

Экспрессия генов сериновых пептидаз и их гомологов на разных стадиях развития большого мучного хрущака *Tenebrio molitor***Научный руководитель – Элпидина Елена Николаевна***Салимгареев Р.С.¹, Терещенкова В.Ф.², Жиганов Н.И.³*

1 - Московский государственный университет имени М.В.Ломоносова, Факультет биоинженерии и биоинформатики, Москва, Россия, *E-mail: s18b1_salimgareev@179.ru*; 2 - Московский государственный университет имени М.В.Ломоносова, Химический факультет, Кафедра химии природных соединений, Москва, Россия, *E-mail: lerunehka_lu@mail.ru*; 3 - Московский государственный университет имени М.В.Ломоносова, Биологический факультет, Кафедра энтомологии, Москва, Россия, *E-mail: nikitoooc@rambler.ru*

Tenebrio molitor, или большой мучной хрущак, принадлежит к семейству жуков-чернотелок (Tenebrionidae), является вредителем запасов зерновых продуктов и признанной биохимической моделью. В отличие от родственного жука *Tribolium castaneum*, являющегося генетической моделью, геном *T. molitor* полностью не секвенирован. В связи с этим особенный интерес представляют новые данные RNA-Seq, позволяющие делать выводы о функционировании генов *T. molitor*. У насекомых из этого семейства выявлены активности единичных пищеварительных цистеиновых и сериновых пептидаз (СП), в то время как общее число транскриптов пептидаз из семейств папаина С1 (цистеиновые) и хитотрипсина S1 (СП) в транскриптоме кишечника личинок *T. molitor* составляет 29 и 223, соответственно.

Целью работы был биоинформатический анализ уровня экспрессии генов СП и их неактивных гомологов (ГСП) из семейства S1 у *T. molitor* на разных стадиях онтогенеза (яйцо, личинки II и IV стадий, куколка, имаго) и, в соответствии с этим, предсказание функций этих белков.

Прочтения РНК были получены путём высокопроизводительного секвенирования второго поколения на платформе Illumina. Выделение РНК было произведено с помощью набора RNEasy Mini, концентрация была измерена флуориметром Qubit, целостность РНК проверена с помощью капиллярного электрофореза на приборе BioAnalyzer 2100. Для подготовки библиотек использован набор NEBNext RNA Library Prep Kit для Illumina, секвенирование проведено на приборе Illumina HiSeq2000 с использованием набора реагентов TruSeq SBS Kit v3. Чтения были приведены к формату fastqс при помощи программы CASAVA версии 1.8.2. После оценки качества прочтений программой FastQC и предварительной обработки в Trimmomatic чтения были картированы на референсный транскриптом *T. molitor* (картировано от 83 % до 93 % прочтений). Сборка транскриптов была выполнена программой Cufflinks, а оценка величины FPKM (Fragments Per Kilobase per Million mapped reads) - StringTie (опция -B). Для подсчёта RPKM (Reads Per Kilobase per Million mapped reads) и FPKM был использован специально составленный сценарий с применением tBLASTn, количество различных мРНК, для которых была оценена экспрессия, составило 282.

Значимый уровень экспрессии транскриптов (RPKM > 100) выявлен на стадии яйца для 9 СП и 5 ГСП, у личинок II стадии - для 20 СП и 6 ГСП, у личинок IV стадии - для 46 СП и 38 ГСП, у куколок - для 19 СП и 16 ГСП, у взрослых насекомых - для 66 СП и 34 ГСП. Полученные значения экспрессии генов белков позволили разделить их на категории в соответствии с паттерном экспрессии. Были выявлены группы СП и ГСП с высокой экспрессией на питающихся стадиях, с постоянной значимой экспрессией, экспрессируемые в основном на непитающихся стадиях и низкоэкспрессируемые. Интересно, что максимальное количество ГСП было выявлено на питающихся стадиях.

Исследование по активным СП поддержано грантом РФФИ (№ 18-04-01221 А), а по ГСП (псевдопептидазам) выполнено за счет гранта Российского научного фонда (проект № 19-74-00080).