Source-Target proximity analysis in protein networks using random walks with restart; applications to cancer drug resistance prediction from single-cell data

Frédéric Cazals* Jérémie Roux† Alain Jean-Marie* Dorian Mazauric*
Guilherme Sales Santa Cruz*

22 November 2019

Summary

A central problem in the multi-omics era is to bridge the gap between genes/proteins of heterogeneous expression (call them sources) and proteins already known to be involved in a specific cell response signaling pathway (call them targets). Insights on this problem are of special interest to unveil the molecular basis of individual cell response that may vary across sister cells [1].

This problem is a particular case of proximity analysis in networks. Random walks on graphs proved of central importance to investigate such questions, with applications to spectral clustering [2], dimensionality reduction [3], community detection, page rank definition, etc.

In this work, we introduce a novel approach to source-target proximity analysis in protein-protein-interaction networks (PPINs). Consider a PPIN graph whose nodes contain the sources S and targets T, and whose edges code interactions between sources, targets, and additional molecules. We code the proximity between a source s and all targets T using the stationary distribution of a suitable random walk with restart. The size of the PPINs considered precluding exact solutions based on linear algebra, we compute stationary distributions numerically using the marmoteCore library [4].

We have previously shown that sister cells in isogenic population treated with Dulanermin (TRAIL), can commit differentially to cell death based on their response dynamics: some cells respond, other resists to the therapy and regrow [5]. We use our framework to analyze single-cell transcriptomic data, obtained from the same population of TRAIL treated HeLa cells to investigate the molecular determinants of tumor cell resistance to Dulanermin. More specifically, we study the relationship between sources (differentially expressed proteins) and targets (proteins involved in the cell death pathway), with applications to drug resistance prediction in tumor cells.

All methods are currently being integrated to the Structural Bioinformatics Library (`https://sbl.inria.fr/applications/`), and will be made available to the community.

Bibliographical references

[1] Navin, N. et al. Tumour evolution inferred by single-cell sequencing. *Nature* 472, 90–94 (2011).

[2] Von Luxburg, U., A tutorial on spectral clustering, *Statistics and Computing*, 2007.

[3] Lafon et al, Diffusion maps and coarse-graining, *IEEE Trans Pattern Anal Mach Intell* 28, 1393–1403 (2006).

[4] A. Jean-Marie et al, marmoteCore: a software platform for Markov modeling, ROADEF, 2016.

[5] Roux et al, Fractional killing arises from cell-to-cell variability in overcoming a caspase activity threshold. *Molecular Systems Biology* 11, 803 (2015).

[1]Inria Sophia Antipolis - Méditerranée
[2]IRCAN - Université Côte d'Azur