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 R.L. Bezbaruah

National Workshop on Computational Biology in Plant Science

November 13 - 14, 2014

National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi-110067

Web-Address: http://www.nipgr.res.in

Biosynthesis:

Bioinformatics bolster a renaissance

Biosynthetic gene clusters encode the enzymatic pathways to make secondary metabolites, molecules of great interest for the pharmaceutical and biotechnology industries. Access to an increasing number of microbial genomes, coupled with efficient bioinformatic tools, is creating new momentum in secondary metabolite research.

Microbial secondary metabolites, natural products, harbor great chemical diversity that results in a broad range of biological activities and are used in agrochemical, food and pharmaceutical industries. Their biosynthetic pathways are encoded by biosynthetic gene clusters (BGCs). BGCs have been characterized for a limited set of bacterial secondary metabolites, but their diversity far exceeds research efforts thus far. The access to an increasing number of microbial genome sequences, recent progress in high-throughput sequencing technologies, has led to a renaissance in natural product research, providing opportunities to mine this biosynthetic potential in a predictive way, in contrast to previous labor- and time-consuming chemical analysis.
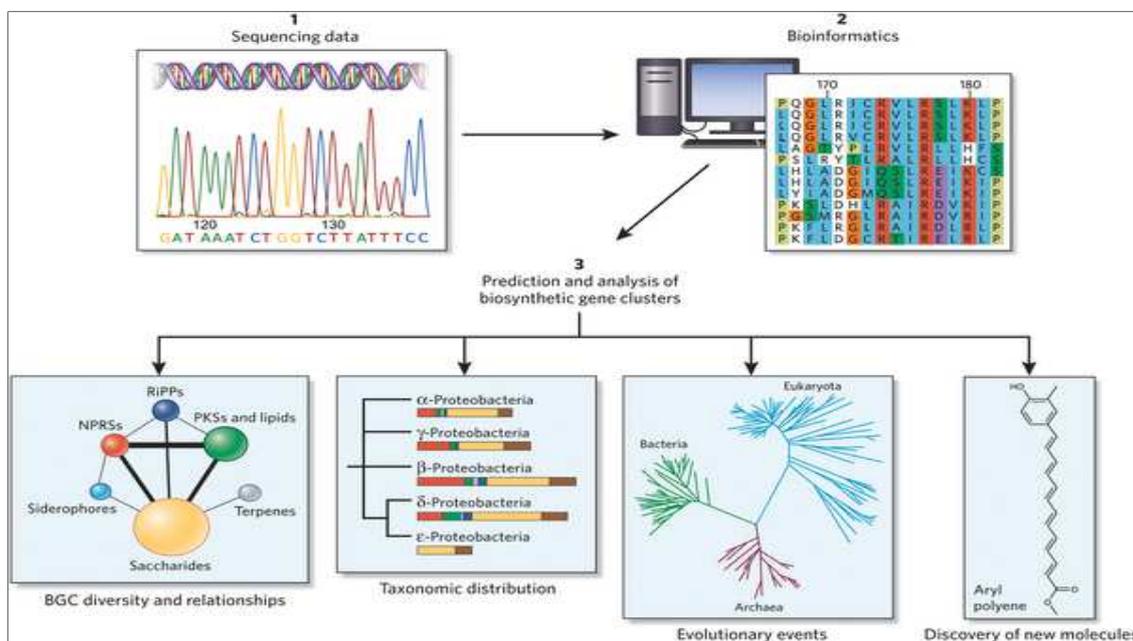
Efficient bioinformatic algorithms are used to detect and analyze biosynthetic gene clusters in microbial genomes. This three-step procedure creates new momentum in natural product research and provides new opportunities to improve our understanding of BGC diversity and taxonomic distribution to resolve evolutionary events and to discover new bioactive molecules. PKS, polyketide synthase; NRPS, nonribosomal peptide synthetase; RiPP, ribosomally synthesized and post-translationally modified peptide.

How gibbons got their swing

Apes' genes explain their flexibility — and why their chromosomes are so peculiar

Asia, a white-cheeked gibbon living in a Virginia zoo, the first of the long-armed, tree-swinging apes to have its genome decoded. Her DNA sequence, and those of seven other gibbons — a total of six different species — help to explain how these apes adapted to life in the trees. They also may explain why gibbons are so diverse compared with the great apes — humans, chimpanzees, gorillas and orangutans. Gibbons are renowned for their dexterity in dense tree canopies, traversing as far as 15 meters a swing and at speeds of more than 55 kilometers an hour .

Geneticists are fascinated by gibbons because they were the first of the apes to branch off from the common ancestor they share with both humans and monkeys. And gibbons have peculiar chromosomes. Compared to other apes, their genomes have many more chromosomal rearrangements, such as duplications, deletions or inversions of large stretches of DNA, that can affect how genes work.

"Chromosomal rearrangements are like earthquakes: one event will completely change the landscape, and you will see it after one generation," says Lucia Carbone, an evolutionary geneticist at Oregon Health and Science University in Portland, who led the gibbon genome effort. "What we found in the gibbon genome was that there were a lot of earthquakes." The authors attribute the explosion in gibbon diversity to a type of 'jumping gene' — a stretch of DNA that can change its position in the genome — that is found in gibbons alone among the primates. This chunk of DNA landed near genes involved in chromosome replication, making the genome more prone to rearrangement, they suggest.

Protein 'map' could lead to potent new cancer drugs

Imperial chemists have gained fresh insights into how a disease-causing enzyme makes changes to proteins and how it can be stopped. The scientists hope their findings will help them to design drugs that could target the enzyme, known as N-myristoyltransferase (NMT), and potentially lead to new treatments for cancer and inflammatory conditions.
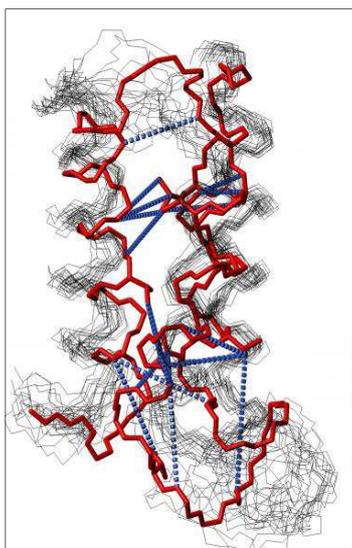
They have already identified a molecule that blocks NMT's activity, and have identified specific protein substrates where this molecule has a potent impact. NMT makes irreversible changes to proteins and is known to be involved in a range of diseases including cancer, epilepsy and Alzheimer's disease. In a study published in the journal *Nature Communications* chemists used living human cancer cells to identify more than 100 proteins that NMT modifies, with almost all these proteins being identified for the very first time in their natural environment. The scientists mapped all of the proteins and also established that a small drug-like molecule can block the activity of NMT and inhibit its ability to modify each of these proteins, suggesting a potential new way to treat cancer.

The researchers spent several years developing a specialised set of tools to identify and examine NMT and the proteins it changes. They used mass spectrometry to quantify the effect of a NMT inhibitor molecule. To examine this interaction, they induced a process called apoptosis, which programmes a cell to die - for example because its DNA has been damaged. This process is essential in cancer chemotherapy, and is very often deactivated in drug resistant cancers.

[http://www.eurekalert.org/pub_releases/2014-09/icl-p092514.php]

MetaProx: the database of metagenomic proximons

MetaProx is the database of metagenomic proximons: a searchable repository of proximon objects conceived with two specific goals. The first objective is to accelerate research involving metagenomic functional interactions by providing a database of metagenomic operon candidates. Proximons represent a special subset of directons (series of contiguous co-directional genes) where each member gene is in close proximity to its neighbours with respect to intergenic distance. As a result, proximons represent significant operon candidates where some subset of proximons is the set of true metagenomic operons. Proximons are well suited for the inference of metagenomic functional networks because predicted functional linkages do not rely on homology-dependent information that is frequently unavailable in metagenomic scenarios. The second objective is to explore representations for semistructured biological data that can offer an alternative to the traditional relational database approach. In particular, we use a serialized object implementation and advocate a *Data as Data* policy where the same serialized objects can be used at all levels (database, search tool and saved user file) without conversion or the use of human-readable markups. MetaProx currently includes 4 210 818 proximons consisting of 8 926 993 total member genes. Database URL: http://metaprox.uwaterloo.ca.

Bi-Force: large-scale bicluster editing

The explosion of the biological data has dramatically reformed today's biological research. The need to integrate and analyze high-dimensional biological data on a large scale is driving the development of novel bioinformatics approaches. Biclustering, also known as 'simultaneous clustering' or 'co-clustering', has been successfully utilized to discover local patterns in gene expression data and similar biomedical data types. Here contribute a new heuristic: 'Bi-Force'. It is based on the weighted bicluster editing model, to perform biclustering on arbitrary sets of biological entities, given any kind of pairwise similarities. The BiForce Toolbox has been developed as a collaboration between MRC Human Genetics Unit and the Finnish Microarray and Sequencing Centre. BiForce Toolbox addresses high throughput analysis of pair-wise epistasis in GWAS of quantitative and disease traits, and provides a solution for all commonly used computer systems. BiForce Toolbox is a stand-alone Java program that integrates bitwise computing with multithreaded parallelisation and thus allows rapid full pair-wise genome scans via a graphical user interface or the command line. Furthermore, it incorporates additional tests of interactions involving SNPs with significant marginal effects, potentially increasing the power of detection of epistasis. BiForce Toolbox is easy to use and has been applied in multiple studies of epistasis in large GWAS datasets, identifying interesting interaction signals and pathways.
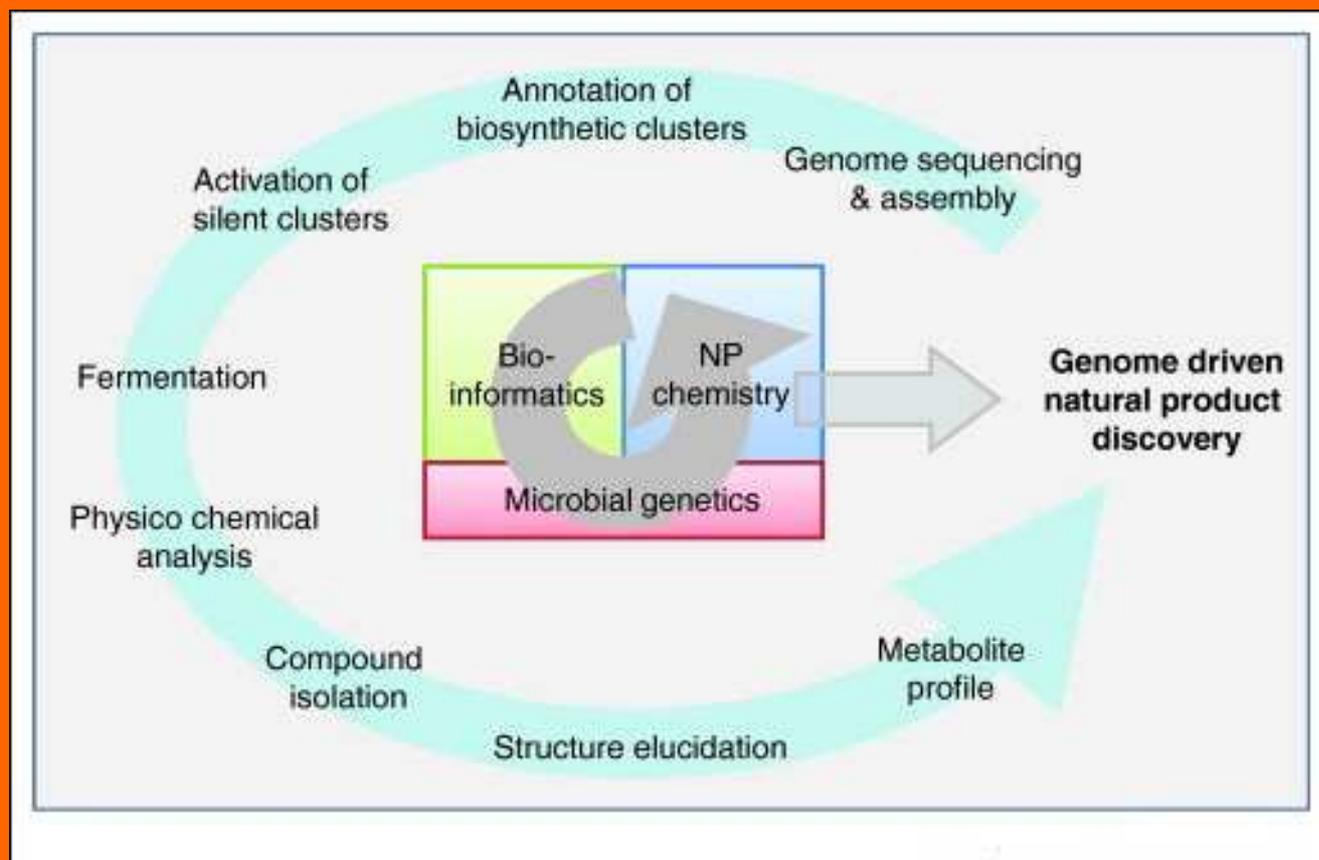
Patent News

Bs and rrbs sequencing-based bioinformatics analysis method and device

WO 2013097061 A1

Inventors: Honglong WU *et al.*

Publication date: Jul 4, 2013

Abstract

Provided in the present invention are a method and a device for analyzing methylation in a genome by bisulfite sequencing or reduced representation bisulfite sequencing. Also provided in the present invention is a bioinformaticsanalysis method based on bisulfite sequencing or reduced representation bisulfite sequencing. The method comprises: detecting DNA methylation, then performing the bioinformatics analysis, wherein the analysis is an analysis for one or more items selected from the following : data output, sequencing fragments alignment, the coverage situation of methylation sites, the methylation level of the methylation sites, the distribution trend of the whole genome methylation data, the whole genome DNA methylation map, the differentially methylated regions, the statistic for the length of insert fragments, and the coverage of CpG sites.

Apoptosomes

The process of programmed cell death, also known as apoptosis, is highly regulated, and the decision to die is made through the coordinated action of many molecules. The apoptosome plays the role of gatekeeper in one of the major processes, termed the intrinsic pathway. It lies between the molecules that sense a problem and the molecules that disassemble the cell once the choice is made. Normally, the many subunits of the apoptosome are separated and inactive, circulating harmlessly through the cell. When trouble occurs, they assemble into a star-shaped complex, which activates protein-cutting caspases that get apoptosis started.
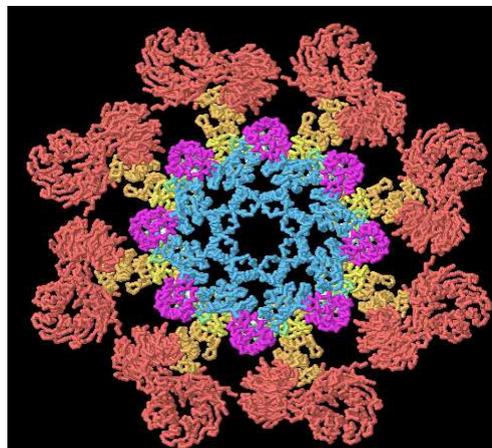
Apoptosome proteins are composed of several functional domains, each with a different role to play. The apoptosome from fruit flies is shown here, with the domains in different colors. At one end of the chain, there is a CARD domain （ magenta ） involved in caspase interaction and a nucleotide-binding domain （ blue ） that assists with assembly of the complex. At the other end of the chain are two "beta-propeller" domains （ red ） that grip cytochrome c and trigger formation of the complex. In between are several smaller domains （ green and yellow ） that coordinate action between the two ends.

Upcoming Events

Conference on Informatics and Integrative Biology

CIIB - 2014

December 17 - 19, 2014 Bose Institute Kolkata

Weblink: http://bicresources.jcbose.ac.in/CIIB/ciib2014/

ICGCW 2014

2nd INDIAN CANCER GENETICS CONFERENCE & WORKSHOP

Advanced Centre for Treatment Research & Education in Cancer (ACTREC)
Tata Memorial Centre (TMC), Navi Mumbai, India

Workshop: 27th – 31st October 2014; Conference: 1st-2nd November, 2014

Weblink: http://www.actrec.gov.in/data%20files/2014/ICGCW-2014-Flyer.PDF

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