EVALUATING THE EXPANSION OF THE EXPERIMENTALLY DETERMINED HUMAN PROTEIN INTERACTOME BASED ON THE PICKLE META-DATABASE

G.N. Dimitrakopoulos[1,2], A. Gioutlakis[1,2], M.I. Klapa[2] and N.K. Moschonas[1,2]

[1] General Biology Laboratory, Department of Medicine, University of Patras, Patras, Greece

[2] Metabolic Engineering & Systems Biology Laboratory, FORTH/ICE-HT, Patras, Greece

ABSTRACT

Introduction: The Protein InteraCtion KnowLedgebasE (PICKLE; www.pickle.gr) is a publicly available meta-database of the experimentally-determined direct protein-protein interaction (PPI) network in human developed by our group[1,2]. PICKLE integrates experimental PPI data from five source databases on the basis of the genetic information ontology network of the UniProtKB/Swiss-Prot-defined reviewed human complete proteome (RHCP), eliminating thus the need for their *a priori* normalization to a selected level of genetic information (i.e. protein (UniProt), gene or mRNA). Its ontological way of PPI data integration from various sources using RHCP as a standardized reference node set, constitutes PICKLE the only PPI meta-database today that enables the evaluation of the expansion of the experimentally determined human protein interactome at both the protein and gene levels. This evaluation, which is the objective of the present study, can be achieved by comparing PICKLE successive releases, which reflect the respective changes in the biological and PPI primary datasets.

Methods: The default ("cross-checked") direct PPI networks at the UniProt level of PICKLE releases 2.1, 2.2 and 2.3, integrating primary biological and PPI data from Feb 2015, Dec 2017 and Nov 2018, respectively, were compared in this study. These networks exclude experimental PPIs that are of low confidence of being direct based on the available experimental evidence[1]. Network analysis was performed using the Cytoscape software[3].

Results and Discussion: The current experimentally determined direct PPI network in human, as reconstructed by PICKLE, comprises 178306 PPIs between 15823 UniProt IDs, supported by 39603 publications. The PICKLE network remains double in number of PPIs from any of the largest primary PPI datasets (BioGRID or IntAct), emphasizing the need for integration of source databases for the reliable reconstruction of the currently known human protein interactome. Interestingly, the increase in the experimentally supported PPIs since PICKLE 2.1 is at least three-fold larger than the respective change in the number of supporting references, reflecting the fact that in the last years PPIs are largely determined by high-throughput experiments. We observed a three times smaller increase in the number of newly added protein-nodes compared to the newly determined PPIs, with most of the latter enriching the interactions of already existing proteins in the human interactome. In addition, as we had predicted[2], the vast majority of the newly added proteins have 1-4 interactions. The interactome becomes denser, further supporting the overlapping between biological processes underlying its robustness and flexibility, the disease co-morbidity and the efforts for drug repurposing. However, its further expansion over the 15% of RHCP proteins that remain with no known protein interactions seems to require more targeted experiments with respect to the proteins to be investigated and the biological processes to which they participate.

Conclusions: Our observations reinforce our previous statement[2] that the structure of the human protein interactome with respect to its hubs has already been largely defined. Identifying its dynamics based on developmental stage-, tissue- and cell-specific omic data in health and disease is the next challenge for network biology and medicine.

Acknowledgements: This work is supported by the NSRF 2014-2020 ELIXIR-GR (MIS 5002780), ΒΙΤΑΔ-ΔΕ (MIS 5002469), EATRIS-GR (MIS 5028091) and INSPIRED (5002550) projects, co-financed by Greece and the European Union (European Regional Development Fund).

REFERENCES

[1] Gioutlakis A, Klapa MI, Moschonas NK. (2017). *PLoS ONE*, 12: e0186039.

[2] Klapa MI, Tsafou K, Theodoridis E, Tsakalidis A, Moschonas NK. (2013). *BMC Syst. Biol.* 7:96.

[3] Shannon P, Markiel A, Ozier O, Baliga NS,..., Ideker T. (2003). *Genome Research*, 13:2498-2504.