

Секция «Биоинженерия и биоинформатика»

Определение таксономии спейсерных последовательностей CRISPR-кассет.

Хабудаев Кирилл Владимирович

Аспирант

Лимнологический Институт СО РАН, Лаборатория ультраструктуры клетки,
Иркутск, Россия

E-mail: khabudaev@gmail.com

Недавно открытые и описанные CRISPR-системы, представляющие собой кассеты с набором идентичных повторов, разделенных переменными последовательностями (спейсерами), имеют важное значение в противофаговом иммунитете прокариот. Согласно ранее выдвинутой теории, спейсеры представляют собой элементы геномов бактериофагов и плазмид, которые однажды заражали клетку [3]. Анализ спейсерной последовательности позволяет определить таксономию организмов, иммунитет к которым имеют исследуемые прокариоты. Однако использование стандартных подходов для анализа таксономии затрудняет небольшая длина спейсера (от 26 до 72 п.о.). Нами был предложен метод исследования спейсеров на основе данных чтений, составляющих сборку метагенома.

В качестве образцов были взяты два метагенома бактериального сообщества, ассоциированного с диатомовой водорослью *Synedra acus* (FE и FK соответственно), различающиеся по видовому составу. Для поиска CRISPR кассет использовался программный инструмент CRT [1], анализ полученных данных проводился с помощью пакета BLAST.

В результате проведенного нами анализа в метагеномах FE и FK были определены 122 и 78 контигов, содержащих CRISPR-кассеты, из которых выделены участки спейсерной последовательности. Обработка полученных спейсеров с помощью процедур blastn и tblastx оказалась невозможной. Следует отметить, что анализ чтений сборки исследуемых метагеномов по белок-кодирующим последовательностям с использованием процедуры tblastx и базы вирусов позволил выделить чтения, представляющие собой вирусную ДНК. Дальнейший анализ выделенных спейсерных последовательностей по этим чтениям с помощью процедуры blastn выявил 100% сходство ($e\text{-value} < 0.01$) у 6 и 7 контигов в метагеноме соответственно, таксономия которых была установлена по GeneBank. Таксономия спейсеров была установлена по сопоставлению с таксономией чтений, имеющих абсолютное сходство с ними.

Предложенный нами метод определения таксономии спейсерной последовательности показал более высокий результат (5% в метагеноме FE и 9% в метагеноме FK) по сравнению с ранее применяемыми методами, где величина идентификации таксономии спейсера по литературным данным находится на уровне 2% [2].

Литература

1. Bland C., Ramsey T.L., Sabree F., Lowe M., Brown K., Kyrpides N.C., Hugenholtz P. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinformatics 2007, 8, 209.
2. Mojica F. J., Diez-Villasenor C., Garcia-Martinez J., Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. 2005, 60, 174–82.

3. Sorek R., Kunin V., Hugenholtz P. CRISPR – a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* 2008, 6, 181-6.