

РАЗРАБОТКА МЕТОДА ДЕТЕКЦИИ ПРОСТРАНСТВЕННЫХ КОНТАКТОВ ПЛАЗМИДНОЙ ДНК С ГЕНОМ В КЛЕТКАХ ЧЕЛОВЕКА

© 2024 А.П Ян^{1,2*}, П.А Сальников^{1,2}, М.М Гридина^{1,2},
П.С. Белокопытова^{1,2}, В.С. Фишман^{1,2*}

¹ Институт цитологии и генетики СО РАН,
630090 Новосибирск, Россия; электронная почта: a.yan@g.nsu.ru; minja-f@yandex.ru

² Новосибирский государственный университет, 630090 Новосибирск, Россия

Поступила в редакцию 11.10.2023

После доработки 01.02.2024

Принята к публикации 02.03.2024

Развитие методов захвата конформации хромосом кардинально изменило наше представление об архитектуре и динамике хроматина. Применение этих технологий для исследования разных организмов позволило раскрыть основные принципы организации хромосом. Однако структурная организация внехромосомных элементов, таких как вирусные геномы или плазмиды, и их взаимодействия с геномом хозяина остаются недостаточно исследованными. В данной работе мы представляем усовершенствованный протокол 4C, предназначенный для изучения взаимодействий ДНК плазмиды с ядерной ДНК. Мы разработали специфический плазмидный вектор и оптимизировали протокол для обеспечения высокой эффективности детекции контактов между плазмидой и ДНК хозяина.

КЛЮЧЕВЫЕ СЛОВА: организация хроматина, 4C, плазида, экстрахромосомные последовательности, взаимодействия плазида–геном, захват конформации хромосом, трансфекция.

DOI: 10.31857/S0320972524040051 EDN: ZFUCFV

ВВЕДЕНИЕ

Недавние достижения в области биохимии и молекулярной биологии, а также появление и развитие секвенирования нового поколения способствовали разработке высокоэффективных методов исследования архитектуры хроматина интерфазного ядра. Установлено, что организация геномов позвоночных реализуется посредством по меньшей мере двух различных механизмов. Первый из них включает в себя взаимодействие между белковыми факторами CTCF и когезином в *cis*-положении [1], в то время как второй связан с организацией геномных компартментов через фазовую сепарацию, опосредованную взаимодействиями белков хроматина в *cis*- и *trans*-положениях [2]. Существуют некоторые исключения [3, 4], однако в большинстве клеток позвоночных эти механизмы действуют совместно, что затрудняет разделение независимых факторов, присущих каждому механизму в отдельности. Изучение формирова-

ния компартментов через фазовую сепарацию представляет особую проблему, поскольку общим термином «компартмент» могут быть названы различные ядерные структуры, каждая из которых характеризуется уникальным составом хроматина. Таким образом, в настоящий момент актуальна проблема полного выявления генетических детерминант, участвующих белков и дополнительных аспектов компартментализации хроматина [5, 6].

Еще одним важным направлением современной 3D-геномики является исследование разнообразия форм архитектуры хроматина в филогенетическом аспекте. Хотя эволюция архитектуры генома у позвоночных [7] и отдельных групп беспозвоночных [8] была относительно хорошо изучена, понимание этого процесса для других видов остается ограниченным. Кроме того, пока недостаточно информации о ядерной организации внеядерной и экзогенной ДНК, включая митохондриальные, пластидные и вирусные геномы, плазмиды, мобильные элементы и так далее. Тем не менее существующие данные указывают на то, что пространственная упаковка этих молекул ДНК

* Адресат для корреспонденции.

играет ключевую роль в их функционировании, и их изучение позволит получить важные знания об общих принципах биологии хроматина.

Вирусные геномы служат наиболее частым примером экзогенной последовательности ДНК. Было установлено, что ДНК вируса гепатита В (HBV), характеризующаяся как эписомальная кольцевая ДНК, преимущественно контактирует с А-компартомента генома. Эти контакты охватывают промоторы, энхансеры и CpG-островки высокоэкспрессирующихся генов, включая те, которые участвуют в клеточном ответе на инфекцию [9]. При этом транскрипционно подавленный HBV ассоциируется с регионами В-компартамента генома хозяина [10]. Кроме того, вирусная ДНК может быть интегрирована в геном хозяина, где она устанавливает новые контакты с соседними участками генома в *cis*-положении. Это может приводить к активации ранее молчавших генов [11] или изменениям нативной 3D-организации генома: либо за счет интеграции в существующие сайты связывания архитектурных белков, либо путем внесения новых сайтов в составе вирусной последовательности. Впоследствии эти изменения могут способствовать нарушению регуляции активности генов и развитию рака [9, 12].

Другой пример пространственных контактов экзогенной ДНК связан с ядерными тельцами промиелоцитарной лейкемии (PML). Эти тельца способны репликации и транскрипции вирусов, что было продемонстрировано для вируса простого герпеса, обезьяньего вируса 40 (SV40), аденовирусов и вируса папилломы человека [13, 14]. Работа Rai et al. [15] посвящена изучению шаперона HIRA, включающего гистон H3.3 в состав нуклеосом. Их исследование показало, что комплекс HIRA связывает в PML-тельцах чужеродную ДНК (в том числе очищенную плазмидную и вирусную) и таким образом участвует в подавлении ее транскрипции. В то же время другие работы указывают на формирование вирусных репликационных компартов рядом с PML-тельцами, где происходят процессы репликации и транскрипции [16, 17]. В целом, существующие данные подчеркивают функциональную важность пространственных контактов между экзогенной ДНК и геномом. Тем не менее биология экстрахромосомной ДНК внутри ядра, особенно в отношении плазмид, требует дальнейших исследований.

Сведения о биологии экстрахромосомной ДНК, в частности, о ее взаимодействии с белками и формировании вирус-ассоциированных компартов, в основном получают методами микроскопии. Однако этот подход имеет ряд серьезных ограничений. В первую очередь разрешение, которое он предлагает, часто недостаточно для небольших структур, таких как некоторые вирусные геномы

и плазмиды, размер которых составляет не более нескольких сотен тысяч пар нуклеотидов (п.н.). Альтернативным подходом к изучению укладки хроматина являются технологии захвата конформации хромосом (3C) [18], разрешение которых может достичь, в случае использования метода Hi-C, сотен или тысяч п.н. и уровня сотен п.н. для кольцевого захвата конформации хроматина (4C). Исследования, выполненные с применением этих методов, предоставляют более детальную картину пространственной организации ДНК, чем в случае микроскопического анализа.

Данная работа посвящена разработке нового метода исследования закономерностей локализации плазмидной ДНК и, в дальнейшей перспективе, других экзогенных молекул ДНК. Этот подход основан на 4C-эксперименте с использованием плазмиды как целевого вектора. Заметное преимущество применения метода 4C с плазмидой заключается в возможности, как мы предполагаем, проводить дальнейшие скрининговые эксперименты – доставлять плазмиды с произвольными вставками последовательностей ДНК в клетку (активные и неактивные регионы генома, GC-богатые и GC-бедные регионы, сайты связывания белков группы Polysomb и т.д.). Это даст возможность наблюдать влияние этих специфических нуклеотидных последовательностей на распределение плазмид по компартаментам внутри ядра.

МАТЕРИАЛЫ И МЕТОДЫ

Плазмидные векторы. Модифицированный вектор pUC19 был сконструирован на основе оригинального вектора pUC19 («Addgene», США) путем изменения двух сайтов рестрикции *NlaIII* путем ПЦР плазмиды с праймерами, содержащими целевую замену, и последующего лигирования продуктов ПЦР в кольцевую молекулу по методу Гибсона. Вектор p4CSCS был синтезирован компанией «Cloning Facility», Россия в виде двух последовательностей (991 п.н. и 975 п.н.), фланкированных гомологичными участками. Эти последовательности были соединены в кольцевую молекулу с помощью метода Гибсона. Карты и последовательности плазмид приведены в Приложении (рис. П1–П3; текст 1).

Плазмиды нарабатывали в клетках TOP10 *Escherichia coli*. Мы заметили, что для размножения вектора p4CSCS потребовалось снизить концентрацию хлорамфеникола до 0,02 мкг на 1 мл среды.

Культивирование клеток человека и трансфекция. Для экспериментов использовали клетки HEK293T, которые культивировали в среде DMEM с добавлением 10% фетальной бычьей сыворотки и 1% Pen Strep (все материалы – «Thermo Fisher

Scientific», США). Клетки поддерживали при температуре 37 °C в атмосфере 5% CO₂.

За день до трансфекции клетки пересаживали в культуральный флакон площадью 75 см² для достижения примерно 70% конфлюэнтности клеток. Трансфекцию проводили в бессывороточной среде DMEM без добавления антибиотиков. При приготовлении смеси для трансфекции смешивали 2,4 мкл Opti-MEM («Thermo Fisher Scientific»), 18 мкг плазмидной ДНК и 36 мкг (2 мкг/мкл) реагента PEI («Sigma», США). Мы убедились, что эти условия обеспечивают эффективную трансфекцию клеток НЕК293Т (>50% трансфицированных клеток), поставив контрольные эксперименты с плазмидой, кодирующей GFP (eGFP-n1), трансфицированной вместе с целевым вектором.

Протокол 4С. Через 24 ч после трансфекции клетки собирали с помощью 0,05%-ного раствора трипсин-ЭДТА и фиксировали 1%-ным формальдегидом, в соответствии с ранее опубликованным протоколом [19]. Клетки подсчитывали, аликвотировали по 3 млн и замораживали. Основные этапы эксперимента, а именно: лизис клеток, гидролиз хроматина, лигирование хроматина, удаление формальдегидных сшивок и очистка ДНК, проводили согласно ранее опубликованному протоколу [19] с небольшими изменениями:

- для переваривания хроматина использовали 50 ед. NlaIII («NEB», США);
- были пропущены этапы биотинилирования и «удаления висящих концов».

Гидролиз очищенной ДНК. Условия гидролиза очищенной ДНК варьировали в зависимости от того, с каким вектором готовили 4С-библиотеки: образцам с векторами на основе p4SCS требуется гидролиз эндонуклеазой рестрикции MseI; образцам с векторами на основе pUC19 – TaiI.

Гидролиз в случае использования векторов на основе p4SCS. Для гидролиза очищенной ДНК на льду смешивали следующие компоненты: 5 мкл 10× CutSmart Buffer («NEB»), 30 ед. MseI, 2 мкг ДНК, ddH₂O до 50 мкл. Смесь инкубировали при 37 °C при встряхивании в течение ночи. После инактивации MseI путем инкубации при 65 °C в течение 20 мин образец очищали с помощью KAPA Pure Beads (1×) в соответствии с рекомендациями производителя («Roche», Швейцария) и добавляли в реакцию лигирования.

Гидролиз в случае использования векторов на основе pUC19. В 4С-экспериментах с модифицированным pUC19 мы проводили гидролиз очищенной ДНК ферментом TaiI: к 250 нг ДНК добавляли 1 мкл 10× рестрикционного буфера («Sibenzyme», Россия), 5 ед. TaiI, ddH₂O до 10 мкл. Рестрикционную смесь инкубировали при 37 °C при постоянном встряхивании в течение ночи, после чего гидролиз останавливали тепловой инактивацией

фермента при 65 °C в течение 20 мин. Весь объем реакционной смеси (10 мкл) добавляли в реакцию лигирования.

Лигирование очищенной ДНК. Для лигирования очищенной ДНК на льду смешивали следующие компоненты: 10 мкл 10%-ного Triton X-100, 10 мкл 10× реакционного буфера для лигазы ДНК T4 («NEB»), 10 мкл 25%-ного ПЭГ, 1 мкл бычьего сывороточного альбумина (10 мг/мл), 1 мкл 100 мМ АТР, 800 ед. лигазы T4, 0,5–2 мкг ДНК, ddH₂O до 100 мкл (конечная концентрация ДНК составляла ~5–20 нг/мкл). Смесь инкубировали при 16 °C в течение ночи, затем очищали с помощью KAPA Pure Beads (0,8×).

Дополнительный гидролиз ферментом DrdI (опционально). В наших первых экспериментах, которые проводились с векторами на основе pUC19 и с гидролизом ферментом TaiI, мы предположили, что доля неинформативных продуктов лигирования может быть уменьшена обработкой эндонуклеазой рестрикции DrdI. Хотя позже мы обнаружили, что это предположение неверно, следует отметить, что в первых экспериментах после лигирования продукты обрабатывали ферментом DrdI в следующих условиях: к смеси после лигирования добавляли 12 ед. эндонуклеазы DrdI и 100 мкл ddH₂O; инкубировали при 37 °C в течение ночи, затем фермент инактивировали инкубацией при 80 °C в течение 20 мин, и очищали ДНК с помощью KAPA Pure Beads (0,8×). При использовании векторов на основе p4SCS этот этап был пропущен.

Подготовка 4С-библиотек для высокопроизводительного секвенирования (NGS). ПЦР с праймерами, содержащими адаптерные последовательности Illumina, проводили в течение 15–20 циклов. Последовательности праймеров приведены в Приложении (текст 2). Продукты ПЦР очищали с помощью KAPA Pure Beads (0,8×). Распределение ДНК-фрагментов в полученных 4С-библиотеках оценивали визуализацией с помощью гель-электрофореза в 1%-ной агарозе.

Концентрацию ДНК измеряли с помощью флуорометра Qubit («Invitrogen», США).

Для экспериментов с модифицированной плазмидой pUC19 мы готовили 4С-библиотеки для NGS методом вложенной ПЦР: проводили первый раунд ПЦР с праймерами i_F1, i_R1 (см. текст 2 в Приложении), затем продукты ПЦР очищали с помощью KAPA Pure Beads (0,8×), и треть от количества продуктов ПЦР служила ДНК-матрицей для второго раунда ПЦР с праймерами F2 и R2 для секвенирования Illumina (текст 2 в Приложении).

Эксперименты с контрольным хроматином и контрольными плазмидами. Клетки НЕК293Т трансфицировали модифицированным вектором pUC19 с использованием PEI и фиксировали

через 24 ч после трансфекции, как описано выше. К фиксированным трансфицированным клеткам НЕК293Т добавляли равное количество фиксированных фибробластов *Mus musculus* и проводили лизис, как описано в 4С-протоколе выше. 200 нг контрольной плазмиды (модифицированная pUC19 с промотором цитомегаловируса (CMV) (рис. П2 в Приложении)) добавляли к образцу перед гидролизом NlaIII. Далее, мы провели 4С-эксперимент для всех этих компонент вместе.

4С-Библиотеки с обработкой фрагментом Кленова ДНК-полимеразы I. Чтобы отличить продукты, образовавшиеся в результате неполного гидролиза ферментом NlaIII от гидролизованных молекул, восстановивших исходную структуру на этапе лигирования, мы включили в описанный выше 4С-протокол этап обработки фрагментом Кленова ДНК-полимеразы I. Для этого после инактивации NlaIII, центрифугирования и удаления супернатанта смешивали на льду следующие компоненты: 20 мкл $\times 10$ NEBuffer 3.1 («NEB»), 1,5 мкл каждого 10 mM dNTP (dATP, dTTP, dCTP, dGTP), 50 ед. фрагмента Кленова, ddH₂O до 200 мкл. Смесь добавляли к осадку, ресуспендировали и инкубировали при 16 °C в течение 4 ч. Затем эксперимент продолжали с этапа «Лигирование тупых концов» протокола Hi-C 2.0 [19] до этапа очистки ДНК включительно и далее – в соответствии с описанным выше протоколом 4С с плазмидой.

Обработка и анализ 4С-данных. Прочтения обрабатывали с помощью cutadapt, а затем выравнивали на референсную последовательность, включающую геном человека версии hg38 и последовательности плазмидных векторов (модифицированный pUC19 или p4CSCS). Для экспериментов с контрольным хроматином в качестве референса использовали геном человека hg38, геном мыши mm10, последовательности модифицированного вектора pUC19 и модифицированного вектора pUC19 с CMV-промотором. Выравнивание проводилось программой BWA mem. Статистику выравнивания анализировали с помощью samtools idxstats. Для визуализации результатов выравнивания использовали программу IGV 2.16.1.

В нашем протоколе должно происходить два события лигирования: первое лигирование по рестрикционному сайту NlaIII в условиях фиксированного хроматина и второе лигирование «в кольцо» по рестрикционному сайту MseI или TaiI. Продукты первой реакции лигирования позволяют получить информацию о пространственной колокализации фрагментов ДНК в ядре, в то время как второе лигирование может происходить либо внутримолекулярно с образованием кольцевой ДНК, либо межмолекулярно между случайными концами ДНК. Таким образом, часть продуктов второго лигирования не несет инфор-

мации о пространственных контактах плазмиды, а отражает случайные столкновения молекул, возникающие в результате процессов диффузии. Поэтому мы не включали в анализ последовательности, лигированные со стороны сайта MseI или TaiI, а использовали только последовательно лигированные по сайту NlaIII. В соответствии с дизайном плазмидного вектора, эти последовательности всегда начинаются с последовательности праймера R2 (рис. П1–П3 и текст 2 в Приложении).

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Дизайн 4С-эксперимента с плазмидой. Цель наших экспериментов заключается в разработке подхода для выявления контактов плазмидной ДНК и генома клетки. Предложенный подход основан на использовании метода 4С и схематично представлен на рис. 1, а. Первоначально клетки трансфицируют плазмидным вектором. Мы предполагаем, что белки хроматина связываются с плазмидной ДНК и обуславливают ее распределение внутри ядра (I). Затем пространственные контакты в хроматине, включая взаимодействия между плазмидой и ДНК, фиксируются формальдегидом (II). В соответствии с 4С-протоколом (подробно описан в разделе «Материалы и методы») образцы фиксированного хроматина последовательно подвергают гидролизу и лигированию, что приводит к образованию химерных молекул между близкорасположенными фрагментами ДНК (III–IV). После этого формальдегидные сшивки удаляют, ДНК гидролизуют вторым ферментом рестрикции и лигируют в растворе (V). В результате образуются кольцевые молекулы, содержащие фрагмент плазмидной ДНК и участок генома, с которым контактировала плазида (VI). Амплификация и секвенирование этих кольцевых молекул с помощью технологии NGS позволяет обнаружить взаимодействие между плазмидной ДНК и геномом.

Согласно описанному дизайну, необходимо, чтобы вектор, используемый в 4С-эксперименте, включал два разных сайта рестрикции: первый гидролизует в фиксированном хроматине (E1; рис. 1, а), второй – после удаления формальдегидных сшивок (E2; рис. 1, а). Принципиально важно разместить области отжига ПЦР-праймеров между этими двумя сайтами рестрикции и как можно ближе к ним. Кроме того, между участками отжига праймеров не должно быть сайтов рестрикции E1 и E2. Также в качестве E1 и E2 следует использовать мелкоцепящие (т.е. с 4-буквенным сайтом узнавания) эндонуклеазы рестрикции, чтобы обеспечить высокое разрешение 4С-метода (определяется частотой встречаемости сайтов E1),

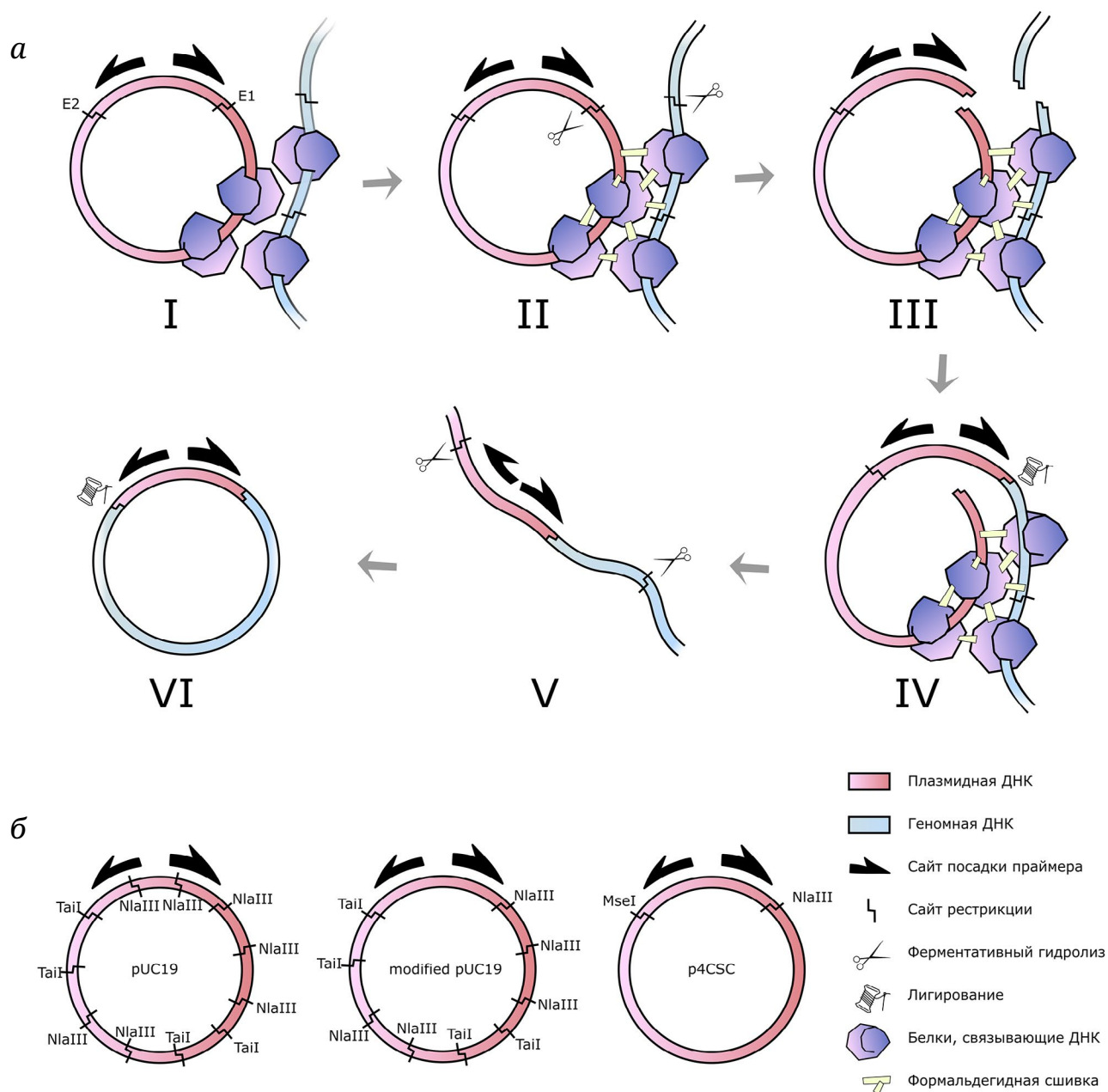


Рис. 1. Дизайн 4С-эксперимента с плазмидным вектором. **а** – Общий дизайн эксперимента (I–VI). **б** – Схема расположения сайтов рестрикции плазмидных векторов, используемых в данной работе. Карты и последовательности плазмид представлены на рис. П1–П3 и в тексте 2 в Приложении

и эффективную амплификацию 4С-фрагментов (определяется частотой сайтов E2). Наконец, необходимо, чтобы фермент E1 был способен гидролизовать хроматин, фиксированный формальдегидом.

Для проверки предложенного дизайна мы провели 4С-эксперименты с плазмидным вектором pUC19. В качестве целевых сайтов E1 и E2 в эксперименте использовали сайты рестрикции NlaIII и TaiI соответственно. Помимо двух указанных сайтов рестрикции, вектор pUC19 содержит дополнительные сайты NlaIII и TaiI (девять и три соответственно), включая те сайты NlaIII, которые

присутствуют в области отжига ПЦР-праймеров (рис. 1, б). Мы ожидали, что дополнительные сайты NlaIII и TaiI за пределами области связывания праймеров, используемых для амплификации библиотеки, не будут мешать проведению 4С-эксперимента, в отличие от сайтов, расположенных между праймерами. Поэтому мы сконструировали модифицированный вектор на основе pUC19, в котором были удалены сайты рестрикции NlaIII (-CATG-), расположенные между праймерами (рис. 1, б).

Оценка количества неспецифических продуктов, образующихся в 4С-протоколе с плазмидой. Дизайн 4С-эксперимента с плазмидой пред-

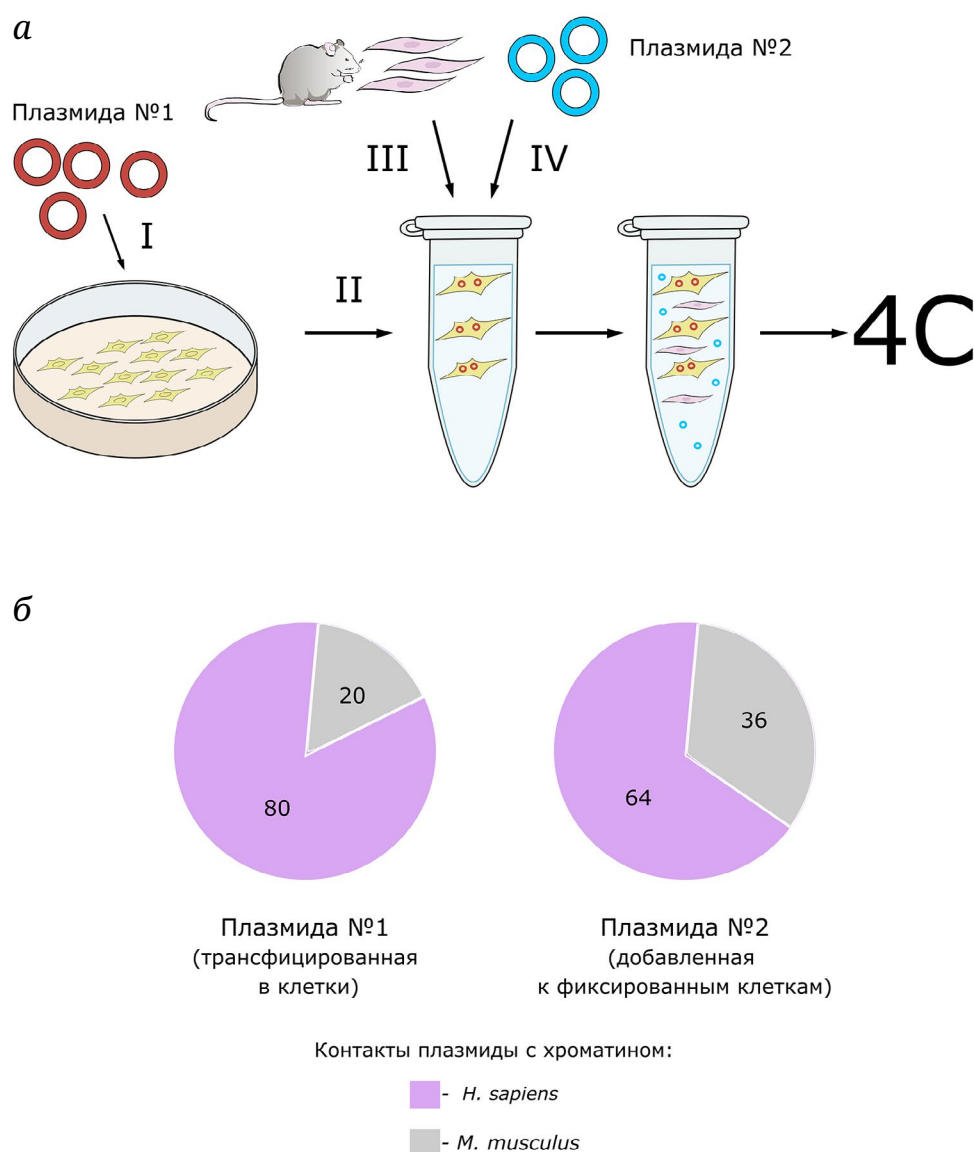


Рис. 2. Оценка доли «случайных» контактов. **а** – Схема эксперимента: плазмиду № 1 трансфицировали в клетки HEK293T (I), затем пространственные контакты ДНК и белков в ядре фиксировали формальдегидом (II). К полученному образцу добавляли фиксированные фибробласты *Mus musculus* (III), плазмиду № 2 (IV) и выполняли 4C-эксперимент. **б** – Плазмида № 1 и плазмида № 2 различаются по соотношению количества контактов с хроматином *H. sapiens* и *M. musculus*. На диаграммах показано процентное соотношение контактов плазмиды с соответствующим хроматином, усредненное по трем независимым экспериментам. Результаты по отдельным репликам представлены на рис. П4 в Приложении. Общее количество контактов принято за 100%

полагает, что этапы рестрикции и лигирования близкорасположенных фрагментов ДНК происходят в условиях фиксированного хроматина, поэтому образующиеся продукты отражают пространственную колокализацию плазмидной и геномной ДНК внутри ядра. Однако не исключено, что часть контактов образуется не внутри ядра, а в растворе в результате диффузионных процессов после фиксации хроматина, приводящих к случайным столкновениям и лигированию фрагментов плазмидной и геномной ДНК. Такие контакты не будут отражать биологических предпочтений распределения плазмид в ядре.

Чтобы оценить частоту таких событий, мы провели контрольный эксперимент, представленный на рис. 2, а. К фиксированным клеткам человека HEK293T, трансфицированным модифицированным вектором pUC19, добавляли контрольный хроматин (фиксированный хроматин другого вида) и контрольную плазмиду (плазмиду с последовательностью, отличной от вектора pUC19). Используя полученный образец, мы подготовили 4C-библиотеки, как описано выше (рис. 1, а). (Детали см. в разделе «Материалы и методы»; «Эксперименты с контрольным хроматином и контрольными плазмидами».)

Учитывая, что контрольный хроматин из клеток отличных от человека видов был добавлен после фиксации, любые лигирования вектора рUC19 (который был трансфицирован в клетки человека НЕК293Т до фиксации) с хроматином других видов представляют собой случайные столкновения молекул. Аналогично лигирование контрольной плазмиды (добавленной после фиксации) с хроматином НЕК293Т и с хроматином других видов также обусловлены диффузией в растворе.

Таким образом, контакты плазмиды № 1 (модифицированной рUC19, трансфицированной в НЕК293Т) отражают как пространственную локализацию плазмиды в живых клетках (при лигировании с хроматином человека), так и возможные случайные процессы диффузии (при лигировании с хроматином человека и других видов). Контакты плазмиды № 2, добавленной в клетки после фиксации с хроматином, не обусловлены биологическими закономерностями и определяются только случайными процессами диффузии (при лигировании с хроматином человека и с хроматином другого вида).

Мы сравнили, как часто плазида № 1, трансфицированная в клетки человека, и плазида № 2, добавленная к клеткам после фиксации, контактируют с хроматином человека и хроматином другого вида. Наши результаты, представленные на рис. 2, б, показывают, что 80% контактов плазмиды № 1 происходит с хроматином человека, в то время как плазида № 2 взаимодействует с хроматином человека только в 64% случаев (представленные значения являются средними для трех повторностей эксперимента; для всех повторностей разница в частоте контактов была статистически значимой, согласно точному тесту Фишера, *p*-значение пренебрежимо мало; подробности см. в Примечании). Эти наблюдения позволяют предположить, что часть обнаруженных в эксперименте контактов между трансфицированной плазмидой и хроматином обусловлена пространственной близостью в ядре, т.е. не может быть объяснена случайными столкновениями плазмидной ДНК с геномной ДНК в растворе. Мы также количественно оценили долю случайных столкновений (см. Примечание).

Оптимизация 4С-вектора. Анализируя данные, полученные в 4С-эксперименте с модифицированным вектором рUC19, мы столкнулись с проблемой: большинство (~90%) прочтений не содержали последовательности генома человека. Вместо этого прочтения выравнивались на разные участки плазмиды. Эта проблема наблюдалась не только при использовании контрольного хроматина и контрольной плазмиды, но и в 4С-экспериментах с клетками НЕК293Т, трансфицированными модифицированным вектором рUC19 без

контроля. Детальный анализ распределения и структуры прочтений показал, что это явление может быть связано с лигированием различных фрагментов плазмиды на этапе внутрихроматинового лигирования. Это происходит потому, что плазмидная ДНК гидролизует не только в предполагаемых сайтах рестрикции *NlaIII* и *Tail*, но и в других сайтах *NlaIII* и *Tail*, присутствующих в плазмиде. В условиях, способствующих лигированию близкорасположенных участков ДНК, фрагменты, происходящие из одной молекулы плазмиды, с большей вероятностью подвергаются лигированию в *cis*-положении, и это явление мы назвали «самолигированием». В соответствии со значительным преобладанием *cis*-взаимодействий над *trans*-, характерным для 3С-данных [20], прочтения, содержащие только плазмидные последовательности, составляют ~90% от общего количества данных.

Высокая частота лигирования между фрагментами плазмидной ДНК, как описано выше, может быть объяснена преобладанием *cis*- (внутри плазмиды) над *trans*- (плазида-геном) пространственными контактами и поэтому неизбежна в 3С-эксперименте. Однако эта проблема может быть решена, если в плазмидном векторе будут отсутствовать множественные сайты гидролиза. В этом случае *cis*-лигирование восстановит исходную молекулу плазмиды, которая из-за своей длины (>1000 п.н.) будет исключена из 4С-библиотеки на этапах ПЦР и секвенирования NGS. Поэтому для решения проблемы самолигирования плазмиды мы стремились модифицировать последовательность плазмиды таким образом, чтобы она содержала только один сайт рестрикции *NlaIII* и *MseI* (в данной конструкции вектора мы используем *MseI* вместо *Tail* на этапе гидролиза очищенной ДНК). Трудность этой задачи в том, что оба фермента являются мелкоцепящими эндонуклеазами рестрикции (что важно для разрешения метода, см. выше), и поэтому во всех функциональных элементах доступных плазмидных векторов присутствуют многочисленные сайты узнавания этих ферментов.

Чтобы преодолеть это ограничение, мы заменили ген устойчивости к ампициллину геном устойчивости к хлорамфениколу, лишенным сайтов рестрикции в промоторной области. Однако ориджин репликации и тело гена хлорамфеникола по-прежнему содержали сайты рестрикции *NlaIII* и *MseI*. Мы внесли однонуклеотидные замены в кодирующую часть гена устойчивости к хлорамфениколу таким образом, чтобы сохранить аминокислотную последовательность неизменной. Выбор альтернативных нуклеотидов основывался на частоте встречаемости кодонов у *E. coli* [21]. В ориджине репликации замены производились

случайным образом, поскольку у нас отсутствовала возможность предсказать их функциональные последствия. В результате мы разработали новую последовательность плазмиды, адаптированную для 4С-эксперимента, содержащую только два целевых сайта рестрикции p4CSCS: (plasmid for 4C with Single Cutter Sites), схематично изображенную на рис. 1, б (карта плазмиды и ее последовательность представлены на рис. ПЗ и в тексте 1 в Приложении). Она содержит регионы связывания праймеров и необходимые для размножения *E. coli* элементы, такие как ориджин репликации и ген устойчивости к хлорамфениколу. Общая длина вектора составляет 1826 п.н.

Мы использовали этот улучшенный плазмидный вектор в 4С-эксперименте на клетках НЕК293Т. Полученные 4С-библиотеки содержали значительную долю контактов между плазмидой и геномом – около 81%. Таким образом, мы получили подтверждение того, что наблюдавшаяся в предыдущих экспериментах высокая частота прочтений, соответствующих плазмидной последовательности, была обусловлена наличием в векторе нескольких сайтов рестрикции. Полученные результаты наглядно демонстрируют преимущество разработанной последовательности плазмиды для 4С-экспериментов.

Хотя контакты между плазмидой и геномом хорошо представлены, в 4С-библиотеках сохраняются прочтения, включающие в себя только плазмидные последовательности. Такие продукты составляют ~20% данных. Согласно визуализации в IGV, большинство таких прочтений затрагивает область, включающую сайт рестрикции *NlaIII*. Это наблюдение поднимает вопрос о происхождении таких прочтений: образуются ли они в результате неполного гидролиза ферментом *NlaIII*, или же молекулы плазмиды были гидролизированы, но затем лигированы обратно с восстановлением молекулы плазмиды?

Чтобы проверить эти гипотезы, мы подготовили 4С-библиотеки, в которых хроматин после гидролиза ферментом *NlaIII* обрабатывали фрагментом Кленова ДНК-полимеразы I *E. coli*. В результате гидролиза ферментом *NlaIII* образуются 3'-концы, которые удаляются 3'→5' экзонуклеазной активностью фрагмента Кленова. Последующее лигирование концов плазмиды создает специфическую последовательность, отличную от последовательности непереваренной плазмиды. Таким образом, при секвенировании продуктов лигирования становится возможным различать молекулы, происходящие из негидролизованной плазмиды (содержащие последовательность 5'...TCTGAC-CATG-AGGAGA...3', включающую сайт рестрикции *NlaIII* 5'...-CATG-...3') и молекулы, которые были гидролизированы *NlaIII*, но затем лигированы обратно

(содержащие укороченную последовательность, после удаления 3'-концов фрагментом Кленова 5'...TCTGAC-AGGAGA...3', без сайта рестрикции *NlaIII*). Данные секвенирования, полученные в шести независимых экспериментах, показали, что количество негидролизированных продуктов (со средней долей 4,82% от всех прочтений) почти в 30 раз превышало количество молекул, которые были гидролизированы и лигированы обратно (средняя доля – 0,17% от всех прочтений). Это указывает на то, что молекулы, содержащие неповрежденную последовательность плазмиды, в основном обусловлены неэффективным гидролизом ферментом *NlaIII*.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В данной работе мы представляем метод, позволяющий выявлять пространственные контакты между плазмидной ДНК, введенной в клетку, и геномом клетки. Наши эксперименты показали многообещающие результаты: 4С-протокол позволяет эффективно детектировать контакты между плазмидной и геномной ДНК. Хотя взаимодействия, представляющие собой контакты между геномными локусами, происходят примерно в миллион раз чаще, чем между геномом и плазмидой (согласно соотношению длины генома человека и приблизительной длины используемых в эксперименте плазмид), значительная часть контактов, обнаруживаемых в 4С-эксперименте, представляет собой взаимодействие геномной и плазмидной ДНК. Оптимизировав дизайн плазмидного вектора, мы добились снижения доли *cis*-контактов, уровень которых высок для 3С-методов из-за внутримолекулярного лигирования молекул, в том числе плазмиды. В перспективе специфичность и информативность этого метода могут быть повышены за счет более эффективного гидролиза хроматина. Это подтверждается нашим наблюдением о том, что часть неинформативных продуктов происходит из негидролизированных плазмидных молекул.

Следует отметить, что в 3С данные, не отражающие смысловые контакты (или «шум»), как ожидается, распределены равномерно и поэтому не должны препятствовать выявлению повышенной частоты контактов между плазмидой и конкретными геномными локусами. Однако высокий уровень «шума» может потребовать более глубокого секвенирования для детектирования специфических взаимодействий.

Мы предлагаем несколько вариантов применения разработанного метода.

- Во-первых, предложенный подход может быть использован для исследования локализации

самих плазмидных молекул. В настоящее время, несмотря на широкое применение плазмид в качестве векторов для экзогенной экспрессии генов в генной инженерии и научных исследованиях, распределение трансфицированных плазмид в ядре клетки остается во многом неясным [22–24].

- Во-вторых, предлагаемый дизайн эксперимента позволяет проверить, как встроенные в плазмиду последовательности ДНК с различными эпигенетическими свойствами (активные и неактивные участки генома, GC-богатые и GC-бедные регионы, сайты связывания белков Polysomb и HP1 и т.д.) будут влиять на локализацию плазмиды. Это может быть применимо в изучении компартментов, образующихся в результате фазовой сепарации ядра. Хотя известно, что в ядре происходит пространственное разделение активного и неактивного хроматина, вопрос о том, сколько всего образуется фаз (или субкомпарментов), остается открытым [25, 26]. Компартментализация ядра активно изучается с точки зрения белков. В то же время это явление может быть рассмотрено и с точки зрения роли мотивов ДНК, с которыми связываются белки, участвующие в фазовом разделении. Несколько таких последовательностей ДНК, необходимых для формирования специфического паттерна контактов локуса, уже обнаружено. Например, с помощью методов Hi-C и ChIP-seq были идентифицированы последовательности ДНК, определяющие профиль контактов для Polysomb-репрессированных локусов в эмбрионах дрозофилы [27]. В остальном генетические детерминанты компартментализации остаются неизвестными. Для их выявления мы предлагаем 4С-метод с плазмидой, применяя его для оценки влияния различных последовательностей ДНК на паттерн контактов плазмиды с геномом. В перспективе метод может быть использован в качестве основы для крупномасштабных скрининговых исследований.

- Метилирование ДНК изменяет трехмерную организацию генома и доступность хроматина, привлекая белки, такие как HP1, которые в результате фазового разделения формируют блоки гетерохроматина [25, 28, 29]. Метилирование сайтов связывания CTCF нарушает его ассоциацию с ДНК, приводя к изменениям в архитектуре генома [30]. В целом, метилирование влияет на способность транскрипционных факторов связываться с ДНК, причем в зависимости от фактора транскрипции метилирование может приводить как к увеличению, так и к уменьшению аффинности связывания *in vitro* [31]. Однако это не было подтверждено *in vivo* [28]. Мы предлагаем использование метода 4С с *in vitro* 5mC-метилированной плазмидной

ДНК как стратегический подход для локализации метилированных локусов ДНК *in vivo* и выяснения, объединяются ли они в отдельные субкомпарменты.

- Метод 4С с плазмидой может быть использован для изучения репарации ДНК. Плазида с модифицированными нуклеотидами, такими как включения 8-оксогуанина, или другими модификациями, вероятно, будет локализована в специфических компартаментах репарации ДНК, сборка которых инициируется активированными PARP1 и FUS [32]. 4С-Эксперимент может быть использован для захвата локусов, привлеченных к компартменту репарации ДНК, а также для отслеживания динамики его формирования.

Полученные результаты дают основания полагать, что усовершенствованный эксперимент 4С с плазмидой будет полезен в областях 3D-геномики, регуляции транскрипции и эпигенетики.

Вклад авторов. В.Ф., П.С. и А.Я. разработали концепцию исследования; А.Я. провела эксперименты при участии П.С. и М.Г.; П.Б. разработала алгоритмы для вычислительного анализа данных; А.Я. выполнила анализ данных при участии П.Б., П.С. и В.Ф.; В.Ф. руководил исследованием и анализировал данные; А.Я., П.С. и В.Ф. написали текст. Все авторы редактировали рукопись и одобрили ее окончательный вариант.

Финансирование. Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (грант № FSUS-2024-0018).

Благодарности. Мы выражаем признательность центру коллективного пользования Института цитологии и генетики СО РАН, 121031800061-7 (Механизмы генетического контроля развития, физиологических процессов и поведения животных) за проведение экспериментов с использованием технологий секвенирования нового поколения. Вычислительный анализ данных был выполнен на высокопроизводительных кластерах Новосибирского государственного университета (поддержано Министерством науки и высшего образования Российской Федерации (грант № FSUS-2024-0018)).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Соблюдение этических стандартов. Настоящая статья не содержит описания каких-либо исследований с участием людей или животных в качестве объектов.

Дополнительные материалы. Приложение к статье опубликовано на сайте журнала «Биохимия» (<https://biochemistrymoscow.com>).

СПИСОК ЛИТЕРАТУРЫ

1. Kabirova, E., Nurislamov, A., Shadskiy, A., Smirnov, A., Popov, A., Salnikov, P., Battulin, N., and Fishman, V. (2023) Function and evolution of the loop extrusion machinery in animals, *Int. J. Mol. Sci.*, **24**, 5017, <https://doi.org/10.3390/ijms24055017>.
2. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N., and Mirny, L. A. (2018) Chromatin organization by an interplay of loop extrusion and compartmental segregation, *Proc. Natl. Acad. Sci. USA*, **115**, E6697-E6706, <https://doi.org/10.1073/pnas.1717730115>.
3. Fishman, V., Battulin, N., Nuriddinov, M., Maslova, A., Zlotina, A., Strunov, A., Chervyakova, D., Korablev, A., Serov, O., and Krasikova, A. (2019) 3D organization of chicken genome demonstrates evolutionary conservation of topologically associated domains and highlights unique architecture of erythrocytes' chromatin, *Nucleic Acids Res.*, **47**, 648-665, <https://doi.org/10.1093/nar/gky1103>.
4. Ryzhkova, A., Taskina, A., Khabarova, A., Fishman, V., and Battulin, N. (2021) Erythrocytes 3D genome organization in vertebrates, *Sci. Rep.*, **11**, 4414, <https://doi.org/10.1038/s41598-021-83903-9>.
5. Razin, S. V., and Gavrilov, A. A. (2020) The role of liquid-liquid phase separation in the compartmentalization of cell nucleus and spatial genome organization, *Biochemistry (Moscow)*, **85**, 643-650, <https://doi.org/10.1134/S0006297920060012>.
6. Kantidze, O. L., and Razin, S. V. (2020) Weak interactions in higher-order chromatin organization, *Nucleic Acids Res.*, **48**, 4614-4626, <https://doi.org/10.1093/nar/gkaa261>.
7. Nuriddinov, M., and Fishman, V. (2019) C-InterSection—a computational tool for interspecies comparison of genome architecture, *Bioinformatics (Oxford, England)*, **35**, 4912-4921, <https://doi.org/10.1093/bioinformatics/btz415>.
8. Lukyanchikova, V., Nuriddinov, M., Belokopytova, P., Taskina, A., Liang, J., Reijnders, J. M. F., Ruzzante, L., Feron, R., Waterhouse, R. M., Wu, Y., Mao, C., Tu, Z., and Sharakhov, I. V. (2022) Anopheles mosquitoes reveal new principles of 3D genome organization in insects, *Nat. Commun.*, **13**, 1960, <https://doi.org/10.1038/s41467-022-29599-5>.
9. Dias, J. D., Sarica, N., Cournac, A., Koszul, R., and Neuveut, C. (2022) Crosstalk between hepatitis B virus and the 3D genome structure, *Viruses*, **14**, 445, <https://doi.org/10.3390/v14020445>.
10. Tang, D., Zhao, H., Wu, Y., Peng, B., Gao, Z., Sun, Y., Duan, J., Qi, Y., Li, Y., Zhou, Z., Guo, G., Zhang, Y., Li, C., Sui, J., and Li, W. (2021) Transcriptionally inactive hepatitis B virus episome DNA preferentially resides in the vicinity of chromosome 19 in 3D host genome upon infection, *Cell Rep.*, **35**, 109288, <https://doi.org/10.1016/j.celrep.2021.109288>.
11. Sokol, M., Wabl, M., Ruiz, I. R., and Pedersen, F. S. (2014) Novel principles of gamma-retroviral insertional transcription activation in murine leukemia virus-induced end-stage tumors, *Retrovirology*, **11**, 36, <https://doi.org/10.1186/1742-4690-11-36>.
12. Razin, S. V., Gavrilov, A. A., and Iarovaia, O. V. (2020) Modification of nuclear compartments and the 3D genome in the course of a viral infection, *Acta Naturae*, **12**, 34-46, <https://doi.org/10.32607/actanaturae.11041>.
13. Everett, R. D. (2013) The spatial organization of DNA virus genomes in the nucleus, *PLoS Pathog.*, **9**, e1003386, <https://doi.org/10.1371/journal.ppat.1003386>.
14. Corpet, A., Kleijwegt, C., Roubille, S., Juillard, F., Jacquet, K., Texier, P., and Lomonte, P. (2020) PML nuclear bodies and chromatin dynamics: catch me if you can! *Nucleic Acids Res.*, **48**, 11890-11912, <https://doi.org/10.1093/nar/gkaa828>.
15. Rai, T. S., Glass, M., Cole, J. J., Rather, M. I., Marsden, M., Neilson, M., Brock, C., Humphreys, I., Everett, R., and Adams, P. (2017) Histone chaperone HIRA deposits histone H3.3 onto foreign viral DNA and contributes to anti-viral intrinsic immunity, *Nucleic Acids Res.*, **45**, 11673-11683, <https://doi.org/10.1093/nar/gkx771>.
16. Schmid, M., Speiseder, T., Dobner, T., and Gonzalez, R. A. (2014) DNA virus replication compartments, *J. Virol.*, **88**, 1404-1420, <https://doi.org/10.1128/JVI.02046-13>.
17. Charman, M., and Weitzman, M. D. (2020) Replication compartments of DNA viruses in the nucleus: location, location, *Viruses*, **12**, 151, <https://doi.org/10.3390/v12020151>.
18. Kempfer, R., and Pombo, A. (2020) Methods for mapping 3D chromosome architecture, *Nat. Rev. Genet.*, **21**, 207-226, <https://doi.org/10.1038/s41576-019-0195-2>.
19. Belaghzal, H., Dekker, J., and Gibcus, J. H. (2017) Hi-C 2.0: an optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation, *Methods*, **123**, 56-65, <https://doi.org/10.1016/j.ymeth.2017.04.004>.
20. Gridina, M., Mozheiko, E., Valeev, E., Nazarenko, L. P., Lopatkina, M. E., Markova, Z. G., Yablonskaya, M. I., Voinova, V. Y., Shilova, N. V., Lebedev, I. N., and Fishman, V. (2021) A cookbook for DNase Hi-C, *Epigenet. Chromatin*, **14**, 15, <https://doi.org/10.1186/s13072-021-00389-5>.
21. Gvritishvili, A. G., Leung, K. W., and Tombran-Tink, J. (2010) Codon preference optimization increases heterologous PEDF expression, *PLoS One*, **5**, e15056, <https://doi.org/10.1371/journal.pone.0015056>.
22. Prajapati, H. K., Kumar, D., Yang, X.-M., Ma, C.-H., Mittal, P., Jayaram, M., and Ghosh, S. (2020) Hitchhiking on condensed chromatin promotes plasmid persistence in yeast without perturbing chromosome function, *bioRxiv*, <https://doi.org/10.1101/2020.06.08.139568>.

23. Gracey Maniar, L. E., Maniar, J. M., Chen, Z.-Y., Lu, J., Fire, A. Z., and Kay, M. A. (2013) Minicircle DNA vectors achieve sustained expression reflected by active chromatin and transcriptional level, *Mol. Ther.*, **21**, 131-138, <https://doi.org/10.1038/mt.2012.244>.
24. Dean, D. A. (1997) Import of plasmid DNA into the nucleus is sequence specific, *Exp. Cell Res.*, **230**, 293-302, <https://doi.org/10.1006/excr.1996.3427>.
25. Mladenova, V., Mladenov, E., and Russev, G. (2009) Organization of plasmid DNA into nucleosome-like structures after transfection in eukaryotic cells, *Biotechnol. Biotechnolog. Equip.*, **23**, 1044-1047, <https://doi.org/10.1080/13102818.2009.10817609>.
26. Hildebrand, E. M., and Dekker, J. (2020) Mechanisms and functions of chromosome compartmentalization, *Trends Biochem. Sci.*, **45**, 385-396, <https://doi.org/10.1016/j.tibs.2020.01.002>.
27. Erdel, F., and Rippe, K. (2018) Formation of chromatin subcompartments by phase separation, *Biophys. J.*, **114**, 2262-2270, <https://doi.org/10.1016/j.bpj.2018.03.011>.
28. Ogiyama, Y., Schuettengruber, B., Papadopoulos, G. L., Chang, J.-M., and Cavalli, G. (2018) Polycomb-dependent chromatin looping contributes to gene silencing during *Drosophila* development, *Mol. Cell*, **71**, 73-88.e5, <https://doi.org/10.1016/j.molcel.2018.05.032>.
29. Mattei, A. L., Bailly, N., and Meissner, A. (2022) DNA methylation: a historical perspective, *Trends Genet.*, **38**, 676-707, <https://doi.org/10.1016/j.tig.2022.03.010>.
30. Rountree, M. R., and Selker, E. U. (2010) DNA methylation and the formation of heterochromatin in *Neurospora crassa*, *Heredity*, **105**, 38-44, <https://doi.org/10.1038/hdy.2010.44>.
31. Phillips, J. E., and Corces, V. G. (2009) CTCF: master weaver of the genome, *Cell*, **137**, 1194-1211, <https://doi.org/10.1016/j.cell.2009.06.001>.
32. Singatulina, A. S., Hamon, L., Sukhanova, M. V., Desforges, B., Joshi, V., Bouhss, A., Lavrik, O. V., and Pastre, D. (2019) PARP-1 activation directs FUS to DNA damage sites to form PARG-reversible compartments enriched in damaged DNA, *Cell Rep.*, **27**, 1809-1821, <https://doi.org/10.1016/j.celrep.2019.04.031>.

TOWARDS DEVELOPMENT OF THE 4C-BASED METHOD DETECTING INTERACTIONS OF PLASMID DNA WITH HOST GENOME

A. P. Yan^{1,2*}, P. A. Salnikov^{1,2}, M. M. Gridina^{1,2},
P. S. Belokopytova^{1,2}, and V. S. Fishman^{1,2*}

¹ Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences,
630090 Novosibirsk, Russia; e-mail: a.yan@g.nsu.ru; minja-f@yandex.ru

² Novosibirsk State University, 630090 Novosibirsk, Russia

Chromosome conformation capture techniques have revolutionized our understanding of chromatin architecture and dynamics at the genome-wide scale. In recent years, these methods have been applied to a diverse array of species, revealing fundamental principles of chromosomal organization. However, structural organization of the extrachromosomal entities, like viral genomes or plasmids, and their interactions with the host genome, remain relatively underexplored. In this work, we introduce an enhanced 4C-protocol tailored for probing plasmid DNA interactions. We design specific plasmid vector and optimize protocol to allow high detection rate of contacts between the plasmid and host DNA.

Keywords: chromatin interactions, 4C, plasmid, extrachromosomal sequences, plasmid-genome interactions, chromosome conformation capture, transfection