

## Изучение распространения и эволюции систем рестрикции-модификации семейства AlwI

Научный руководитель – Алексеевский Андрей Владимирович

*Макарикова Ольга Леонидовна*

*Студент (специалист)*

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

*E-mail: omakarikova@bk.ru*

Системы рестрикции-модификации (Р-М) - защитные системы бактерий и архей против чужеродной ДНК, представленные белками с метилтрансферазной и эндонуклеазной активностями. Системы Р-М типа II в силу простого устройства и понятного механизма действия получили широкое применение в молекулярной биологии и генной инженерии, в частности, системы с эндонуклеазой рестрикции (ЭР) семейства AlwI, к которому относятся некоторые никирующие ЭР (никазы). Целью данной работы является изучение распространения и эволюции (горизонтальные переносы) систем Р-М, содержащих белки этого семейства.

Семейство ЭР AlwI представлено в Pfam доменом RE\_AlwI. На основании разметки доменов по пространственной структуре никазы N.BspD6I обнаружено, что профиль домена RE\_AlwI охватывает лишь малую часть каталитического домена. По выравниванию каталитических доменов проверенной и не-избыточной выборки (87 последовательностей), из REBASE (содержащей информацию о всех системах Р-М) с доменом RE\_AlwI, был построен новый профиль семейства, названный cat\_AlwI. Он был сравнен с профилем Pfam на 1097 полных протеомах прокариот из базы данных Uniprot, содержащих хотя бы одну ЭР с доменом RE\_AlwI. Проверено, что 84 последовательности с доменом RE\_AlwI, не найденные профилем cat\_AlwI, найдены профилем RE\_AlwI за счет сходства ДНК-связывающего и линкерного доменов, а не каталитического.

По данным REBASE о системах Р-М с ЭР с доменом RE\_AlwI найдено 20 классов систем Р-М, содержащих метилтрансферазы из 3-х разных семейств. Классом считается группа систем Р-М с одинаковыми каталитическими доменами ЭР и метилтрансфераз. Самым представленным оказался класс с одной ЭР и двумя метилтрансферазами семейства Pfam MethyltransfD12. Для этого класса были получены координаты в геноме для генов белков ЭР, найденных профилем cat\_AlwI в упомянутых выше 1097 протеомах, и метилтрансфераз, найденных в тех же 1097 протеомах профилем Pfam MethyltransfD12. В результате были отобраны белки, находящиеся в одном локусе и наиболее вероятно ассоциированные с одной системой Р-М. Для таких белков ЭР (80 последовательностей) и метилтрансфераз (155 последовательностей) были построены филогенетические деревья с целью дальнейшего анализа эволюции этого класса систем Р-М.

Для класса, включающего 24 системы Р-М с ЭР семейства AlwI и двумя МТазы семейства N6\_N4\_mtase показано, что на филогенетическом дереве 48 МТазы разбиваются на две клады по 24 листа так, что одна МТазы из системы принадлежит одной клade а вторая - другой. Предположительно, различия в последовательностях МТазы из двух клад связаны с непалиндромными сайтами узнавания.

Для каждого класса систем Р-М с ЭР семейства AlwI планируется провести филогенетический анализ с целью определения эволюционных событий и, в частности, изучить разделение на клады для метилтрансфераз из других классов систем Р-М с двумя метилтрансферазами.

Работа поддержана грантом РФФИ № 21-14-00135.