doi 10.18699/bgrs2024-3.2-25

An analysis of virus-host interaction: case study for HIV-infection

Tarasova O.*, Biziukova N., Rudik A., Ivanov S., Poroikov V.
Institute of Biomedical Chemistry, Moscow, Russia
* *olga.a.tarasova@gmail.com*

Key words: virus-host interactions; HIV; bioinformatics; systems biology; data analysis; knowledgebase

Motivation and Aim: Analysis of virus-host interactions is essential for deep understanding of the molecular interplay that is caused by a viral infection. The large amount of experimental data available on people with viral infections can be used to develop computational models of virus-host interactions [1]. Such models can help to determine the mechanisms of viral infection progression and the efficacy of antiviral therapy, in order to improve existing and develop new treatment strategies including those that take into account the individual characteristics of a patient. Systems biology and bioinformatics approaches allow us to analyze the entire set of experimental data in silico and generate new knowledge.
Methods and Algorithms: The developed approach includes a set of methods for automated selection of relevant data from different sources and their integration. Based on the retrieved data, we performed a search for key regulators of HIV infection progression and carried out their experimental validation. We have developed algorithms to estimate the effectiveness of antiretroviral therapy, taking into account i.e. amino acid and nucleotide sequences of the virus. In particular, we used the in-house developed algorithms of named entity recognition and extraction of associations between extracted entities [2, 3] to perform the text and data mining of virus-host interactions. The developed algorithms of machine-learning based HIV drug resistance prediction [4, 5] was used for building web-based models aimed at drug resistance prediction [5]. Transcriptome-based bioinformatics analysis of peripheral blood mononuclear cells (PBMC) obtained from HIV-infected people (PLWH) with various viral infection progression and viral infection control allowed us to identify the molecular mechanisms of HIV infection progression that are influenced by individual characteristics of patient's reply on the infection [1]. An analysis of relationship between differential expression in genes of patients with HIV-infection and the therapeutic outcome was compared with experimental data resulting from clinical study on eleven people observed before starting antiretroviral therapy and 24 weeks after.
Results: The developed methods are integrated into the hiv-host platform (https://www.way2drug.com/hiv-host), which includes (1) databases on the interaction of HIV macromolecules with the human body and information on the history of HIV therapy, and (2) models and web applications for estimating viral drug resistance and efficacy of antiretroviral therapy, synergy of antiretroviral drugs, models for identifying master regulators of infection progression.
As a result of gene enrichement analysis performed for the samples obtained from PLWH, it was found that the elevated expression of proapoptotic and innate immunity genes before taking antiretroviral therapy can be associated with higher effectiveness of antiretroviral therapy. Some of observations found as a result of text and data mining

analysis [6] and bioinformatics evaluation of gene expression in PBMC of PLWH were confirmed in the clinical study.

Based on the developed methodology we expanded the data analysis on a set of viruses that included Hepatitis B and C viruses, SARS-CoV-2, influenza A and B, and herpes simplex virus. We collected data on interaction for three interactions on these viruses with (1) the host (human body), (2) with potential antivirals and (3) on the interaction of the host proteins with chemical compounds. Based on the performed data analysis we created freely available knowledgebase on the interaction of chemical compounds with viral proteins and their host targets.

Conclusions: The developed platform integrates information about the interaction of viruses and the host at the molecular level and provides the opportunity for further bioinformatic searches in order to obtain new knowledge about the pathogenesis of viral infections and the rational design of antiviral drugs.

Funding: The study is supported by the Program for Basic Research in the Russian Federation for a long-term period 2021–2030 (project No. 122030100170-5).

References

1. Ivanov S.M., Tarasova O.A., Poroikov V.V. Transcriptome-based analysis of human peripheral blood reveals regulators of immune response in different viral infections. *Front Immunol.* 2023;14:1199482. doi 10.3389/fimmu.2023.1199482
2. Tarasova O.A., Rudik A.V., Biziukova N.Y. et al. Chemical named entity recognition in the texts of scientific publications using the naïve Bayes classifier approach. *J Cheminform.* 2022;14:55. doi 10.1186/s13321-022-00633-4
3. Biziukova N. et al. Automated Extraction of Information From Texts of Scientific Publications: Insights Into HIV Treatment Strategies. *Front Genet.* 2020;11:618862. doi 10.3389/fgene.2020.618862
4. Tarasova O. et al. A Computational Approach for the Prediction of HIV Resistance Based on Amino Acid and Nucleotide Descriptors. *Molecules.* 2018;23(11):2751. doi 10.3390/molecules23112751
5. Paremskaia A.I. et al. Web Service for HIV Drug Resistance Prediction Based on Analysis of Amino Acid Substitutions in Main Drug Targets. *Viruses.* 2023;15(11):2245. doi 10.3390/v15112245
6. Tarasova O. et al. Identification of Molecular Mechanisms Involved in Viral Infection Progression Based on Text Mining: Case Study for HIV Infection. *Int J Mol Sci.* 2023;24(2):1465. doi 10.3390/ijms24021465