

## ОПТИМИЗАЦИЯ МЕТОДА ФИЛОГЕНЕТИЧЕСКИХ ПРОФИЛЕЙ ДЛЯ ПРЕДСКАЗАНИЯ ВЗАИМОСВЯЗАННЫХ БЕЛКОВ

Пятницкий М.А., Лисица А.В., Арчаков А.И.

ГУ НИИ БМХ им. В.Н.Ореховича РАМН, Россия, 119121, Москва, ул. Погодинская 10,  
Тел.: (495) 246-37-31, E-mail: mpyat@bioinformatics.ru

Взаимодействующие белки определяют большинство процессов в клетке. Изучение взаимосвязанных белков сейчас особенно актуально благодаря успехам крупномасштабных проектов по секвенированию геномов различных организмов. Одним из наиболее используемых подходов для предсказания взаимосвязанных белков является метод филогенетических профилей, основанный на анализе совместной коэволюции белков [1,2]. Два белка считаются структурно-функционально взаимосвязанными в том случае, если они достаточно часто совместно встречаются или отсутствуют у ряда организмов. Каждый белок представляется в виде бинарного вектора – филогенетического профиля (ФП). Применение методов кластерного анализа позволяет определять группы взаимосвязанных белков и проводить аннотацию белков с неизвестными функциями. В качестве целевого организма была выбрана *E.coli*, сравнение предсказаний проводилось с помощью базы данных KEGG. Использовался набор из 120 полных геномов прокариот. Был проведен анализ по предсказанию взаимодействующих белков с помощью представленной методики. Использовались стандартные методы кластеризации, в том числе различные способы иерархической кластеризации, метод *k*-ближайших соседей, РАМ. Впервые для филогенетических профилей был применен метод ART1, использующий нейронные сети и специально разработанный для кластеризации бинарных векторов. Также получена формула для вероятности случайного совпадения филогенетических профилей и оценено распределение этой величины в зависимости от объема кластера путем компьютерного моделирования. Использование методов сравнения разбиений позволило получить методику для количественного сравнения различных методов предсказаний. Применение кластерного анализа к филогенетическим профилям белков исследуемого организма позволяет заключить, что предложенный метод способен с достаточной степенью надежности идентифицировать группы структурно-функционально взаимодействующих белков.

### Литература

1. Wu J., Kasif S., DeLisi C. Identification of functional links between genes using phylogenetic profiles // *Bioinformatics* **19**, 12, 2003. Стр. 1524-30.
2. Snitkin E.S. Gustafson A.M., Mellor J., Wu J., DeLisi C. Comparative assessment of performance and genome dependence among phylogenetic profiling methods // *BMC Bioinformatics*, 7,12, 2006.