Секция «Биоинформатика»

Свойства данных пробинга РНК – SHAPE

Научный руководитель – Миронов Андрей Александрович

Погорельская Александра Михайловна

Студент (специалист)

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

E-mail: alexapogorelskaya@gmail.com

Предсказание вторичной структуры РНК является фундаментальной задачей вычислительной биологии. Самый распространенный метод предсказания - термодинамический подход, основанный на модели Тернера и алгоритме Зукера [3]. У этого подхода существует набор принципиальных недостатков, из-за которых возникает низкое качество предсказаний, например, термодинамическая модель предсказывает плотный энергетический спектр вторичных структур и не учитывает кинетику формирования структур.

В последние годы большую популярность получил экспериментальный метод SHAPE определения конформаций отдельных рибонуклеотидов, позволяющий значительно улучшить качество предсказания вторичной структуры РНК. Он основан на взаимодействии специального реагента с 2'-OH группой рибозы, если она находится в специальной конформации. Результатом эксперимента является интенсивность модификации реагентом (реакционная способность), которая коррелирует со спаренностью нуклеотида во вторичной структуре. Комбинирование реакционных способностей SHAPE с алгоритмом Зукера позволяет значительно улучшить качество предсказания вторичной структур, что говорит о содержании информации о вторичной структуре в данных SHAPE [1].

Для усиления сигнала вторичной структуры в экспериментальных данных мы разработали модель для итеративной коррекции позиций, в которых вероятностная модель вторичной структуры не согласуется с данными SHAPE. В результате применения этого подхода собственная предсказательная способность данных SHAPE выросла. Однако для предсказания алгоритма, комбинирующего экспериментальные данные и термодинамическую модель, значительного улучшения добиться не удалось.

Чтобы оценить свойства нуклеотидов, реактивность которых модифицируется в результате работы алгоритма, мы использовали 10 РНК с известной пространственной структурой. С помощью базы данных структур РНК URS DataBase [2] мы получили трехмерные параметры нуклеотидов: количество контактов разной природы взаимодействий. На основе этих данных мы оценили вклад взаимодействий в реактивность нуклеотида.

Источники и литература

1) CW Leonard, NV Dokholyan, KM Weeks. Principles for understanding the accuracy of SHAPE-directed RNA structure modeling. Biochemistry 52, 588-595 (2013).

2) Eugene Baulin, Victor Yacovlev, Denis Khachko, Sergei Spirin, and Mikhail Roytberg. URS DataBase: universe of RNA structures and their motifs. Database (Oxford), 2016 May 30.

3) Zuker Prediction of RNA secondary structure by energy minimization. M. Methods Mol Biol. 1994;25:267-94