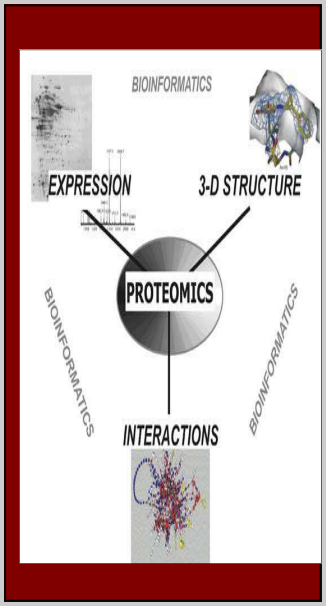
Volume 8, Issue 6 June 2015

Bioinformatics up to Date

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

Inside…..



A new approach to complex, large-scale genome analysis , mSet algorithm

Researchers at EMBL-EBI have developed a new approach to studying the effect of multiple genetic variations on different traits. The new algorithm, published in *Nature Methods*, makes it possible to perform genetic analysis of up to 500,000 individuals and many traits - at the same time.

The relationship between genes and specific traits is more complicated than simple one-to-one relationships between genes and diseases. Genome-wide association studies (GWAS) show that many genetic factors are at play for any given trait, but scientists are just beginning to explore how, specifically, genetic variations affect health and disease. Two major statistical challenges to finding these connections involve analysing associations between many different genetic variants and multiple traits, and making the best use of data from large cohorts that include hundreds of thousands of individual.

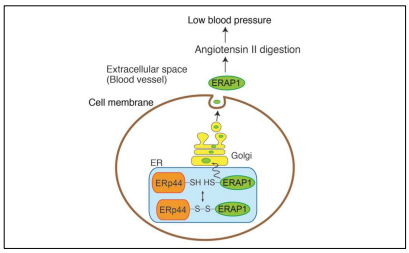
The researchers tested their algorithm on data from two studies from public repositories, and compared the results with existing state-of-the-art tools. Their study of four lipid-related traits (LDL and HDL cholesterol levels, C-reactive protein, triglycerides) proved that the new method is substantially faster, and can explain a larger proportion of these traits in terms of the genetics that drive them.

Using the new method, GWAS researchers can explore several variants of a gene at once while comparing them with several related phenotypes. This makes it much easier to pinpoint which genes - or locations on genes - are involved in a particular function, such as lipid regulation. The new algorithm provides much-needed methods for genomics, making large-scale, complex analysis a manageable and practical endeavour.

"Our method, which we call mSet, provides a principled approach to testing for statistical relationships between multiple genetic variants and groups of traits. These methods will help researchers determine which specific aspects of our biology are inherited, and uncover new insights into the genetics behind our countless biological processes."

Blood Pressure Regulation By a Stress-Sensitive Gatekeeper

Researchers at the RIKEN Brain Science Institute have uncovered a new mechanism for the regulation of blood pressure. Published in Molecular Cell, the study links events at the single-cell level to a system-level effect, showing that blood pressure can drop dramatically if the protein ERAP1 is released from cells and enters the blood stream.



Because high blood pressure is a primary risk factor for stroke, heart disease, and diabetes, understanding how our bodies naturally regulate blood pressure is essential for developing treatments that help keep it at normal levels. Towards this end, the RIKEN team began their investigation when they saw that mice lacking the protein ERp44 had lower than normal blood pressure. "

A direct investigation showed that angiotensin II—a peptide hormone vital for maintaining blood pressure—in these ERp44 knockout mice was removed from circulation faster than in normal mice, explaining the drop in blood pressure. ERp44 is a multi-functional protein located in the endoplasmic reticulum—the place where proteins are folded into their proper shapes before being released into the rest of a cell. To determine why angiotensin II did not remain in the blood of these mice, researchers searched for proteins that bind to ERp44 inside cells and are also able to leave cells and enter the blood stream where they can interact with angiotensin.

[***ERp44 exerts redox-dependent control of blood pressure at the ER.*** *Molecular Cell (2015)*]

Malaria's Doorway to Infect Blood Cells Identified

Scientists have identified a protein on the surface of human red blood cells that serves as an essential entry point for invasion by the malaria parasite. This discovery opens up a promising new avenue for the development of therapies to treat and prevent malaria.

Researchers at Harvard T. H. Chan School of Public Health and the Broad Institute have identified the presence of this protein, called CD55, found to be critical to the Plasmodium falciparum parasite ' s ability to attach itself to the red blood cell surface during invasion.

The five-year study was carried out in collaboration with labs at Harvard Medical School and the Broad Institute. It appears online May 7, 2015 in Science.

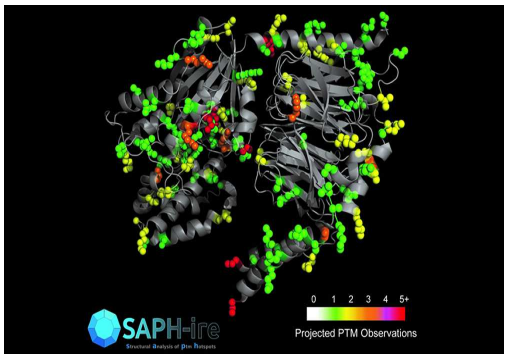
The researchers transformed stem cells into red blood cells, which allowed them to conduct a genetic screen for host determinants of P. falciparum infection. They found that malaria parasites failed to attach properly to the surface of red blood cells that lacked CD55. The protein was required for invasion in all tested strains of the parasite, including those developed in a laboratory as well as those isolated from patients, making it a primary candidate for intervention.

" The discovery of CD55 as an essential host factor for P. falciparum raises the intriguing possibility of host-directed therapeutics for malaria, as is used in HIV, " said Egan. " CD55 also gives us a hook with which to search for new parasite proteins important for invasion, which could serve as vaccine targets. "

[***A forward genetic screen identifies erythrocyte CD55 as essential for Plasmodium falciparum invasion.*** *Science (May 7, 2015)*]

New Informatics Tool Helps Scientists Prioritize Protein Modification Research

Researchers have developed a new informatics technology that analyzes existing data repositories of protein modifications and 3D protein structures to help scientists identify and target research on "hotspots" most likely to be important for biological function.



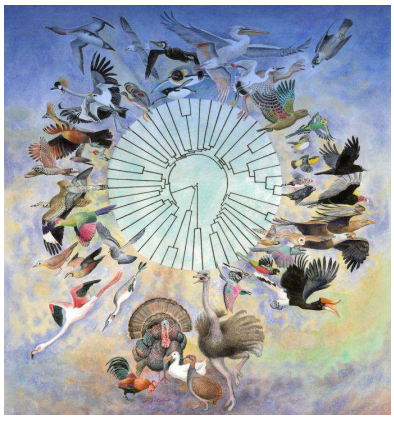
Known as SAPH-ire (Structural Analysis of PTM Hotspots), the tool could accelerate the search for potential new drug targets on protein structures, and lead to a better understanding of how proteins communicate with one another inside cells. SAPH-ire has been tested on a well-studied class of proteins involved in cellular communication, where it correctly predicted a previously-unknown regulatory element.

"SAPH-ire predicts positions on proteins that are likely to be important for biological function based on how many times those parts of the proteins have been found in a chemically-modified state when they are taken out of a cell," explained Matthew Torres, an assistant professor in the School of Biology at the Georgia Institute of Technology. "SAPH-ire is a tool for discovery, and we think it will lead to a new understanding of how proteins are connected in cells."

The tool and its proof-of-concept testing were reported June 12 in the journal Molecular and Cellular Proteomics. The research was supported by the National Institutes of Health's National Institute of General Medical Sciences (NIGMS) and Georgia Tech.

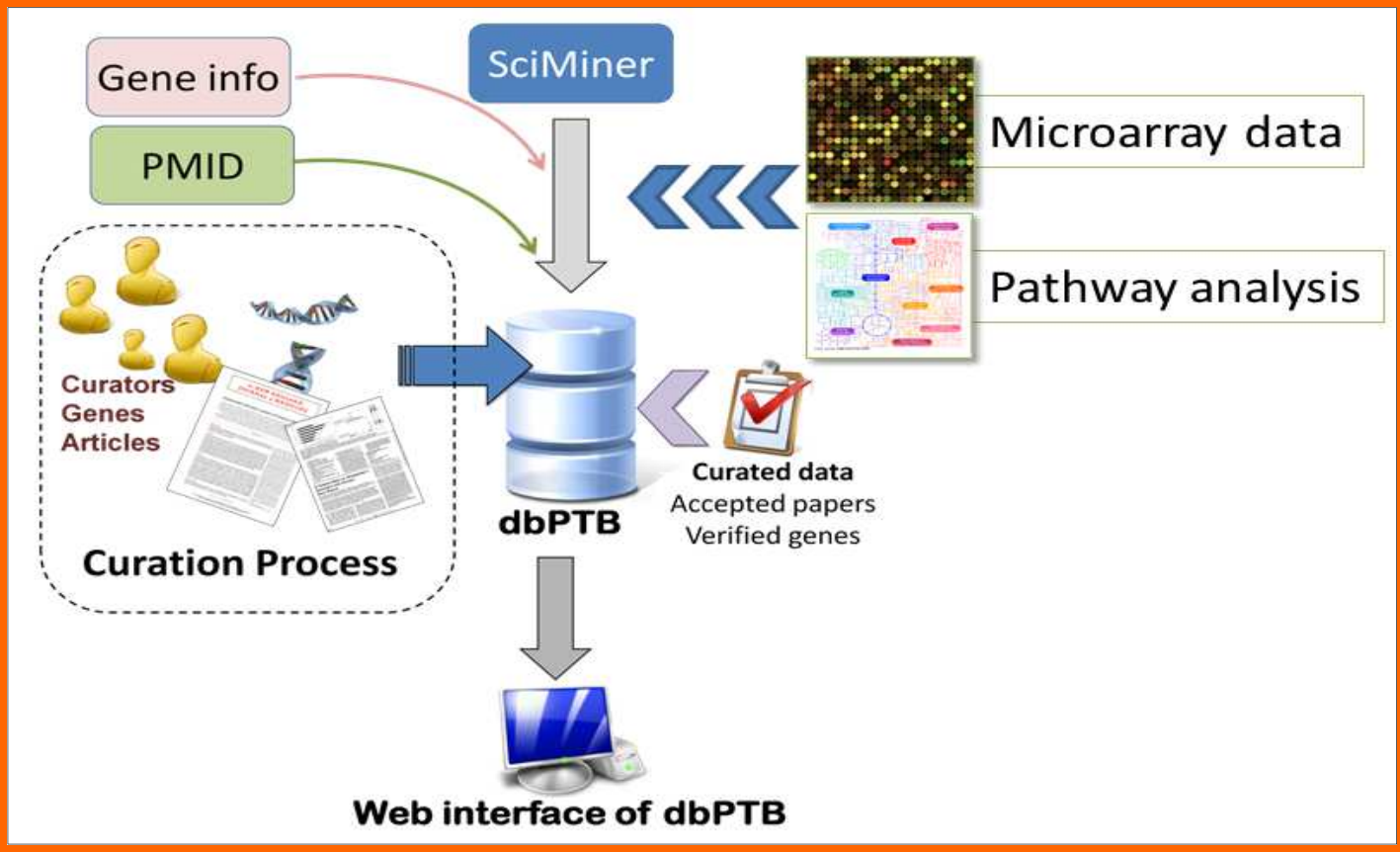
B10K -- Toward decoding all bird genomes

The Avian Phylogenomics Consortium formally announces the launch of the Bird 10,000 genomes (B10K) project, an initiative to generate representative draft genome sequences from all extant bird species within the next five years. This will be the first attempt to sequence the genomes of all living species of a vertebrate class. The establishment of this project is built on the success of the previous ordinal level project, which provided the first proof of concept for carrying out large-scale sequencing of multiple representative species across a vertebrate class and a window into the types of discoveries that can be made with such genomes (1). The announcement of the B10K project is published online today in *Nature*.



The B10K project will allow the completion of a genomic level tree of life of the entire living avian class, decode the link between genetic variation and phenotypic variation, uncover the correlation of genetic evolutionary and biogeographical and biodiversity patterns across a wide-range of species, evaluate the impact of various ecological factors and human influence on species evolution, and unveil the demographic history of an entire class of organisms. Given all these aims, the consortium is carrying out the project in three phases. Each phase focuses on the completion of milestones at hierarchical levels of avian classification (Fig. 2). They envision this project will have significant scientific and public impact that will change the understanding of avian biology and evolution, which in turn will affect the understanding of other organisms and open doors to new areas of research.

Bioinfo. Animation



Patent News

Modular bioinformatics platform

US 20030177143 A1

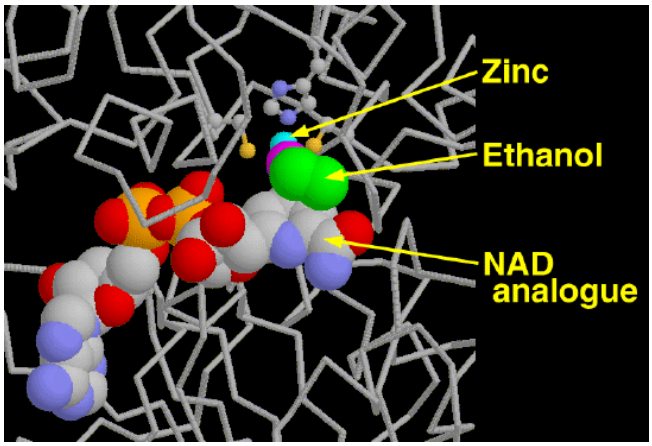
Inventors : Steve Gardner

Abstract

A bioinformatics system and method is provided for integrated processing of biological data. According to one embodiment, the invention provides an interlocking series of target identification, target validation, lead identification, and lead optimization modules in a discovery platform oriented around specific components of the drug discovery process. The discovery platform of the invention utilizes genomic, proteomic, and other biological data stored in structured as well as unstructured databases. According to another embodiment, the invention provides overall platform/architecture with integration approach for searching and processing the data stored in the structured as well as unstructured databases. According to another embodiment, the invention provides a user interface, affording users the ability to access and process tasks for the drug discovery process.

Alcohol Dehydrogenase

Alcohol dehydrogenase provides a line of defense against a common toxin in our environment. But this protection carries with it some dangers. Alcohol dehydrogenase also modifies other alcohols, often producing dangerous products. For instance, methanol, which is commonly used to "denature" ethanol rendering it undrinkable, is converted into formaldehyde by alcohol dehydrogenase. The formaldehyde then does the damage, attacking proteins and embalming them. Small amounts of methanol cause blindness, as the sensitive proteins in the retina are attacked, and larger amounts, perhaps a glassful, lead to widespread damage and death.



Alcohol dehydrogenase uses two molecular "tools" to perform its reaction on ethanol. The first is a zinc atom, which is used to hold and position the alcoholic group on ethanol. The second is a large NAD cofactor (constructed using the vitamin niacin), which actually performs the reaction. PDB entry 1adc, shown here, contains ethanol molecules bound to the two active sites. A slightly-modified version of NAD was used in the structure analysis, so that the enzyme would not immediately attack the ethanol.

Upcoming Events

7th Annual Meeting of Proteomics Society–India (2015)

Venue: VIT Univeristy, Vellore, Tamil Nadu, India
Dates: December, 3rd-6th Dec. 2015
Web: www.psivellore2015.org

Workshop on Protein Structure Prediction and Computer-aided Drug Designing July 24-25, 2015

Excellence in Bioinformatics Bioinformatics Infrastructure Facility
Department of Biochemistry University of Lucknow
Lucknow-226007

Kindly send us your feedback to

Robin Das
Research Fellow; BIF, Biotech Division.
CSIR-North East Institute of Science and Technology, Jorhat, Assam
E-mail: robindas460@gmail.com
Ph No-07399923578