

Translocatome: a novel tool for the functional analysis of protein translocation between cellular organelles

Authors: Levente Dobronyi^{1*}, Péter Mendik¹, Péter Csermely¹, Dániel V. Veres^{1,2}

¹ Semmelweis University, Department of Medical Chemistry, Molecular Biology and Pathobiochemistry, Faculty of Medicine, Hungary

² Turbine Ltd.

* dlevi1717@gmail.com

Introduction: Localization of proteins in subcellular compartments has a key role in cellular regulation and function. In eukaryotic cells, organelles are well distinguished cellular components with different microenvironments and specific molecular interactions. Translocation is considered as a process when functionally active proteins change their subcellular localization in a regulated manner. Although, how individual e.g. transcription factors translocate from the cytoplasm to the nucleus resulting in an alteration in cellular behaviour is well known, systematic analysis of this phenomenon is still missing.

Objectives: Our goal is to create a database that enables to predict the probability of protein translocation with a machine learning algorithm. With the help of this database, the role of translocation in certain diseases could be better understood. Further investigation by dynamic network simulations can be used on a Glioblastoma Multiforme specific cell model to find correlation between translocating proteins and malignancy.

Materials and Methods: Data are gathered from related articles by manual curation and from our in-house developed ComPPI database (<http://comppi.linkgroup.hu>), the GO database [<http://geneontology.org>]. Relevant localization, structural and regulation related properties of the proteins are collected.

Results and Discussion: We have gathered detailed information for more than 100 human translocating proteins by manual curation. From the ComPPI interaction data for more than 13000 proteins have been imported as well and made into a protein-protein interaction (PPI) network that enables us to predict the probability of protein translocation. Further properties (such as information on biological processes from GO) are annotated to the proteins for more precise predictions. Utilizing this data, the machine learning algorithm is able to adjust translocation probability scores to the proteins based on the identified differences between a positive (translocating) and a negative (not translocating) training set.

Conclusion: With the help of our database, the role of the translocation can be better understood and new drug targets can be discovered based on the systematic analysis of the phenomenon.

References:

Veres, D. V. et al. ComPPI: a cellular compartment-specific database for protein-protein interaction network analysis. *Nucleic Acids Res.* 43, D485–D493 (2015).

Szalay, K. Z., Nussinov, R. & Csermely, P. Attractor structures of signaling networks: Consequences of different conformational barcode dynamics and their relations to network-based drug design. *Mol. Inform.* 33, 463–468 (2014).

The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.