

Распределение генов митохондрий грибов в пространстве частот триплетов

Федотовская В.Д.^{1*}, Шпагина Т.О.¹, Колесникова А.И.¹, Садовский М.Г.^{1, 2}

¹ Сибирский федеральный университет, Институт фундаментальной биологии и биотехнологии, Красноярск, Россия

² Институт вычислительного моделирования СО РАН, Красноярск, Россия

* e-mail: viktoriia.fedotovskaia@gmail.com

Ключевые слова: структура, функция, нуклеотидные последовательности, упругие карты, метод динамических ядер, митохондриальные геномы грибов

Мотивация и цель: Выявление связи между структурой нуклеотидной последовательности, функцией, кодируемой ей, и таксономией ее носителя – одна из фундаментальных задач биоинформатики. Такие исследования зависят от выбранного материала. Ранее указанная связь исследовалась на геномах хлоропластов [1] и на последовательности РНК генов бактерий [2].

Методы и алгоритмы: В данной работе связь между структурой, функцией и таксономией изучалась на примере трех генов митохондрий грибов (*atp6*, *atp8* и *atp9*) пяти таксономических единиц – отделов. Из полных митохондриальных геномов были выделены указанные гены в двух версиях: исходные последовательности с интронами и последовательности без интронов. Каждая версия преобразовывалась в частотный словарь триплетов с единичным или тройным шагом рамки считывания. Далее на упругих картах рассматривалось распределение генов в 64-мерном пространстве, где осями являлись частоты триплетов. Кроме того, была проведена кластеризация генов методом динамических ядер. Словари разбивались на кластеры для $2 \leq K \leq 5$, результат кластеризации проверялся на устойчивость.

Результаты: Распределение генов на упругой карте оказалось упорядоченным – выделялись кластеры, точки в которых соответствуют одинаковым генам. Связи с таксономией найдено не было. Результаты кластеризации методом динамических ядер были представлены в виде слоистого графа, где было показано перераспределение генов при увеличении числа кластеров. В этом случае также выделялись геноспецифичные кластеры, не было найдено связи с таксономией.

Заключение и доступность: Таким образом, при использовании двух методов кластеризации (линейного и нелинейного) на примере геномов митохондрий грибов было доказано преобладание функции над таксономией. Данный генетический материал является специфичным, поэтому мы не можем утверждать, что обнаруженный эффект универсален.

Список литературы

1. M. Sadovsky et al. Genome structure of organelles strongly relates to taxonomy of bearers. *Bioinformatics Biomedical Engineering*. 2015:481-490.
2. Горбань А.Н. Классификация нуклеотидных последовательностей по частотным словарям обнаруживает связь между их структурой и таксономическим положением организмов. *Журнал общей биологии*. 2003;64(1):65-77.