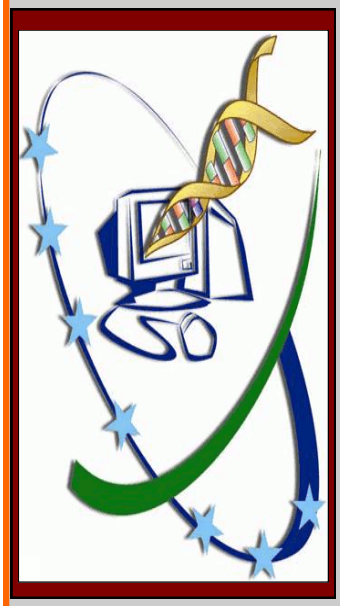
Volume 10, Issue 7 August 2017

Bioinformatics up to Date

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

Inside…..



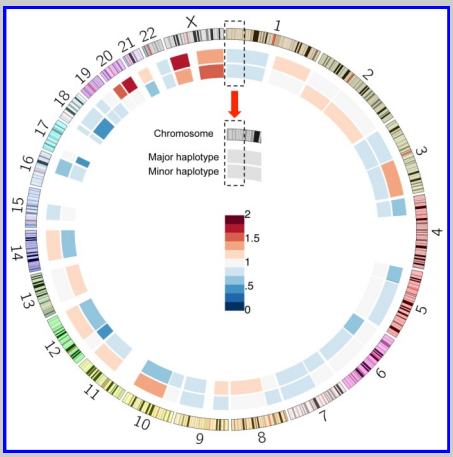
About us

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.

Chromosome-scale mega-haplotypes enable digital karyotyping of cancer aneuploidy

Genomic instability is a frequently occurring feature of cancer that involves large-scale structural alterations. These somatic changes in chromosome structure include duplication of entire chromosome arms and aneuploidy where chromosomes are duplicated beyond normal diploid content

Recent advances in sequencing technology allow the characterization of haplotypes that extend



megabases along the human genome using high molecular weight (HMW) DNA. We developed a method that leverages haplotypes to identify chromosomal segmental alterations in cancer and uses this information to join haplotypes together, thus extending the range of phased variants. With this approach, we identified mega-haplotypes that encompass entire chromosome arms. We characterized the chromosomal arm changes and aneuploidy events in a manner that offers similar information as a traditional karyotype but with the benefit of DNA sequence resolution. We applied this approach to characterize aneuploidy and chromosomal alterations from a series of primary colorectal cancers.

[***Source:*** *https://doi.org/10.1093/nar/gkx712* John M. Bell *et al.* Bioinformatics oxford (August 2017)]

paraGSEA: a scalable approach for large-scale gene expression profiling

More studies have been conducted using gene expression similarity to identify functional connections among genes, diseases and drugs. Gene Set Enrichment Analysis (GSEA) is a powerful analytical method for interpreting gene expression data.

The source code of paraGSEA is licensed under the GPLv3 and available at http://github.com/ysycloud/paraGSEA.

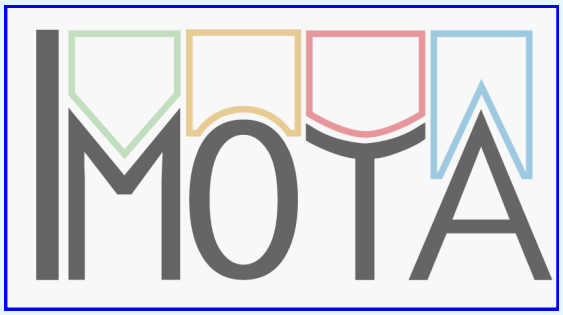
However, due to its enormous computational overhead in the estimation of significance level step and multiple hypothesis testing step, the computation scalability and efficiency are poor on large-scale datasets. We proposed *paraGSEA* for efficient large-scale transcriptome data analysis. By optimization, the overall time complexity of paraGSEA is reduced from $O(mn)$ to $O(m+n)$, where m is the length of the gene sets and n is the length of the gene expression profiles, which contributes more than 100-fold increase in performance compared with other popular GSEA implementations such as GSEA-P, SAM-GS and GSEA2. By further parallelization, a near-linear speed-up is gained on both workstations and clusters in an efficient manner with high scalability and performance on large-scale datasets. The analysis time of whole LINCS phase I dataset (GSE92742) was reduced to nearly half hour on a 1000 node cluster on Tianhe-2, or within 120 hours on a 96-core workstation.

[***Source:*** https://doi.org/10.1093/nar/gkx679 Shaoliang Peng *et al.* Bioinformatics oxford (31 july 2017)]

IMOTA: an interactive multi-omics tissue atlas for the analysis of human miRNA–target interactions

Web repositories for almost all 'omics' types have been generated—detailing the repertoire of representatives across different tissues or cell types. A logical next step is the combination of these valuable sources. With IMOTA (interactive multi omics tissue atlas), we developed a database that includes 23 725 relations between miRNAs and 23 tissues, 310 932 relations between mRNAs and the same tissues as well as 63 043 relations between proteins and the 23 tissues in *Homo sapiens*

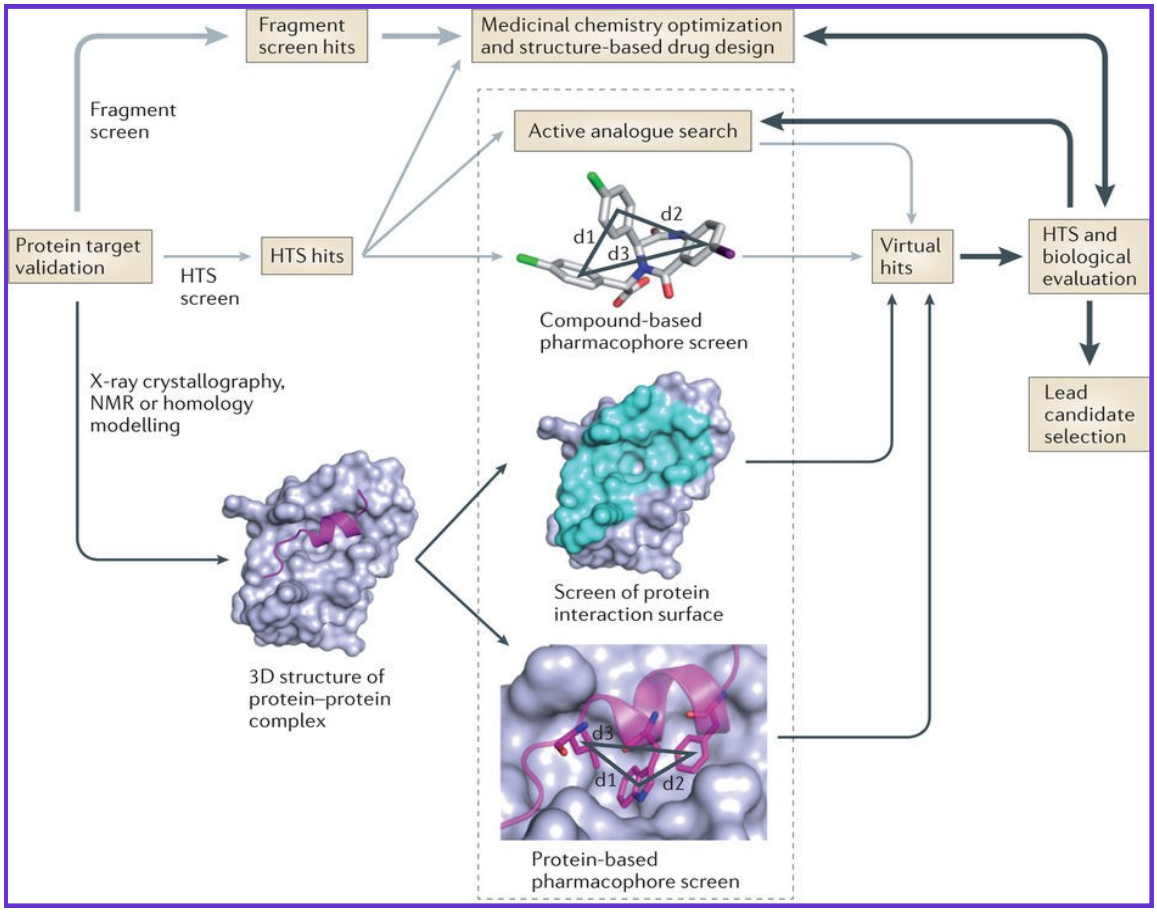
The IMOTA repository is freely available online at https://ccb-web.cs.uni-saarland.de/imota/.



IMOTA relies on the work of different research groups. Generally, the data resources can be divided in two parts. First, background databases that contain general information on genes, proteins, miRNAs, targets of miRNAs or tissues. Second, databases that store the actual expression profiles of the omics data in tissues. IMOTA also contains data on tissue-specific interactions, e.g. information on 331 413 miRNAs and target gene pairs that are jointly expressed in the considered tissues.

[***Source:*** https://doi.org/10.1093/nar/gkx701 Valeria Palmieri *et al.* Bioinformatics oxford (August 2017)]

Bioinfo. Animation

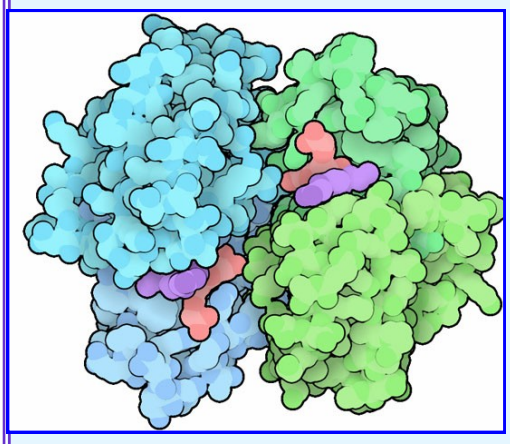


Structure based drug designing

Molecule of the month

Glutathione Transferases

Glutathione transferase tags toxic molecules, making them easy to recognize and remove. Cells are filled with a confusing jumble of small molecules. Sometimes dangerous molecules get introduced in the mix, and cells need a way to find and remove them. In a first line of defense, enzymes like cytochrome p450 modify slippery carbon-rich toxins, making them more soluble.



In a second line of defense, the glutathione transferases attach a convenient handle to unwanted molecules, which is then recognized by the cell's export machinery. A collection of just over 20 glutathione transferases work together to scour each cell for toxins. These twenty enzymes protect us from many different perils, including toxins made by bacteria and fungi, reactive molecules formed during the cooking of food or smoking, and a variety of environmental pollutants. Because of this, each glutathione transferase typically recognizes a variety of foreign molecules, attaching them all to glutathione. This job is so important that in some cells, such as liver cells, glutathione transferase can make up 10% of the total protein content of the cell.

[Source: http://pdb101.rcsb.org/motm/212]

Upcoming events

Patents

Information System for Biological and Life Sciences Research

US 20170076231 A1

Inventors : Ramin Cyrus

ABSTRACT

An online life science research environment and virtual community with a focus on design and analysis of biological experiments includes a life sciences laboratory system employing at least one networked computer system that defines a virtual research environment. Users access the system through a portal associated with the networked computer system (s). The virtual research environment has a data coupling mechanism by which the user designates a set of user-specified data for bioinformatics processing. A processor(s) associated with the networked computer system (s) performs bioinformatics services upon the user-specified data. In one embodiment, the data coupling mechanism enables transfer of the user-specified data to a memory space that is mediated or accessed by the processor performing the bioinformatics processing. Users may thus exploit bioinformatics processing resources that are not deployed on users' local computer environments, and to store and organize information relating to life sciences research in a secure, online workspace.

Kindly send us your feedback to

Dr Yumnam Silla Devi
BIF Center, Biotechnology Group, BSTD
CSIR-North East Institute of Science and Technology, Jorhat, Assam
E-mail: bio.sillayumnam@gmail.com.

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