Jung S., Baranello L., Huang S.C., El Khattabi L., Dose M., Pruett N., et al. // *Cell.* 2018. V. 175. № 1. P. 292–294.
110. Vietri Rudan M., Barrington C., Henderson S., Ernst C., Odum D.T., Tanay A., Hadjur S. // *Cell Rep.* 2015. V. 10. № 8. P. 1297–1309.
111. de Wit E., Vos E.S., Holwerda S.J., Valdes-Quezada C., Versteeg M.J., Teunissen H., Splinter E., Wijchers P.J., Krijger P.H., de Laat W. // *Mol. Cell.* 2015. V. 60. № 4. P. 676–684.
112. Guo Y., Xu Q., Canzio D., Shou J., Li J., Gorkin D.U., Jung I., Wu H., Zhai Y., Tang Y., et al. // *Cell.* 2015. V. 162. № 4. P. 900–910.
113. Davidson I.F., Bauer B., Goetz D., Tang W., Wutz G., Peters J.M. // *Science.* 2019. V. 366. № 6471. P. 1338–1345.
114. Kim Y., Shi Z., Zhang H., Finkelstein I.J., Yu H. // *Science.* 2019. V. 366. № 6471. P. 1345–1349.
115. Stigler J., Camdere G.O., Koshland D.E., Greene E.C. // *Cell Rep.* 2016. V. 15. № 5. P. 988–998.
116. Ulianov S.V., Khrameeva E.E., Gavrilov A.A., Flyamer I.M., Kos P., Mikhaleva E.A., Penin A.A., Logacheva M.D., Imakaev M.V., Chertovich A., et al. // *Genome Res.* 2016. V. 26. № 1. P. 70–84.
117. Luzhin A.V., Flyamer I.M., Khrameeva E.E., Ulianov S.V., Razin S.V., Gavrilov A.A. // *J. Cell. Biochem.* 2019. V. 120. № 3. P. 4494–4503.
118. Holzmann J., Politi A.Z., Nagasaka K., Hantsche-Grininger M., Walther N., Koch B., Fuchs J., Durnberger G., Tang W., Ladurner R., et al. // *eLife.* 2019. V. 8. P. e46269.
119. Kentepozidou E., Aitken S.J., Feig C., Stefflova K., Ibarra-Soria X., Odum D.T., Roller M., Flicek P. // *Genome Biol.* 2020. V. 21. № 1. P. 5.
120. Haarhuis J.H.I., van der Weide R.H., Blomen V.A., Yanez-Cuna J.O., Amendola M., van Ruiten M.S., Krijger P.H.L., Teunissen H., Medema R.H., van Steensel B., et al. // *Cell.* 2017. V. 169. № 4. P. 693–707 e614.
121. Gassler J., Brandao H.B., Imakaev M., Flyamer I.M., Ladstatter S., Bickmore W.A., Peters J.M., Mirny L.A., Tachibana K. // *EMBO J.* 2017. V. 36. № 24. P. 3600–3618.
122. Zhang H., Emerson D.J., Gilgenast T.G., Titus K.R., Lan Y., Huang P., Zhang D., Wang H., Keller C.A., Giardine B., et al. // *Nature.* 2019. V. 576. № 7785. P. 158–162.
123. Owens N., Papadopoulou T., Festuccia N., Tachtsidi A., Gonzalez I., Dubois A., Vandormael-Pournin S., Nora E.P., Bruneau B.G., Cohen-Tannoudji M., et al. // *eLife.* 2019. V. 8. P. e47898.
124. Lambert S.A., Jolma A., Campitelli L.F., Das P.K., Yin Y., Albu M., Chen X., Taipale J., Hughes T.R., Weirauch M.T. // *Cell.* 2018. V. 172. № 4. P. 650–665.
125. Lambert S.A., Yang A.W.H., Sasse A., Cowley G., Albu M., Caddick M.X., Morris Q.D., Weirauch M.T., Hughes T.R. // *Nat. Genet.* 2019. V. 51. № 6. P. 981–989.

126. Imbeault M., Helleboid P.Y., Trono D. // *Nature*. 2017. V. 543. № 7646. P. 550–554.
127. Schmitges F.W., Radovani E., Najafabadi H.S., Barazandeh M., Campitelli L.F., Yin Y., Jolma A., Zhong G., Guo H., Kanagalingam T., et al. // *Genome Res*. 2016. V. 26. № 12. P. 1742–1752.
128. Barazandeh M., Lambert S.A., Albu M., Hughes T.R. // *G3 (Bethesda)*. 2018. V. 8. № 1. P. 219–229.
129. Platt R.N., 2nd, Vandeweye M.W., Ray D.A. // *Chromosome Research*. 2018. V. 26. № 1–2. P. 25–43.
130. Bruno M., Mahgoub M., Macfarlan T.S. // *Annu. Rev. Genet.* 2019. V. 53. P. 393–416.
131. Emerson R.O., Thomas J.H. // *J. Virol.* 2011. V. 85. № 22. P. 12043–12052.
132. Okumura K., Sakaguchi G., Naito K., Tamura T., Igarashi H. // *Nucleic Acids Res.* 1997. V. 25. № 24. P. 5025–5032.
133. Rohrmoser M., Kluge M., Yahia Y., Gruber-Eber A., Maqbool M.A., Forne I., Krebs S., Blum H., Greifengberg A.K., Geyer M., et al. // *Nucleic Acids Res.* 2019. V. 47. № 2. P. 700–715.
134. Diehl A.G., Ouyang N., Boyle A.P. // *Nat. Commun.* 2020. V. 11. № 1. P. 1796.
135. Herz H.M., Garruss A., Shilatfard A. // *Trends in Biochemical Sciences*. 2013. V. 38. № 12. P. 621–639.
136. Maeda T. // *Int. J. Hematol.* 2016. V. 104. № 3. P. 310–323.
137. Al Chiblak M., Steinbeck F., Thiesen H.J., Lorenz P. // *BMC Molecular and Cell Biology*. 2019. V. 20. № 1. P. 60.
138. Schumacher C., Wang H., Honer C., Ding W., Koehn J., Lawrence Q., Coulis C.M., Wang L.L., Ballinger D., Bowen B.R., et al. // *J. Biol. Chem.* 2000. V. 275. № 22. P. 17173–17179.
139. Yang P., Wang Y., Macfarlan T.S. // *Trends Genet.* 2017. V. 33. № 11. P. 871–881.
140. Francis M., Cheng H., Ma P., Grider A. // *Biol. Trace Elem. Res.* 2019. V. 192. № 2. P. 83–90.
141. Ogo O.A., Tyson J., Cockell S.J., Howard A., Valentine R.A., Ford D. // *Mol. Cell Biol.* 2015. V. 35. № 6. P. 977–987.
142. Kino T., Pavlatou M.G., Moraitis A.G., Nemery R.L., Raygada M., Stratakis C.A. // *J. Clin. Endocrinol. Metab.* 2012. V. 97. № 8. P. E1557–1566.
143. Fadda A., Syed N., Mackeh R., Papadopoulou A., Suzuki S., Jithesh P.V., Kino T. // *Scientific Reports*. 2017. V. 7. P. 41598.
144. Wagner S., Hess M.A., Ormonde-Hanson P., Malandro J., Hu H., Chen M., Kehrer R., Frodsham M., Schumacher C., Beluch M., et al. // *J. Biol. Chem.* 2000. V. 275. № 21. P. 15685–15690.
145. Frieze S., Lan X., Jin V.X., Farnham P.J. // *J. Biol. Chem.* 2010. V. 285. № 2. P. 1393–1403.
146. Brix D.M., Bundgaard Clemmensen K.K., Kallunki T. // *Cells*. 2020. V. 9. № 1. P. 223.
147. Galli G.G., Mulhaupt H.A., Carrara M., de Lichtenberg K.H., Christensen I.B., Linnemann D., Santoni-Rugiu E., Calogero R.A., Lund A.H. // *Oncogene*. 2014. V. 33. № 25. P. 3342–3350.
148. Noll L., Peterson F.C., Hayes P.L., Volkman B.F., Sander T. // *Leukemia Research*. 2008. V. 32. № 10. P. 1582–1592.
149. Peterson F.C., Hayes P.L., Waltner J.K., Heisner A.K., Jensen D.R., Sander T.L., Volkman B.F. // *J. Mol. Biol.* 2006. V. 363. № 1. P. 137–147.
150. Helleboid P.Y., Heusel M., Duc J., Piot C., Thorball C.W., Coluccio A., Pontis J., Imbeault M., Turelli P., Aebersold R., et al. // *EMBO*. 2019. V. 38. № 18. P. e101220.
151. Duan Z., Person R.E., Lee H.H., Huang S., Donadieu J., Badolato R., Grimes H.L., Papayannopoulou T., Horwitz M.S. // *Mol. Cell Biol.* 2007. V. 27. № 19. P. 6889–6902.
152. Galli G.G., Carrara M., Francavilla C., de Lichtenberg K.H., Olsen J.V., Calogero R.A., Lund A.H. // *Mol. Cell Biol.* 2013. V. 33. № 22. P. 4504–4516.
153. Myslinski E., Krol A., Carbon P. // *J. Biol. Chem.* 1998. V. 273. № 34. P. 21998–22006.
154. Schuster C., Myslinski E., Krol A., Carbon P. // *EMBO*. 1995. V. 14. № 15. P. 3777–3787.
155. Bailey S.D., Zhang X., Desai K., Aid M., Corradin O., Cowper-Sal Lari R., Akhtar-Zaidi B., Seacheri P.C., Haibe-Kains B., Lupien M. // *Nat. Commun.* 2015. V. 2. P. 6186.
156. Heidari N., Phanstiel D.H., He C., Grubert F., Jahanbani F., Kasowski M., Zhang M.Q., Snyder M.P. // *Genome Res*. 2014. V. 24. № 12. P. 1905–1917.
157. Myslinski E., Gerard M.A., Krol A., Carbon P. // *J. Biol. Chem.* 2006. V. 281. № 52. P. 39953–39962.
158. Ngondo-Mbongo R.P., Myslinski E., Aster J.C., Carbon P. // *Nucleic Acids Res.* 2013. V. 41. № 7. P. 4000–4014.
159. Schaub M., Krol A., Carbon P. // *Nucleic Acids Res.* 2000. V. 28. № 10. P. 2114–2121.
160. Schaub M., Myslinski E., Krol A., Carbon P. // *J. Biol. Chem.* 1999. V. 274. № 35. P. 25042–25050.
161. Sathyan K.M., McKenna B.D., Anderson W.D., Duarte F.M., Core L., Guertin M.J. // *Genes Dev.* 2019. V. 33. № 19–20. P. 1441–1455.
162. Mourad R., Cuvier O. // *Nucleic Acids Res.* 2018. V. 46. № 5. P. e27.
163. Wen Z., Huang Z.T., Zhang R., Peng C. // *Cell Biol. Toxicol.* 2018. V. 34. № 6. P. 471–478.
164. Yang Y., Zhang R., Singh S., Ma J. // *Bioinformatics*. 2017. V. 33. № 14. P. i252–i260.
165. Raab J.R., Chiu J., Zhu J., Katzman S., Kurukuti S., Wade P.A., Haussler D., Kamakaka R.T. // *EMBO*. 2012. V. 31. № 2. P. 330–350.
166. van Bortle K., Phanstiel D.H., Snyder M.P. // *Genome Biol.* 2017. V. 18. № 1. P. 180.
167. Layat E., Probst A.V., Tourmente S. // *Biochim. Biophys. Acta*. 2013. V. 1829. № 3–4. P. 274–282.
168. Smith D.R., Jackson I.J., Brown D.D. // *Cell*. 1984. V. 37. № 2. P. 645–652.
169. Matthews N.E., White R. // *BioEssays*. 2019. P. e1900048.
170. Schwartz Y.B., Cavalli G. // *Genetics*. 2017. V. 205. № 1. P. 5–24.
171. Zolotarev N., Maksimenko O., Kyrchanova O., Sokolinskaya E., Osadchiy I., Girardot C., Bonchuk A., Ciglar L., Furlong E.E.M., Georgiev P. // *Nucleic Acids Res.* 2017. V. 45. № 21. P. 12285–12300.
172. Chung H.R., Schafer U., Jackle H., Bohm S. // *EMBO Reports*. 2002. V. 3. № 12. P. 1158–1162.
173. Chung H.R., Lohr U., Jackle H. // *Mol. Biol. Evol.* 2007. V. 24. № 9. P. 1934–1943.
174. Mackeh R., Marr A.K., Fadda A., Kino T. // *Nuclear Receptor Signaling*. 2018. V. 15. P. 1550762918801071.
175. Ecco G., Imbeault M., Trono D. // *Development*. 2017. V. 144. № 15. P. 2719–2729.
176. Jauch R., Bourenkov G.P., Chung H.R., Urlaub H., Reidt U., Jackle H., Wahl M.C. // *Structure*. 2003. V. 11. № 11. P. 1393–1402.
177. Zolotarev N., Fedotova A., Kyrchanova O., Bonchuk A., Penin A.A., Lando A.S., Eliseeva I.A., Kulakovskiy I.V., Maksimenko O., Georgiev P. // *Nucleic Acids Res.* 2016. V. 44. № 15. P. 7228–7241.
178. Gaszner M., Vazquez J., Schedl P. // *Genes Dev.* 1999. V. 13. № 16. P. 2098–2107.
179. Page A.R., Kovacs A., Deak P., Torok T., Kiss I., Dario P., Bastos C., Batista P., Gomes R., Ohkura H., et al. // *EMBO*. 2005. V. 24. № 24. P. 4304–4315.
180. Baxley R.M., Soshnev A.A., Koryakov D.E., Zhimulev I.F.,

- Geyer P.K. // *Dev. Biol.* 2011. V. 356. № 2. P. 398–410.
181. Mohan M., Bartkuhn M., Herold M., Philippen A., Heinel N., Bardenhagen I., Leers J., White R.A., Renkawitz-Pohl R., Saumweber H., et al. // *EMBO.* 2007. V. 26. № 19. P. 4203–4214.
182. Gambetta M.C., Furlong E.E.M. // *Genetics.* 2018. V. 210. № 1. P. 129–136.
183. Baxley R.M., Bullard J.D., Klein M.W., Fell A.G., Morales-Rosado J.A., Duan T., Geyer P.K. // *Nucleic Acids Res.* 2017. V. 45. № 8. P. 4463–4478.
184. Maksimenko O., Kyrchanova O., Klimenko N., Zolotarev N., Elizarova A., Bonchuk A., Georgiev P. // *Biochim. Biophys. Acta Gene Regul. Mech.* 2020. V. 1863. № 1. P. 194446.
185. Li J., Gilmour D.S. // *EMBO.* 2013. V. 32. № 13. P. 1829–1841.
186. 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. № 11. P. 2188–2198.
187. Soshnev A.A., He B., Baxley R.M., Jiang N., Hart C.M., Tan K., Geyer P.K. // *Nucleic Acids Res.* 2012. V. 40. № 12. P. 5415–5431.
188. Maksimenko O., Bartkuhn M., Stakhov V., Herold M., Zolotarev N., Jox T., Buxa M.K., Kirsch R., Bonchuk A., Fedotova A., et al. // *Genome Res.* 2015. V. 25. № 1. P. 89–99.
189. Negre N., Brown C.D., Shah P.K., Kheradpour P., Morrison C.A., Henikoff J.G., Feng X., Ahmad K., Russell S., White R.A., et al. // *PLoS Genet.* 2010. V. 6. № 1. P. e1000814.
190. Negre N., Brown C.D., Ma L., Bristow C.A., Miller S.W., Wagner U., Kheradpour P., Eaton M.L., Loriaux P., Sealfon R., et al. // *Nature.* 2011. V. 471. № 7339. P. 527–531.
191. Baumann D.G., Gilmour D.S. // *Nucleic Acids Res.* 2017. V. 45. № 18. P. 10481–10491.
192. Soshnev A.A., Baxley R.M., Manak J.R., Tan K., Geyer P.K. // *Development.* 2013. V. 140. № 17. P. 3613–3623.
193. Melnikova L., Elizar'ev P., Erokhin M., Molodina V., Chetverina D., Kostyuchenko M., Georgiev P., Golovnin A. // *Sci. Rep.* 2019. V. 9. № 1. P. 5314.
194. Kyrchanova O., Maksimenko O., Stakhov V., Ivlieva T., Parshikov A., Studitsky V.M., Georgiev P. // *PLoS Genet.* 2013. V. 9. № 7. P. e1003606.
195. Kyrchanova O., Chetverina D., Maksimenko O., Kullyev A., Georgiev P. // *Nucleic Acids Res.* 2008. V. 36. № 22. P. 7019–7028.
196. Maeda R.K., Karch F. // *Chromosoma.* 2015. V. 124. № 3. P. 293–307.
197. Kyrchanova O., Mogila V., Wolle D., Magbanua J.P., White R., Georgiev P., Schedl P. // *Mech. Dev.* 2015. V. 138. Pt 2. P. 122–132.
198. Gruzdeva N., Kyrchanova O., Parshikov A., Kullyev A., Georgiev P. // *Mol. Cell Biol.* 2005. V. 25. № 9. P. 3682–3689.
199. Barges S., Mihaly J., Galloni M., Hagstrom K., Muller M., Shanower G., Schedl P., Gyurkovics H., Karch F. // *Development.* 2000. V. 127. № 4. P. 779–790.
200. Iampietro C., Gummalla M., Muterio A., Karch F., Maeda R.K. // *PLoS Genet.* 2010. V. 6. № 12. P. e1001260.
201. Bender W., Lucas M. // *Genetics.* 2013. V. 193. № 4. P. 1135–1147.
202. Bowman S.K., Deaton A.M., Domingues H., Wang P.I., Sadreyev R.I., Kingston R.E., Bender W. // *eLife.* 2014. V. 3. P. e02833.
203. Kyrchanova O., Mogila V., Wolle D., Deshpande G., Parshikov A., Cleard F., Karch F., Schedl P., Georgiev P. // *PLoS Genet.* 2016. V. 12. № 7. P. e1006188.
204. Kyrchanova O., Sabirov M., Mogila V., Kurbidavaeva A., Postika N., Maksimenko O., Schedl P., Georgiev P. // *Proc. Natl. Acad. Sci. USA.* 2019. V. 116. № 27. P. 13462–13467.
205. Wang Q., Sun Q., Czajkowsky D.M., Shao Z. // *Nat. Commun.* 2018. V. 9. № 1. P. 188.
206. Ramirez F., Bhardwaj V., Arrigoni L., Lam K.C., Gruning B.A., Villaveces J., Habermann B., Akhtar A., Manke T. // *Nat. Commun.* 2018. V. 9. № 1. P. 189.
207. Chathoth K.T., Zabet N.R. // *Genome Res.* 2019. V. 29. № 4. P. 613–625.
208. Dorsett D. // *Trends Genet.* 2019. V. 35. № 7. P. 542–551.
209. Perez-Rico Y.A., Barillot E., Shkumatava A. // *iScience.* 2020. V. 23. № 5. P. 101046.