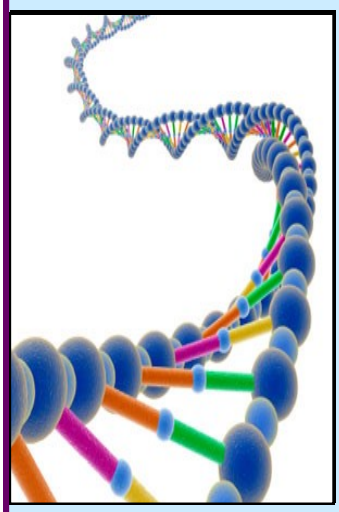
Volume 11, Issue 06 June 2018

Bioinformatics up to Date

(Bioinformatics Infrastructure Facility, Biotechnology Division)
North-East Institute of Science & Technology
Jorhat -785006, Assam
(http://www.rrljorhat.res.in/biotechnology.html)

Inside…..

About us	1
Cover story	1
Computers for Biologist	2
Bioserver/softwares/tools	2
Bioinfo. Animation	3
Upcoming Events	3
Molecule of the month	4
Contact Us	4



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.

Third-generation Sequencing Reveals Extensive Polycistronism and Transcriptional Overlapping in a Baculovirus

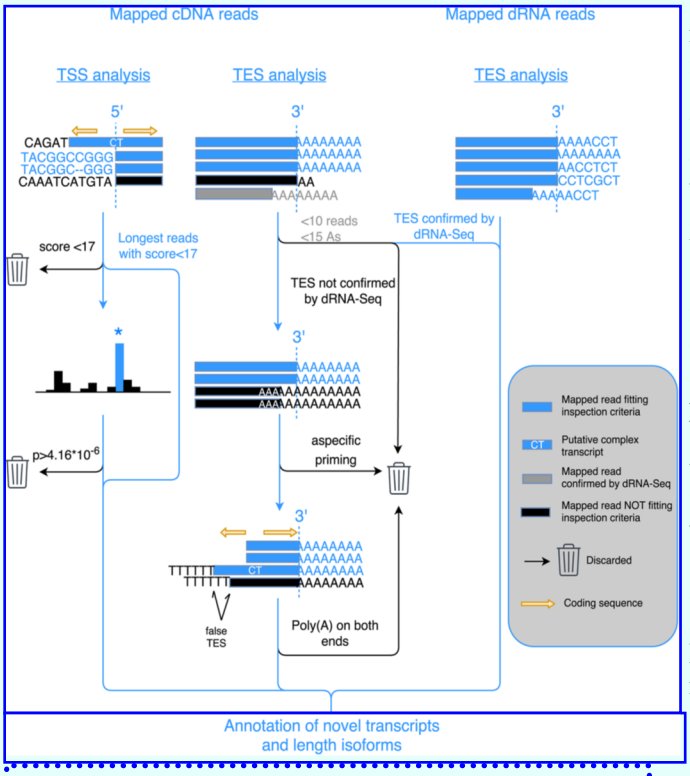


Fig:The schematic representation of the work-

The Autographa californica multiple nucleopolyhedrovirus (AcMNPV) is an insect-pathogen baculovirus. In this study, we applied the Oxford Nanopore Technologies platform for the analysis of the polyadenylated fraction of the viral transcriptome using both cDNA and direct RNA sequencing methods. We identified and annotated altogether 132 novel transcripts and transcript isoforms, including 4 coding and 4 non-coding RNA molecules, 47 length variants, 5 splice isoforms, as well as 23 polycistronic and 49 complex transcripts. All of the identified novel protein-coding genes were 5′-truncated forms of longer host genes. In this work, we demonstrated that in the case of transcript start site isoforms, the promoters and the initiator sequence of the longer and shorter variants belong to the same kinetic class. Long-read sequencing also revealed a complex meshwork of transcriptional overlaps, the function of which needs to be clarified. Additionally, we developed bioinformatics methods to improve the transcript annotation and to eliminate the non-specific transcription reads generated by template switching and false priming.

Source: Norbert Moldován *et al. 2018, J Scientific reports*

PEA: an integrated R toolkit for plant epitranscriptome analysis

PEA, an integrated R toolkit to facilitate the analysis of plant epitranscriptome data. The PEA toolkit contains a comprehensive collection of functions required for read mapping, CMR calling, motif scanning and discovery, and gene functional enrichment analysis. PEA also takes advantage of machine learning technologies for transcriptome-scale CMR prediction, with high prediction accuracy, using the Positive Samples Only Learning algorithm, which addresses the two-class classification problem by using only positive samples (CMRs), in the absence of negative samples (non-CMRs). Hence PEA is a versatile epitranscriptome analysis pipeline covering CMR calling, prediction, and annotation, and we describe its application to predict N6-methyladenosine (m6A) modifications in *Arabidopsis thaliana*. Experimental results demonstrate that the toolkit achieved 71.6% sensitivity and 73.7% specificity, which is superior to existing m6A predictors. PEA is potentially broadly applicable to the in-depth study of epitranscriptomics.

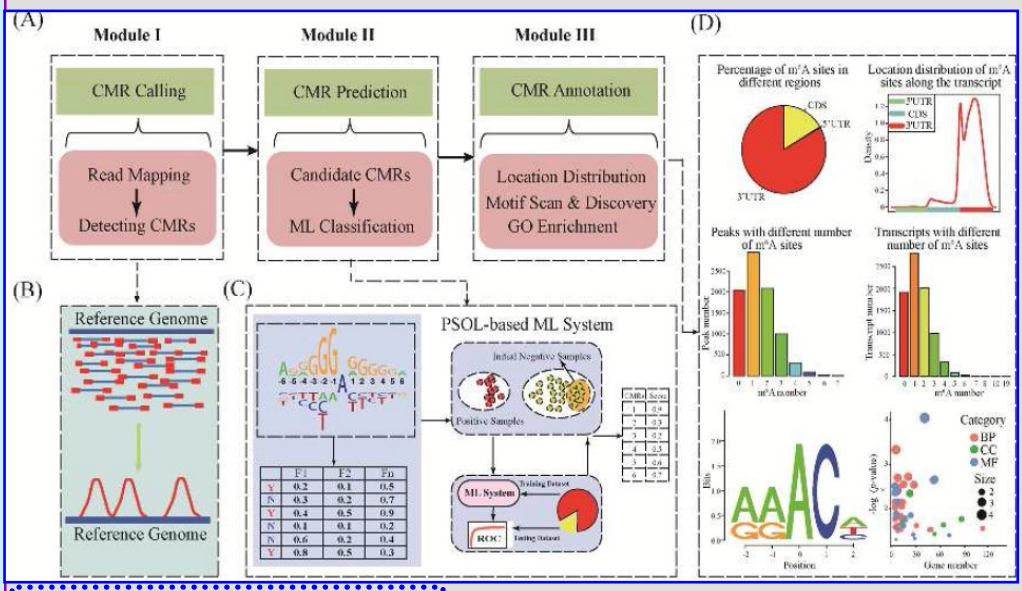


Fig:The schematic overview of PEA

Availability of this tool: PEA Docker image is available at https://hub.docker.com/r/malab/pea, source codes and user manual are available at https://github.com/cma2015/PEA.

Source: Jingjing Zhai *et al.* 2018 *„J Oxford Bioinformatics*

MIDORI server: a webserver for taxonomic assignment of unknown metazoan mitochondrialencoded sequences using a curated database

MIDORI server, a user-friendly platform to facilitate taxonomic classification of mitochondrial-encoded gene sequences with MIDORI. The server currently performs taxonomic assignments with three algorithms that predict taxonomy using k-mer similarity: SPINGO, RDP classifier and SINTAX. A maximum of 10,000 sequences in a FASTA format can be uploaded at once, and all of them must be shorter than 4,000 base pairs. Each algorithm can be run using two versions of each of the fifteen mitochondrial-encoded gene reference datasets: MIDORI-Unique and MIDORI-Longest. MIDORI-Unique contains all haplotypes of every species while MIDORI-Longest contains a single haplotype per species, the longest one. MIDORI-Longest for the COI gene contains the longest sequence for every species represented in the COI dataset. In thi study, using 1336 zooplankton sequences (Machida et al. 2009, 500 bp), they estimated the time required for assignments using the three algorithms with default settings (reference: COI-Longest). As a result, relatively longer calculation time was required for RDP classifier (630 sec.), compared to SPINGO (90 sec.) and SINTAX (100 sec.). Assigned phyla were compared between the results obtained from RDP classifier and SINTAX. The result indicated that about 10% of assignments were inconsistent between the results (most likely the groups with fewer referene sequences).

MIDORI available at this web site(http://www.referencemidori.info/download.php.

Source: Matthieu Leray *et al. 2018,J Oxford Bioinformatics*

Bioinfo. Animation (Source :Man S. Kim *et al. 2018,J Scientific Reports)*

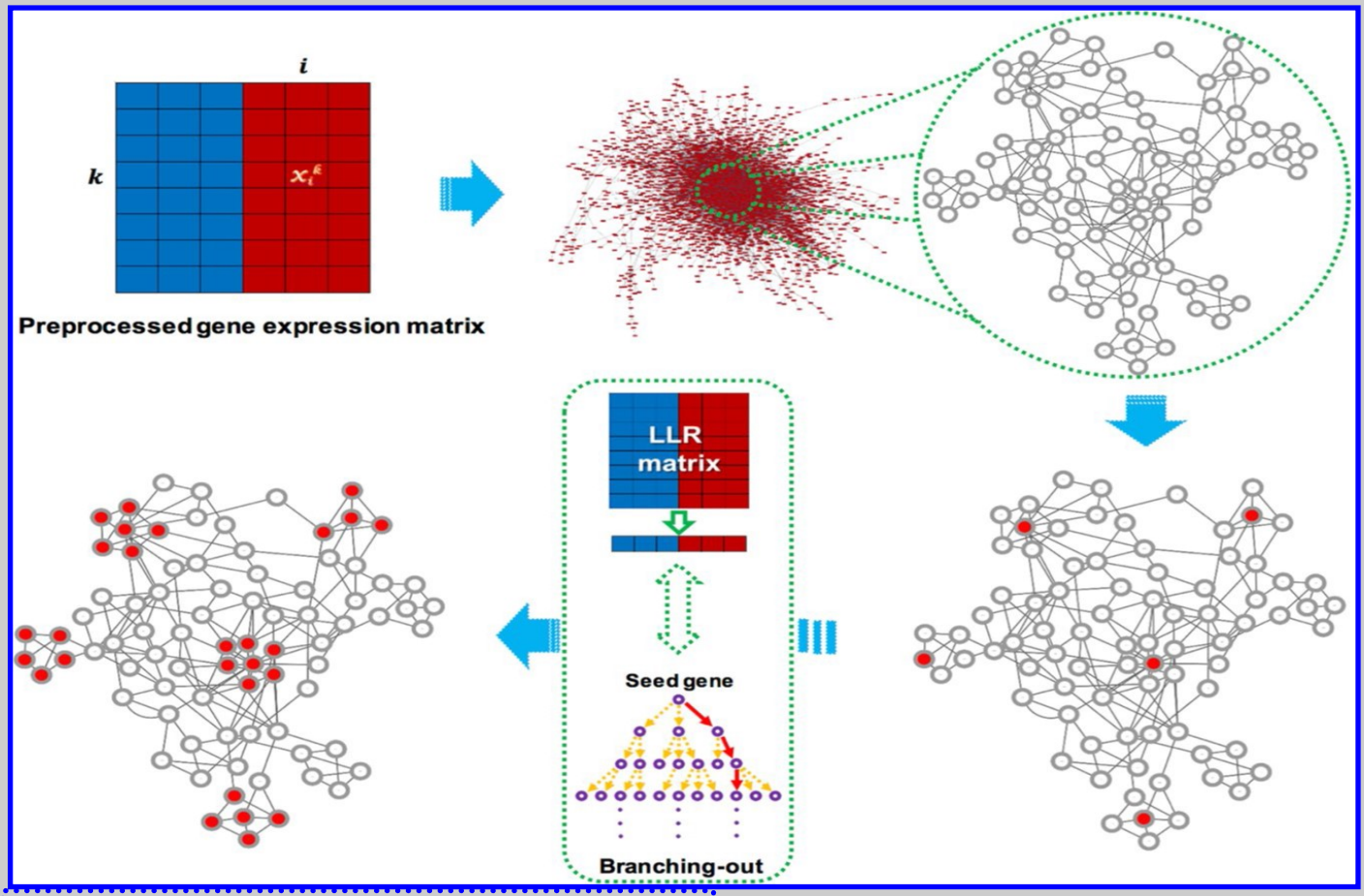


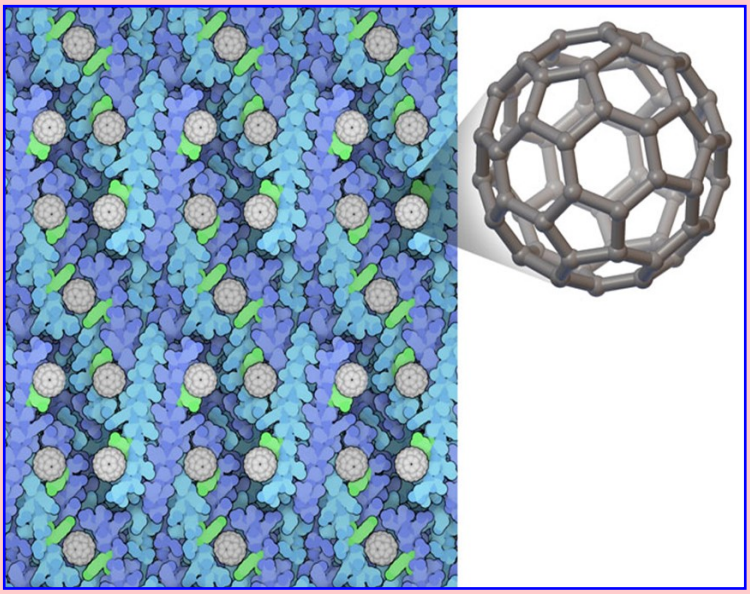
Fig:Schematic overview of our network-based NGS data analysis.

Upcoming event

1.http://hibit2018.org/
2.http://www.sgrfconferences.org/2018/NGBT/

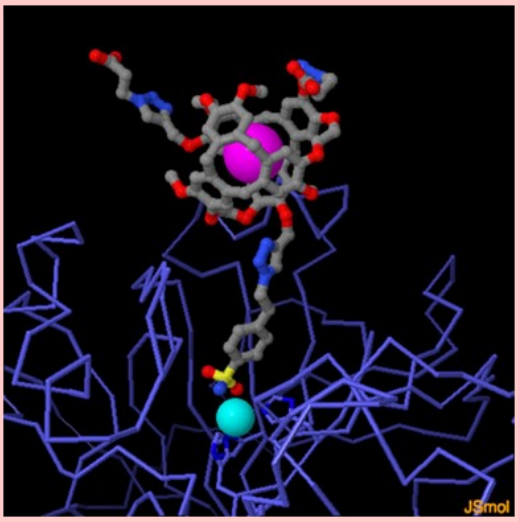
Proteins and Nanoparticles

Pure carbon is found in many forms: a perfect three-dimensional lattice of carbon atoms provides the strength of diamonds, and two-dimensional sheets of carbon atoms provide the smooth, slippery feeling of graphite. Researchers discovered a new form of carbon in 1985, with 60 atoms arranged to form a symmetrical hollow sphere. Because of their similarity to the geodesic domes designed by Buckminster Fuller, these new molecules were called "fullerenes". Like diamond and graphite, this new form of carbon has its own properties: fullerenes, and tube-shaped variants called nanotubes, can conduct electricity or heat, but at a much smaller scale than typical copper wires. Because of their perfect, symmetrical shape and surprising properties, fullerenes have become a favorite subject in nanotechnology. However, fullerenes have a major drawback for life sciences applications: they are insoluble in water. Scientists are now designing new interactions between proteins and fullerenes (and other nanoparticles) for novel applications in nanotechnology and medicine.



Biosensor for MRI

An isotope of xenon (129-Xe) is used as a contrast agent for magnetic resonance imaging (MRI) in medical diagnostic testing, both as a gas to image airspaces in the lung, and dissolved in body fluids to image the bloodstream and tissues. Designed cryptophanes are being developed as a way to target the xenon to specific proteins. Cryptophanes are hollow shells, but with openings on the sides that allow entry and exit of host molecules. The cryptophane shown here is the right size for a single xenon atom (shown in magenta), and is coupled to a specific inhibitor of carbonic anhydrase, which recognizes a zinc atom (turquoise) in the enzyme active site. Upon binding of the designed cryptophane to the target protein, xenon displays a distinctive MRI spectrum. To explore the structure of this cryptophane bound to human carbonic anhydrase II (PDB entry 3cyu), click on the image for an interactive JSmol.



Source:http://pdb101.rcsb.org/motm/222

Kindly send us your feedback to

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

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**