

# Bioinformatics up to Date

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



## Inside.....

About us	1
Our Focus	1
Genomics	2
Proteomics	2
Bioserver	3
Computers for Biologists	3
Bioinfo.	
Animation	4
Bioinfo. Patent	4
Molecule of the month	5
Upcoming Events	5
Contact Us	5

**Advisor:**  
Dr D Ramaiah

**Editors:**  
Mr Robin Das  
Dr R L Bezbaruah  
Dr H P Deka Baruah



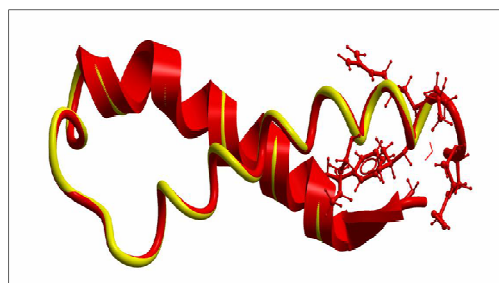
## About Us

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.

## Our Focus

**Comparative Modeling, Structure Comparison & Docking Studies for AVPR2 Protein with Some Anti-Diabetic Plant Compounds from North East India**

Arginine Vessopressin Protein Receptor 2 (AVPR2) plays an important role in water homeostasis normally but certain mutation of AVRP2 leads to Nephrogenic Diabetes Insipidus (NDI) in human. Herein, an attempt has

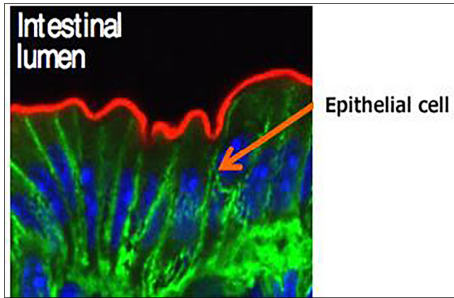


been made to predict an ideal homology 3D model of this protein and docking studies were done against some of anti-diabetic compounds from plants from North East India. The best structural protein model was found with highest sequence identity and least RMSD of 0.1414. The Molecular Docking and ADME results suggest two potential agents as mutant AVPR2 protein inhibitor. The two compounds namely Marmesiline and Caffeic acid which having greater binding affinities ( $-129.040 \text{ kJmol}^{-1}$  and  $-95.981 \text{ kJmol}^{-1}$ ) to the receptor protein.

[ published in 3D- Comparative Modelling, Structure Comparison & Docking Studies for AVPR2 Protein with Some Anti-Diabetic Plant Compounds from North East India; IGJPS, 2014; 4(1): 8-17 ]

## Novel Protein Against Bowel Inflammation

Professor Matozaki Takashi and Associate Professor Murata Yoji at the Kobe University Graduate School of Medicine Division of Molecular and Cellular Signaling, were first demonstrate the role of stomach cancer-associated protein tyrosine phosphatase (SAP)-1 in the pathogenesis and prevention of Crohn's disease, ulcerative colitis, and other inflammatory bowel disorders. Their findings, published online on July 20, 2015, by the Proceedings of the National Academy of Sciences of the United States of America. The findings are expected to accelerate the development of targeted therapies for inflammatory gastrointestinal diseases.



Previously, Prof. MATOZAKI, Assistant Prof. MURATA, and their colleagues found that SAP-1 localizes to the microvilli of the brush border in gastrointestinal epithelial cells. Here, they showed that SAP-1 ablation in a mouse model of inflammatory bowel disease resulted in a marked increase in the incidence and severity of bowel inflammation, suggesting that SAP-1 plays a protective role against colitis. In addition, carcinoembryonic antigen-related cell adhesion molecule (CEACAM) 20, an intestinal microvillus-specific membrane protein, was identified as the target of SAP-1 tyrosine dephosphorylation. Suppression of CEACAM20 functions via dephosphorylation was suggested to contribute to preventing colitis. By shedding light on the anti-inflammatory mechanism of the intestinal epithelial cells, Prof. MATOZAKI and colleagues believe that their findings will drive the development of drugs that target SAP-1 and CEACAM20 to overcome intractable inflammatory bowel diseases.

*[Protein tyrosine phosphatase SAP-1 protects against colitis through regulation of CEACAM20 in the intestinal epithelium. Murata, Y et al. PNAS (2015)]*

## A New Genetic Mechanism for Colorectal Cancer Progression

A Genetics researchers team from Case Western Reserve School of Medicine have identified a novel long non-coding RNA (lncRNA), dubbed DACOR1, that has the potential to stymie the growth of tumour cells in the second-most deadly form of cancer in the U.S. -- colorectal cancer. The research findings published in journal *Human Molecular Genetics*, August 2015.



The researchers found that this lncRNA is present in cells of healthy colons, but becomes suppressed in those carrying the disease. More importantly, this lncRNA interacts with a key enzyme known as DNMT1 that has important functions in all healthy cells of the body. Thus, the authors applied a name to this novel lncRNA -- DACOR1, which stands for DNMT1-Associated Colon Cancer Repressed lncRNA-1.

The scientists' next challenge is to determine how to deliver DACOR1 to tumours where it may be able to slow, or even stop, the spread of malignant cells. The researchers' initial findings appeared in this month's edition of *Human Molecular Genetics*.

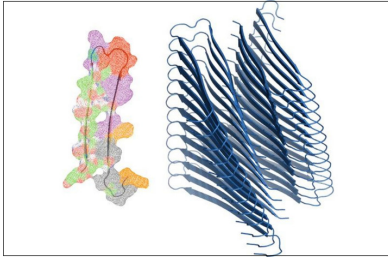
"We found that the metabolism of cancer cells slows when we put DACOR1 back in," said senior author Dr. Ahmad M. Khalil, an assistant professor of genetics and genome sciences. "If we could figure out a way to deliver DACOR1 to tumours, we could change the methylation patterns in cancer cells to either destroy or at least regress tumours."

During their research, investigators found that specific lncRNAs regulate DNA methylation in specific human genes. Researchers also sought clues on how lncRNAs affect normal tissue versus cancerous tissue and characterized one particular lncRNA, DACOR1, which is present in healthy colon tissue but missing from colon cancer tissue.

*[DNMT1-associated long non-coding RNAs regulate global gene expression and DNA methylation in colon cancer. Merry, CR et al. Human Molecular Genetics (August 25, 2015)]*

## New Software Program that Could Accelerate Drug Discovery for Diabetes and Other Diseases

Led by Prof. Jérôme Waldispühl the School of Computer Science McGill University researchers, have created a suite of computer programs that should speed up the process of drug discovery for diseases like Alzheimer kind. The programs are designed to scan the fibrils ( or misfolded proteins ) looking for weak spots. The idea is to then design helpful genetic muta-



tions to dissolve the bonds that hold the fibrils together - a bit like finding the right strand of wool to tug on to unravel a whole knotted ball.

But for the Fibrilizer, as McGill has dubbed its suite of computer tools, a name that hints at the super heroic nature of the programs they have developed, the task is of a very different order. "Within the space of a week, by using our programs and a supercomputer, we were able to look at billions of possible ways to weaken the bonds within these toxic protein strands. We narrowed it down to just 30 - 50 possibilities that can now be explored further," says Mohamed Smaoui, a McGill postdoctoral fellow and the first author on three recent papers on the research. "Typically biochemists can spend months or years in the lab trying to pinpoint these promising mutations."

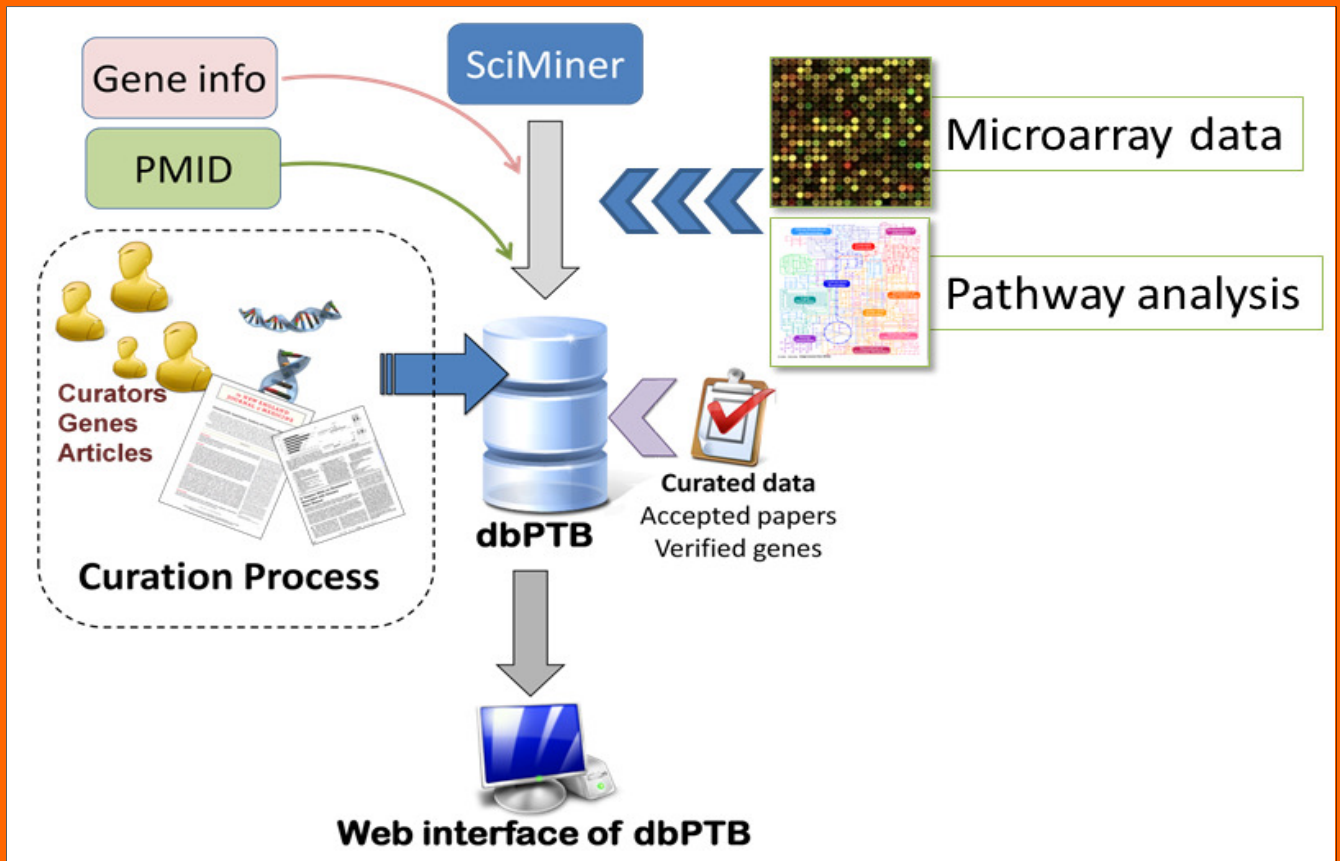
The researchers tested their program on a medical compound that scientists have been trying to improve for the last couple decades. The compound is administered as part of a drug that is used by diabetes patients to boost the performance of insulin and is sold under the name Symlin. The synthetic compound is based on a version of the protein amylin, yet is known to be toxic to the pancreas over the long-term, creating amyloid fibrils. The McGill team were able to use Fibrilizer to pinpoint a limited number of possible genetic modifications to the compound that would act to reduce its toxicity.

[<https://scicasts.com/bio-it/1844-bioinformatics/10102-new-software-program-that-could-accelerate-drug-discovery-for-diabetes-and-other-diseases>]

## CLIPZ: a database and analysis environment for experimentally determined binding sites of RNA-binding proteins.

*The stability, localization and translation rate of mRNAs are regulated by a multitude of RNA-binding proteins (RBPs) that find their targets directly or with the help of guide RNAs. Among the experimental methods for mapping RBP binding sites, cross-linking and immunoprecipitation (CLIP) coupled with deep sequencing provides transcriptome-wide coverage as well as high resolution. However, partly due to their vast volume, the data that were so far generated in CLIP experiments have not been put in a form that enables fast and interactive exploration of binding sites. To address this need, we have developed the CLIPZ database and analysis environment. Binding site data for RBPs such as Argonaute 1-4, Insulin-like growth factor II mRNA-binding protein 1-3, TNRC6 proteins A-C, Pumilio 2, Quaking and Polypyrimidine tract binding protein can be visualized at the level of the genome and of individual transcripts. Individual users can upload their own sequence data sets while being able to limit the access to these data to specific users, and analyses of the public and private data sets can be performed interactively. CLIPZ, available at <http://www.clipz.unibas.ch>, aims to provide an open access repository of information for post-transcriptional regulatory elements.*

[<http://www.ncbi.nlm.nih.gov/pubmed/21087992>]



### Patent News

## Bioinformatics analyzing method for redundancy of SSR (Simple Sequence Repeat) molecular marker

CN102156824B

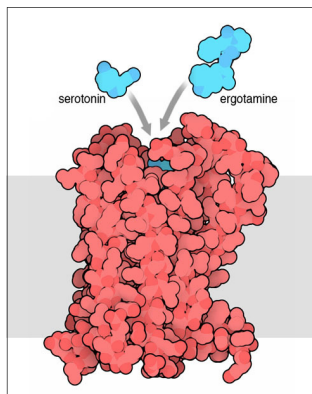
CN Grant

Abstract

The present invention discloses a SSR markers redundancy bioinformatics analysis method comprising the steps of: A1, download the relevant public databases or own SSR markers developed SSR markers; A2, on the SSR markers pretreatment, converted into FASTA format; A3, good FASTA file backup to handle a file suffix ".bk", use the backup file as input, for each species SSR markers respectively, than right, query similarity sequence; A4, A3 from the results obtained in a manner similar to the matching score of not less than 81%; and there is no gap; to filter a pair of primers, and then extract a similar primer number; A5, all similar primers write a line, output the final result File out. list. The researchers developed the same time cause the development of different species in the same SSR markers there is redundancy, the use of the method of the present invention can be achieved to redundancy purposes.

## Serotonin Receptor

Serotonin, a small molecule made from the amino acid tryptophan, was discovered for its role in the constriction of blood vessels. Most of the serotonin in your body is found in the digestive system where it helps to control the motions needed for



digestion. However, the most dramatic effects of serotonin are in the brain. Less than one in a million neurons uses serotonin for communication, but these neurons help to control our emotions, moods and thoughts. Serotonin is released from vesicles in nerve cells and picked up by receptors on the target cell surface. There are 15 different forms of these receptors in our bodies, most of which are G protein-coupled receptors like the one shown here from PDB entry 4iar. Serotonin binds to the portion of the receptor on the outside of the cell (shown here at the top of the picture). This induces subtle changes in the shape of the protein and sends a signal to G proteins inside the cell. In some cases, this leads to an excitatory response in the cells, and in other cases it is inhibitory, all depending on the particular receptor and its individual cellular context.

Many different drugs bind to serotonin receptors and modify the normal signaling process. The structure shown here has the drug ergotamine bound in the site normally occupied by serotonin. Ergotamine activates the receptor and has been used to control the painful swelling of blood vessels that cause migraines.

Once serotonin has finished its signaling task, it is transported back into the nerve cell by a serotonin transporter protein. This transporter is found in the cell membrane, and transports a sodium ion and a chloride ion along with each molecule of serotonin.

[<http://www.rcsb.org/pdb/101/motm.do?momID=164>]

### Upcoming Events

**CONFERENCE**  
ON  
**Bioinformatics and Computer Aided Drug Design**  
7<sup>th</sup> December, 2015  
Organize by  
DBT-Bioinformatics Infrastructure facility  
Department of Biosciences and Bioengineering  
Indian Institute of Technology Guwahati-781039 Assam India

 **IEEE BIBM 2015**  
IEEE International Conference on Bioinformatics & Biomedicine  
Washington D.C., Nov 9 – 12, 2015

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*