

УДК575

Распространенность сайтов связывания miRNA и lncRNA вокруг некоторых генов, ассоциированных с ожирением

Липович Леонард¹, Романов Дмитрий Евгеньевич², Лянгасова Ольга Владиславовна², Амелина Мария Александровна^{2,3}, Шкурат Михаил Алексеевич², Бутенко Елена Викторовна², Алаа Хашим Абд Али², Деревянчук Екатерина Григорьевна², Шкурат Татьяна Павловна²

¹*Center for Molecular Medicine & Genetics, Wayne State University, Detroit, MI, United States*

²*Южный федеральный университет, Ростов-на-Дону, Россия; tshkurat@yandex.ru*

³*Ростовский государственный медицинский университет, Ростов-на-Дону, Россия*

DOI: 10.18522/2308-9709-2020-32-5

Аннотация:

Изучена локализация сайтов связывания miRNA и lncRNA вокруг генов *FTO*, *APOC1*, *LIPC*, *LPL* и *APOE*. Показано, что наибольшее число сайтов связывания с miRNA и lncRNA находится в цис-регуляторных участках и интронах гена *FTO*. Представлен перечень мотивов miRNA и lncRNA.

Выдвинута гипотеза, что количество сайтов связывания miRNA и lncRNA в интронах и цис-регуляторных районах генов коррелирует с числом тканей, в которых экспрессируется ген, и с числом функций его белка.

Ключевые слова: miRNA, lncRNA, ожирение, *FTO*, has-miR-619, lnc-SNAPC3-5:1.

Prevalence of miRNA and lncRNA binding sites around obesity associated genes

DOI: 10.18522/2308-9709-2020-32-5

Lipovich Leonard¹, Romanov Dmitry E.², Lyangasova Olga V.², Amelina Maria A.^{2,3}, Shkurat Mikhail A.², Butenko Elena V.², Makarenko Maxim S.², Alaa Khashim Abd Ali², Derevyanchuk Ekaterina G.², Shkurat Tatyana P.²

¹*Center for Molecular Medicine & Genetics, Wayne State University, Detroit, MI, United States*

²*South Federal University, Rostov-on-Don, Russia; tshkurat@yandex.ru*

³*Rostov State Medical University, Rostov-on-Don, Russia*

Abstract:

The localization of miRNA and lncRNA binding sites around *APOC1*, *LIPC*, *LPL*, *APOE* and *FTO* genes was studied. It was shown that the largest number of miRNA and lncRNA binding sites is located in cis-regulatory regions and introns of *FTO* gene. The list of miRNA and lncRNA motifs is presented.

It was hypothesized that the number of miRNA and lncRNA binding sites in introns and cis-regulatory regions of genes correlates with the number of tissues in which the gene is expressed, and with the number of functions of its protein.

Key words: miRNA, lncRNA, obesity, *FTO*, has-miR-619, lnc-SNAPC3-5:1.

Введение

Исследования транскриптома человека показало, что большая половина транскриптома млекопитающих состоит из некодирующей РНК [11]. Некодирующие РНК делятся на короткие некодирующие РНК (включая miRNA) и длинные некодирующие РНК (lncRNA). Длинные некодирующие РНК (lncRNA), длиной в среднем около 200 нуклеотидов, представляют собой новый класс регуляторов, участвующих в различных биологических процессах [6, 8, 9]. Недавние исследования показали, что многие lncRNA играют ключевую роль в регуляции развития адипоцитов и ожирения [12]. MiRNA являются эволюционно консервативными некодирующими одноцепочечными РНК, длиной в среднем около 18–24 нуклеотидов, которые регулируют экспрессию генов на посттранскрипционном уровне путем типичного связывания с 3'-нетранслируемой областью (3'-UTR) специфической мРНК-мишени последовательности, что приводит к снижению экспрессии белка путем блокирования трансляции мРНК, и / или стимулирования деградации мРНК [4]. Более 60 % всех кодирующих белок генов напрямую регулируется микроРНК (miRNA). Одна микроРНК может связываться и регулировать более чем одну мишень, иногда как часть одного и того же сигнального пути. И наоборот, мРНК может содержать несколько различных сайтов связывания miRNA в своем 3'-UTR, добавляя несколько уровней регуляции. Таким образом, miRNA являются «точными настройками» паттернов экспрессии генов и служат важными регуляторами ряда клеточных эффектов и молекулярных сигнальных путей, участвующих, в том числе, и в ожирении. В последние годы показано, что многие miRNA кодируются в межгенных участках ДНК и обнаружены в интронах [2, 7]. Интронные miRNA (in-miRNA) локализованы в интронах белок-кодирующих генов и транскрибируются с ДНК автономно от транскрипции pre-mRNA при наличии собственного промотора, либо вырезаются из pre-mRNA в виде pre-miRNA. Открытия в течение последнего десятилетия показывают, что РНК работает не только как функциональный посланник между ДНК и белком, но также участвует в регуляции и организации генома, в регуляции экспрессии генов. И ее роль увеличивается с усложнением организма. Показана важная роль микроРНК в липидном и жировом обмене. Некоторые из микроРНК

рассматриваются как потенциальные чувствительные диагностические маркеры ожирения.

Целью нашей работы было изучение локализации сайтов связывания микроРНК с матричной РНК в цис-регуляторных районах генов *FTO*, *LIPC*, *LPL*, *APOC3*, *APOE*, ассоциированных с ожирением.

Материал и методы

Поиск мотивов гомологичных микроРНК осуществлялся в цис-регуляторных районах изучаемых генов с помощью биоинформационного пакета MEME Suite. Известные miRNA взяли согласно данным баз данных miRBase (<http://mirbase.org/>). Нуклеотидные последовательности цис-регуляторных районов и интронов генов взяты нами из электронных баз данных NCBI (<http://www.ncbi.nlm.nih.gov/>) с помощью наборов скрипта, IFITSN, которые были разработаны специально для автоматизированного получения информации от последовательностей NCBI. Под окрестностью гена понимается участок ДНК, содержащий как последовательность самого гена, так и фланкирующие участки ДНК до следующего гена или длиной 50000 н.п. до и после гена (таблица 1).

Таблица 1 – Число нуклеотидов (длина) в окрестностях некодирующей ДНК в изучаемых генах

| Общепринятое название | Длина окрестности гена (нуклеотиды) | Число нуклеотидов перед геном (от предыдущего) | Число нуклеотидов в гене | Число нуклеотидов после гена (до сл гена) |
|-----------------------|-------------------------------------|--|--------------------------|---|
| <i>FTO</i> | 517978 | 50000 | 417978 | 50000 |
| <i>APOC1</i> | 17409 | 4852 | 5103 | 7454 |
| <i>LIPC</i> | 345225 | 160774 | 158121 | 26330 |
| <i>LPL</i> | 261259 | 86993 | 28189 | 146077 |
| <i>APOE</i> | 10632 | 2093 | 3612 | 4927 |

Данные из базы данных «miRBase» анализировали при помощи программы «mscanner» для автоматического поиска сайтов связывания с использованием программы «mscanner» [10].

Итоги первого запроса отфильтровывали с получением только определенных последовательностей, в которых степень гомологии нуклеотидов была больше 90 %, как для пре-микроРНК, так и для зрелой микроРНК.

В работе использована программа dotolog [1], которая предназначена для построения точечной матрицы гомологий (дотплота) для двух и более последовательностей. Данная программа позволяет проводить множественное выравнивание последовательностей, полученных по результатам визуального анализа. Интерфейс программы реализует

концепцию «as Google Maps», предоставляя удобство ориентирования по последовательностям. При отображении широко используется цветовое кодирование, в частности, мотивы (непрерывные последовательности совпадений) подсвечиваются зеленым цветом. Пользовательские аннотации подсвечиваются цветом, сгенерированным по уникальному идентификатору элемента. Реализована возможность фильтрации отображаемых элементов. Таким образом, предоставляемый программой dotolog функционал позволяет эффективно выявлять исследуемые элементы генома.

Результаты исследования и их обсуждение

Результаты изучения локализации сайтов связывания микроРНК с матричной РНК в цис-регуляторных районах генов и в кодирующих последовательностях ДНК представлены в таблице 2. Анализ *in silico* показал, что в генах и межгенных пространствах исследуемых генов находится следующее количество мотивов, гомологичных зрелым микроРНК.

Таблица 2 – Количество сайтов зрелых микроРНК (*mat_hsa_mir*) вокруг генов липидного обмена

| Наименование гена | Всего | Перед геном | Внутри гена | После гена |
|-------------------|-------|-------------|-------------|------------|
| <i>APOC1</i> | 5 | 2 | 0 | 3 |
| <i>FTO</i> | 43 | 1 | 35 | 7 |
| <i>LPL</i> | 6 | 1 | 0 | 5 |
| <i>APOE</i> | 4 | 1 | 1 | 2 |
| <i>LIPC</i> | 113 | 58 | 27 | 28 |

На рисунках 1 и 2 показана плотность распределения сайтов связывания с мотивами зрелых miRNA в окрестности генов *FTO* и *LIPC*.

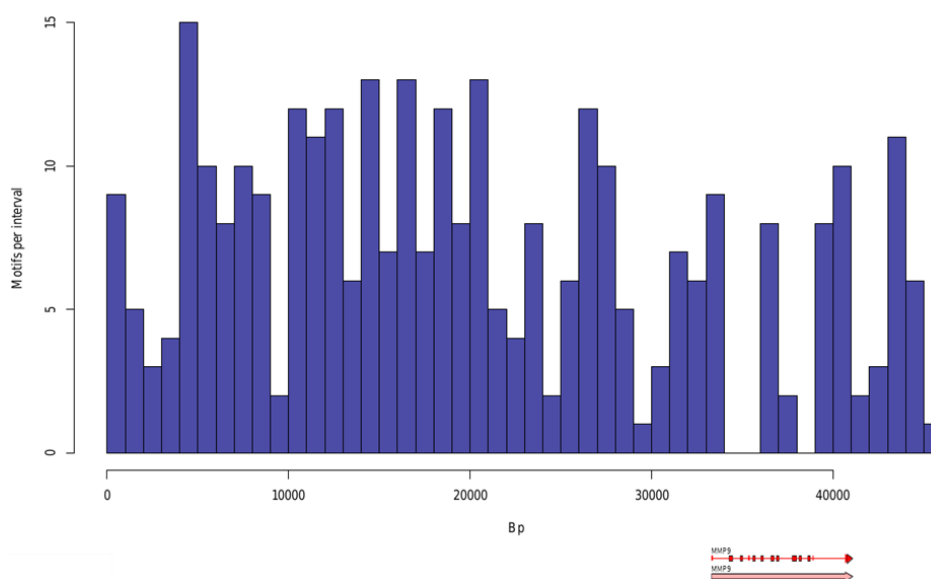


Рис. 1 – Плотность распределения сайтов связывания зрелых miRNA

в окрестности гена *FTO*

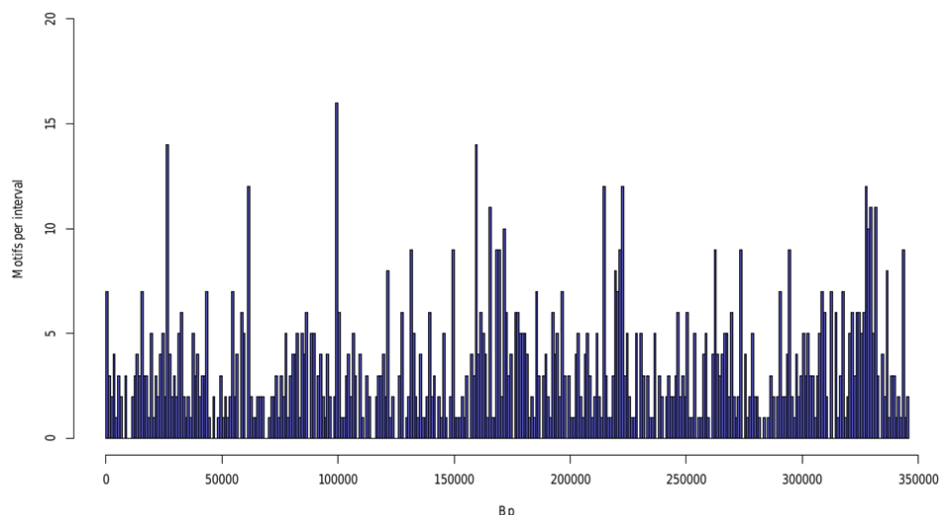


Рис. 2 – Плотность распределения сайтов связывания зрелых miRNA в окрестности гена *LIPC*

График плотности имеет волнообразный характер в цис-регуляторных районах генов и интронах, при этом повышенная плотность наблюдается в длинных интронах.

Согласно нашим данным биоинформационного анализа, в окрестностях гена *FTO* локализованы следующие мотивы: hsa-miR-619-5p (16), hsa-miR-5585-3p (2), hsa-miR-1285-3p (2), hsa-miR-3613-5p (1), hsa-miR-6788-5p (1), hsa-miR-3134 (1), hsa-miR-1256 (1), hsa-miR-320d (1), hsa-miR-6845-3p (1), hsa-miR-3133 (1), hsa-miR-649 (1), hsa-miR-1261 (1), hsa-miR-4496 (1), hsa-miR-4433b-5p (1), hsa-miR-297 (1), hsa-miR-183-5p (1), hsa-miR-3162-3p (1), hsa-miR-7110-5p (1), hsa-miR-4445-5p (1), hsa-miR-877-3p (1), hsa-miR-205-3p (1), hsa-miR-708-5p (1), hsa-miR-3140-5p (1), hsa-miR-2052 (1), hsa-miR-3661 (1), hsa-miR-1972 (1). В скобках указано количество сайтов.

Из них мотивы, перекрывающиеся с интронами в гене *FTO*: hsa-miR-619-5p (14), hsa-miR-7110-5p (1), hsa-miR-4445-5p (1), hsa-miR-6845-3p (1), hsa-miR-1261 (1), hsa-miR-3613-5p (1), hsa-miR-6788-5p (1), hsa-miR-1285-3p (1), hsa-miR-320d (1), hsa-miR-3140-5p (1), hsa-miR-5585-3p (1), hsa-miR-4433b-5p (1), hsa-miR-205-3p (1), hsa-miR-708-5p (1), hsa-miR-183-5p (1), hsa-miR-2052 (1), hsa-miR-4496 (1), hsa-miR-649 (1), hsa-miR-3661 (1), hsa-miR-3134 (1), hsa-miR-1972 (1), hsa-miR-1256 (1).

Как видно, наиболее часто в интронах гена *FTO* встречается hsa-miR-619. В частности, 16 сайтов локализовано вокруг гена *FTO*, из них 14 – внутри. Ранее hsa-miR-619 была названа уникальной miRNA (umiRNA), поскольку было обнаружено, что она способна связывать более 600, 1200 и 1000 генов соответственно [3]. Этот уникальный мотив микроРНК имеет сайты связывания в кодирующих последовательностях и нетранслируемых областях мРНК. Предполагается, что эти umiRNAs являются координаторами

экспрессии генов, которые участвуют во многих основных биологических процессах [7].

В таблице 3 представлены данные о количестве сайтов зрелых длинных РНК (lncRNA) вокруг исследуемых генов.

Таблица 3 – Количество сайтов зрелых длинных РНК (lncRNA) вокруг генов липидного обмена

| Наименование гена | Всего | Перед геном | Внутри гена | После гена |
|-------------------|-------|-------------|-------------|------------|
| <i>APOC1</i> | 9 | 0 | 0 | 9 |
| <i>FTO</i> | 133 | 30 | 86 | 17 |
| <i>LPL</i> | 89 | 89 | 0 | 0 |
| <i>APOE</i> | 1 | 0 | 1 | 0 |

Как видно из таблицы, максимальное количество сайтов связывания с длинными РНК регистрировалось в окрестностях гена *FTO* – это были следующие последовательности перед геном *FTO*: lnc-GPATCH4-2:1 (3), lnc-SMIM10L2A-2:1 (1), lnc-GGT5-3:1 (1), lnc-ZMYM1-3:1 (1), lnc-B3GALNT2-3:1 (1), lnc-CSE1L-4:1 (1), lnc-SYT15-5:1 (1), lnc-LIPI-20:1 (1), lnc-LIPI-21:1 (1), lnc-CSF1-4:1 (1), lnc-TEFM-4:1 (1), lnc-PTEN-3:1 (1), lnc-BCKDK-2:1 (1), lnc-HYAL2-1:1 (1), lnc-SNAPC3-5:1 (1), lnc-API5-12:1 (1), lnc-ATP6AP1L-5:1 (1), lnc-LHFPL6-9:1 (1), lnc-C1orf105-3:1 (1), lnc-C1orf105-4:1 (1), lnc-TEAD3-3:1 (1), lnc-MYL2-6:1 (1), lnc-RGS8-3:1 (1), lnc-ADGRD2-4:1 (1), lnc-CHD9-13:1 (1), lnc-HIST1H4J-1:1 (1), lnc-BBS5-6:1 (1), lnc-SNX33-4:1 (1).

В окрестностях гена *FTO* указаны мотивы, представленные более, чем в одной копии: lnc-SNAPC3-5:1 (4), lnc-SYT15-5:1 (3), lnc-LIPI-20:1 (3), lnc-LHFPL6-9:1 (3), lnc-C1orf105-4:1 (3), lnc-ADGRD2-4:1 (3), lnc-CSF1-4:1 (3), lnc-CHD9-13:1 (3), lnc-ZMYM1-3:1 (3), lnc-LIPI-21:1 (3), lnc-C1orf105-3:1 (3), lnc-GPATCH4-2:1 (3), lnc-PTEN-3:1 (2), lnc-PTEN-2:1 (2), lnc-CCL2-9:1 (2).

Из всех исследованных нами генов максимальную длину имел ген *FTO*, он также имел максимальное количество сайтов связывания miRNA и lncRNA как в цис-регуляторных районах, так и в интронах. С увеличением размера генома в ходе эволюции эукариот интроны демонстрируют тенденцию к увеличению в геноме, выполняя при этом регуляторную функции, как при сплайсинге, так и экспрессии генов в различных тканях. На рисунках 3 и 4 представлены результаты экспрессии гена *FTO*, полученных в проекте «HRA RNA-seq normal tissues» при изучении транскриптомов генов, кодирующих белки в 27 дифференцированных нормальных тканях от 95 индивидуумов [5]. Как видно из рисунков 3 и 4 ген *FTO* транскрибируется во всех изученных тканях, в то время как ген *APOC1* – только в четырех из 27.

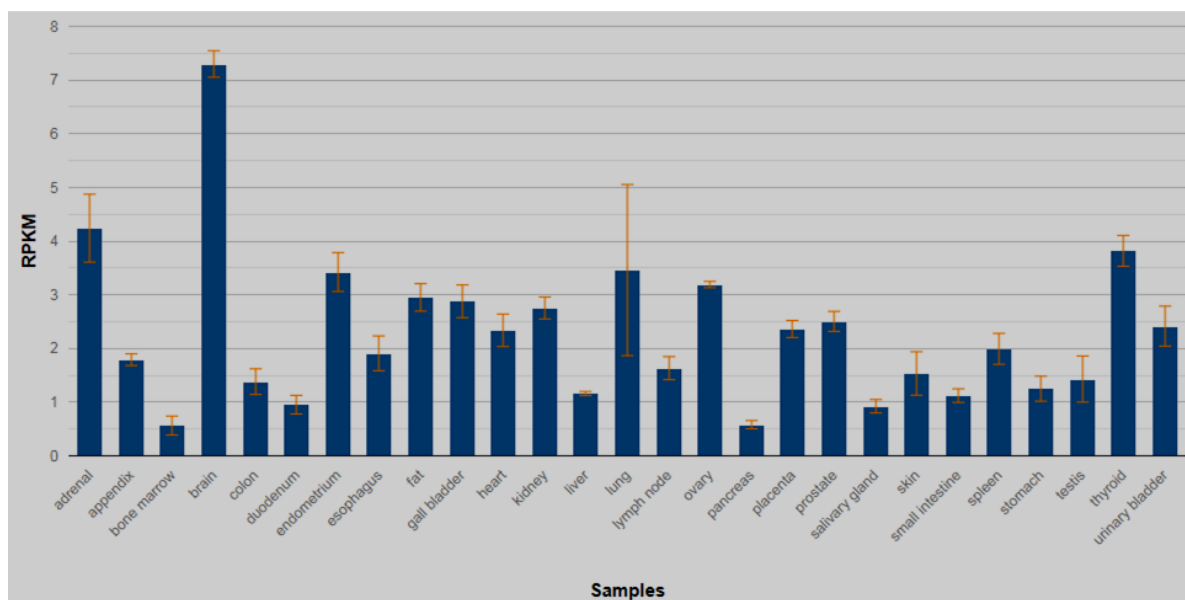


Рис. 3 – Результаты экспрессии гена *FTO* в различных тканях здоровых людей. Источник: NCBI. *FTO* [Homo sapiens (human)] Gene ID: 79068 <https://www.ncbi.nlm.nih.gov/gene/79068/?report=expression>

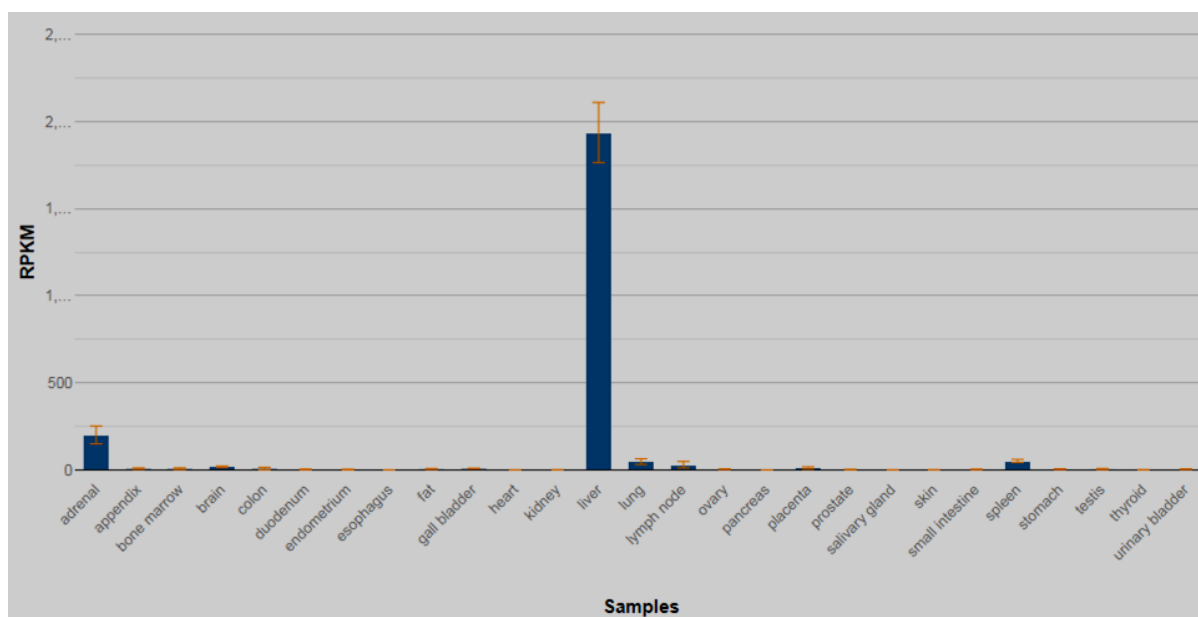


Рис. 4 – Результаты экспрессии гена *APOC1* в различных тканях здоровых людей. *APOC1* apolipoprotein C1 [Homo sapiens (human)] Источник: NCBI Gene ID: 341 <https://www.ncbi.nlm.nih.gov/gene/341/?report=expression>

Ген *FTO* осуществляет следующие функции:

- регулирование многоклеточного роста организма;
- процесс регуляции дыхательной системы;
- регуляция пролиферации жировых клеток;
- стимуляция клеточного ответа на повреждение ДНК;

- окислительное деметилирование;
- окислительное деметилирование одноцепочечной РНК;
- регулирование хранения липидов;
- РНК репарация;
- развитие жировой ткани;
- ДНК деметилирование;
- регулирование дифференциации клеток коричневого жира;
- гомеостаз;
- окислительно-восстановительный процесс;
- ДНК репарация;
- ДНК деалкилирование.

APOC1 очень короткий ген, и его белок выполняет небольшое количество функций: является активатором лецитинхолестеринацилтрансферазы, ингибирует эфиры холестерина-переносящий белок, а также печёночную липазу и фосфолипазу 2.

Чем больше экспрессия одной мРНК, тем большее количество miRNAs связывается с этой мРНК. Это позволяет одной miRNA поддерживать определенный баланс экспрессии соответствующих генов-мишеней. [3]. Наименьшее количество сайтов связывания микроРНК и длинных некродирующих РНК локализовано в цис-регуляторных последовательностях гена *APOC1*. Количество локализованных сайтов связывания коррелирует с числом тканей, в которых идет транскрипция гена.

Заключение

Наша гипотеза: количество сайтов связывания miRNA и lncRNA в интронах и цис-регуляторных районах этих генов коррелирует с числом тканей, в которых экспрессируется ген, а также с числом функций его белка.

Исследование выполнено при финансовой поддержке Министерства науки и высшего образования РФ в рамках государственного задания в сфере научной деятельности № 0852-2020-0028.

Список литературы

1. Романов Д. Е., Шкурят Т. П. Свидетельство о государственной регистрации программы для ЭВМ №2016663454. Dotolog: Программа для автоматизации визуального анализа дотплот-изображений нуклеотидных последовательностей ДНК – Заявка №2016661011. Зарегистрировано в Реестре программ для ЭВМ 07 декабря 2016 г.
2. Шкурят Т. П., Романов Д. Е., Пономарева Н. С., Бахтадзе Г. Б., Лянгасова О. В., Тимофеева С. В., Батталов Д. В., Панич А. Е., Капранов Ф. Распространенность микроРНК в интронах и

- окрестностях генов CGA, FSHB, LHB, TSHB у млекопитающих с формированием различного числа доминирующих фолликулов // Живые и биокосные системы, 2015. 14. URL: <http://www.jbks.ru/archive/issue-14/article-13>.
3. Atambayeva S., Niyazova R., Ivashchenko A., Pyrkova A., Pinsky I., Akimniyazova A., Labeit S. The binding sites of miR-619-5p in the mRNAs of human and orthologous genes // BMC genomics, 2017. 18(1). P. 428.
 4. Bartel D. P. MicroRNAs: Genomics, biogenesis, mechanism and function // Cell, 2004. 116. P. 281—297.
 5. Fagerberg L., Hallström B. M., Oksvold P., Kampf C., Djureinovic D., Odeberg J., Asplund A. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics // Molecular & Cellular Proteomics, 2014. 13(2). P. 397—406.
 6. Hon C. C., Ramilowski J. A., Harshbarger J., Bertin N., Rackham O. J., Gough J., Lizio M. An atlas of human long non-coding RNAs with accurate 5' ends // Nature, 2017. 543(7644). P. 199—204.
 7. Ivashchenko A., Berillo O., Pyrkova A., Niyazova R., Atambayeva S. The properties of binding sites of miR-619-5p, miR-5095, miR-5096 and miR5585-3p in the mRNAs of human genes // BioMed Res. Int., 2014. P. 1674—1684. DOI: 10.1155/2014/720715.
 8. Latorre J., Fernández-Real J. M. LncRN as in adipose tissue from obese and insulin-resistant subjects: new targets for therapy? // EBioMedicine, 2018. 30. P. 10—11.
 9. Milligan M. J., Lipovich L. Pseudogene-derived lncRNAs: emerging regulators of gene expression // Frontiers in genetics, 2015. 5. P. 476.
 10. Shkurat T., Romanov D., Pshenichnyy E., Ponomareva N., Aleksandrova A., Bakhtadze G., Lomteva S. Prevalence of mirnas in introns and cis-regulatory regions of genes of the somatotrophic axis in mammals // American Journal of Applied Sciences, 2015. 12(1). P. 1.
 11. The Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature, 2012. 489. P. 57—74.
 12. Wei S., Du M., Jiang Z. et al. Long noncoding RNAs in regulating adipogenesis: new RNAs shed lights on obesity // Cellular and Molecular Life Sciences, 2016. 73(10). P. 2079—2087.

Spisok literaturey

1. Romanov, D. E., Shkurat T. P. Svidetel'stvo o gosudarstvennoj registracii programmy dlja JeVM №2016663454. Dotolog: Programma dlja avtomatizacii vizual'nogo analiza dotplot-izobrazhenij nukleotidnyh

- posledovatel'nostej DNK – Zayavka №2016661011. Zaregistrirvano v Reestre programm dlja JeVM 07 dekabrja 2016 g.
2. Shkurat, T. P., Romanov, D. E., Ponomareva, N. S., Bahtadze, G. B., Ljangasova, O. V., Timofeeva, S. V., Battalov, D. V., Panich, A. E., Kapranov, F. Rasprostranennost' mikroRNK v intronah i okrestnostjakh genov CGA, FSHB, LHB, TSHB u mlekopitajushhih s formirovanem razlichnogo chisla dominirujushhih follikulov // Zhivye i biokosnye sistemy, 2015. 14. URL: <http://www.jbks.ru/archive/issue-14/article-13>.
 3. Atambayeva, S., Niyazova, R., Ivashchenko, A., Pyrkova, A., Pinsky, I., Akimniyazova, A., Labeit, S. The binding sites of miR-619-5p in the mRNAs of human and orthologous genes // BMC genomics, 2017. 18(1). P. 428.
 4. Bartel, D. P. MicroRNAs: Genomics, biogenesis, mechanism and function // Cell, 2004. 116. P. 281-297.
 5. Fagerberg, L., Hallström, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Asplund, A. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics // Molecular & Cellular Proteomics, 2014. 13(2). P. 397-406.
 6. Hon, C. C., Ramilowski, J. A., Harshbarger, J., Bertin, N., Rackham, O. J., Gough, J., Lizio, M. An atlas of human long non-coding RNAs with accurate 5' ends // Nature, 2017. 543(7644). P. 199-204.
 7. Ivashchenko, A., Berillo, O., Pyrkova, A., Niyazova, R., Atambayeva S. The properties of binding sites of miR-619-5p, miR-5095, miR-5096 and miR5585-3p in the mRNAs of human genes // BioMed Res. Int., 2014. P. 1674-1684. DOI: 10.1155/2014/720715.
 8. Latorre, J., Fernández-Real, J. M. LncRNAs in adipose tissue from obese and insulin-resistant subjects: new targets for therapy? // EBioMedicine, 2018. 30. P. 10-11.
 9. Milligan, M. J., Lipovich, L. Pseudogene-derived lncRNAs: emerging regulators of gene expression // Frontiers in genetics, 2015. 5. P. 476.
 10. Shkurat, T., Romanov, D., Pshenichnyy, E., Ponomareva, N., Aleksandrova, A., Bakhtadze, G., Lomteva, S. Prevalence of mirnas in introns and cis-regulatory regions of genes of the somatotropic axis in mammals // American Journal of Applied Sciences, 2015. 12(1). P. 1.
 11. The Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature, 2012. 489. P. 57–74.
 12. Wei, S. Du, M., Jiang, Z. et al. Long noncoding RNAs in regulating adipogenesis: new RNAs shed lights on obesity // Cellular and Molecular Life Sciences, 2016. 73(10). P. 2079-2087.