

Using a network approach to unravel biological pathways involved in cancer

C Rioualen^{1,2,3,4}, Q Da Costa^{1,2,3,4}, R El-Helou^{1,2,3,4}, E Charafe-Jauffret^{1,2,3,4}, C Ginestier^{1,2,3,4}, S Vasseur^{1,2,3,4}, F Guillaumond-Marchai^{1,2,3,4}, E Mas^{3,5}, I Crenon^{3,5}, G Pinna⁶, A Harel-Bellan⁶, G Bidaut^{1,2,3,4}

Cancer Research Center of Marseille
 1. Inserm U1068
 2. CNRS UMR7258
 3. Aix-Marseille Université
 4. Institut Paoli-Calmettes
 Center for Research in Oncobiology and Oncopharmacology
 5. Inserm UMR 911
 CEA SACLAY, Gif-sur-Yvette
 6. Plateforme ARN interference PArL

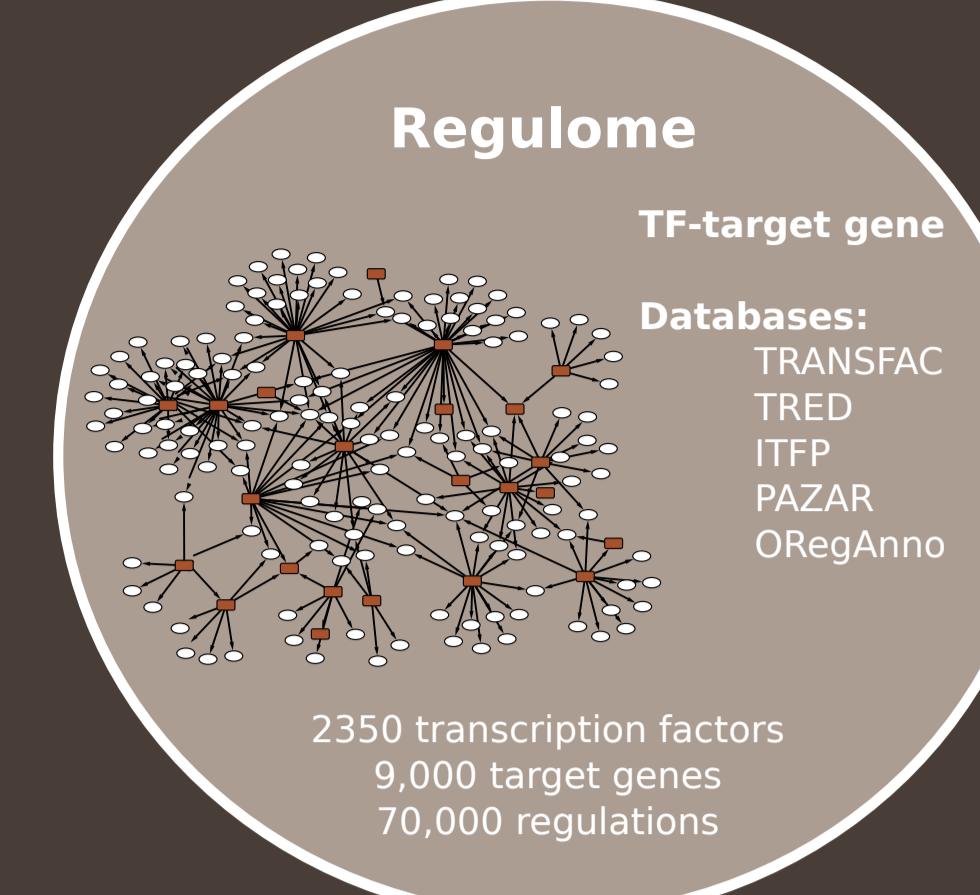
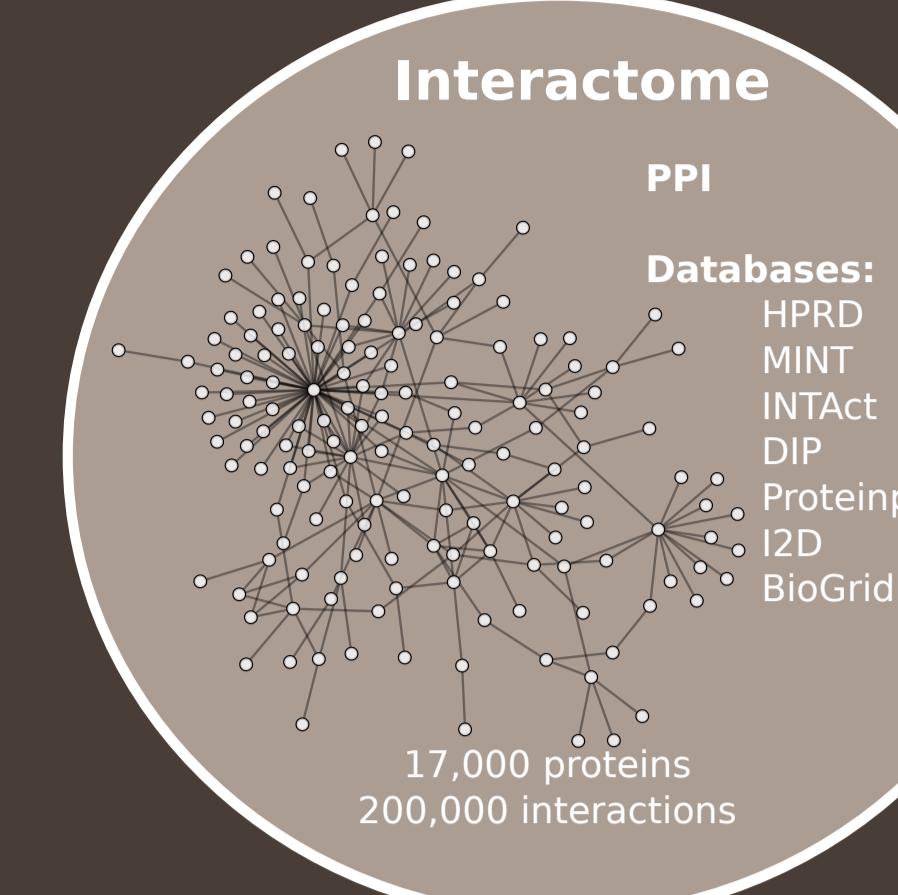
Introduction

Complex diseases, such as cancer, rarely arise from single-gene irregularities, but rather from complex interplays between intra- and inter-cellular components (1).

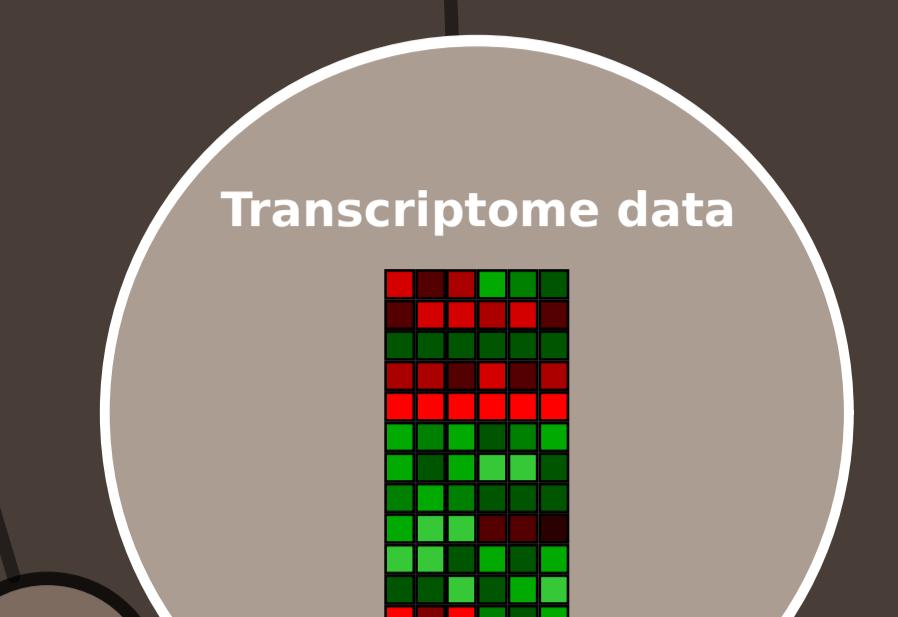
Networks approaches, including interaction data such as protein-protein interactions and transcription regulatory networks, have the ability to shed light on dynamic mechanisms within the cell.

Identifying signaling pathways involved in cancer can be the key leading to novel and more efficient therapeutic strategies, and to the improvement of clinical outcomes.

Networks building from public databases

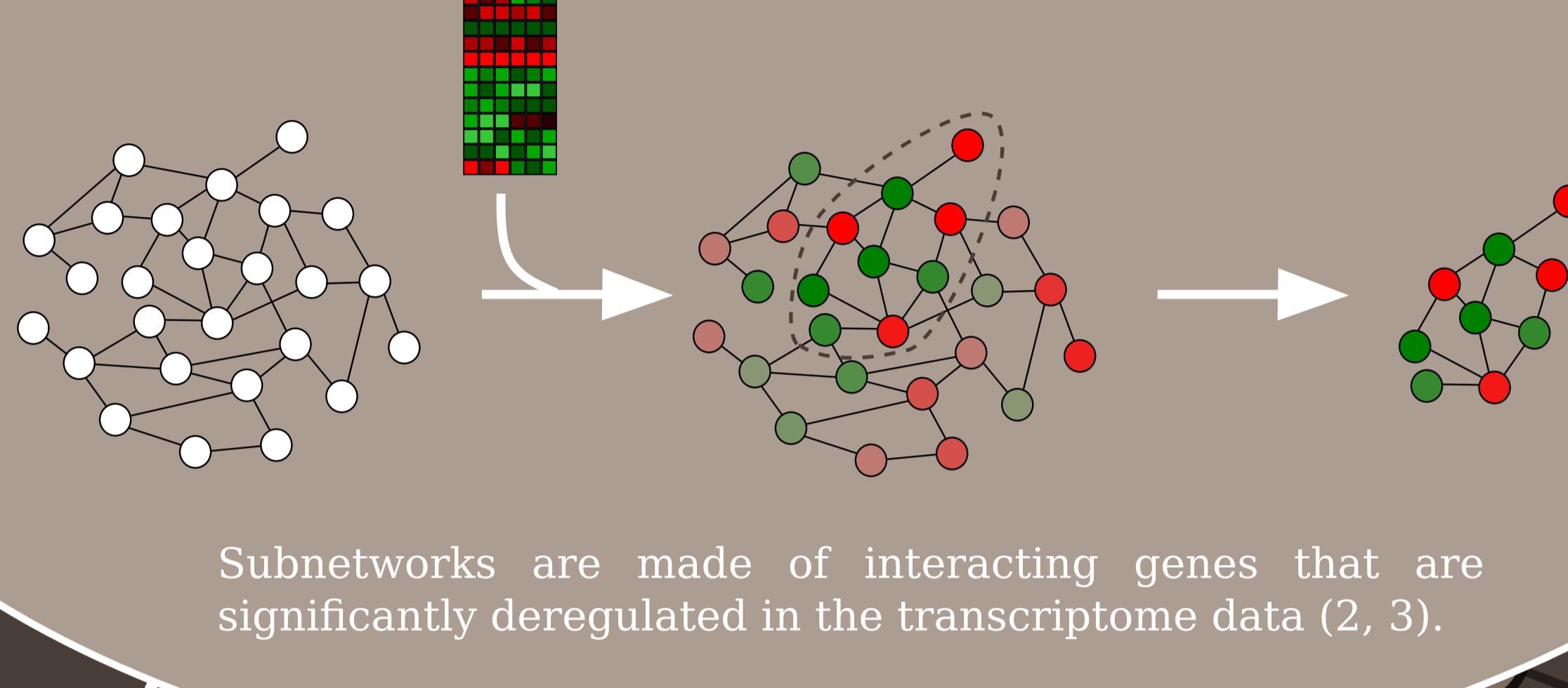


Input biological data



Gene ID	Z-Score
79789	0.26146437
608	-1.275269379
19357	-0.03599317
84187	0.117722377
57402	3.206966068
54737	0.04671515
90711	0.547270054

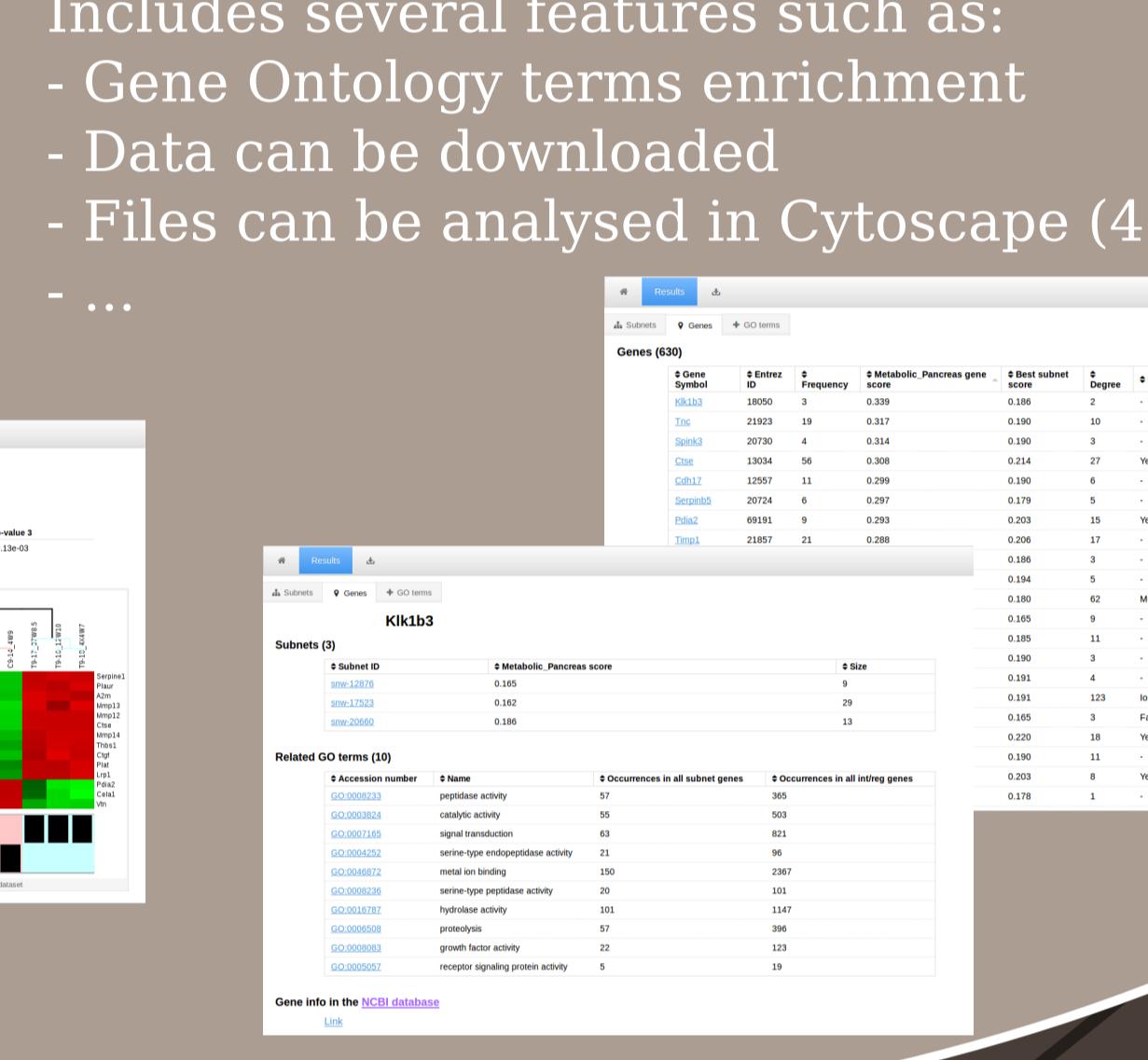
Subnetworks detection algorithm with ITI



Subnetworks integration (opt.)

Subnetworks from the interactome and the regulome are integrated according their number of shared nodes.

HTML visualisation



A. Pancreatic cancer project

This study in mouse aims to identify metabolic pathways involved in the pancreatic cancer development (tumorigenesis) over time. Therefore, we analysed transcriptome data from 3 different stages of disease development (4, 6 and 9 weeks old diseased mice and the age-paired normal mice; diseased=T for "tumor" and healthy=C for "control", early stage=4 for 4 weeks, intermediate stage=6 for 6 weeks and late stage=9 for 9 weeks).

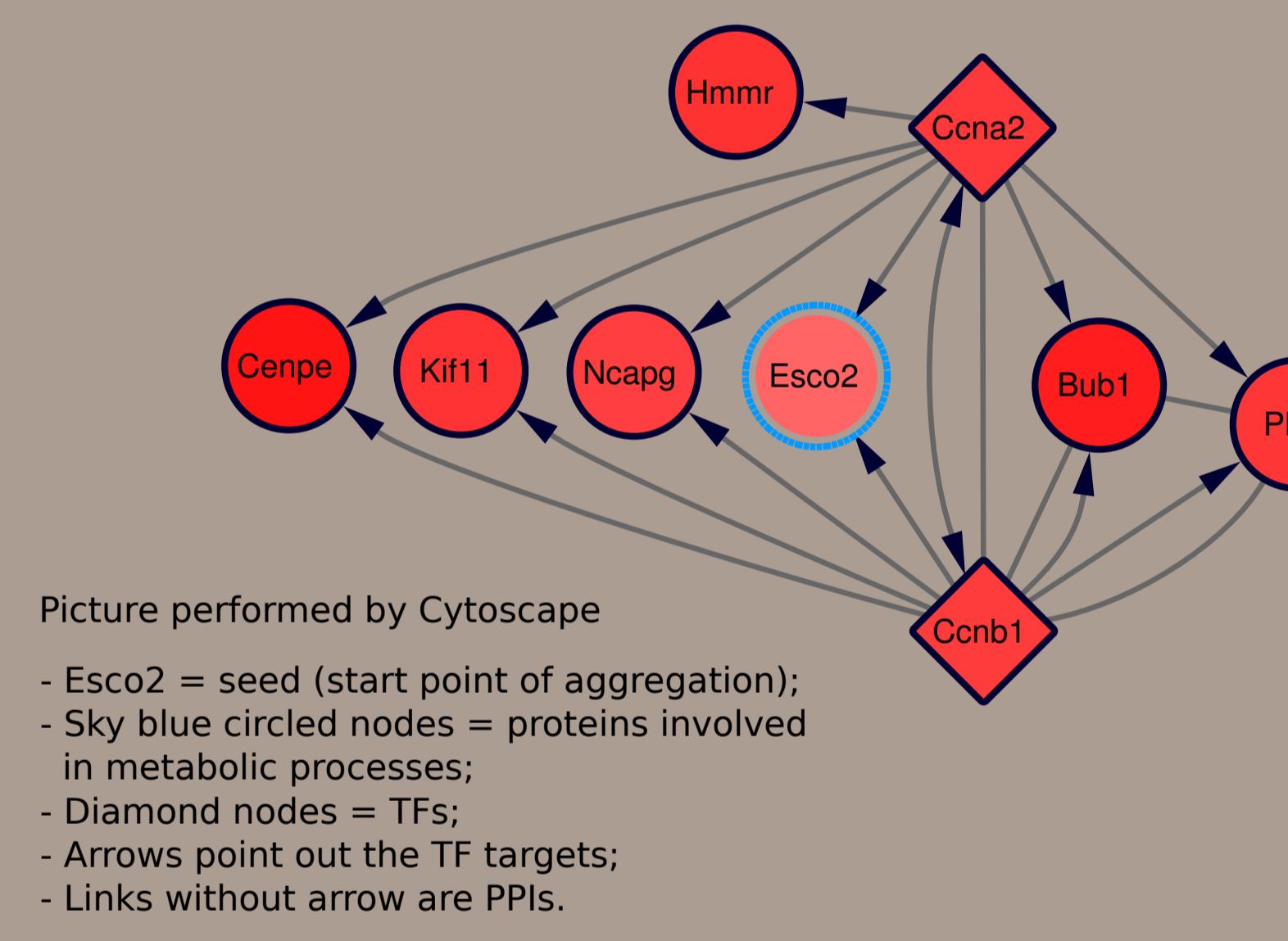
The mouse network used is constituted by 15,000 proteins, 125,000 PPIs, 1,700 TFs, 34,000 regulations.

In these ITI analyses, we used only proteins involved in metabolism as seed (start point of aggregation) of each subnetwork.

We made 3 analyses to isolate subnetworks for each time. We will shortly compare and combine subnetworks to visualize the gene expression evolution over time.

- Bold circled node (Igfbp3) = seed;
- Double circled nodes = proteins involved in metabolic processes;
- Rectangle nodes = TFs;
- Blue arrows point out the TF targets;
- Purple links are PPIs.

A subnetwork from the 4T versus 4C analysis



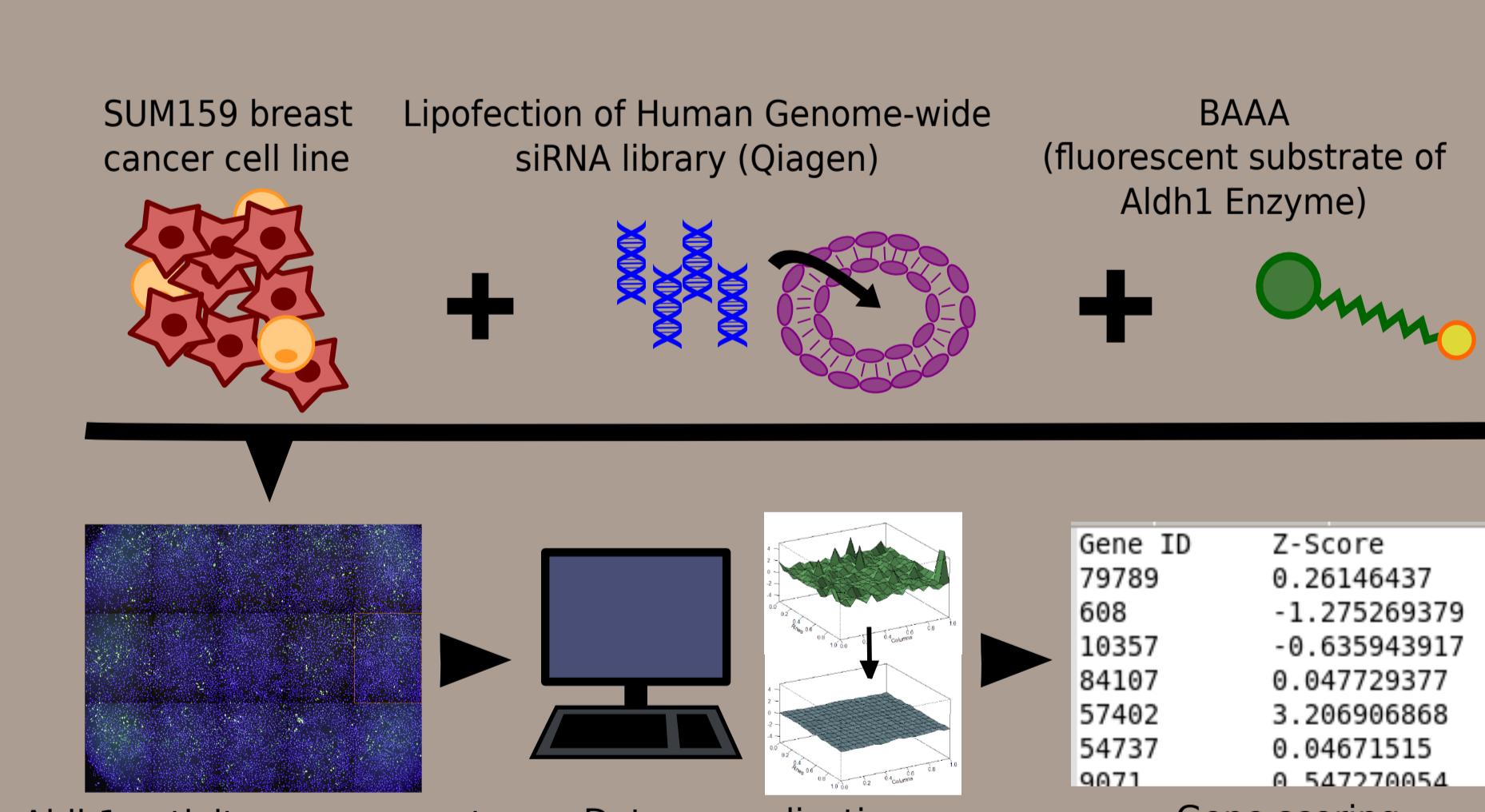
A subnetwork from the 9T versus 9C analysis

Picture performed by Graphviz (software used in the HTML report making)

B. Breast cancer stem cells project

Background. Breast cancer has a 5-year survival rate of ~90%, but remaining 10% are associated with lethal metastatic relapses. **Cancer stem cells** (CSC) might be the cause for these relapses, for they tend to resist conventional treatments, and have the ability to generate metastases thanks to their properties: **self-renewal** and **differentiation** (5, 6). The aim of this study is to better understand their biology, in order to build a new drug which could target them specifically.

Methods.



A genome-wide siRNA screening was performed on cancer cells, and the Aldh1 activity was measured. A low aldh+/aldh- ratio is associated with a low proportion of CSCs, while a high ratio is associated with a high proportion of CSCs (7). All genes are scored accordingly.

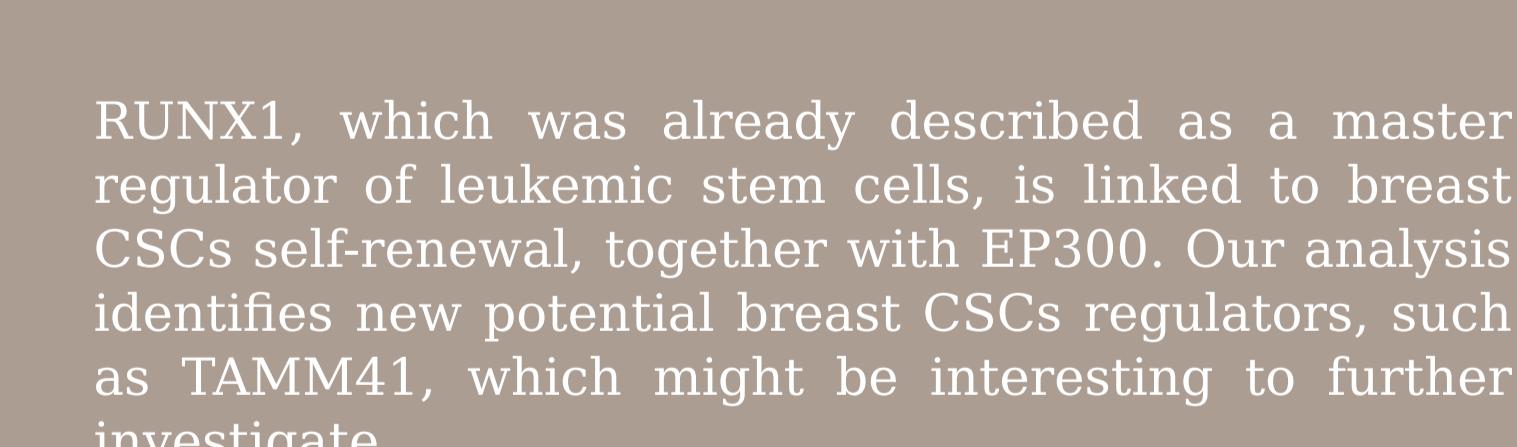
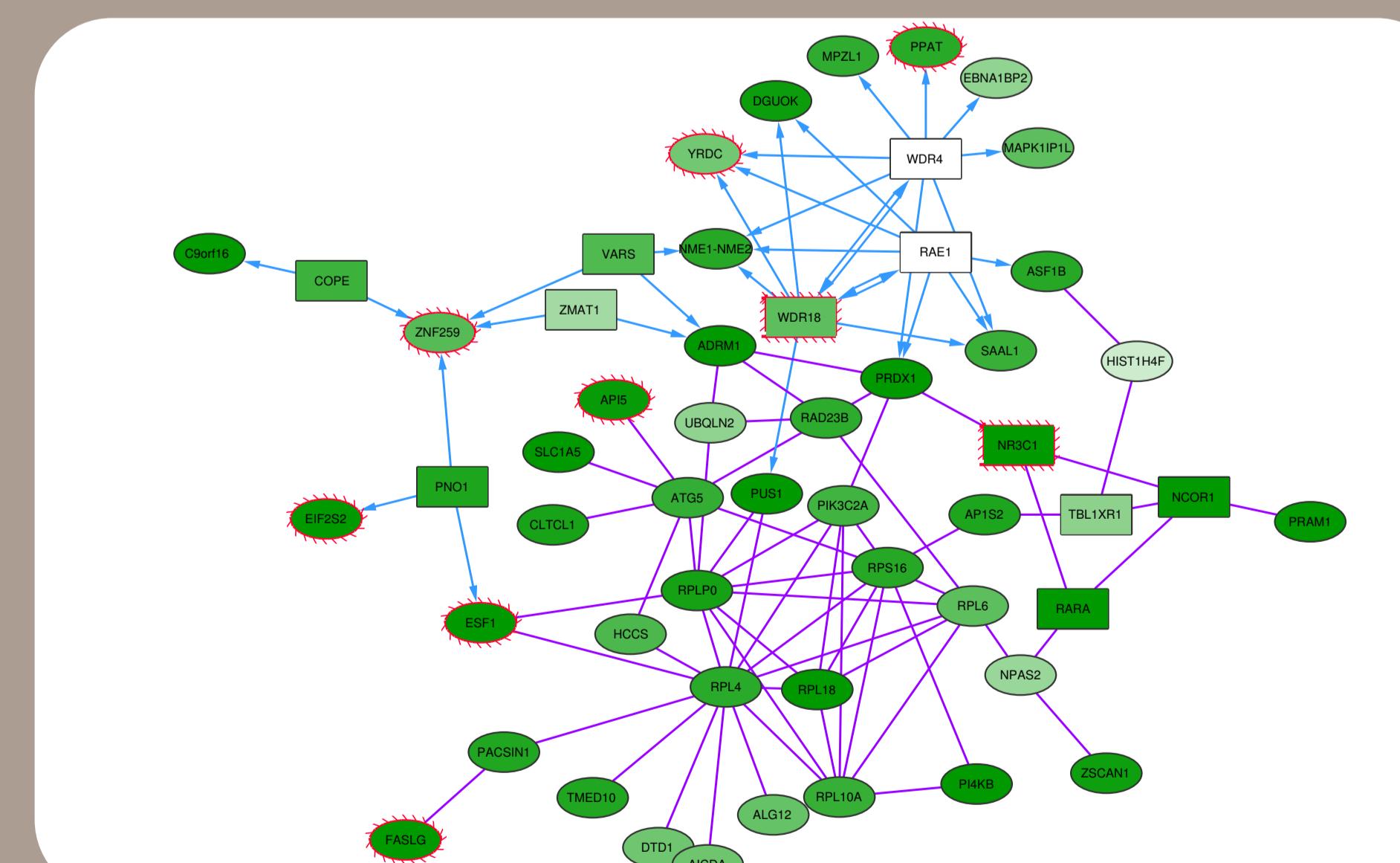
Subnetworks are built by aggregation of genes using their score.

Results.

Subnetworks detection

62 subnetworks were found in total. Double-circled are the genes found to be interesting after a primary screening. Green nodes are proteins coded by genes associated with a growth of the CSC population, while red ones are associated with CSC population decrease. Squares are transcription factors.

Subnetworks integration



RUNX1, which was already described as a master regulator of leukemic stem cells, is linked to breast CSCs self-renewal, together with EP300. Our analysis identifies new potential breast CSCs regulators, such as TAMM41, which might be interesting to further investigate.

Analysis of genes linked with CSC population decrease (fig on the left) revealed the potential implication of the MED family (mediator complex), as well as the RPL family (ribosomal proteins). NCOR1/RARA (nuclear corepressor complex), known for regulating Aldh1 expression, were used as a positive control.

Conclusion

Network biology seems to be a good approach in order to understand complex diseases such as cancer. Thus we developed the ITI pipeline in the Integrative Bioinformatics team.

Network data integration tends to confirm the implication of genes identified in vitro, but also allows the discovery of new genes, and their connections altogether. These connections can lead us to building whole pathways involved in a given clinical condition.

These findings can potentially better treatments in several ways: limit secondary effects due to lack of specificity, improve efficiency through synergistic effects. They can be used in drug repositioning or adjuvant drug building (1), thus improving the disease's prognosis and clinical outcome.

Acknowledgements

A. Additional support for QDC obtained from Plan Cancer 2013-2016 Action 3.6 to GB.

B. Support to this project is coming from the Institut National du Cancer grant Number INCA_5911 to CG.

References

- From protein interaction networks to novel therapeutic strategies. Jaeger et al., JUBMB Life. 2012 Jun;64(6):529-37.
- Interactome-transcriptome integration for predicting distant metastasis in breast cancer. Garcia et al., Bioinformatics. 2012 Mar 1;28(5):672-8.
- 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.
- QDC: a web-based platform for the detection of biological pathways. Shannon P et al., Genome Research. 2003 Nov; 13(11):2498-504.
- Targeting breast cancer stem cells. Lin S & Wicha MS. J Clin Oncol. 2010 Sep 1;28(25):4006-12. doi: 10.1200/JCO.2009.27.5368.
- Stem cells, cancer and cancer stem cells. Reya et al., Nature. 2001 Nov 1;414(6859):105-11.
- 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):55-67. doi: 10.1016/j.stem.2007.08.014.