

УДК 577.214.5

Тип клеток как один из основных факторов, определяющих распределение R-петель в геноме

К. Ю. Олейникова^{1,2}, Н. А. Жигалова¹, Э. П. Хатчинс³, А. С. Рузов^{1*}¹Институт биоинженерии, ФИЦ Биотехнологии РАН, Москва, 117312 Россия²Центр инфохимии, Университет ИТМО, Санкт-Петербург, 197101 Россия³Южный университет науки и технологий, кафедра системной биологии, факультет естественных наук, Шэньчжэнь, 518055 Китай

*E-mail: alexey.ruzov@gmail.com

Поступила в редакцию 26.09.2025

Принята к печати 27.01.2026

DOI: 10.32607/actanaturae.27833

РЕФЕРАТ R-петли – молекулярные структуры, состоящие из ДНК:РНК-гибрида и неспаренной одноцепочечной ДНК, – участвуют как в нормальной жизнедеятельности клеток, так и в патогенезе многочисленных заболеваний. Разработано несколько подходов к картированию R-петель в геноме, однако результаты, полученные разными методами, часто противоречат друг другу. В недавнем исследовании Чедина и соавт. сравнили геномное распределение R-петель в нормальных и опухолевых клеточных линиях, определенное с использованием различных методологий. Важно заметить, что авторы этого исследования исходили из неартикулированного постулата о высокой степени сходства между геномным распределением R-петель в разных типах клеток. В настоящей работе сравнили наборы данных ДНК:РНК-иммунопреципитации (DNA:RNA IP, DRIP), полученные с использованием одного и того же протокола на разных клеточных линиях, и показали, что только 26% пиков R-петель являются общими для клеток HAP1, происходящих из хронического миелоидного лейкоза, и плюрипотентных стволовых клеток человека. В то же время в линиях клеток с нокаутом двух генов, полученных от родительской линии HAP1, намного более значительные доли пиков R-петель идентичны как друг другу (подавляющее большинство), так и пикам R-петель их родительской линии (71 и 57.7%). Сделан вывод, что тип клеток является одним из основных факторов, определяющих геномное распределение R-петель, и, следовательно, только систематическое сравнение большого массива наборов данных, полученных из разных типов клеток/тканей, может устранить несоответствия между различными методами картирования этих структур.

КЛЮЧЕВЫЕ СЛОВА R-петли, ДНК:РНК-гибриды, ДНК-РНК-иммунопреципитация (DRIP), клеточная дифференцировка, hPSC, клетки HAP1.

СПИСОК СОКРАЩЕНИЙ DRIP – ДНК-РНК-иммунопреципитация; HNRNPA2B1 – гетерогенный ядерный рибонуклеопротеин A2/B1; hPSC – плюрипотентные стволовые клетки человека.

ВВЕДЕНИЕ

R-молекулярные структуры, состоящие из ДНК:РНК-гибрида и неспаренной одноцепочечной ДНК, широко распространены в геноме и могут участвовать в регуляции широкого спектра биологических процессов, таких как терминация транскрипции, репарация ДНК, гомеостаз теломер и рекомбинация с переключением класса иммуноглобулинов [1]. В то же время нерегулярные или патологические R-петли могут нарушать транскрипцию и репликацию, приводя к накоплению двухцепочечных разрывов ДНК и представляя собой основной источ-

ник генетического стресса и нестабильности в клетках млекопитающих [1]. Учитывая связь нестабильности генома с онкотрансформацией, интерес к регуляции распределения R-петель в различных системах привел к разработке экспериментальных методов картирования R-петель в геноме [2]. Некоторые из этих методов основаны на использовании антител, которые специфически распознают РНК:ДНК-гибриды (антитела S9.6) [3]. Эти методики включают иммунопреципитацию ДНК:РНК (DRIP) вместе с ее многочисленными вариантами (DRIPc-seq; ssDRIP-seq и т.д.) и метод, основан-

ный на расщеплении ДНК под целевыми участками с последующей тагментацией (R-Loop Cleavage Under Targets & Tagmentation) [4, 5]. Другие методы основаны либо на картировании одноцепочечной ДНК в составе R-петель с помощью бисульфита [6], либо на использовании РНКазы H1, фермента, ответственного за распознавание и последующее расщепление РНК:ДНК-гибридов [7]. Важно отметить, что наборы R-петель, полученные с использованием разных (а иногда и аналогичных) методов на разных клеточных линиях, часто могут иметь существенные различия [8].

Для устранения этих противоречий Чедин (Chedin) и соавт. сравнили наборы данных геномного распределения R-петель, полученных разными методами на разных клеточных линиях, чтобы оценить степень их сходства [8]. Стоит отметить, что в этой работе не полностью учитываются типы клеток, на которых получены рассматриваемые наборы данных. Неявное предположение заключалось в том, что существует принципиальное сходство в распределении R-петель в генах домашнего хозяйства в различных типах клеток. Тем не менее, неявно, насколько обосновано это предположение. Чедин и соавт. не оценивают и не обсуждают потенциальную степень различий между геномным распределением R-петель в разных клеточных системах, таких как плюрипотентные стволовые клетки человека (hPSC) и линии опухолевых клеток (U2OS или HeLa), и используют сравнение профилей покрытия сигнала соответствующих экспериментов DRIP из различных типов клеток для выявления наборов данных, которые они считают «диссонансными» (discordant) [8]. Мы полагаем, что такой подход, неявно предполагающий существование сходства в геномном распределении R-петель в клеточных линиях различного происхождения, может привести к потенциальному искажению интерпретации экспериментальных данных.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Клеточные линии и клеточная культура, DRIP и подготовка библиотек для высокопроизводительного секвенирования

Клетки HNRNPA2B1 KO, YTHDF2 KO и их изогенные родительские клетки HAP1 дикого типа (Horizon Discovery, № HZGHC007378c010, № HZGHC006678c001 и № C631) культивировали на среде DMEM/F12 (Gibco Life Technologies, США № 11320033). ДНК-РНК-иммунопреципитацию (DRIP) проводили согласно ранее опубликованной методике, включая контроль специфичности сигнала DRIP с помощью обработки образцов РНКазой H

[9]. Геномные библиотеки были подготовлены с помощью набора NEBNext Ultra II DNA Library Prep Kit для Illumina (NEB, США № E7645) согласно инструкции производителя.

Биоинформатический анализ данных DRIP-секвенирования

Параллельно анализировали наборы данных HAP1 (WT, HNRNPA2B1 KO и YTHDF2 KO) (PRJNA1250978) и ранее опубликованный набор данных hPSC человека (PRJNA474076). Пики определяли с помощью MACS2 peak caller. Обнаружение широких пиков выполняли с параметрами: `-format BAM -g hs -keep-dup all`. Консенсусные пики идентифицировали с помощью инструмента bedtools intersect со стандартными параметрами для минимального пересечения пиков на входе (`-f 1E-9 -F 1E-9`). Детальное описание биоинформатического анализа доступно на сайте (https://github.com/katerinaoleynikova/human_samples_paper_25).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Чтобы оценить уровень сходства между геномным распределением R-петель в клетках разного происхождения, сравнили наборы данных DRIP, полученные с использованием ранее опубликованного нами протокола [9] на трех различных клеточных моделях: hPSC, клетках HAP1 «дикого типа» (HAP1 WT), полученных из линии клеток хронического миелоидного лейкоза KBM-7 [10], и двух изогенных линиях клеток HAP1 с нокаутом генов YTHDF2 или HNRNPA2B1 с помощью генетического редактирования (клеточные линии HAP1 YTHDF2 KO и HAP1 2B1 KO соответственно). Для определения геномных областей, обогащенных R-петлями в этих клеточных линиях, нами выполнено DRIP-секвенирование, прочтения картированы на геном человека, выявлены пики R-петель, после чего были идентифицированы пики, достоверные для каждого образца, путем сравнения двух соответствующих реплик. Показано, что, хотя в hPSC и HAP1 дикого типа выявлено сходное количество достоверных пиков (4482 и 4458 соответственно), только 1187 пиков (около 26%) были общими в этих двух линиях клеток (рис. 1А). Интересно, что, в то время как нокаут генов YTHDF2 и HNRNPA2B1 привел к значительному снижению общего количества пиков (1496 и 950 соответственно) по сравнению с клетками HAP1 дикого типа, наборы данных HAP1 YTHDF2 KO и HAP1 2B1 KO были сходными; при этом подавляющее большинство пиков HAP1 2B1 KO (873 из 950) перекрывалось с пиками, выявленными в клетках HAP1 YTHDF2 KO (рис. 1Б). Более того, 71% (681) пиков, идентифицированных в клетках

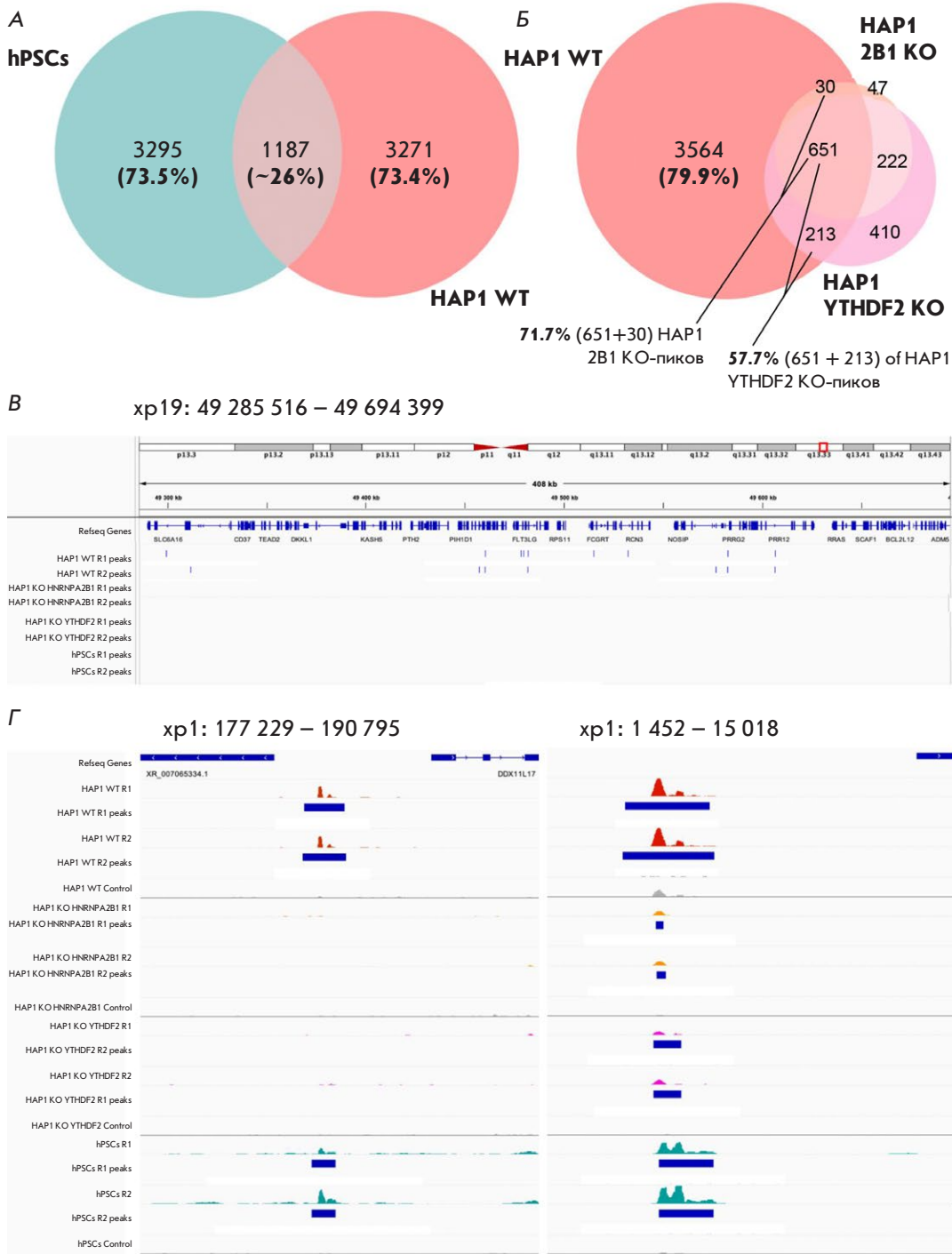


Рис. 1. Тип клеток является основным фактором, определяющим распределение R-петель в геноме. Диаграммы Венна, демонстрирующие перекрывание наборов данных пиков R-петель, полученных с помощью DRIP в hPSC и клетках HAP1 WT (А), а также в клетках HAP1 WT и HAP1 с генетическим нокаутом генов *YTHDF2* и *HNRNPA2B1* (Б). На диаграммах указано количество пиков R-петель в каждой категории. В – распределение пиков R-петель (показаны синими вертикальными черточками) в наборах данных DRIP, полученных на указанных клеточных линиях в области вокруг гена домашнего хозяйства *RPL13A* и обозначенной Чединым и соавт. как «золотой стандарт» для сравнения наборов данных. Г – профили R-петель (синие прямоугольники) в наших наборах данных и в контрольных образцах в двух репрезентативных геномных локусах (иллюстрации получены с помощью геномного браузера)

HAP1 2B1 KO, и 57.7% (864) пиков в HAP1 YTHDF2 KO были идентичны пикам R-петель, обнаруженных в их родительской клеточной линии HAP1 WT (рис. 1Б). Интересно, что геномная область, расположенная вокруг гена домашнего хозяйства *RPL13A*, использованная Чединым и соавт. для сравнения наборов данных и названная «регионом золотого стан-

дарта», содержала только пики HAP1 дикого типа, но не пики R-петель, полученные на hPSC или двух других клеточных линиях с нокаутом, протестированных в нашем исследовании (рис. 1В). Нами идентифицированы другие локусы, где пики R-петель присутствовали во всех четырех протестированных клеточных линиях (рис. 1Г, левая панель). Однако,

по нашему мнению, обозначение каких-либо геномных регионов как областей «золотого стандарта» требует создания наборов данных на дополнительных клеточных линиях различного происхождения, поскольку геномное распределение R-петель, по-видимому, в высокой степени специфично для типа клеток. Таким образом, нами показано, что наборы R-петель, полученные в клеточных линиях разного происхождения с использованием одной и той же методологии, существенно различаются, при этом лишь около четверти пиков R-петель совпадают в hPSC и линией HAP1, происходящей из клеток хронического миелоидного лейкоза. В то же время распределение R-петель в клетках HAP1 с нокаутом генов *YTHDF2* и *HNRNPA2B1*, и в клетках HAP1 дикого типа, очевидно, более сходно между собой, чем с их распределением в hPSC. Учитывая, что в число недавно выявленных факторов, участвующих в регуляции R-петель, входят структура хроматина [10], позиционирование нуклеосом [11] и модификации РНК [9], которые значительно различаются в разных типах клеток, эти результаты не кажутся нам особенно неожиданными [12]. Необходимо отметить, что интерпретация наших результатов и сделанные из них выводы ограничены

выбранными моделями. Поэтому для более убедительного обобщения необходимо проанализировать наборы других дифференцированных и опухолевых линий с различным гистогенезом.

ЗАКЛЮЧЕНИЕ

Подводя итог, мы приходим к выводу о клеточной специфичности геномного распределения R-петель. Следовательно, только систематическое и стандартизированное сравнение большого массива наборов данных о распределении R-петель в геноме различных типов клеток может устранить несоответствия между экспериментальными результатами, полученными с помощью различных методов картирования R-петель. ●

Работа финансировалась в рамках проекта Российского научного фонда (грант № 22-65-00022).

Работа К.Ю. Олейниковой по анализу набора данных hPSC человека была поддержана Министерством науки и высшего образования РФ. Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

- Aguilera P, Aguilera A. R-loop homeostasis in genome dynamics, gene expression and development. *Current Opinion in Genetics & Development*. 2025;92:102325. doi: 10.1016/j.gde.2025.102325
- Yadav C, Yadav R, Nanda S, Ranga S, Ahuja P. The hidden architects of the genome: a comprehensive review of R-loops. *Molecular Biology Reports*. 2024;51(1):1095.
- Boguslawski SJ, Smith DE, Michalak MA, Mickelson KE, Yehle CO, Patterson WL, Carrico RJ. Characterization of monoclonal antibody to DNA:RNA and its application to immunodetection of hybrids. *J Immunol Methods*. 1986;89:123-130. doi: 10.1016/0022-1759(86)90040-2
- García-Rubio M, Soler-Oliva ME, Aguilera A. Genome-Wide Analysis of DNA–RNA Hybrids in Yeast by DRIPc-Seq and DRIP-Seq. In *R-Loops: Methods and Protocols* New York, NY: Springer US; 2022:429-443.
- Wang H, Li C, Liang K. Genome-wide native R-loop profiling by R-loop cleavage under targets and tagmentation (R-Loop CUT&Tag). In *R-Loops: Methods and Protocols* New York, NY: Springer US; 2022:345-357.
- Malig M, Hartono SR, Giafaglione JM, Sanz LA, Chedin F. Ultra-deep coverage single-molecule R-loop footprinting reveals principles of R-loop formation. *Journal of Molecular Biology*. 2020;432(7):2271-2288. doi: 10.1016/j.jmb.2020.02.014
- Cerritelli SM, Sakhuja K, Crouch RJ. RNase H1, the gold standard for R-loop detection. *R-Loops: Methods and Protocols*. 2022:91-114. doi: 10.1007/978-1-0716-2477-7_7
- Chédin F, Hartono SR, Sanz LA, Vanoosthuysse V. Best practices for the visualization, mapping, and manipulation of R-loops. *EMBO Journal*. 2021;40(4):e106394. doi: 10.15252/embj.2020106394
- Abakir A, Giles TC, Cristini A, Foster JM, Dai N, Starczak M, et al. N6-methyladenosine regulates the stability of RNA: DNA hybrids in human cells. *Nature Genetics*. 2020;52(1):48-55. doi: 10.1038/s41588-019-0549-x
- Bayona-Feliu A, Barroso S, Muñoz S, Aguilera A. The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription–replication conflicts. *Nature Genetics*. 2021;53(7):1050-1063. doi: 10.1038/s41588-021-00867-2
- Werner M, Trauner M, Schauer T, Ummethum H, Márquez-Gómez E, Lalonde M, et al. Transcription-replication conflicts drive R-loop-dependent nucleosome eviction and require DOT1L activity for transcription recovery. *Nucleic Acids Research*. 2025;53(4):gkaf109. doi: 10.1093/nar/gkaf109
- Carette JE, Guimaraes CP, Varadarajan M, Park AS, Wuehrlich I, Godarova A, et al. Haploid genetic screens in human cells identify host factors used by pathogens. *Science*. 2009;326(5957):1231-1235. doi: 10.1126/science.1178955