

GeNetDB: a Software Platform for the Study of Genetic Regulatory Networks

Aurélien Mazurie, Julien Lambert

Laboratoire de génétique de la neurotransmission UMR CNRS 7091, Bat. CERVI 5ème étage,
Hôpital Pitié-Salpêtrière, 83 bd. de l'Hopital, 75013 Paris - France
aurelien.mazurie@free.fr

Keywords. Gene networks, Databases, Data analysis, Graph algorithms, Knowledge representation

Introduction

The software platform GeNetDB (*Genetic Network DataBase*) has been conceived to simplify interdisciplinary studies of gene networks modeling and inference, thanks to a reasoned and exhaustive access to relevant data. The system contains a databank of main published genetic networks and information classes such as expression profiles, and biological properties of their components. Our software allows one to search, visualize, and combine such information by means of a graphical interface or by shell scripts. Outputs can easily be parameterized to allow non biologists to easily access them, within the scope of research in bioinformatics or biomathematics.

GeNetDB is available on our web site, at <http://aurelien.mazurie.free.fr/genetdb.html>

The GeNetDB platform

1. Study of genetic regulatory networks

Genetic regulatory networks (GRN) are composed by all of the regulatory links among genes: a gene can produce a transcription factor that modulates the transcription rate of a target gene, as well as modulate other aspects of the target's activity (gene accessibility, post-transcriptional modifications, translation, post-translational modifications).

The study of those networks must permit the identification of regulation anomalies in pathologies with genetic component, as well as the search for structures having a biological function [1, 2]. More generally this study aims at discovering a gene level organization which could explain the main cell's behavior: growth, differentiation, environment adaptation.

For a long time such studies have remained in the field of theoretical biology, because of the impossibility to observe on a large scale the *in vivo* behavior of the networks components, namely genes and their transcripts. Thanks to technologies like DNA microarrays (for expression data) or ChIP-chips (for genetic networks), such information is now at our disposal. The other main difficulty was to piece together genetic networks without extensive reading of literature: they remain scattered in specialized databases, not very accessible to the non specialist, which moreover are often little or non-formatted, and only hardly available for consultation by automated querying tools.

The goal of GeNetDB is to simplify those studies by proposing a unique and homogeneous source for this data, in an annotated and standardized form, and to permit easy application of any processing, research and visualization algorithm that can be developed by its users.

2. The tool

GeNetDB is organized into a client/server architecture; the client software, written in Java, is installed on the user's computer and has access to the server (MySQL, hosted in our laboratory) or to a stand-alone version of the database (installed by the user). The client software can be modified and adapted either by its source code (available under open-source GNU licence), by creating plugins, or finally by writing shell scripts (in Python, and soon in Java, JavaScript and Tcl). All these elements are gathered together under a single, simple graphical interface, usable by a non computer scientist.

3. State of the project

At this time GeNetDB is still under development. However the data model is finalized, and several genetic networks [3, 4, 5] and expression data sets of the yeast *Saccharomyces cerevisiae* have been integrated. The active phase of the development is currently directed towards client software, especially visualization and processing (motif finding, networks comparison, etc.) functions.

A working version will be presented at the ECCB forum in September 2003.

References

- [1] S. Shen-Orr, R. Milo, S. Mangan and U. Alon, Network motifs in the transcriptional regulation network of *Escherichia coli*, *Nature Genetics*, 31:64-68, 2002
- [2] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, Network motifs: simple building blocks of complex networks, *Science*, 298:824-827, 2002
- [3] N. Guelzim, S. Bottani, P. Bourguin, F. Kepes, Topological and causal structure of the yeast transcriptional regulatory network, *Nature Genetics*, 31:60-63, 2002
- [4] TI. Lee, NJ. Rinaldi, F. Robert, DT. Odom, Z. Bar-Joseph, GK. Gerber, NM. Hannett, CT. Harbison, CM. Thompson, I. Simon, J. Zeitlinger, EG. Jennings, HL. Murray, DB. Gordon, B. Ren, JJ. Wyrick, JB. Tagne, TL. Volkert, E. Fraenkel, DK. Gifford, RA. Young, Transcriptional regulatory networks in *Saccharomyces cerevisiae*, *Science*, 298:799-804, 2002
- [5] I. Simon I, J. Barnett, N. Hannett, CT. Harbison, NJ. Rinaldi, TL. Volkert, JJ. Wyrick, J. Zeitlinger, DK. Gifford, TS. Jaakkola, RA. Young, Serial regulation of transcriptional regulators in the yeast cell cycle, *Cell*, 106:697-708, 2001