

Регуляторные последовательности в геномах бактерий

Научный руководитель – Комиссаров Алексей Сергеевич

Куреев Николай Алексеевич

Студент (магистр)

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург, Россия

E-mail: nikkureev@rambler.ru

Изучение бактериальных геномов представляет особый интерес для современной биологии. Они значительно отличаются от геномов других организмов, например человека, но имеют не менее сложную систему регуляции. Это особенно важно учитывая наличие у бактерий сложных структур — оперонов. В ходе работы над проектом нами было проанализировано множество уже существующих программ, созданных для поиска последовательностей, регулирующих работу генома [1]. Однако большинство из этих программ показали не лучшие результаты и нуждаются в усовершенствовании.

Для решения проблемы поиска регуляторных последовательностей нами было решено провести комплексный анализ уже существующих подходов. Наиболее эффективными и часто используемыми являются алгоритмы с использованием элементов машинного обучения (логистическая регрессия, метод опорных векторов, деревья решений) [2] и глубокого обучения (сверточные и рекуррентные нейронные сети) [3]. Большинство признаков, используемых для работы моделей, подбираются основываясь на знании биологического устройства оперонов бактерий. Наша работа посвящена поиску наиболее точной стратегии аннотации регуляторных последовательностей в бактериальных геномах и дальнейшей разработке программного обеспечения, способного решить данную задачу.

На данный момент мы обобщили большое количество теоретической информации и создали несколько моделей, которые показывают многообещающие результаты. Полученные нами данные в будущем можно будет использовать для более глубокой аннотации бактериальных геномов, в том числе на уровне оперонов. Такая информация будет крайне полезной в медицинской сфере при разработке новых методов борьбы с патологическими штаммами.

Источники и литература

- 1) Cassiano M. H. A., Silva-Rocha R. Available Benchmarking Bacterial Promoter Prediction Tools: Potentialities and Limitations //Msystems. – 2020. – Т. 5. – №. 4.
- 2) Liu B. et al. Computational prediction of sigma-54 promoters in bacterial genomes by integrating motif finding and machine learning strategies //IEEE/ACM transactions on computational biology and bioinformatics. – 2018. – Т. 16. – №. 4. – С. 1211-1218.
- 3) Umarov R. K., Solovyev V. V. Recognition of prokaryotic and eukaryotic promoters using convolutional deep learning neural networks //PloS one. – 2017. – Т. 12. – №. 2. – С. e0171410.