doi 10.18699/bgrs2024-2.1-15

Reconstruction and computer analysis of the structural and functional organization of the gene network regulating cholesterol biosynthesis in humans and the evolutionary characteristics of genes involved in the network

Ignatieva E.V.[1]*, Mikhailova A.D.[2], Lashin S.A.[1]

[1] *Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

[2] *Novosibirsk State University, Novosibirsk, Russia*

* *eignat@bionet.nsc.ru*

Key words: cholesterol biosynthesis; gene network; feedback loops; evolution; gene variability; phylostratigraphic analysis; gene age; stabilizing selection; SREBP

Motivation and Aim: Cholesterol is a structural component of cell membranes, a precursor to steroid hormones and other biochemical substances in the animal body. Cholesterol is consumed by animals with food, and is also synthesized in the cells *de novo*. Previously, the gene network of regulating intracellular cholesterol level was reconstructed [1]. This network comprised a number of regulatory loops, each of which included transcription factors of the SREBP family (SREBP-1 and SREBP2). The activity of transcription factors of the SREBP family (SREBPs) is regulated inversely depending on the intracellular cholesterol level. The functioning of this mechanism is carried out by several proteins, which, in combination with SREBPs, may be called "cholesterol sensor" (that is, SCAP, INSIG1, INSIG2, SREBP1, SREBP2). The other mechanisms regulating SREBPs functioning are known, due to which a precise modulation of SREBPs activity is carried out depending on the state of the cell and the organism as a whole [1]. It is also known that elevated cholesterol level may be a risk factor for the development of cardiovascular diseases (atherosclerosis, coronary heart disease), and may increase the severity of other pathological processes. Systematization of data on the molecular mechanisms controlling the activity of SREBPs and cholesterol biosynthesis in the format of a gene network and obtaining new knowledge about the gene network as a single object is extremely important in the context of understanding the involvement of this system in the diseases progression. The aim of the current study is to build an extended gene network regulating cholesterol biosynthesis in humans and to analyze the structural and functional organization of the network and the evolutionary characteristics of genes.

Methods and Algorithms: We used ANDSystem [2] to reconstruct and to analyze the gene network regulating cholesterol biosynthesis. Evolutionary metrics of genes (phylostratigraphic age index (PAI) and divergence index (DI)) have been calculated as it was described in [3] taking into account sequences of orthologous genes that are 50 % or more identical to one under consideration.

Results: Using ANDSystem we built a new extended version of the gene network regulating cholesterol biosynthesis. It includes (1) proteins that make up the "cholesterol sensor" and the genes encoding them; (2) proteins regulating the activity of genes and pro-

teins of the “cholesterol sensor”; (3) cholesterol and enzymes involved in its biosynthesis; (4) genes regulated by transcription factors of the SREBP family, and the proteins encoded by them.

Regulatory circuits have been identified that control the activity of SREBPs. All these regulatory circuits include proteins encoded by the SREBPs target genes. The four shortest circuits show the regulatory effects of PPARG or NR0B2/SHP1 or FAS or LPIN1 on the activity of SREBP-1 or SREBP-2. Within the other five circuits, the regulation of SREBP-1 or SREBP-2 activity is carried out indirectly, due to the effects of FAS or PPARG or AR on the proteins functioning within “cholesterol sensor” (INSIG1 or SCAP), which, in turn, affect the activity of SREBP-1 or SREBP-2.

The analysis of the evolutionary characteristics of genes using PAI and DI metrics revealed that: (1) the genes involved in the gene network regulating cholesterol biosynthesis are characterized by decreased PAI values compared to the set of all human protein-coding genes, that may mean that they are on average more “ancient”, in addition, these genes have significantly decreased DI values (for most of them DI is less than 0.4), indicating a stronger pressure of stabilizing selection on these genes compared to the pressure of stabilizing selection on all human protein-coding genes as a whole; (2) the subgroup of genes, encoding enzymes of cholesterol biosynthesis contains significantly more genes with a greater phylostratigraphic age (PAI $\leq$ 2, cellular organisms (the root of the phylogenetic tree), and the stages of eukaryota and metazoa divergence) than protein-coding genes as a whole, that means that they are more “ancient”; (3) the subgroup of genes, encoding regulatory proteins that affect cholesterol sensor activity, contains significantly more genes with a greater phylostratigraphic age (PAI < 5, all stages before vertebrata divergence) than protein-coding genes as a whole, that means that they are more “ancient”; (4) the subgroup of genes, encoding regulatory proteins that affect cholesterol sensor, contains significantly more genes having low DI values (less than 0.4), indicating a stronger pressure of stabilizing selection on these genes compared to the pressure of stabilizing selection on all human protein-coding genes as a whole.

Conclusion: Analysis of the gene network regulating cholesterol biosynthesis has revealed a complex nature of the molecular-genetic mechanisms controlling this process. We identified new genes involved in feedback loops and revealed features of evolutionary characteristics of individual subgroups of genes involved in the network. Taking into account these findings may be useful in the search for new pharmacological targets for the treatment of pathologies associated with cholesterol level deviations.

Funding: The study is supported by publicly funded project No. FWNR-2022-0020 of the Federal Research Center ICG SB RAS.

References

1. Kolchanov N.A., Ignatieva E.V., Podkolodnaya O.A., Likhoshvai V.A., Matushkin Yu.G. Gene Networks *Vavilov Journal of Genetics and Breeding*. 2013;17(4/2):833-850
2. Ivanisenko V.A., Demenkov P.S., Ivanisenko T.V., Mishchenko E.L., Saik O.V. A new version of the ANDSystem tool for automatic extraction of knowledge from scientific publications with expanded functionality for reconstruction of associative gene networks by considering tissue-specific gene expression. *BMC Bioinformatics*. 2019;20(Suppl. 1):34. doi 10.1186/s12859-018-2567-6
3. Mustafin Z.S., Lashin S.A., Matushkin Yu.G. Phylostratigraphic analysis of gene networks of human diseases. *Vavilov Journal of Genetics and Breeding*. 2021;25(1):46-56. DOI 10.18699/VJ21.006