Volume 9, Issue 12 December 2016

Bioinformatics up to Date

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

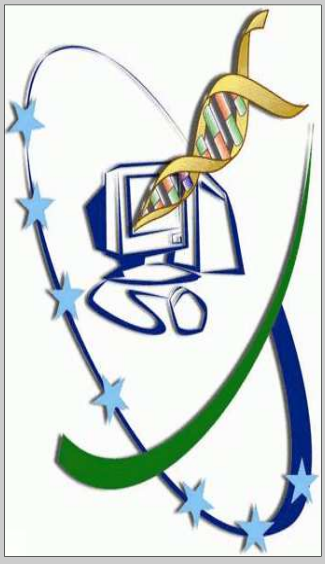
Inside…..

Advisor:

Dr D Ramaiah

Editors:

Mr Robin Das
Dr Y S Devi
Dr R Saikia
Dr H P Deka Baruah



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.

OmniPath: New tool of literature curated human signaling pathways

Omnipath a new software tool developed by researchers in the UK and Germany that helps researchers see biological signalling pathways with unprecedented accuracy. The work published in journal *Nature Methods,* OmniPath offers a comprehensive, unified collection of literature-curated signalling pathways based on an analysis of 41 000 scientific papers.

The data in OmniPath are primarily based on small-scale experiments, but its Pypath software makes it possible to add datasets obtained from large screening experiments or converted from reactions. Pypath (a Python module) lets users build custom signalling networks and combine them with other data. It is a powerful tool for incorporating pathways into bioinformatics workflows, and makes the analysis behind OmniPath fully open source, transparent and easily reproducible.

OmniPath covers approximately three times more proteins (7,984) and four times more interactions (36,557) than the largest causal resource it contains. It covers ~39% of the human proteome, 61% of disease–gene associations, >80% of cancer-related genes and 54% of druggable proteins (as compared to 13%, 42%, 55% and 22%, respectively, in the largest casual resource. OmniPath encompasses 41,237 references from 1,132 journals. On average, each interaction is supported by 2.88 references. OmniPath integrates additional information on the structure and mechanism of the interactions, drug targets, functional annotation, tissue-specific expression and mutations to increase its applicability.

[D Turei, T Korcsmaros and J Saez-Rodriguez (2016) OmniPath: guidelines and gateway for literature-curated signaling pathway resources. Nature Methods 13(12)]

Researchers Develop Machine Learning Technique that Helps Identify Cancer Cell Types

The researchers from Brown University have developed a new image analysis technique to distinguish two key cancer cell types associated with tumour progression. The epithelial-mesenchymal transition(EMT) is a process by which more docile epithelial cells transform into more aggressive mesenchymal cells. Tumours with higher numbers of mesenchymal cells are often more malignant and more resistant to drug therapies. The new technique combines microscopic imaging with a machine learning algorithm to better identify and distinguish between the two cell types in laboratory samples. The training was done by using an epithelial cell line cultured in a petri dish that serves as a model for human breast cancer. The researchers activated a transcription factor called Snail that is well known to cause these cells to quickly undergo an extreme form of EMT. The researchers showed that, after training, the algorithm was able to categorize individual cells as either epithelial or mesenchymal with greater than 92 percent accuracy.

The team then used the algorithm to analyze sets of cells that undergo EMT triggered by pathways less well studied than that used in the training set. They treated epithelial cells with a compound TGF-beta1 which promotes rapid cell growth and is also thought to induce EMT. They showed that the growth factor induced EMT more slowly than in the training set, and produced changes in cell shape that were subtler. Then the researchers looked at epithelial cells treated with the chemotherapy drug Taxol. Recent research has suggested that Taxol and other drugs, when delivered in sub-lethal doses, could induce EMT in the cells they fail to kill. In that way, the drugs may actually prime the tumours to become more drug resistant.

It's a preliminary finding that will require much more study to fully understand, Wong says. But it could shed light on how tumours become resistant to Taxol and other drugs. With more development, the researchers think their technique could provide a new means to screen the effectiveness of cancer drugs.

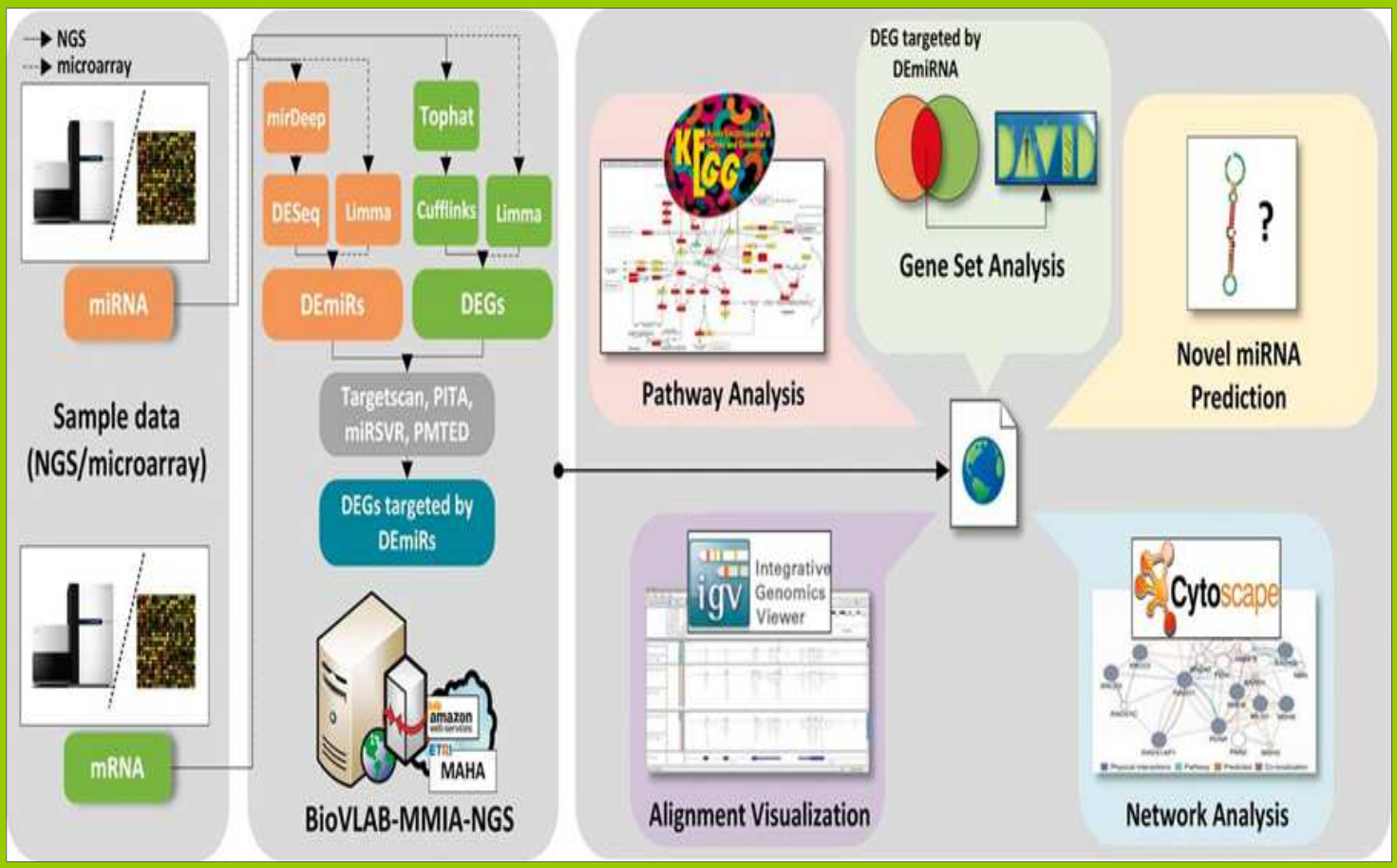
[SOURCE: ***Morphological single cell profiling of the epithelial–mesenchymal transition.*** *Susan E. Leggett et al. Integr. Biol. (2016):oct 12]*

HEDD: the human epigenetic drug database

Epigenetic drugs are chemical compounds that target disordered post-translational modification of histone proteins and DNA through enzymes, and the recognition of these changes by adaptor proteins. Epigenetic drug-related experimental data such as gene expression probed by high-throughput sequencing, co-crystal structure probed by X-RAY diffraction and binding constants probed by bio-assay have become widely available. The mining and integration of multiple kinds of data can be beneficial to drug discovery and drug repurposing. HEMD and other epigenetic databases store comprehensively epigenetic data where users can acquire segmental information of epigenetic drugs. However, some data types such as high-throughput datasets are not provide by these databases and they do not support flexible queries for epigenetic drug-related experimental data. Therefore, in reference to HEMD and other epigenetic databases, a relatively comprehensive database has been developed for human epigenetic drugs. The human epigenetic drug database (HEDD) focuses on the storage and integration of epigenetic drug datasets obtained from laboratory experiments and manually curated information. The latest release of HEDD incorporates five kinds of datasets: (i) drug, (ii) target, (iii) disease, (vi) high-throughput and (v) complex. In order to facilitate data extraction, flexible search options were built in HEDD, which allowed an unlimited condition query for specific kinds of datasets using drug names, diseases and experiment types. Database URL:http://hedds.org/

[Source: Database (Oxford). 2016 Dec 26;2016. pii: baw159. doi: 10.1093/database/baw159]

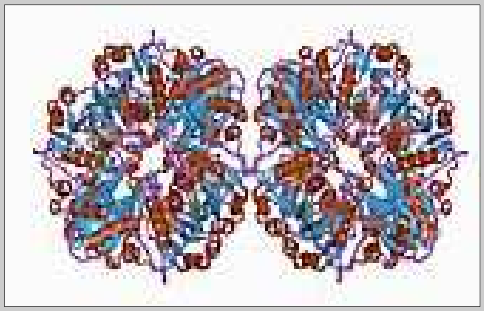
Bioinfo. Animation



Molecule of the month

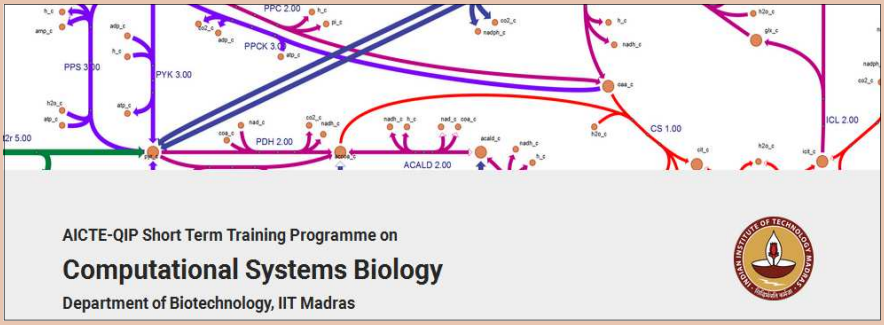
Glyceraldehyde 3-Phosphate Dehydrogenase

Glyceraldehyde 3-phosphate dehydrogenase (abbreviated as GAPDH or less commonly as G3PDH) (EC 1.2.1.12) is an enzyme of ~37kDa that catalyzes the sixth step of glycolysis and thus serves to break down glucose for energy and carbon molecules. In addition to this long established metabolic function, GAPDH has both glyceraldehyde-3-phosphate dehydrogenase and nitrosylase activities, thereby playing a role in glycolysis and nuclear functions, respectively. Participates in nuclear events including transcription, RNA transport, DNA replication and apoptosis. Glyceraldehyde-3-phosphate dehydrogenase is a key enzyme in glycolysis that catalyzes the first step of the pathway by converting D-glyceraldehyde 3-phosphate (G3P) into 3-phospho-D-glyceroyl phosphate. Component of the GAIT (gamma interferon-activated inhibitor of translation) complex which mediates interferon-gamma -induced transcript-selective translation inhibition in inflammation processes. Upon interferon-gamma treatment assembles into the GAIT complex which binds to stem loop-containing GAIT elements in the 3'-UTR of diverse inflammatory mRNAs (such as ceruplasmin) and suppresses their translation.



Under normal cellular conditions, cytoplasmic GAPDH exists primarily as a tetramer. This form is composed of four identical 37-kDa subunits containing a single catalytic thiol group each and critical to the enzyme's catalytic function.[5][6] Nuclear GAPDH has increased isoelectric point (pI) of pH 8.3–8.7.[6] Of note, the cysteine residue C152 in the enzyme's active site is required for the induction of apoptosis by oxidative stress

Upcoming events



AICTE STTP on Computational Systems Biology 2017

February 6-11, 2017

Training–cum–Workshop on Biological Data Analysis Using R-Package

January 5–6, 2017
CSIR-North East Institute of Science and Technology
Jorhat 785006, Assam, India

Patents

E-genechip online web service for data mining bioinformatics

WO 2003008963 A1

Inventors Eugenia Wang, William Christopher Hall, Xuechun Zhao

Abstract

A method for data analysis of microarrays. The method includes accessing a software program that performs a multistage analysis of the biological image, through an internet webserver. The analysis includes at least a step of comparing the digitized quantitative data for each sample. The method may further include digitizing the data for each sample and quantitating the data for each sample. Further, the method may include using a software program which quantitates the intensity and size of each sample; compares the quantitative value for each sample with the quantitative value of one or more controls; captures data, quantitates data, analyses data, or stores data; removes the background based on negative controls; averages or adjusts intensities as a function of the number of the biological images; generates reports; and stores data. The microarray can be processed to display the samples using an assay such as chromogenic assay or fluorescent assay, using a radioactive label imaged on radiographic film or using any other means in the art. The samples can be oligonucleotides such as DNA or mRNA or proteomics. The biological images can be in any form, preferably in the form of immunoassays, dot blots, Northern assays, Southern assay, Western assay, and electrophoretic gels.

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