

УДК 577.218

# Белки с цинковыми пальцами типа C2H2 – самый многочисленный и наименее изученный класс транскрипционных факторов высших эукариот

А. А. Федотова, А. Н. Бончук, В. А. Могила, П. Г. Георгиев\*

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

\*E-mail: georgiev\_p@mail.ru

Поступила в редакцию 27.10.2016

Принята к печати 22.02.2017

**РЕФЕРАТ** С появлением методов полногеномного анализа стало возможным широкомасштабное изучение распределения транскрипционных факторов на регуляторных элементах в геноме (энхансерах, промоторах, сайленсерах и инсуляторах), а также выявление ключевых принципов организации хромосом. Однако белковые факторы, отвечающие за архитектуру хромосом и организацию регуляторных доменов, остаются почти неисследованными. В настоящем обзоре предпринята попытка обобщить данные, которые касаются обширного класса белков, имеющих кластеры доменов цинковых пальцев C2H2-типа (C2H2-белки). К этому классу относится хорошо исследованный консервативный белок CTCF, ключевой, согласно современным представлениям, для формирования архитектуры хромосом позвоночных. Отличительными особенностями C2H2-белков являются специфичное и эффективное связывание с уникальными длинными последовательностями ДНК и быстрое распространение в процессе эволюции в пределах таксонов. Приведенные в обзоре данные позволяют предложить модель, согласно которой многие C2H2-белки выполняют функции, сходные с CTCF в организации архитектуры хромосом.

**КЛЮЧЕВЫЕ СЛОВА** KRAB-домен, ZAD-домен, SCAN-домен, CTCF, архитектурные белки, транскрипционные факторы.

**СОКРАЩЕНИЯ** а.о. – аминокислотный остаток (при числе); C2H2-белки – белки, содержащие кластеры цинковых пальцев C2H2-типа; ТАД – топологически ассоциированные домены; ТФ – транскрипционный фактор.

## ВВЕДЕНИЕ

В результате проведенных за последние годы полногеномных исследований внутри- и межхромосомных взаимодействий установлено, что хромосомы человека, мыши и дрозофилы организованы в большие топологические домены (ТАД) [1–4]. Внутри топологических доменов могут происходить дистанционные взаимодействия между промоторами, энхансерами и сайленсерами, что определяет регуляцию экспрессии генов [5, 6]. Однако механизмы, обеспечивающие организацию и поддержание архитектуры хромосом, остаются почти неизученными [7]. Предполагается, что существует особый класс архитектурных белков, инактивация которых значительно влияет на распределение меж- и внутрихромосомных контактов [8, 9].

У позвоночных описан высококонсервативный транскрипционный фактор (ТФ) CTCF, который

считается основным архитектурным белком хромосом [10, 11]. Вместе с когезиновым комплексом CTCF участвует в организации границ топологических доменов, а также поддерживает дистанционные взаимодействия между регуляторными элементами внутри доменов [12–14]. Белок CTCF содержит кластер доменов цинковых пальцев C2H2-типа, часть из которых определяет высокоспецифичное связывание белка с ДНК. Белки с цинковыми пальцами C2H2-типа (C2H2-белки) появились рано в процессе эволюции и найдены у многих эукариотических организмов [15, 16]. Многие из них структурно сходны с CTCF. C2H2-белки можно разделить на три группы [17]: 1) белки, имеющие один-два или много C2H2-доменов, распределенных неупорядоченно, 2) белки, содержащие три C2H2-домена, организованные в кластер на C-конце, 3) белки, содержащие более трех

C2H2-доменов в одном или нескольких кластерах. Наиболее хорошо описана группа консервативных ТФ с тремя C2H2-доменами, многие из которых играют ключевую роль в регуляции экспрессии генов у всех высших эукариот [18, 19]. Настоящий обзор посвящен намного хуже исследованной группе ТФ, содержащих более трех C2H2-доменов.

### СТРУКТУРА И ФУНКЦИОНАЛЬНАЯ РОЛЬ ДОМЕНА C2H2-ТИПА

Цинковые пальцы типа C2H2 (Cys2-His2) – одни из наиболее часто встречаемых доменов в составе ТФ высших эукариот. Классический C2H2-домен длиной 28–30 а.о. состоит из β-шпильки (антипараллельный β-лист, состоящий из двух β-тяжей) и следующей за ней α-спирали, которые формируют левозакрученную ββα-структуру (рис. 1А). Стабилизация структуры цинкового пальца достигается с помощью координационной связи атома цинка с двумя консервативными остатками цистеина с одного конца β-листа и двумя консервативными остатками гистидина с С-концевой части α-спирали. Пары цистеинов и гистидинов являются консервативными, так же как гидрофобное ядро, формирующее α-спираль. Остальные аминокислотные остатки в составе C2H2-доменов отличаются значительной вариабельностью.

Одной из первых была расшифрована структура комплекса из трех тандемно расположенных C2H2-доменов белка Zif268 млекопитающих [20]. Показано, что три цинковых пальца образуют полукруг, располагаясь в большой бороздке ДНК (рис. 1А). Каждый из трех C2H2-доменов связывается с тремя-четырьмя нуклеотидами ДНК, используя аминокислоты

в одних и тех же позициях α-спирали (рис. 1Б) – аргинин в позиции –1, а также аминокислотные остатки в позициях 2, 3 и 6. Биохимическое и структурное изучение C2H2-доменов подтвердило ключевую роль аминокислот в этих позициях для специфичного связывания с ДНК. Согласно канонической модели, аминокислоты в положении 6, 3 и –1 отвечают за узнавание первого, второго и третьего нуклеотида с 5'-конца соответственно, а аминокислота в позиции 2 узнает четвертый нуклеотид, который находится на комплементарной цепи (рис. 1Б).

Структурные исследования C2H2-доменов позволили описать новый принцип распознавания ДНК. Отличительной особенностью белков с C2H2-доменами является их специфическое связывание с длинными последовательностями ДНК (20–40 п.н.), что отличает этот класс белков от всех других ТФ, которые обычно узнают сравнительно короткие выроджденные ДНК. Обычно тандемные C2H2-домены, участвующие в узнавании ДНК, разделены консервативными участками, состоящими из 5 а.о. [21]. Разработаны алгоритмы, которые позволяют с высокой степенью вероятности предсказывать последовательность, с которой может связываться кластер C2H2-доменов, и, наоборот, подбирать комбинации C2H2-доменов, узнающих целевые последовательности ДНК [22, 23]. Однако в больших кластерах (более трех C2H2-доменов) интерференция между соседними C2H2-доменами усложняет точное предсказание сайта связывания [24].

В отличие от универсального механизма взаимодействия C2H2-доменов с ДНК, контакты с белками и РНК могут осуществляться с помощью разных ком-

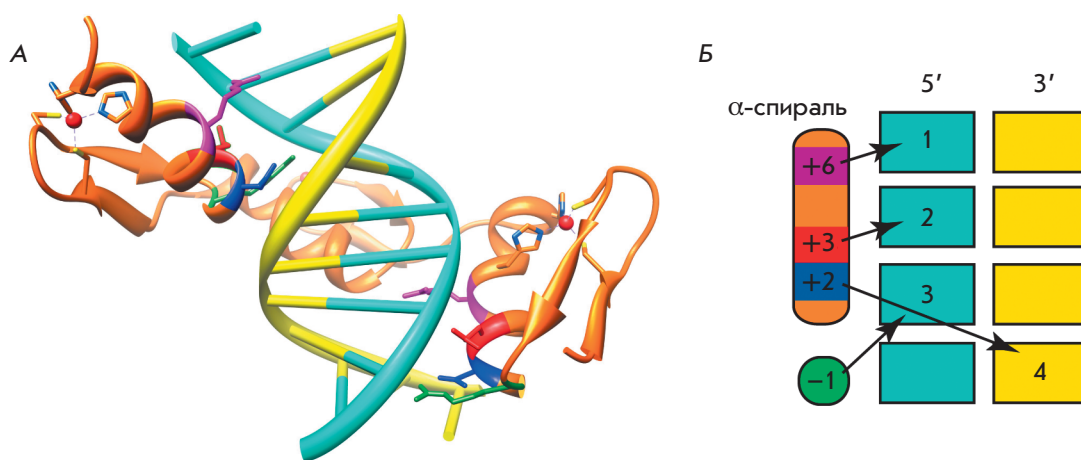
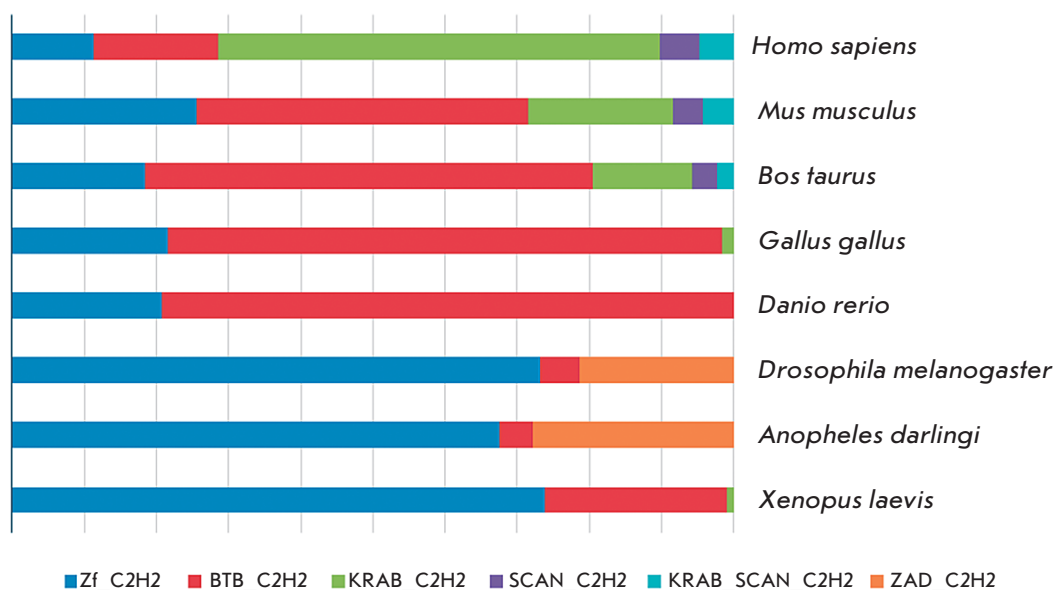


Рис. 1. Модель специфичного связывания C2H2-доменов с ДНК. А – кристаллическая структура комплекса трех цинковых пальцев белка Zif268 и ДНК [20]. Цветом обозначены аминокислотные остатки, участвующие в специфичном узнавании ДНК (–1 – зеленый, +2 – синий, +3 – красный, +6 – фиолетовый). Б – модель узнавания нуклеотидов сайта связывания аминокислотами α-спирали. Адаптировано из [24]



**Рис. 2.** Соотношение различных вариантов С2Н2-белков у различных видов высших эукариот: человек (*Homo sapiens*), мышь (*Mus musculus*), дикий бык (*Bos taurus*), курица (*Gallus gallus*), данио (*Danio rerio*), дрозофила (*Drosophila melanogaster*), малярийный комар (*Anopheles darlingi*), гладкая шпорцевая лягушка (*Xenopus laevis*). Данные получены из базы Uniprot

бинаций аминокислот, что описано в других обзорах [25, 26]. Обычно С2Н2-кластеры находятся в центре или на С-конце белка. Большая часть белков, содержащих С2Н2-кластер в центральном положении, не имеют других консервативных доменов. В то же время белки с кластером на С-конце часто содержат дополнительные N-концевые домены (рис. 2). У позвоночных можно выделить KRAB- и SCAN-домены, а у насекомых – ZAD-домен [27, 28]. Небольшая группа С2Н2-белков имеет на N-конце консервативный ВТВ/POZ-домен. Этот домен часто встречается в составе разных классов белков, поэтому мы исключили описание этой группы С2Н2-белков из настоящего обзора. Обобщенную информацию по ВТВ-содержащим белкам можно найти в подробных обзорах [29, 30].

### ТРАНСКРИПЦИОННЫЕ ФАКТОРЫ, СОДЕРЖАЩИЕ ТОЛЬКО КЛАСТЕР С2Н2-ДОМЕНОВ

Группа ТФ, которые содержат только кластер С2Н2-доменов, включает наиболее изученный и консервативный белок СТСФ (СССТС-связывающий фактор) [31], впервые описанный как негативный регулятор экспрессии гена *myc* [32]. Позже сайт связывания СТСФ был найден в составе инсулятора, находящегося на границе β-глобинового локуса курицы [33]. Сайты связывания СТСФ часто расположены на границах хромосомных районов с разным эпигенетическим статусом и транскрипционной активностью, а также на границах топологически ассоциированных доменов (ТАД), пространственно разделяющих хромосомы на районы,

внутри которых происходит взаимодействие регуляторных элементов [34–37].

Белок СТСФ – один из немногих консервативных белков, содержащих кластер С2Н2-доменов. Гомолог СТСФ у дрозофилы (dСТСФ) также часто находится на границах ТАД и в составе инсуляторов [38, 39]. В модельных трансгенных системах dСТСФ поддерживает дистанционные взаимодействия между промотором репортерного гена и активатором GAL4 [40, 41]. На N-конце белка dСТСФ обнаружен домен, способный к гомодимеризации (рис. 3А), который, вероятно, нужен для поддержания дистанционных взаимодействий между удаленными друг от друга сайтами связывания белка dСТСФ [42]. Попытки найти аналогичный димеризующийся домен в белках СТСФ позвоночных пока не привели к результату. В *in vitro* экспериментах показано, что С-концевой домен СТСФ связывается с кластером С2Н2-доменов [43]. Однако специфичность такого взаимодействия не доказана.

Согласно общепринятой модели, когезиновый комплекс, необходимый для спаривания гомологичных хромосом в процессе клеточного деления [44], связывается с белком СТСФ и участвует в поддержании специфичных дистанционных взаимодействий между его сайтами в интерфазных хромосомах (рис. 3Б). В С-концевом домене СТСФ человека картирован участок, который взаимодействует с одним из белков когезинового комплекса [44].

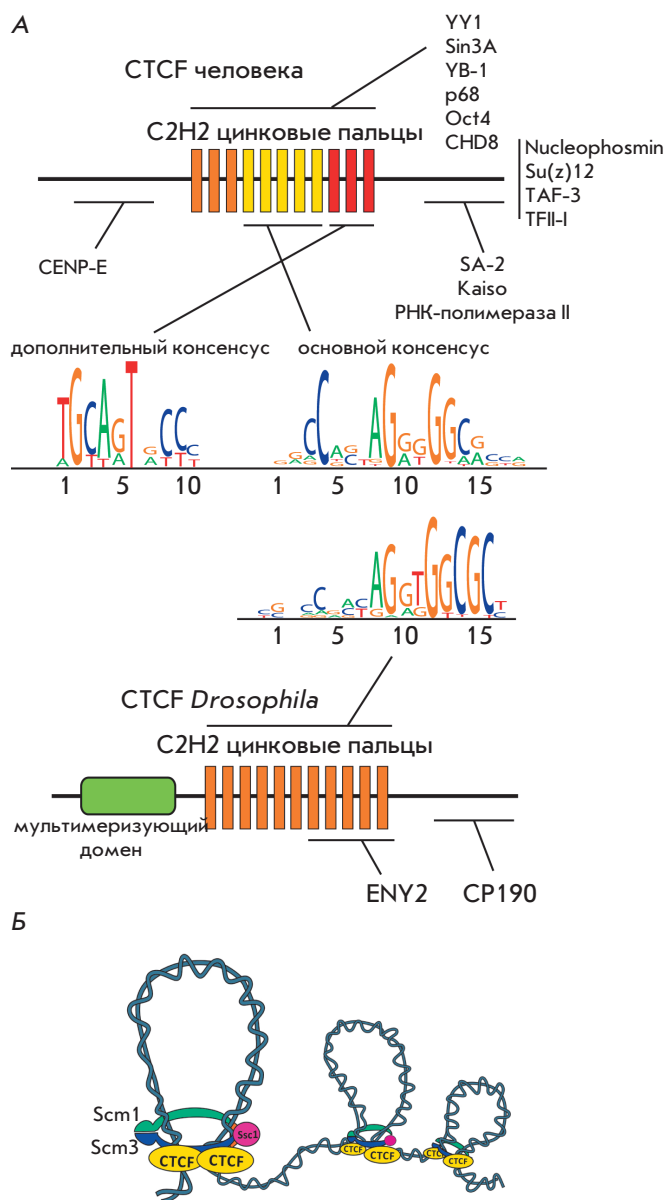
Связывание СТСФ с ДНК, консервативное даже между насекомыми и млекопитающими, подробно изучено у многих высших эукариот [15, 45]. С2Н2-домены 4–7 белка СТСФ (рис. 3А) участвуют в свя-

зывании с основным консенсусным сайтом [46, 47]. Примерно 20% сайтов содержат второй мотив из 10 п.н., для связывания с которым нужны С2Н2-домены 9–11 [47, 48]. Предполагается, что второй мотив, расположенный на расстоянии 5–6 п.н., увеличивает стабильность связывания СТСФ с ДНК.

Транскрипционный фактор СТСФ вовлечен во многие процессы, такие, как эмбриональное развитие, инактивация X-хромосомы у самок, регуляция рекомбинации между генными кластерами при сборке зрелых генов иммуноглобулинов, регуляция альтернативного сплайсинга [34–37, 49–51]. Показано, что СТСФ взаимодействует с большим количеством белков (рис. 3А), таких, как Smad [52], факторы основного комплекса транскрипции TFII-I [53] и TAF-3 [54], хеликаза р68, содержащая домен с DEAD-боксом [55], нуклеофосмин, Kaiso [56], ТФ YB-1, YY1 и Oct4 [57–59], хеликаза CHD8 [60], Su(z)12 – компонент Polycomb-комплекса 2 (PRC2) [61], компонент Sin3A деацетилазного комплекса [62], CENP-E [48] и многие другие белки [49]. В большинстве случаев в белок-белковых взаимодействиях участвуют отдельные С2Н2-домены белка СТСФ [49]. Вероятно, участие СТСФ в различных процессах (активация и репрессия транскрипции, организация дистанционных взаимодействий, формирование ТАД) во многом определяется формированием альтернативных комплексов с белками-партнерами.

Получены экспериментальные данные, показывающие, что с СТСФ связывается большое количество РНК, модулирующих его активность. РНК-связывающий домен СТСФ представляет собой комбинацию участка С-концевого домена и двух С2Н2-доменов (10 и 11), которые *in vitro* неспецифично узнают РНК [63, 64]. Предполагается, что взаимодействие с некоторыми РНК может увеличивать способность СТСФ формировать мультимерные комплексы [63] или снижать стабильность взаимодействия СТСФ с ДНК [11]. Активность СТСФ также регулируется различными посттрансляционными модификациями: поли(ADP)-рибозилированием [65], фосфорилированием [66] и сумоилированием [67].

Белок СТСФ достаточно хорошо изучен и служит примером ТФ, включающего кластер С2Н2-доменов и неструктурированные N- и С-концевые участки. Большая часть других С2Н2-белков имеет аналогичную структуру, но при этом их функции и свойства пока не исследованы. Можно предположить, что некоторые С2Н2-белки могут выполнять функции, сходные с СТСФ. Интересно, что у дрозофилы мутанты по гену *ctcf* доживают до взрослой стадии, что предполагает присутствие в геномах насекомых других транскрипционных факторов, замещающих функции СТСФ [42].



**Рис. 3.** Сравнение структуры и свойств белков СТСФ человека и дрозофилы. А – доменная структура белков СТСФ человека и *D. melanogaster*. Отмечены домены, участвующие в белок-белковых взаимодействиях и узнавании ДНК. Белки СТСФ дрозофилы [135] и человека [46] имеют похожие консенсусные сайты связывания. Б – механизм дистанционных взаимодействий при участии когезинового комплекса и белка СТСФ

**ТРАНСКРИПЦИОННЫЕ ФАКТОРЫ С КЛАСТЕРОМ С2Н2-ДОМЕНОВ И N-КОНЦЕВЫМ КРАВ-ДОМЕНОМ**

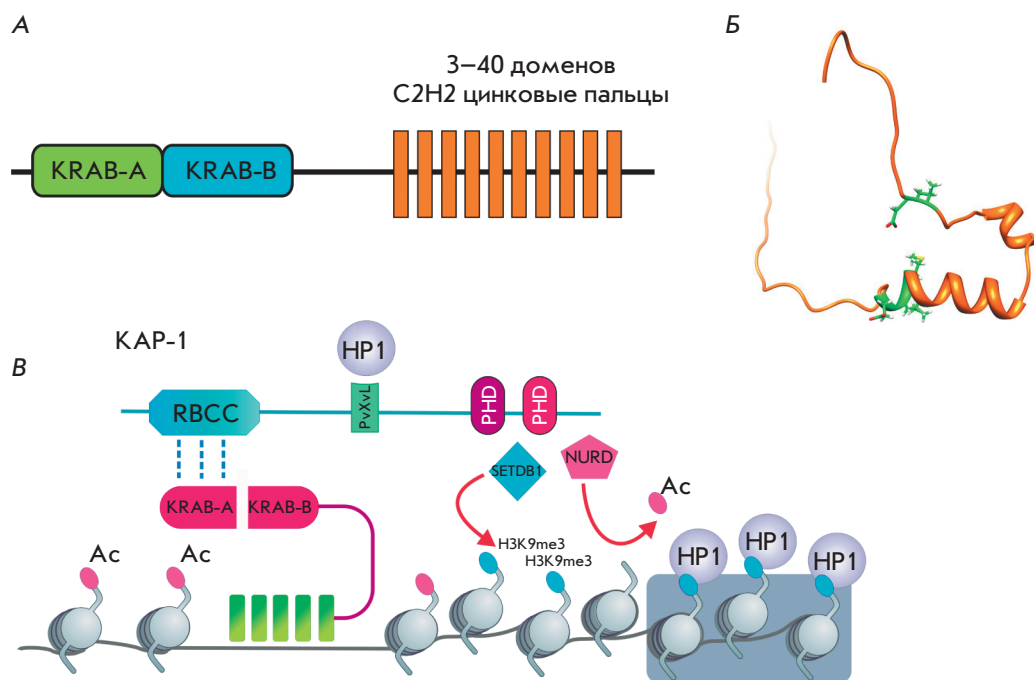
Около одной трети белков человека с кластером С2Н2-доменов содержат на N-конце КРАВ-домен (Krüppel-associated box) (рис. 4А) [68]. В общей сложности у человека найдено 742 разных С2Н2-белка



с КРАВ-доменом, которые кодируются 423 генами [69]. При этом 384 гена сгруппированы в 25 хромосомных кластеров, и только 39 КРАВ-С2Н2-белков кодируются одиночно расположенными генами. Белки с КРАВ-доменами обнаружены только у четвероногих животных. Кластеризация на хромосомах и распространение в пределах крупных таксонов позволяют предположить, что данное семейство генов произошло в результате дупликаций, которые закрепились в процессе эволюционного отбора [70]. КРАВ-домен состоит приблизительно из 75 а.о. и структурно может быть разделен на два субдомена – А и В, которые, как предсказано, сворачиваются в две амфипатические  $\alpha$ -спирали (рис. 4Б). Субдомены КРАВ-А и КРАВ-В всегда кодируются отдельными экзонами. В результате альтернативного сплайсинга образуются мРНК, кодирующие либо только субдомен КРАВ-А, либо одновременно оба субдомена – КРАВ-А и КРАВ-В, разделенные спейсером переменной длины. У человека белки КРАВ могут содержать от 2 до 40 С2Н2-доменов. В отличие от генов других семейств, С2Н2-домены КРАВ-белков чаще всего кодируются одним экзоном [71].

Белки КРАВ-С2Н2 широко представлены в геномах четвероногих и многие из них участвуют в репрессии транскрипции [70]. Универсальный и хорошо изученный механизм репрессии связан с привлечением белка КАР-1 (КРАВ-Associated Protein-1), единственного описанного кофактора всех исследованных КРАВ-белков, репрессирующих транс-

крипцию. КРАВ-А-домен непосредственно взаимодействует с КАР-1, который, в свою очередь, служит платформой для привлечения репрессорных комплексов (рис. 4В). Пять аминокислот (рис. 4Б), консервативных во всех КРАВ-А-доменах млекопитающих (DV в положении 6,7 и MLE в положениях 36–38), нужны для связывания КАР-1 [72, 73]. Функциональная роль КРАВ-В-субдомена остается неизученной. Предполагается, что этот домен увеличивает эффективность привлечения КАР-1-зависимого репрессорного комплекса [74]. На N-конце белка КАР-1 находится домен RBCC (Ring Finger/Box/Coiled-Coil), который обеспечивает связывание КАР-1 с КРАВ-доменом. Центральная часть КАР-1 содержит гидрофобный пентапептид, взаимодействующий с Chromo-Shadow (CS)-доменом белка HP1. На С-конце белка расположены два PHD-домена, которые рекрутируют комплексы, участвующие в деацетилировании (NURD) и метилировании (SETDB1) гистонов [70, 75–77]. Репрессия, инициируемая в районе связывания КРАВ-С2Н2-белков, может распространяться в окружающие участки генома на десятки тысяч нуклеотидов в результате последовательного внесения метки H3K9me3 и связывания с ней репрессора HP1 [78–80]. Пик экспрессии белка КАР-1 приходится на ранние эмбриональные стадии, и репрессия транскрипции, осуществляемая КРАВ-С2Н2-белками, критична для раннего эмбрионального развития. На более поздних стадиях репрессия может поддерживаться в соматических



**Рис. 4.** Структура и свойства КРАВ-домена. А – типичная доменная организация КРАВ-С2Н2-белков. Б – структура КРАВ-А – зеленым отмечены 5 а.о., консервативных во всех КРАВ-А-доменах млекопитающих (DV в положении 6,7 и MLE в положениях 36–38) и необходимых для связывания КАР-1 [PDB 1V65]. В – схема привлечения КАР-1 и репрессорных комплексов

клетках с помощью эпигенетических механизмов без непосредственного участия KRAB-C2H2-белков [81, 82].

Большая часть KRAB-C2H2-белков видов и родоспецифичны. У некоторых позвоночных, птиц, ящериц и лягушек KRAB-A-домен имеет множественные замены аминокислот, необходимых для взаимодействия с KAP-1 [31, 72]. Это можно объяснить либо тем, что KRAB-домен у них выполняет другие функции, либо эти классы позвоночных имеют измененный KAP-1, который сохраняет способность взаимодействовать с KRAB-доменом. В целом, эволюционный анализ консервативности KRAB-C2H2-белков показал, что у позвоночных каждого класса происходило независимое формирование семейства генов KRAB-C2H2, что подтверждает высокую эволюционную скорость появления новых генов этого класса.

Известны только три гена KRAB-C2H2, входящих в один кластер и общих для всех исследованных видов позвоночных. Эти гены кодируют C2H2-белки, содержащие KRAB-домен с измененной структурой, который перестает связывать репрессор KAP-1 и участвует в стимуляции транскрипции [31, 83, 84]. Наиболее интересен высококонсервативный ген *Meisetz* (PRDM9), который имеет не только модифицированный KRAB-домен, но и домен SET [85, 86]. С помощью совместного действия SET-и KRAB-доменов происходит рекрутирование H3K4-метилтрансферазы, отвечающей за триметилирование гистона H3 по лизину 4 (H3K4me3). Модификация H3K4me3 в области промоторов обычно коррелирует с активной транскрипцией. Биоинформатический анализ показал, что часть KRAB-домена, кодируемого геном *Meisetz*, имеет гомологию с KRI-мотивом, присутствующим в геномах всех хорошо изученных эукариот, включая арабидопсис, рис, грибы и дрожжи [31, 86]. Широкое распространение мотива KRI предполагает, что KRAB-домен белка *Meisetz* мог произойти из этого мотива в результате присоединения дополнительных аминокислотных остатков. Приобрести функции репрессора KRAB-домен мог в результате случайных мутаций, которые позволили репрессору связываться с KAP-1, что закрепилось в процессе эволюции.

Экспериментально показано, что у позвоночных TF класса KRAB-C2H2 играют важную роль во многих процессах эмбрионального развития, клеточной дифференцировки и пролиферации, регуляции клеточного цикла и апоптоза [70, 73]. Сайты связывания KRAB-C2H2-белков коррелируют с участками открытого (свободного от нуклеосом) хроматина, что объясняется связыванием с активными регуляторными районами генов [87]. Полногеномные исследова-

ния показывают, что белки KRAB-C2H2 связываются с энхансерами и промоторами генов и могут в некоторых случаях активировать транскрипцию [88–90]. Способность KRAB-C2H2-белков активировать транскрипцию должна коррелировать с супрессией взаимодействия KRAB-домена с белком KAP-1. Механизм такой супрессии пока остается неисследованным, но, вероятно, связан с обратимыми модификациями аминокислотных остатков KRAB-домена. Важную роль могут играть и C2H2-домены, которые потенциально способны рекрутировать отдельные TF и целые комплексы, позитивно/негативно регулирующие транскрипцию.

Интересной функцией обладает уже описанный выше белок *Meisetz* (PRDM9), который экспрессируется только в гонадах млекопитающих [91]. Оказалось, что большая часть горячих точек рекомбинации у млекопитающих содержит вероятный сайт связывания белка PRDM9 [92]. Быстрые эволюционные изменения в первичной структуре C2H2-доменов и их количестве привели к тому, что у разных млекопитающих PRDM9 связывается с разными нуклеотидными последовательностями ДНК [91, 93–96]. Связывание PRDM9 приводит к формированию свободного от нуклеосом района и модификации H3K4me3 окружающих нуклеосом [97]. Предполагается, что комплекс SPO11, индуцирующий двухцепочечные разрывы, одновременно узнает гистоновую метку H3K4me3 и непосредственно связывается с белком PRDM9 [98].

Недавно была открыта новая функциональная роль KRAB-C2H2-белков в репрессии транскрипции чужеродной ДНК, прежде всего эндогенных ретровирусов, мобильных элементов LINE и SINE [79, 87, 99, 100]. Мобильные элементы составляют значительную часть генома млекопитающих, и репрессия их транскрипции жизненно необходима [101]. Разные KRAB-C2H2-белки связываются с регуляторными областями мобильных элементов и ретровирусов определенного типа и индуцируют их эпигенетическую репрессию. Существует гипотеза, согласно которой вновь возникшие белки KRAB-C2H2 закрепляются эволюционным отбором, так как играют критическую роль в подавлении экспрессии новых мобильных элементов, в то время как более консервативные KRAB-C2H2 участвуют в регуляции экспрессии клеточных генов [79].

Другим объяснением быстрой эволюции генов, кодирующих KRAB-C2H2-белки, может быть их ключевая роль в контроле экспрессии генов, определяющих развитие нервной [102] и кровеносной [103] систем у млекопитающих. Например, многие гены, кодирующие KRAB-C2H2-белки, специфичные для человека и приматов, активно транскрибируются

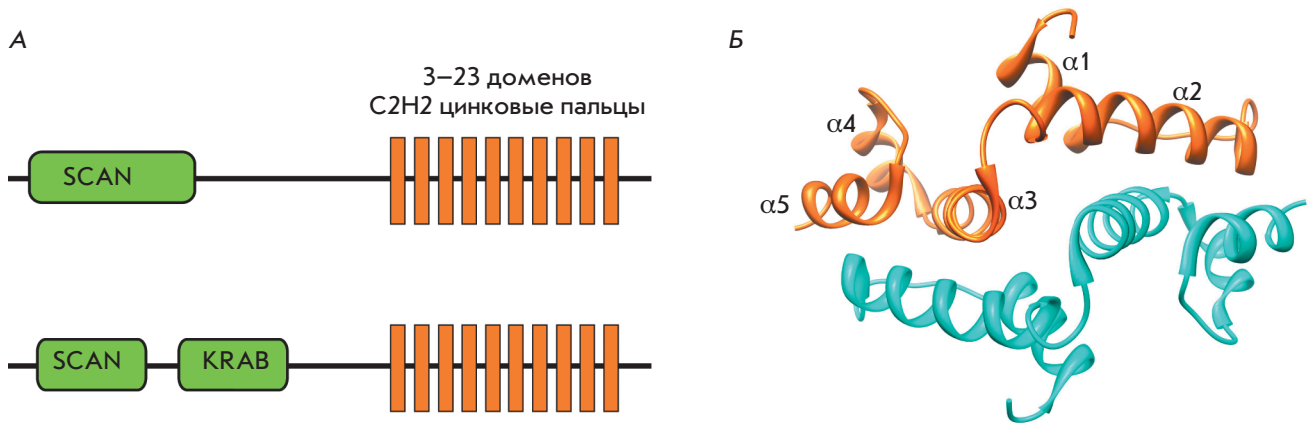


Рис. 5. Структура и свойства SCAN-домена. А – типичная доменная организация SCAN-C2H2- и SCAN-KRAB-C2H2-белков. Б – кристаллическая структура димера SCAN-домена белка Zfp206 [110]

в мозге [102]. Однако не существует прямой корреляции между количеством генов KRAB-C2H2 и усложнением организма. Так, например, опоссум имеет в 2 раза больше генов KRAB-C2H2, чем человек [31]. Можно надеяться, что появление новых технологий получения специфичных антител, полногеномного анализа и мутаций отдельных генов с помощью системы CRISPR/Cas9 позволит в ближайшем будущем прояснить функциональную роль KRAB-C2H2-белков.

#### ТРАНСКРИПЦИОННЫЕ ФАКТОРЫ С КЛАСТЕРОМ C2H2-ДОМЕНОВ И N-КОНЦЕВЫМ SCAN-ДОМЕНОМ

SCAN-домен был впервые описан в составе ТФ ZNF174 человека [114] (рис. 5А). В дальнейшем белки с доменами такого типа обнаружили и у представителей ряда других классов позвоночных [104]. У человека, мыши и коровы найдено 71, 38 и 28 SCAN-C2H2-белков соответственно [94]. Гены, кодирующие SCAN-белки, обычно располагаются в геноме небольшими кластерами (по два–семь) [104]. Находящиеся в кластерах SCAN-домены имеют более высокую степень гомологии между собой, что предполагает их появление в результате дубликации генов и последующей адаптивной эволюции. Примерно половина генов кодирует одновременно SCAN- и KRAB-домены и обычно находится в больших кластерах вместе с генами, кодирующими только KRAB-C2H2-белки (рис. 5А) [27, 94].

Структура SCAN-домена имеет высокую степень сходства с С-концевыми доменами белка капсида вируса иммунодефицита человека [105] и белка Gag ретротранспозонов семейства *gypsy/Ty3* [27]. На основе этих данных предположили, что исходно SCAN-домены возникли из капсидных белков ретровирусов

у низших позвоночных, далее в процессе эволюции этот домен приобрел новую функцию в составе ТФ, содержащих кластеры C2H2-доменов [106]. KRAB-домен в комбинации со SCAN-доменом присутствует у млекопитающих и ящериц, но его нет у курицы и лягушки.

На сегодняшний день описаны пространственные структуры SCAN-доменов белков Zfp206 [107], PEG3 [108], ZNF24, ZNF174 [105] и MZF-1 [109], которые имеют высокую степень сходства (рис. 5Б). Особенности пространственной структуры можно рассмотреть на примере SCAN-домена белка Zfp206 [107], который формирует антипараллельный гомодимер. Каждый мономер в составе гомодимера состоит из пяти  $\alpha$ -спиралей. Ядро внутренней поверхности гомодимера образуется упаковкой второй  $\alpha$ -спирали одного мономера вокруг третьей и пятой  $\alpha$ -спиралей противоположного мономера и наоборот. N-Концевая первая  $\alpha$ -спираль обеспечивает дополнительные контакты одного мономера с третьей  $\alpha$ -спиралью противоположного мономера (рис. 5Б). Все SCAN-домены могут гомодимеризоваться, но только некоторые SCAN-домены способны образовывать гетеродимеры [104, 111–113]. Первая  $\alpha$ -спираль SCAN-домена отличается наибольшей вариабельностью последовательности гидрофобных аминокислотных остатков и рассматривается как возможный кандидат, детерминирующий образование гетеродимеров между различными SCAN-доменами. Так, показана способность SCAN-домена белка Zfp206 к гетеродимеризации с аналогичным доменом Zfp110 [107]. Замена первой  $\alpha$ -спирали на  $\alpha$ -спираль гетерологичного SCAN-домена ZNF174 или ее удаление приводят к потере способности этих SCAN-доменов к гетеродимеризации.

Известно, что KRAB-A-домен рекрутирует репрессорный комплекс, тогда как данные о влиянии SCAN-домена на транскрипцию отсутствуют [104, 111]. Существуют только отрывочные сведения о функциональной роли ТФ SCAN-C2H2. Например, ТФ ZNF263 человека, содержащий SCAN- и KRAB-домены и 9 C2H2-доменов, преимущественно связывается с промоторными областями и способен участвовать как в активации, так и в репрессии транскрипции [89]. Другой представитель семейства – белок ZNF658, также содержит SCAN- и KRAB-домены и участвует в активации экспрессии генов рРНК, которые транскрибируются РНК-полимеразой I [115]. Можно предположить, что основная функция SCAN-белков связана со способностью к формированию гомо- и гетеродимеров между SCAN-доменами.

### ТРАНСКРИПЦИОННЫЕ ФАКТОРЫ С КЛАСТЕРОМ C2H2-ДОМЕНОВ И N-КОНЦЕВЫМ ZAD-ДОМЕНОМ

ZAD (zinc-finger-associated domain)-домен (рис. 6А) обнаружен на N-конце C2H2-белков многих членистоногих [28]. У позвоночных найден только один белок, содержащий на N-конце структуру, имеющую сходство с ZAD-доменом [116]. В геномах *Anopheles gambiae*, *Drosophila melanogaster* и *Apis mellifera* (медоносная пчела) найдено 147, 98 и 29 ZAD-C2H2-белков соответственно [116], тогда как у ракообразных (*Daphnia pulex*) обнаружено только четыре гена, кодирующих ZAD-подобные домены. Обычно гены, кодирующие ZAD-домены с более высокой гомологией, образуют небольшие кластеры. Предполагается, что эти гены возникли в результате множественных дубликаций исходных копий с последующим их закреплением в результате положительного отбора [28, 116]. Вероятно, эволюционный процесс шел очень бы-

стро, так как только у нескольких белков ZAD-C2H2 найдены очевидные гомологи у удаленных друг от друга видов дрозофилы [116].

Размер ZAD-домена варьирует в пределах 71–97 а.о. Множественное выравнивание последовательностей 32 членов семейства показало, что этот домен состоит из четырех консервативных блоков, соединенных варьирующими по длине районами [28]. Отличительной особенностью ZAD-доменов является присутствие двух инвариантных пар цистеинов, участвующих в координации иона цинка.

К настоящему моменту получена кристаллическая структура только одного ZAD-домена белка Grauzone (Grau) [117], которая может быть прототипом для всех ZAD-структур (рис. 6Б). N-Концевая часть ZAD-домена представляет собой глобулярную структуру вокруг иона цинка, а C-концевой стебель сформирован длинной  $\alpha$ -спиралью 2 ( $\alpha 2$ ), которая составляет почти одну треть от всех аминокислот в ZAD-домене. Сворачивание ZAD-домена в значительной степени зависит от координации ионом цинка двух пар цистеинов, разделенных приблизительно 50 а.о., что приводит к стягиванию участков  $\beta 2$ - $\alpha 2$  и N-конца к центру домена.

В кристалле две молекулы ZAD-домена ассоциированы в виде антипараллельного димера. Большая часть аминокислотных остатков, консервативных в ZAD-семействе [28], отвечает за контакты между двумя субъединицами. ZAD-домен Grau имеет отрицательно заряженную поверхность, что указывает на его неспособность связывать ДНК [117]. Предполагается, что основной функцией ZAD-домена является образование гомодимеров C2H2-белков [118]. ZAD-домены также участвуют в регуляции ядерной локализации некоторых ZAD-C2H2-белков [119].

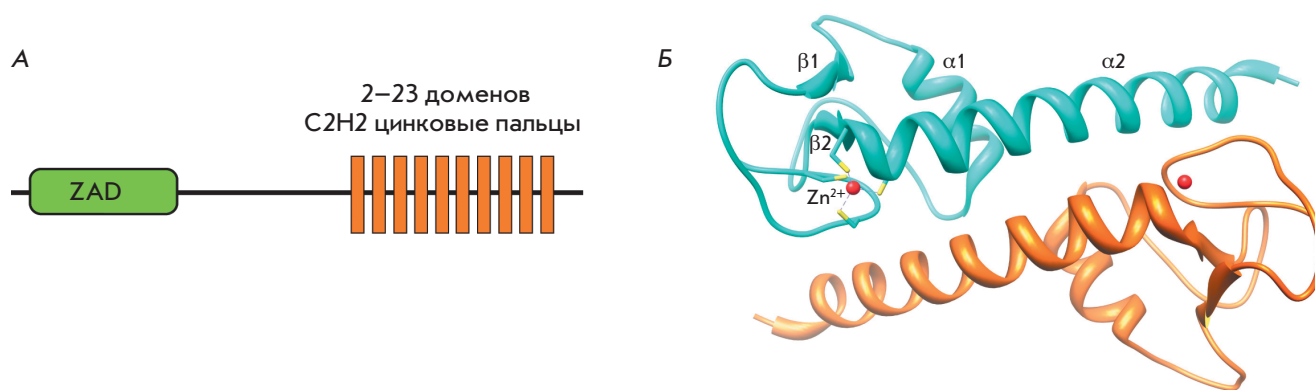


Рис. 6. Структура и свойства ZAD-домена. А – типичная доменная организация ZAD-C2H2-белков. Б – кристаллическая структура димера ZAD-домена белка Grau [117]



Белки с ZAD-доменами составляют примерно одну треть от общего числа белков с C2H2-кластерами и одну десятую часть от всех ТФ в геноме *D. melanogaster* [28]. К настоящему моменту изучены функции только небольшой части ТФ ZAD-C2H2. Большинство ZAD-C2H2-белков экспрессируются в процессе оогенеза и в раннем эмбриогенезе [116]. На важную функциональную роль ZAD-C2H2-белков в процессе развития дрозофилы указывают результаты ряда исследований.

Белок, связывающий мотив 1 (Motif 1 Binding Protein, M1BP), экспрессируется на высоком уровне во всех тканях и на всех стадиях развития дрозофилы и является ключевым в организации архитектуры более 2000 промоторов дрозофилы, имеющих характерный мотив (T/C)GG(T/C)CACACTG [120].

В трансгенных линиях дрозофилы три ZAD-C2H2-белка – Pita, ZIPIC и Zw5 – проявляют свойства инсуляторных/архитектурных белков: блокируют взаимодействие между энхансером и промотором и поддерживают дистанционные взаимодействия [118, 121–124]. ZAD-домены этих белков формируют только стабильные гомодимеры [118]. Интересно, что фрагменты ДНК, содержащие сайты связывания разных ZAD-C2H2-белков не поддерживают дистанционные взаимодействия, что предполагает ключевую роль димеризации ZAD-доменов в организации специфичных контактов между удаленными участками хроматина. Действительно, ZAD-домен белка ZIPIC нужен для поддержания дистанционных взаимодействий между активатором GAL4 и промотором репортерного гена в дрожжах *Saccharomyces cerevisiae* [118]. Как и в случае M1BP, сайты связывания белков ZIPIC, Pita и Zw5 преимущественно находятся в непосредственной близости от стартов транскрипции [118, 125], что предполагает архитектурную функцию этих белков в организации промоторов. Нуль-мутации в генах, кодирующих белки Pita и Zw5, приводят к поздней эмбриональной – ранней личиночной летальности, что указывает на важную роль этих белков в процессе эмбрионального развития [121, 126].

Белок Grau экспрессируется на всех стадиях развития дрозофилы, он обнаруживается в ядрах питающих и фолликулярных клеток, окружающих ооцит [127, 128]. Мутации гена этого белка приводят к остановке оогенеза на стадии II мейоза, что связано с ролью Grau в активации промотора гена *cortex*, регулирующего мейоз в ооцитах [127, 128]. Белок Serendipity delta (Sry  $\delta$ ) связывается с промотором ключевого гена раннего эмбриогенеза *bicoid* и стимулирует его транскрипцию [129]. Нуль-мутации в гене *sry*  $\delta$  проявляются как эмбриональные летали, что указывает на значимость Sry  $\delta$  в раннем развитии дрозофилы [130].

Белок Trade Embargo (Trem) экспрессируется преимущественно в половых клетках дрозофилы и, вероятно, выполняет сходную функцию с белком PRDM9 у млекопитающих [95, 97, 131]. Белок Trem определяет места посадки белка Mei-P22, который участвует в индукции разрывов на мейотических хромосомах [131]. Белок Mei-P22 вместе со своим партнером, Mei-W68, участвует в образовании двухцепочечных разрывов, инициирующих кроссинговер в процессе мейоза [132–134]. Согласно модели, Trem совместно с партнерами создает области открытого хроматина, которые привлекают с помощью специфичных белок-белковых взаимодействий комплекс Mei-P22/Mei-W68, индуцирующий двухцепочечные разрывы [131].

В целом, имеющиеся данные показывают, что группа ТФ ZAD-C2H2 играет важную роль в организации структуры и функциональной активности промоторов, рекрутировании белковых комплексов и формировании архитектуры хромосом.

## ЗАКЛЮЧЕНИЕ

В настоящее время нерешенными остаются многие вопросы, связанные с регуляцией транскрипции генов, организацией структуры регуляторных элементов и механизмами дистанционных взаимодействий. Также совершенно очевидно, что CTCF позвоночных не может быть единственным и ключевым ДНК-связывающим белком, определяющим архитектуру хромосом позвоночных [7].

Отличие от хорошо изученных ТФ других классов, белки C2H2 специфично связываются с длинными участками ДНК, достигающими нескольких десятков пар нуклеотидов. Белки C2H2 могут эффективно связываться с ДНК в виде мономеров, в отличие от большинства других ТФ, которые связываются с короткими палиндромными последовательностями как гомо- или гетеродимеры. Часть C2H2-доменов в комбинации с неструктурированными участками C2H2-белков может обеспечивать многообразные взаимодействия с белковыми комплексами, отдельными ТФ и РНК. Таким образом, C2H2-белки можно рассматривать в качестве перспективных кандидатов на роль организаторов архитектуры регуляторных элементов, таких, как промоторы, энхансеры, инсуляторы и сайленсеры. К сожалению, отсутствие достаточного количества экспериментальных данных не позволяет подтвердить правильность этого предположения для позвоночных. С другой стороны, всесторонне изученный белок CTCF позвоночных обладает рядом свойств (специфичное связывание с сайтом на ДНК, формирование открытых участков хроматина, рекрутирование белковых комплексов, организация дис-

танционных взаимодействий), которые могут быть экстраполированы на другие С2Н2-белки.

Наконец, многие С2Н2-белки имеют гомодимеризующиеся домены. Интересно, что у членистоногих и позвоночных произошло распространение разных доменов – ZAD и SCAN соответственно. Основное объединяющее свойство ZAD- и SCAN-доменов – их способность к преимущественному формированию гомодимеров. Показано, что гомодимеризующиеся ZAD-домены трех ZAD-С2Н2-белков (Pita, ZIPIC, Zw5) определяют специфичность дистанционных взаимодействий [118]. Вероятно, аналогичными свойствами обладают и другие ZAD-С2Н2-белки. Пока получены только отдельные данные о роли SCAN-С2Н2-белков в организации активных промоторов позвоночных. Возможно, что, помимо ZAD- и SCAN-доменов, С2Н2-белки имеют и другие домены, способные к мультимеризации, как, например, N-концевой домен белка dCTCF дрозофилы [42].

Таким образом, имеющиеся фрагментарные данные уже сейчас позволяют предложить модель, согласно которой С2Н2-белки играют роль посредников в передаче информации от нуклеотидной последовательности регуляторных элементов (промоторов, энхансеров, сайленсеров) к белковым комплексам, которые определяют свойства регуляторных элементов. Предполагается, что исследование отдельных представителей этого обширного класса ТФ, выяснение функциональной роли ZAD-, SCAN- и КРАВ-доменов, идентификация новых белков-партнеров и димеризующихся доменов позволит оценить реальный вклад С2Н2-белков в формирование архитектуры хромосом и структуры регуляторных элементов. ●

*Работа выполнена при поддержке Российского  
научного фонда (проект № 14-24-00166).*

#### СПИСОК ЛИТЕРАТУРЫ

- Dixon J.R., Selvaraj S., Yue F., Kim A., Li Y., Shen Y., Hu M., Liu J.S., Ren B. // *Nature*. 2012. V. 485. P. 376–380.
- Nora E.P., Lajoie B.R., Schulz E.G., Giorgetti L., Okamoto I., Servant N., Piolot T., van Berkum N.L., Meisig J., Sedat J., et al. // *Nature*. 2012. V. 485. P. 381–385.
- Razin S.V., Gavrilov A.A., Vassetzky Y.S., Ulianov S.V. // *Transcription*. 2016. V. 7. P. 84–90.
- Dekker J., Rippe K., Dekker M., Kleckner N. // *Science*. 2002. V. 295. P. 1306–1311.
- 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. P. 1665–1680.
- Lupianez D.G., Kraft K., Heinrich V., Krawitz P., Brancati F., Klopocki E., Horn D., Kayserili H., Opitz J.M., Laxova R., et al. // *Cell*. 2015. V. 161. P. 1012–1025.
- Maksimenko O., Georgiev P. // *Front. Genet.* 2014. V. 5. P. 28.
- Bouwman B.A., de Laat W. // *Genome Biol.* 2015. V. 16. P. 154.
- Dekker J., Mirny L. // *Cell*. 2016. V. 164. P. 1110–1121.
- Merkenschlager M., Nora E.P. // *Ann. Rev. Genom. Human Genet.* 2016. V. 17. P. 17–43.
- Ghirlando R., Felsenfeld G. // *Genes Dev.* 2016. V. 30. P. 881–891.
- Handoko L., Xu H., Li G., Ngan C.Y., Chew E., Schnapp M., Lee C.W., Ye C., Ping J.L., Mulawadi F., et al. // *Nat. Genet.* 2011. V. 43. P. 630–638.
- Gibcus J.H., Dekker J. // *Mol. Cell*. 2013. V. 49. P. 773–782.
- Zuin J., Dixon J.R., van der Reijden M.I., Ye Z., Kolovos P., Brouwer R.W., van de Corput M.P., van de Werken H.J., Knoch T.A., van IJcken W.F., et al. // *Proc. Natl. Acad. Sci. USA*. 2014. V. 111. P. 996–1001.
- Heger P., Marin B., Bartkuhn M., Schierenberg E., Wiehe T. // *Proc. Natl. Acad. Sci. USA*. 2012. V. 109. P. 17507–17512.
- Разин С.В., Борунова В.В., Максименко О.Г., Кантидзе О.Л. // *Биохимия*. 2012. Т. 77. С. 277–288.
- Iuchi S. // *Cell. Mol. Life Sci.* 2001. V. 58. P. 625–635.
- Pearson R.C., Funnell A.P., Crossley M. // *IUBMB Life*. 2011. V. 63. P. 86–93.
- Swamyathan S.K. // *Hum. Genomics*. 2010. V. 4. P. 263–270.
- Pavletich N.P., Pabo C.O. // *Science*. 1991. V. 252. P. 809–817.
- Wolfe S.A., Nekludova L., Pabo C.O. // *Annu. Rev. Biophys. Biomol. Struct.* 2000. V. 29. P. 183–212.
- Garton M., Najafabadi H.S., Schmitges F.W., Radovani E., Hughes T.R., Kim P.M. // *Nucl. Acids Res.* 2015. V. 43. P. 9147–9157.
- Persikov A.V., Singh M. // *Nucl. Acids Res.* 2014. V. 42. P. 97–108.
- Vandevenne M., Jacques D.A., Artuz C., Nguyen C.D., Kwan A.H., Segal D.J., Matthews J.M., Crossley M., Guss J.M., Mackay J.P. // *J. Biol. Chem.* 2013. V. 288. P. 10616–10627.
- Brayer K.J., Segal D.J. // *Cell Biochem. Biophys.* 2008. V. 50. P. 111–131.
- Brayer K.J., Kulshreshtha S., Segal D.J. // *Cell Biochem. Biophys.* 2008. V. 51. P. 9–19.
- Emerson R.O., Thomas J.H. // *J. Virol.* 2011. V. 85. P. 12043–12052.
- Chung H.R., Schafer U., Jackle H., Bohm S. // *EMBO Repts.* 2002. V. 3. P. 1158–1162.
- Perez-Torrado R., Yamada D., Defossez P.A. // *Bioessays*. 2006. V. 28. P. 1194–1202.
- Stogios P.J., Chen L., Prive G.G. // *Protein Sci.* 2007. V. 16. P. 336–342.
- Liu H., Chang L.H., Sun Y., Lu X., Stubbs L. // *Genome Biol. Evol.* 2014. V. 6. P. 510–525.
- Lobanenkova V.V., Nicolas R.H., Adler V.V., Paterson H., Klenova E.M., Polotskaja A.V., Goodwin G.H. // *Oncogene*. 1990. V. 5. P. 1743–1753.
- Bell A.C., West A.G., Felsenfeld G. // *Cell*. 1999. V. 98. P. 387–396.
- Chaumeil J., Skok J.A. // *Curr. Opin. Immunol.* 2012. V. 24. P. 153–159.
- Herold M., Bartkuhn M., Renkawitz R. // *Development*. 2012. V. 139. P. 1045–1057.
- Holwerda S., de Laat W. // *Front. Genet.* 2012. V. 3. P. 217.
- Merkenschlager M., Odom D.T. // *Cell*. 2013. V. 152. P. 1285–1297.
- Sexton T., Yaffe E., Kenigsberg E., Bantignies F., Leblanc B., Hoichman M., Parrinello H., Tanay A., Cavalli G. // *Cell*. 2012. V. 148. P. 458–472.

39. Pirrotta V., Li H.B. // *Curr. Opin. Genet. Dev.* 2012. V. 22. P. 101–109.
40. Kyrchanova O., Toshchakov S., Podstreshnaya Y., Parshikov A., Georgiev P. // *Mol. Cell. Biol.* 2008. V. 28. P. 4188–4195.
41. Kyrchanova O., Ivlieva T., Toshchakov S., Parshikov A., Maksimenko O., Georgiev P. // *Nucl. Acids Res.* 2011. V. 39. P. 3042–3052.
42. 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.
43. Pant V., Kurukuti S., Pugacheva E., Shamsuddin S., Mariano P., Renkawitz R., Klenova E., Lobanenkov V., Ohlsson R. // *Mol. Cell Biol.* 2004. V. 24. P. 3497–3504.
44. Xiao T., Wallace J., Felsenfeld G. // *Mol. Cell. Biol.* 2011. V. 31. P. 2174–2183.
45. Vietri Rudan M., Barrington C., Henderson S., Ernst C., Odom D. T., Tanay A., Hadjur S. // *Cell Rep.* 2015. V. 10. P. 1297–1309.
46. Schmidt D., Schwalie P.C., Wilson M.D., Ballester B., Goncalves A., Kutter C., Brown G.D., Marshall A., Flicek P., Odom D.T. // *Cell.* 2012. V. 148. P. 335–348.
47. Nakahashi H., Kwon K.R., Resch W., Vian L., Dose M., Stavreva D., Hakim O., Pruett N., Nelson S., Yamane A., et al. // *Cell Rep.* 2013. V. 3. P. 1678–1689.
48. Xiao T., Wongtrakongate P., Trainor C., Felsenfeld G. // *Cell Rep.* 2015. V. 12. P. 1704–1714.
49. Ohlsson R., Bartkuhn M., Renkawitz R. // *Chromosoma.* 2010. V. 119. P. 351–360.
50. Wendt K.S., Yoshida K., Itoh T., Bando M., Koch B., Schirghuber E., Tsutsumi S., Nagae G., Ishihara K., Mishiro T., et al. // *Nature.* 2008. V. 451. P. 796–801.
51. Proudhon C., Hao B., Raviram R., Chaumeil J., Skok J.A. // *Adv. Immunol.* 2015. V. 128. P. 123–182.
52. Bergstrom R., Savary K., Moren A., Guibert S., Heldin C. H., Ohlsson R., Moustakas A. // *J. Biol. Chem.* 2010. V. 285. P. 19727–19737.
53. 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. P. E677–686.
54. Liu Z., Scannell D.R., Eisen M.B., Tjian R. // *Cell.* 2011. V. 146. P. 720–731.
55. Yao H., Brick K., Evrard Y., Xiao T., Camerini-Otero R.D., Felsenfeld G. // *Genes Dev.* 2010. V. 24. P. 2543–2555.
56. Defosse P.A., Kelly K.F., Filion G.J., Perez-Torrado R., Magdinier F., Menoni H., Nordgaard C.L., Daniel J.M., Gilson E. // *J. Biol. Chem.* 2005. V. 280. P. 43017–43023.
57. Klenova E., Scott A.C., Roberts J., Shamsuddin S., Lovejoy E.A., Bergmann S., Bubb V.J., Royer H.D., Quinn J.P. // *J. Neurosci.* 2004. V. 24. P. 5966–5973.
58. Donohoe M.E., Zhang L.F., Xu N., Shi Y., Lee J.T. // *Mol. Cell.* 2007. V. 25. P. 43–56.
59. Donohoe M.E., Silva S.S., Pinter S.F., Xu N., Lee J.T. // *Nature.* 2009. V. 460. P. 128–132.
60. Ishihara K., Oshimura M., Nakao M. // *Mol. Cell.* 2006. V. 23. P. 733–742.
61. Li T., Hu J.F., Qiu X., Ling J., Chen H., Wang S., Hou A., Vu T.H., Hoffman A.R. // *Mol. Cell. Biol.* 2008. V. 28. P. 6473–6482.
62. Lutz M., Burke L.J., Barreto G., Goeman F., Greb H., Arnold R., Schultheiss H., Brehm A., Kouzarides T., Lobanenkov V., et al. // *Nucl. Acids Res.* 2000. V. 28. P. 1707–1713.
63. Saldana-Meyer R., Gonzalez-Buendia E., Guerrero G., Narendra V., Bonasio R., Recillas-Targa F., Reinberg D. // *Genes Dev.* 2014. V. 28. P. 723–734.
64. 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. P. 361–375.
65. Guastafierro T., Catizone A., Calabrese R., Zampieri M., Martella O., Bacalini M.G., Reale A., Di Girolamo M., Miccheli M., Farrar D., et al. // *Biochem. J.* 2013. V. 449. P. 623–630.
66. Klenova E.M., Chernukhin I.V., El-Kady A., Lee R.E., Pugacheva E.M., Loukinov D.I., Goodwin G.H., Delgado D., Filippova G.N., Leon J., et al. // *Mol. Cell. Biol.* 2001. V. 21. P. 2221–2234.
67. MacPherson M.J., Beatty L.G., Zhou W., Du M., Sadowski P.D. // *Mol. Cell. Biol.* 2009. V. 29. P. 714–725.
68. Bellefroid E.J., Poncelet D.A., Lecocq P.J., Revelant O., Martial J.A. // *Proc. Natl. Acad. Sci. USA.* 1991. V. 88. P. 3608–3612.
69. Huntley S., Baggott D.M., Hamilton A.T., Tran-Gyamfi M., Yang S., Kim J., Gordon L., Branscomb E., Stubbs L. // *Genome Res.* 2006. V. 16. P. 669–677.
70. Lupo A., Cesaro E., Montano G., Zurlo D., Izzo P., Costanzo P. // *Curr. Genomics.* 2013. V. 14. P. 268–278.
71. Shannon M., Kim J., Ashworth L., Branscomb E., Stubbs L. // *DNA Seq.* 1998. V. 8. P. 303–315.
72. Margolin J.F., Friedman J.R., Meyer W.K., Vissing H., Thiesen H. J., Rauscher F. J., 3rd // *Proc. Natl. Acad. Sci. USA.* 1994. V. 91. P. 4509–4513.
73. Urrutia R. // *Genome Biol.* 2003. V. 4. P. 231.
74. Vissing H., Meyer W.K., Aagaard L., Tommerup N., Thiesen H.J. // *FEBS Lett.* 1995. V. 369. P. 153–157.
75. Lechner M.S., Begg G.E., Speicher D.W., Rauscher F.J., 3rd // *Mol. Cell. Biol.* 2000. V. 20. P. 6449–6465.
76. Schultz D.C., Ayyanathan K., Negorev D., Maul G.G., Rauscher F.J., 3rd // *Genes Dev.* 2002. V. 16. P. 919–932.
77. Sripathy S.P., Stevens J., Schultz D.C. // *Mol. Cell. Biol.* 2006. V. 26. P. 8623–8638.
78. Groner A.C., Meylan S., Ciuffi A., Zangger N., Ambrosini G., Denervaud N., Bucher P., Trono D. // *PLoS Genet.* 2010. V. 6. P. e1000869.
79. Wolf G., Macfarlan T.S. // *Cell.* 2015. V. 163. P. 30–32.
80. Ying L., Lin J., Qiu F., Cao M., Chen H., Liu Z., Huang Y. // *FEBS J.* 2015. V. 282. P. 174–182.
81. Wiznerowicz M., Jakobsson J., Szulc J., Liao S., Quazzola A., Beermann F., Aebischer P., Trono D. // *J. Biol. Chem.* 2007. V. 282. P. 34535–34541.
82. Rowe H.M., Friedli M., Offner S., Verp S., Mesnard D., Marquis J., Aktas T., Trono D. // *Development.* 2013. V. 140. P. 519–529.
83. Okumura K., Sakaguchi G., Naito K., Tamura T., Igarashi H. // *Nucl. Acids Res.* 1997. V. 25. P. 5025–5032.
84. Conroy A.T., Sharma M., Holtz A.E., Wu C., Sun Z., Weigel R.J. // *J. Biol. Chem.* 2002. V. 277. P. 9326–9334.
85. Hayashi K., Yoshida K., Matsui Y. // *Nature.* 2005. V. 438. P. 374–378.
86. Birtle Z., Ponting C.P. // *Bioinformatics.* 2006. V. 22. P. 2841–2845.
87. Najafabadi H.S., Albu M., Hughes T.R. // *Bioinformatics.* 2015. V. 31. P. 2879–2881.
88. Losson R., Nielsen A.L. // *Biochim. Biophys. Acta.* 2010. V. 1799. P. 463–468.
89. Frieze S., Lan X., Jin V.X., Farnham P.J. // *J. Biol. Chem.* 2010. V. 285. P. 1393–1403.
90. Yang L., Wang H., Kornblau S.M., Graber D.A., Zhang N., Matthews J.A., Wang M., Weber D.M., Thomas S.K., Shah J.J., et al. // *Oncogene.* 2011. V. 30. P. 1329–1340.
91. Parvanov E.D., Petkov P.M., Paigen K. // *Science.* 2010. V. 327. P. 835.

92. Smagulova F, Gregoretto I.V., Brick K., Khil P., Camerini-Otero R.D., Petukhova G.V. // *Nature*. 2011. V. 472. P. 375–378.
93. Oliver P.L., Goodstadt L., Bayes J.J., Birtle Z., Roach K.C., Phadnis N., Beatson S.A., Lunter G., Malik H.S., Ponting C.P. // *PLoS Genet*. 2009. V. 5. P. e1000753.
94. Thomas J.H., Emerson R.O., Shendure J. // *PLoS One*. 2009. V. 4. P. e8505.
95. Baudat F., Buard J., Grey C., Fledel-Alon A., Ober C., Przeworski M., Coop G., de Massy B. // *Science*. 2010. V. 327. P. 836–840.
96. Berg I.L., Neumann R., Lam K.W., Sarbajna S., Odenthal-Hesse L., May C.A., Jeffreys A.J. // *Nat. Genet*. 2010. V. 42. P. 859–863.
97. Baker C.L., Walker M., Kajita S., Petkov P.M., Paigen K. // *Genome Res*. 2014. V. 24. P. 724–732.
98. Brick K., Smagulova F., Khil P., Camerini-Otero R.D., Petukhova G.V. // *Nature*. 2012. V. 485. P. 642–645.
99. Castro-Diaz N., Ecco G., Coluccio A., Kapopoulou A., Yazdanpanah B., Friedli M., Duc J., Jang S.M., Turelli P., Trono D. // *Genes Dev*. 2014. V. 28. P. 1397–1409.
100. Turelli P., Castro-Diaz N., Marzetta F., Kapopoulou A., Raclot C., Duc J., Tieng V., Quenneville S., Trono D. // *Genome Res*. 2014. V. 24. P. 1260–1270.
101. Richardson S.R., Doucet A.J., Kopera H.C., Moldovan J.B., Garcia-Perez J.L., Moran J.V. // *Microbiol Spectr*. 2015. V. 3. P. MDNA3-0061-2014.
102. Nowick K., Gernat T., Almaas E., Stubbs L. // *Proc. Natl. Acad. Sci. USA*. 2009. V. 106. P. 22358–22363.
103. Barde I., Rauwel B., Marin-Florez R.M., Corsinotti A., Laurenti E., Verp S., Offner S., Marquis J., Kapopoulou A., Vanicek J., et al. // *Science*. 2013. V. 340. P. 350–353.
104. Sander T.L., Haas A.L., Peterson M.J., Morris J.F. // *J. Biol. Chem*. 2000. V. 275. P. 12857–12867.
105. Ivanov D., Stone J.R., Maki J.L., Collins T., Wagner G. // *Mol. Cell*. 2005. V. 17. P. 137–143.
106. Sander T.L., Stringer K.F., Maki J.L., Szauter P., Stone J.R., Collins T. // *Gene*. 2003. V. 310. P. 29–38.
107. Liang Y., Huimei Hong F., Ganesan P., Jiang S., Jauch R., Stanton L.W., Kolatkar P.R. // *Nucl. Acids Res*. 2012. V. 40. P. 8721–8732.
108. Rimsa V., Eadsforth T.C., Hunter W.N. // *PLoS One*. 2013. V. 8. P. e69538.
109. 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. P. 137–147.
110. Liang Y., Choo S.H., Rossbach M., Baburajendran N., Palasingam P., Kolatkar P.R. // *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun*. 2012. V. 68. P. 443–447.
111. Williams A.J., Blacklow S.C., Collins T. // *Mol. Cell. Biol*. 1999. V. 19. P. 8526–8535.
112. 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. P. 17173–17179.
113. Porsch-Ozcurumez M., Langmann T., Heimerl S., Borsukova H., Kaminski W.E., Drobnik W., Honer C., Schumacher C., Schmitz G. // *J. Biol. Chem*. 2001. V. 276. P. 12427–12433.
114. Stone J.R., Maki J.L., Blacklow S.C., Collins T. // *J. Biol. Chem*. 2002. V. 277. P. 5448–5452.
115. Ogo O.A., Tyson J., Cockell S.J., Howard A., Valentine R.A., Ford D. // *Mol. Cell. Biol*. 2015. V. 35. P. 977–987.
116. Chung H.R., Lohr U., Jackle H. // *Mol. Biol. Evol*. 2007. V. 24. P. 1934–1943.
117. Jauch R., Bourenkov G.P., Chung H.R., Urlaub H., Reidt U., Jackle H., Wahl M.C. // *Structure*. 2003. V. 11. P. 1393–1402.
118. Zolotarev N., Fedotova A., Kyrchanova O., Bonchuk A., Penin A.A., Lando A.S., Eliseeva I.A., Kulakovskiy I.V., Maksimenko O., Georgiev P. // *Nucl. Acids Res*. 2016. V. 44. P. 7228–7241.
119. Zolotarev N.A., Maksimenko O.G., Georgiev P.G., Bonchuk A.N. // *Acta Naturae*. 2016. V. 8. № 3. P. 97–102.
120. Li J., Gilmour D.S. // *EMBO J*. 2013. V. 32. P. 1829–1841.
121. Gaszner M., Vazquez J., Schedl P. // *Genes Dev*. 1999. V. 13. P. 2098–2107.
122. Blanton J., Gaszner M., Schedl P. // *Genes Dev*. 2003. V. 17. P. 664–675.
123. Schwartz Y.B., Linder-Basso D., Kharchenko P.V., Tolstorukov M.Y., Kim M., Li H.B., Gorchakov A.A., Minoda A., Shanower G., Alekseyenko A.A., et al. // *Genome Res*. 2012. V. 22. P. 2188–2198.
124. Kyrchanova O., Leman D., Parshikov A., Fedotova A., Studitsky V., Maksimenko O., Georgiev P. // *PLoS One*. 2013. V. 8. P. e62690.
125. 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. P. 89–99.
126. Laundrie B., Peterson J.S., Baum J.S., Chang J.C., Fileppo D., Thompson S.R., McCall K. // *Genetics*. 2003. V. 165. P. 1881–1888.
127. Chen B., Harms E., Chu T., Henrion G., Strickland S. // *Development*. 2000. V. 127. P. 1243–1251.
128. Chu T., Henrion G., Haegeli V., Strickland S. // *Genesis*. 2001. V. 29. P. 141–152.
129. Payre F., Noselli S., Lefrere V., Vincent A. // *Development*. 1990. V. 110. P. 141–149.
130. Crozatier M., Kongsuwan K., Ferrer P., Merriam J.R., Lengyel J.A., Vincent A. // *Genetics*. 1992. V. 131. P. 905–916.
131. Lake C.M., Nielsen R.J., Hawley R.S. // *PLoS Genet*. 2011. V. 7. P. e1002005.
132. McKim K.S., Hayashi-Hagihara A. // *Genes Dev*. 1998. V. 12. P. 2932–2942.
133. Jang J.K., Sherizen D.E., Bhagat R., Manheim E.A., McKim K.S. // *J. Cell. Sci*. 2003. V. 116 (Pt 15). P. 3069–3077.
134. Liu H., Jang J.K., Kato N., McKim K.S. // *Genetics*. 2002. V. 162. P. 245–258.