

Experimental design over multi-layer cellular networks

Supervisors: Mohamed Elati, Wajdi Dhifi, Pauline Trebulle

Contact: mohamed.elati@univ-lille.fr

I3-BioNet group (<http://i3-bionet.issb.genopole.fr/>),

INSERM U908, CPAC, Cell Plasticity and Cancer

59655 Villeneuve d'Ascq - Lille

November 21, 2018

1 Context and subject

Scientists now routinely measure, characterize and localize an ever-growing number of molecules at the level of entire biological systems. Thus, despite the continuous expansion of “omic” approaches contributing to the elucidation of systems-level networks, we still know little about the organization of discrete biological activities in space and time, and their integration into larger systems and coherent phenotypes. We propose to rationalize the comprehension of complex genotype-phenotype interactions through a “computational regulatory network” cycle, that we call “I3-BioNet”, composed of three main building “blocks”: the inference (data-learn) [2, 3], interrogation (model-query)[1, 4], and implementation (design-test) of regulatory networks (Fig1).

2 Objective

One of the most important applications of the I3-BioNet concept is the formulation of new hypotheses and the generation of in-silico predictions. This can be used to guide the design of new experiments and thus to kick-start the I3-BioNet engineering cycle. The latter objective could be elaborated by making use of the knowledge generated by the inference/interrogation blocks to introduce a computer-aided experimental design component into the I3-BioNet

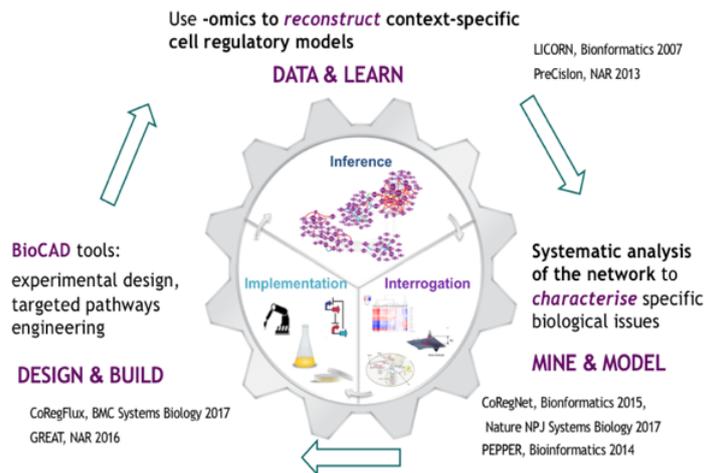


Figure 1: The I3-BioNet concept: A semi-automated framework to elucidate and then investigate cell regulatory networks.

cycle, making it possible to formulate new hypotheses and to generate in silico predictions. The resultant hypotheses and predictions can then be used to guide the design of new high-quality functional experiments (*e.g.*, siRNA, CRISPR/CAS9 system, drug monitoring, *etc*) for scientists (INSERM-ITMO LIONS project: <http://lions.issb.genopole.fr/>) or for a machine acting like a robot scientist (EU AdaLab project: www.adalab-project.org). This objective requires the development, adaptation and combination of state-of-the-art active learning, experimental design and optimization algorithms. Active learning can drastically reduce the cost of performing measurements, and thus significantly reduce the number of iterations (and cost) of the engineering cycle. The second challenge in synthetic biology is the design and engineering of large, self-adaptive, coupled regulatory/metabolic systems at the whole-genome scale, for practical purposes, such as the optimization of versatile chassis Yeast strains for production of valuable Lipid and aromatic compounds (EU CHASSY project: <http://chassy.eu/>). To this end, computer-aided design (BioCAD) has become an indispensable tool for synthetic biology projects. Here, the aim of the implementation module of the I3-BioNet is to provide a tool: i) to optimize and design regulatory circuits associated with pathways or a phenotype of interest, and ii) to integrate the architecture of the genome into the re-factoring of chromosome to ensure the biological stability of the circuit. We try local and global optimization strategies, including stochastic optimization, and integrate the optimization routines with the simulation code to form the core of a single analysis and optimization algorithm.

3 Requirements

Bioinformatics or Statistics / machine learning / Systems Biology, with a strong interest in biology application. Knowledge of either R or Python is essential, and knowledge of both languages would be a plus. Depending on background and interests, the student could work more model-oriented or more machine learning oriented.

References

- [1] Daniel Banos Trejo, Pauline Trébulle, and Mohamed Elati. Integrating transcriptional activity in genome-scale models of metabolism. *BMC systems biology*, 11(7):134, 2017.
- [2] Mohamed Elati, Rémy Nicolle, Ivan Junier, David Fernández, Rim Fekih, Julio Font, and François Képès. PreCisIon: PREdiction of CIS-regulatory elements improved by gene’s positION. *Nucleic Acids Research*, 41(3):1406–1415, February 2013.
- [3] Rémy Nicolle, François Radvanyi, and Mohamed Elati. Coregnet: reconstruction and integrated analysis of co-regulatory networks. *Bioinformatics*, 2015.
- [4] Pauline Trébulle, Jean-Marc Nicaud, Christophe Leplat, and Mohamed Elati. Inference and interrogation of a coregulatory network in the context of lipid accumulation in *Yarrowia lipolytica*. *NPJ systems biology and applications*, 3(1):21, 2017.