Interactome–regulome–transcriptome integrative approach: a mean to disclose cancer stem cells regulatory circuits

C Rioualen[1,2,3,4], R El-Helou[1,2,3,4], Q Da Costa[1,2,3,4], E Charafe-Jauffret[1,2,3,4], C Ginestier[1,2,3,4] and G Bidaut [1,2,3,4]
contact: claire.rioualen@inserm.fr

Centre de Recherche en Cancérologie de Marseille
1. Inserm U1068
2. CNRS UMR7258
3. Aix-Marseille Université
4. Institut Paoli-Calmettes

www.iti.sourceforge.net

Breast cancer is the deadliest cancer in women worldwide[1]. One rising hypothesis is that **cancer stem cells** (CSC) could explain relapses, by **resisting to conventional treatments** and being able to generate metastases thanks to **self-renewal and differentiation properties**[2].

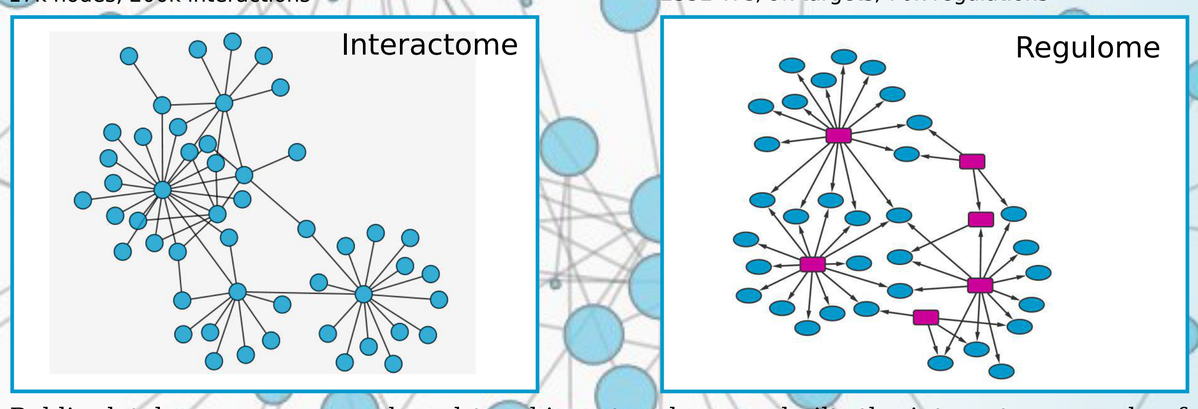
An miRNA screening revealed miR-600 could be a potential regulator of CSCs. We decided to investigate the effect of miR-600 knock up and kock down on CSCs, and to incorporate a **network-based approach** in order to identify key signaling pathways.

Input data

Networks building

HPRD, MINT, INTAct, DIP, Proteinpedia, I2D , BioGrid
17k nodes, 200k interactions

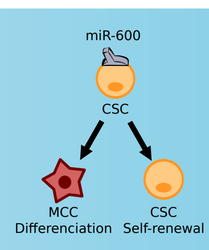
TRANSFAC, TRED, ITFP, PAZAR, ORegAnno
2352 TFs, 9k targets, 70k regulations



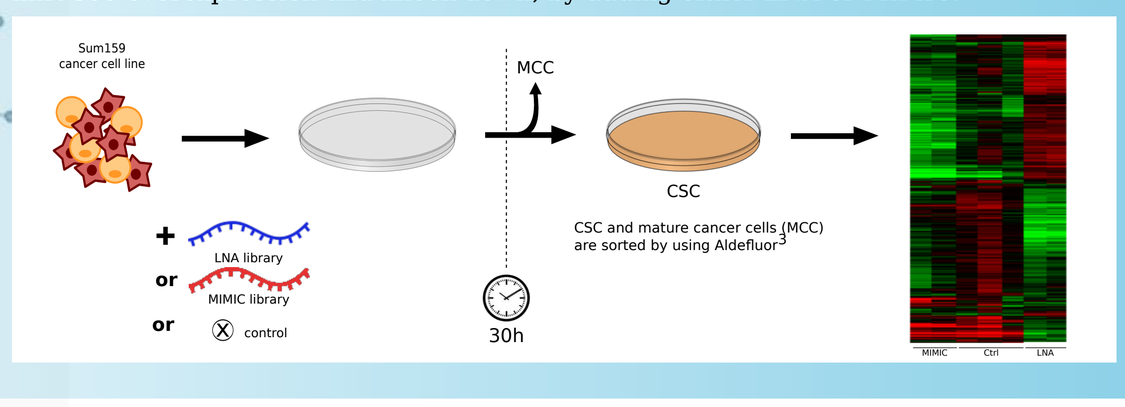
Public databases were parsed, and two big networks were built: the interactome, made of protein-protein interactions, and the regulome, made of TF-target relationships. Above are subsets of these networks.

Transcriptome data

After a genome-wide miRNA screening, miR-600 was identified as a potential "switcher" for CSCs: its activation by MIMIC and its inhibition by LNA could have a mirror effect on CSC pathways, by favoring their differentiation and preventing their self-renewal at the same time.



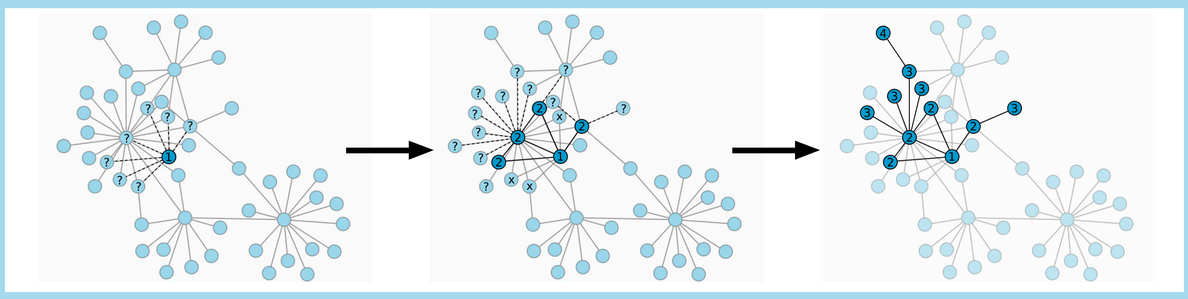
Transcriptome data was generated by measuring gene expression in CSCs after miR-600 overexpression and knock-down, by adding either LNA or MIMIC.



Data integration & Networks analysis

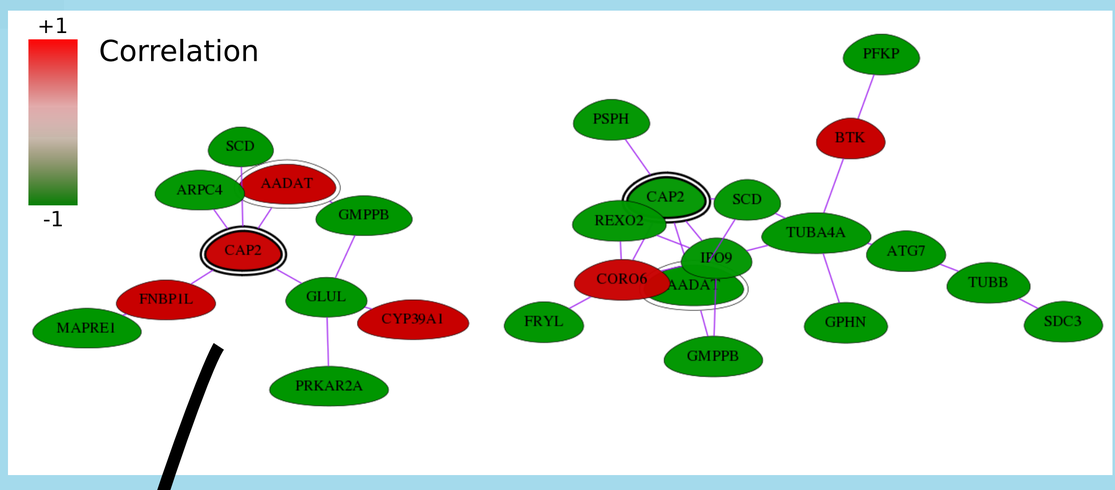
Subnetworks detection algorithm

A **subnetwork is recursively aggregated** around a "seed", a gene from the network, by using the ITI algorithm[4,5]. Its score is calculated **by averaging correlations between genes expressions** (LNA vs control, MIMIC vs control).
Neighbouring nodes are added if they improve the score of the subnetwork. This procedure is repeated for each gene of the two networks (interactome and regulome).
Resulting **subnetworks are statistically validated** by randomizing interaction data, expression data and subnetworks interactions.
The procedure is repeated separately in the interactome and the regulome.



Switchers identification

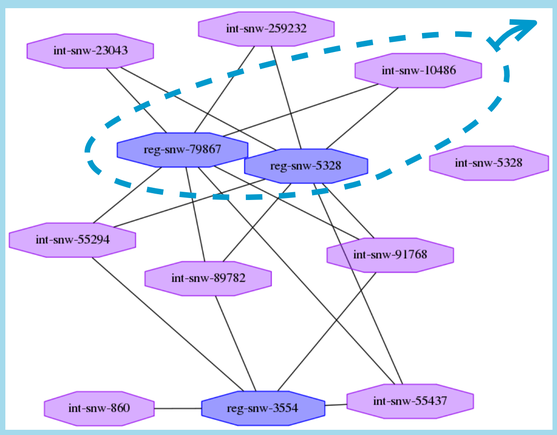
42 subetworks are found ine the 4 different conditions (LNA-interactome, LNA-regulome, MIMIC-interactome, MIMIC-regulome).
35 genes were found to act like switchers in the subnetworks (double circles): they were both being strongly over-expressed in MIMIC and under-expressed in LNA, confirming the biological experiments.

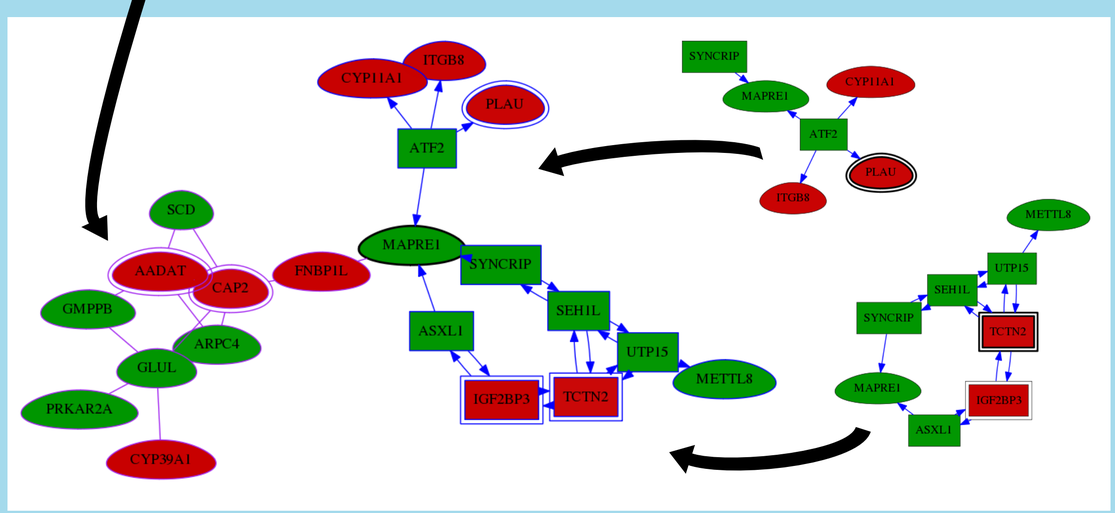


Subnetworks integration

In order to **visualize pathways of regulation** that involve several levels of regulation, we integrated the interactome subnetworks (groups of interacting proteins) with regulome subnetworks (transcription factors and targets).

Most subnetworks interact with each other, showing a possible implication in common pathways regulating cancer stem cells.





The integration sheds light on potential **pathways that are connecting switching genes,** thus regulating CSC differentiation and proliferation. It could also enable the discovery of new switchers, whose expression is only slightly affected by miR-600 overexpression or knock-down.

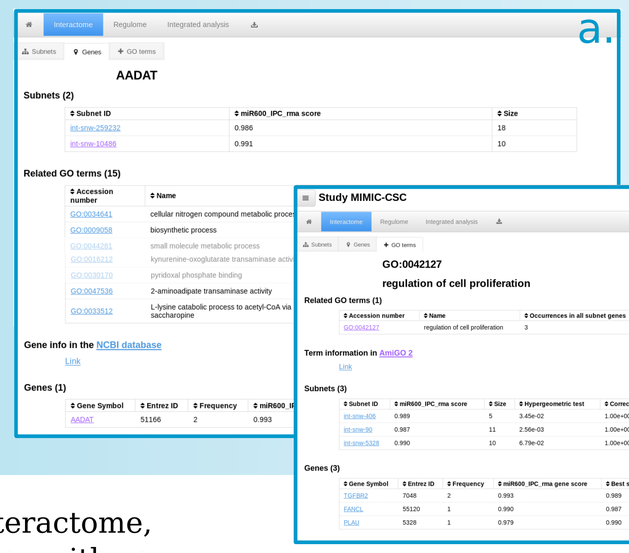
Targeting a pathway, rather than several genes separately, could prove to be more efficient and less money-consuming.
A Gene Ontology enrichment search revealed the presence of several terms related to **cell proliferation**.

Online database of subnetworks

Our pipeline includes the possibility of generating full **HTML reports**.

Detected subnetworks are stored and can be investigated gene by gene (a). This visualization also includes all the information on Gene Ontology terms enriched in the subnetworks (b). Each subnetwork has its own page, compiling all the related information (c).

All the data can be downloaded from the site in flat format, and subnetwork files are formatted in *.nnf format, thus enabling their analysis in Cytoscape.



We developed a new **multi-level integrative approach** mixing interactome, regulome, transcriptome data and post-translational information with a candidate approach.

The integrative analysis confirmed the potential implication of miR-600 in CSC differentiation or self-renewal. Using a pathway approach to explore the biology of cancer stem cells is a good way to develop **new targets in breast cancer treatment**.

Ultimately, elaborating a new drug targeting CSCs could greatly improve breast cancer clinical outcome.

1. World Cancer Report. International Agency for Research on Cancer. 2008.
2. Targeting breast cancer stem cells. Liu S & Wicha MS. J Clin Oncol. 2010 Sep 1;28(25):4006-12. doi: 10.1200/JCO.2009.27.5388.
3. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Ginestier C et al., Cell Stem Cell. 2007 Nov;1(5):555-67. doi: 10.1016/j.stem.2007.08.014.
4. Interactome-transcriptome integration for predicting distant metastasis in breast cancer. Garcia et al., Bioinformatics. 2012 Mar 1;28(5):672-8.
5. Detection of driver protein complexes in breast cancer metastasis by large-scale transcriptome-interactome integration. Garcia et al., Methods Mol Biol. 2014;1101:67-85.
6. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Matys et al., Nucleic Acids Res. 2006 Jan 1;34(Database issue):D108-10.

Acknowledgments
Support to this project is coming from the Institut National du Cancer grant Number INCA_5911 to CG. Additional support for QDC obtained from Plan Cancer 2013-2016 Action 3.6 to GB.

Instituts thématiques Inserm Institut national de la santé et de la recherche médicale
cnrs dépasser les frontières
Aix*Marseille université
INSTITUT PAOLI-CALMETTES
CRCM
CRCM's Integrative Bioinformatics
http://cibi.marseille.inserm.fr/
Institut National du Cancer
Roche