Volume 7, Issue 5 May 2014

Bioinformatics up to Date

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

Inside…..

Protein Data Bank

As of Tuesday May 27, 2014 at 5 PM PDT there are 100547 Structures

Atlas Shows How Genes Affect Our Metabolism

In the most comprehensive exploration of the association between genetic variation and human metabolism, researchers have provided unprecedented insights into how genetic variants influence complex disease and drug response through metabolic pathways.

The team has linked 145 genetic regions with more than 400 molecules involved in human metabolism in human blood. This new compendium of associations between genetic regions and metabolite levels provides a powerful tool to identify genes that could be used in drug and diagnostic tests for a wide range of metabolic disorders. The team measured the levels of a large number of metabolites, both those already known and many as yet uncharacterized, from many different metabolic pathways. They found 90 new genetic associations, trebling the figure of known genetic associations with metabolites. In many of the cases where metabolites were known, the team were able to link the molecule to gene function. They mapped genes to their likely substrates or products and linked these to a number of conditions, including hypertension, cardiovascular disease and diabetes.

They also developed an open-access database that allows researchers to easily search through the findings, to understand genetic variants associated with metabolism one metabolite at a time and in the context of the complete metabolic network. This database will facilitate drug discovery for metabolic disorders and also help researchers to understand the biology behind disease.

"This work provides an important new window into the genetic variation underlying human metabolism," said Dr. Eric Fauman, study co-author and Associate Research Fellow from Pfizer Inc. "Through targeted Precision Medicine and by linking human disease genes to in vivo biological markers, we hope to enhance our ability to deliver impactful new medicines for patients across a variety of disorders.

Bioinfo. Carrier

1. **Walk-in Interviews on May 23, 2014 at 10.0 AM for the posts of Research Associate and Senior Research Fellows at NATIONAL RESEARCH CENTRE ON PLANT BIOTECHNOLOGY LBS, CENTRE, PUSA CAMPUS, IARI NEW DELHI**
2. **Systems biology/bioinformatics PhD position, Braga, PORTUGAL; http://www.silicolife.com/jobs-phd-sysbio**

New Gene Expression Mechanism of PRRS Virus

A collaborative study involving Kansas State University researchers has discovered a new gene expression mechanism in porcine reproductive and respiratory syndrome, or PRRS, virus — an important swine pathogen that costs the U.S. pork industry more than $600 million a year. The discovery provides a new avenue for scientists to explore strategies to control and prevent the disease.

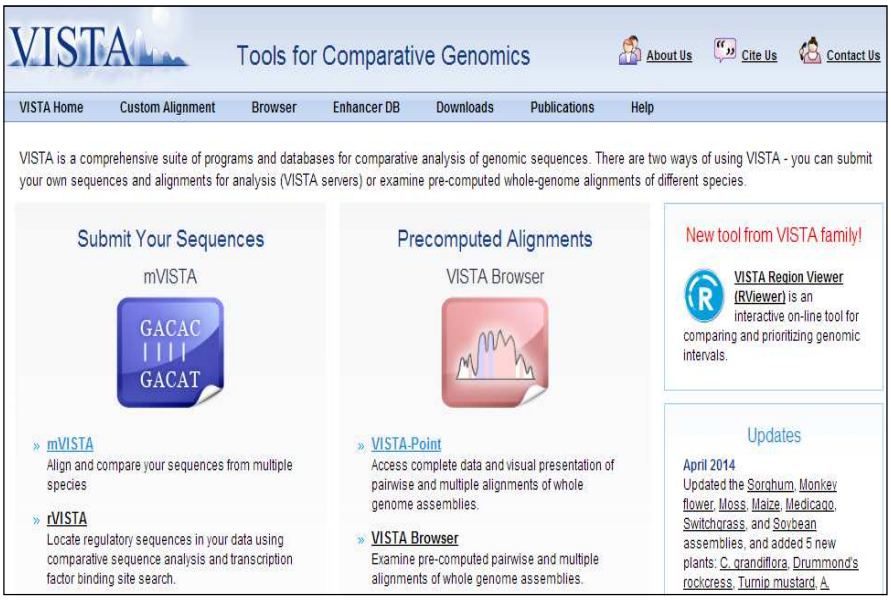
Dr. Ying Fang, associate professor of diagnostic medicine and pathobiology at Kansas State University, led the study that looked at the unique gene expression mechanism of the PRRS virus. She and colleagues found a new protein in the virus, nsp2TF, was generated through novel ribosomal frameshifting signals. The gene expression mechanism called -2 ribosomal frameshifting.

"Frameshifting occurs when a ribosome encounters a 'slippery' sequence and downstream signal in messenger RNA," Fang said. "This causes the ribosome to shift two nucleotides backward, which results in all the genetic codons downstream of the shifted site to be read differently and produce a new protein that has a different function."

With the most recent study, Fang and colleagues have shown that this -2 frameshifting requires a PRRS virus protein, nsp1beta. It is the first time a virus's genetic mechanism has been found to require the action of a transacting viral protein rather than a RNA structure to induce a ribosomal frameshifting, which is novel in the protein translation field.

GenomeVISTA – an integrated software package for whole-genome alignment and visualization

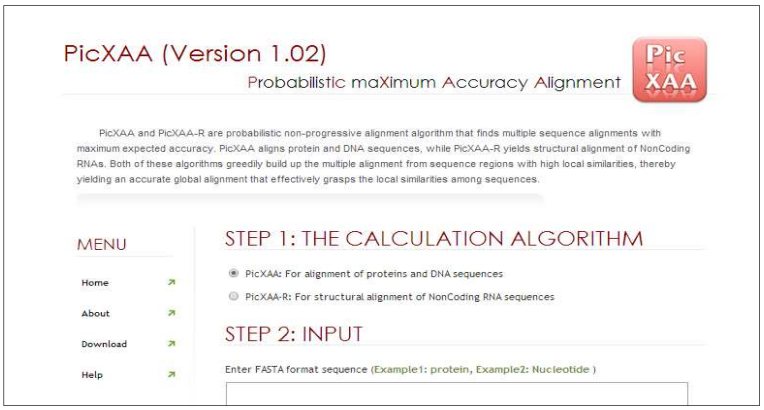
With the ubiquitous generation of complete genome assemblies for a variety of species, efficient tools for whole-genome alignment along with user-friendly visualization are critically important. Our VISTA family of tools for comparative genomics, based on algorithms for pairwise and multiple alignments of genomic sequences and whole-genome assemblies have become one of the standard techniques for comparative analysis. Most of the VISTA programs have been implemented as Web-accessible servers and are extensively used by the biomedical community. In this manuscript we introduce GenomeVISTA: a novel implementation that incorporates most features of the VISTA family – fast and accurate alignment, visualization capabilities, GUI, and analytical tools within a standalone software package. GenomeVISTA thus provides flexibility and security for users who need to conduct whole genome comparisons on their own computers.



Availability and implementation: Implemented in Perl, C/C++ and Java, the source code is freely available for download at the VISTA website:http://genome.lbl.gov/vista/.

PicXAA-Web

PicXAA-Web, a web-based platform for accurate probabilistic alignment of multiple biological sequences. The core of PicXAA-Web consists of PicXAA, a multiple protein/DNA sequence alignment algorithm, and PicXAA-R, an extension of PicXAA for structural alignment of RNA sequences. Both PicXAA and PicXAA-R are probabilistic non-progressive alignment algorithms that aim to find the optimal alignment of multiple biological sequences by maximizing the expected accuracy.

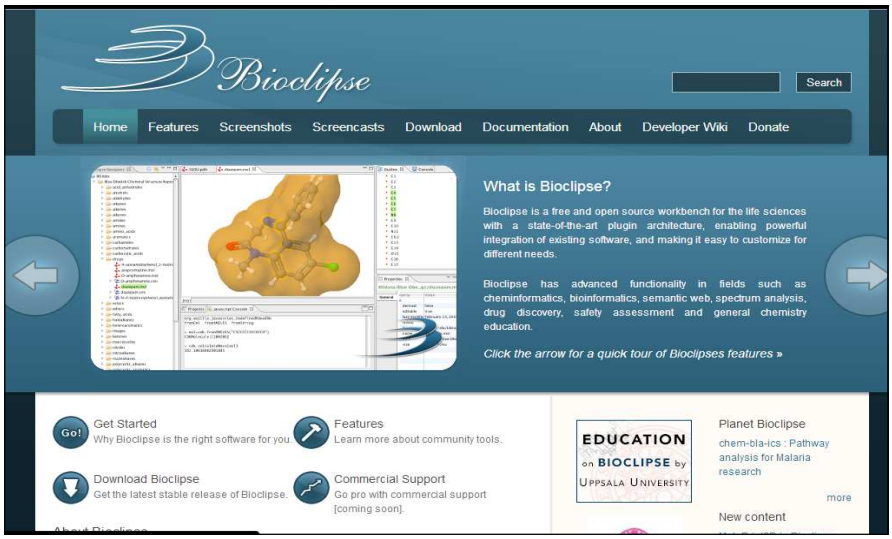


PicXAA and PicXAA-R greedily build up the alignment from sequence regions with high local similarity, thereby yielding an accurate global alignment that effectively captures local similarities among sequences. PicXAA-Web integrates these two algorithms in a user-friendly web platform for accurate alignment and analysis of multiple protein, DNA and RNA sequences.

The main goal of PicXAA is to find the MSA that maximizes the expected number of correctly aligned residue pairs. To this aim, it first computes the posterior pairwise alignment probability $Pa(x_i \sim y_j \mid x, y)$ between residues $x_i \in x$ and $y_j \in y$ for all sequence pairs (x, y) in the input sequence set. PicXAA-Web can be freely accessed at http://gsp.tamu.edu/picxaa/.

Bioclipse

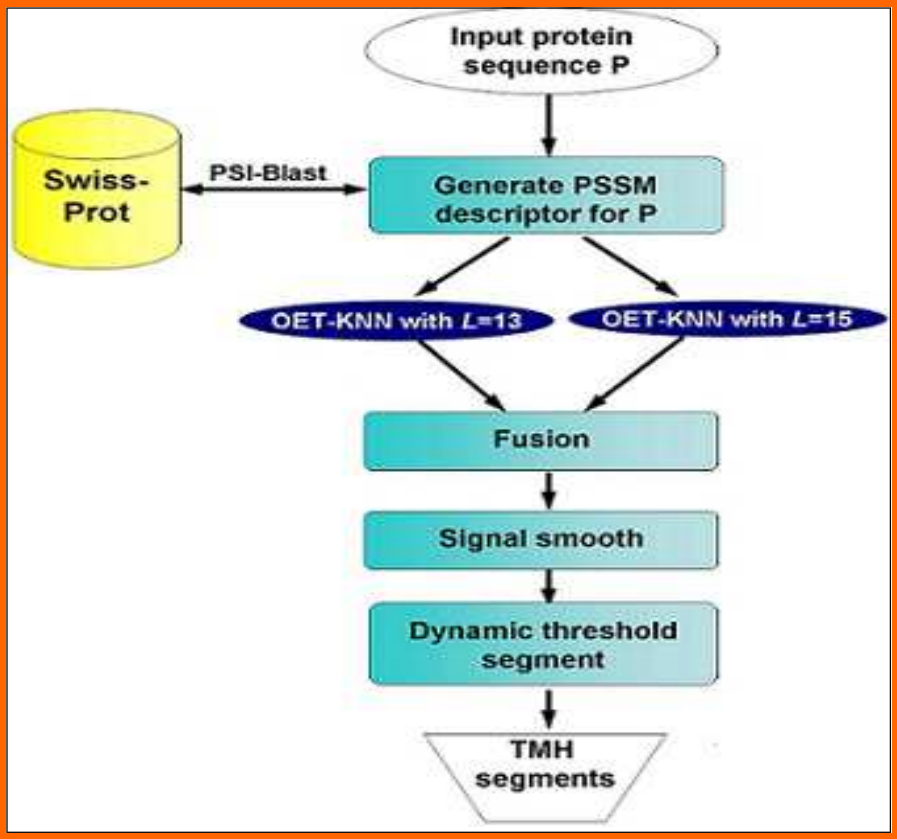
The Bioclipse project is a Java-based, open source, visual platform for chemo- and bioinformatics based on the Eclipse Rich Client Platform (RCP). It gained scripting functionality in 2009. Like any RCP application, Bioclipse uses a plugin architecture that inherits basic functionality and visual interfaces from Eclipse, such as help system, software updates, preferences, cross-platform deployment etc. Via its plugins, Bioclipse provides functionality for chemo- and bioinformatics, and extension points that easily can be extended by other, possibly proprietary, plugins to provide additional functionality.



The first stable release of Bioclipse includes a Chemistry Development Kit (CDK) plugin to provide a chemoinformatic backend, a Jmol plugin for 3D-visualization of molecules, and a BioJava plugin for sequence analysis. Recently, the R platform, using StatET, and OpenTox were added.

Bioclipse is developed as a collaboration between the Proteochemometric Group, Dept. of Pharmaceutical Biosciences, Uppsala University, Sweden, the Christoph Steinbeck Group at the European Bioinformatics Institute, and the Analytical Chemistry Department at Leiden University, but also includes extensions developed at other academic institutes, including the Karolinska Institutet and Maastricht University. The development is backed up by the International Bioclipse Association.

Bioinfo. Animation



Flowchart of TMH (Transmembrane helix) prediction in MemBrain

Patent News

Method for capturing target region and method and system for processing bioinformatics thereof

CN 103547681 A

Inventors: Yuqi ZHU, Xin Jin, Ying Li, Yingrui Li
Applicant: Bgi Shenzhen Co., Limited

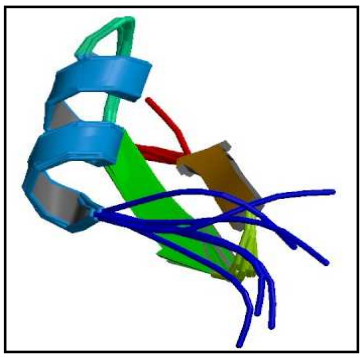
ABSTRACT

The invention discloses a method for capturing target region, and a method and a system for processing bioinformatics thereof. Said method for capturing target region includes: capturing the gene fragments of the species related to the species for which the target region capturing chip is designed with a target region capturing chip, and obtaining the sequences of the gene fragments by sequencing the gene fragments of the related species. The methods and system of the invention help to screen the genome-wide data of the related species by using the gene chip designed for a single species in the target region capturing of other related species.

NMR solution structure of the GS-TAMAPIN MUTATION R6A

PubMed Abstract:

The scorpion toxin tamapin displays the most potent and selective blockage against KCa2.2 channels known to date. In this work, we report the biosynthesis, three dimensional structure, and cytotoxicity on cancer cell lines (Jurkat E6-1 andhuman mammary breast cancer MDA-MB-231) of recombinant tamapin and five related peptides bearing mutations onresidues (R6A,R7A, R13A, R6A-R7A, and GS-tamapin) that were previously suggested to be important for tamapin'sactivity. The indicated cell lines were used as they constitutively express KCa2.2 channels. The studied toxin-like peptidesdisplayed lethal responses on Jurkat T cells and breast cancer cells; their effect is dose- and time-dependent with IC50values in the nanomolar range. The order of potency is r-tamapin > GS-tamapin > R6A > R13A > R6A-R7A > R7A for JurkatT cells and r-tamapin > R7A for MDA-MB-231 breast cancer cells. Our structural determination by NMR demonstrated thatr-tamapin preserves the folding of the αKTx5 subfamily and that neither single nor double alanine mutations affect thethree-dimensional structure of the wild-type peptide. In contrast, our activity assays show that changes in cytotoxicityare related to the chemical nature of certain residues. Our results suggest that the toxic activity of r-tamapin on Jurkatand breast cancer cells could be mediated by the interaction of charged residues in tamapin with KCa2.2 channels via theapoptotic cell death pathway.



Weight:	**3528.20**
Molecule:	**Potassium channel toxin alpha-KTx 5.4**
Polymer:	**1**
Type:	**protein**
Length:	**33**
Chains:	**A**
Mutation:	**R8A**
Organism	**Mesobuthus tumulus**
Release:	**2014-05-28**

[PubMed: 24821061 **DOI:** 10.1021/tx4004193]

Upcoming Events

29th to 30th August 2014
Istanbul, Turkey
Website: http://rdsp-scoop.org/iwbe

ANNA UNIVERSITY, Chennai - 600 025

Training Program on

Next Generation Sequencing (NGS) - Bioinformatics and Data Analysis (2014)

PROGRAMME CONDUCTED BY **AU-KBC Research Centre**

M.I.T Campus of Anna University, Chromepet, Chennai – 600 044. India.
Tel : 91 - 44 – 2223 2711 Fax : 91 – 44 – 2223 2711.

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