

Перевод на английский язык <https://vavilov.elpub.ru/jour>

## Контекстные сигналы в митохондриальных микроРНК млекопитающих

О.В. Вишнеvский<sup>1</sup>, П.С. Ворожейкин<sup>2</sup>✉, И.И. Титов<sup>1, 2, 3</sup>

<sup>1</sup> Федеральньй исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

<sup>2</sup> Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

<sup>3</sup> Курчатовский геномный центр ИЦиГ СО РАН, Новосибирск, Россия

✉ pavel.vorozheykin@gmail.com

**Аннотация.** МикроРНК – это малые некодирующие РНК, которые регулируют экспрессию генов на пост-транскрипционном уровне в цитоплазме, и, таким образом, играют важную роль в большом числе биологических процессов. Последние исследования обнаружили присутствие последовательностей микроРНК не только в цитоплазме, но и внутри митохондрий. Такие микроРНК (так называемые митомиры, mitomiRs) могут иметь ядерное или митохондриальное происхождение, при этом для некоторых из них установлена роль в регуляции функций митохондриальных генов, а для большинства – она пока неизвестна. Выявление нуклеотидных сигналов, уникальных для митомиров, может помочь определить эту роль. В нашей работе составлена выборка экспериментально обнаруженных митомиров человека, мыши и крысы. С целью выделения сигналов, которые могут быть ответственны за функционирование митомиров и за их транспортировку в митохондрии или из них, осуществлен контекстный анализ для полученных последовательностей митомиров. Для трех видов в группе данных митомиры/не-митомиры и в группе всех микроРНК из базы miRBase выявлены статистически перепредставленные 8-буквенные мотивы (уровень значимости  $p < 0.01$ , с учетом поправки Бонферрони на множественность сравнения). Для этих мотивов обнаружены закономерности их локализации в функционально значимых участках для разных типов микроРНК. Для рассматриваемой группы митомиры/не-митомиры также обнаружены статистически значимые особенности нуклеотидного состава последовательностей микроРНК возле границ разрезания комплексами Drosha/Dicer (критерий независимости  $\chi^2$  Пирсона для первых трех позиций микроРНК с уровнем значимости  $p < 0.05$ ). Наблюдаемые частоты нуклеотидов, предположительно, могут указывать на наличие у митомиров (в сравнении с не-митомирами) более однородного разрезания прай-миРНК комплексом Drosha при формировании 5'-конца последовательностей. Результаты работы могут быть полезными для выявления сигналов, принимающих участие в возникновении, процессинге и функциях митомиров. Ключевые слова: микроРНК; пре-миРНК; митомир; митохондрия.

**Для цитирования:** Вишнеvский О.В., Ворожейкин П.С., Титов И.И. Контекстные сигналы в митохондриальных микроРНК млекопитающих. *Вавиловский журнал генетики и селекции*. 2022;26(8):819-825. DOI 10.18699/VJGB-22-99

## The context signals of mitochondrial miRNAs (mitomiRs) of mammals

O.V. Vishnevsky<sup>1</sup>, P.S. Vorozheykin<sup>2</sup>✉, I.I. Titov<sup>1, 2, 3</sup>

<sup>1</sup> Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

<sup>2</sup> Novosibirsk State University, Novosibirsk, Russia

<sup>3</sup> Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia

✉ pavel.vorozheykin@gmail.com

**Abstract.** MicroRNAs (miRNAs) are small non-coding RNAs that regulate gene expression at the post-transcriptional level in the cytoplasm and play an important role in a wide range of biological processes. Recent studies have found that the miRNA sequences are presented not only in the cytoplasm, but also in the mitochondria. These miRNAs (the so-called mitomiRs) may be the sequences of nuclear or mitochondrial origin; some of them are involved in regulation of the mitochondrial gene functions, while the role of others is still unknown. The identification of nucleotide signals, which are unique to mitomiRs, may help to determine this role. We formed a dataset that combined the experimentally discovered mitomiRs in human, rat and mouse. To isolate signals that may be responsible for the mitomiRs' functions or for their translocation from or into mitochondria a context analysis was carried out for the sequences. For three species in the group mitomiRs/non-mitomiRs and the group of all miRNAs from the miRBase database statistically overrepresented 8-letter motifs were identified ( $p$ -value  $< 0.01$  with Bonferroni correction for multiple comparisons), for these motifs the patterns of the localization in functionally important regions for different types of miRNAs were found. Also, for the group mitomiRs/non-mitomiRs we found the statistically significant features of the miRNA nucleotide context near the Dicer and Drosha cleavage

sites (Pearson's  $\chi^2$  test of independence for the first three positions of the miRNA,  $p$ -value < 0.05). The observed nucleotide frequencies may indicate a more homogeneous pri-miRNA cleavage by the Drosha complex during the formation of the 5' end of mitomiRs. The obtained results can help to determine the role of the nucleotide signals in the origin, processing, and functions of the mitomiRs.

Key words: miRNA; pre-miRNA; mitomiR; mitochondrion.

**For citation:** Vishnevsky O.V., Vorozheykin P.S., Titov I.I. The context signals of mitochondrial miRNAs (mitomiRs) of mammals. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2022;26(8):819-825. DOI 10.18699/VJGB-22-99

## Введение

К настоящему времени достаточно хорошо исследованы основные пути биогенеза микроРНК, которые начинаются в ядре клетки и заканчиваются в цитоплазме (Bartel, 2018). При изучении нуклеотидного состава микроРНК и их предшественников (прайм-/пре-миРНК) установлено наличие сигналов, которые могут влиять на функции микроРНК и на процесс созревания микроРНК на отдельных стадиях биогенеза. Нуклеотидная последовательность микроРНК может как непосредственно определять функции микроРНК, так и оказывать влияние на точность разрезания комплексами процессинга Drosha/Dicer 5'-конца микроРНК и, таким образом, формировать сайт-специфически модифицированные микроРНК со сдвигом так называемого seed region (Starega-Roslan et al., 2015a, b; Rolle et al., 2016) (seed region – это участок со 2-го по 7-й нуклеотид микроРНК, который отвечает за ее адресацию).

Наличие мотивов в одноцепочечных концах (UG; CNNC) или в основании (CUC/GHG) прайм-миРНК, в терминальной петле (GU) шпильки пре-миРНК может приводить к блокированию или, наоборот, к содействию процессинга микроРНК (Auyeung et al., 2013; Fang, Bartel, 2015; Nguyen et al., 2015; Starega-Roslan et al., 2015a, b; Rolle et al., 2016; Vorozheykin, Titov, 2020). Помимо ядра и цитоплазмы, эти малые РНК-последовательности, а также белки комплексов для их процессинга обнаружены в органеллах, например в митохондриях (Kren et al., 2009; Bandiera et al., 2011; Wang et al., 2015). Эти наблюдения свидетельствуют как о вероятном существовании новых путей биогенеза микроРНК внутри митохондрий, так и о возможной транспортировке зрелых микроРНК между цитоплазмой и митохондриями пока неизвестными транспортными комплексами. Существование таких митохондриальных микроРНК (так называемых митомиров, mitomiRs) поднимает вопросы об их эволюционном происхождении и их функциях внутри и вне органелл. Присутствуют ли у них структурные особенности, обеспечивающие их функции и транспортировку внутрь или из митохондрий?

В настоящей статье собрана информация из опубликованных работ об экспериментально наблюдаемых микроРНК в митохондриях (митомиров). Для отобранных митомиров выполнена оценка контекстных особенностей последовательностей на предмет возможного влияния нуклеотидных сигналов на происхождение, процессинг и функции митомиров.

## Материалы и методы

В работе использованы последовательности микроРНК *Homo sapiens*, *Mus musculus*, *Rattus norvegicus* из базы данных miRBase (<http://miRBase.org>, выпуск 22.1) (Kozo-

mara et al., 2019). Общее количество микроРНК для трех рассматриваемых видов составляет 5398 последовательностей.

Информация о митомирах получена из статей с экспериментальным изучением (методами RT-qPCR, microarray, qRT-PCR) локализации микроРНК внутри и вне митохондрий для разных организмов и тканей (Kren et al., 2009; Bian et al., 2010; Bandiera et al., 2011; Barrey et al., 2011; Mercer et al., 2011; Das et al., 2012; Sripada et al., 2012; Wang et al., 2015). По этим работам для человека, мыши и крысы сформированы две выборки последовательностей: митомиры (652 последовательности микроРНК, наблюдаемые в митохондриях) и все остальные микроРНК из базы miRBase (мы будем называть их не-митомиры, 4766 последовательностей).

Для изучения особенностей последовательностей разных групп микроРНК выполнен поиск статистически перепредставленных (уровень значимости  $p < 0.01$ , с учетом поправки Бонферрони на множественность сравнения) олигонуклеотидных мотивов с помощью программы ARGO (Vishnevsky, Kolchanov, 2005) для *de novo* поиска мотивов в 15-буквенном коде для группы выборок митомиры/не-митомиры и группы всех микроРНК трех рассматриваемых видов из базы miRBase. При поиске мотивов для всех микроРНК из базы miRBase программа ARGO оценивает ожидаемую долю случайных последовательностей с мононуклеотидным частотным составом как у анализируемой выборки, содержащих мотив по случайным причинам.

Для полученных мотивов выполнена оценка совпадения мотивов в рамках каждой из рассмотренных групп и между двумя группами. Для каждой пары мотивов определяется коэффициент сходства (мера) Jaccard, в нашем случае  $\frac{N_{\text{совпад.}}}{N_{\text{общее}}}$ , где  $N_{\text{совпад.}}$  – число 4-буквенных нуклеотидных последовательностей, которые совпадают у двух мотивов;  $N_{\text{общее}}$  – число всех 4-буквенных нуклеотидных последовательностей, соответствующих обоим мотивам. Данная мера принимает значение от 0 до 1, где значение 0 демонстрирует полное различие, а значение 1 – полное совпадение двух мотивов.

Для оценки вероятности получения коэффициента Jaccard по случайным причинам использован метод, предложенный в работе (Real, Vargas, 1996), в которой случайное значение коэффициента Jaccard предполагается распределенным по биномиальному закону (с точностью до нормировки). Для выявленных мотивов определена локализация мотивов в последовательности микроРНК и анализ нуклеотидного состава с целью обнаружения неоднородности вырезания микроРНК из предшественника комплексами Drosha и Dicer. При анализе локализации

мотивов для группы всех микроРНК из базы miRBase в качестве «контрастной» выборки использованы случайные позиции, выбранные внутри последовательностей микроРНК изучаемой выборки.

### Результаты и обсуждение

При обработке публикаций с экспериментально наблюдаемыми микроРНК в митохондриях нами выявлено упоминание 652 уникальных идентификаторов микроРНК. При этом 272 последовательности из найденных микроРНК (митомиров) можно охарактеризовать как высокодостоверные, так как они или верифицированы дополнительно методами RT-qPCR/qRT-PCR, или в данных microarray-экспериментов наблюдаются в большем количестве внутри митохондрий, нежели вне их.

Отдельно нужно отметить семь митомиров, последовательности которых полностью присутствуют в митохондриальном геноме человека: hsa-miR-1974, hsa-miR-1977, hsa-miR-1978, hsa-miR-4461, hsa-miR-4463, hsa-miR-4484, hsa-miR-4485-3p, что может служить дополнительным подтверждением их достоверности. При этом ввиду наложения последовательности микроРНК на тРНК митохондрий к текущему моменту из базы miRBase удалены упоминания следующих микроРНК: hsa-miR-1974,

hsa-miR-1977, hsa-miR-1978. Из базы также удалена микроРНК hsa-miR-4461; экспериментально полученные для нее данные не удовлетворяют требованиям, предъявляемым при аннотации микроРНК. Таким образом, из базы miRBase были исключены последовательности, которые не соответствуют текущим известным путям биогенеза микроРНК, но могут формироваться неизвестными неканоническими путями.

Для последующего изучения и сравнения характеристик митомиров в работе также использована выборка не-митомиров общим количеством 4766 последовательностей. В нее вошли все микроРНК человека, мыши и крысы из рассматриваемой базы miRBase, за исключением митомиров.

С помощью программы ARGO (Vishnevsky, Kolchanov, 2005) проанализированы группы выборок микроРНК: группа из двух выборок митомиров/не-митомиров и группа всех микроРНК из базы miRBase. Для каждой из групп отобрано 40 (табл. 1) и 44 (табл. 2) 8-нуклеотидных мотивов, которые имеют статистически значимое различие встречаемости в выборках микроРНК в каждой из групп (уровень значимости  $p < 0.01$ , с учетом поправки Бонферрони на множественность сравнения). Для мотивов внутри каждой из групп, а также для мотивов из разных

**Таблица 1.** Мотивы, имеющие статистически значимое различие встречаемости в выборках последовательностей митомиров и не-митомиров

№ п/п	Мотив	% митомиров	% не-митомиров	Уровень значимости, $p$	№ п/п	Мотив	% митомиров	% не-митомиров	Уровень значимости, $p$
1	CAKTSNAN	8.44	0.87	$3.8 \times 10^{-26}$	21	MWCMBAVH	9.82	2.86	$9.7 \times 10^{-8}$
2	KTGCANDK	8.90	1.26	$5.6 \times 10^{-21}$	22	RKTGYWBH	11.81	3.98	$1.2 \times 10^{-7}$
3	HASHWSBD	28.53	11.89	$5.8 \times 10^{-21}$	23	GYHSHBDG	18.10	7.97	$1.7 \times 10^{-7}$
4	WMAGKGCD	6.29	0.54	$1.8 \times 10^{-20}$	24	YWCMCTBT	5.21	0.85	$3.8 \times 10^{-7}$
5	MNTVCANK	13.96	3.40	$3.6 \times 10^{-20}$	25	KKVAACMH	5.98	1.17	$8.9 \times 10^{-7}$
6	HRVRNTSH	34.97	17.14	$1.4 \times 10^{-18}$	26	CTNVRBTS	9.66	3.13	$2.1 \times 10^{-5}$
7	KBAGGTWG	5.21	0.41	$8.5 \times 10^{-17}$	27	CTRKNBVW	14.88	6.34	$2.4 \times 10^{-5}$
8	AGSAVCWY	5.21	0.41	$8.7 \times 10^{-17}$	28	RCABCMHN	6.13	1.40	$5.7 \times 10^{-5}$
9	RHASHWSB	20.86	7.93	$9.1 \times 10^{-16}$	29	YCMYWMHM	6.29	1.48	$7.5 \times 10^{-5}$
10	RCADTSDH	9.97	2.13	$7.9 \times 10^{-15}$	30	SAGVAMHN	8.13	2.45	$2.0 \times 10^{-4}$
11	RSTRDRT	8.13	1.44	$4.5 \times 10^{-14}$	31	WKMYCMKA	5.21	1.06	$2.1 \times 10^{-4}$
12	WMDWSCWB	15.49	5.08	$1.2 \times 10^{-13}$	32	NMYASDGS	10.43	3.79	$3.5 \times 10^{-4}$
13	HSVYDGDN	44.02	26.58	$1.9 \times 10^{-12}$	33	KGARNMCM	5.52	1.22	$4.0 \times 10^{-4}$
14	WRMACWTB	6.13	0.85	$3.5 \times 10^{-12}$	34	TSRGWSDG	5.98	1.46	$9.8 \times 10^{-4}$
15	CCHKBWGD	9.36	2.20	$2.3 \times 10^{-11}$	35	WCCHBTHS	6.60	1.77	$1.3 \times 10^{-3}$
16	GBYWYWG	12.12	3.63	$6.5 \times 10^{-11}$	36	SAVWSSCW	6.13	1.59	$2.9 \times 10^{-3}$
17	KGYNASW	10.74	3.03	$4.8 \times 10^{-10}$	37	STRHDGTT	5.06	1.11	$3.5 \times 10^{-3}$
18	CADKGNTD	8.13	1.79	$5.8 \times 10^{-10}$	38	NGGCWMDS	7.06	2.07	$4.1 \times 10^{-3}$
19	GWGSTNVY	9.66	2.60	$4.2 \times 10^{-9}$	39	HCBRRCT	5.37	1.26	$5.3 \times 10^{-3}$
20	WCAKSWR	6.44	1.24	$3.5 \times 10^{-8}$	40	YSTSRSTS	5.98	1.55	$5.8 \times 10^{-3}$

Примечание. % митомиров – доля последовательностей митомиров, содержащих мотив; % не-митомиров – доля последовательностей не-митомиров, содержащих мотив.

Здесь и в табл. 2: в качестве результата отобраны мотивы с уровнем значимости  $p < 0.01$  (с учетом поправки Бонферрони на множественность сравнения).

**Таблица 2.** Мотивы, имеющие статистически значимое различие встречаемости в выборке всех микроРНК базы miRBase в сравнении со случайными последовательностями с тем же мононуклеотидным составом, что и анализируемая выборка микроРНК

№ п/п	Мотив	% miRBase	% случайных	Уровень значимости, <i>p</i>	№ п/п	Мотив	% miRBase	% случайных	Уровень значимости, <i>p</i>
1	YNCKBYCB	12.09	5.36	$1.1 \times 10^{-71}$	23	HRHABYRC	5.63	2.99	$2.8 \times 10^{-15}$
2	BYNCYKYC	11.26	4.94	$2.0 \times 10^{-67}$	24	GCKSVKBK	6.28	3.46	$3.1 \times 10^{-15}$
3	AWRYRHWY	6.33	2.10	$1.3 \times 10^{-59}$	25	KTGYABDD	5.63	3.01	$1.0 \times 10^{-14}$
4	RHARHRHW	11.79	5.90	$1.0 \times 10^{-50}$	26	GTWDWHYV	5.15	2.76	$9.9 \times 10^{-13}$
5	NCKKYCBV	11.09	5.45	$1.2 \times 10^{-49}$	27	RHBTKTGH	5.94	3.36	$1.8 \times 10^{-12}$
6	WDYAYDKW	9.21	4.16	$2.1 \times 10^{-49}$	28	HWYVYAYR	6.11	3.49	$3.0 \times 10^{-12}$
7	RHAWWYRY	5.06	1.63	$4.2 \times 10^{-48}$	29	NRRMRSSA	8.40	5.29	$4.4 \times 10^{-12}$
8	VGGMDVNG	11.92	6.10	$4.8 \times 10^{-48}$	30	RVKGGMRV	7.88	4.88	$5.0 \times 10^{-12}$
9	YRTANANV	5.04	1.86	$3.9 \times 10^{-37}$	31	CBKCYCNV	5.76	3.33	$2.2 \times 10^{-10}$
10	NHYVCAG	9.47	4.81	$8.4 \times 10^{-37}$	32	KCCNBKBC	5.89	3.48	$2.0 \times 10^{-9}$
11	BTBYCYKY	9.21	4.66	$5.7 \times 10^{-36}$	33	HATHNYWY	5.70	3.35	$3.1 \times 10^{-9}$
12	YKHCTYYH	7.97	3.89	$2.5 \times 10^{-33}$	34	NKGWTDTH	5.06	2.89	$9.3 \times 10^{-9}$
13	ANBGHWDH	16.11	10.11	$4.4 \times 10^{-33}$	35	ASDHAVWW	5.37	3.15	$2.6 \times 10^{-8}$
14	CDGKVNNN	38.28	30.07	$7.4 \times 10^{-29}$	36	BCDGTKHY	5.30	3.11	$5.5 \times 10^{-8}$
15	RRMDGNAR	9.63	5.33	$5.0 \times 10^{-28}$	37	WGDRMHKG	8.99	6.10	$1.0 \times 10^{-7}$
16	GRGRHDGD	9.10	4.94	$6.7 \times 10^{-28}$	38	WWWTYRBD	5.41	3.27	$7.6 \times 10^{-7}$
17	DYAYDGTN	6.02	2.82	$6.2 \times 10^{-26}$	39	TBTMMYHY	5.30	3.30	$5.1 \times 10^{-5}$
18	WHAYAHNS	6.24	3.07	$2.0 \times 10^{-23}$	40	KSRGNBAG	6.31	4.12	$6.0 \times 10^{-5}$
19	NSDTNTHT	9.10	5.32	$1.5 \times 10^{-20}$	41	HMCMKYCH	5.44	3.50	$6.9 \times 10^{-4}$
20	TVYNYVCA	6.39	3.38	$1.3 \times 10^{-18}$	42	GWSGVDMN	7.88	5.54	$1.9 \times 10^{-3}$
21	DRYBTKTG	5.35	2.72	$1.0 \times 10^{-16}$	43	GWGHKBAB	5.08	3.26	$4.1 \times 10^{-3}$
22	TGBRRWKW	5.70	2.98	$1.3 \times 10^{-16}$	44	TWVTDWRH	5.19	3.37	$9.7 \times 10^{-3}$

Примечание. % miRBase – доля последовательностей базы miRBase, содержащих мотив; % случайных – оцененная программой ARGO доля случайных последовательностей, содержащих данный мотив по случайным причинам.

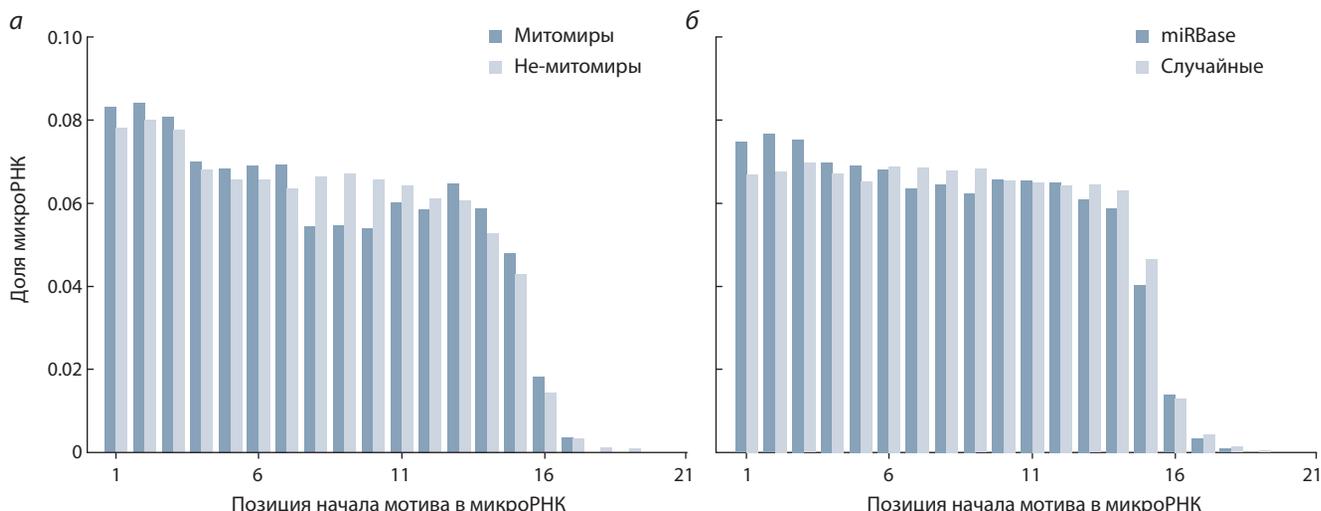
групп среднее значение коэффициента сходства Jaccard (исключая из расчета нулевые значения) для всех трех вычислений не превышает 0.02.

Для двух мотивов, KTGCANDK из группы митомиров/не-митомиров и KTGYABDD из группы всех микроРНК, отмечаются максимальный коэффициент (0.3) и минимальная вероятность наблюдения такого коэффициента по случайным причинам (0.81). Данные мотивы встречаются в 193 последовательностях из группы митомиров/не-митомиров и в 315 последовательностях группы всех микроРНК из базы miRBase. За исключением этих двух мотивов, коэффициент Jaccard для всех остальных рассматриваемых пар мотивов не превышает 0.13, вероятность наблюдения соответствующих коэффициентов по случайным причинам менее 0.001. Эти данные свидетельствуют о низкой степени совпадения мотивов как внутри каждой из групп, так и между двумя группами.

Такие различия нуклеотидного состава между выборками митомиров и не-митомиров могут выступать в качестве специфических сигналов для процессинга митомиров, например для распознавания и перемещения

последовательностей в/из митохондрий транспортными комплексами или для реализации специфических функций митомиров через прямое связывание с мишенями в митохондриальной или клеточной ДНК. В то же время мотивы, обнаруженные в выборке всех микроРНК, могут соответствовать сигналам, общим для процессинга и функций микроРНК, независимо от их локализации.

Для обеих рассматриваемых групп выборок первая позиция мотивов имеет тенденцию располагаться в начале микроРНК (рис. 1): максимальная доля последовательностей наблюдается для мотивов с началом в позициях 1–3 микроРНК. В полученных наблюдениях присутствует статистически значимая зависимость типа последовательностей микроРНК от позиций локализации начала мотива (критерий независимости  $\chi^2$  Пирсона, уровни значимости *p*:  $4.46 \times 10^{-2}$  и  $6.58 \times 10^{-5}$  соответственно для группы митомиров/не-митомиров и для группы все микроРНК из базы miRBase/случайные позиции в микроРНК). При этом для митомиров, в отличие от остальных выборок, наблюдается значительное сокращение числа микроРНК, в которых начало мотива располагается в по-



**Рис. 1.** Доля микроРНК в зависимости от позиций начала мотивов, встречающихся в последовательностях митомиров и не-митомиров (а) и во всех микроРНК из базы miRBase (б).

Для каждой микроРНК с мотивом определяется начальная с 5'-конца позиция мотива в последовательности. Если один мотив встречается в микроРНК несколько раз или несколько мотивов встречаются в одной микроРНК, каждый случай рассматривается независимо и порождает структуру данных (микроРНК, позиция мотива). График нормирован на общее количество полученных структур по всем мотивам. Спад наблюдений в позициях с номерами больше 15 обусловлен вариабельностью длин последовательностей микроРНК от 15 до 28 нуклеотидов.

зиях 8–10 микроРНК. Возможная причина в том, что для всех микроРНК (и митомиров, и не-митомиров) так называемый seed region является наиболее консервативным и значимым участком с функциональной точки зрения, поэтому рассматриваемые 8-буквенные консервативные мотивы часто «охватывают» эту область.

В противовес этому мотивы, начинающиеся с позиций 8–10 в микроРНК, располагаются в области так называемого дополнительного seed region (~13–16), который, предположительно, менее консервативен в митомирах и реже участвует в процессе связывания микроРНК с мишенью по сравнению с не-митомирами (см. рис. 1, а). Выборка всех микроРНК большей частью состоит из не-митомиров, поэтому результаты наблюдений для нее примерно совпадают с результатами наблюдений для не-митомиров.

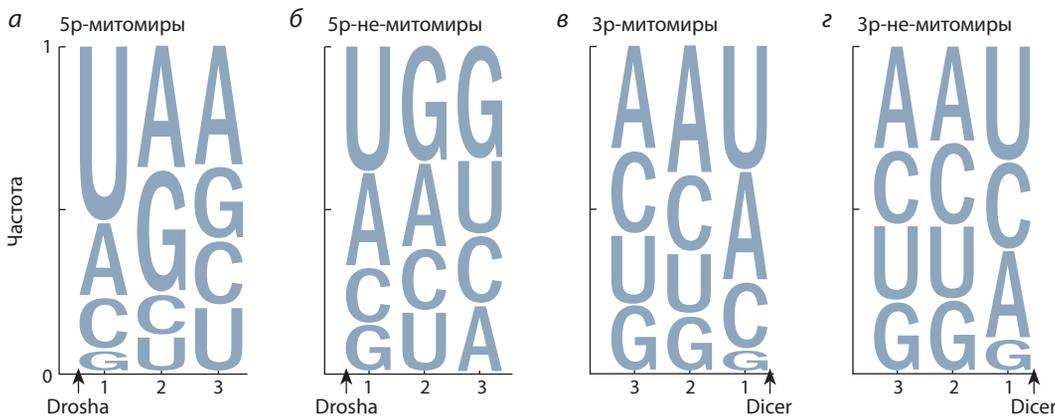
Для обнаруженных мотивов отмечаются разные паттерны локализации внутри последовательностей микроРНК. Один мотив может быть как в нескольких разных микроРНК с разным местоположением внутри последовательностей (например, мотив KTG CANDK с уровнем значимости  $p = 5.6 \times 10^{-21}$  встречается с позиции 14 с 5'-конца митомира hsa-miR-92a-1-5p и с позиции 2 с 5'-конца митомира mmi-miR-19b-3p), так и внутри одной микроРНК, в том числе не пересекаясь друг с другом (например, для митомира hsa-miR-33a-5p мотив KTG CANDK встречается дважды, начиная с позиции 1 и с позиции 12 с 5'-конца).

Таким образом, вариабельность локализации мотивов в микроРНК может свидетельствовать как о функциональной важности этих нуклеотидных сигналов для отдельных микроРНК, так и о возможном участии сигналов в процессинге микроРНК, в частности в отборе и перемещении последовательностей митомиров между митохондрией и цитоплазмой.

Для рассматриваемых групп микроРНК отмечается увеличение доли последовательностей, в которых мотивы начинаются в позициях 1–3 (см. рис. 1), поэтому далее мы осуществили контекстный анализ этого участка только для тех последовательностей митомиров и не-митомиров, в которых ранее были обнаружены мотивы. Для первых трех позиций 5'-конца 5p- и 3p-микроРНК вычислены попозиционные частоты встречаемости нуклеотидов (рис. 2).

Для митомиров из 5p-ветви пре-миРНК в первой позиции преимущественно наблюдается U и совсем редко – G, при этом в позиции 2 преимущественно встречается A или G (см. рис. 2, а). У не-митомиров отмечается увеличение в первой позиции числа G и A и уменьшение числа U (см. рис. 2, б). Для сайта разрезания Drosha наблюдается инверсия в позициях 2–3 между G в не-митомирах и A – в митомирах. В каждой из трех позиций частоты нуклеотидов демонстрируют зависимость от типа микроРНК (критерий независимости  $\chi^2$  Пирсона, уровни значимости  $p$ :  $2.89 \times 10^{-31}$ ,  $1.03 \times 10^{-28}$ ,  $1.79 \times 10^{-42}$  для позиций 1, 2 и 3 соответственно), при этом позиция 3 демонстрирует наиболее значимое различие частот между типами микроРНК, в отличие от позиций 1 и 2.

Сравнивая наблюдаемый нуклеотидный состав 5'-концов микроРНК для митомиров/не-митомиров с результатами исследования точности разрезания пре-миРНК комплексами Drosha и Dicer из работы (Starega-Roslan et al., 2015b), можно предположить, что для митомиров из 5p-ветви пре-миРНК разрезание Drosha осуществляется более точно, нежели для не-митомиров. То есть для митомиров формируются более однородный 5'-конец и соответствующий ему так называемый seed region, что может свидетельствовать о большей консервативности функций митомиров в сравнении с не-митомирами. Обнаруженные сигналы однородности разрезания для



**Рис. 2.** Частоты встречаемости нуклеотидов первых трех позиций начиная с 5'-конца микроРНК из 5p- и 3p-ветвей пре-миРНК для выборок митомиров (а, в) и не-митомиров (б, г), в последовательностях которых обнаружены мотивы.

Размеры букв пропорциональны частотам встречаемости нуклеотидов. По оси X отображаются номера позиций в микроРНК, отсчет от 5'-конца. Стрелками показаны области разрезания комплексами Dicer или Drosha. Для позиций 1–3 в 5p-микроРНК и для позиций 1–2 в 3p-микроРНК присутствует статистически значимая зависимость типа последовательностей микроРНК от частот встречаемости нуклеотидов в рассматриваемой позиции (критерий независимости  $\chi^2$  Пирсона, уровни значимости  $p$ :  $2.89 \times 10^{-31}$ ,  $1.03 \times 10^{-28}$ ,  $1.79 \times 10^{-42}$ ,  $1.17 \times 10^{-9}$ ,  $3.23 \times 10^{-10}$  для соответствующих позиций).

митомиров могут говорить как о возможном существовании в митохондриях более точного Drosha-подобного комплекса для процессинга микроРНК, так и о вероятном компенсировании неточности разрезания нуклеотидным составом последовательностей прай-миРНК, выбираемых для процессинга комплексом Drosha.

Для 5p-не-митомиров контекст смещается в сторону неоднородного разрезания (т. е. более активной сайт-специфической модификации микроРНК), и в таком случае не-митомиров могут выступать в качестве фактора функциональной изменчивости. Можно предположить, что митохондрии не «терпят» вариабельности «своих» микроРНК и в процессе эволюции устранили изоформы регуляторных последовательностей, а наблюдаемые митомиров могут являться оставшимися консервативными последовательностями, ведущими свое происхождение со времен доместикации предков митохондрий.

Для митомиров и не-митомиров из 3p-ветви пре-миРНК не наблюдается заметных отличий в нуклеотидном составе, за исключением инверсии в позиции 1 второго и третьего по популярности нуклеотидов (см. рис. 2, в и г). Статистически значимую зависимость позиционных нуклеотидных частот от типа микроРНК демонстрируют только позиции 1 и 2 (критерий независимости  $\chi^2$  Пирсона, уровни значимости  $p$ :  $1.17 \times 10^{-9}$  и  $3.23 \times 10^{-10}$  соответственно). Сравнение наблюдаемых частот нуклеотидов с результатами статьи (Starega-Roslan et al., 2015b) не дает однозначных выводов о качестве разрезания 5'-границы 3p-микроРНК комплексом Dicer.

### Заключение

В настоящей работе сформирована выборка экспериментально подтвержденных митомиров и произведен нуклеотидный анализ их последовательностей. Для группы данных митомиров/не-митомиров и для группы всех микроРНК из базы miRBase найдены статистически перепредставленные 8-буквенные мотивы внутри последо-

вательностей микроРНК. Эти мотивы демонстрируют, что последовательности митомиров могут представлять новый, неканонический класс микроРНК. Мотивы для группы митомиров/не-митомиров могут выступать в качестве сигналов для процессинга митомиров (например, участие при перемещении митомиров в/из митохондрий), для реализации функций митомиров (через связывание с мишенями в мтДНК или клеточной ДНК), в то время как мотивы группы всех микроРНК могут соответствовать сигналам, общим для процессинга и функций микроРНК, независимо от их локализации в клетке.

Нуклеотидный состав митомиров (в сравнении с не-митомирами) возле 5'-границ разрезания комплексами Drosha/Dicer, предположительно, может свидетельствовать о более однородном формировании 5'-конца последовательностей митомиров и, следовательно, о более консервативном функционале этих последовательностей.

### Список литературы / References

- Auyeung V.C., Ulitsky I., McGeary S.E., Bartel D.P. Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. *Cell*. 2013;152(4):844-858. DOI 10.1016/j.cell.2013.01.031.
- Bandiera S., Rüberg S., Girard M., Cagnard N., Hanein S., Chrétien D., Munnich A., Lyonnet S., Henrion-Caude A. Nuclear outsourcing of RNA interference components to human mitochondria. *PLoS One*. 2011;6(6):e20746. DOI 10.1371/journal.pone.0020746.
- Barrey E., Saint-Auret G., Bonnamy B., Damas D., Boyer O., Gidrol X. Pre-microRNA and mature microRNA in human mitochondria. *PLoS One*. 2011;6(5):e20220. DOI 10.1371/journal.pone.0020220.
- Bartel D.P. Metazoan microRNAs. *Cell*. 2018;173(1):20-51. DOI 10.1016/j.cell.2018.03.006.
- Bian Z., Li L.-M., Tang R., Hou D.-X., Chen X., Zhang C.-Y., Zen K. Identification of mouse liver mitochondria-associated miRNAs and their potential biological functions. *Cell Res*. 2010;20(9):1076-1078. DOI 10.1038/cr.2010.119.
- Das S., Ferlito M., Kent O.A., Fox-Talbot K., Wang R., Liu D., Raghavachari N., Yang Y., Wheelan S.J., Murphy E., Steenberg C.

- Nuclear miRNA regulates the mitochondrial genome in the heart. *Circ. Res.* 2012;110(12):1596-1603. DOI 10.1161/CIRCRESAHA.112.267732.
- Fang W., Bartel D.P. The menu of features that define primary microRNAs and enable *de novo* design of microRNA genes. *Mol. Cell.* 2015;60(1):131-145. DOI 10.1016/j.molcel.2015.08.015.
- Kozomara A., Birgaoanu M., Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* 2019;47(D1):D155-D162. DOI 10.1093/nar/gky1141.
- Kren B.T., Wong P.Y.-P., Sarver A., Zhang X., Zeng Y., Steer C.J. MicroRNAs identified in highly purified liver-derived mitochondria may play a role in apoptosis. *RNA Biol.* 2009;6(1):65-72. DOI 10.4161/rna.6.1.7534.
- Mercer T.R., Neph S., Dinger M.E., Crawford J., Smith M.A., Shearwood A.-M.J., Haugen E., Bracken C.P., Rackham O., Stamatoyannopoulos J.A., Filipovska A., Mattick J.S. The human mitochondrial transcriptome. *Cell.* 2011;146(4):645-658. DOI 10.1016/j.cell.2011.06.051.
- Nguyen T.A., Jo M.H., Choi Y.-G., Park J., Kwon S.C., Hohng S., Kim V.N., Woo J.-S. Functional anatomy of the human microprocessor. *Cell.* 2015;161(6):1374-1387. DOI 10.1016/j.cell.2015.05.010.
- Real R., Vargas J. M. The probabilistic basis of Jaccard's index of similarity. *Syst. Biol.* 1996;45(3):380-385. DOI 10.1093/sysbio/45.3.380.
- Rolle K., Piwecka M., Belter A., Wawrzyniak D., Jeleniewicz J., Barciszewska M.Z., Barciszewski J. The sequence and structure determine the function of mature human miRNAs. *PLoS One.* 2016; 11(3):e0151246. DOI 10.1371/journal.pone.0151246.
- Sripada L., Tomar D., Prajapati P., Singh R., Singh A.K., Singh R. Systematic analysis of small RNAs associated with human mitochondria by deep sequencing: detailed analysis of mitochondrial associated miRNA. *PLoS One.* 2012;7(9):e44873. DOI 10.1371/journal.pone.0044873.
- Starega-Roslan J., Galka-Marciniak P., Krzyzosiak W.J. Nucleotide sequence of miRNA precursor contributes to cleavage site selection by Dicer. *Nucleic Acids Res.* 2015a;43(22):10939-10951. DOI 10.1093/nar/gkv968.
- Starega-Roslan J., Witkos T., Galka-Marciniak P., Krzyzosiak W. Sequence features of Drosha and Dicer cleavage sites affect the complexity of isomiRs. *Int. J. Mol. Sci.* 2015b;16(12):8110-8127. DOI 10.3390/ijms16048110.
- Vishnevsky O.V., Kolchanov N.A. ARGO: a web system for the detection of degenerate motifs and large-scale recognition of eukaryotic promoters. *Nucleic Acids Res.* 2005;33(Web Server Iss.):W417-W422. DOI 10.1093/nar/gki459.
- Vorozheykin P.S., Titov I.I. Erratum to: How animal miRNAs structure influences their biogenesis. *Russ. J. Genet.* 2020;56(8):1012-1024. DOI 10.1134/S1022795420220019.
- Wang W.-X., Visavadiya N.P., Pandya J.D., Nelson P.T., Sullivan P.G., Springer J.E. Mitochondria-associated microRNAs in rat hippocampus following traumatic brain injury. *Exp. Neurol.* 2015;265:84-93. DOI 10.1016/j.expneurol.2014.12.018.

---

#### ORCID ID

I.I. Titov orcid.org/0000-0002-2691-3292

**Благодарности.** Работа ОВВ и ИИТ выполнена при финансовой поддержке бюджетного проекта № FWNR-2022-0020. Работа ИТ поддержана Курчатовским геномным центром ИЦиГ СО РАН (075-15-2019-1662).

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 07.09.2022. После доработки 17.11.2022. Принята к публикации 17.11.2022.