

Классификация последовательностей ДНК-метилтрансфераз

Научный руководитель – Алексеевский Андрей Владимирович

Серебrenникова Мария Юрьевна

Студент (специалист)

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

E-mail: mserebrennikova@list.ru

ДНК клеток прокариот и эукариот, их вирусов, часто модифицируется путем ферментативного метилирования, которое у прокариот выполняет роль меток сайтов рестрикции систем рестрикции-модификации, для предотвращения их расщепления родственными эндонуклеазами, а также может принимать участие в репарации ДНК, регуляции экспрессии генов и контроле времени репликации ДНК [1].

Существует несколько классификаций ДНК-метилтрансфераз (МТаз): по метилированной группе атомов ДНК; по последовательности консервативных мотивов; по типам рестрикционно-модификационных систем; по особенностям процесса метилирования. Однако общепринятой классификации на основе сходства последовательностей не существует.

Для разработки такой классификации мы загрузили все 132835 последовательностей МТаз из REBASE [2]. В Pfam, базе данных семейств белковых доменов, нами были обнаружены девять семейств, аннотированных как каталитические домены ДНК-метилтрансфераз: PF05869, PF00145, PF07669, PF13651, PF02086, PF05063, PF02384, PF01555, PF12564.

С помощью профилей HMM от Pfam в каждой из 132835 последовательностей МТаз была определена доменная архитектура, то есть порядок всех Pfam доменов. Все МТазы были разделены на три группы: группа (1) включает 25318 МТаз без Pfam доменов; группа (2) включает 6116 МТаз, без каталитического домена из нашего списка; группа (3) включает 101401 МТаз, имеющих каталитический домен, некоторые из них включают более одного каталитического домена.

МТазы групп (2) и (3) были разделены на подгруппы с одинаковыми доменными архитектурами. Для получившихся ста доменных архитектур, состоящих из более чем 10 последовательностей МТаз каждая (всего 106752 МТаз), было построено множественное выравнивание.

В дальнейших планах - получить ответы на следующие вопросы:

Почему последовательности группы (1) вообще не имеют доменов Pfam? Являются ли они МТазами?

Правда ли, что последовательности группы (2) действительно имеют домены, не учтенные в списке из 9 установленных каталитических доменов ДНК-метилтрансфераз? Если такие домены есть, то необходимо найти их.

Каков источник МТаз из группы (3), имеющей более одного каталитического домена? Одним из предполагаемых источников является частая ошибка аннотации в Pfam: домен в белковой последовательности содержится в Pfam как два последовательных одинаковых домена, соответствующих N-концевой и C-концевой частям.

Являются ли МТазы с одинаковой доменной архитектурой гомологичными? Мы планируем проверить, образуют ли домены из последовательностей с одинаковой доменной архитектурой клады в филогенетическом дереве этого каталитического домена.

Работа поддержана РНФ грантом 16-14-10319.

Источники и литература

- 1) Bujnicki JM. Sequence permutations in the molecular evolution of DNA methyltransferases. BMC evolutionary biology. 2002 Dec;2(1):3.
- 2) Roberts RJ, Vincze T, Posfai J, Macelis D. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res. 2015 (Database issue):D298-9.