

Transcriptome-based analysis of the diversity of membrane-bound lectins in Baikal amphipods *Eulimnogammarus* sp. and the Holarctic amphipod *Gammarus lacustris*

Lubyaga Y.A.^{1,4*}, Dolgikh A.V.^{2,3}, Drozdova P.B.^{1,4}, Nazarova A.A.^{1,4}, Timofeyev M.A.¹

¹ Irkutsk State University, Institute of Biology, Lenin St., 3 Irkutsk, 664025, Russian Federation

² Dept. of Genetics and Biotechnology St. Petersburg State University, Universitetskaya Emb., 7-9, St. Petersburg, 199034, Russian Federation

³ Bioinformatics Institute, Kantemirovskaya St., 2A, St. Petersburg, 197342, Russian Federation

⁴ Baikal Research Center, Irkutsk, Lenin St., 21, 664003, Russian Federation

ABSTRACT. The aim of this study was transcriptome-based analysis of the diversity of membrane-bound lectins in Baikal amphipods *Eulimnogammarus verrucosus* and *E. cyaneus* and a Holarctic amphipod *Gammarus lacustris*. The results of this study show that studied species of amphipods possess quite similar repertoires of lectins. The obtained lectin sequence data can be used in further molecular-level exploration of amphipod physiology.

Keywords: Amphipods, Baikal, lectins, gene expression, transcriptome-based analysis

1. Introduction

The functioning of the immune system remains a poorly studied area of modern physiology of aquatic invertebrates. At the same time, the peculiarities of the immune response of different species against certain pathogens are important for their fate in the ecosystem.

Lake Baikal is the largest freshwater reservoir inhabited by a diverse fauna of invertebrates with the high level of endemism. The dominant organisms of Baikal, both in composition and in number of species, are communities of numerous benthic macroinvertebrates, in particular, amphipods (Amphipoda, Crustacea) (Takhteev, 2000). The extraordinary ecological and morphological differentiation of Baikal amphipods during their evolution could lead to a significant variety of receptors responsible for the recognition of foreign bodies and the triggering of the immune response, especially due to the possible co-evolution of certain types of amphipods with their specific pathogens. Thus, different amphipod species may have significant differences in their sensitivity to certain pathogens. Lectins are a very ancient and versatile group of proteins containing carbohydrate-binding domains. Lectins are important in cellular interactions and tissue development and are also key in recognizing pathogens. Thus, these proteins are both an element

of normal physiological processes of the body and participants in protective reactions (Gupta et al., 2012). In this work, we focused on a factor important for the immune system of crustaceans, the composition and expression of lectins.

2. Material and methods

The objects of this study were two Baikal endemic species of amphipods *Eulimnogammarus verrucosus* and *E. cyaneus*, as well as a Holarctic amphipod *Gammarus lacustris*.

To check the presence of transcripts encoding putative lectins, we searched the available transcriptome assemblies of amphipods of the Baikal region (Naumenko et al., 2017; Drozdova et al., 2019) with known sequences of these proteins found in other species using the blastp algorithms of the ncbi-blast (Camacho et al., 2009) package. Protein sequences were predicted with TransDecoder (Haas et al., 2013). The expression level of the found transcripts was evaluated using Trinity (Grabherr et al., 2011) where the Bowtie2 tool was used for alignment, and RSEM (Li and Dewey, 2011) was used as a utility to count the expression level of transcripts in samples. Analysis of protein sequences for the presence of transmembrane domains, a signal peptide and the intracellular or extracellular location

*Corresponding author.

E-mail address: yuliya.a.lubyaga@gmail.com (Y.A. Lubyaga)

of the main part of the protein molecule was carried out using the web interfaces Protter and TMHMM-2.0.

3. Results

The results of the search for lectins were similar for two endemic amphipod species of Lake Baikal (*E. cyaneus* and *E. verrucosus*) and a potential invader species, *G. lacustris*. In total, we found over 1000 sequences in the studied species that were homologous to the C-type lectins, 25 transcripts homologous to the S-type lectins, 27 transcripts homologous to the L-type lectins, 65 transcripts homologous to the M-type and 30 transcripts the P-type lectins.

Further exploration of the expression levels of these transcripts showed that transcripts of the C-type lectins have the highest expression level (>10 transcripts per million) compared to other types of lectins in all three species. This result may mean that some of the C-type lectins are expressed constitutively and/or in multiple tissue types, while the other types are expressed only in particular types of cells or only in response to infection. However, it was shown that sequences with a high level of expression do not have transmembrane domains, or the predicted protein sequences are based on incomplete transcripts. Thus, protein sequences of 9 C-type lectins, 3 L-type lectins, 6 M-type lectins and 4 P-type lectins having a transmembrane domain, a start and stop codon, and extracellular localization of the main part of the protein were predicted. Nonetheless, the expression of these transcripts does not exceed 1 transcript per million.

4. Conclusions

The analysis of amphipod transcriptomes helped us to discover many putative lectin sequences, which can be used in further molecular-level exploration in studies of Baikal amphipods. The main criterion for the selection of sequences will be their similarity with the studied lectins of other species, and their expression level will also be considered. Nevertheless, it should

be noted, that lectins present on the certain type of cells may not have a high level of expression at the level of the transcriptome of the whole organism. In addition, their expression level may be low under control conditions and increase when pathogens enter the body.

This study was partially supported by grants Russian foundation for basic research (projects N 20-44-383007) and Russian science foundation (project N 20-64-46003).

References

- Camacho C., Coulouris G., Avagyan V. et al. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10. DOI: 10.1186/1471-2105-10-421
- Drozdova P., Rivarola-Duarte L., Bedulina D. et al. 2019. Comparison between transcriptomic responses to short-term stress exposures of a common Holarctic and endemic Lake Baikal amphipods. *BMC Genomics* 20: 1-14. DOI: 10.1186/s12864-019-6024-3
- Grabherr M.G., Haas B.J., Yassour M. et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29. DOI: 10.1038/nbt.1883
- Gupta G.S. 2012. Lectins: an overview. In: *Animal lectins: form, function and clinical applications*. Wien, pp. 3-25. DOI: 10.1007/978-3-7091-1065-2_1
- Haas B.J., Papanicolaou A., Yassour M. et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8: 1494-1512. DOI: 10.1038/nprot.2013.084
- Li B., Dewey C.N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12. DOI: 10.1186/1471-2105-12-323
- Naumenko S.A., Logacheva M.D., Popova N.V. et al. 2017. Transcriptome-based phylogeny of endemic Lake Baikal amphipod species flock: fast speciation accompanied by frequent episodes of positive selection. *Molecular Ecology* 26: 536-553. DOI: 10.1111/mec.13927
- Takhteev V.V. 2000. *Ocherki o bokoplavakh ozera Baikal: sistematika, sravnitel'naya ekologiya, evolyutsiya* [Essays on the amphipods of Lake Baikal: systematics, comparative ecology, evolution]. Irkutsk: Irkutsk State University Press. (in Russian)