

УДК 577.21

ПРОСТРАНСТВЕННАЯ ОРГАНИЗАЦИЯ ГЕНОВ ДОМАШНЕГО ХОЗЯЙСТВА В ИНТЕРФАЗНЫХ ЯДРАХ

© 2014 г. Е. С. Гушчанская¹, А. А. Артемов^{2,3}, С. В. Ульянов¹, А. А. Пенин²,
М. Д. Логачева², С. В. Разин^{1,4}, А. А. Гаврилов^{1*}

¹Институт биологии гена Российской академии наук, Москва 119334

²Факультет биоинформатики и биоинженерии Московского государственного университета им. М.В. Ломоносова, Москва 119992

³Институт проблем передачи информации им. А.А. Харкевича Российской академии наук, Москва 127994

⁴Биологический факультет Московского государственного университета им. М.В. Ломоносова, Москва 119992

Поступила в редакцию 10.07.2014 г.

Принята к печати 14.07.2014 г.

Пространственная организация эукариотического генома тесно связана с его функционированием. В частности, взаимодействие промоторов генов с удаленными энхансерными элементами в составе активаторных хроматиновых блоков, а также привлечение генов к общим транскрипционным фабрикам играет важную роль в активации транскрипции генов. Большинство работ, посвященных анализу пространственного взаимодействия регуляторных элементов генома, выполнено на моделях тканеспецифичных генов. Между тем остается не ясным, какой вклад в пространственную организацию хромосом вносят гены домашнего хозяйства, которые транскрибируются в большинстве клеточных типов и числом заметно превосходят число транскрибирующихся тканеспецифичных генов. С целью ответить на этот вопрос мы провели полногеномный анализ пространственных взаимодействий генов домашнего хозяйства курицы *CARHSP1* и *TRAP1*, расположенных на хромосоме 14, используя для этого технологию 4C. Установлено, что для промоторов этих генов повышена частота ассоциации с транскрипционно активными областями хромосом и с участками, обогащенными сайтами связывания общего транскрипционного фактора Sp1 и CpG-островками, которые маркируют промоторы генов домашнего хозяйства. В то же время, при анализе взаимодействий участка хромосомы 14 из обедненной генами области не обнаружено подобных предпочтений. Указания на ассоциацию промоторов генов домашнего хозяйства получены и в независимых цитологических экспериментах по визуализации неметилированных CpG-островков в индивидуальных ядрах клеток человека. В них выявлена кластеризация CpG-островков в ядерном пространстве. В результате проведенных исследований получены доказательства того, что взаимодействие промоторов генов домашнего хозяйства – важный фактор, определяющий пространственную организацию интерфазных хромосом.

Ключевые слова: гены домашнего хозяйства, пространственная организация генома, транскрипционная фабрика, 4C.

SPATIAL ORGANIZATION OF HOUSE-KEEPING GENES IN INTERPHASE NUCLEI, by E. S. Gushchanskaya¹, A. V. Artemov^{2,3}, S. V. Ulyanov¹, A. A. Penin², M. D. Logacheva², S. V. Razin^{1,4}, A. A. Gavrilov^{1*} (¹Institute of Gene Biology, Russian Academy of Sciences, 119334 Moscow, *e-mail: aleksey.gavrilov@mail.ru; ²Department of Bioengineering and Bioinformatics, Moscow State University, Moscow, 119991 Russia; ³Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, 127994 Russia; ⁴Department of Biology, Moscow State University, Moscow, 119991 Russia). Spatial organization of the eukaryotic genome is tightly connected to its functioning. In particular, the interaction of gene promoters with remote enhancer elements in active chromatin hubs, as well as the recruitment of genes to the common transcription factories plays an important role in regulation of gene transcription. Most of works related to the analysis of spatial interaction of genome regulatory elements relies on models of tissue-specific genes. Meanwhile, it remains unclear to which extent the spatial organization of chromosomes is guided by

Принятые сокращения: 3C (chromosome conformation capture) – фиксация конформации хромосомы; 4C (chromosome conformation capture on chip) – фиксация конформации хромосомы–чип; DAPI (4',6-diamidino-2-phenylindole) – 4',6-диамидино-2-фенилиндол; FISH (fluorescence in situ hybridization) – флуоресцентная гибридизация *in situ*.

* Эл. почта: aleksey.gavrilov@mail.ru

house-keeping genes that are transcribed in most of cell types and outnumber the transcribed tissue-specific genes. To address this question, we used the 4C technique to characterize genome-wide the spatial contacts of the chicken house-keeping genes *CARHSP1* и *TRAP1* situated on chromosome 14. The promoters of these genes had an increased frequency of interaction with chromosome regions enriched in CpG islands and binding motifs for the ubiquitous transcription factor Sp1, both of which mark promoters of house-keeping genes, and overall with transcriptionally active regions. By contrast, the analysis of interaction of a gene poor region of chromosome 14 revealed no such preferences. The evidence for the interaction of house-keeping gene promoters were also obtained in independent cytological experiments aimed at visualization of non-methylated CpG islands in individual nuclei of human cells, which showed clustering of CpG islands in the nuclear space. Altogether, the results of our work suggest that the interaction of house-keeping genes constitutes an important factor that determines the spatial organization of interphase chromosomes.

Keywords: house-keeping genes, genome spatial organization, transcription factory, 4C.

DOI: 10.7868/S0026898414060056

ВВЕДЕНИЕ

В последнее время появляется все больше наблюдений, свидетельствующих о тесной связи между пространственной организацией генома и его функционированием [1–3]. В многочисленных цитологических и биохимических исследованиях показано, что для регуляции транскрипции важно взаимное расположение генов и их регуляторных последовательностей внутри клеточного ядра. На основании результатов, полученных с использованием технологии 3C [4] при изучении альфа- и бета-глобиновых генов позвоночных, создана модель активаторного хроматинового блока. Согласно этой модели, пространственное сближение удаленных регуляторных элементов с регулирующими тканеспецифичными генами обеспечивает активацию экспрессии этих генов [5, 6]. Однако свидетельства о пространственной сближенности удаленных элементов генома не решают вопроса о механизмах, которые обеспечивают это сближение. Предполагается, что удаленные регуляторные элементы генома удерживаются в составе единого комплекса посредством взаимодействий между белками, связанными с этими регуляторными элементами. Но до сих пор не получено прямых доказательств этой гипотезы. Недостаточно изучены не только механизмы поддержания взаимодействий удаленных областей генома, но и, собственно, принципы функционально-зависимой пространственной организации протяженных областей генома.

В последнее время значительную роль в установлении и поддержании трехмерной организации интерфазного генома приписывают так называемым “транскрипционным фабрикам” — структурам, в которых сосредоточены молекулы РНК-полимеразы и факторы транскрипции и куда привлекаются различные гены, подлежащие транскрипции [7–9]. В известном смысле, транскрипционная фабрика может рассматриваться и как активаторный хроматиновый блок, так как в ее составе присутствуют промоторы и регуляторные элементы ряда генов [10–12]. Несмотря на огромный интерес к

проблеме, число экспериментальных работ, посвященных изучению транскрипционных фабрик, остается небольшим, а результаты этих работ нередко противоречат друг другу. Так как число транскрибирующихся генов существенно превышает число транскрипционных фабрик [13], логично предположить, что в составе одной транскрипционной фабрики находятся разные гены. Случайно ли распределение транскрибирующихся генов между индивидуальными транскрипционными фабриками? Некоторыми авторами показано, что функционально связанные гены преимущественно привлекаются в общие транскрипционные фабрики [14]. Но известны и такие транскрипционные фабрики, которые содержат как тканеспецифичные гены, так и гены домашнего хозяйства [15–17]. В связи с этим следует отметить, что даже в дифференцированных клетках число транскрибирующихся генов домашнего хозяйства заметно превышает число транскрибирующихся тканеспецифичных генов, поэтому есть основания полагать, что именно наличие в составе транскрипционных фабрик генов домашнего хозяйства может определять пространственную организацию интерфазных хромосом.

Для проверки этого предположения нами проанализирован полногеномный спектр удаленных взаимодействий генов домашнего хозяйства кур, *CARHSP1* и *TRAP1*, расположенных на 14-ой хромосоме. Используя технологию 4C (chromosome conformation capture on chip [18–20]) с последующим полногеномным секвенированием, мы картировали полный набор фрагментов ДНК, контактирующих с промоторами генов *CARHSP1* и *TRAP1* в культивируемых куриных эритроцитах (линия HD3). Полученные результаты подтвердили наше предположение о том, что в ядерном пространстве промоторы генов домашнего хозяйства ассоциированы между собой. Данные 4C-анализа косвенно подтверждены результатами, полученными в цитологических экспериментах по детекции двунитевых разрывов ДНК, вносимых сайт-специфической эндонуклеазой ре-

стрикции в неметирированные последовательности CpG-островков, которые, в большинстве случаев, маркируют промоторы генов домашнего хозяйства [21, 22]. В этих экспериментах выявлена кластеризация точек разрывов – а значит и кластеризация CpG-островков в ядерном пространстве. В целом, в результате проведенного исследования показано, что пространственное взаимодействие промоторов генов домашнего хозяйства играет важную роль в создании специфической пространственной архитектуры интерфазных хромосом.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Клеточные линии. Для проведения экспериментов по 4C-анализу использованы клетки линии HD3. Изначально эти клетки получали путем трансформации эритроидных клеток ранней стадии дифференцировки из селезенки кур вирусом птичьего эритробластома, содержащего онкоген *v-erbA* и мутантный температуро-чувствительный онкоген *v-erbB* [23].

В экспериментах по визуализации кластеров CpG-островков использованы клетки линии MCF-7 (клетки рака молочной железы человека), содержащие стабильно трансфицированную генетическую конструкцию pTet-on, которая необходима для функционирования индуцируемой системы Tet-on, использованной в экспериментах (см. раздел “Результаты исследования”).

4C-анализ. 4C-анализ выполняли, следуя протоколу, описанному в работе [18], с небольшими изменениями.

Первые стадии 4C совпадают с процедурой 3C-анализа [24, 25]. Образцы 3C-ДНК, приготовленные с использованием эндонуклеазы рестрикции HindIII, дополнительно очищали с использованием набора QIAEX II Gel Extraction Kit (“Qiagen”). Концентрацию ДНК определяли с помощью флуориметрического метода (Qubit, “Invitrogen”). 3C-матрицу (50 мкг, концентрация 100 нг/мкл) обрабатывали не чувствительной к метилированию у эукариот эндонуклеазой рестрикции DpnII (200 ед., “NEB”) в течение 16 ч при 37°C. Реакцию рестрикции останавливали нагреванием до 65°C в течение 20 мин, добавляли ацетат натрия (рН 5.2) до конечной концентрации 0.2 М, экстрагировали раствор равным объемом смеси фенол-хлороформ (1 : 1) и осаждали ДНК этанолом. Осадок ДНК растворяли в 10 мМ Трис (рН 8.0), добавляли 14 мл лигазного буфера (“Thermo Scientific”), 100 ед. ДНК-лигазы фага T4 (“Thermo Scientific”) и инкубировали 4.5 ч при 16°C и 30 мин при комнатной температуре. После этого экстрагировали ДНК равным объемом фенола, добавляли гликоген до 20 мкг/мл, ацетат натрия (рН 5.2) до конечной концентрации 0.2 М и два

объема холодного этанола, выдерживали сутки при –70°C и осаждали ДНК центрифугированием в течение 20 мин при 16 100 × g и 4°C. Очистку ДНК проводили с помощью набора QIAquick PCR purification kit (“Qiagen”). Для линейаризации кольцевых молекул ДНК обрабатывали рестриктазами (в количестве 200 ед., в соответствующих буферах для рестрикции): StuI для якорного фрагмента *CARHSP1*, NcoI для якорного фрагмента в генной пустыне и якорного фрагмента *TRAP1*. Инкубацию проводили в течение 16 ч при 37°C. Для очистки ДНК использовали набор QIAquick PCR purification kit (“Qiagen”), каждый образец растворяли в 100 мкл 10 мМ Трис (рН 8.0).

Аmplификацию проводили с помощью набора Expand Long Template PCR System (“Roche”), в состав которого входят термостабильная ДНК-полимераза *Taq* и термостабильная ДНК-полимераза *Tgo*, которая обладает 3’–5’ экзонуклеазной активностью. Для амплификации использованы следующие последовательности праймеров (5’–3’): *CARHSP1*-якорный фрагмент CTCTGAGAG-CAGCCCCAGTC и TCCTAAGGAGATTTAT-GTTTTGAGAA; якорный фрагмент в генной пустыне AATTTGTGAAGCAGTTGTATGTAGTC и TCTTCTCCACATAATCCCACAC; *TRAP1*-якорный фрагмент CCAGAGATTCTCAAATCAGCA и СТАТGGGGACAAGTGGAACAG (для всех пар первым указан праймер со стороны HindIII-сайта, вторым – праймер со стороны DpnII-сайта).

После амплификации полученные 4C-библиотеки смешивали в равных массовых соотношениях, и суммарный образец наносили на 1%-ный агарозный гель и разделяли электрофорезом при напряжении 5–7 В/см. Продукты амплификации вырезали из геля, разделяя их на две фракции: 70–400 п.н. (короткие фрагменты) и 400–1500 п.н. (длинные фрагменты). Выделяли ДНК из геля, используя набор QIAquick gel extraction kit (“Qiagen”). Фракцию, содержащую длинные фрагменты, обрабатывали ультразвуком – чтобы уменьшить размер фрагментов до 100–500 п.н. – на приборе Covaris S220 (параметры: 300 с, режим 10%, максимальная мощность 23 Вт). После этого фракции подготавливали к лигированию с адаптерами (аденирование тупых концов геномной ДНК) и проводили лигирование, используя набор TruSeq DNA sample prep kit v.2 (“Illumina”); полученные продукты разделяли в агарозном геле и отбирали фрагменты длиной 200–600 п.н. После очистки измеряли концентрацию ДНК (Qubit “Invitrogen”) и проводили ПЦР с праймерами, комплементарными дистальным регионам адаптеров Illumina (I-qPCR-1.1: 5’-AATGATACGGCGACCACCGAGAT-3’ и I-qPCR-2.1: 5’-CAAGCAGAAGACGGCATAC-GA-3’). Продукты амплификации двух фракций смешивали в равных массовых соотношениях, полученную смесь разводили до общей концентрации фрагментов ДНК 2 нМ, денатурировали, ис-

пользуя 0.1 М NaOH, и разводили буфером HT1 (“Illumina”) до конечной концентрации 10 пМ. Кластеры генерировали с помощью системы CBot и набора TruSeq PE Cluster Kit v. 3 (“Illumina”). Секвенирование производили на приборе HiSeq2000, используя набор TruSeq SBS Kit v. 3 (“Illumina”). Длина прочтений составляла 101 н. с каждого конца.

Отбор парноконцевых чтений, полученных с помощью прибора Illumina, производили путем сравнения с частью последовательности, входящей в состав якорного фрагмента. Из пары концевых прочтений выбирали только относящиеся к HindIII-сайту. Прочтения, содержащие последовательности, относящиеся к якорному фрагменту, считали целевыми, т.е. происходящими из продуктов лигирования любого фрагмента генома с якорным. Такие последовательности картировали на куриный геном сборки galGal4 (UCSC genome browser) с помощью программы bowtie [26]. Куриный геном (сборка galGal4) разбивали на HindIII-фрагменты *in silico*, и регионы между сайтами рестрикции использовали для дальнейшего анализа. Число прочтений, попадающих на левый или правый концы одного HindIII-фрагмента, суммировали, и конечное значение считали 4C-сигналом для данного рестриктового фрагмента. Нужно отметить, что при анализе распределения плотности прочтений не учитывается зависимость среднего покрытия генома от длины рестриктового фрагмента (очевидно, что при равном количестве прочтений, приходящихся на один рестриктоный фрагмент, их плотность будет тем меньше, чем больше размер этого фрагмента). С целью учесть этот фактор при анализе геном был разбит на интервалы размером по 100 т.п.н., в которых подсчитывали распределение 4C-сигнала.

Поиск CpG-островков для анализа корреляций с 4C-сигналом проводили, используя таблицу cpGIslandEx для куриного генома сборки galGal4 [27]. Для анализа корреляций плотности сайтов связывания фактора транскрипции Sp1 с 4C-сигналом предполагаемые участки связывания (“мотивы”) фактора Sp1 идентифицировали с использованием алгоритма PWM.

Трансфекция клеток эукариот плазмидной ДНК. Трансфекцию клеток линии MCF-7 плазмидным вектором проводили с помощью реактива для трансфекции Turbofect. Накануне трансфекции клетки рассаживали на покровные стекла, помещенные в 12-луночные планшеты таким образом, чтобы плотность клеток через сутки составила 60–70%. В день трансфекции клетки промывали фосфатно-солевым буфером (PBS) и добавляли 2 мл полной среды DMEM без тетрациклина (среда DMEM, 10%-ная фетальная бычья сыворотка). Отдельно готовили смесь ДНК с реактивом: для этого к бессывороточной среде DMEM

объемом 200 мкл добавляли смесь раствора плазмиды (до концентрации 10 мкг/мл) и реактива Turbofect в соотношении 1 : 2. Инкубировали 15–20 минут при комнатной температуре и добавляли в клеточную культуру. Затем инкубировали клетки 1 ч при 37°C в CO₂-инкубаторе, после чего отбирали среду, содержащую реагент для трансфекции и плазмидную ДНК, чтобы уменьшить токсический эффект, и добавляли 3 мл свежей среды. После этого в клеточную культуру добавляли доксициклин до концентрации 10 мкг/мл — для индукции системы Tet-on.

Иммуоцитохимия. Для иммуоокрашивания клетки фиксировали 4%-ным параформальдегидом 10 мин при равномерном покачивании при комнатной температуре, затем трижды промывали PBS и пермеабилizировали 1%-ным раствором Triton X-100 в течение 10 мин при комнатной температуре. После фиксации и пермеабилзации клетки помещали в блокирующий раствор (1% BSA, 0.05% Tween, PBS) на 30 мин. Инкубацию с первичными антителами, разведенными в блокирующем растворе, проводили в течение 1 ч при комнатной температуре во влажной камере. Затем клетки промывали трижды по 5 мин PBS с 0.25% BSA и 0.05% Tween-20. Первичные антитела проявляли антителами против соответствующих иммуноглобулинов, конъюгированными с различными флуорохромами в течение 30 мин при комнатной температуре в темноте; после этого клетки снова трижды промывали PBS с 0.25% BSA и 0.05% Tween-20 и окрашивали ДНК флуоресцентным красителем DAPI в течение 10 мин при комнатной температуре в темноте.

Препараты анализировали на сканирующем конфокальном микроскопе Leica, и полученные изображения обрабатывали с помощью программного обеспечения ImageJ.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Применение метода 4C для анализа пространственных взаимодействий в геноме кур

Исследование взаимодействий с каждым якорным фрагментом проводили в двух биологических повторах. 4C-библиотеки анализировали с помощью массивного параллельного секвенирования (~30 × 10⁶ парноконцевых чтений, т.е. 15 миллионов пар на один образец, длина прочтения составляла 101 н.), после чего отбирали прочтения со стороны HindIII-сайта, соответствующие одному из трех якорных фрагментов, и картировали на куриный геном (galGal4, UCSC). Учитывая тот факт, что любой из HindIII-концов искомого фрагмента может быть объединен с якорным в процессе лигирования, мы суммировали число прочтений, картированных независимо на оба конца HindIII-фрагментов, и рассматрива-

ли эту сумму как параметр, отвечающий вероятности лигирования данного HindIII-фрагмента с якорным.

Прочтения со стороны DpnII-сайта не включали в анализ по той причине, что некоторые фрагменты, детектированные со стороны DpnII-сайта, отражают неспецифические продукты вторичного лигирования. Дальнейший анализ проводили по описанной ранее методике [19].

В результате типичного 4С-эксперимента получают набор рестриктных фрагментов, который отражает частоту сшивки якорного фрагмента с геномным окружением. На рис. 1 показано распределение картированных фрагментов вдоль 14-й хромосомы для трех якорных фрагментов и двух биологических повторов. Распределение прочтений по хромосоме 14 свидетельствует о наличии нескольких взаимодействующих областей для каждого из якорных фрагментов (рис. 1а).

При визуальном анализе профили распределения прочтений для якорных фрагментов, содержащих промоторы генов домашнего хозяйства (*CARHSP1*, *TRAP1*), кажутся более схожими между собой, нежели с профилем для якорного фрагмента в генной пустыне (рис. 1а).

Промоторы генов домашнего хозяйства взаимодействуют между собой в ядре

Для якорных фрагментов “*CARHSP1*”, “*TRAP1*” и “генная пустыня” отсекуено ответственно около 1, 6 и 2.5 млн. химерных продуктов лигирования, отвечающих взаимодействию этих якорных фрагментов с другими фрагментами генома, расположенными на той же и других хромосомах. Сходство биологических реплик было очень высоким для экспериментов с якорными фрагментами “*CARHSP1*” и “генная пустыня” (коэффициенты корреляции Пирсона ~ 0.99; приведены в табл. 1 и 2), поэтому прочтения, полученные для двух биологических реплик, объединили и далее анализировали как единый массив данных. Принимая во внимание факт недостаточного сходства для двух биологических реплик в случае с якорным фрагментом “*TRAP1*” (см. таблицы 1 и 2), мы решили анализировать только результаты, полученные для первой биологической реплики.

Для всех якорных фрагментов большая часть партнеров по взаимодействию (от 50 до 80%) приходится на 14-ю хромосому, т.е. на хромосому, в которой расположены якорные фрагменты; при этом наибольшую частоту лигирования детектировали для фрагментов поблизости от якорного. Эти результаты хорошо согласуются с данными литературы [18, 20, 28], где показано, что внутри-

Таблица 1. Коэффициенты корреляции Пирсона для двух биологических реплик (HD3_1 и HD3_2), полученные при сравнении полногеномных профилей 4С-сигналов (выделены серым цветом)

Якорный фрагмент	CARHSP1 (HD3_1)	Генная пустыня (HD3_1)	TRAP1 (HD3_1)	CARHSP1 (HD3_2)	Генная пустыня (HD3_2)	TRAP1 (HD3_2)
CARHSP1 (HD3_1)	1.000	0.011	0.011	0.999	0.012	0.005
Генная пустыня (HD3_1)	0.011	1.000	0.004	0.011	0.999	0.003
TRAP1 (HD3_1)	0.011	0.004	1.000	0.011	0.004	0.643
CARHSP1 (HD3_2)	0.999	0.011	0.011	1.000	0.012	0.005
Генная пустыня (HD3_2)	0.012	0.999	0.004	0.012	1.000	0.003
TRAP1 (HD3_2)	0.005	0.003	0.643	0.005	0.003	1.000

Таблица 2. Коэффициенты корреляции Пирсона для двух биологических реплик (HD3_1 и HD3_2), полученные при сравнении профилей 4С-сигналов для хромосомы 14 (выделены серым цветом)

Якорный фрагмент	CARHSP1 (HD3_1)	Генная пустыня (HD3_1)	TRAP1 (HD3_1)	CARHSP1 (HD3_2)	Генная пустыня (HD3_2)	TRAP1 (HD3_2)
CARHSP1 (HD3_1)	1.000	-0.012	-0.019	0.999	-0.011	-0.021
Генная пустыня (HD3_1)	-0.012	1.000	-0.031	-0.012	0.999	-0.028
TRAP1 (HD3_1)	-0.019	-0.031	1.000	-0.019	-0.032	0.629
CARHSP1 (HD3_2)	0.999	-0.012	-0.019	1.000	-0.012	-0.021
Генная пустыня (HD3_2)	-0.011	0.999	-0.032	0.012	1.000	-0.028
TRAP1 (HD3_2)	-0.021	-0.028	0.629	-0.021	-0.028	1.000

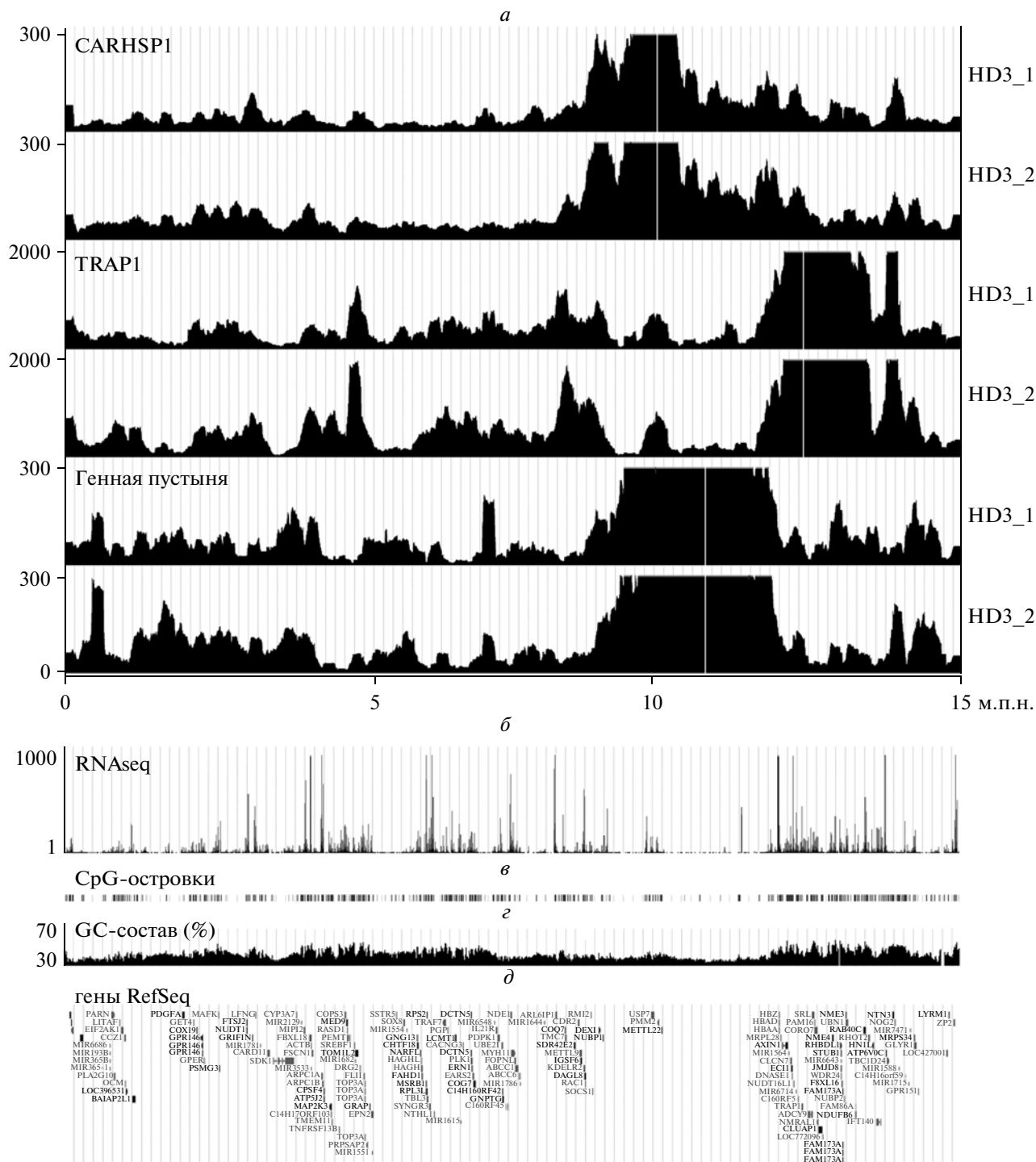


Рис. 1. Распределение 4C-сигнала, RNAseq-сигнала, CpG-островков и генов (RefSeq) вдоль хромосомы 14 кур. *а* – 4C-сигнал для якорных фрагментов “CARHSP1”, “TRAP1” и фрагмента в генной пустыне. По оси *x* отмечены координаты хромосомы 14 (млн. п.н.). По оси *y* отложены количества прочтений, картированные на концы HindIII-фрагментов куринового генома сборки galGal4 (UCSC), как описано в разделе “Экспериментальная часть”. Положения якорных фрагментов обозначены белыми вертикальными линиями. Эксперименты с каждым якорным фрагментом представлены для двух повторов (HD3_1 и HD3_2, указано справа на рисунке). *б* – Распределение сигнала RNAseq вдоль хромосомы 14. *в* – Распределение аннотированных CpG-островков вдоль хромосомы 14. *г* – CG-состав хромосомы 14. *д* – Распределение генов (RefSeq) вдоль хромосомы 14.

хромосомные взаимодействия происходят чаще межхромосомных и что в пределах одной хромосомы частота взаимодействия двух участков обратно пропорциональна расстоянию между ними в линейной последовательности ДНК.

С целью выяснить вопрос о возможной ассоциации промоторов генов домашнего хозяйства между собой мы проанализировали корреляции между частотами взаимодействия якорных фрагментов, несущих промоторы генов *CARHSP1* и *TRAP1* (гомологи одноименных генов домашнего хозяйства человека [29–31]), с удаленными областями генома и наличием в этих областях активных генов (по данным RNAseq), аннотированных CpG-островков и общего фактора транскрипции Sp1. Необходимо отметить, что мы использовали данные RNAseq, полученные для легочной ткани (депонированы в базе данных GEO, GSE57616); для других клеточных типов курицы транскриптом не проанализирован. Наиболее значительные корреляции выявлены между уровнем 4C-сигнала и плотностью CpG-островков. Такая картина была видна особенно отчетливо для всей совокупности хромосом (данные не показаны) и для всех хромосом при исключении 14-й (рис. 2а). Корреляции на 14-й хромосоме также заметны, но разброс значений в этом случае достаточно высок, вероятно, из-за малого числа сегментов, на которые разбивается при анализе относительно небольшая 14-я хромосома (данные не показаны). Это наблюдение подтвердилось при сравнении позиций максимальных значений (пиков) 4C-сигнала с максимальными значениями (пиками) плотности CpG-островков на 14-й хромосоме (рис. 3а). Принимая во внимание тот факт, что промоторы генов домашнего хозяйства (в частности, генов *CARHSP1* и *TRAP1*), как правило, располагаются в пределах CpG-островков, мы заключили, что эти промоторы, расположенные как на одной, так и на разных хромосомах, имеют тенденцию кластеризоваться в интерфазном ядре.

Эти наблюдения подтверждаются также положительной корреляцией между интенсивностью 4C-сигнала и плотностью сайтов связывания фактора транскрипции Sp1, участки узнавания которого маркируют промоторы генов домашнего хозяйства (рис. 2б и 3б), а также значительной положительной корреляцией с транскрипционно активными участками генома (рис. 2в).

Важно отметить, что подобных корреляций не наблюдали в экспериментах 4C, где использовался якорный фрагмент, лежащий в области, обедненной генами (рис. 2 – якорный фрагмент “генная пустыня”).

Неметилированные CpG-островки кластеризованы в интерфазном ядре

Нами проведена визуализация распределения неметилированных CpG-островков в интерфазном ядре, что позволило оценить характер распределения промоторов генов домашнего хозяйства в индивидуальных ядрах. Для этого генетическую конструкцию, содержащую кодирующую последовательность гена рестриктазы HpaII, трансфицировали в клетки человека линии MCF-7. Рестриктаза HpaII разрезает ДНК по последовательностям CCGG в том случае, когда цитозин неметилирован. В границах CpG-островков число мишеней для этой нуклеазы существенно превосходит таковое в среднем по геному. В этой связи в условиях недостатка экспрессированной в ядре нуклеазы HpaII разрывы происходят преимущественно в CpG-островках. Эти разрывы стимулируют образование фокусов фосфорилированного гистона H2AX (γ H2AX) [32], которые визуализировали с использованием иммунофлуоресцентных методов.

Кодирующую последовательность гена рестриктазы HpaII (любезно предоставлена компанией “СибЭнзим”, Россия), клонировали в плазмидный вектор pTRE-tight (часть доксициклин-индуцируемой системы экспрессии генов в клетках млекопитающих Tet-on). Генетическую конструкцию трансфицировали в клетки линии MCF-7 с помощью реагента Turbofect. Клетки выращивали с использованием полной среды DMEM без тетрациклина. После трансфекции в клеточную культуру добавляли доксициклин (полусинтетический аналог тетрациклина) до концентрации 10 мкг/мл для индукции системы Tet-on. Двунитевые разрывы визуализировали с помощью иммуноокрашивания антителами против модифицированной формы гистона γ H2AX [32].

Оптимальное время – с момента индукции доксициклином до появления фокусов γ H2AX – подбирали экспериментально, оно составило 15–16 ч. При анализе обработанных клеток среднее число фокусов на ядро составляло ~100 (рис. 4а). В контрольных экспериментах, проведенных на клетках, которые не трансфицировали плазмидной конструкцией, но обрабатывали реактивом для трансфекции и доксициклином, не обнаружено фокусов γ H2AX (рис. 4б).

Известно, что в геноме человека присутствует несколько десятков тысяч CpG-островков [33]. Следовательно, обнаруженные нами кластеры могут включать в себя от нескольких десятков до нескольких сотен CpG-островков. Если же учесть, что неметилированные CpG-островки часто маркируют промоторы генов домашнего хозяйства [21, 22], эти данные не только подтверждают вывод об ассоциации промоторов генов домашнего хо-

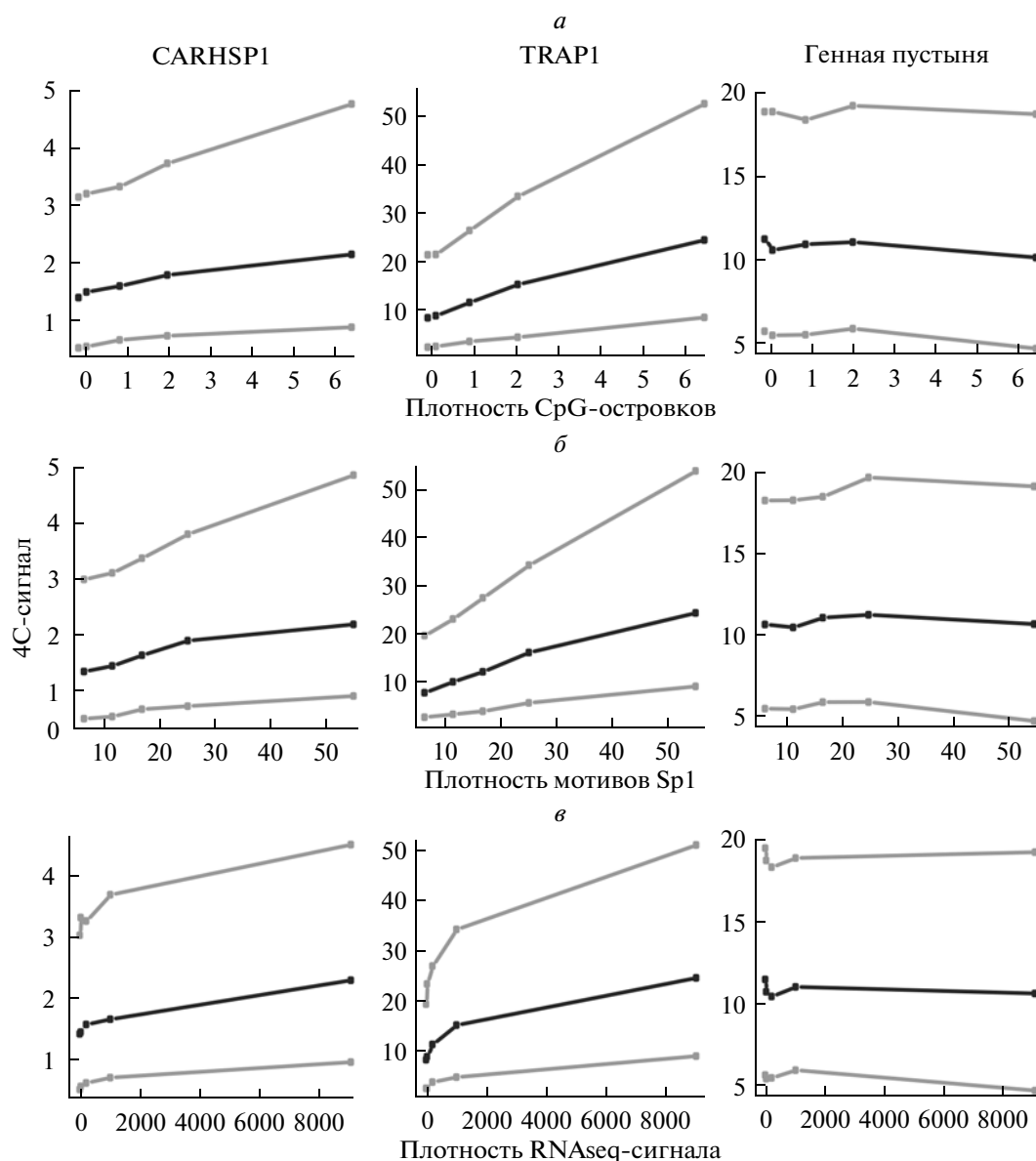


Рис. 2. Корреляция интенсивности 4C-сигнала и плотности CpG-островков, мотивов транскрипционного фактора Sp1 и RNAseq-сигнала для якорных фрагментов “CARHSP1”, “TRAP1” и фрагмента в генной пустыне. На графиках приведена усредненная зависимость интенсивности 4C-сигнала (суммарное количество прочтений по краям рестриктных фрагментов) от плотности CpG-островков (а), мотивов фактора Sp1 (б) и RNAseq-сигнала (в) в геномных интервалах длиной 100 т.п.н. Для данного анализа весь геном (за исключением хромосомы 14) был разделен на неперекрывающиеся интервалы означенной длины; в каждом интервале определяли суммарный 4C-сигнал и плотность интересующих геномных элементов (CpG-островков или Sp1-мотивов) или RNAseq-сигнала; далее интервалы были разделены на 5 равных групп по плотности элементов/RNAseq-сигнала. Для каждой группы изображен медианный сигнал 4C (средняя кривая), а также нижний (25%) и верхний (75%) квантили. Координата точки *x* соответствует средней плотности элементов генома/RNAseq-сигнала в данной группе.

зайства в интерфазных ядрах, но и впервые позволяю дать количественную оценку этим взаимодействиям.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В последние годы активно обсуждается роль транскрипционных фабрик в пространственной

организации генома. И чаще всего говорят о так называемых “специализированных” транскрипционных фабриках [34], в составе которых транскрибируются тканеспецифичные гены. Между тем, как уже говорилось выше, число транскрибирующихся генов домашнего хозяйства существенно превышает число специализированных генов, транскрибирующихся в любой клетке [33].

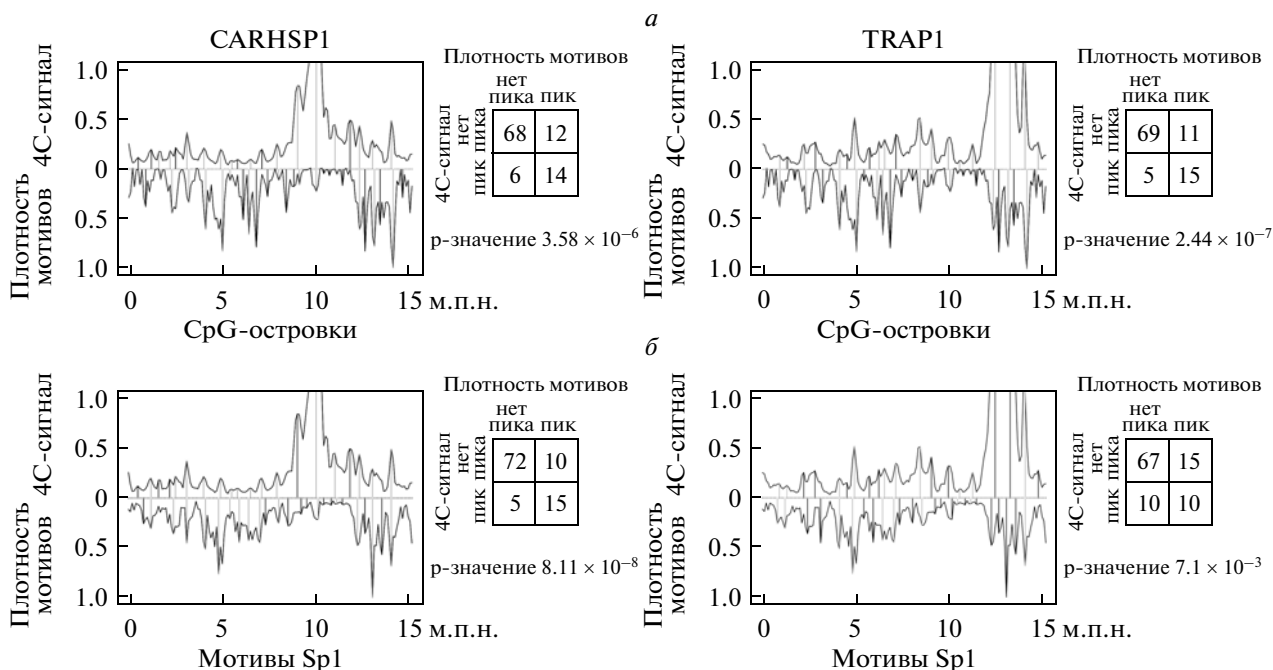


Рис. 3. Колокализация 4С-пиков и участков повышенной плотности CpG-островков и мотивов транскрипционного фактора Sp1 для якорных фрагментов “CARHSP1” и “TRAP1”. На графиках над осью x приведены кривые, отражающие интенсивность 4С-сигнала, а под осью x, в перевернутом виде, — кривые, отражающие плотность CpG-островков (а) и мотивов фактора Sp1 (б) в интервалах длиной 100 т.п.н. Пиком считали интервал, значение в котором превышало значение в соседних двух интервалах (2 предшествующих и 2 последующих). Пики выделены черными линиями; ко-локализация пиков 4С-сигнала и пиков плотности геномных элементов выделена серыми линиями (критерием ко-локализации считали попадание пиков в одни и те же или соседние интервалы). Для каждой пары значений сигнала и плотности справа от графика приведена таблица сопряженности, отражающая соотношение двух факторов: отсутствие или наличие пика 4С-сигнала; отсутствие или наличие пика плотности геномного элемента. Сопряженность определялась по тесту Фишера.

Наши наблюдения дают основания полагать, что ассоциация CpG-островков играет важную роль в пространственной сегрегации активных регионов генома. Действительно, обнаружены взаимодействия якорных фрагментов, содержащих CpG-островки, с регионами высокой плотности CpG-островков на той же и других хромосомах. С другой стороны, с якорным фрагментом, расположенным в области геной пустыни, преимущественно взаимодействовали регионы, не содержащие CpG-островков.

В этой связи уместно напомнить о том, что CpG-островки часто ассоциированы с промоторами генов домашнего хозяйства, которые, в свою очередь, составляют основную часть активно транскрибирующихся генов для каждого типа клеток [22]. Это утверждение согласуется с данными о том, что экспрессирующиеся тканеспецифичные гены могут привлекаться к ранее сформированным транскрипционным фабрикам, содержащим гены домашнего хозяйства [17, 35].

В настоящий момент нет оснований предполагать, что существует некий специфический принцип ассоциации определенных промоторов генов домашнего хозяйства в ядерном пространстве.

Вполне вероятно, определяющим фактором для объединения промоторов в кластеры служит их исходная сближенность в пространстве ядра, задаваемая хромосомальной позицией. При 4С-анализе выявлено наличие довольно обширных взаимодействующих регионов, содержащих промоторы генов домашнего хозяйства/CpG-островки. Результаты иммуноцитохимического исследования подтвердили кластеризацию большого числа CpG-островков.

Однако довольно сложно представить одновременное взаимодействие нескольких десятков CpG-островков в трехмерном пространстве. Можно предположить, что области хроматина, содержащие промоторы активно транскрибируемых генов, представляют собой структурно-функциональные компартменты, в стабилизации которых задействованы различные механизмы [36, 37]. Прежде всего, роль стабилизаторов могут играть транскрипционные факторы. Так, ранее показано, что транскрипционный фактор Sp1 опосредует взаимодействия между удаленными элементами генома [38]. Этот вывод подтвержден в нашем исследовании — выявлены значительные положительные корреляции между интенсивностью 4С-сигнала и

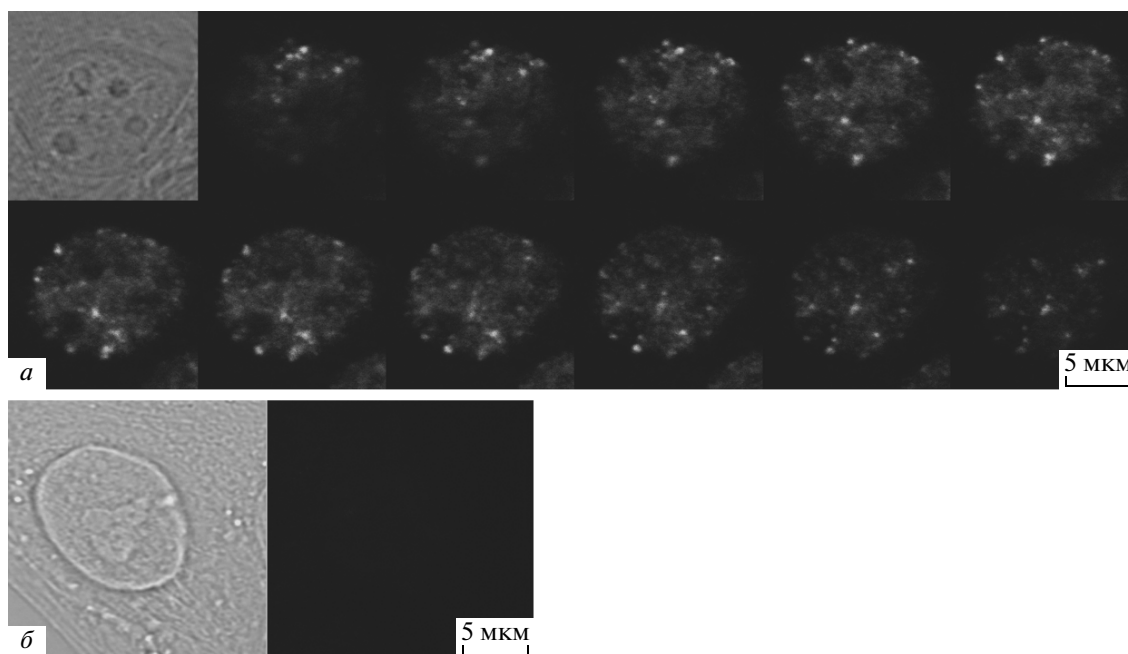


Рис. 4. Кластеризация неметилированных CpG-островков в интерфазном ядре. Клетки трансфицировали генетической конструкцией pTRE-tight, содержащей ген эндонуклеазы рестрикции HpaII, и индуцировали его экспрессию доксициклином. После фиксации и пермеабиллизации клеток наличие гистона γ H2AX выявляли иммуноцитохимическим методом. Фотографии получены на сканирующем конфокальном микроскопе Leica. *а* – Очертания ядра (фазовый контраст) и серия конфокальных срезов (шаг 0.3 мкм), на которых видно появление фокусов во всем объеме клеточного ядра. *б* – Эксперименты с контрольными клетками, трансфицированными реагентом Turbofect без генетической конструкции (приведен один из типичных конфокальных срезов) – показано отсутствие фокусов гистона γ H2AX.

плотностью сайтов связывания фактора Sp1. Поочередное привлечение промоторов в состав активного хроматинового компартмента может быть опосредовано ассоциацией промоторов с пре-инициаторными транскрипционными комплексами, содержащими активную (фосфорилированную) форму РНК-полимеразы II [39].

Итак, на основании полученных результатов можно сделать вывод о том, что кластеризация CpG-островков и промоторов генов домашнего хозяйства – одна из важных детерминант, определяющих характер макроупаковки хроматиновой фибриллы в интерфазном ядре.

Работа выполнена при финансовой поддержке Российского научного фонда (проект № 14-14-01088).

СПИСОК ЛИТЕРАТУРЫ

- Holwerda S., de Laat W. 2012. Chromatin loops, gene positioning, and gene expression. *Front. Genet.* **3**, 217.
- de Wit E., de Laat W. 2012. A decade of 3C technologies: insights into nuclear organization. *Genes Dev.* **26**, 11–24.
- Palstra R.J. 2009. Close encounters of the 3C kind: long-range chromatin interactions and transcriptional regulation. *Brief. Funct. Genomic Proteomic.* **8**, 297–309.
- Gavrilov A.A., Razin S.V., Iarovaia O.V. 2012. C-methods to study 3D organization of the eukaryotic genome. *Biopolymers Cell.* **28**, 245–251.
- Palstra R.J., Tolhuis B., Splinter E., Nijmeijer R., Grosveld F., de Laat W. 2003. The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* **35**, 190–194.
- de Laat W., Grosveld F. 2003. Spatial organization of gene expression: the active chromatin hub. *Chromosome Res.* **11**, 447–459.
- Sutherland H., Bickmore W.A. 2009. Transcription factories: gene expression in unions? *Nat. Rev. Genet.* **10**, 457–466.
- Razin S.V., Gavrilov A.A., Pichugin A., Lipinski M., Iarovaia O.V., Vassetzky Y.S. 2011. Transcription factories in the context of the nuclear and genome organization. *Nucl. Acids Res.* **39**, 9085–9092.
- Papantonis A., Cook P.R. 2013. Transcription factories: genome organization and gene regulation. *Chem. Rev.* **113**, 8683–705.
- Osborne C.S., Chakalova L., Brown K.E., Carter D., Horton A., Debrand E., Goyenechea B., Mitchell J.A., Lopes S., Reik W., Fraser P. 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* **36**, 1065–1071.
- Faro-Trindade I., Cook P.R. 2006. Transcription factories: structures conserved during differentiation and evolution. *Biochem. Soc. Trans.* **34**, 1133–1137.

12. Carter D.R., Eskiw C., Cook P.R. 2008. Transcription factories. *Biochem. Soc. Trans.* **36**, 585–589.
13. Martin S., Failla A.V., Spori U., Cremer C., Pombo A. 2004. Measuring the size of biological nanostructures with spatially modulated illumination microscopy. *Mol. Biol. Cell.* **15**, 2449–2455.
14. Schoenfelder S., Sexton T., Chakalova L., Cope N.F., Horton A., Andrews S., Kurukuti S., Mitchell J.A., Umlauf D., Dimitrova D.S., Eskiw C.H., Luo Y., Wei C.L., Ruan Y., Bieker J.J., Fraser P. 2010. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet.* **42**, 53–61.
15. Osborne C.S., Chakalova L., Mitchell J.A., Horton A., Wood A.L., Bolland D.J., Corcoran A.E., Fraser P. 2007. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol.* **5**, e192.
16. Philonenko E.S., Klochkov D.B., Borunova V.V., Gavrilov A.A., Razin S.V., Iarovaia O.V. 2009. TMEM8 – a non-globin gene entrapped in the globin web. *Nucl. Acids Res.* **37**, 7394–7406.
17. Zhou G.L., Xin L., Song W., Di L.J., Liu G., Wu X.S., Liu D.P., Liang C.C. 2006. Active chromatin hub of the mouse alpha-globin locus forms in a transcription factory of clustered housekeeping genes. *Mol. Cell. Biol.* **26**, 5096–5105.
18. Simonis M., Klous P., Splinter E., Moshkin Y., Willemssen R., de Wit E., van Steensel B., de Laat W. 2006. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* **38**, 1348–1354.
19. Tolhuis B., Blom M., Kerkhoven R.M., Pagie L., Teunissen H., Nieuwland M., Simonis M., de Laat W., van Lohuizen M., van Steensel B. 2011. Interactions among Polycomb domains are guided by chromosome architecture. *PLoS Genet.* **7**, e1001343.
20. Splinter E., de Wit E., van de Werken H.J., Klous P., de Laat W. 2012. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. *Methods.* **58**, 221–230.
21. Kundu T.K., Rao M.R. 1999. CpG islands in chromatin organization and gene expression. *J. Biochem.* **125**, 217–222.
22. Deaton A.M., Bird A. 2011. CpG islands and the regulation of transcription. *Genes Dev.* **25**, 1010–1022.
23. Beug H., Palmieri S., Freudenstein C., Zentgraf H., Graf T. 1982. Hormone-dependent terminal differentiation in vitro of chicken erythroleukemia cells transformed by ts mutants of avian erythroblastosis virus. *Cell.* **28**, 907–919.
24. Tolhuis B., Palstra R.J., Splinter E., Grosveld F., de Laat W. 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell.* **10**, 1453–1465.
25. Gavrilov A., Eivazova E., Priozhkova I., Lipinski M., Razin S., Vassetzky Y. 2009. Chromosome conformation capture (from 3C to 5C) and its ChIP-based modification. *Methods Mol. Biol.* **567**, 171–188.
26. Langmead B., Schatz M.C., Lin J., Pop M., Salzberg S.L. 2009. Searching for SNPs with cloud computing. *Genome Biol.* **10**, R134.
27. Gardiner-Garden M., Frommer M. 1987. CpG islands in vertebrate genomes. *J. Mol. Biol.* **196**, 261–282.
28. Splinter E., de Wit E., Nora E.P., Klous P., van de Werken H.J., Zhu Y., Kaaij L.J., van Ijcken W., Gribnau J., Heard E., de Laat W. 2011. The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes Dev.* **25**, 1371–1383.
29. Groblewski G.E., Yoshida M., Bragado M.J., Ernst S.A., Leykam J., Williams J.A. 1998. Purification and characterization of a novel physiological substrate for calcineurin in mammalian cells. *J. Biol. Chem.* **273**, 22738–22744.
30. Felts S.J., Owen B.A., Nguyen P., Trepel J., Donner D.B., Toft D.O. 2000. The hsp90-related protein TRAP1 is a mitochondrial protein with distinct functional properties. *J. Biol. Chem.* **275**, 3305–3312.
31. Song H.Y., Dunbar J.D., Zhang Y.X., Guo D., Donner D.B. 1995. Identification of a protein with homology to hsp90 that binds the type 1 tumor necrosis factor receptor. *J. Biol. Chem.* **270**, 3574–3581.
32. Paull T.T., Rogakou E.P., Yamazaki V., Kirchgessner C.U., Gellert M., Bonner W.M. 2000. A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr. Biol.* **10**, 886–895.
33. Antequera F., Bird A. 1993. Number of CpG islands and genes in human and mouse. *Proc. Natl. Acad. Sci. USA.* **90**, 11995–11999.
34. Xu M., Cook P.R. 2008. Similar active genes cluster in specialized transcription factories. *J. Cell Biol.* **181**, 615–623.
35. Gavrilov A.A., Zukher I.S., Philonenko E.S., Razin S.V., Iarovaia O.V. 2010. Mapping of the nuclear matrix-bound chromatin hubs by a new M3C experimental procedure. *Nucl. Acids Res.* **38**, 8051–8060.
36. Gavrilov A.A., Gushchanskaya E.S., Strelkova O., Zhironkina O., Kireev I.I., Iarovaia O.V., Razin S.V. 2013. Disclosure of a structural milieu for the proximity ligation reveals the elusive nature of an active chromatin hub. *Nucl. Acids Res.* **41**, 3563–3575.
37. Razin S.V., Gavrilov A.A., Ioudinkova E.S., Iarovaia O.V. 2013. Communication of genome regulatory elements in a folded chromosome. *FEBS Lett.* **587**, 1840–1847.
38. Nolis I.K., McKay D.J., Mantouvalou E., Lomvardas S., Merika M., Thanos D. 2009. Transcription factors mediate long-range enhancer-promoter interactions. *Proc. Natl. Acad. Sci. USA.* **106**, 20222–20227.
39. Marenduzzo D., Finan K., Cook P.R. 2006. The depletion attraction: an underappreciated force driving cellular organization. *J. Cell Biol.* **175**, 681–686.