Reconstruction and analysis of the human protein-protein interaction network involved in response to tick-borne encephalitis virus infection.

Elena V. Ignatieva

The Federal Research Center Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia; Novosibirsk State University, Novosibirsk 630090, Russia, eignat@bionet.nsc.ru

Alexander V. Igoshin

The Federal Research Center Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia, igoshin@bionet.nsc.ru

Nikolay S. Yudin

The Federal Research Center Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia; Novosibirsk State University, Novosibirsk 630090, Russia, yudin@bionet.nsc.ru

Tick-borne encephalitis (TBE) is a serious human neurological disease caused by TBE virus (TBEV). In humans, infections by TBEV may result in encephalitis, meningitis and haemorrhagic fevers with mortality rates as high as 20–30% [1]. At present, the mechanisms of TBEV-caused pathogenesis remain unclear and the contributions of host proteins during TBEV infection are poorly understood. As well, very little is known about genetic factors predisposing to diseases caused by TBEV. In this work, we investigate the genetic basis of virus-host interactions basing on theoretical analysis of the network involving TBEV RNA or proteins and proteins of the host cell. The objectives of this study were: (1) to compile a catalog of human genes involved in response to TBEV infection; (2) to construct and analyze networks formed by associations between genes/proteins from the catalog; (3) to reveal additional genes/proteins which may be involved into the network through protein-protein interactions; (3) to prioritize genes/proteins according to the number of their neighbors in the network.

Performing queries to COREMINE system (https://www.coremine.com/medical/#search) and PubMed (https://www.ncbi.nlm.nih.gov/pubmed/), we created a catalog of genes involved in response to TBEV infection. The current version of the catalog comprised 35 genes. Among them, 17 genes (Sublist A) encoded proteins involved in direct physical interactions with the TBEV proteins or RNA. Eleven proteins (Sublist B) participated in nonspecific response to viral infection (unfolded protein response pathway and other pathways). Seven genes/proteins

(Sublist C) had allelic variants associated with severe forms of the diseases caused by TBEV [2].

To obtain more characteristics of genes involved in response to TBEV infection and to predict new genes which may be relevant to this system, we constructed network presenting pairwise physical interactions between proteins encoded by genes from the catalog. Data on pairwise physical interactions between 35 proteins/genes from catalog and 50 additional proteins (Sublist D) were obtained from STRING [3] and uploaded into Cytoscape [4]. Only high-confidence edges with STRING scores greater than 0.4 were included. We also uploaded into the network two types of data extracted from scientific publications: (1) interactions between viral proteins or RNA and host proteins; (2) interactions between proteins within nonspecific response pathways. As a result, we obtained network involving 96 nodes and 719 edges.

To characterize the functional domains of the network involved in response to TBEV infection, we explored clusters identified by the MCODE tool [5]. We revealed one extremely dense cluster formed by 35 nodes with very high score exceeding 30. This cluster comprised 30 ribosomal proteins and 5 other proteins that also had high number of associations with other nodes in cluster. Among 35 genes/proteins involved in cluster, we found *RPSA*, encoding ribosomal protein SA that had direct physical interactions with viral E protein [6]. *RPSA* encodes protein with dual function: (1) this protein is also known as laminin binding protein (LBP) located on the cell membrane and functioning as a cell receptor for TBEV [6, 7]; (2) RPSA (ribosomal protein SA) was shown to be involved in binding with 40S ribosomal subunits, as well as in the organization of a binding site for the specific structural element of hepatitis C virus RNA (IRES element), on the 40S [8]. Revealing highly connected cluster involving ribosomal proteins is in good agreement with the notion that translation is an essential step in the TBEV life cycle [9].

At the next step we prioritized the proteins from four previously described groups of genes/proteins (Sublists A, D, C, and D) according to the number of neighbors in the network. The top proteins were: (1) for Sublist A - RPSA, CASK, and TJP2 (35, 6, and 5 neighbors respectively); (2) for Sublist B - ATF6, ERN1, and ICAM1 (7, 5, and 4 neighbors); (3) for

Sublist C - OAS2, OAS3, and TLR3 (10, 9, and 4 neighbors); (4) for Sublist D - UBC, RPS27A, and RPS3 (61, 38, and 37 neighbors).

Conclusion: A catalog comprising 35 human genes for which there are indications of their direct or indirect relevance to response to TBEV infection was designed, and the network involving physical interactions between proteins was constructed. Analysis of the human protein-protein interaction network involved in response to TBEV infection revealed extremely dense cluster formed by ribosomal proteins and identified proteins with maximal number of neighborhoods (RPSA, ATF6, OAS2, UBC, etc.). We propose to keep in mind these proteins and genes as potential candidates for investigating the genetic factors predisposing to severe forms of the diseases caused by TBEV and as potential drug targets in tick-borne encephalitis treatment.

This work was supported by the Russian Science Foundation (project no. 16-15-00127)

References

1. K.Werme et al. (2008) Tick-borne encephalitis virus NS5 associates with membrane protein scribble and impairs interferon-stimulated JAK-STAT signaling, *Cell Microbiol*, **10**:696-712.

2. N.S.Yudin et al. (2017) Genetic predisposition to diseases caused by flaviviruses, *Mol Biol (Mosk)*, **51**:

3. D.Szklarczyk, et al. (2011) The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored, *Nucleic Acids Res*, **39**:D561–D568.

4. M.E.Smoot et al. (2011) Cytoscape 2.8: new features for data integration and network visualization, *Bioinformatics* **27**:431-2.

5. G.D.Bader, C.W.Hogue (2003) An automated method for finding molecular complexes in large protein interaction networks, *BMC Bioinformatics*, **4**:2

6. B.N.Zaitsev et al. (2014) Force-induced globule-coil transition in laminin binding protein and its role for viral-cell membrane fusion, *J Mol Recognit*, **27**:727-38

7. E.V.Protopopova et al. (1997) Isolation of a cellular receptor for tick-borne encephalitis virus using anti-idiotypic antibodies, *Vopr Virusol*, **42**:264-8

8. A.A.Malygin et al. (2009) C-terminal fragment of human laminin-binding protein contains a receptor domain for venezuelan equine encephalitis and tick-borne encephalitis viruses, *Biochemistry (Mosc)*, **74**:1328-36

9. H.Rouha et al. (2011) Mutational analysis of three predicted 5'-proximal stem-loop structures in the genome of tick-borne encephalitis virus indicates different roles in RNA replication and translation, *Virology*, **417**:79-86