

## ВЕБ-СЕРВИСЫ ДЛЯ ПРЕДСКАЗАНИЯ МИШЕНЕЙ МИКРОРНК С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

Аристархов М.А.<sup>1</sup>, Дергилев А.И.<sup>2,3</sup>, Потапова А.Ю.<sup>1</sup>, Иванов-Ростовцев П.А.<sup>1</sup>, Орлов Ю.Л.<sup>1,2,3</sup>

<sup>1</sup> Первый МГМУ им. И.М. Сеченова Минздрава России (Сеченовский Университет)  
ул. Трубецкая, 8-2, г. Москва, 119991, РФ; e-mail: max-a2000@mail.ru

<sup>2</sup> Институт цитологии и генетики СО РАН  
просп. ак. Лаврентьева, 10, г. Новосибирск, 630090, РФ

<sup>3</sup> Новосибирский государственный университет  
ул. Пирогова, 1, г. Новосибирск, 630090, РФ

Поступила в редакцию 10.08.2023. DOI: 10.29039/rusjbpс.2023.0641

**Аннотация.** Поиск генов мишеней микроРНК требует разработки новых программных средств и веб-сервисов. МикроРНК – короткие некодирующие молекулы РНК – играют важную роль в регуляции метаболизма, в ответе на стрессовые воздействия окружающей среды у растений, регулируют экспрессию генов. Понимание функций микроРНК, исследование их генов-мишеней, может помочь в разработке новых лекарственных препаратов, решении биотехнологических задач. Исследование и определение мишеней микроРНК в геноме связано с техническими проблемами. МикроРНК способствует деградации мРНК или подавляет ее трансляцию, и этот процесс может происходить без полной комплементарности мишени. Таким образом определение мишени по принципу комплементарности не однозначно. Кроме того, одна молекула микроРНК может соответствовать сразу нескольким генам-мишеням. Решением является использование больших объемов данных и методов машинного обучения, нейронных сетей. Нейросети в биоинформатике используются для различных задач, таких как анализ биомедицинских данных, диагностика, прогнозирование, классификация и сегментация нуклеотидных последовательностей. Поиск и предсказание мишеней микроРНК с помощью методов машинного обучения активно развивается в настоящее время. Был проведен сравнительный анализ современных нейронных сетей для данной задачи. Разработан веб-сервис для предсказания микроРНК с использованием нейронной сети. С помощью языка программирования Python и библиотеки Flask была разработана серверная часть сервиса. Использовалась нейронная сеть Mitaq, основанная на глубоком обучении, которая способна предсказывать мишени для микроРНК с более высокой точностью. Будут продолжены исследования с целью повышения эффективности и расширения функционала разработанной программной системы.

**Ключевые слова:** биоинформатика, микроРНК, гены-мишени, предсказание, нейронные сети, медицинская информатика, веб-сервис.

### ВВЕДЕНИЕ

Поиск генов-мишеней микроРНК является важной задачей биоинформатики, требующей разработки новых программных средств и веб-сервисов [1-3]. Актуальность задачи связана с тем, что микроРНК регулируют экспрессию генов, играют важную роль в метаболизме, контролируя многие клеточные процессы. МикроРНК являются объектом активного исследования, в том числе в биологии растений, исследовании ответа на стрессовые воздействия окружающей среды [4-7]. Изменение экспрессии микроРНК связано с различными заболеваниями, включая сахарный диабет, ожирение, болезни печени и рак [8].

МикроРНК представляют из себя короткие некодирующие молекулы РНК длиной 18–25 нуклеотидов (в среднем 22). МикроРНК были открыты в конце 20-го века в ходе исследований на модельном организме нематоды *Caenorhabditis elegans* [9]. В ходе исследований было обнаружено, что ген lin-14 кодирует короткий РНК-транскрипт, который играет важную роль в развитии организма. Это привело к открытию нового класса РНК, получивших название микроРНК. В дальнейшем было показано, что микроРНК являются ключевыми регуляторами экспрессии генов [2]. Они играют важную роль в регуляции метаболизма, контролируя многие процессы, такие как регуляция глюкозного метаболизма, метаболизма липидов, аминокислот и железа, восстановления после повреждений тканей и воспаления. МикроРНК играют важную роль практически во всех биологических процессах [10]. Исследования дифференциальной экспрессии микроРНК у организма-хозяина во время инфекции подтверждают, что они могут играть ключевую роль в реакции хозяина на вторгающиеся патогены [7].

В целом, микроРНК являются объектом активного исследования в биологии, биотехнологии и медицине [8,11]. Понимание их биологических функций может помочь в разработке новых лекарственных препаратов и методов лечения различных заболеваний [12].

В мире активно развиваются компьютерные базы данных микроРНК (<https://mirdb.org/>). Необходимо дальнейшее детальное исследование микроРНК, разработка новых программных средств для их анализа, определения их мишеней в геноме [13]. Существующие компьютерные решения связаны с теоретическими и

техническими проблемами. Молекула микроРНК может быть не полностью комплементарной молекуле-мишени мРНК, но все равно вызывает регуляцию экспрессии генов. МикроРНК способствует деградации мРНК или подавляет ее трансляцию, и этот процесс может происходить без полной комплементарности мишени. Это затрудняет однозначное определение мишени молекулы, в связи с невозможностью точно определить мишень по принципу комплементарности [14]. Кроме того, одна молекула микроРНК может соответствовать сразу нескольким генам-мишеням [11].

Решением является использование современных методов биоинформатики – нейронных сетей. Нейросети – это компьютерные системы, которые могут обрабатывать и анализировать данные, используя методы машинного обучения [15]. В медицине нейросети используются для различных задач, таких как диагностика, прогнозирование, классификация и сегментация.

Для нашего веб-сервиса использовалась нейронная сеть *Mitar*, основанная на глубоком обучении (DL-based), которая способна предсказывать мишени для микроРНК с более высокой точностью. Подход объединяет сверточные нейронные сети (CNN), которые преуспевают в изучении пространственных объектов, и рекуррентные нейронные сети (RNN), которые распознают последовательные объекты.

Разработан веб-сервис для предсказания микроРНК с использованием нейронной сети. Проведен сравнительный анализ современных нейронных сетей для данной задачи. Выполнен анализ существующей литературы в области баз данных в Интернете, обращая внимание на основные модели прогнозирования мишеней микроРНК. Среди известных аналогов есть только сервисы, использующие экспериментально подтвержденные данные, а использующие нейронные сети – отсутствуют. Использование предложенной программы поможет улучшить эффективность поиска микроРНК, сократить затраты на экспериментальные исследования и разработку лекарственных средств [16,17].

## МАТЕРИАЛЫ И МЕТОДЫ

Материалами исследования являлись геномные последовательности и их функциональная разметка. Использовались данные из открытых баз данных NCBI (<https://www.ncbi.nlm.nih.gov/>) и UCSC Genome Browser (<https://genome.ucsc.edu/>).

Для разработки использовался PyCharm — интегрированная среда разработки на языке Python, созданная компанией JetBrains (<https://www.jetbrains.com/ru-ru/pycharm/>). Он является инструментом для разработки с использованием современных веб-фреймворков, например, Django, Flask, Google App Engine, Pyramid и web2py (<https://flathub.org/ru/apps/com.jetbrains.PyCharm-Professional>). Для разработки серверной части использован веб-фреймворк Flask языка Python. Flask — это упрощенная платформа Python для веб-приложений, которая обеспечивает основные возможности маршрутизации URL-адресов и визуализации страниц. Сервис работает на архитектуре REST (REpresentational State Transfer) (<https://otus.ru/journal/arhitektura-rest-1-api/>). Для работы с данным сервисом разработаны несколько HTML страниц, которые будут возвращаться ответом на запрос в сервис. HTML - язык маркировки гипертекстов, который не зависит от типа компьютерной платформы.

С помощью языка программирования Python (v. 3.11.4, <https://www.python.org/>) и библиотеки Flask (v. 2.3.2, <https://flask.palletsprojects.com/>) была разработана серверная часть сервиса. Она включает в себя набор функций, которые должны вызываться при поступлении HTTP запроса. Главная страница сервиса содержит форму для вставки последовательности микроРНК. Выполняется проверка на заполненность поля для ввода последовательности, и, если поле заполнено, создается HTTP запрос, который отправляется на серверную часть. Серверная часть, получает запросы в формате Json.

## РЕЗУЛЬТАТЫ

Техническим результатом работы является разработка компьютерного кода. Программа дает оценку присутствия микроРНК для нуклеотидной последовательности.

Рассмотрим механизмы действия микроРНК, полученные в ходе обзора литературы и послужившие основой собственной разработки. У растительных микроРНК практически полное соответствие с их целевыми (мишеневыми) мРНК вызывает подавление генов за счёт расщепления транскриптов. В отличие от этого, микроРНК у животных распознают свои мишеневые мРНК, используя всего 6-8 нуклеотидов (затравочную область) на 5'-конце микроРНК, что недостаточно для расщепления мРНК. Конкретная микроРНК может иметь несколько мРНК мишени, и множество микроРНК может регулировать альтернативную мишень [11].

Оценка количества уникальных мишеней для типичной микроРНК варьирует в зависимости от используемых методов. Однако несколько подходов показали, что у микроРНК млекопитающих есть множество уникальных мишеней. Например, анализ микроРНК у высококонсервативных позвоночных показывает, что каждая имеет в среднем около 400 консервативных мишеней. Похожим образом, эксперименты показывают, что один вид микроРНК может снижать стабильность сотен уникальных информационных РНК. Другие эксперименты показывают, что один вид микроРНК может подавлять продукцию сотен белков, но эта репрессия обычно бывает относительно слабой (меньше, чем в 2 раза) [9].

МикроРНК-опосредованная репрессия включается в процессы регуляции почти всех клеточных процессов. Паттерны регуляции нескольких других микроРНК уже были выявлены у многих видов животных, включая человека.

Согласно последней версии miRBase (<https://mirbase.org/>) (версия 16.0), геном человека содержит 1048 аннотированных микроРНК последовательностей, и возможны дополнительные аннотации в будущем. Согласно литературе, около трети этих микроРНК находится в 113 кластерах генов и, экспериментальные данные профилирования микроРНК в различных тканях и клеточных линиях, показывают, что эти кластеры в основном экспрессируются совместно. Это наблюдение вызвало вопросы о том, как регулируются гены, находящиеся в кластерах, и как это влияет на экспрессию микроРНК [18]. Считается, что если одна микроРНК из кластера теряет свою регуляторную функцию, то это может повлиять на другие микроРНК из того же кластера. Поэтому важно исследовать, можно ли регулировать одну микроРНК в кластере независимо от других, особенно в отношении тех микроРНК, которые играют роль в патологии человека. МикроРНК нацелены на примерно треть генов, кодирующих белки, и одна микроРНК может одновременно взаимодействовать с примерно 200 транскриптами благодаря различным паттернам связывания [18].

Существует три основных механизма регуляции генов, опосредованных микроРНК: репрессия трансляции, прямая деградация мРНК и опосредованный микроРНК распад мРНК. Важно отметить, что последние данные свидетельствуют о том, что механизм репрессии в основном определяется снижением устойчивости мРНК-мишеней.

Разработан сервис для предсказания микроРНК с использованием нейронной сети. Был проведен сравнительный анализ современных нейронных сетей для данной задачи.

С помощью языка программирования Python и библиотеки Flask была разработана серверная часть сервиса. Она включает в себя набор функций, которые должны вызываться при поступлении HTTP запроса.

Для работы пользователя было разработано несколько HTML страниц и css файлов для корректного отображения информации.

Главная страница (рис. 1) содержит форму для вставки последовательности микроРНК. При нажатии на кнопку ввода выполняется проверка на заполненность поля для ввода последовательности и создается HTTP запрос, который отправляется на серверную часть.

Получив тело запроса, серверная часть обрабатывает его нейронной сетью Mitar. Получив ответ от нейронной сети, серверная часть формирует ответ на HTTP запрос. Ответ может содержать в себе предсказанные последовательности мишеней либо сообщение об ошибке, если что-то пошло не так.

Серверная часть, получая запрос, разбирает его на составляющие, чтобы получить тело запроса. Запрос приходит в формате Json. Для корректной работы с сервисом лучше всего использовать Google Chrome, Mozilla Firefox, Microsoft Edge либо любой другой браузер, поддерживающий новые версии html и css.

Получив ответ, разработанная программа (отображает его после заголовка “Результаты” (рис. 3) в виде нескольких (в данный момент настроено на 3) последовательностей мишеней микроРНК с указанием точности их определения.

Кроме того, было реализовано меню для взаимодействия с другими страницами сервиса.

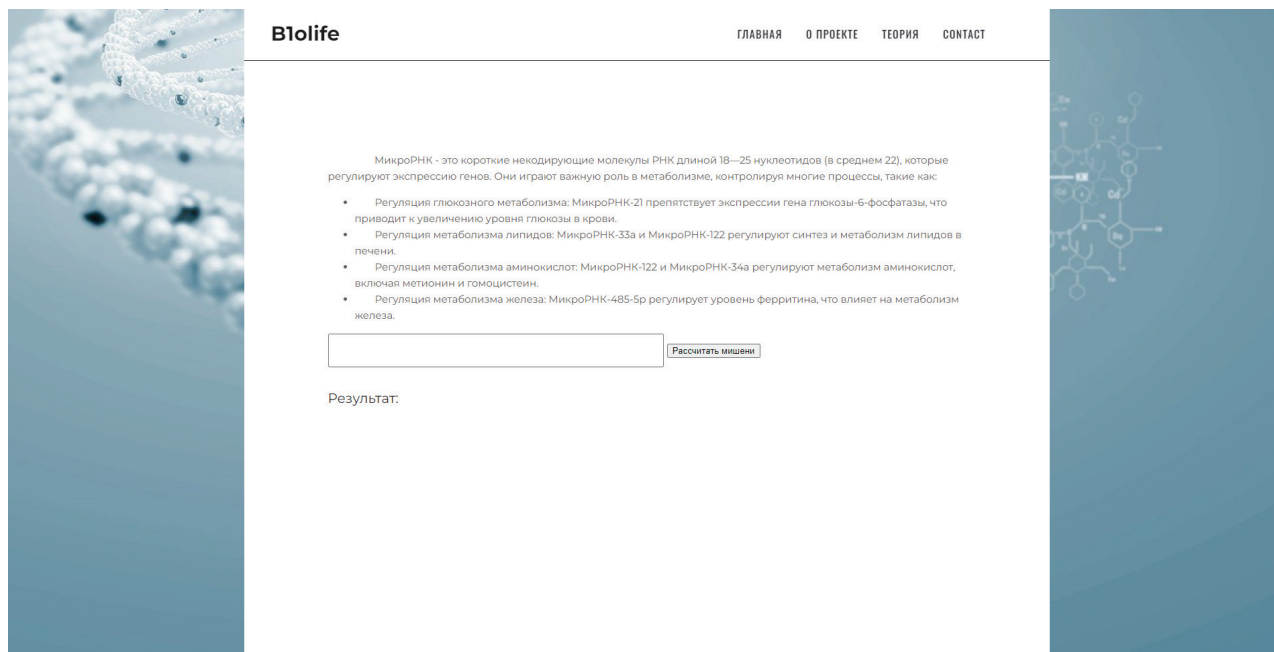


Рисунок 1. Главная страница разработанного сервиса



Рисунок 2. Страница с теоретическими данными



Рисунок 3. Пример запроса предсказания с выводом результата

## ЗАКЛЮЧЕНИЕ

Представлены ключевые компоненты разработанной программы, ее возможности и ограничения, а также примеры ее использования для анализа последовательностей полных генов микроорганизмов, их эволюции [19]. Ранее нами был разработан набор программ для контекстного анализа нуклеотидных последовательностей, в том числе на основе стохастических деревьев [20–22], сложности текста [23], лингвистической сложности [24]. Контекстные сигналы в нуклеотидной последовательности могут быть связаны не только с повторами текста, но и расположением микроРНК, которые имеют шпильчатую структуру [14,25].

Недавние исследования показали существование повторяющихся нуклеотидных последовательностей Short Interrupted Repeat Cassette (SIRC) в геноме арабидопсиса [26]. Компьютерный поиск таких вырожденных нуклеотидных последовательностей технически также может использовать нейронные сети. Ранее было показано влияние микроРНК MIR-203, MIR-129-2, MIR-34A и MIR-34B/C на гены-мишени онкогена TP53 [8,27]. Исследование регуляции экспрессии генов с учетом факторов транскрипции, альтернативного сплайсинга и микроРНК имеет большое значение для биоинформатики, требует интеграции с существующими базами данных [28–31]. Планируется продолжение исследований на основе данной работы с целью повысить эффективность работы сервиса и добавить больше функциональных возможностей, связать его с другими отечественными программными ресурсами биоинформатики, такими как ICGenomics [32,33], расширить доступ к базам данным используя инструмент BioGraph [34].

*Благодарности. Работа поддержана грантом бюджетным проектом ИЦиГ СО РАН FWNR-2022-0020 "Системная биология и биоинформатика: реконструкция, анализ и моделирование структурно-функциональной организации и эволюции генных сетей человека, животных, растений и микроорганизмов". Авторы выражают благодарность В.Н. Бабенко за помощь в работе.*

### Список литературы / References:

1. Chao H., Zhang S., Hu Y., Ni Q., Xin S., Zhao L., Ivanisenko V.A., Orlov Y.L., Chen M. Integrating omics databases for enhanced crop breeding. *J. Integr Bioinform.*, 2023, p. 20230012, doi: 10.1515/jib-2023-0012.

2. Quillet A., Saad C., Ferry G., Anouar Y., Vergne N., Lecroq T., Dubessy C. Improving Bioinformatics Prediction of microRNA Targets by Ranks Aggregation. *Front. Genet.*, 2020, vol. 10, p. 1330, doi: 10.3389/fgene.2019.01330.
3. Roberts J.T., Borchert G.M. Computational Prediction of MicroRNA Target Genes, Target Prediction Databases, and Web Resources. Bioinformatics in MicroRNA Research. *Methods in Molecular Biology*, NY: Springer, 2017, pp. 109-122, doi: 10.1007/978-1-4939-7046-9\_8.
4. Orlov Y.L., Babenko V.N., Dergilev A.V., Galieva A.G., Dobrovolskaya O.B., Chen M. Databases and computer resources on plant miRNA to study its role in abiotic stress response. In book: *Plant Genetics, Genomics, Bioinformatics, and Biotechnology (PlantGen2019) Abstracts*. Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, 2019, p. 150, doi: 10.18699/PlantGen2019-132.
5. Orlov Y.L., Dobrovolskaya O., Yuan C.H., Afonnikov D.A., Zhu Y., Chen M. Integrative computer analysis of antisense transcripts and miRNA targets in plant genomes. *Journal of Stress Physiology & Biochemistry*, 2012, vol. 8, no. 3, p. S7.
6. Орлов Ю.Л., Цуканов А.В., Богомолов А.Г., Добровольская О.Б. Методы биоинформатики для поиска некодирующих РНК, связанных с устойчивостью растений к засухе. *Сборник тезисов республиканской научной конференции современные проблемы генетики, геномики и биотехнологии*. Академия Наук Республики Узбекистан, Центр Геномики и Биоинформатики. 18 мая 2018 года. Ташкент, 2018, с. 139-142 [Orlov Yu.L., Tsukanov A.V., Bogomolov A.G., Dobrovolskaya O.B. Bioinformatics methods for searching for non-coding RNAs associated with plant drought resistance. Collection of abstracts of the republican scientific conference modern problems of genetics, genomics and biotechnology. Academy of Sciences of the Republic of Uzbekistan, Center for Genomics and Bioinformatics. May 18, 2018. Tashkent, 2018, pp. 139-142 (In Russ.)].
7. Wang J., Meng X., Dobrovolskaya O.B., Orlov Y.L., Chen M. Non-coding RNAs and Their Roles in Stress Response in Plants. *Genomics Proteomics Bioinformatics*, 2017, vol. 15, no. 5, pp. 301-312, doi: 10.1016/j.gpb.2017.01.007.
8. Voropaeva E.N., Pospelova T.I., Orlov Y.L., Churkina M.I., Berezina O.V., Gurazheva A.A., Ageeva T.A., Seregina O.B., Maksimov V.N. The Methylation of the p53 Targets the Genes MIR-203, MIR-129-2, MIR-34A and MIR-34B/C in the Tumor Tissue of Diffuse Large B-Cell Lymphoma. *Genes*, 2022, vol. 13, no. 8, pp. 1401, doi: 10.3390/genes13081401.
9. Bartel D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 2004, vol. 116, no. 2, pp. 281-297, doi: 10.1016/s0092-8674(04)00045-5.
10. Klimontov V.V., Koshechkin K.A., Orlova N.G., Sekacheva M.I., Orlov Y.L. Medical Genetics, Genomics and Bioinformatics-2022. *International Journal of Molecular Sciences*, 2023, vol. 24, no. 10, p. 8968, doi: 10.3390/ijms24108968.
11. Matsuyama H., Suzuki H.I. Systems and Synthetic microRNA Biology: From Biogenesis to Disease Pathogenesis. *International Journal of Molecular Sciences*, 2019, vol. 21, pp. 132, doi: 10.3390/ijms21010132.
12. Anashkina A.A., Leberfarb E.Y., Orlov Y.L. Recent Trends in Cancer Genomics and Bioinformatics Tools Development. *Int. J. Mol. Sci.*, 2021, 22, p. 12146, doi: 10.3390/ijms222212146.
13. Riffo-Campos L., Riquelme I., Brebi-Mieville P. Tools for Sequence- Based miRNA Target Prediction: What to Choose? *International Journal of Molecular Sciences*, 2016, vol. 17, no. 12, p. 1987, doi: 10.3390/ijms17121987.
14. Putta P., Orlov Y.L., Podkolodny N.L., Mitra C.K. Relatively conserved common short sequences in transcription factor binding sites and miRNA. *Vavilov Journal of Genetics and Breeding*, 2011, vol. 15, no. 4, pp. 750-756.
15. Schmidhuber J. Deep Learning. *Scholarpedia*, 2015, vol. 10, no. 11, p. 32832, doi: 10.4249/scholarpedia.32832.
16. Berg M.M.J. van den, Krauskopf J., Ramaekers J.G., Kleinjans J.C.S., Prickaerts J., Bried J.J. Circulating microRNAs as potential biomarkers for psychiatric and neurodegenerative disorders. *Progress in Neurobiology*, 2020, vol. 185, p. 101732, doi: 10.1016/j.pneurobio.2019.101732.
17. Vasu S., Kumano K., Darden C.M., Rahman I., Lawrence M.C., Naziruddin B. MicroRNA Signatures as Future Biomarkers for Diagnosis of Diabetes States. *Cells*, 2019, vol. 8, no. 12, p. 1533, doi: 10.3390/cells8121533.
18. Yuan Y., Weidhaas J.B. Functional microRNA binding site variants. *Molecular Oncology*, 2019, vol. 13, no. 1, pp. 4-8, doi: 10.1002/1878-0261.12421.
19. Orlov Y.L., Potapov V.N. Complexity: an internet resource for analysis of DNA sequence complexity. *Nucleic Acids Res.*, 2004, iss. 32, pp. 628-33, doi: 10.1093/nar/gkh466.
20. Витяев Е.Е., Орлов Ю.Л., Вишнеvский О.В., Беленок А.С., Колчанов Н.А. Компьютерная система "GENE DISCOVERY" для поиска закономерностей организации регуляторных последовательностей эукариот. *Молекулярная биология*, 2001, т. 35, № 6, с. 952-960 [Vityaev E.E., Orlov Yu.L., Vishnevsky O.V., Belenok A.S., Kolchanov N.A. Computer system "GENE DISCOVERY" for searching for patterns of organization of regulatory sequences of eukaryotes. *Molecular Biology*, 2001, vol. 35, no. 6, pp. 952-960 (In Russ.)].
21. Vityaev E.E., Orlov Y.L., Vishnevsky O.V., Pozdnyakov M.A., Kolchanov N.A. Computer system "Gene Discovery" for promoter structure analysis. *In Silico Biol.*, 2002, vol. 2, pp. 257-62.
22. Orlov Y.L., Filippov V.P., Potapov V.N., Kolchanov N.A. Construction of stochastic context trees for genetic texts. *In Silico Biol.*, 2002, vol. 2, no. 3, pp. 233-247.
23. Orlov Y.L., Boekhorst R., Abnizova I.I. Statistical measures of the structure of genomic sequences: entropy, complexity, and position information. *J Bioinform Comput Biol.*, 2006, vol. 4, pp. 523-536, doi: 10.1142/s0219720006001801.

24. Митина А.В., Орлов Ю.Л. Оценка лингвистической сложности генетических последовательностей штаммов SARS-CoV-2. *Сборник научных трудов VII Съезда Биофизиков России*: в 2 томах, том 1 – Краснодар: Типография ФГБОУ ВО «КубГТУ», 2023, с. 330, doi: 10.26297/SbR6.2023.001 [Mitina A.V., Orlov Yu.L. Assessment of linguistic complexity of genetic sequences of SARS-CoV-2 strains. *Collection of scientific papers of the VII Congress of Biophysicists of Russia*: in 2 volumes, vol. 1 - Krasnodar: Printing house of FGBOU VO “KubGTU”, 2023, p. 330 (In Russ.)].
25. Дергилев А.И., Орлова Н.Г., Митина А.В., Орлов Ю.Л. Применение методов оценки сложности текста к анализу геномных кластеров сайтов связывания транскрипционных факторов. *Сборник научных трудов VII Съезда Биофизиков России*: в 2 томах, том 1 - Краснодар: Типография ФГБОУ ВО «КубГТУ», 2023, с. 335-336, doi: 10.26297/SbR6.2023.001 [Dergilev A.I., Orlova N.G., Mitina A.V., Orlov Yu.L. Application of methods for assessing text complexity to the analysis of genomic clusters of transcription factor binding sites. *Collection of scientific papers of the VII Congress of Biophysicists of Russia*: in 2 volumes, vol.1 - Krasnodar: Printing house of FGBOU VO “KubGTU”, 2023, pp. 335-336 (In Russ.)].
26. Gorbenko I.V., Petrushin I.S., Shcherban A.B., Orlov Y.L., Konstantinov Y.M. Short Interrupted Repeat Cassette (SIRC)-Novel Type of Repetitive DNA Element Found in Arabidopsis thaliana. *Int J Mol Sci.*, 2023, vol. 24, no. 13, p. 11116, doi: 10.3390/ijms241311116.
27. Voropaeva E.N., Pospelova T.I., Voevoda M.I., Maksimov V.N., Orlov Y.L., Seregina O.B. Clinical aspects of TP53 gene inactivation in diffuse large B-cell lymphoma. *BMC Med Genomics*, 2019, vol. 12, suppl. 2, p. 35, doi: 10.1186/s12920-019-0484-9.
28. Babenko V.N., Bragin A.O., Spitsina A.M., Chadaeva I.V., Galieva E.R., Orlova G.V., Medvedeva I.V., Orlov Y.L. Analysis of differential gene expression by RNA-seq data in brain areas of laboratory animals. *J Integr Bioinform.*, 2016, vol.13, no. 4, p. 292, doi: 10.2390/biecoll-jib-2016-292.
29. Orlov Y.L., Baranova A.V. Editorial: Bioinformatics of Genome Regulation and Systems Biology. *Front Genet.*, 2020, vol. 11, pp. 625, doi: 10.3389/fgene.2020.00625.
30. Orlov Y., Dobrovolskaya O., Chen M., Hofstaedt R. Bioinformatics of genome regulation and structure – 2020 papers collection. *Journal of Integrative Bioinformatics*, 2020 vol. 17, no. 4, p. 20200038, doi: 10.1515/jib-2020-0038.
31. Спицина А.М., Орлов Ю.Л., Подколodная Н.Н., Свичкарев А.В., Дергилев А.И., Чен М., Кучин Н.В., Черных И.Г., Глинский Б.М. Суперкомпьютерный анализ геномных и транскриптомных данных, полученных с помощью технологий высокопроизводительного секвенирования ДНК. *Программные системы: теория и приложения*, 2015, т. 6, № 1(24), с. 157-174, doi: 10.25209/2079-3316-2015-6-1-157-174 [Spitsina A.M., Orlov Y.L., Podkolodnaya N.N. et al. Supercomputer analysis of genomics and transcriptomics data revealed by high-throughput DNA sequencing. *Program systems: theory and applications*, 2015, vol. 6, no. 1(24), pp. 157-174 (In Russ.)].
32. Orlov Y.L., Bragin A.O., Babenko R.O., Dresvyannikova A.E., Kovalev S.S., Shaderkin I.A., Orlova N.G., Naumenko F.M. Integrated Computer Analysis of Genomic Sequencing Data Based on ICGenomics Tool. In: *Advances in Intelligent Systems, Computer Science and Digital Economics. CSDEIS 2019, AISC 1127. International Journal of Intelligent Systems and Applications (IJISA)*, 2020, pp. 154-164, doi: 10.1007/978-3-030-39216-1\_15.
33. Орлов Ю.Л., Брагин А.О. и др. ICGenomics: программный комплекс анализа символьных последовательностей геномики. *Вавиловский журнал генетики и селекции*, 2012, т. 16, № 4/1, с. 732-741 [Orlov Y.L., Bragin A.O. et al. ICGenomics: a program complex for analysis of symbol sequences in genomics. *Vavilov Journal of Genetics and Breeding*, 2012, vol. 16, no. (4/1), pp. 732-741 (In Russ.)].
34. Veljkovic A.N., Orlov Y.L., Mitic N.S. BioGraph: Data Model for Linking and Querying Diverse Biological Metadata. *International Journal of Molecular Sciences*, 2023, vol. 24, no. 8, pp. 6954, doi: 10.3390/ijms24086954.

## WEB-SERVICES FOR MICRORNA TARGET PREDICTION USING NEURAL NETWORKS

Aristarkhov M.A.<sup>1</sup>, Dergilev A.I.<sup>2,3</sup>, Potapova A.Y.<sup>1</sup>, Ivanov-Rostovtsev P.A.<sup>1</sup>, Orlov Y.L.<sup>1,2,3</sup><sup>1</sup> I.M. Sechenov First Moscow State Medical University of the Ministry of Health of Russia (Sechenov University)

Trubetskaya 8-2, Moscow, 119991, Russia; e-mail: max-a2000@mail.ru

<sup>2</sup> Institute of Cytology and Genetics SB RAS

Lavrentieva str., 10, Novosibirsk, 630090, Russia

<sup>3</sup> Novosibirsk State University

Pirogova str., 1, Novosibirsk, 630090, Russia

Received 10.08.2023. DOI: 10.29039/rusjbp.2023.0641

**Abstract.** The pursuit of microRNA target genes necessitates the creation of novel software and web services. MicroRNAs, abbreviated as short non-coding RNA molecules, hold a pivotal role in metabolic regulation, plant responses to environmental stress, and gene expression. Gaining insights into microRNA functions and investigating their target genes can advance drug development and address biotechnological challenges. However, the study and identification of microRNA targets within the genome present technical obstacles. MicroRNA molecules may not exhibit complete complementarity with their mRNA targets. These molecules either contribute to mRNA degradation or inhibit translation, and this process can transpire without full target complementarity. Consequently, the delineation of targets solely based on the principle of complementarity lacks unequivocal clarity. Moreover, a single microRNA molecule can correspond to multiple target genes simultaneously. The solution entails harnessing substantial datasets, employing machine learning techniques, and leveraging neural networks. In bioinformatics, neural networks serve a variety of functions, encompassing the analysis of biomedical data, diagnostics, prediction, classification, and nucleotide sequence segmentation. The pursuit and anticipation of microRNA targets through machine learning methods are currently undergoing vigorous development. A comparative assessment of contemporary neural networks for this task has been executed. A neural network-driven web service for microRNA prediction has been created. The server aspect of the service was developed using the Python programming language and the Flask library. The Mitar neural network, founded on deep learning, was employed. This network demonstrates heightened precision in predicting microRNA targets. We deliberate on the applications of miRNA prediction in gene expression analysis. Sustained research efforts are imperative to enhance the efficiency and broaden the capabilities of the developed computer system.

**Key words:** bioinformatics, microRNA, recognition, neural networks, medical informatics, web-service.