

## Применение генетического алгоритма в задаче реконструкции филогении

Научный руководитель – Спирин Сергей Александрович

*Сигорских Андрей Иванович*

*Студент (специалист)*

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

*E-mail: carolusrex8547@gmail.com*

Большинство современных алгоритмов реконструкции филогении сводятся к оптимизации критерия соответствия реконструируемого дерева поданному на вход выравниванию. Чаще всего для этой задачи используются такие методы оптимизации, как Nearest neighbor interchange (NNI) и Subtree pruning and regrafting (SPR). Первый из них работает достаточно быстро, однако часто не находит оптимального решения. Второй более аккуратен, однако требует значительного времени.

Целью данной работы было исследование возможности применения в задаче оптимизации филогенетического дерева какого-либо эволюционного алгоритма. Эволюционные алгоритмы [4] - семейство стратегий оптимизации, имитирующих процесс естественного отбора. В множестве эволюционных алгоритмов был выбран алгоритм Genitor. Попытки применить различные эволюционные алгоритмы к задаче реконструкции филогении предпринимались и раньше [3], однако данную работу отличает введение нового оператора скрещивания деревьев, основанного на нахождении максимального общего поддерева. Алгоритм нахождения максимального общего поддерева был описан в работе [2] и реализован нами в прошлом году.

Оптимизация филогенетического дерева на основе Genitor была реализована в виде модуля к разработанной в нашей группе программе реконструкции филогении PQ [1]. Было проведено тестирование и сравнение с традиционными методами оптимизации. Показано, что такая оптимизация существенно превосходит SPR как в скорости, так и в эффективности работы. Получены оценки рекомендуемых параметров запуска алгоритма.

Работа поддержана грантом РНФ 16-14-10319. Автор выражает благодарность научному руководителю С.А.Спирину.

### Источники и литература

- 1) Пензар Д.Д. PQ: новая программа филогенетической реконструкции // Материалы Международного молодежного научного форума «ЛОМОНОСОВ-2016». М.: МАКС Пресс, 2016. Секция «Биоинженерия и биоинформатика», подсекция «Биоинформатика», номер 28.
- 2) Dong S. Calculation, visualization, and manipulation of masts (maximum agreement subtrees) // Computational Systems Bioinformatics Conference, 2004. Proceedings. 2004 IEEE.
- 3) Lewis P.O. A genetic algorithm for maximum-likelihood phylogeny inference using nucleotide sequence data // Mol. Biol. Evol. 1998. 15(3). Pp. 277–283.
- 4) Whitley D. A genetic algorithm tutorial // Statistics and Computing. 1994. Vol. 4. Pp. 65–85.