THE BROAD PHENOTYPE-SPECIFIC APPLICATIONS OF THE NETWORK-BASED SWIM TOOL

Federica Conte [1] and Paola Paci [2]

[1] Institute for Systems Analysis and Computer Science "A. Ruberti" (IASI) National Research Council (CNR) of Rome, Rome, Italy, (e-mail: federica.conte@iasi.cnr.it)

[2] Department of Computer, Control and Management Engineering, Sapienza University of Rome, Rome, Italy, and Karolinska Institutet, 17177 Stockholm, Sweden (e-mail: paci@diag.uniroma1.it)

ABSTRACT: SWIM is a recently developed network-based tool that fulfils the criteria of the new quickly emerging field of Network Medicine in finding disease-associated genes, called switch genes. Here, a brief summary of the promising results obtained by applying SWIM in different biological contexts is presented.

KEYWORDS: network medicine, network theory, disease genes

1 Introduction

Recently, we developed a new promising methodology, called SWIM (SWItch Miner), which integrates different network-based methods to analyse the correlation network arising from large-scale gene expression data Paci *et al.*, 2017. Considering the topological properties of the nodes and assessing their functional roles according to their ability to convey information within and between modules in the network, SWIM identifies a small pool of genes (called switch genes) that are associated with intriguing patterns of molecular co-abundance and play a crucial role in the observed phenotype.

The phenotype-specific applications of SWIM are broad and include the identification of switch genes in grapevine berry maturation Palumbo *et al.*, 2014, in human cancers Paci *et al.*, 2017, including glioblastoma multiforme (GBM) Fiscon *et al.*, 2018 and in chronic obstructive pulmonary disease (COPD) Paci *et al.*, 2020. More recently, SWIM has been applied within the framework of Network Medicine to study the interplay between switch genes and human diseases in the human interactome (i.e., the cellular network of all physical molecular interactions) Paci *et al.*, 2021.

In the following, a detailed description of the more recent applications of SWIM to complex diseases is provided.

2 Methods

SWIM is a freely downloadable network-based tool, developed both in MAT-LAB Paci *et al.*, 2017 and in R language Paci & Fiscon, 2022, which predicts important (switch) genes that are strongly associated with drastic changes in cell phenotype. SWIM encompasses several steps detailed in Paci *et al.*, 2017.

3 Results and Discussions

3.1 Glioblastoma

Glioblastoma is the most aggressive and frequent brain tumour, with a median survival time of 12–15 months from diagnosis Young *et al.*, 2015. This tumour is resistant to the standard therapies and its aggressiveness seems to be due to the presence of cancer stem-like cells Gimple *et al.*, 2019. Thus, targeting cancer stem-like cells could pave the way for new therapeutic strategies.

A recent study identified 19 neurodevelopmental transcription factors (TFs) that are selectively expressed in glioblastoma stem-like cells to maintain their stem-like phenotype and prevent differentiation Suva *et al.*, 2014. A subset of only four of them (named 4-core TFs), SOX2, OLIG2, POU3F2, and SALL2, has been shown to be sufficient to fully reprogram differentiated cells into glioblastoma stem-like cells Suva *et al.*, 2014.

In order to identify switch genes related to the stem-like phenotype, SWIM was applied to GBM dataset of Suva *et al.*, 2014 and then the further dataset of Schulte *et al.*, 2011 was used to validate the results Fiscon *et al.*, 2018. Among the common switch genes obtained by running SWIM on the these two GBM datasets, there is FOSL1. It is up-regulated in differentiated glioblastoma cells and this up-regulation highly correlates with the over-expression of genes involved in cell-cell communications. It is down-regulated in stem-like cells and this down-regulation highly correlates with the up-regulation of the 4-core of TFs. To investigate a possible co-regulation of the 4-core of TFs, their promoter regions were inspected to search for enriched motifs and they were found to harbour a consensus binding site for FOSL1.

Altogether these findings suggest FOSL1 as possible therapeutic biomarker of glioblastoma, which could promote the differentiation of cancer stem-like cells by repressing the 4-core TFs. This hypothesis has been partially experimentally validated in Pecce *et al.*, 2021.

3.2 COPD

COPD is a heterogeneous and complex syndrome influenced by both genetic and environmental determinants, and is one of the main causes of morbidity and mortality worldwide.

By applying SWIM on COPD Paci *et al.*, 2020, the correlation network turned out to be formed by three well-characterised modules: i) one populated by switch genes, all up-regulated in COPD cases and involved in COPD-related pathways, like B cell receptor signalling pathway; ii) one populated by negative interactors of switch genes, down-regulated in COPD cases, including well-known GWAS genes like AGER and CAVIN1; iii) one populated by well-recognised immune signature genes, all up-regulated in COPD cases. Switch genes appear to form localised connected subnetworks displaying an intriguingly common pattern of up-regulation in COPD cases compared with controls. A more sophisticated analysis revealed that they were not only topologically related, but also functionally relevant to the observed phenotype as witnessed by their enrichment in the regulation of inflammatory and immune responses. Finally, SWIM was applied on another severe lung disease with an inflammatory component, i.e., the acute respiratory distress syndrome (ARDS), demonstrating that, even though different diseases can share similar endophenotypes, the molecular network determinants responsible for them are disease-specific.

3.3 Network Medicine

Network Medicine is a new emerging paradigm in medicine, where disease proteins are assumed not to be randomly scattered, but agglomerate in specific regions of the molecular interactome, suggesting the existence of specific disease network modules for each disease Barabási *et al.*, 2011. To quantify the interplay between switch genes and human diseases in the human interactome, the results obtained by the pan-cancer Paci *et al.*, 2017 and COPD Paci *et al.*, 2020 SWIM-based analysis were complemented with the application of SWIM tool on two cardiac disorders (i.e., ischemic and non-ischemic cardiomyopathy) and on Alzheimer's disease (AD) Paci *et al.*, 2021. Switch genes associated with specific disorders were found to be not randomly scattered but they form localised connected subnetworks. These subnetworks overlap between similar diseases (like cancers or cardiac disorders) and are situated in different neighbourhoods for pathologically distinct phenotypes (like AD and COPD), showing a direct relation between the pathobiological similarity of diseases and their relative distance in the human interactome. Finally, the first SWIM-

informed Human Disease Network was built, where nodes correspond to distinct disorders and a link occurs between two diseases if they share a substantial number of switch genes. Clustering of nodes belonging to the same disease class means that similar pathophenotypes have a higher probability of sharing switch genes than do pathophenotypes that belong to different disease classes. These findings support the hypothesis that SWIM-based correlation network, when integrated with an interactome-based network analysis, not only identifies novel candidate disease genes, but also may offer useful tool by which to elucidate the molecular underpinnings of human disease and reveal commonalities between seemingly unrelated diseases.

References

BARABÁSI, ALBERT-LÁSZLÓ, GULBAHCE, NATALI, & LOSCALZO, JOSEPH. 2011. Network medicine: a network-based approach to human disease. *Nature Reviews. Genetics*, **12**(1), 56–68.

FISCON, GIULIA, CONTE, FEDERICA, LICURSI, VALERIO, NASI, SERGIO, & PACI, PAOLA. 2018. Computational identification of specific genes for glioblastoma stem-like cells identity. *Scientific Reports*, **8**(1), 7769.

GIMPLE, RYAN C., BHARGAVA, SHRUTI, DIXIT, DEOBRAT, & RICH, JEREMY N. 2019. Glioblastoma stem cells: lessons from the tumor hierarchy in a lethal cancer. *Genes & Development*, **33**(11-12), 591–609.

PACI, PAOLA, & FISCON, GIULIA. 2022. SWIMmeR: an R-based software to unveiling crucial nodes in complex biological networks. *Bioinformatics*, **38**(2), 586–588.

PACI, PAOLA, COLOMBO, TERESA, FISCON, GIULIA, GURTNER, AYMONE, PAVESI, GIULIO, & FARINA, LORENZO. 2017. SWIM: a computational tool to unveiling crucial nodes in complex biological networks. *Scientific Reports*, **7**(Mar.), srep44797.

PACI, PAOLA, FISCON, GIULIA, CONTE, FEDERICA, LICURSI, VALERIO, & OTHERS. 2020. Integrated transcriptomic correlation network analysis identifies COPD molecular determinants. *Scientific Reports*, **10**(1), 1–18. Number: 1 Publisher: Nature Publishing Group.

PACI, PAOLA, FISCON, GIULIA, CONTE, FEDERICA, WANG, RUI-SHENG, FARINA, LORENZO, & LOSCALZO, JOSEPH. 2021. Gene co-expression in the interactome: moving from correlation toward causation via an integrated approach to disease module discovery. *npj Systems Biology and Applications*, **7**(1), 1–11. Number: 1 Publisher: Nature Publishing Group.

PALUMBO, MARIA CONCETTA, ZENONI, SARA, FASOLI, MARIANNA, MASSONNET, MÉLANIE, FARINA, LORENZO, CASTIGLIONE, FILIPPO, PEZZOTTI, MARIO, & PACI, PAOLA. 2014. Integrated Network Analysis Identifies Fight-Club Nodes as a Class of Hubs Encompassing Key Putative Switch Genes That Induce Major Transcriptome Reprogramming during Grapevine Development. *The Plant Cell Online*, Dec., tpc.114.133710.

PECCE, VALERIA, VERRIENTI, ANTONELLA, FISCON, GIULIA, SPONZIELLO, MARIALUISA, CONTE, FEDERICA, DURANTE, COSIMO, FARINA, LORENZO, FILETTI, SEBASTIANO, & PACI, PAOLA. 2021. FOSL1 in the stemness mechanism: An in vitro study. *Scientific Reports*.

SCHULTE, ALEXANDER, GÜNTHER, HAUKE S., PHILLIPS, HEIDI S., & OTHERS. 2011. A distinct subset of glioma cell lines with stem cell-like properties reflects the transcriptional phenotype of glioblastomas and overexpresses CXCR4 as therapeutic target. *Glia*, **59**(4), 590–602.

SUVA, MARIO L, RHEINBAY, ESTHER, GILLESPIE, SHAWN M, & OTHERS. 2014. Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. *Cell*, **157**(3), 580–594.

YOUNG, RICHARD M, JAMSHIDI, ARIA, DAVIS, GREGORY, & SHERMAN, JONATHAN H. 2015. Current trends in the surgical management and treatment of adult glioblastoma. *Annals of translational medicine*, **3**(9).