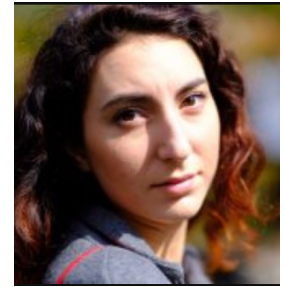# Gaia Ceddia

PhD candidate, 3rd year. Adv: **Marco Masseroli**
Thesis submission: October 2020

Predicting biological interactions for genomic applications using complex network algorithms

Relevant publications/submissions

- **Ceddia G**, Pinoli P, Ceri S, Masseroli M. "Non-negative Matrix Tri-Factorization for data integration and network-based drug repositioning". **IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)**, 2019.

- Dissez G, **Ceddia G**, Pinoli P, Ceri S, Masseroli M, "Drug Repositioning Predictions by Non-Negative Matrix Tri-Factorization of Integrated Association Data". Proceedings of the 10th **ACM International Conference on Bioinformatics, Computational Biology and Health Informatics**, 2019.

- **Ceddia G**, Pinoli P, Ceri S, Masseroli M. "Matrix Factorization-based Technique for Drug Repurposing Predictions". **Under review @ IEEE Journal of Biomedical and Health Informatics**.

- **Ceddia G**, Martino LN, Parodi A, Secchi P, Campaner S, Masseroli M. "Association rule mining to identify transcription factor interactions in genomic regions". **Bioinformatics**, 36 (4), 1007-1013, 2019.

- **Ceddia G**, Pidò S, Masseroli M. "Computational analysis and comparison of gene networks from TCGA normal and cancer data". To be submitted @ **Lecture Notes in Bioinformatics,** 2020.

Interaction networks of Transcription Factors
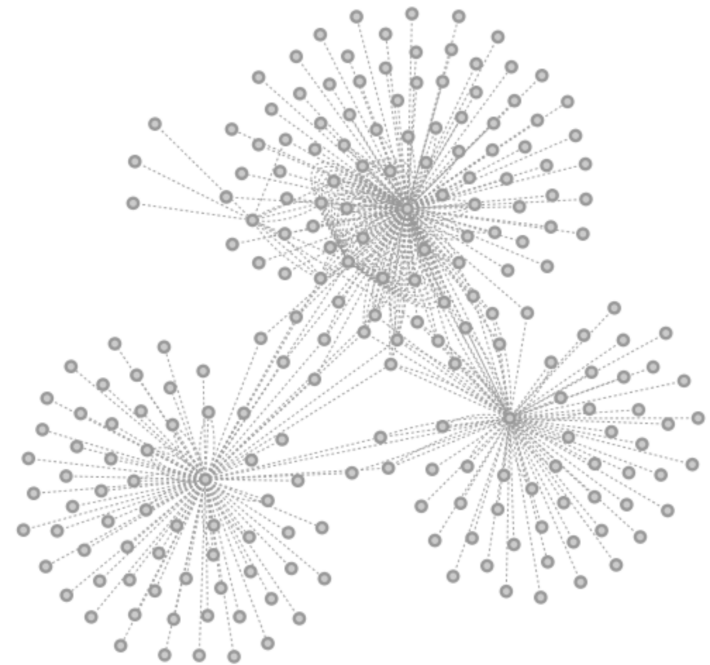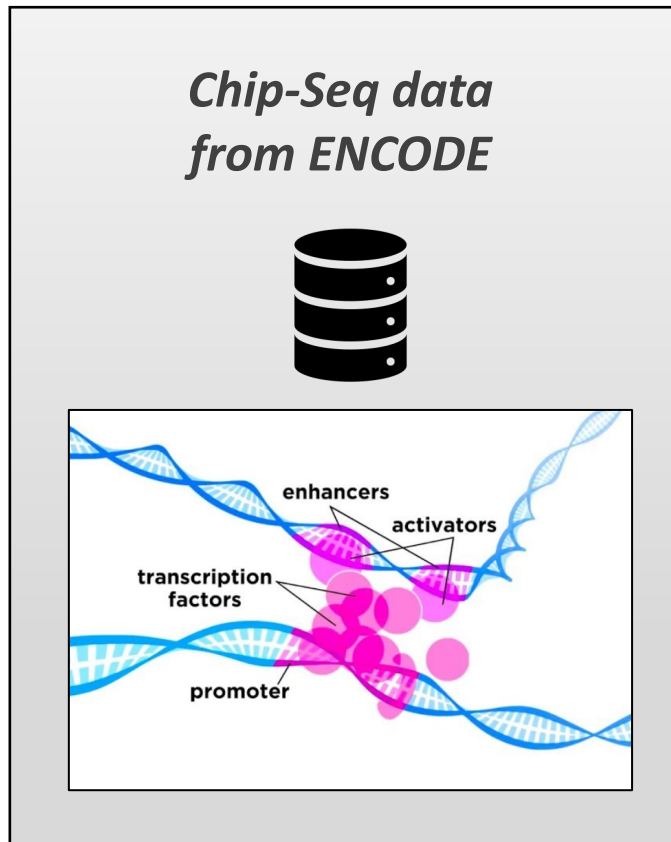
Chip-Seq data from ENCODE

enhancers
activators
transcription factors
promoter

Ceddia G, *Martino LN, Parodi A, Secchi P, Campaner S, Masseroli M. "Association rule mining to identify transcription factor interactions in genomic regions".* **Bioinformatics**, *36 (4), 1007-1013, 2019.*

Complex network integration for drug repurposing

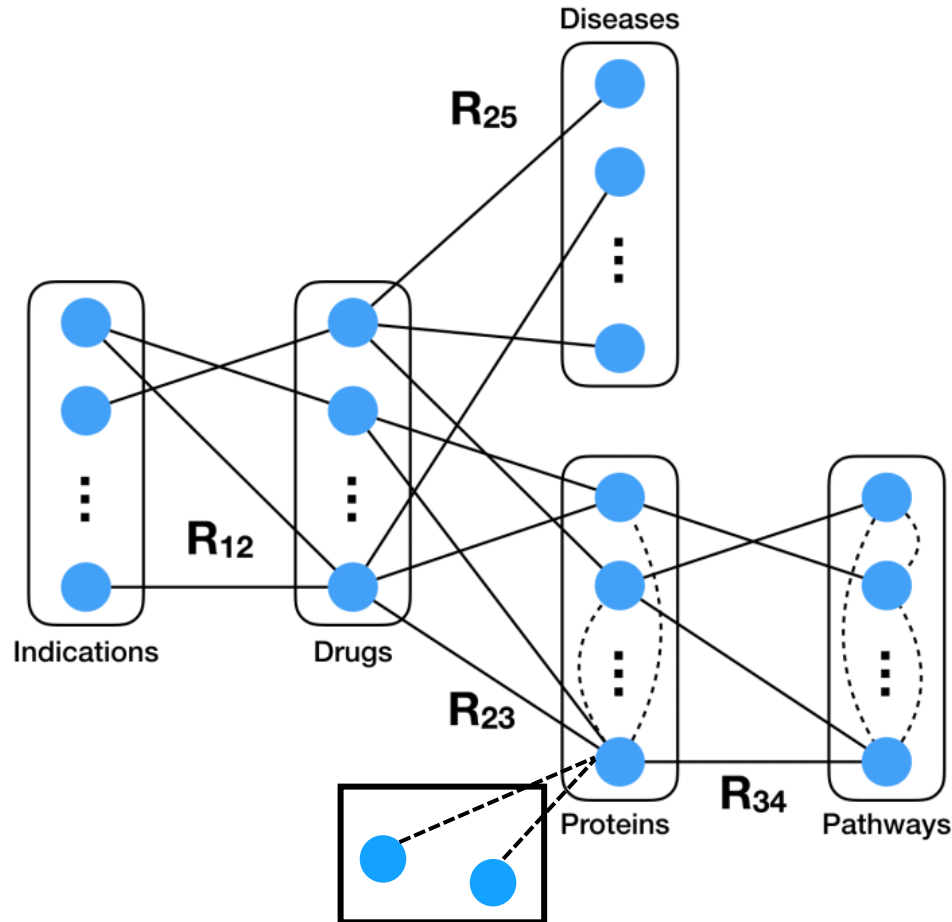*Integration of heterogeneous data from different sources*

Complex network integration for drug repurposing
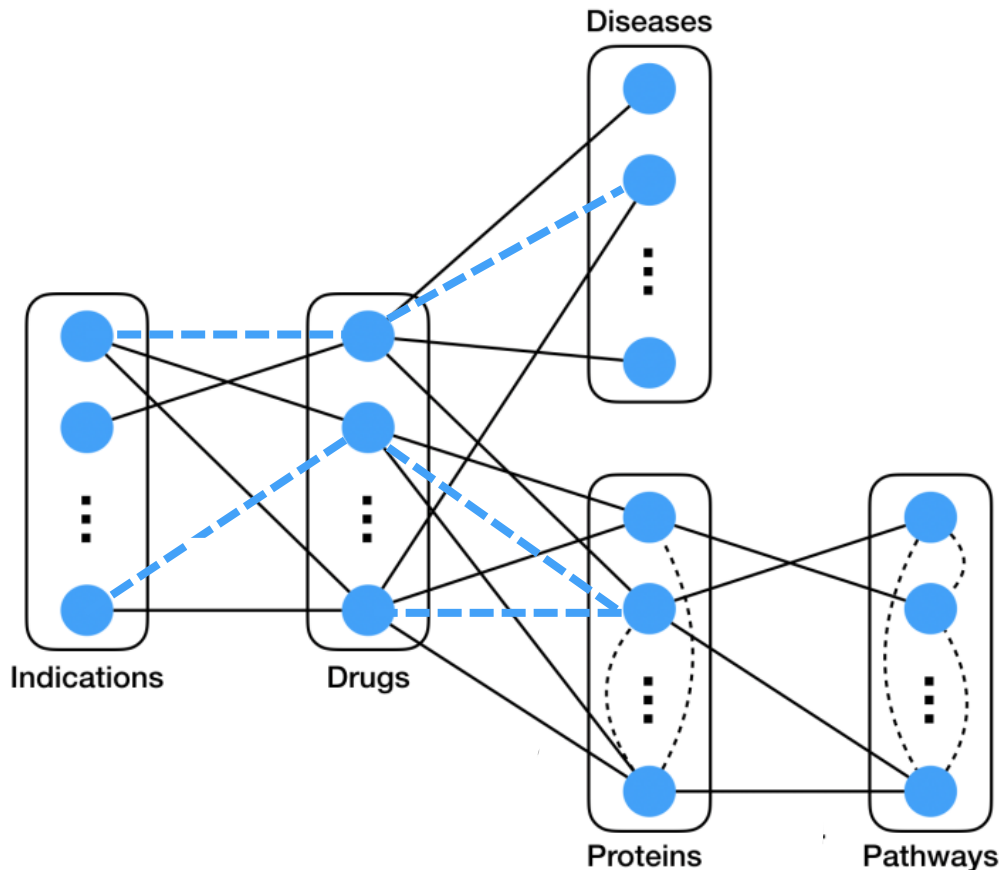
Complex network integration for drug repurposing

*We used the **__Non-Negative Matrix Tri-Factorization__** (NMTF) method simultaneously on R_{12}, R_{23}, R_{25} and R_{34} to **predict Indications-Drugs, Drugs-Proteins and Drugs-Diseases links**.*
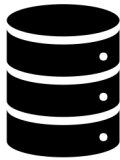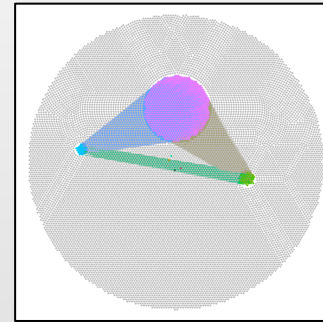
Open problems and questions

- *The computational **validation** shows interesting results and performances; however **literature validation** is weak and can be easily biased.*

- ***Generalization** is not an easy task. Parameter choice is done by using formulas or by grid search of the clustering dispersion coefficient.*

- ***Speed** is different depending by the application.*

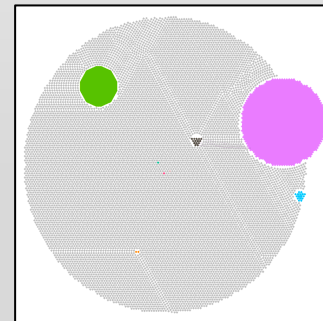# Gene set extraction using complex network algorithms
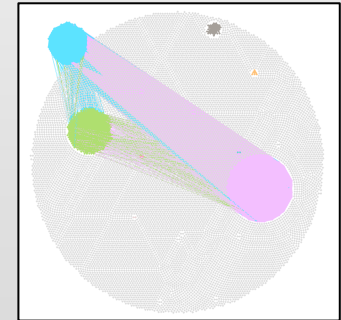
TCGA RNA-Seq and miRNA-Seq data

Gene co-expression networks

prad

kirc

lihc

Open problems and questions

- *We found communities by integrating Normal and Cancer co-expression networks.*

- *Extracted gene sets are **validated** by Normal-Cancer classification. Is it too simple as a problem?*

- *Are there other ways to **biologically** validate their relevance?*