

РНК-хроматиновые взаимодействия. Анализ данных*

Г.К. Рябых, А.А. Жарикова, И.С. Ильницкий, А.А. Миронов

В статье рассмотрены подходы к исследованию данных по РНК-хроматиновым взаимодействиям. Описаны источники данных и разработанные нами методы и протокол обработки данных. Описана также база данных RNA-Chrom (<https://rnachrom2.bioinf.fbb.msu.ru>). Рассмотрены некоторые особенности данных и представлены планы на будущее.

Ключевые слова: РНК, хроматин, РНК-хроматиновые взаимодействия.

*Работа выполнена при финансовой поддержке РФФИ (проект №20-04-00459).

Введение

С момента возникновения молекулярной биологии рибонуклеиновым кислотам отводилась вспомогательная роль переносчика информации от ДНК к белку. Выделялись три типа РНК: матричная (информационная) РНК – собственно переносчик информации – и два типа РНК, необходимые для синтеза белка, – транспортная РНК и рибосомальная РНК. Однако вскоре выяснилось, что роль РНК в клетке гораздо более разнообразна. РНК может выполнять ферментативную роль, катализируя некоторые биохимические реакции, структурную роль, участвуя в формировании различных структур в клетке, регуляторную роль, участвуя в регуляции экспрессии генов. Например, длинная некодирующая РНК XIST ответственна за инактивацию одной из копий X-хромосом у самок млекопитающих. Некодирующие РНК MALAT1 и NEAT1 участвуют в формировании структур в ядрах клеток – спеклов и параспеклов. Малые ядерные РНК необходимы для сплайсинга,

малые ядрышковые РНК участвуют в созревании рибосомных РНК. Прорыв, связанный с секвенированием генома человека и других организмов и анализом того, какие РНК транскрибируются в клетках, показал, что есть тысячи РНК, про которые никто до этого не подозревал, – некодирующие РНК. Биологическая роль этих РНК оставалась в большинстве случаев до сих пор неясна. Кроме того, известно, что многие регуляторные белки, такие как белки, участвующие в регуляции экспрессии генов, в частности в развитии организмов, взаимодействуют с РНК. Например, белки репрессии системы Polycomb, архитектурный белок CTCF, белки, метилирующие ДНК, имеют РНК-связывающий домен и выполняют свою функцию только в присутствии РНК. Для некоторых РНК показана их ассоциация с раковой трансформацией клеток.

В связи с этим возникла задача – найти на хроматине места, с которыми взаимодействуют те или иные РНК, быть может, в комплексе с белками. Для решения этих задач был разработан ряд экспериментальных подходов, которые можно разделить условно на два типа. Эксперименты первого типа направлены на построение карт контактов с хроматином для конкретных, заранее известных РНК [1–6], которые в дальнейшем мы будем называть «один-против-всех». Основная идея этих подходов заключается в следующем. Синтезируются меченные биотином олигонуклеотиды, комплементарные к целевой РНК. Исследуемые клетки фиксируются формальдегидом или другим сшива-



РЯБЫХ
Григорий Кириллович
Институт проблем передачи информации РАН



ЖАРИКОВА
Анастасия Александровна
Московский государственный университет им. М.В. Ломоносова



ИЛЬНИЦКИЙ
Иван Сергеевич
Институт проблем передачи информации РАН



МИРОНОВ
Андрей Александрович
Московский государственный университет им. М.В. Ломоносова

ющим агентом. Далее ДНК фрагментируется, с помощью зондов отбираются комплексы РНК–ДНК, фрагменты ДНК секвенируются. В результате получают те участки ДНК, с которыми взаимодействует целевая РНК, – карта контактов.

Другой тип экспериментов – «все-против-всех» – ищет все контакты всех РНК [7–12]. В этих экспериментах тоже сначала клетки фиксируются и ДНК фрагментируется. Далее добавляется специальный синтетический линкер, с которым сшиваются РНК и ДНК, оказавшиеся в одном комплексе. Далее химеры ДНК–линкер–РНК секвенируются. В результате получается большая карта контактов для многих РНК (см. обзор [13]).

Настоящая работа посвящена компьютерному анализу карт контактов, полученных в разных экспериментах.

Сбор и подготовка данных

Были отобраны все доступные данные, полученные в экспериментах по поиску контактов РНК с хроматином. Всего получен 191 набор данных – 20 наборов данных типа «все-против-всех» и 171 набор данных типа «один-против-всех» для человека и мыши. В каждой из работ, опубликованных по этой проблеме, предлагался свой протокол обработки данных. Для того чтобы обработанные данные можно было сравнивать, мы обработали их единообразно своим протоколом. Обратим внимание на то, что данные типа «один-против-всех» содержат только прочтения фрагментов ДНК, в то время как при секвенировании в экспериментах «все-против-всех» получают контактные пары, содержащие как ДНК-, так и РНК-последовательности (в дальнейшем ДНК- и РНК-части).

Единый протокол обработки данных заключался в следующем. Из данных были удалены ПЦР-дубликаты с помощью программы FastUniq, затем данные были обрезаны по качеству прочтений с помощью программы Trimimatic. Далее данные были картированы на референсные геномы – на геном человека hg38 и на геном мыши mm10 с помощью программы HISAT2, при этом сохранялись только уникальные картирования. Для экспериментов типа «все-против-всех» была проведена аннотация РНК-частей генами.

Для экспериментов типа «все-против-всех» важным этапом является аннотация РНК-частей. Для аннотации использовалась разметка генов GENCODE (v35 для человека и M35 для мыши) и некоторые другие источники. При аннотации иногда возникают коллизии – один и тот же локус принадлежит двум или более генам. Это возникает иногда потому, что в интронах некоторых генов закодированы микроРНК или другие некодирующие РНК. Для того чтобы однозначно определить, к какому гену следует отнести ту или

иную РНК-часть в таблице контактов, мы применили процедуру голосования. Выбирается тот ген, у которого плотность РНК-частей наибольшая. В результате мы получили однозначное назначение контакта к гену.

При аннотации была проведена проверка ориентации РНК-частей. Для этого мы отобрали около 50 высокоэкспрессируемых генов рибосомных белков и проверили коллинеарность РНК-частей аннотациям этих генов. Оказалось, что в разных экспериментах представлены разные ориентации РНК-частей. В одной группе экспериментов MARGI (mapping RNA-genome interactions) РНК генов рибосомных белков представлены в обеих ориентациях в равной степени, поэтому выбрать общую ориентацию РНК-частей не представляется возможным. Эти данные были исключены из дальнейшего рассмотрения. Осталось 187 наборов данных: 16 для данных «все-против-всех» и 171 для данных «один-против-всех». Далее все реплики для каждого эксперимента были объединены. В результате обработки данных для каждого эксперимента получены карты контактов РНК–ДНК.

Отметим важную особенность протокола обработки – значительное уменьшение количества прочтений. Например, в одном из экспериментов GRID (GSM2396700) на входе было 157 029 755 прочтений, а после всех этапов обработки осталось лишь 39 180 118 контактов, то есть только 25% от всех исходных данных. В экспериментах типа «один-против-всех» ситуация намного лучше – остается от 20% до 70% контактов. Основная доля потерь данных связана с картированием, а точнее, с требованием уникальности картирования данных на геном. В данных типа «все-против-всех» по дизайну эксперимента фрагменты ДНК- и РНК-частей значительно короче, поэтому программе картирования реже удастся найти уникальные прочтения.

Экспериментальные данные имеют ряд смещений. Один из возможных источников смещения данных

связан с доступностью хроматина. Для плотно упакованного хроматина мы видим пониженную плотность контактов, и это связано не столько с тем, что с плотно упакованным хроматином меньше реальных контактов, но и с тем, что плотно упакованный хроматин хуже поддается фрагментации. Для учета этого рода смещений проводят нормализацию. В данных типа «один-против-всех» используется так называемый «инпут». Для получения этих данных проводят точно такой же эксперимент, но без отбора фрагментов РНК с помощью зондов. В результате этого эксперимента получается набор прочтений, плотность которых характеризует ожидаемую величину смещения данных. Для нормализации данных в экспериментах «все-против-всех» используют внутреннюю нормализацию. При этой нормализации предполагается, что белок-кодирующие матричные РНК не имеют специфических контактов с хроматином. Поэтому вместо инпута строится трек контактов мРНК с хроматином. Чтобы избежать полимеразного следа, рассматриваются только контакты с хромосомами, на которых нет гена соответствующей мРНК. Здесь мы исключаем из рассмотрения контакты 50 самых высококонтактируемых и 1000 самых низкоконтактируемых мРНК. В результате мы получаем трек контактов всех отобранных мРНК с «чужими» хромосомами, который играет роль инпута. Имея треки инпута, можно разделить наблюдаемые плотности контактов на ожидаемые плотности контактов из инпута и получить нормированные контакты.

В экспериментах типа «все-против-всех» не все РНК-части принадлежали известным аннотированным генам, некоторое количество РНК-прочтений оказывались вне известных генов. Мы провели кластеризацию этих РНК-частей и нашли несколько тысяч новых, ранее неизвестных, потенциальных некодирующих РНК, которые в дальнейшем мы будем называть «Х-РНК».

Для экспериментов типа «один-против-всех» мы искали области с повышенной плотностью контактов (пики) с помощью программы MACS2. Для экспериментов типа «все-против-всех» мы пики не искали, поскольку в этих данных для каждой конкретной РНК, за редким исключением, число контактов невелико.

База данных RNA-Chrom

Все обработанные данные были собраны в аналитическую базу данных RNA-Chrom с веб-интерфейсом (<https://rnachrom2.bioinf.fbb.msu.ru>) [14]. Наши данные содержат более миллиарда контактов. Поэтому наша база данных основана на системе управления базами данных ClickHouse, которая позволяет эффективно работать с большими объемами данных. Наша база данных содержит, наряду с контактами РНК с хроматином, также метаданные экспериментов, в том числе: тип клеток; способ обработки клеток; экспериментальный протокол; ссылку на статью; ссылку на источник данных; процесс обработки данных и количество оставшихся прочтений на каждом этапе обработки.

Веб-интерфейс ориентирован на два типа запросов. Первый тип – получить все контакты во всех экспериментах для данной РНК (запрос «от РНК»). Пользователь либо задает имя РНК, которая его интересует, либо выбирает РНК из списка. В результате он получает список контактов, полученных в разных экспериментах. Здесь можно, во-первых, посмотреть графическое представление распределения контактов по хромосомам и распределение плотности контактов в зависимости от расстояния от гена, кодирующего рассматриваемую РНК. Эксперименты типа «все-против-всех» позволяют также проанализировать, какие области самой РНК участвуют в контактах.

Наш интерфейс предоставляет также возможность отправить контакты в геномный браузер (UCSC Genome Browser) для более детального визуального анализа и сравнения с другими треками, например, с треками эпигеномных разметок. Наконец, можно посмотреть, с какими генами контактирует выбранная РНК в разных экспериментах. При этом можно выбрать не только тело гена, но также и окружающие области.

Например, если мы хотим проанализировать контакты РНК Meg3, то мы увидим из графического представления, что практически во всех экспериментах типа «все-против-всех» эта РНК преимущественно контактирует с локусом ДНК, где она закодирована. А в эксперименте «один-против-всех», где исследовалась именно эта РНК, контакты достаточно равномерно распределены по хромосомам. Далее, если мы посмотрим гены-мишени для эксперимента iMARGI (*id*=6), то увидим, что эта РНК преимущественно контактирует с генами разнообразных неко-

дирующих РНК. Кликнув на идентификатор эксперимента, можно получить полную информацию об эксперименте, количестве контактов и прочем.

Другой вид запроса – найти все РНК, которые контактируют с заданным локусом ДНК (запрос «от ДНК»). При этом не обязательно указывать координаты. Достаточно задать имя гена. Например, задаем ген TP53 и его регуляторную область – 50 тыс. нуклеотидов перед ним. Сразу получаем список РНК, контакты которых были найдены в разных экспериментах. Для каждой из показанных РНК можно посмотреть, в каких экспериментах наблюдались контакты с интересующим локусом, и получить всю информацию об этой РНК так же, как это делалось при анализе «от РНК».

Разработанная нами база данных предоставляет широкие возможности для визуального анализа данных по РНК-хроматиновым контактам.

Некоторые особенности данных РНК-хроматиновых контактов

Отметим некоторые особенности данных. Во-первых, во многих работах по экспериментам типа «все-против-всех» отмечалось, что чем выше экспрессия гена той или иной РНК, тем больше контактов она имеет. Поскольку уровень экспрессии белок-кодирующих мРНК заметно выше уровня экспрессии некодирующих РНК (за редким исключением), то и контактов они имеют значительно больше.

Поскольку есть такой тренд, мы ввели понятие «хроматинового потенциала». Мы его строим следующим образом. Для белок-кодирующих мРНК строим график распределения отношения числа контактов к уровню экспрессии, полученному в эксперименте по тотальному РНК-секвенированию без поли(А)-обогащения. Мы при этом исключаем 10% самых высокоэкспрессируемых и 25% самых низкоэкспрессируемых генов. Для этого распределения находим среднее и среднеквадратичное отклонения. Далее, опираясь на эти числа, для каждой РНК мы можем вычислить *z*-скор. Эта величина характеризует родство РНК к хроматину, которую мы назовем хроматиновым потенциалом. Если поставить порог на величину хроматинового потенциала, то мы сможем отобрать те РНК, которые имеют большую склонность контактировать с хроматином. Например, если отобрать РНК с хроматиновым потенциалом больше 1, то доля мРНК среди отобранных РНК сократится почти вдвое. К сожалению, применение такого подхода имеет ограничения. Не все эксперименты «все-против-всех» сопровождаются данными по РНК-секвенированию в тех же условиях и на тех же клеточных линиях, для которых были получены контакты.

Второй тренд, который наблюдается практически во всех данных, как типа «все-против-всех», так и типа «один-против-всех», – это увеличение плотности ДНК-контактов поблизости от гена, с которого происходит транскрипция РНК. Эта зависимость может быть описана степенным законом, и, следуя терминологии Hi-C, мы называем эту зависимость «скейлингом». В нашей базе для РНК в том числе представлены графики скейлинга. Наличие скейлинга говорит нам о том, что значительная доля контактов не является специфической. Это явление можно объяснить следующим образом. Молекула РНК отрицательно заряжена. Хроматин, кроме отрицательно заряженной ДНК, содержит большое количество положительно заряженных белков – гистонов. Причем N-концы этих белков неструктурированы и содержат большое количество положительно заряженных аминокислот – лизина и аргинина. Если мы примем во внимание способ фиксации формальдегидом или другими агентами, которые «сшивают» по аминокислотным группам, то мы увидим большое количество зафиксированных неспецифических контактов. Возможных точек неспецифических контактов на хроматине – миллионы. Если специфические (регуляторные) контакты имеют аффинность, даже в 100 раз превышающую неспецифические контакты, мы без дополнительного анализа данных их все равно не заметим.

Заключение

В данной работе мы описали данные по РНК-хроматиновым взаимодействиям. Мы разработали протокол компьютерной обработки данных, который из «сырых» данных строит таблицы контактов. Были описаны некоторые не очевидные заранее особенности данных и способов их обработки. Анализ данных показал, что данные содержат очень высокий уровень шума. Дальнейшая работа с этими данными предполагает более тщательный анализ и фильтрацию

шума, в частности, необходимо разработать методы поиска скоплений контактов (пиков) в данных типа «все-против-всех», провести детальное сравнение экспериментов с учетом особенности данных. А также привлечь другие (внешние) данные для сравнительного анализа. К сожалению, очень трудно найти внешние данные, например, по РНК-секвенированию, поскольку требуется, чтобы эти данные соответствовали данным по РНК-хроматиновым контактам, то есть были получены для тех же клеток и в тех же условиях. Кроме того, они должны быть по тотальному секвенированию (с деплецией рРНК), а не поли(А)-секвенированию.

Дальнейшее развитие этого направления исследований предполагается вести в следующих направлениях. Во-первых, предполагается более подробно изучить явление скейлинга – увеличения плотности контактов с приближением к гену РНК. Во-вторых, предполагается разработать новый метод поиска скопления контактов на хроматине (пиков) с учетом скейлинга. Далее предполагается формализовать процедуру обработки данных в виде единого конвейера. С момента окончания работы по проекту было опубликовано заметное количество наборов данных. Предполагается обработать эти данные и внести их в базу данных RNA-Chrom. Наконец, предполагается разработать следующую версию базы данных и включить в нее пики контактов, ортологи РНК, функциональную аннотацию генов, а также интегрировать ее с другими ресурсами.

Литература

1. J.M. Engreitz, A. Pandya-Jones, P. McDonel, A. Shishkin, K. Sirokman, C. Surka, S. Kadri, J. Xing, A. Goren, E.S. Lander, K. Plath, M. Guttman
Science, 2013, **341**, 1. DOI: 10.1126/science.1237973.
2. M.D. Simon, C.I. Wang, P.V. Kharchenko, J.A. West, B.A. Chapman, A.A. Alekseyenko, M.L. Borowsky, M.I. Kuroda, R.E. Kingston
Proc. Natl. Acad. Sci., 2011, **108**, 20497.
DOI: 10.1073/pnas.1113536108.
3. C. Chu, K. Qu, F.L. Zhong, S.E. Artandi, H.Y. Chang
Mol. Cell, 2011, **44**(4), 667. DOI: 10.1016/j.molcel.2011.08.027.
4. J.J. Quinn, I.A. Ilik, K. Qu, P. Georgiev, C. Chu, A. Akhtar, H.Y. Chang
Nat. Biotechnol., 2014, **32**, 933. DOI: 10.1038/nbt.2943.
5. T. Mondal, S. Subhash, R. Vaid, S. Enroth, S. Uday, B. Reinius, S. Mitra, A. Mohammed, A.R. James, E. Hoberg, A. Moustakas, U. Gyllenstein, S.J.M. Jones, C.M. Gustafsson, A.H. Sims, F. Westerlund et al.
Nat. Commun., 2015, **6**, 1. DOI: 10.1038/ncomms8743.
6. H.P. Chu, C. Cifuentes-Rojas, B. Kesner, E. Aeby, H. Lee, C. Wei, H.J. Oh, M. Boukhali, W. Haas, J.T. Lee
Cell, 2017, **170**, 86. DOI: 10.1016/j.cell.2017.06.017.
7. B. Sridhar, M. Rivas-Astroza, T.C. Nguyen, W. Chen, Z. Yan, X. Cao, L. Hebert, S. Zhong
Curr. Biol., 2017, **27**, 602. DOI: 10.1016/j.cub.2017.01.011.
8. X. Li, B. Zhou, L. Chen, L.T. Gou, H. Li, X.D. Fu
Nat. Biotechnol., 2017, **35**, 940. DOI: 10.1038/nbt.3968.
9. J.C. Bell, D. Jukam, N.A. Teran, V.I. Risca, O.K. Smith, W.L. Johnson, J.M. Skotheim, W.J. Greenleaf, A.F. Straight
Elife, 2018, **7**, 1. DOI: 10.7554/eLife.27024.
10. Z. Yan, N. Huang, W. Wu, W. Chen, Y. Jiang, J. Chen, X. Huang, X. Wen, J. Xu, Q. Jin, K. Zhang, Z. Chen, S. Chien, S. Zhong
Proc. Natl. Acad. Sci., 2019, **116**, 3328.
DOI: 10.1073/pnas.1819788116.
11. A. Bonetti, F. Agostini, A.M. Suzuki, K. Hashimoto, G. Pascarella, J. Gimenez, L. Roos, A.J. Nash, M. Ghilotti, C.J.F. Cameron, M. Valentine, Y.A. Medvedeva, S. Noguchi, E. Agirre, K. Kashi, Samudyata
Nat. Commun., 2020, **11**, 1.
DOI: 10.1038/s41467-020-14337-6.
12. A.A. Gavrillov, A.A. Zharikova, A.A. Galitsyna, A.V. Luzhin, N.M. Rubanova, A.K. Golov, N.V. Petrova, M.D. Logacheva, O.L. Kantidze, S.V. Ulianov, M.D. Magnitov, A.A. Mironov, S.V. Razin
Nucl. Acids Res., 2020, **48**, 6699.
DOI: 10.1093/nar/gkaa457.
13. G.K. Ryabykh, D.E. Mylarshchikov, S.V. Kuznetsov, A.I. Sigorskikh, T.Y. Ponomareva, A.A. Zharikova, A.A. Mironov
Mol. Biol., 2022, **56**, 210.
DOI: 10.1134/S0026893322020121.
14. G.K. Ryabykh, S.V. Kuznetsov, Y.D. Korostelev, A.I. Sigorskikh, A.A. Zharikova, A.A. Mironov
Database, 2023, **2023**, baad025. DOI: 10.1093/database/baad025.

[illegible]

RNA-Chromatin Interactions. The Data Analysis*

Grigory K. Ryabykh

Institute for Information Transmission
Problems, RAS
19-1 Bolshoy Karetny Lane,
Moscow, 127051, Russia
ryabykhgrigory@gmail.com

Anastasiya A. Zharikova

Lomonosov Moscow State University
1-73 Leninskie Gory,
Moscow, 119991, Russia
azharikova89@gmail.com

Ivan S. Ilnitskiy

Institute for Information Transmission
Problems, RAS
19-1 Bolshoy Karetny Lane,
Moscow, 127051, Russia
nfsus96@gmail.com

Andrey A. Mironov

Professor,
Lomonosov Moscow State University
1-73 Leninskie Gory,
Moscow, 119991, Russia
mironov@bioinf.fbb.msu.ru

Abstract

This article reviews approaches to the study of data on RNA-chromatin interactions. Data sources and the data processing methods and protocol we developed are described. The RNA-Chrom database (<https://rnachrom2.bioinf.fbb.msu.ru>) is also described. Some features of the data are reviewed and future plans are presented.

Keywords: RNA, chromatin, RNA-chromatin interactions.

**The work was financially supported by RFBR (project 20-04-00459).*

References

1. J.M. Engreitz, A. Pandya-Jones, P. McDonel, A. Shishkin, K. Sirokman, C. Surka, S. Kadri, J. Xing, A. Goren, E.S. Lander, K. Plath, M. Guttman
Science, 2013, **341**, 1. DOI: 10.1126/science.1237973.
2. M.D. Simon, C.I. Wang, P.V. Kharchenko, J.A. West, B.A. Chapman, A.A. Alekseyenko, M.L. Borowsky, M.I. Kuroda, R.E. Kingston
Proc. Natl. Acad. Sci., 2011, **108**, 20497.
DOI: 10.1073/pnas.1113536108.
3. C. Chu, K. Qu, F.L. Zhong, S.E. Artandi, H.Y. Chang
Mol. Cell, 2011, **44**(4), 667. DOI: 10.1016/j.molcel.2011.08.027.
4. J.J. Quinn, I.A. Ilik, K. Qu, P. Georgiev, C. Chu, A. Akhtar, H.Y. Chang
Nat. Biotechnol., 2014, **32**, 933. DOI: 10.1038/nbt.2943.
5. T. Mondal, S. Subhash, R. Vaid, S. Enroth, S. Uday, B. Reinius, S. Mitra, A. Mohammed, A.R. James, E. Hoberg, A. Moustakas, U. Gyllenstein, S.J.M. Jones, C.M. Gustafsson, A.H. Sims, F. Westerlund et al.
Nat. Commun., 2015, **6**, 1. DOI: 10.1038/ncomms8743.
6. H.P. Chu, C. Cifuentes-Rojas, B. Kesner, E. Aebly, H. Lee, C. Wei, H.J. Oh, M. Boukhali, W. Haas, J.T. Lee
Cell, 2017, **170**, 86. DOI: 10.1016/j.cell.2017.06.017.
7. B. Sridhar, M. Rivas-Astroza, T.C. Nguyen, W. Chen, Z. Yan, X. Cao, L. Hebert, S. Zhong
Curr. Biol., 2017, **27**, 602. DOI: 10.1016/j.cub.2017.01.011.
8. X. Li, B. Zhou, L. Chen, L.T. Gou, H. Li, X.D. Fu
Nat. Biotechnol., 2017, **35**, 940. DOI: 10.1038/nbt.3968.
9. J.C. Bell, D. Jakam, N.A. Teran, V.I. Risca, O.K. Smith, W.L. Johnson, J.M. Skotheim, W.J. Greenleaf, A.F. Straight
Elife, 2018, **7**, 1.
DOI: 10.7554/eLife.27024.
10. Z. Yan, N. Huang, W. Wu, W. Chen, Y. Jiang, J. Chen, X. Huang, X. Wen, J. Xu, Q. Jin, K. Zhang, Z. Chen, S. Chien, S. Zhong
Proc. Natl. Acad. Sci., 2019, **116**, 3328.
DOI: 10.1073/pnas.1819788116.
11. A. Bonetti, F. Agostini, A.M. Suzuki, K. Hashimoto, G. Pascarella, J. Gimenez, L. Roos, A.J. Nash, M. Ghilotti, C.J.F. Cameron, M. Valentine, Y.A. Medvedeva, S. Noguchi, E. Agirre, K. Kashi, Samudiyata
Nat. Commun., 2020, **11**, 1.
DOI: 10.1038/s41467-020-14337-6.
12. A.A. Gavrilov, A.A. Zharikova, A.A. Galitsyna, A.V. Luzhin, N.M. Rubanova, A.K. Golov, N.V. Petrova, M.D. Logacheva, O.L. Kantidze, S.V. Ulianov, M.D. Magnitov, A.A. Mironov, S.V. Razin
Nucl. Acids Res., 2020, **48**, 6699.
DOI: 10.1093/nar/gkaa457.
13. G.K. Ryabykh, D.E. Mylarshchikov, S.V. Kuznetsov, A.I. Sigorskikh, T.Y. Ponomareva, A.A. Zharikova, A.A. Mironov
Mol. Biol., 2022, **56**, 210. DOI: 10.1134/S0026893322020121.
14. G.K. Ryabykh, S.V. Kuznetsov, Y.D. Korostelev, A.I. Sigorskikh, A.A. Zharikova, A.A. Mironov
Database. 2023. **2023**, baad025. DOI: 10.1093/database/baad025.