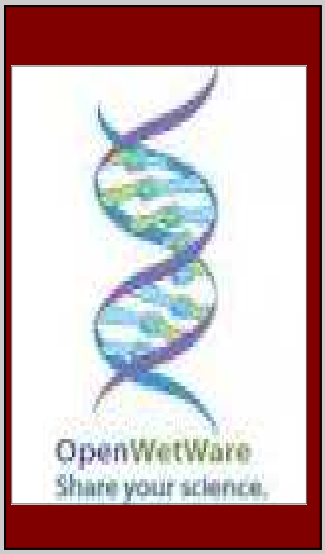
Volume 9, Issue 1 January 2016

Bioinformatics up to Date

(Bioinformatics Infrastructure Facility, Biotechnology Division)
North-East Institute of Science & Technology
Jorhat - 785 006, Assam

Inside…..



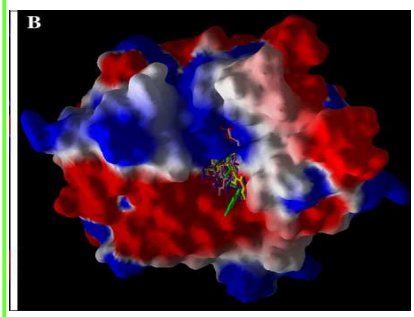
About Us

The Bioinformatics Infrastructure Facility (BIF) at Biotechnology division, CSIR NEIST, Jorhat runs under the Biotechnology Information System Network (BTISnet) programme of DBT, Ministry of Science & Technology, and Government of India. The Centre was established on 2nd February, 2008 to promote innovation in Biological research and education through Bioinformatics accomplishment. The main goal is to facilitate and expose students and researchers from different academic institutions of North East India in Bioinformatics. The center conduct training and workshops for enlightening the use of bioinformatics applications in biological research and development. The Centre has access to global information through 24 hour high speed internet facility, and also e-journal facilities with DeLCON, Science Direct etc. To date the Centre has profoundly extended support in R & D work with a great intensity to different biological discipline including medicinal chemistry, computer aided drug design, genomics and proteomic data analysis etc.

Our Focus

Molecular Docking and In Silico Studies on Analogues of 2-Methylheptyl Isonicotinate with DHDPS Enzyme of *Mycobacterium tuberculosis*

Mycobacterium tuberculosis is one of the most dangerous pathogen infecting one third of the world's population. The cell wall of mycobacteria is characterized by high diaminopimelic acid (DAP) content—an intermediate of the (S)-lysine biosynthetic pathway



and dihydrodipicolinate synthase (DHDPS) enzyme catalyses the first unique reaction of this biosynthesis. Interestingly, the gene knockout experiment demonstrates the essentiality of the DAP pathway, where the absence of DAP results in cell lysis and death. Because of this importance, any inhibitor of DHDPS enzyme may indicate a new class of anti-tubercular agent. In this perspective, the aim of our study was to focus on the molecular docking analysis of DHDPS enzyme against the analogues of 2-methylheptyl isonicotinate—a compound having strong antibacterial property against *Mycobacterium tuberculosis*. The analogues used in the present study were retrieved from the NCBI PubChem database. Further, the top docked compounds at the active site of the DHDPS enzyme were also analyzed for ADME-Toxicity prediction.

[source: Med Chem Res; DOI 10.1007/s00044-013-0488-5]

Algorithm to Automate Process of Searching for Genes

A team of scientists from Germany, USA, and Russia, including Dr. Mark Borodovsky, a Chair of the Department of Bioinformatics at MIPT, have proposed an algorithm, BRAKER1 to automate the process of searching for genes, making it more efficient. The new development combines the advantages of the most advanced tools for working with genomic data. The new method will enable scientists to analyze DNA sequences faster and more accurately and identify the full set of genes in a genome. The paper describing the algorithm published recently in the journal *Bioinformatics* 2015.

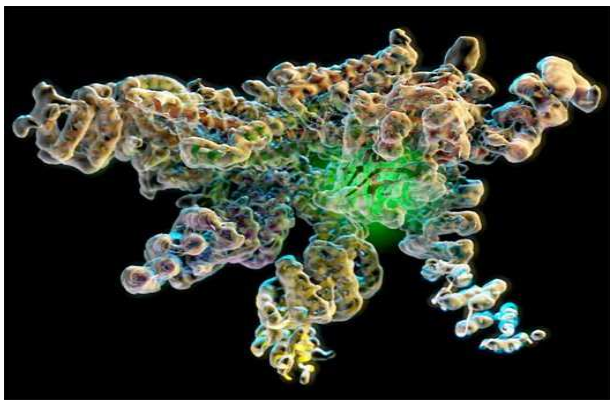
BRAKER1 a pipeline for unsupervised RNA-Seq-based genome annotation that combines the advantages of GeneMark -ET and AUGUSTUS. As input BRAKER1 requires a genome assembly file and a file in bam-format with spliced alignments of RNA-Seq reads to the genome. First, GeneMark-ET performs iterative training and generates initial gene structures. Second, AUGUSTUS uses predicted genes for training and then integrates RNA-Seq read information into final gene predictions. In their experiments, they observed that BRAKER1 was more accurate than MAKER2 when it is using RNA-Seq as sole source for training and prediction. BRAKER1 does not require pre-trained parameters or a separate expert-prepared training step.

The algorithm proposed by the scientists determines which regions in the DNA are genes and which are not. A Markov chain studied in known genes can be used for this. The states of the chain in this case are either nucleotides or nucleotide words (k-mers). The algorithm determines the most probable division of a genome into coding and noncoding regions, classifying the genomic fragments in the best possible way according to their ability to encode proteins or RNA.

[Source: ***BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS.*** *Katharina J. Hoff et al. Bioinformatics (2015)*]

Rpn11: A Deubiquitinase Enzyme

A new study led by scientists at The Scripps Research Institute (TSRI) and the University of California (UC), Berkeley shows how a crucial molecular enzyme starts in a tucked-in somersault position and flips out when it encounters the right target. The new findings, published recently in the journal *eLife* 2016, give scientists a clearer picture of the process through which cells eliminate proteins that promote diseases such as cancer and Alzheimer's.



The new research is the first study in almost 20 years to solve a large component of the proteasome at near-atomic resolution. Lander one of the researcher, said the breakthrough was possible with recent advances in cryo-electron microscopy (EM), an imaging technique in which a sample is bombarded with an electron beam, producing hundreds of thousands of protein images that can be consolidated into a high-resolution structure. Using cryo-EM, scientists investigated part of the proteasome that contains a deubiquitinase enzyme called Rpn11. Rpn11 performs a crucial function called deubiquitination, during which it cleaves molecular tags from proteins scheduled for recycling in the proteasome. This is a key step in proteasomal processing--without Rpn11, the protein tags would clog the proteasome and the cell would die. Rpn11 and its surrounding proteins latch onto the proteasome to form a sort of lid. "The lid complex wraps around the proteasome like a face-hugger in the movie 'Alien,'" said Lander.

Going forward, the researchers hope to use the same cryo-EM techniques to investigate other components of the proteasome--and figure out exactly how it recognizes and destroys proteins.

[source: ***Atomic structure of the 26S proteasome lid reveals the mechanism of deubiquitinase inhibition.*** *Corey M Dambacher, Evan J Worden, Mark A Herzik, Andreas Martin, Gabriel C Lander. eLife (2016)*]

New Computational Framework for Direct Reprogramming Between Human Cell Types

An algorithm developed by the an international team of researchers from the Duke-NUS Medical School (Duke-NUS), the University of Bristol, Monash University and RIKEN that can predict the factors required to convert one human cell type to another. The findings recently published online on 18 January 2016 in the journal *Nature Genetics*, have significant implications for regenerative medicine and lay the groundwork for further research into cell reprogramming.

This approach was brought to the fore by Shinya Yamanaka, whose Nobel prize-winning work involved the reprogramming of fibroblast cells from the skin to induced pluripotent stem cells (iPS). Determining the unique set of cellular factors that is needed to be manipulated for each cell conversion is a long and costly process that involved much trial and error. As a result, this first step of identifying the key set of cellular factors for cell conversion is the major obstacle researchers and doctors face in the field of cell reprogramming.

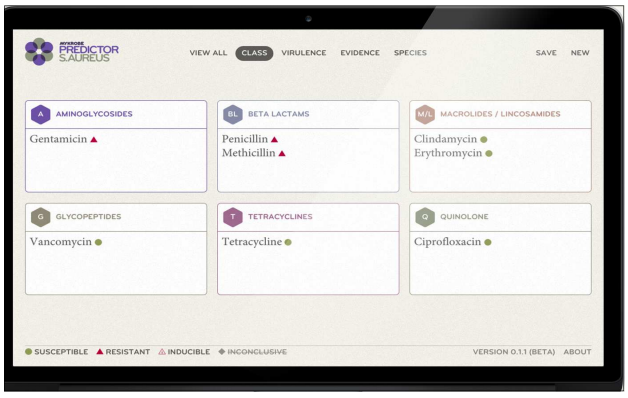
In order to overcome this obstacle, Duke-NUS Senior Research Fellow Dr. Owen Rackham worked for five years to develop a computational algorithm to predict the cellular factors for cell conversions. The algorithm, called Mogrify, is able to predict the optimal set of cellular factors required for any given cell conversion.

To further validate Mogrify's predictive ability, the team conducted two novel cell conversions in the laboratory using human cells, and these were successful in both attempts solely using the predictions of Mogrify.

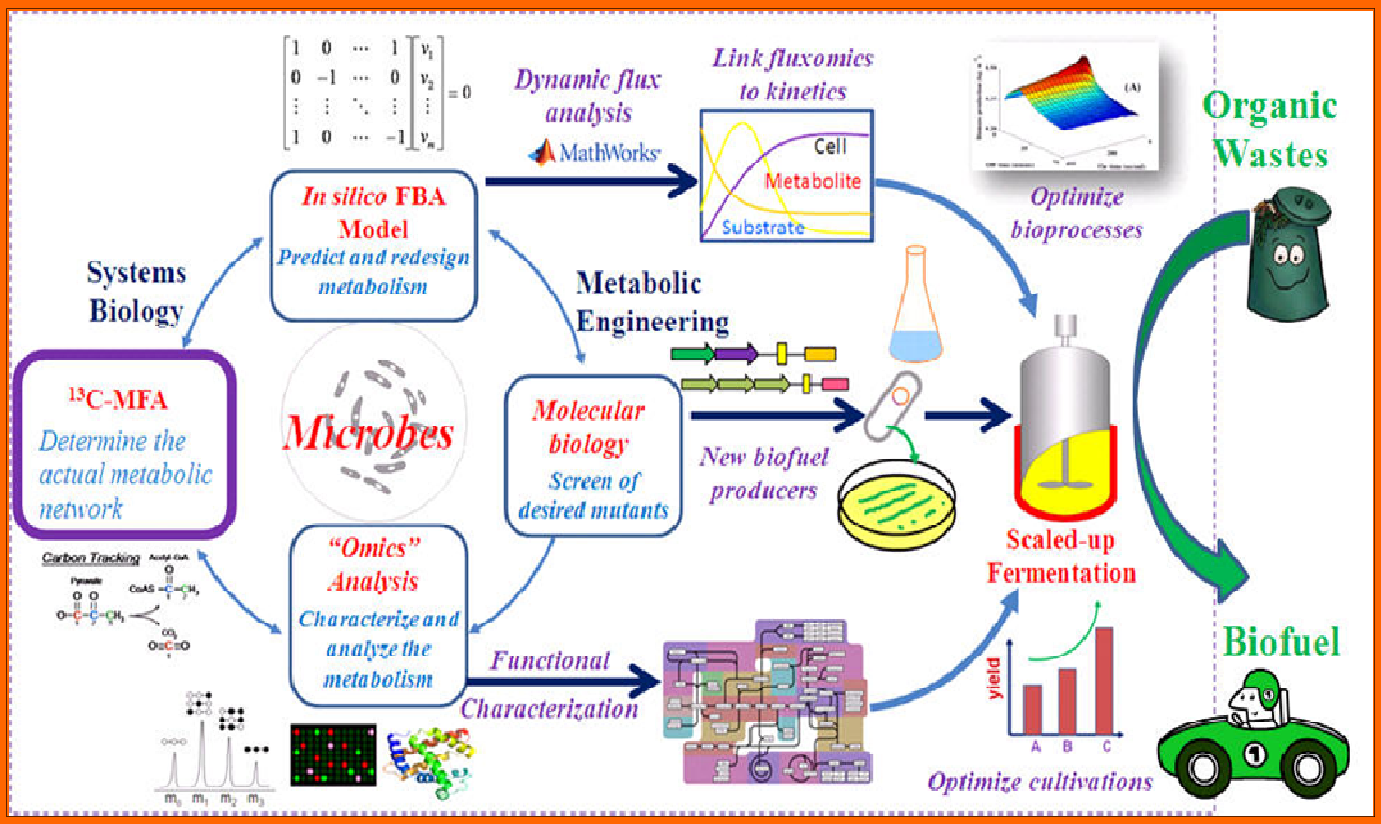
*[**A predictive computational framework for direct reprogramming between human cell types.** Owen J L Rackham et al. Nature Genetics (2016)]*

Mykrobe Predictor : Smart Tools for Microbial Genomes

The Mykrobe predictor is a smart biological tool designed for microbiologists and doctors, providing information needed in order to choose the best treatment. It analyses the whole genome of a bacterial sample, all within a couple of minutes, and predicts which drugs the infection is resistant to. No expertise is needed to run or interpret it, and it works on a standard desktop or laptop. Mykrobe Predictor streamlines this process by automating genome analysis, cross-checking the bacterium's DNA sequence with previous strains to look for resistance-causing mutations and presenting information about the bug in an easy-to-understand format. The Mykrobe Predictor software, developed by DrZaminIqbal and colleagues at the Wellcome Trust Centre for Human Genetics, University of Oxford, runs on a standard laptop or tablet without the need for any specialist expertise. The program can analyse the entire genetic code of a bacterium in under 3 minutes, once a bacterial sample has been cultured and its DNA sequenced.



The software is now being evaluated in hospitals in Oxford, Brighton and Leeds in a project led by Professor Derrick Crook, which in collaboration with parallel programmes at UCL and Cambridge University aims to develop whole genome sequencing as a routine tool for the diagnosis and control of infections within the NHS. The programme is funded by the Department of Health and the Wellcome Trust through the joint Health Innovation Challenge Fund.



patent

Bioinformatics Computation Using A Maprreduce-Configured Computing System

US20080133474A1

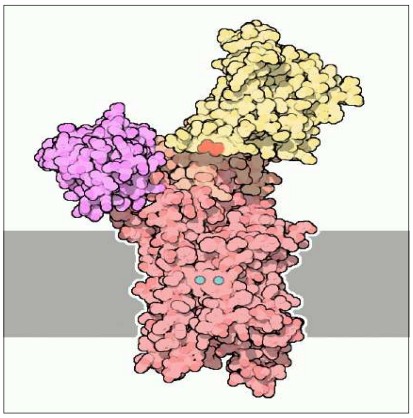
Inventor: Ruey-Lung Hsiao Ali Dasdan Hung-Chih Yang

Abstract

A MapReduce architecture may be utilized for sequence alignment algorithm processing (such as BLAST or BLAST-like algorithms). In addition, a MapReduce architecture may be extended such that memory of the computing devices of a MapReduce-configured system may be shared between different jobs of sequence alignment and/or other bioinformatics algorithm processing, thereby reducing overhead associated with executing such jobs using the MapReduce-configured system.

Calcium Pump

The calcium pump is an amazing machine with several moving parts. It is found in the membrane, as shown here from PDB entry 1su4 . It has a big domain poking out on the outside of the sarcoplasmic reticulum, and a region that is embedded in the membrane, forming a tunnel to the other side. For each ATP broken, it transfers two calcium ions (shown here in blue) through the membrane, and two or three hydrogen ions back in the opposite direction. As shown on the next page, the calcium pump bends and flexes during the pumping cycle.



The calcium pump allows muscles to relax after this frenzied wave of calcium-induced contraction. The pump is found in the membrane of the sarcoplasmic reticulum. In some cases, it is so plentiful that it may make up 90% of the protein there. Powered by ATP, it pumps calcium ions back into the sarcoplasmic reticulum, reducing the calcium level around the actin and myosin filaments and allowing the muscle to relax. The calcium pump goes through a cycle of changes in the process of pumping.

[source: pdb.rcsb.org/molecule of the month]

Upcoming Events

DBT Institutional Biotech Hub Workshop on

Training on Basic Molecular Biology Techniques for Microbial Genotyping

February 24-25, 2016

Biotechnology Group, CSIR-NEIST, Jorhat

6th International Symposium on Current Trends in Drug Discovery & Research

CSIR-Central Drug Research Institute, Lucknow

Kindly send us your feedback to

Dr Ratul Saikia, Robin Das
BIF Center, Biotechnology Group, BSTD
CSIR-North East Institute of Science and Technology, Jorhat, Assam
E-mail: rsaikia19@gmail.com, robindas460@gmail.com