

БАЗА ДАННЫХ НЕКОДИРУЮЩИХ ЭЛЕМЕНТОВ ДНК МОРСКИХ ЕЖЕЙ

Савина Ю.А. (Университет ИТМО)

Научный руководитель – в.н.с., к.б.н. Комиссаров А.С. (Университет ИТМО)

Некоторые виды морских ежей являются важными модельными организмами в биомедицинских исследованиях. Однако до сих пор большая часть ныне живущих видов иглокожих остается малоизученной. С помощью создания базы данных некодирующих элементов ДНК хорошо изученных видов морских ежей, возможно усовершенствовать процесс сборки и аннотации геномов новых видов. Проект будет иметь большую ценность для научного сообщества как ресурс для изучения эволюции генома, развития и экологии иглокожих.

Введение. Морские ежи – это класс морских беспозвоночных, насчитывающий свыше 1000 современно живущих видов. Они занимают уникальную позицию на филогенетическом древе и имеют общего предка с человеком, с которым насчитывает приблизительно 8000 общих генов. Это позволяет считать данную группу организмов важной моделью в биомедицинских исследованиях. Хотя за последние 14 лет были созданы такие геномные и транскриптомные ресурсы, как EchinoBase и SpBase, до сих пор большая часть ныне живущих видов остается малоизученной. С целью усовершенствования процесса сборки и аннотации геномов морских ежей было предложено создать базу данных некодирующих элементов их ДНК. Описание этих регуляторных элементов и понимание того, как они развиваются, предоставляет беспрецедентные возможности для определения функционально важных элементов.

Основная часть. Для работы были отобраны все имеющиеся в открытом доступе геномные сборки морских ежей: *Strongylocentrotus purpuratus*, *Hemicentrotus pulcherrimus*, *Lytechinus variegatus* и *Eucidaris tribuloides*. Для оценки их качества использовался инструмент оценки сборки генома (QUAST 5.0.2). Программа оценивает качество сборки генома путем вычисления различных метрик, таких как число контигов и скаффолдов и их длины. Поскольку геном пурпурного морского ежа собирался несколько раз в период с 2006 года, последний опубликованный вариант сборки его генома обладает самой высокой степенью качества, по сравнению с остальными видами. С помощью инструментов RepeatModeler и RepeatMasker исследуемые последовательности ДНК были проверены на наличие перемежающихся повторов и последовательностей ДНК низкой сложности. Кодированные геномные области были предсказаны с помощью AUGUSTUS – инструмента для поиска генов, кодирующих белки, и их экзон-интронной структуры. Найденные гены были проаннотированы с помощью инструмента eggNOG-mapper. На заключительном этапе исследования все геномные сборки были проанализированы с целью идентификации тандемных повторов. В настоящий момент ведется работа по сведению всех полученных геномных данных в единую базу данных.

Выводы. База данных некодирующих элементов ДНК морских ежей является ценным ресурсом для анализа генов и геномных последовательностей иглокожих. Детальное изучение геномов морских ежей поможет ответить на вопросы о влиянии генов на фенотип, а также даст возможность инициировать открытие новых элементов ДНК, новых веществ и новых материалов.

Савина Ю.А. (автор)

Подпись

Комиссаров А.С. (научный руководитель)

Подпись