Differential network analysis

Background

When a living and self-organizing system becomes complex, simplifying strategies are needed to ensure communication between the system elements, stability of the system, and organization but also system versatility and growing at optimal energy utilization. Modularity is such a simplifying strategy.

Genes are linked to networks; gene network analysis compares networks under different conditions and identifies associations. Whereas in differential network analysis, the focus is on the underlying differential gene regulatory network topology and the modular network structure. (Langfelder & Horvath, 2008, & Zhang, 2008). To identify modules is essential for understanding the whole network architecture. (Yip et al., 2006)

The Spatial transcriptomics (ST) technique (Stål et al., 2016) makes it possible to investigate gene expression under different conditions like healthy, inflamed or cancerous. Expression data created with this novel technique will be used to study network structure and modularity under specific and changing conditions in tissue samples of a prostate cancer patient.

Project aim

The project aims to identify modules (sets of tightly interconnected genes) in gene networks of prostate cancer. Further, modular differences will be studied in healthy and cancerous regions of tissue samples to reveal changed genes and pathways.

Intermediate steps/goals

1. Applying correlation analysis on groups of ST spots within cancerous and healthy regions to identify network modules
2. Comparing the network structures and identifying modules of cancerous and healthy regions
3. Performing functional analysis of highly differential modules

Data and Methods

Gene expression data analysed with ST is the source data for the project. More detailed, the data comprises the expression in approximately 1,000 spots (10 to 100 cells in each spot) of an annotated human prostate tissue sample. This results in very detailed picture of transcription levels in cancerous, healthy and PIN regions.

Based on the expression data of a specific region, a gene network will be inferred. Applying co-expression analysis and clustering will reveal modules. Co-expressed means that the expressions of two genes are significantly correlated; modules are defined as the branches of the hierarchical clustering tree (Figure 1).

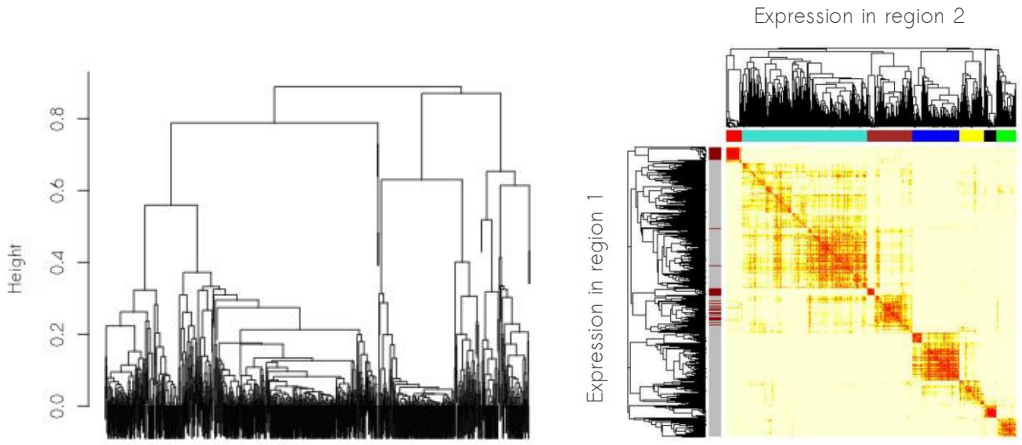

Figure1: Hierarchical tree with branches showing modules and heatmap showing modules and co-regulated genes (Yip & Horvath, 2006)

To compare the modules in two networks inferred from the expression in different regions, the degree of similarity or dissimilarity has to be measured; this will be done with a link overlap measure.
Finally, significantly different modules will be investigated concerning their biological function and compared with known pathways.
Different statistical, phylogeny and clustering methods will be applied; additionally, several databases (e.g. FunCoup (Schmitt et al., 2013)) will be used.
To realise the project, Python (Van Rossum, 2003) and R (R Core Team, 2017) will be utilised.

References

Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics, 9(1), 559.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Schmitt, T., Ogris, C., & Sonnhammer, E. L. (2013). FunCoup 3.0: database of genome-wide functional coupling networks. Nucleic acids research, 42(D1), D380-D388.

Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., ... & Mollbrink, A. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science, 353(6294), 78-82.

Van Rossum, G., & Drake, F. L. (2003). Python language reference manual (p. 144). Network Theory.

Yip, A. M., & Horvath, S. (2006). The Generalized Topological Overlap Matrix for Detecting Modules in Gene Networks. In BIOCOMP (pp. 451-457).

Zhang, B., Li, H., Riggins, R. B., Zhan, M., Xuan, J., Zhang, Z., ... & Wang, Y. (2008). Differential dependency network analysis to identify condition-specific topological changes in biological networks. Bioinformatics, 25(4), 526-532.