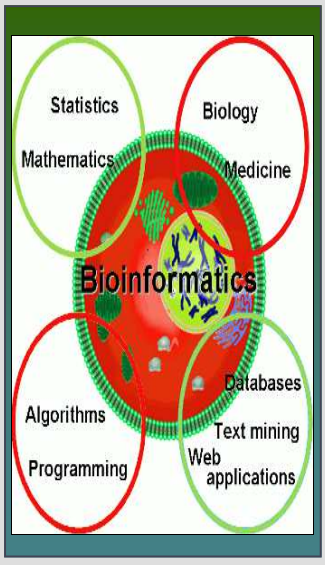
Volume 8, Issue 1 January 2015

Bioinformatics up to Date

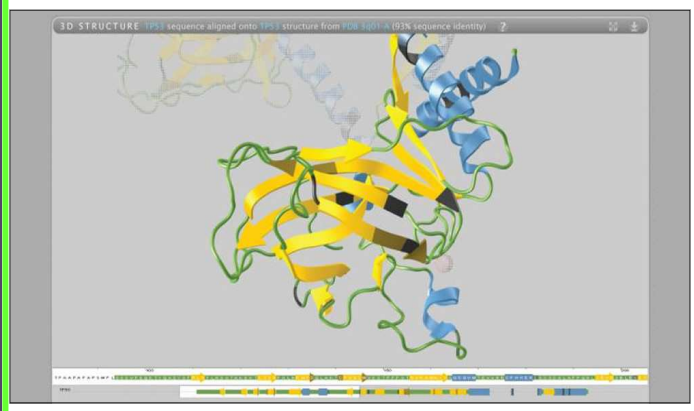
(Bioinformatics Infrastructure Facility, Biotechnology Division)
North-East Institute of Science & Technology
Jorhat - 785 006, Assam

Inside…..



Aquaria: simplifying discovery and insight from protein structure

Life scientists now have access to a publicly available web resource that streamlines and simplifies the process of gleaning insight from 3D protein structures known as Aquaria.



The Aquaria project was led by Dr. Seán O'Donoghue, from The Garvan Institute of Medical Research and CSIRO, in collaboration with Dr Andrea Schafferhans from the Technical University of Munich. The project started in 2009 and involved an international team of around a dozen programmers and bioinformatics experts.

Aquaria is built upon the Protein Data Bank, which contains just over 100,000 protein structures.

"The Protein Data Bank is a fantastic resource containing a wealth of detail about the molecular processes of life, but we were aware that few biologists take full advantage of it," said O'Donoghue.

"Aquaria is fast, it comes with an easy-to-use interface and contains twice as many models as all other similar resources combined. It also allows users to view additional information - such as genetic differences between individuals - mapped onto 3D structures.

"For example, you can add Single Nucleotide Polymorphisms, or 'SNPs', that cause protein changes, then visualise exactly where those changes occur in the protein structure. This provides valuable insight into why proteins sometimes completely change their function as a result of one small change in the DNA code.

Aquaria will be useful to a broad range of life scientists, from medical researchers - at institutes like Garvan - to scientists studying agriculture, biosecurity, ecology and nutrition at institutes like CSIRO. Aquaria's flexibility and extensibility allows information to be combined in completely new ways - quickly and easily. All a scientist needs to do is enter the name of their favourite protein, and then navigate a brave new world of possibility.

Publication: ***Aquaria: simplifying discovery and insight from protein structures.*** *Seán I O' Donoghue et al. Nature Methods (January 2015)*

Population Genomics Unveil Seahorse Domain

In a finding vital to effective species management, a team including City College of New York biologists has determined that the lined seahorse（Hippocampus erectus）is more a permanent resident of the western mid-Atlantic Ocean than a vagrant. Researchers including PhD student J.T. Boehm and Dr. Michael Hickerson of City College decided to test the alternative hypotheses of historical persistence versus the ephemerality of a northern Virginia Province population. They used a dataset consisting of 11,708 randomly sampled spots from the genomes of individuals collected from the eastern Gulf of Mexico to Long Island, N.Y.

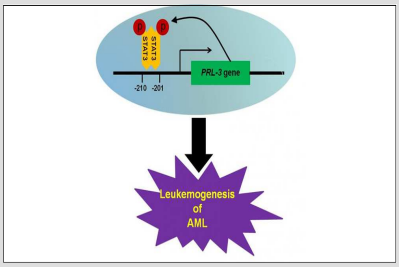
"Concordant results from genomic analyses all infer three genetically divergent subpopulations, and strongly support Virginia Province inhabitants as a genetically diverged and a historically persistent ancestral gene pool," said Boehm.

The results suggested that individuals that emerge in coastal areas during the warm season can be considered "local" and support offshore migration during the colder months.

ARESite

ARESite is an online resource for the investigation of AU-rich elements（ARE）in vertebrate mRNA UTR sequences hosted at the Institute of Theoretical Chemistry, University of Vienna. AREs are one of the most prominent cis-acting regulatory elements found in 3' vertebrate, untranslated regions of mRNAs. Various ARE-binding proteins that possess RNA stabilizing or destabilizing functions are recruited by sequence-specific motifs. This online resource allows detailed investigation of these functional elements by analysis of the phylogentic conservation and the structural context these motifs are embedded in. Moreover, ARESite provides information about experimentally validated targets from extensive literature search. The database is publicly available at: http://rna.tbi.univie.ac.at/ARESite.

New Therapeutic Target for Treatment of Acute Myeloid Leukemia Discovered



A study by the Cancer Science Institute of Singapore (CSI Singapore) at the National University of Singapore (NUS) has found new interactions between two molecules involved in acute myeloid leukemia (AML), STAT3 and PRL-3, which may offer a new therapeutic target for cancer treatment.

The scientists discovered that STAT3, a molecule which has the potential to cause cancer, associates with and regulates the levels of PRL-3, a gene which has been implicated in various types of cancers. Led by Associate Professor Chng Wee Joo, Deputy Director and Senior Principal Investigator at CSI Singapore and Director of the National University Cancer Institute, Singapore, the research team found a connection between PRL-3 and STAT3 for the first time and showed that the STAT3-PRL-3 regulatory loop contributes to the development of AML. They discovered that STAT3, a transcription factor, binds and promotes the production of PRL-3 in cells. A decrease in STAT3 levels led to a corresponding decrease in the levels of PRL-3, and diminished the malignant properties of leukemia cells. The scientists concluded that a disruption of this regulatory loop may offer an attractive anti-AML therapeutic strategy.

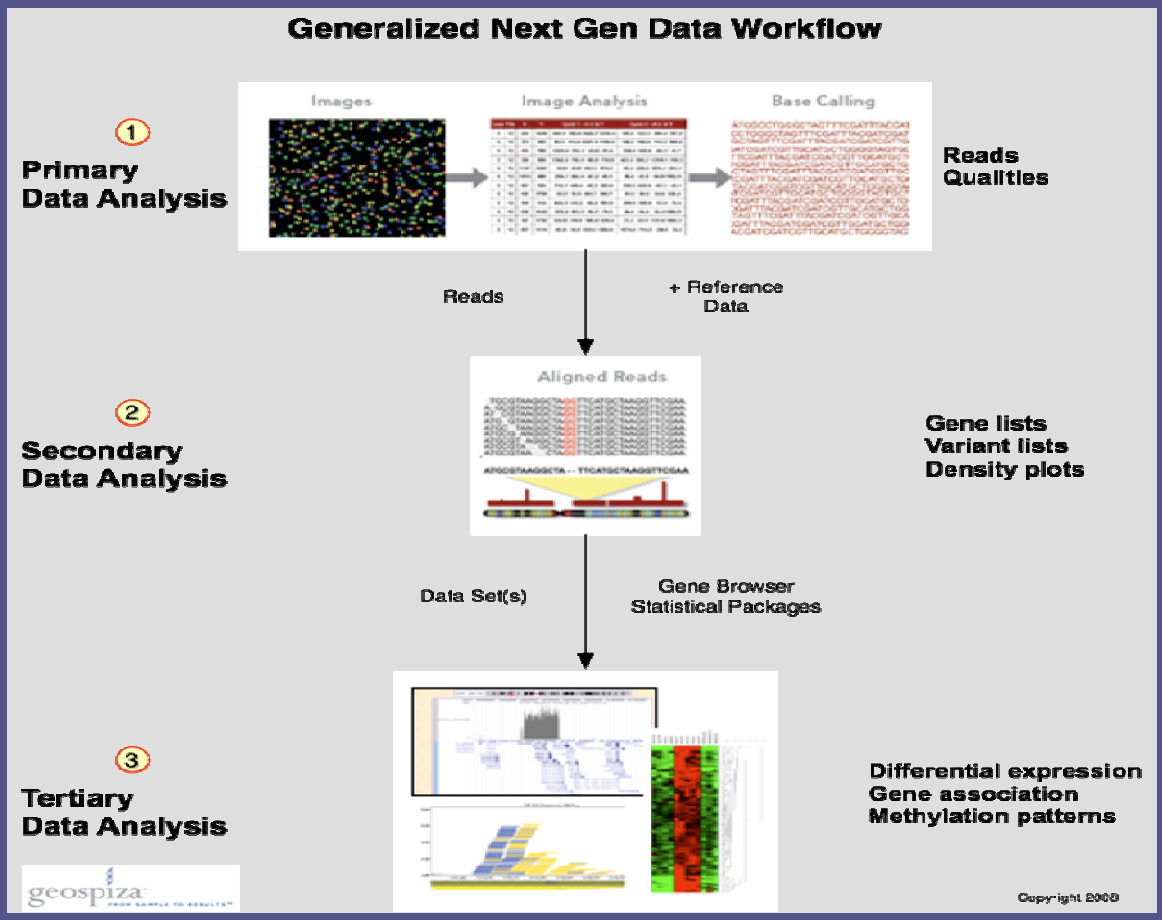
FastPCR

FastPCR is an integrated tool for PCR primers or probe design, *in silico* PCR, oligonucleotide assembly and analyses, alignment and repeat searching. The FastPCR software provides comprehensive and professional facilities for designing any kind of PCR primers for standard, long distance, inverse, real-time PCR (LUX and self-reporting), multiplex PCR, group-specific (universal primers for genetically related DNA sequences) or unique (specific primers for each from genetically related DNA sequences), overlap extension PCR (OE-PCR) multi-fragments assembling cloning and LAMP (Loop-mediated Isothermal Amplification).

FastPCR has the capacity to handle long sequences and sets of nucleic acid or protein sequences and it allowed the individual task and parameters for each given sequences and joining several different tasks for single run. It also allows sequence editing and databases analysis.

The program includes various bioinformatics tools for analysis of sequences with GC or AT skew, CG content and purine-pyrimidine skew, the linguistic sequence complexity; generation random DNA sequence, restriction I-II-III types enzymes and homing endonucleases analysis, find or create restriction enzyme recognition sites for coding sequences and supports the clustering of sequences and consensus sequence generation and sequences similarity and conservancy analysis.

Bioinfo. Animation



Bioinfo. patent

Modular bioinformatics platform

US 20030177143 A1

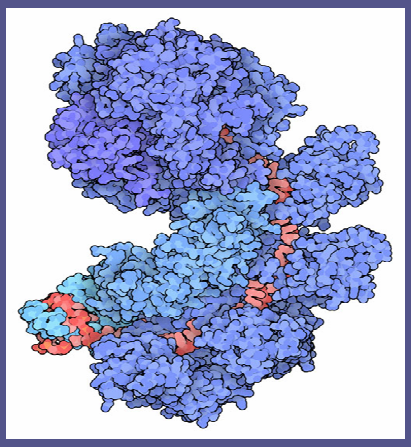
Inventors: Steve Gardner

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.

CRISPR

Living organisms are under constant attack by viruses and have evolved an effective set of weapons to fight them. Bacteria and archaea take several approaches. They have several hard-



wired systems that fight the most common attackers. Bacteria use CRISPR sequences, stored in their genome, to identify attacking viruses. The name stands for "clustered regularly interspaced short palindromic repeats," which refers to the unusual pattern of sequences found in CRISPR DNA. They are composed of many small pieces of viral DNA harvested from viruses that attacked in the past, separated by a distinctive repeated sequence used to create the archive. Remarkably, new sequences are added at the beginning of this collection, so we can read the CRISPR sequence to get a history of viruses that have attacked the bacterial population in the past. A system of Cas proteins (short for "CRISPR-associated proteins") use this stored information to fight the viruses if they try to infect the bacteria again. The center of this system is the large complex Cascade. It carries an RNA transcript of the CRISPR sequence and searches through the cell for matching viral DNA from an infection. If it finds viral DNA, it unwinds it and mobilizes nucleases to cut it up.

International Conference on Chemical and Biological Sciences (ICCBS 2015)

19th to 20th March 2015
Florence, Italy
Website: http://www.iccbs.org/

International Conference on Medical Information and Bioengineering ICMIB 2015

6th to 7th March 2015
Melaka, Malaysia
Website: http://www.icmib. org/

Kindly send us your feedback to

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