Predictive approaches for biological systems engineering

by the group BIOSS and the GDR biosynsys

The objective of this session is to present current research and open problems in synthetic biology that are relevant to the computational systems biology and bioinformatics communities. We will focus more specifically on approaches that aim at improving our capacities to engineer gene expression and metabolism in a predictable manner at the cellular level.

15:00-15:10 Gregory Batt (*Inria/Institut Pasteur*) Introduction

15:10-16:00 Jean-Loup Faulon (*Micalis INRA/AgroParisTech/Univ Paris Saclay, and Manchester Institute of Biotechnology / Univ of Manchester*) Tutorial: Design-Build-Test-Learn paradigm in synthetic biology

Synthetic biology borrows concepts from engineering to design build and test biologically based parts, devices and systems that do not exist in the natural world. Recently, a learning component has been introduced to perform experimental designs and close the loop of the engineering cycle. The first part of the tutorial will show how computational methods can help the synthetic biology process. The 4 steps of the Design-Build-Test-Learn (DBTL) cycle will be unrolled on a practical example of strain engineering for bioproduction. An emphasis will be given to bioinformatics and cheminformatics methods relevant to the Design and Learn steps. Machine learning techniques used to drive the cycle toward a fully automated process will also be outlined. The second part of the tutorial will address the problem of engineering information processing devices for programing living systems. It will be illustrated by the engineering of a metabolic perceptron for neural computing in biological systems.

16:00-16:30 Pierre Millard (*LISBP, Université de Toulouse, CNRS, INRA, INSA*) Improving predictions of metabolic models with isotopic data

Synthetic biology brings engineering tools and perspectives to the design of living systems. Kinetic modeling emerged as a promising road to better understand the functioning of metabolism and ultimately support rational design of synthetic pathways. Kinetic models are typically constructed based on the network topology and calibrated using proteomics and metabolomics data. Isotopic data collected in labeling experiments contain important functional information on metabolism (e.g. they can be used to identify the network topology and to quantify in vivo fluxes), but they cannot be integrated into kinetic models due to the lack of relevant modeling frameworks. We will present novel computational approaches that can be used to improve and validate the predictions of kinetic models using isotopic data, illustrate fundamental and applicative outcomes, and discuss limitations that should be addressed to improve our capacity to control metabolic systems.

16:30-17:00 Coffee break

17:00-17:30 Hidde de Jong (*INRIA Grenoble - Rhône-Alpes and Université Grenoble-Alpes*) Reengineering bacterial metabolism using synthetic biology and optimal control theory

Microorganisms have evolved complex strategies for controlling the distribution of available resources over cellular functions. Biotechnology and synthetic biology aim at interfering with these strategies, so as to optimize the production of metabolites and other compounds of interest, by (re)engineering the underlying regulatory networks of the cell. The resulting reallocation of resources can be described by simple, so-called self-replicator models and the maximization of the synthesis of a product of interest formulated as a dynamic optimal control problem. I will illustrate this approach for the maximization of metabolite production in the case where microbial growth can be switched off through an external control signal. More generally, I intend to show the potential of optimal control theory for better understanding and improving biotechnological production processes.

17:30-18:00 Guillaume Cambray (*DGIMI, INRA/Université de Montpellier*) Probing sequence spaces and biological activities with synthetic biology

Comparative analyses of natural and mutated sequences have been used to probe molecular mechanisms, but contingent evolutionary history and small sample sizes may produce biased outcomes.
DNA editing technologies are providing approaches to overcome these limitations. High-throughput chemical synthesis can be used to program design-of-experiments directly into DNA, thereby producing massive samples tailored for subsequent statistical analyses. High-throughput sequencing can then be used to rapidly and quantitatively characterize these large libraries.
In this talk, I will first show how we designed 244,000 DNA sequences to optimize translation efficiency in *Escherichia coli* and minimize its cost for the cell. In a second part, I will present ongoing efforts to comprehensively dissect entire viral genomes and build holistic descriptions of their functioning.