Lawrence Berkeley National Laboratory
Recent Work

Title
Web-based Tool for Fast and Accurate de novo Inference of Regulons in the Sets of Closely Related Bacterial Genomes

Permalink
https://escholarship.org/uc/item/0ch5q4ph

Authors
Novichkov, Pavel S.
Stavrovskaya, Elena D.
Gelfand, Mikhail s.
et al.

Publication Date
2009-10-30
Web-based tool for fast and accurate de novo inference of regulons in the sets of closely related bacterial genomes

Pavel S. Novichkov1,6, Elena D. Stavrovskaya1,4, Mikhail S. Gelfand3,5, Andrey M. Mironov6, Inna Dubchak1,5,6, Dmitry A. Rodionov1,4

1Lomonosov Moscow State University, Moscow, 119992, Russia; 2Monash University, Melbourne, Victoria 3145, Australia; 3Institute of Genomic and Bioinformatics, Moscow State University, Moscow, 119992, Russia; 4Rutgers University, Newark, New Jersey 07102, USA; 5Institute of Embryology, Russian Academy of Sciences, Moscow, 119992, Russia; 6Virtual Institute for Microbial Stress and Survival, http://vimss.lit.gov

Introduction

One of the major challenges for the bioinformatics community in view of constantly growing number of complete genomes is providing effective tools to enable high-quality reconstruction of transcriptional regulatory networks (TRN). Definition of a particular TRN includes specification of which transcription factors (TFs) bind to TF-binding sites (TBS) in the promoter regions of which genes and what is the integrated effect of all these TFs on the expression of all these genes. Reconstruction of TRNs helps to better understand the metabolism and functions of bacteria.

Among different approaches that are used for TRN reconstruction are an expression data-driven approach, and comparative genomics approaches that are either computing-driven, or subsystem (pathway)-driven. DNA microarrays, reporting gene expression, continue to be an important tool for high-throughput measurements on transcriptional levels, and machine-learning approaches were used to identify TRN (without a TBS component) from a compendium of microarray expression profiles. However, in many cases the complexity of the interactions between regulons makes it difficult to distinguish between direct and indirect effects on transcription. Availability of a large number of complete genomes opens an opportunity to apply modern approaches of comparative genomics to expand the known regulons to yet uncharacterized organisms and to predict and describe new regulons with high precision.

Comparative genomics

Threshold selection problem

Web Based GUI

Evolutionary regulon conservation score

Testing the platform for de novo regulon inference

Acknowledgments

We are grateful to Mike Wiberg for useful discussions, Minnie Kaiser, Olga Laskina, and Anna Laskina for contributions to regulon reconstructions.

This work was part of the Virtual Institute for Microbial Stress and Survival (http://vimss.lit.gov) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Geothermal Energy Systems (contract DE-AC02-05CH11231), Lawrence Berkeley National Laboratory and the U.S. Department of Energy.