

Deciphering cancer stem cells regulatory circuits through an interactome–regulome–transcriptome integrative approach

Claire Rioualen, Rita El-Helou, Quentin Da Costa,
Emmanuelle Charafe-Jauffret, Christophe Ginestier, Ghislain Bidaut

Centre de Recherche en Cancérologie de Marseille – Bioinformatics platform

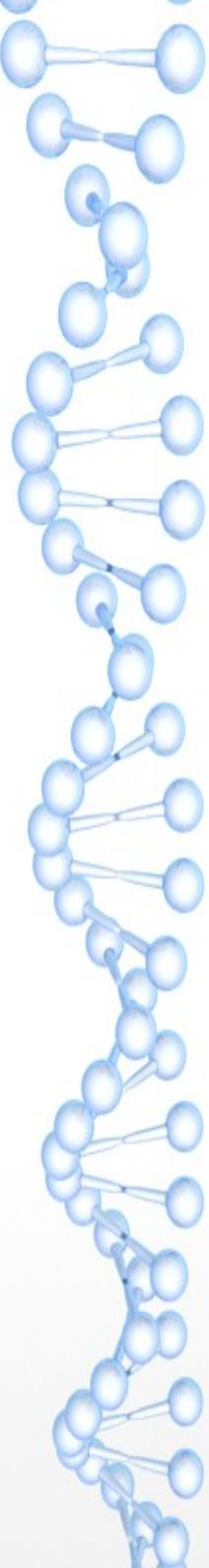
Inserm U1068
CNRS UMR7258
Aix-Marseille Université
Institut Paoli-Calmettes



111000101010111011010010010
1010101110101111110001100
0010101011100011000111101
01011100101110111110010
10000111011100101001

CRCM's Integrative BioInformatics



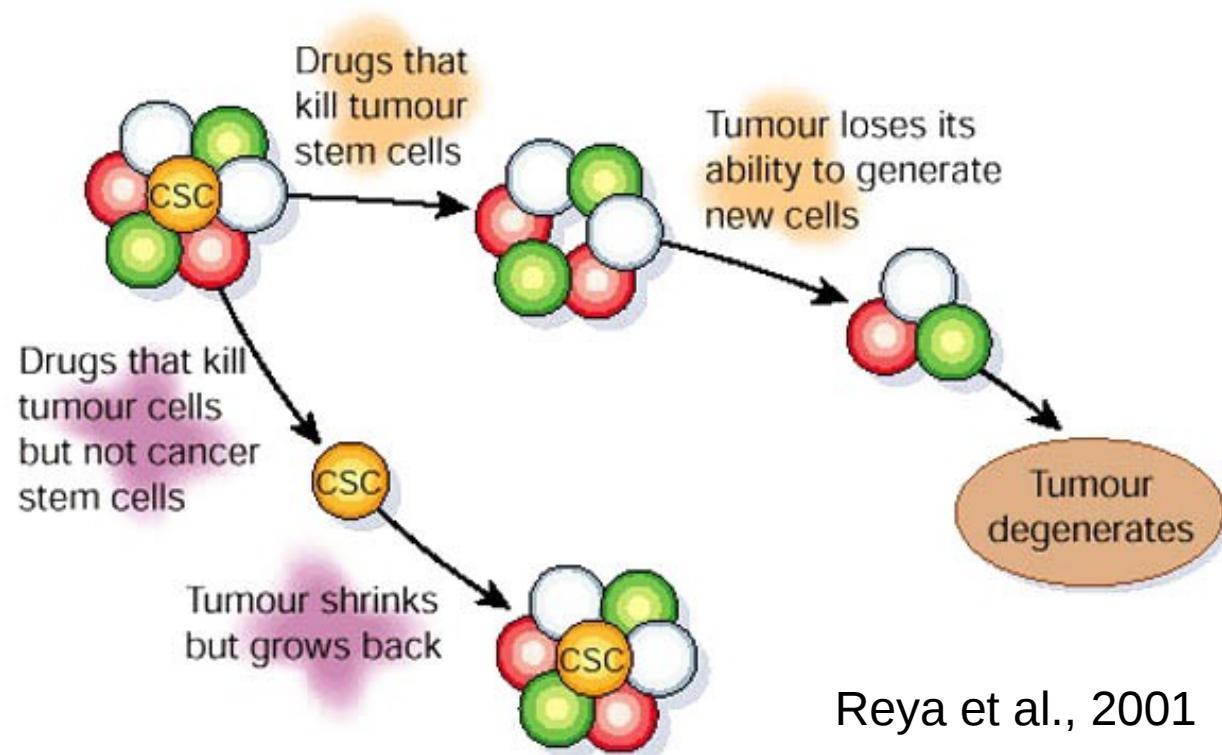


Breast Cancer

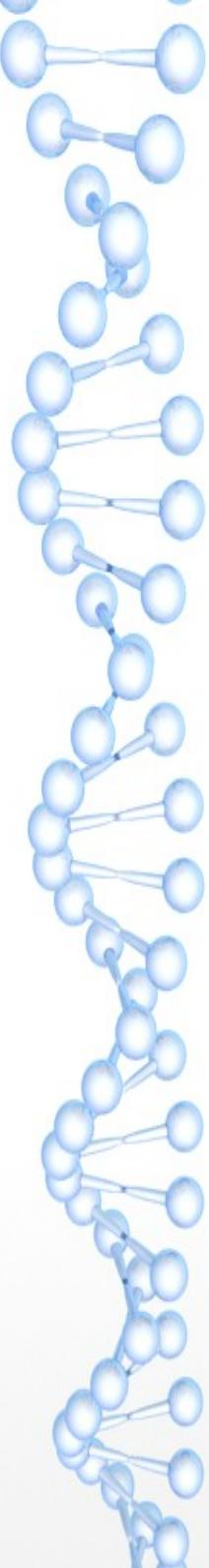
- Deadliest cancer in women worldwide
- 5-year survival rate of 85% in developed countries
- Most deaths are caused by metastatic relapses
- **Cancer stem cells** could explain **relapses**

Cancer Stem Cells (CSC)

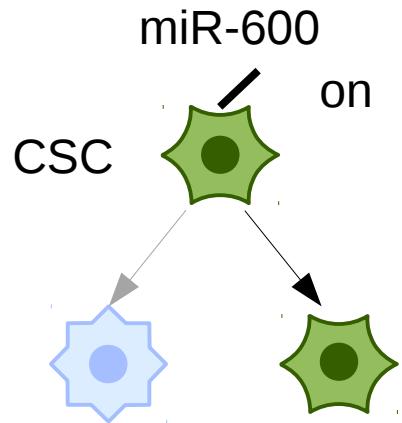
- CSC show resistance to conventional treatments
- In order to build new treatments targeting CSC, we need to better understand their biology



Reya et al., 2001



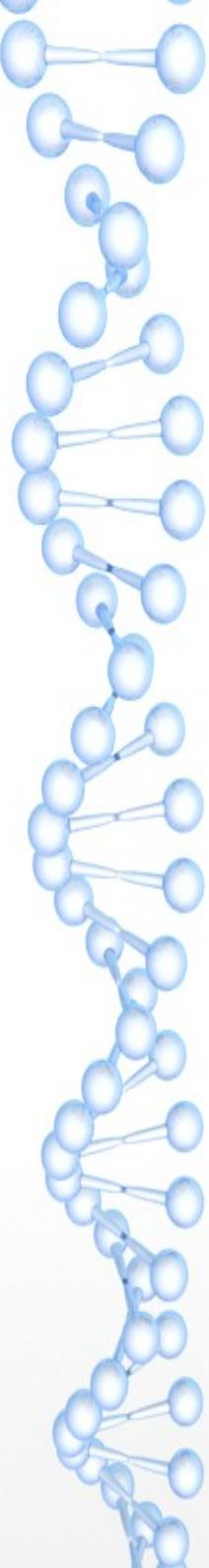
Cancer Stem Cells (CSC)



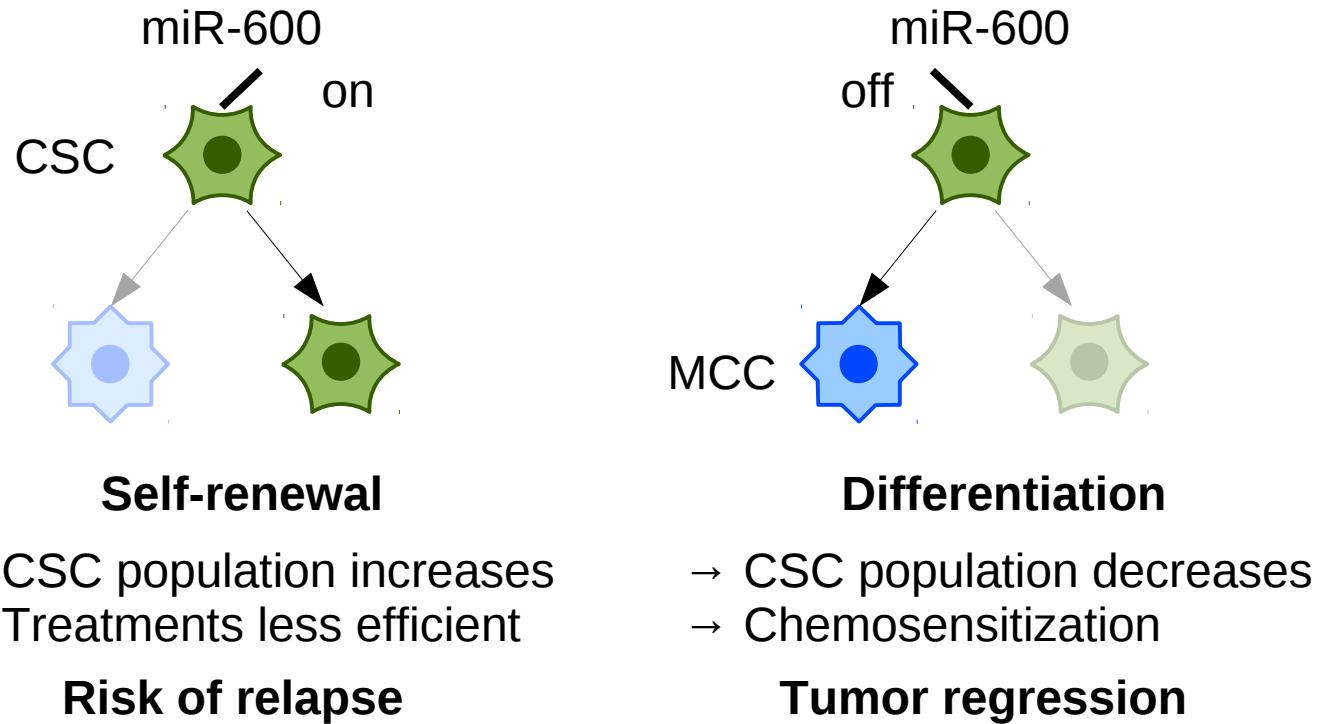
Self-renewal

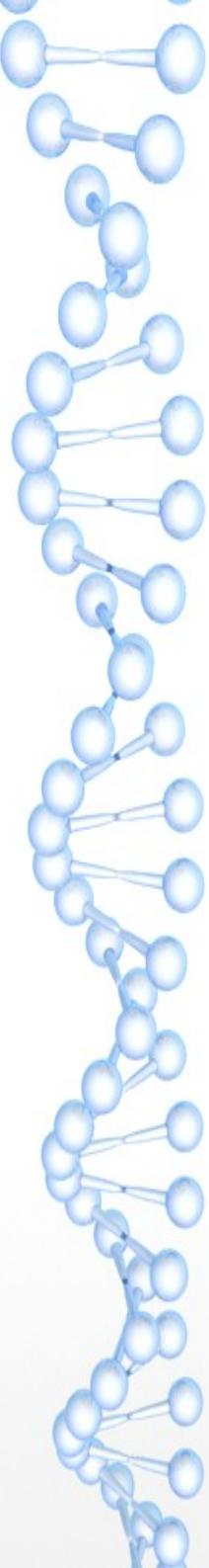
- CSC population increases
- Treatments less efficient

Risk of relapse

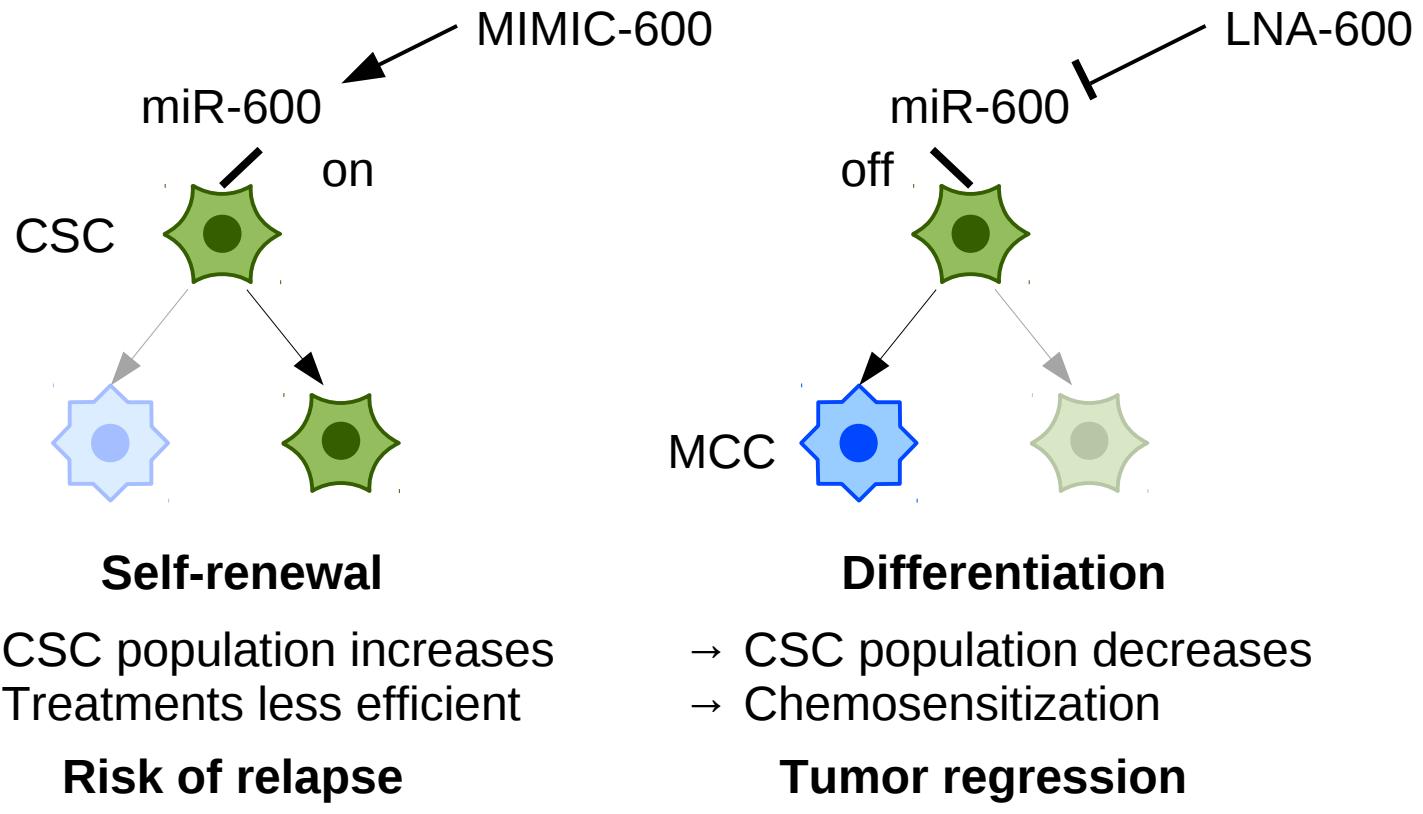


Cancer Stem Cells (CSC)

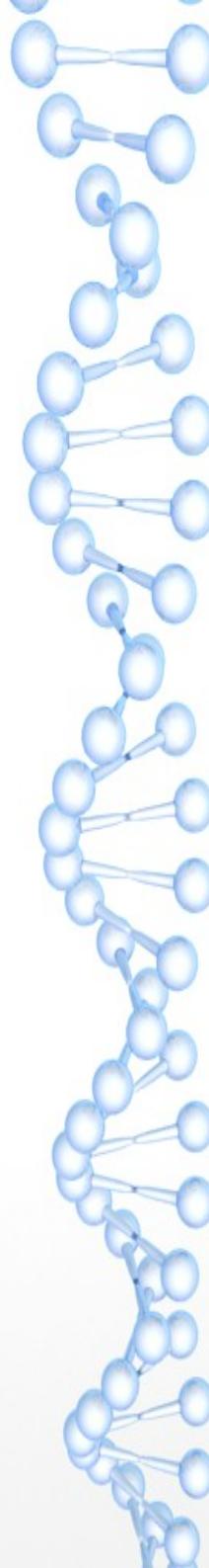




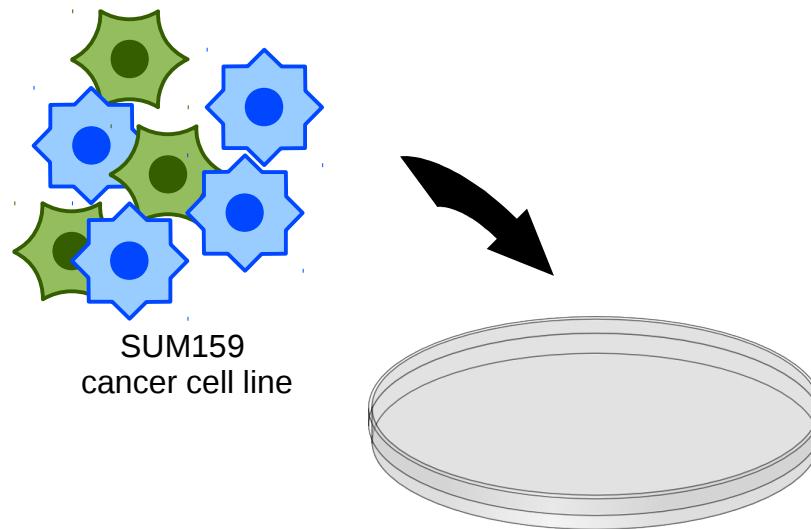
Cancer Stem Cells (CSC)



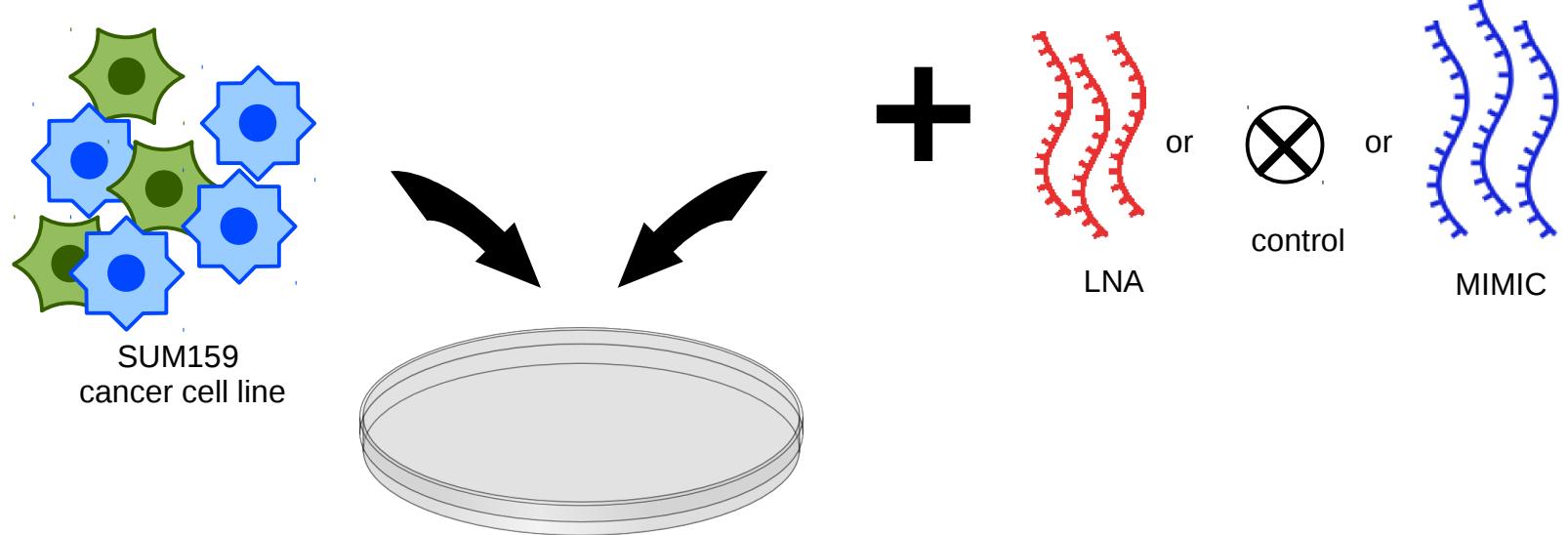
- Objective:
 - Identify “mirror” genes up/down-regulated by miR-600
 - Unravel pathways involved in CSC regulation



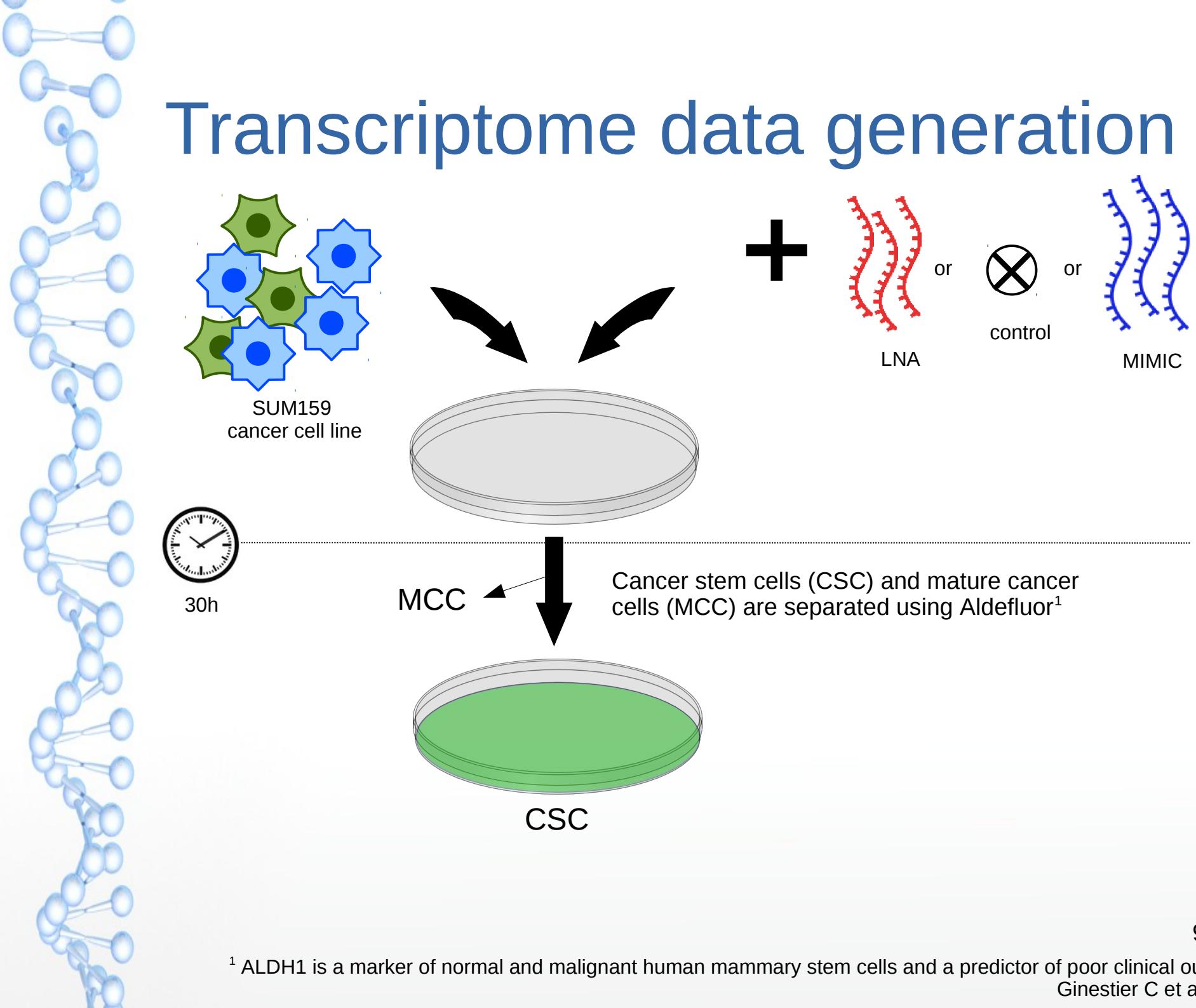
Transcriptome data generation



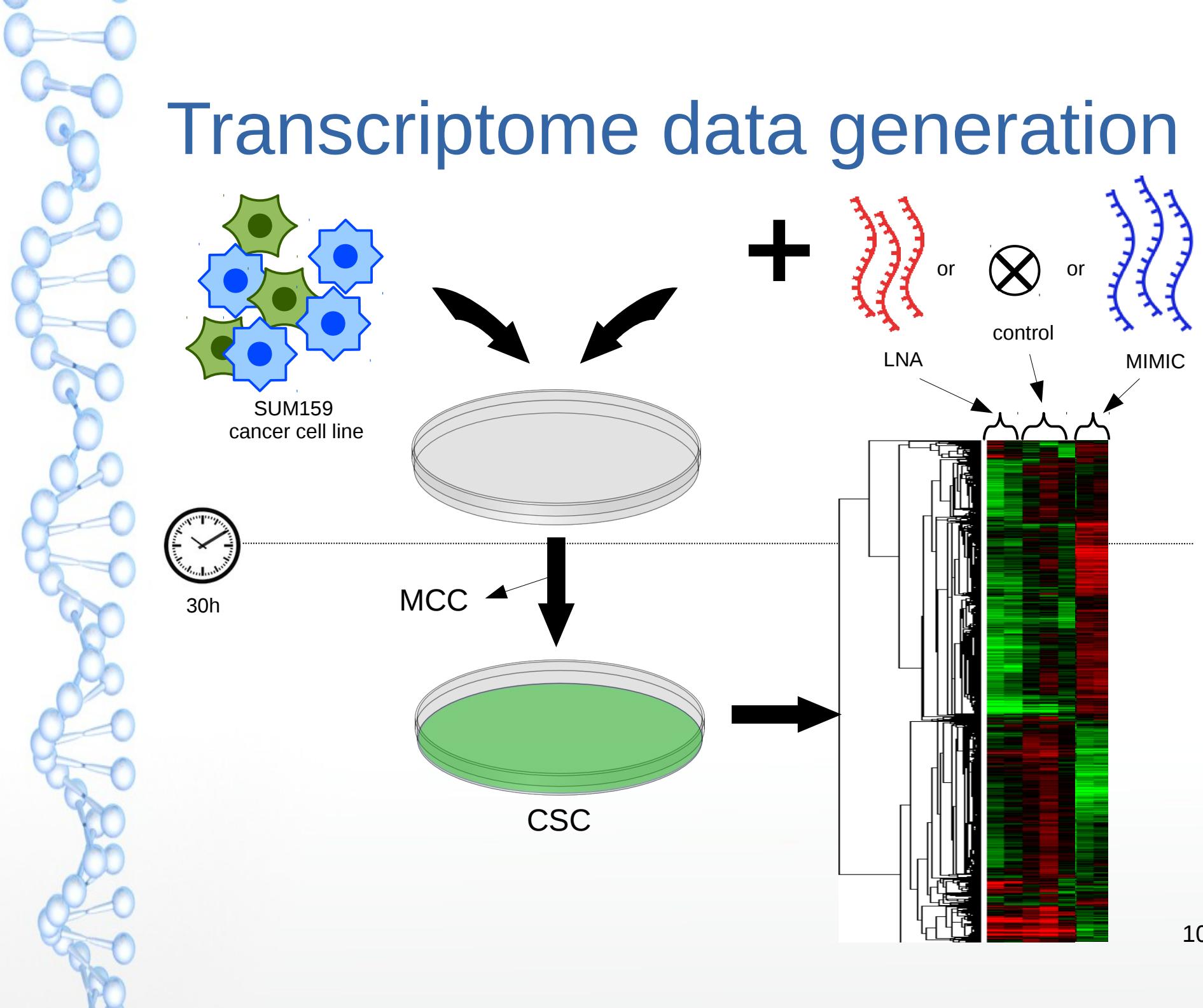
Transcriptome data generation



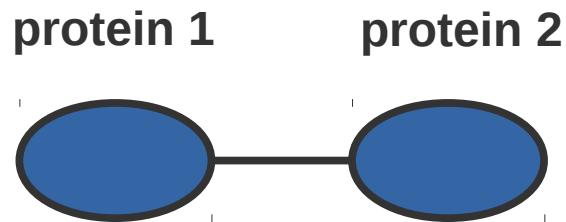
Transcriptome data generation



Transcriptome data generation



Network data: interactome & regulome



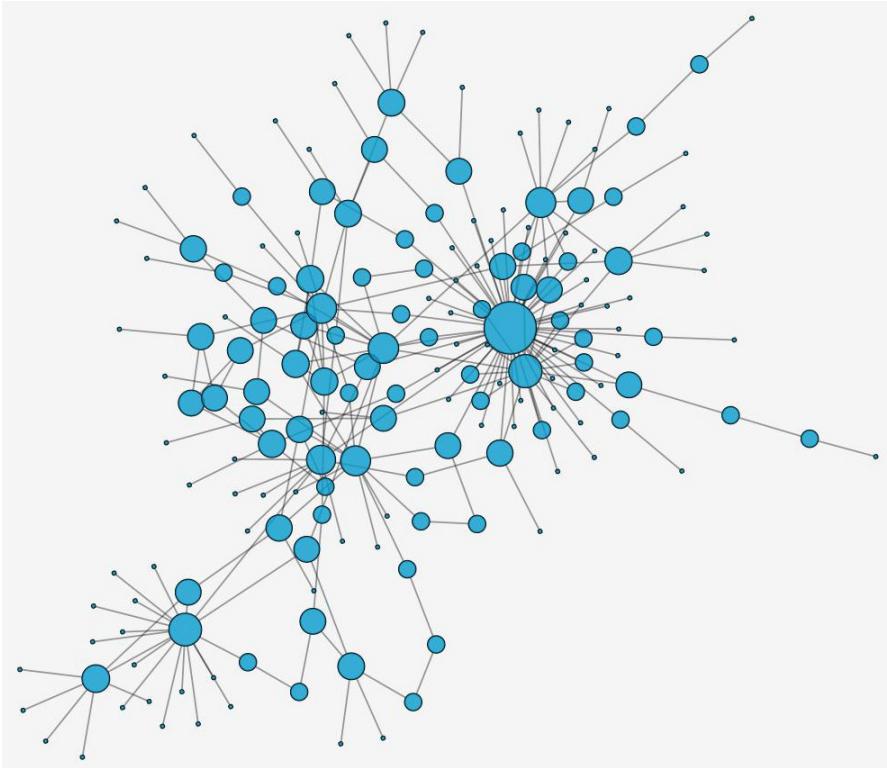
Interactome:

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

Regulome:

- TRANSFAC, TRED, ITFP, PAZAR, OregAnno
- 2352 transcription factors
- 9k targets
- 70k regulations

Network data: interactome & regulome



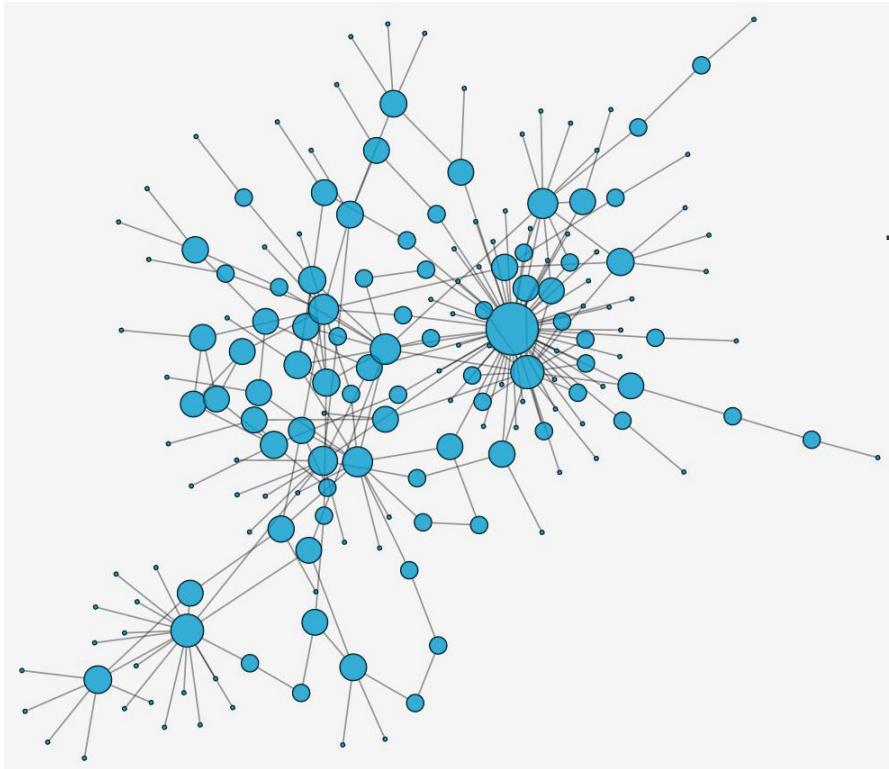
Interactome:

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

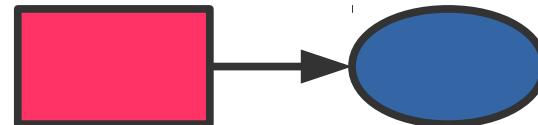
Regulome:

- TRANSFAC, TRED, ITFP, PAZAR, OregAnno
- 2352 transcription factors
- 9k targets
- 70k regulations

Network data: interactome & regulome



transcription factor



target gene

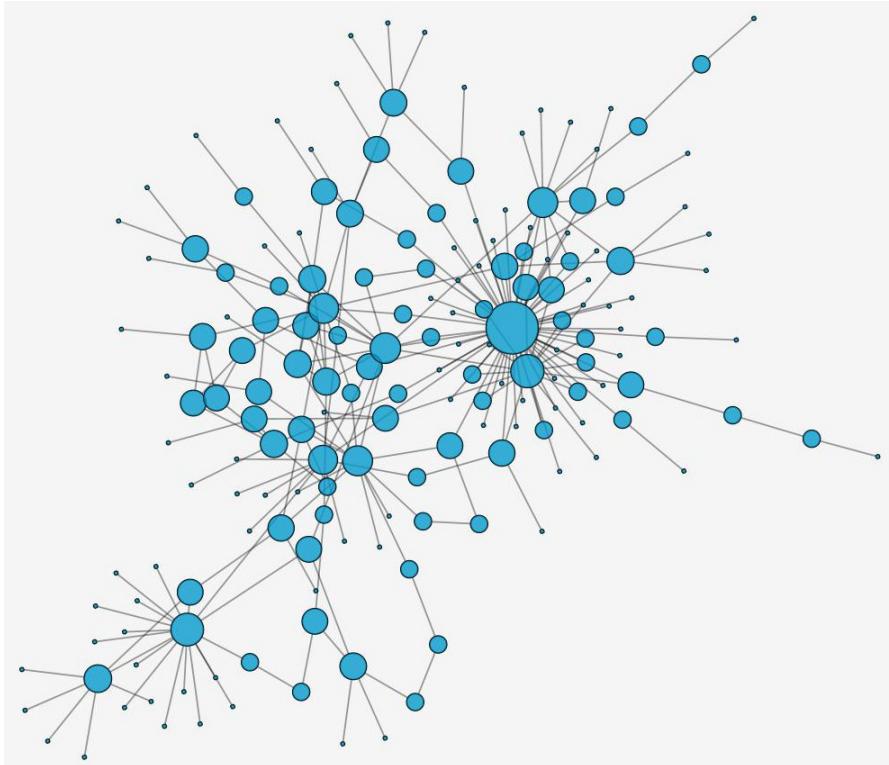
Interactome:

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

Regulome:

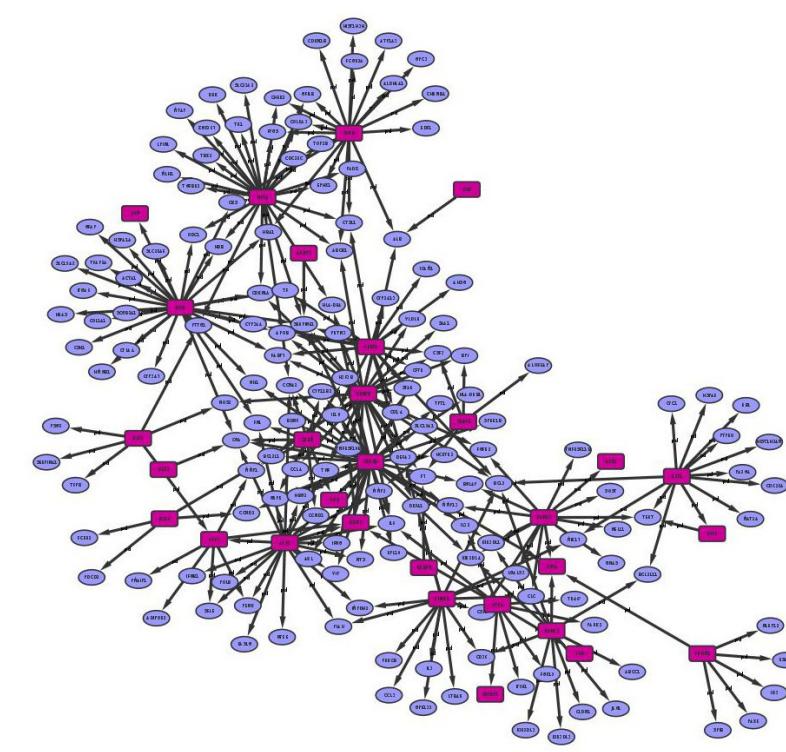
- TRANSFAC, TRED, ITFP, PAZAR, OregAnno
- 2352 transcription factors
- 9k targets
- 70k regulations

Network data: interactome & regulome



Interactome:

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

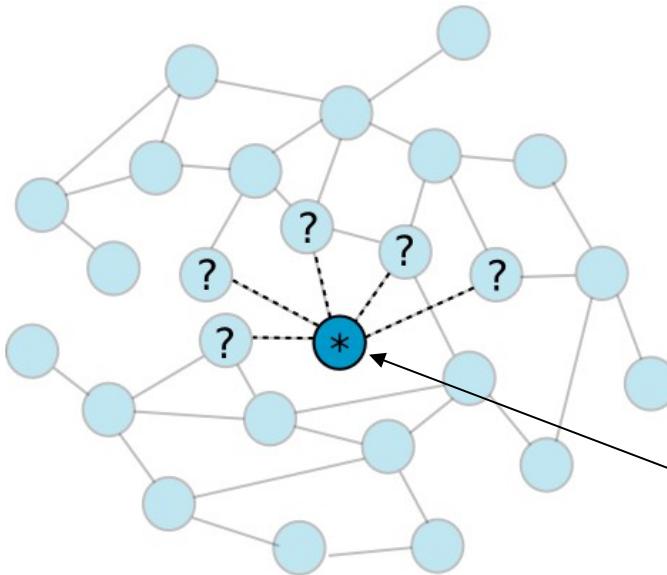


Regulome:

- TRANSFAC, TRED, ITFP, PAZAR, OregAnno
- 2352 transcription factors
- 9k targets
- 70k regulations

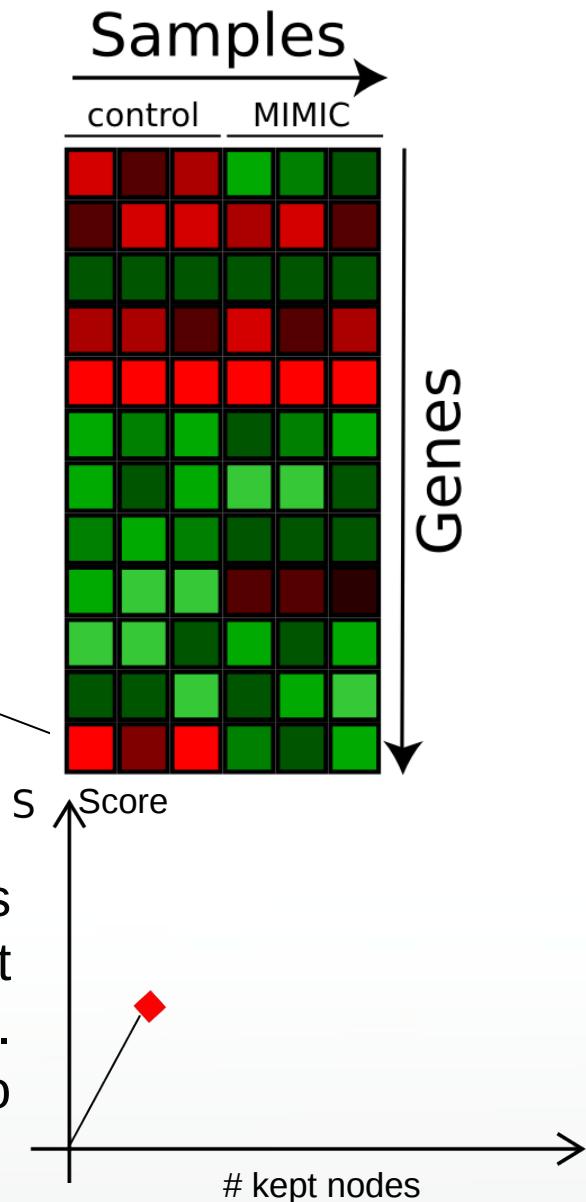
ITRI pipeline

Part 1: subnetworks detection



1. Subnetwork seed detection with ITI¹

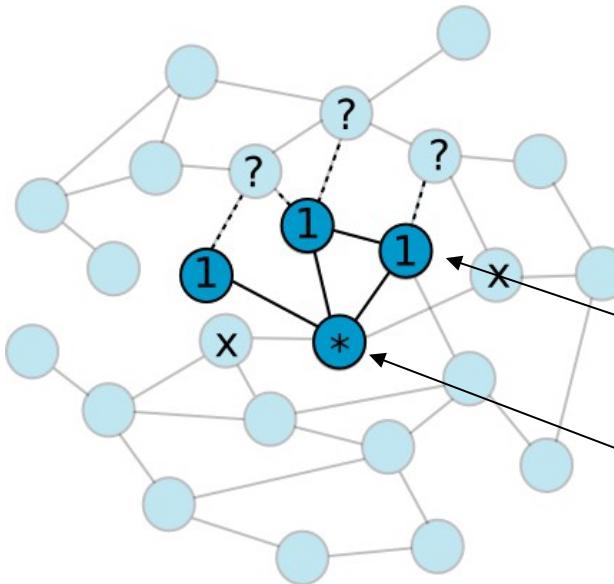
Node * is called the seed. If the gene is differentially expressed with LNA or MIMIC, it is kept and its neighbors are investigated next. The seed is assigned a score according to the expression.



¹ Interactome-transcriptome integration for predicting distant metastasis in breast cancer.
Garcia et al., 2012.

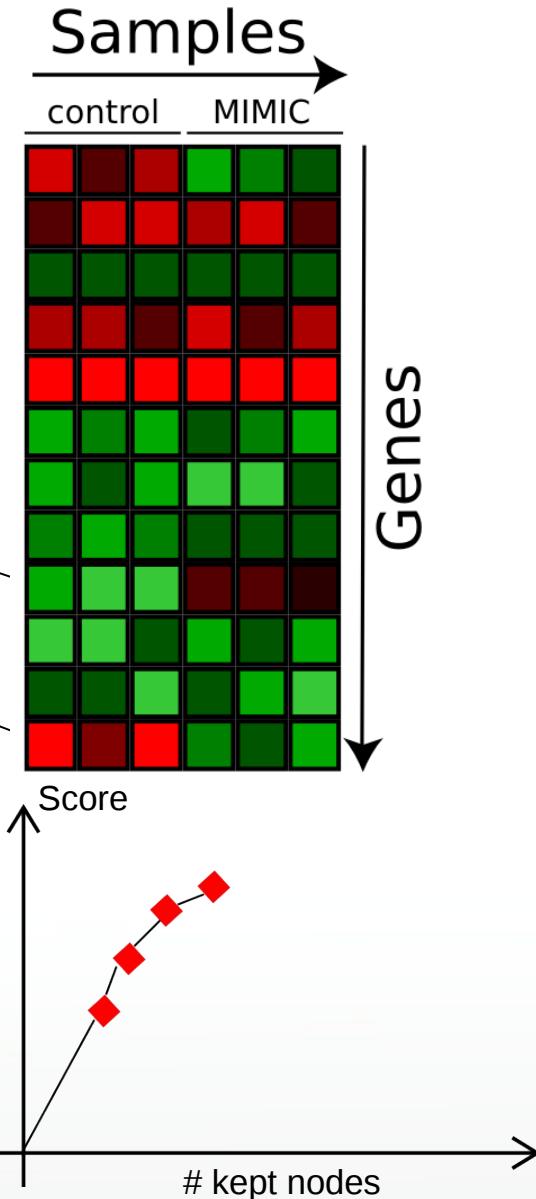
ITRI pipeline

Part 1: subnetworks detection



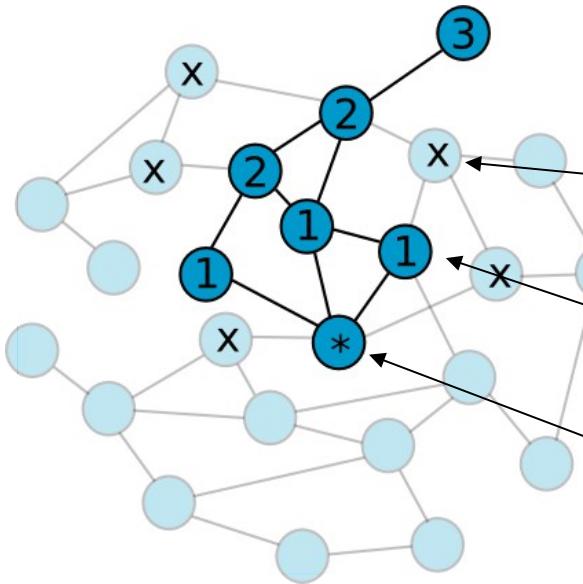
2. Neighbors exploration

Nodes that increase the average score of the subnetwork are kept. Their neighbors are then investigated recursively too.



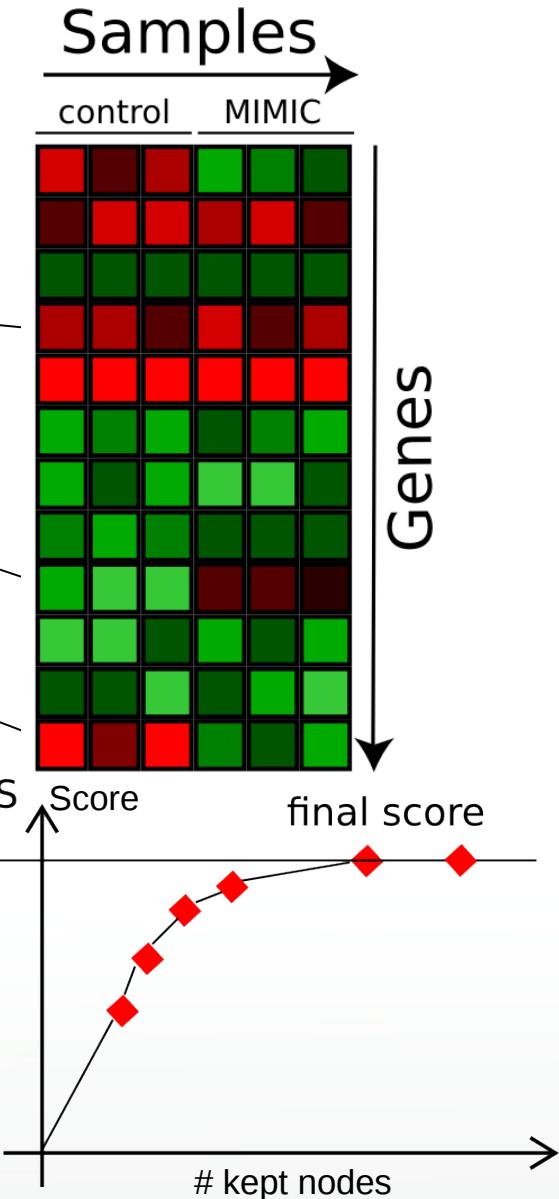
ITRI pipeline

Part 1: subnetworks detection



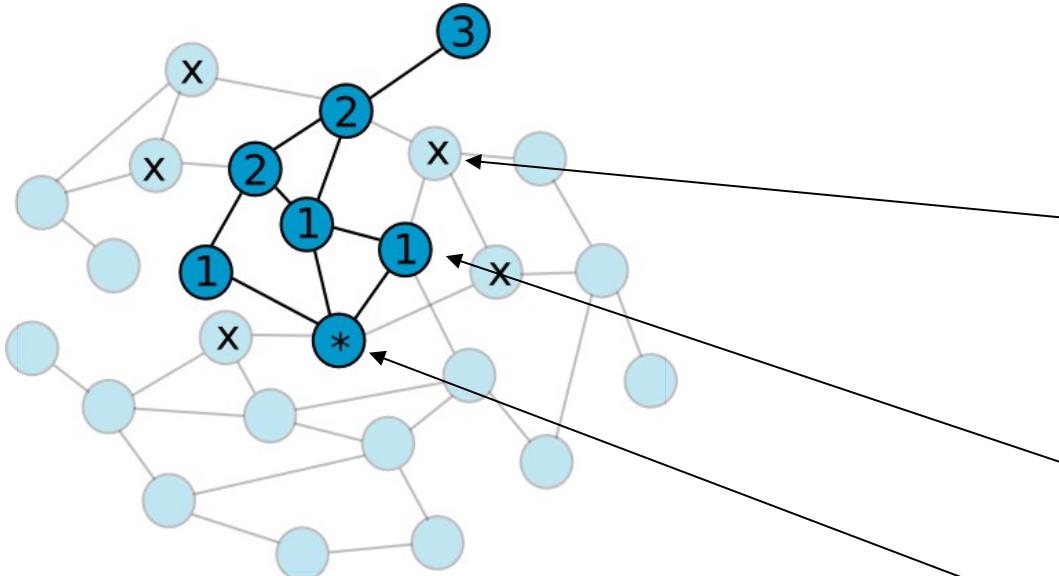
3. Subnetwork completion

A node that is not correlated stops the path of expansion for the subnetwork. Once the score cannot be improved, the subnetwork is considered complete.



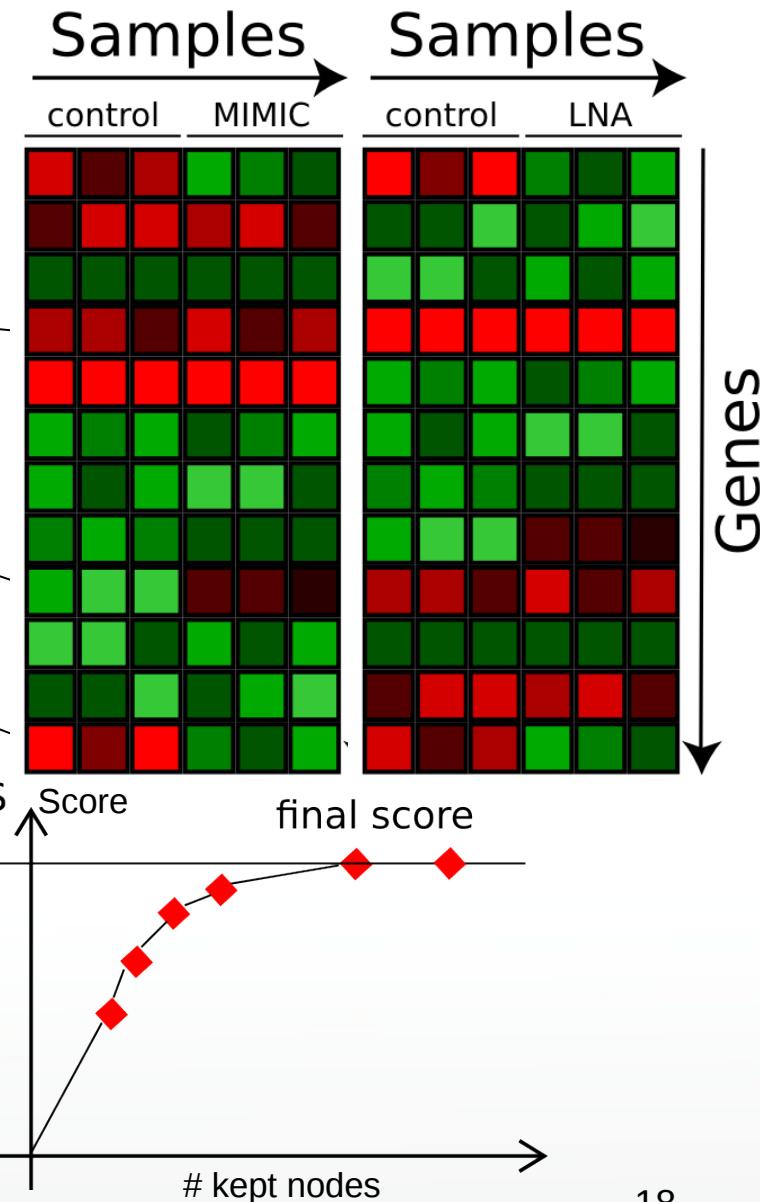
ITRI pipeline

Part 1: subnetworks detection



3. Subnetwork completion

A node that is not correlated stops the path of expansion for the subnetwork. Once the score cannot be improved, the subnetwork is considered complete.

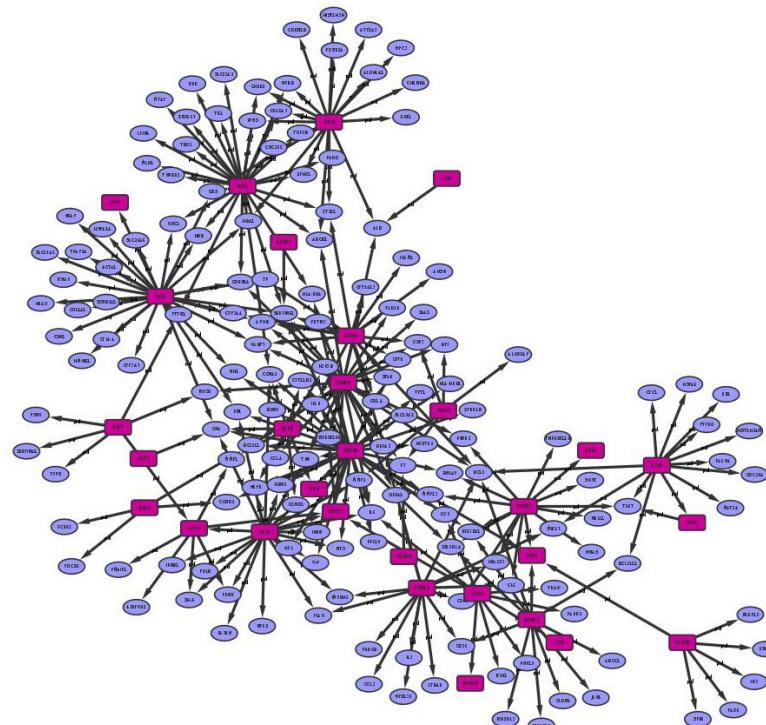


ITRI pipeline

Part 2: subnetworks integration



Integration



Interactome:

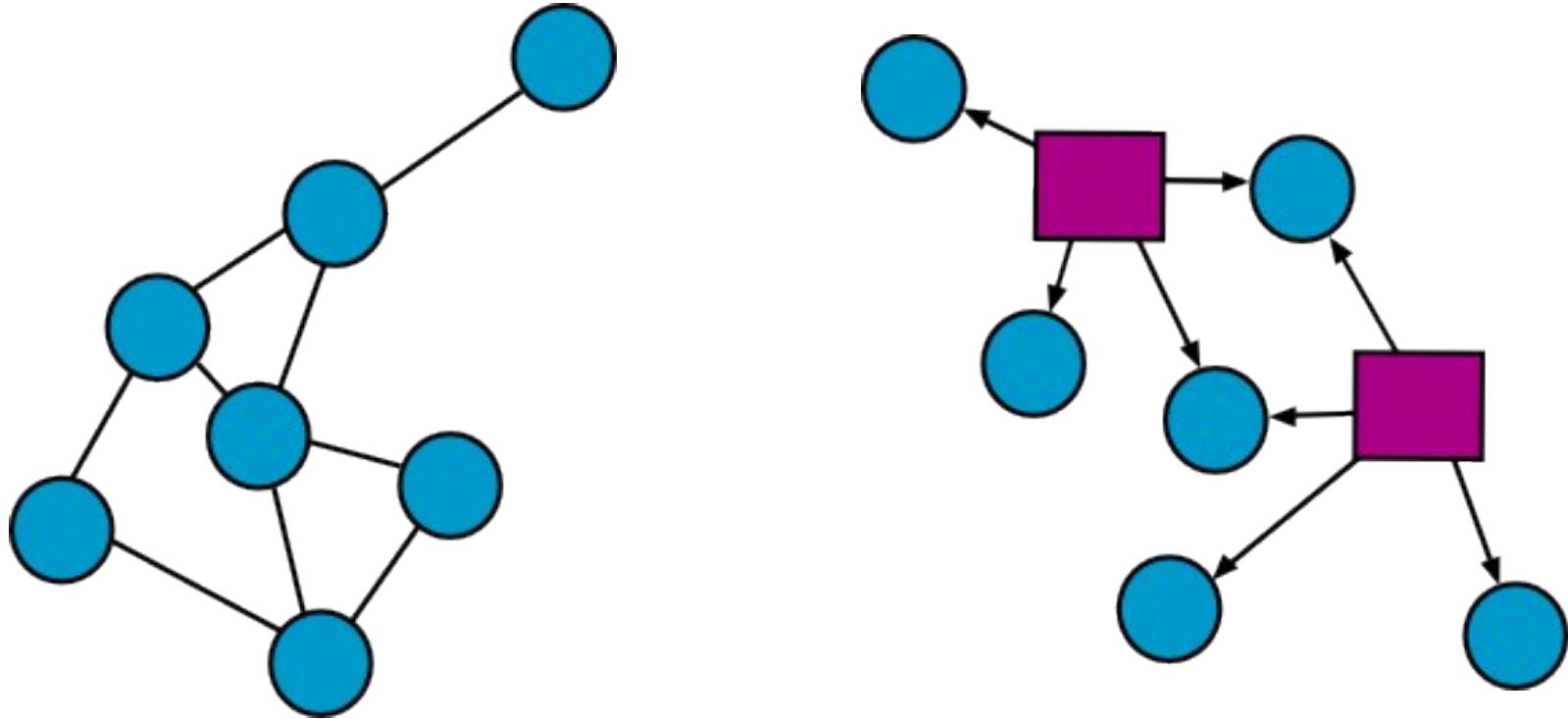
- HPRD, MINT, INTAct, DIP, Proteinpedia, I2D, BioGrid
- 17k nodes
- 200k interactions

Regulome:

- TRANSFAC, TRED, ITFP, PAZAR, OregAnno
- 2352 transcription factors
- 9k targets
- 70k regulations

ITRI pipeline

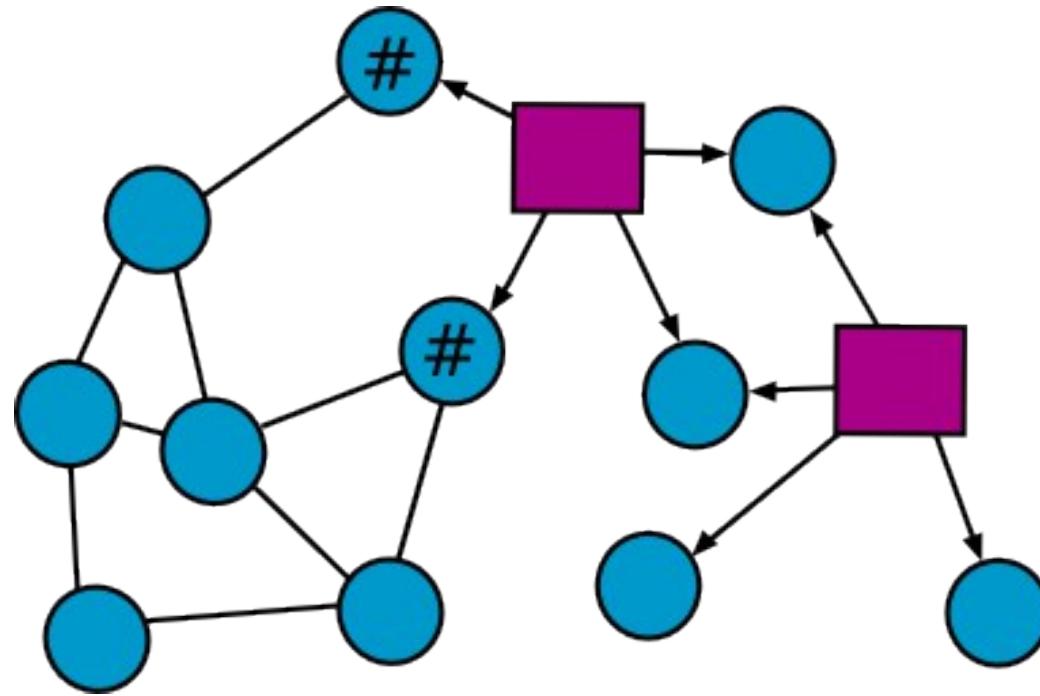
Part 2: subnetworks integration



- 2 different types of subnetworks are obtained:
 - Protein-protein interactions
 - TF-target genes

ITRI pipeline

Part 2: subnetworks integration

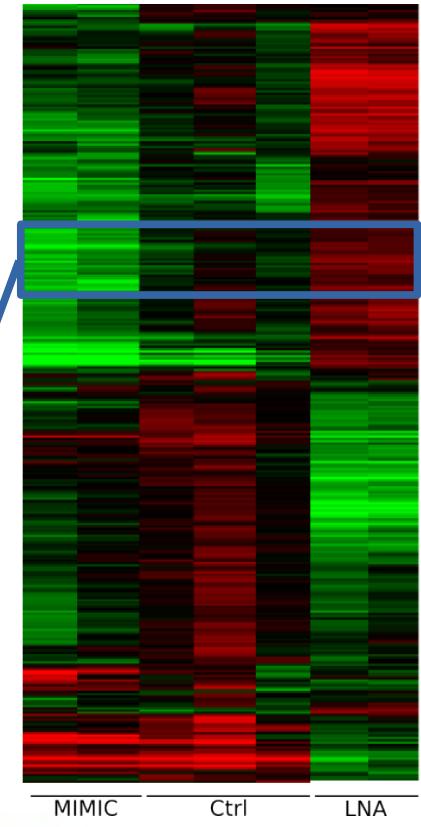
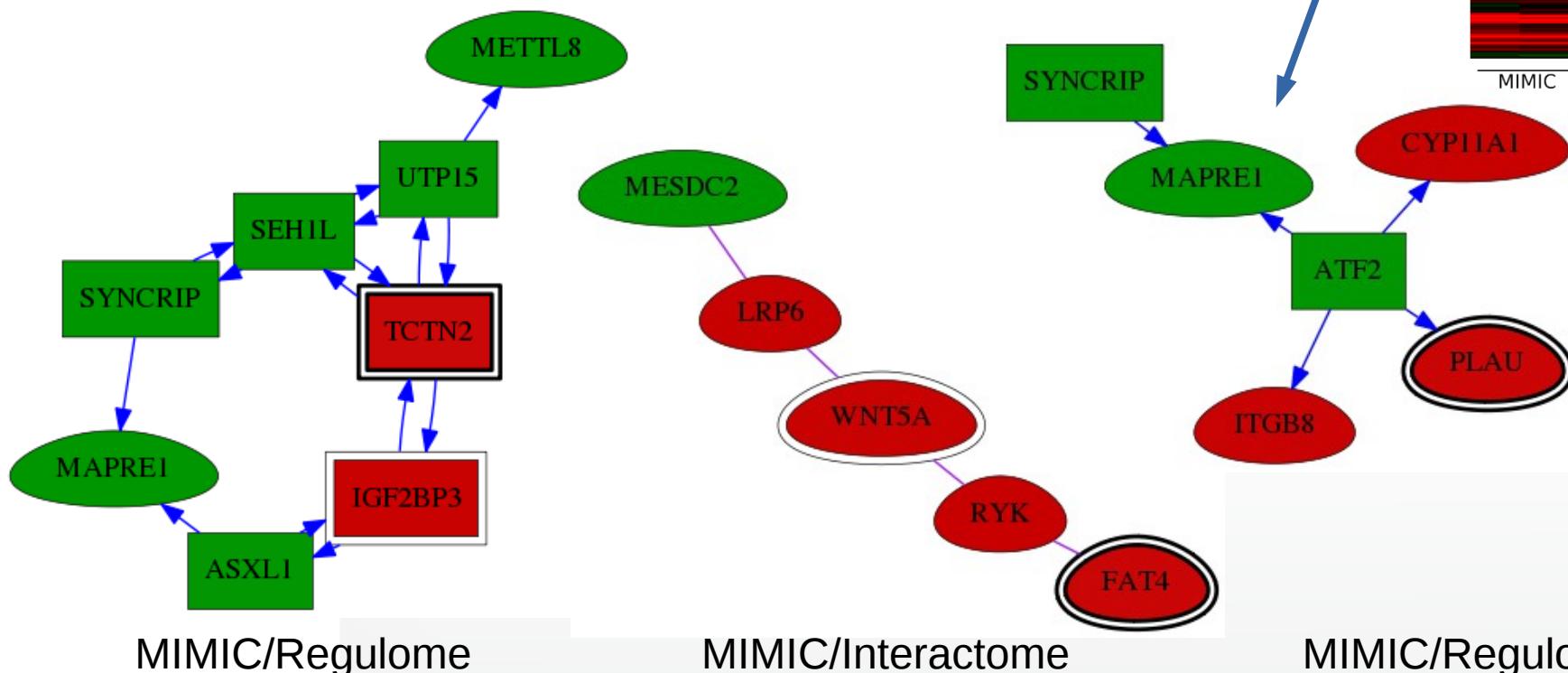


- Subnetworks are merged into “meta-subnets” according to the number of common nodes
- Proteins are assimilated as genes

Results

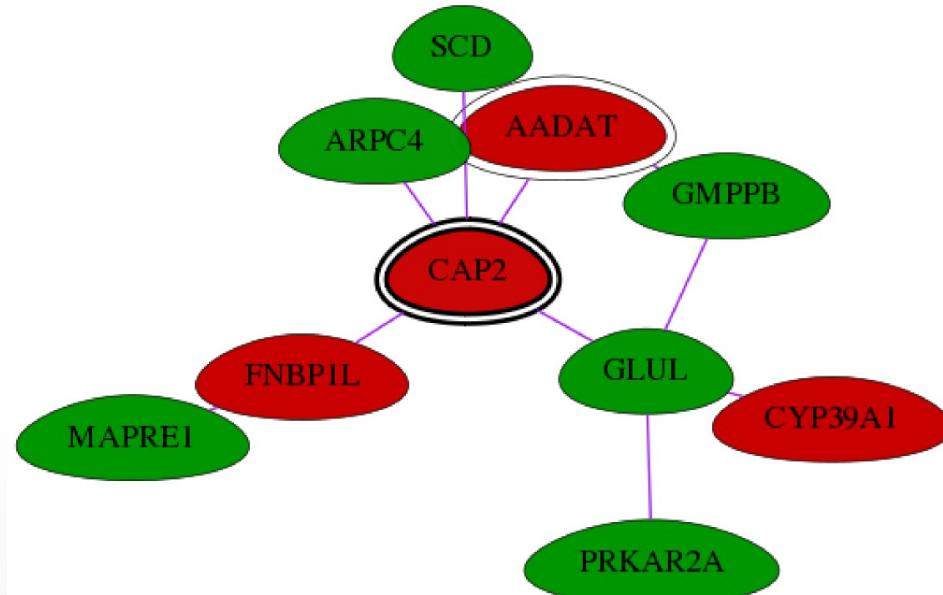
- 42 subnetworks were differentially expressed

| | Regulome | Interactome |
|-------|----------|-------------|
| LNA | 9 | 18 |
| MIMIC | 3 | 12 |

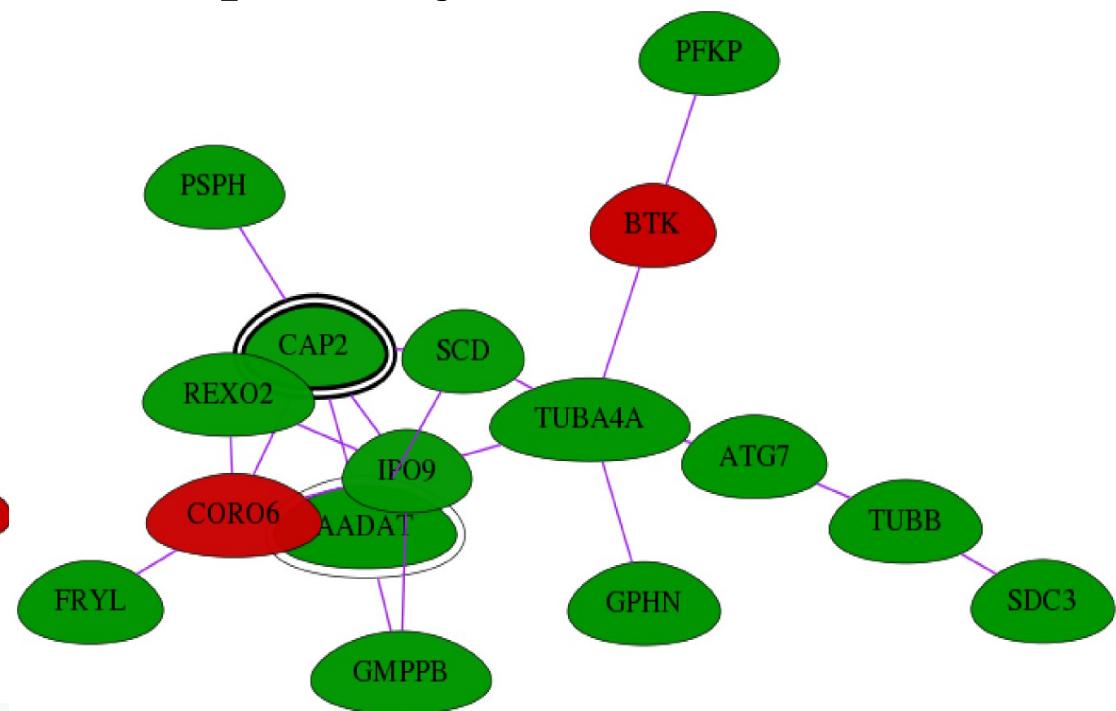


Results

- Mirror genes were found in LNA and MIMIC data, reinforcing the hypothesis that miR-600 could be a switcher
- Genes could be involved in CSC pathways

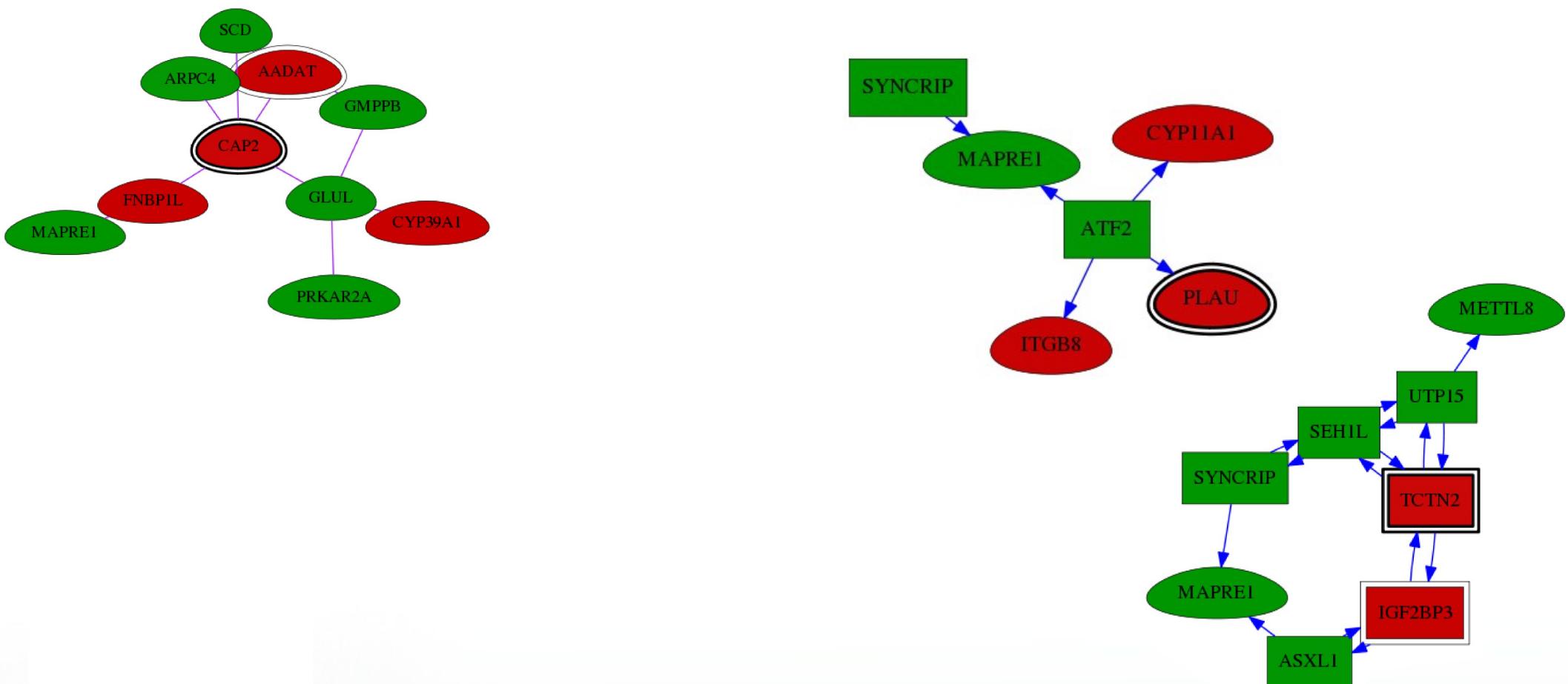


MIMIC

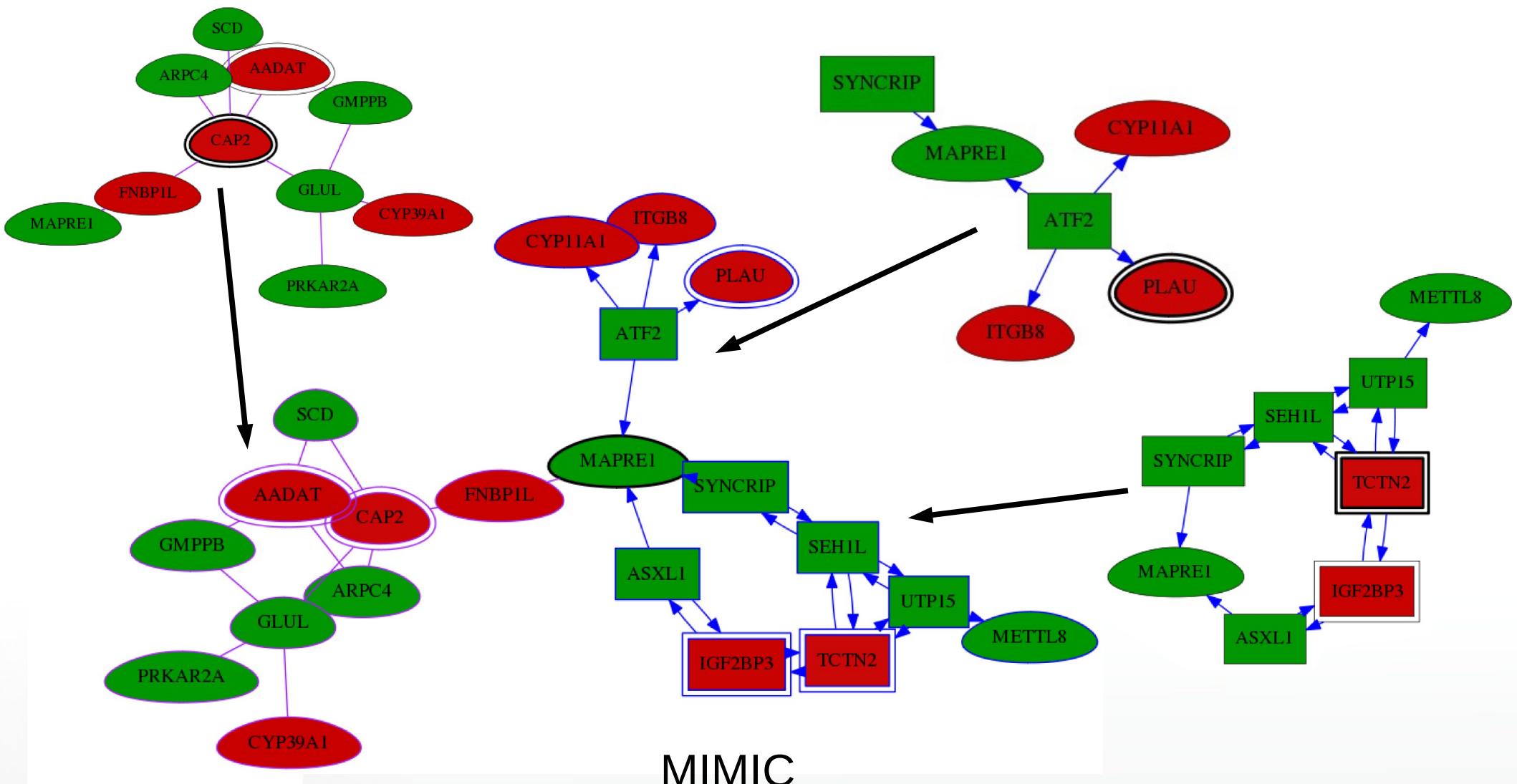


LNA

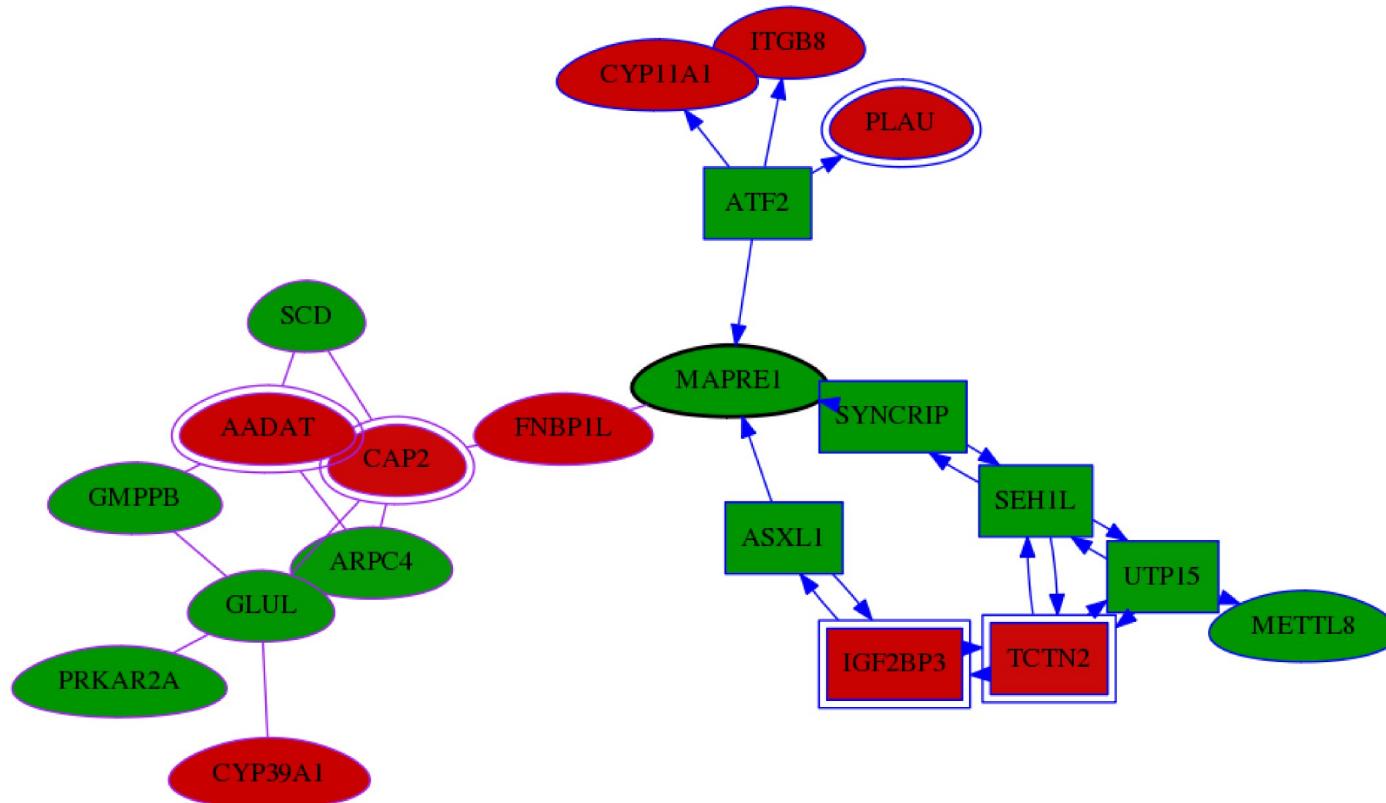
Regulome-Interactome integration



Regulome-Interactome integration



Regulome-Interactome integration



- Mirror genes can be integrated into switching pathways
- Enrichment in GO terms related to cell proliferation

HTML report

Interactome Regulome Integrated analysis

Subnets Genes GO terms

int-snw-10486

Statistics

| Dataset | Score | p-value 1 | p-value 2 | p-value 3 |
|----------------|-------|-----------|-----------|-----------|
| miR600_IPC_rma | 0.991 | 1.33e-01 | 1.51e-01 | 1.39e-01 |

int-snw-10486 subnetwork with miR600_IPC_rma dataset

int-snw-10486 heatmap from miR600_IPC_rma dataset

Genes (10)

| Gene Symbol | Entrez ID | Frequency | miR600_IPC_rma gene score | Best subnet score | Degree | genes miroir |
|-------------------------|-----------|-----------|---------------------------|-------------------|--------|--------------|
| MAPRE1 | 22919 | 2 | -0.998 | 0.991 | 163 | - |
| AADAT | 51166 | 2 | 0.993 | 0.991 | 87 | Yes |
| CAP2 | 10486 | 1 | 0.972 | 0.991 | 83 | Yes |
| GMPPB | 29925 | 3 | -0.996 | 0.991 | 239 | - |
| SCD | 6319 | 1 | -0.993 | 0.991 | 69 | - |
| ARPC4 | 10093 | 1 | -0.981 | 0.991 | 54 | - |
| FNBP1L | 54874 | 1 | 0.999 | 0.991 | 82 | - |
| CYP39A1 | 51302 | 2 | 0.996 | 0.991 | 4 | - |
| PRKAR2A | 5576 | 3 | -0.996 | 0.991 | 179 | - |
| GLUL | 2752 | 2 | -0.991 | 0.991 | 96 | - |

Related GO terms (127)

| Accession number | Name | Hypergeometric test | Corrected p-value | Occurrence in subnet | Occurrences in all subnet genes | Occurrences in all int/reg genes |
|----------------------------|----------------------------------|---------------------|-------------------|----------------------|---------------------------------|----------------------------------|
| GO_0044281 | small molecule metabolic process | 4.848479e-03 | 1.000000e+00 | 4 | 18 | 1099 |

HTML report

Interactome Regulome Integrated analysis

Subnets Genes + GO terms

int-snw-10486

Statistics

Interactome Regulome Integrated analysis

Subnets Genes + GO terms

AADAT

Subnets (2)

| Subnet ID | miR600_IPC_rma score | Size |
|----------------|----------------------|------|
| int-snw-259232 | 0.986 | 18 |
| int-snw-10486 | 0.991 | 10 |

Related GO terms (15)

| Accession number | Name | Occurrences in all subnet genes | Occurrences in all int/reg genes |
|------------------|---|---------------------------------|----------------------------------|
| GO:0034641 | cellular nitrogen compound metabolic process | 3 | 176 |
| GO:0009058 | biosynthetic process | 1 | 9 |
| GO:0044281 | small molecule metabolic process | 18 | 1099 |
| GO:0016212 | kyurenone-oxoglutarate transaminase activity | 1 | 4 |
| GO:0030170 | pyridoxal phosphate binding | 1 | 49 |
| GO:0047536 | 2-aminoacidate transaminase activity | 1 | 1 |
| GO:0033512 | L-lysine catabolic process to acetyl-CoA via saccharopine | 1 | 3 |

Gene info in the NCBI database

[Link](#)

Genes (1)

| Gene Symbol | Entrez ID | Frequency | miR600_IPC_rma gene score | Best subnet score | Degree | genes miroir |
|-------------|-----------|-----------|---------------------------|-------------------|--------|--------------|
| AADAT | 51166 | 2 | 0.993 | 0.991 | 87 | Yes |

Related GO terms (127)

| Accession number | Name | Hypergeometric test | Corrected p-value | Occurrence in subnetwork | Occurrences in all subnet genes | Occurrences in all int/reg genes |
|------------------|----------------------------------|---------------------|-------------------|--------------------------|---------------------------------|----------------------------------|
| GO:0044281 | small molecule metabolic process | 4.848479e-03 | 1.000000e+00 | 4 | 18 | 1099 |

HTML report

Interactome Regulome Integrated analysis

Subnets Genes + GO terms

int-snw-10486

Statistics

Interactome Regulome Integrated analysis

Subnets Genes + GO terms

AADAT

Subnets (2)

Related G

GO:0042127

regulation of cell proliferation

Related GO terms (1)

| Accession number | Name | Occurrences in all subnet genes | Occurrences in all int/reg genes |
|------------------|----------------------------------|---------------------------------|----------------------------------|
| GO:0042127 | regulation of cell proliferation | 3 | 101 |

Term information in AmiGO 2
[Link](#)

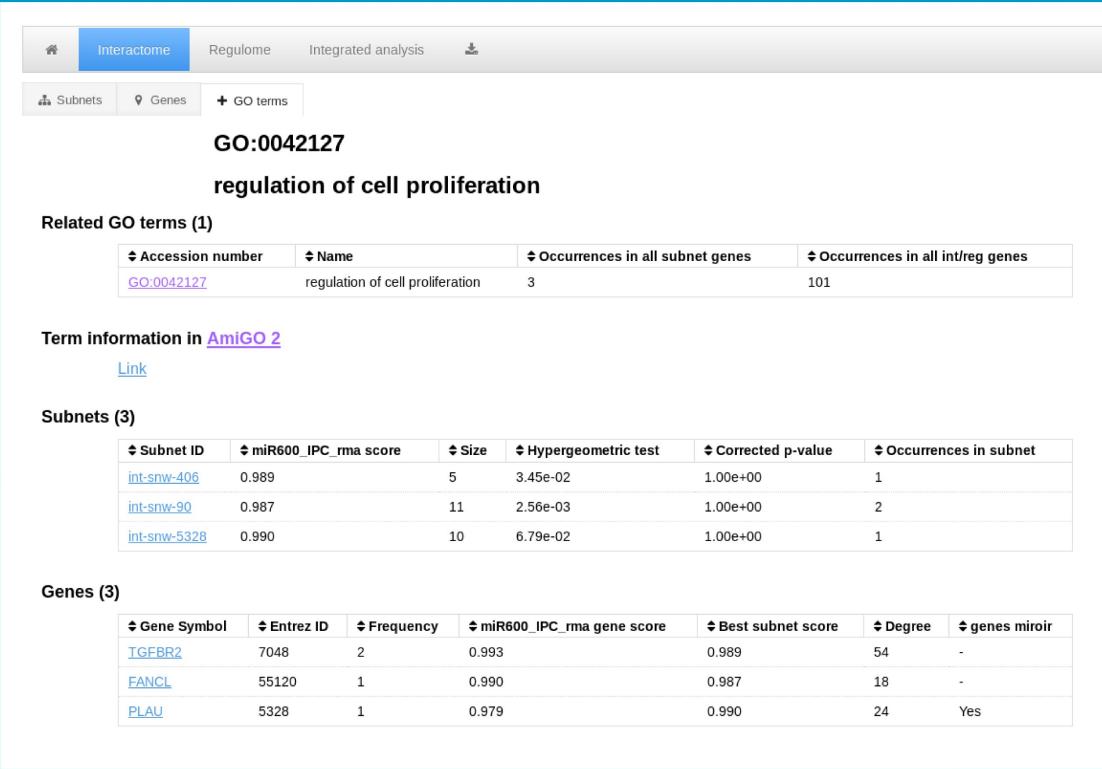
Subnets (3)

| Subnet ID | miR600_IPC_rma score | Size | Hypergeometric test | Corrected p-value | Occurrences in subnet |
|--------------|----------------------|------|---------------------|-------------------|-----------------------|
| int-snw-406 | 0.989 | 5 | 3.45e-02 | 1.00e+00 | 1 |
| int-snw-90 | 0.987 | 11 | 2.56e-03 | 1.00e+00 | 2 |
| int-snw-5328 | 0.990 | 10 | 6.79e-02 | 1.00e+00 | 1 |

Genes (3)

| Gene Symbol | Entrez ID | Frequency | miR600_IPC_rma gene score | Best subnet score | Degree | genes miroir |
|-------------|-----------|-----------|---------------------------|-------------------|--------|--------------|
| TGFBR2 | 7048 | 2 | 0.993 | 0.989 | 54 | - |
| FANCL | 55120 | 1 | 0.990 | 0.987 | 18 | - |
| PLAU | 5328 | 1 | 0.979 | 0.990 | 24 | Yes |

Related GO terms (127)
Accession number
GO:0044281 process



HTML report

The screenshot displays a hierarchical structure of analysis results, likely from a bioinformatics tool. The main title is "int-snw-10486".

- Statistics**:
 - AADAT**
 - Subnets (2)**
 - Related**
 - GO:0042127**
regulation of cell proliferation
- Genes (1)**
 - Term in**
 - Subnet**
 - Gene in**
 - Genes (**
 - Genes (**
- Related GO terms (12)**
 - Accession number**
 - GO:0044281** process

Conclusion

- ITRI: a new **multi-level integrative approach** mixing interactome, regulome, transcriptome data and post-transcriptional information with a candidate approach
- The integrative analysis confirmed the potential **involvement of miR-600 in CSC** differentiation or self-renewal
- Using a system approach to explore the biology of cancer stem cells shows connections between identified switching genes
- Ultimately, elaborating a **new drug targeting CSC pathways** could greatly improve breast cancer treatments & clinical outcome
- See poster A38 for more info!

Thank you

- Integrative Bioinformatics Platform
 - Ghislain Bidaut
 - Quentin Da Costa
 - Samuel Granjeaud
 - Samad El Kaoutari



- Molecular Oncology Team
 - Christophe Ginestier
 - Rita El Helou
- DISC platform
 - Bernard Chetrit
- Plan Cancer 2013-2016 Action 3.6

11100010101011110110100010
1010101110101111110001100
0010101011110011000111101
01011100101110111110010
10000111011100101001



Institut national
de la santé et de la recherche médicale



alliance nationale
pour les sciences de la vie et de la santé