

# The chymotrypsin-like S1 peptidase family of Tenebrionidae beetles

P-01.3-19

N. Zhiganov <sup>\*I</sup>, V. Tereshchenkova <sup>\*II</sup>, R. Salimgareev <sup>III</sup>, I. Filippova <sup>II</sup>, E. Elpidina <sup>IV</sup>

<sup>I</sup>Faculty of Biology, Lomonosov Moscow State University, Moscow, 119991, Russia, <sup>II</sup>Chemical Faculty, Lomonosov Moscow State University, Moscow, Russia,

<sup>III</sup>Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russia, <sup>IV</sup>A.N.Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russia

The chymotrypsin S1 family is one of the most numerous families of peptidases. We studied and characterized in detail the set of predicted S1 serine peptidases (SP) in the transcriptomes and genome of two model Tenebrionidae insects, *Tenebrio molitor* and *Tribolium castaneum*. S1 SP provide up to 60% of the total digestive proteolytic activity in *T. molitor* and up to 27% in *T. castaneum* larvae. The search for sequences of SP was carried out in the gut and whole larvae RNA-Seq transcriptomes of *T. molitor* lacking sequenced genome, and in the genome and larval gut transcriptome of *T. castaneum*. *T. molitor* SP were manually annotated and GenBank annotation of *T. castaneum* was manually curated if needed. *T. molitor* larval gut transcriptome contained 196 predicted sequences of proteins homologous to S1 SP, and *T. castaneum* gut transcriptome - 157 sequences. Among them, there were 97 predicted active enzymes in *T. molitor* and 102 in *T. castaneum*, containing typical catalytic residues. The remaining sequences, 99 and 55 respectively, lacked conservation in the active site residues and were classified as inactive SP homologs. Additional 138 sequences were found in *T. molitor* whole larvae transcriptome. The sets of SP from both insects contained predicted sequences of trypsins, chymotrypsin-like peptidases, elastases, collagenases, and 22-26 non-annotated sequences contained unusual residues of the S1 substrate binding subsite. The most numerous group in both insects was trypsins (34–44 sequences), while collagenases were the smallest group (5–4). Only 8-10 of active SP had a high level of mRNA expression, lacked additional regulatory domains, and can presumably be related to digestive peptidases. At the same time, almost half of the trypsins with a low level of mRNA expression contained a regulatory clip domain and presumably belonged to the regulatory SP. A phylogenetic analysis of S1 SP was carried out.

This research was funded by the RFBR, project # 20-54-56044 Iran\_t.

\* The authors marked with an asterisk equally contributed to the work.