

# Bioinformatics up to Date

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



## Inside.....

Cover story	1
Bioinfo. Carrier	1
Genomics	2
Proteomics	2
Bioserver	3
Computers for Biologists	3
Bioinfo.	4
Animation	4
Bioinfo. Patent	4
Molecule of the month	5
Upcoming Events	5
Contact Us	5

**Advisor:**  
 Dr. D. Ramaiah

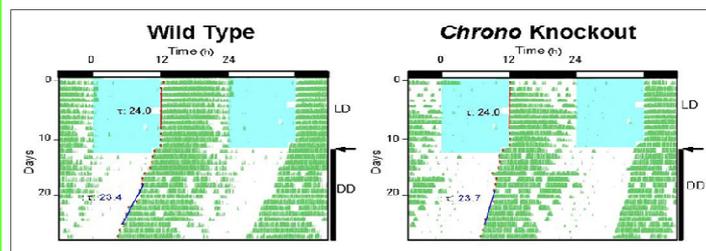
**Editors:**  
 Mr Robin Das  
 Dr R.L. Bezbaruah

**Protein Data Bank**

As of Tuesday Apr 22, 2014 at 5 PM PDT there are 99624 Structures

## Bioinformatics Profiling Identifies a New Mammalian Clock Gene

Over the last few decades researchers have characterized a set of clock genes that drive daily rhythms of physiology and behaviour in all types of species, from flies to humans. Over 15 mammalian clock proteins have been identified, but researchers surmise there are more. To accelerate clock-gene discovery, the investigators, led by John Hogenesch, PhD, professor of Pharmacology and first author Dr. Ron Anafi, an instructor in the department of Medicine, used a computer-assisted approach to identify and rank candidate clock components. This approach found a new core clock gene, which the team named CHRONO.



Their findings appear in PLOS Biology.

Hogenesch likens their approach to online profiling of movie suggestions for customers: “Think of Netflix. Based on your personalized movie profile, it predicts what movies you may want to watch in the future based on what you watched in the past.” He thought the team could use this approach to identify new clock genes, given criteria already established from the “behaviour” of known clock genes identified in the past two decades:

Clock genes cause oscillations at the messenger RNA and protein level.

Clock proteins physically interact with other clock proteins to form complexes that control daily rhythm inside cells.

Disruption of clock genes in cell models cause changes in observable behavioural and metabolic traits on a 24-hour cycle.

Clock genes are conserved across 600 million years of evolution from fruitflies to humans.

They found that several of the genes they identified physically interact with known clock proteins and modulate the daily rhythm of cells. One candidate, dubbed Gene Model 129, interacted with BMAL1, a well-known core clock component, and repressed the key driver of molecular rhythms, the BMAL1/CLOCK protein complex that guides the daily transcription of other proteins in a complicated system of genes that switch on and off over the course of the 24-hour day.

### Bioinfo. Carrier

1. IASST, Guwahati invites application for the following posts in BIF facility, Division of Life Science, IASST: Research Associate, Traineeship and Studentship . [<http://1.bp.blogspot.com/-JdckOGXO2Fc/U1tlikxLAgI/AAAAAAAAARek/5cgLwI7FDRk/s1600/IASST+Guwahati+Jobs+2014.jpg>]
2. CSIR-institute of Genomics and Integrative Biology (CSIR-IGIB) invites applications for admission to PhD. Program-2014. [Http://www.igib.res.in/sites/default/files/fulladvtd.phd\\_2014.pdf](http://www.igib.res.in/sites/default/files/fulladvtd.phd_2014.pdf)

## **GAIM-General Amyloid Interaction Motif**

Researchers from NeuroPhage Pharmaceuticals have engineered a series of molecules with the potential to treat most neurodegenerative diseases that are characterized by misfolded proteins, such as Alzheimer's, Parkinson's and Huntington's diseases. These molecules are based on what the Company calls a general amyloid interaction motif, or GAIM, which recognizes a characteristic common to many toxic, misfolded proteins, not just one type of misfolded protein. This approach provides NeuroPhage with an array of therapeutic targets, so that a number of pathologies, such as amyloid beta plaques, tau tangles and alpha-synuclein Lewy bodies, can all be addressed simultaneously with a single drug candidate.

"The research published on April 22 describes GAIM, NeuroPhage's unique approach to treat diseases characterized by misfolded proteins. GAIM has the potential to provide a more robust response than previous therapies because it enables the simultaneous targeting of multiple pathologies within a single disease," said Dr. Richard Fisher, Chief Scientific Officer at NeuroPhage. The method was published online in the Journal of Molecular Biology.

The discovery of GAIM has led to the creation of NeuroPhage's lead candidate, NPT088, which is the GAIM motif fused to a portion of a human antibody. The result is a potential therapeutic that can be easily delivered to patients. NeuroPhage has accumulated extensive preclinical data on this candidate, demonstrating its efficacy across disease models of Alzheimer's, Parkinson's and related diseases characterized by aggregation of the tau protein.

## **Proteins Discovered in Gonorrhoea May Offer New Approach to Treatment**

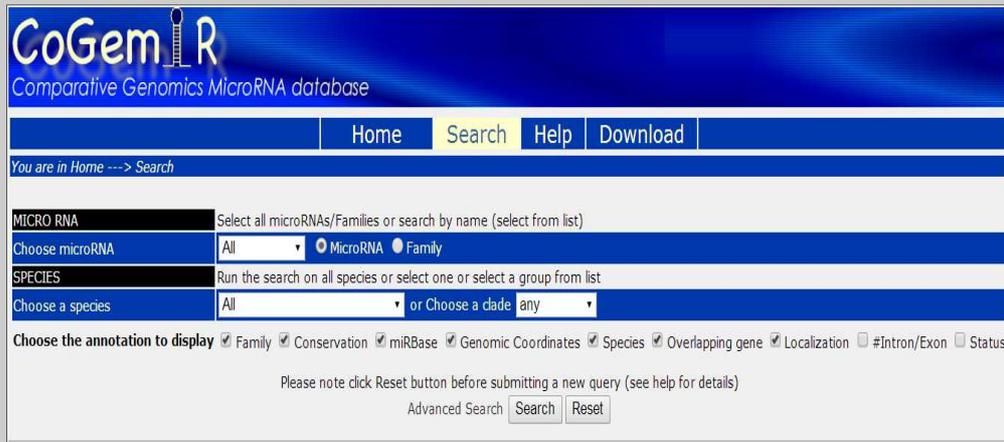
Researchers at Oregon State University have discovered novel proteins in, or on the surface of the bacteria that causes gonorrhoea, which offer a promising new avenue of attack against a venereal disease that is showing increased resistance to the antibiotics used to treat it. Only a single, third-generation cephalosporin antibiotic still shows good efficacy against gonorrhoea, creating a race against time to find some alternative way to treat this disease that can have serious health effects. It's the second most commonly reported infectious disease in the United States.

Investigations based on these proteins might lead to new ways to combat the disease, including a vaccine, new types of drugs to block the growth of the bacteria, or even restoring the efficacy of some older antibiotics that have lost their usefulness, said Aleksandra Sikora, an assistant professor in the OSU College of Pharmacy.

"This could be a milestone in finding new ways to treat a global problem," Sikora said. "It appears that one or more of these proteins, either within the bacterial cell envelope or on its surface, are essential to its growth and survival. Now we have a new target to aim at." The new findings were just published in *Molecular and Cellular Proteomics*, by researchers from OSU and the University of Washington. The research has been supported by OSU and the Medical Research Foundation of Oregon. Other proteins on the bacteria surface also help it attach to the host. The membrane vesicles are spherical structures that contain proteins and DNA, and are involved in antibiotic resistance, microbe communication and delivery of factors important for infection.

## CoGemiR

CoGemiR is a publicly available microRNA-centered database whose aim is to offer an overview of the genomic organization of microRNAs and of its extent of conservation during evolution in different metazoan species. The database collects information on genomic



The screenshot shows the CoGemiR web interface. At the top, there is a navigation bar with links for Home, Search, Help, and Download. Below this, there is a search form with several sections: 'MICRO RNA' with a dropdown menu for selecting all microRNAs/Families or searching by name; 'Choose microRNA' with radio buttons for 'All', 'MicroRNA', and 'Family'; 'SPECIES' with a dropdown menu for selecting all species or a group from a list; 'Choose a species' with a dropdown menu for selecting a species or a date; and 'Choose the annotation to display' with checkboxes for Family, Conservation, miRBase, Genomic Coordinates, Species, Overlapping gene, Localization, #Intron/Exon, and Status. There are also 'Advanced Search', 'Search', and 'Reset' buttons.

location, conservation and expression data of both known and newly predicted microRNAs and displays the data by privileging a comparative point of view. The database also includes a microRNA prediction pipeline to annotate microRNAs in recently sequenced genomes. This information is easily accessible via web through a user-friendly query page.

The information contained in CoGemiR are both collected from pre-existing databases and newly generated via a processing pipeline we implemented in order to determine

microRNA genomic organization and conservation across a number of metazoan species.

The database web interface allows verifying whether a given microRNA is reported to be intragenic or intergenic in one or more species. For intragenic microRNAs, it is possible to verify whether or not a particular microRNA is located within the same gene in different species as well as to establish whether its position within the host gene (e.g., exonic, intronic, etc.) is also conserved.

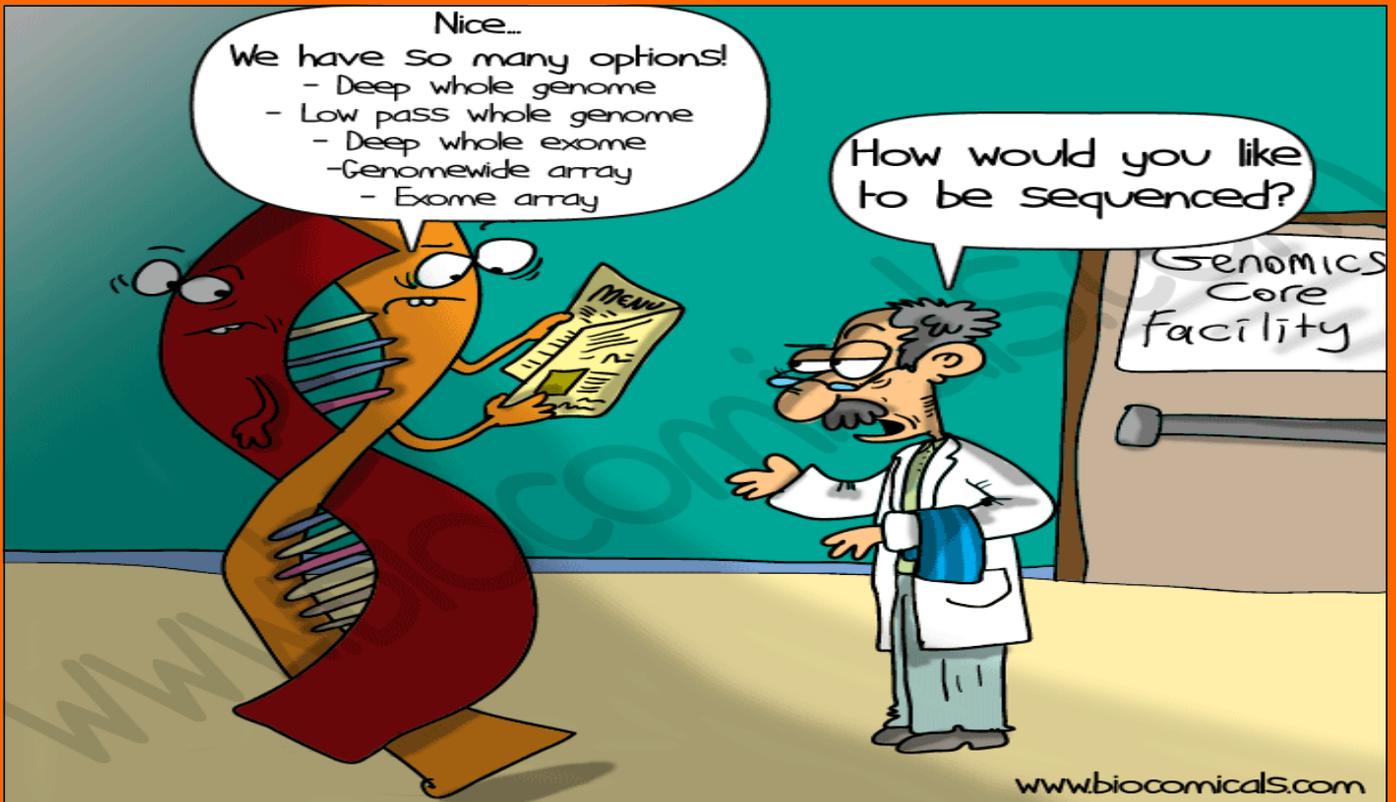
## Phevor (Phenotype Driven Variant Ontological Re-ranking tool)

A computational tool Phevor (Phenotype Driven Variant Ontological Re-ranking tool), developed at the University of Utah (U of U) has successfully identifies undiagnosed illnesses and unknown gene mutations by analyzing the exomes, or areas of DNA where proteins that code for genes are made, in individual patients and small families.

Mark Yandell, Ph.D, professor of human genetics, led the research. Phevor represents a major advance in personalized health care, according to Lynn B. Jorde, Ph.D., U of U professor and chair of human genetics and also a co-author on the study. As the cost of genome sequencing continues to drop, Jorde expects it to become part of standardized health care within a few years, making diagnostic tools such as Phevor more readily available to clinicians.

Phevor works by using algorithms that combine the probabilities of gene mutations being involved in a disease with databases of phenotypes, or the physical manifestation of a disease, and information on gene functions. By combining those factors, Phevor identifies an undiagnosed disease or the most likely candidate gene mutation for causing a disease. It is particularly useful when clinicians want to identify an illness or gene mutation involving a single patient or the patient and two or three other family members, which is the most common clinical situation for undiagnosed diseases.

Yandell, the lead developer of the software, describes Phevor as the application of mathematics to biology. "Phevor is a way to try to get the most out of a child's genome to identify diseases or find disease-causing gene mutations," Yandell said.



Patent News

**Data mining platform for bioinformatics and other knowledge discovery**

**WO 2002103954 A2**

**Publication date:** Dec 27, 2002

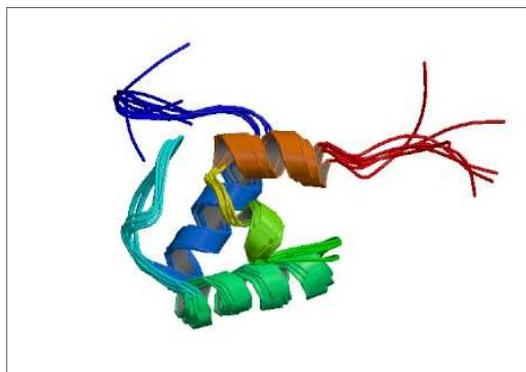
**Inventors:** Rene Doursat, Isabelle Guyon, David Lewis, Edward Reiss, Jason Weston

**ABSTRACT**

The data mining platform comprises a plurality of system modules (500, 550), each formed from a plurality of components. Each module has an input data component (502, 552), a data analysis engine (504, 554) for processing the input data, an output data component (506, 556) for outputting the results of the data analysis, and a web server (510) to access and monitor the other modules within the unit and to provide communication to other units. Each module processes a different type of data, for example, a first module processes microarray (gene expression) data while a second module processes biomedical literature on the Internet for information supporting relationships between genes and diseases and gene functionality

## Solution Structure of an Active Site Mutant Peptidyl Carrier Protein

**PubMed Abstract:** Phosphopantetheine transferases represent a class of enzymes found throughout all forms of life. From a



structural point of view, they are subdivided into three groups, with transferases from group II being the most widespread. They are required for the posttranslational modification of carrier proteins involved in diverse metabolic pathways. Here the crystal structure of the group II phosphopantetheine transferase Sfp from *Bacillus* in complex with a substrate carrier protein in the presence of coenzyme A and magnesium, and observed two protein-protein interaction sites. Mutational analysis showed that only the hydrophobic contacts between the carrier protein's second helix and the C-terminal domain of Sfp are essential for their productive interaction. Comparison with a similar structure of a complex of human proteins suggests that the mode of interaction is highly conserved in all domains of life.

Molecule: Tyrocidine synthase 3  
Polymer: 1  
Length: 90  
Chains: A  
Structure Weight: 9972.53  
Method: SOLUTION NMR  
Release: 2014-04-23  
UniProtKB: 030409

### Upcoming Events



188<sup>th</sup> OMICS Group Conference  
4<sup>th</sup> International Conference on  
**Proteomics & Bioinformatics**  
August 04-06, 2014 Hilton Chicago/ Northbrook, Chicago, USA  
Français | English

**Website:** <http://www.proteomicsconference.com>

## 2014 International Conference on Biomedical and Bioinformatics Engineering (ICBBE 2014)

26<sup>th</sup> to 27<sup>th</sup> August 2014

Taipei, Syria

**Website:** <http://www.icbbe.com/>

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](mailto:robindas460@gmail.com)  
Ph No-07399923578