A Network Approach to Cancer: Clustering Cell Lines by Gene Dependency Rank Distances

Jane Adams[1], Celestin Coquide[2], Diana Garcia[3], Rodrigo Migueles Ramirez[4]

[1] Vermont Complex Systems Center, University of Vermont, United States [2] Institut UTINAM, OSU THETA, University of Bourgogne Franche-Comté, France, [3] National Institute of Genomics Medicine, Mexico City, Mexico, [4] Quantitative Life Sciences Interfaculty Program, McGill University, Montreal, Canada

Specificity is one of the biggest challenges of current cancer treatments. Evidence has shown that while cancerous cells bear alterations that confer them malignant phenotypes, these key alterations can be used to target cancer cells with increased specificity. Here, we hypothesize that a systematic evaluation and comparison of these key genes across different cancerous cell lines can allow us to detect unexpected similarities in dependency profiles thus providing insight for combinatory therapies and drug repurposing. Using the Achilles Dataset from the DepMap project, which evaluates cell line-specific per-gene dependency profiles, we compare the dependency probability rankings across 625 different cancer cell lines by calculating Spearman's rank correlation coefficients. The use of a backbone sparsification method resulted in a network composed by 126 nodes and 183 links distributed in 15 components. About half of the components are composed of heterogeneous cancer types and 80% of the cell lines (including aggressive cancers such as lung and pancreatic cancer) are preferentially linked with cell lines of other types, as measured by balanced homophily. To detect thematic cellular functions associated with these key gene dependencies, communities were detected using InfoMap, and the pooled 5% of the highest ranked gene sets per community were mapped to enriched biological processes using Gene Ontology. Communities shared enriched biological processes associated with cancer related processes such as cell division, regulation of the immune response and key signaling pathways. The biggest community has 35 cell lines of 9 different cancer types, where 77% correspond to pancreatic or lung cancer, and it is the only community with DNA repair as an enriched GO term. These results suggest that this approach could allow us to identify key genes and biological processes to which different cancer cell types are similarly sensitive to, thus providing hints for combinatorial therapies and novel applications for existing therapeutic strategies.

[1] Tsherniak, A. et al. *Defining a Cancer Dependency Map.* Cell 170, 564-576.e16 (2017). DOI: 10.1016/j.cell.2017.06.010.

[2] Serrano, M. et al. *Extracting the Multiscale Backbone of Complex Weighted Networks.* Proceedings of the National Academy of Sciences Apr 2009, 106 (16) 6483-6488; DOI: 10.1073/pnas.0808904106

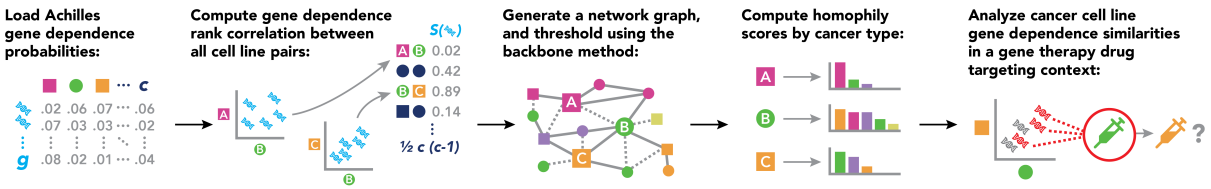

Figure 1: Representative graphic describing the data analysis performed to generate hypotheses about heterogeneity in genetic vulnerability networks of cancer cell lines