

## Построение алгоритма глубокого анализа данных масс-спектрометрии белков с учётом пост-трансляционных модификаций

Научный руководитель – Спирин Сергей Александрович

*Строгов Юрий Юрьевич*

*Студент (специалист)*

Московский государственный университет имени М.В.Ломоносова, Москва, Россия

*E-mail: 17234998yura@gmail.com*

Протеомный анализ представленности белков и их количества в биологических образцах с использованием масс-спектрометрических (МС) детекторов всё чаще применяется в практике биохимических и молекулярно-биологических лабораторий. Данный метод анализа является высокопроизводительным и точным, способным однозначно идентифицировать тысячи молекул пептидов и белков в одном эксперименте. В результате такой стандартной обработки можно получить качественный и количественный анализ белкового состава образца.

Существуют методы углубленного анализа МС данных. Примерами таких методов являются идентификация пост-трансляционных модификаций (ПТМ) белков и определение белок-белковых взаимодействий (ББВ) идентифицированных ранее белковых компонентов. Идентификация ПТМ является важной задачей, так как пост-трансляционные модификации могут оказывать большое влияние на структуру и функцию белка в биологической системе. Но поиск ПТМ стандартными методами создает дополнительные сложности в идентификации пептидов с ростом количества ПТМ, на которые нужно проверить образец <sup>[1]</sup>.

В основе разработанного нами алгоритма лежит использование информации о ПТМ белков из базы данных UniProt и использование стандартного метода идентификации белковых групп по белковой базе данных программой MaxQuant. Основные идеи такого алгоритма в том, чтоб автоматизированно проводить поиск модифицированных белков по каждой ПТМ отдельно, в котором сложность поиска возрастает линейно, и уменьшить базу поиска при идентификации по каждой отдельной ПТМ. Этот подход должен минимизировать проблемы, возникающие в стандартных методах поиска.

Вторая важная задача при анализе протеомных данных — это исследования ББВ. Существуют эффективные молекулярные методы определения ББВ, такие как коиммунопреципитация. Однако разработка вычислительных методов для решения этой задачи позволяет получать дополнительную информацию без использования экспериментальных методик, как минимум на начальных этапах анализа. Определение белков-партнеров может помочь в определении функций и ролей исследуемых белков.

Метод исследования ББВ, представленный в данной работе, основывается на использовании экспериментальных данных о ББВ между белками из базы данных UniProt и использовании метода предсказания белок-белкового комплекса с помощью докинга, реализованного в программе LightDock <sup>[2]</sup>. Конечный результат работы программы — интерактивная сеть, которая далее обрабатывается различными методами анализа графов.

Таким образом, в данной работе представлены два алгоритма углубленного анализа данных, позволяющие обогатить информацию, получаемую при исследовании белковых составов экспериментальных образцов.

### Источники и литература

- 1) Bugyi F., Szabó D., Szabó G. et al. Influence of Post-Translational Modifications on Protein Identification in Database Searches. ACS Omega. 2021;6(11):7469-7477.
- 2) Jiménez-García B., Roel-Touris J., Romero-Durana M. et al. LightDock: a new multi-scale approach to protein–protein docking, Bioinformatics. 2018;34(1):49–55.