

Lawrence Berkeley National Laboratory

Lawrence Berkeley National Laboratory

Title

Progress in the development of the RegTransBase database and the comparative analysis system

Permalink

<https://escholarship.org/uc/item/0tk6s3sw>

Author

Cipriano, Michael J.

Publication Date

2009-02-08

Progress in the development of the RegTransBase database and the comparative analysis system

Michael J. Cipriano¹, Pavel Novichkov^{1,4}, Alexei E. Kazakov², Dmitry Ravcheev², **Adam P. Arkin**^{1,3,4}, Mikhail S. Gelfand^{2,5}, **Inna Dubchak**^{1,4,6}

¹Lawrence Berkeley National Laboratory, 1 Cyclotron Road. Berkeley, CA 94720, USA; ²Institute for Information Transmission Problems, RAS. Bolshoi Karetny pereulok 19, Moscow, 127994, Russia; ³Department of Bioengineering, University of California, Berkeley, CA, 94710; ⁴Virtual Institute of Microbial Stress and Survival <http://vimss.lbl.gov/>, Berkeley, CA, 94710; ⁵Faculty of Bioengineering and Bioinformatics, Moscow State University. Vorobievsky Gory 1-73, Moscow 119992, Russia; ⁶ Department of Energy Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Acknowledgements

This work was part of the Virtual Institute for Microbial Stress and Survival (<http://VIMSS.lbl.gov>) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics: GTL program through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.

RegTransBase, a database describing regulatory interactions in prokaryotes, has been developed as a component of the MicrobesOnline/RegTransBase framework successfully used for interpretation of microbial stress response and metal reduction pathways. It is manually curated and based on published scientific literature. RegTransBase describes a large number of regulatory interactions and contains experimental data which investigates regulation with known elements. It is available at <http://regtransbase.lbl.gov>.

Over 1000 additional articles were annotated last year resulting in the total number of 5118 articles. We specifically focused on annotating the facts of regulation in bioenergy-related bacteria such as: Clostridia, Thermoanaerobacter, Geobacillus stearothermophilus, Zymomonas, Fibrobacter, Ruminococcus, Prevotella, Acetobacter, Anaeromyxobacterium, Streptomyces, Ralstonia.

Currently, the database describes close to 12000 experiments (30% growth in the last year) in relation to 531 genomes. It contains data on the regulation of ~39000 genes and evidence for ~10000 interactions with ~1130 regulators. We removed redundancy in the list of Effectors (currently the database contains about 630 of them) and turned them into controlled vocabulary.

RegTransBase additionally provides an expertly curated library of 150 alignments of known transcription factor binding sites covering a wide range of bacterial species. Each alignment contains information as to the transcription factor which binds the DNA sequence, the exact location of the binding site on a published genome, and links to published articles. RegTransBase builds upon these alignments by providing a set of computational modules for the comparative analysis of regulons among related organisms.

The new tool - “advanced browsing” was developed to allow a user to search the data contained in RegTransBase in a step-by-step manner. Different types of classifications, such as taxonomy, effectors, the type of experimental result, a phenotype, and genome relevance allow for creating and applying complex search criteria. We are planning to include additional classifications such as metabolic pathways and types of experimental techniques in this scheme.

There is an increasingly tight coupling of RegTransBase with MicrobesOnline in reporting cis-regulatory sites and regulatory interactions, and integrating RegTransBase searches into MicrobesOnLine cart functions.