

Биоинформатический анализ суперсемейств белков на уровне 3D-структуры с использованием методов машинного обучения

Научный руководитель – Суплатов Дмитрий Андреевич

Тимонина Дарья Сергеевна

Аспирант

Московский государственный университет имени М.В.Ломоносова, Факультет
биоинженерии и биоинформатики, Москва, Россия

E-mail: darik93@yandex.ru

Понимание взаимосвязи между последовательностью/структурой белка и его биологической функцией - одна из самых сложных проблем современной биологии. В процессе эволюции белков от общего предка одно функциональное свойство может сохраняться, в то время как другие могут изменяться в результате мутаций, приводящих к функциональному разнообразию. Сравнительный анализ гомологов, реализующих различные свойства в рамках общей структуры суперсемейства, может помочь понять взаимосвязь между структурой, функцией и регуляцией белка, но представляет собой методологическую и вычислительную проблему, поскольку при изучении эволюционно удаленных родственных белков необходимо учитывать как последовательности, так и их 3D-структуры. В этой работе предложен новый подход к биоинформатическому анализу суперсемейств белков на уровне 3D-структуры по двум направлениям: (1) анализ 3D-мотивов, которые структурно схожи во всех исследуемых гомологах и, таким образом, являются 3D-эквивалентами «консервативных позиций» во множественном выравнивании и потенциальными детерминантами общих свойств [1]; (2) анализ «специфических участков подсемейств» (subfamily-specific regions или SSR), которые схожи в подсемействах белков, но различаются между ними и, таким образом, являются 3D-эквивалентами «специфических позиций» и потенциальными детерминантами функционального разнообразия [2]. Проведен биоинформатический анализ особенностей организации локальной 3D-структуры белков, реализующих различные свойства в рамках общей укладки в суперсемействах - нейраминидаз GN34, β/α -Barrel-fold Basic Amino Acid Decarboxylase, α/β -гидролаз, MAP-киназ, металло-бета-лактамаз и других. Сделан вывод о том, что специфически организованные участки локальной 3D-структуры среди гомологов должны рассматриваться как одни из ключевых детерминант функционального разнообразия в суперсемействе, наряду с паттернами специфических замен в аминокислотной последовательности. Соответствующие SSR могут играть ключевую роль в проявлении разных каталитических активностей в разных родственных белках, могут быть ответственны за различия в субстратной специфичности, а также могут характеризовать конформационную пластичность центров связывания лигандов [2]. Разработаны соответствующие веб-приложения, свободно доступные в сети интернет.

Источники и литература

- 1) Suplatov, D., Timonina, D., Sharapova, Y., & Švedas, V. (2019). Yosshi: a web-server for disulfide engineering by bioinformatic analysis of diverse protein families. *Nucleic acids research*, 47(W1), W308-W314.
- 2) Timonina, D., Sharapova, Y., Švedas, V., & Suplatov, D. (2021). Bioinformatic analysis of subfamily-specific regions in 3D-structures of homologs to study functional diversity and conformational plasticity in protein superfamilies. *Computational and Structural Biotechnology Journal*, 19, 1302-1311.