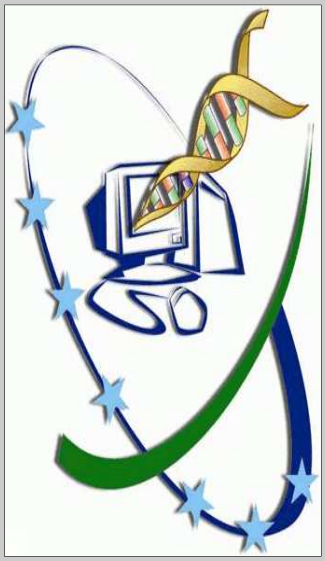
Volume 9, Issue 10 October 2016

Bioinformatics up to Date

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

Inside…..



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.

Methods to detect the effects of alternative splicing and transcription on proteins

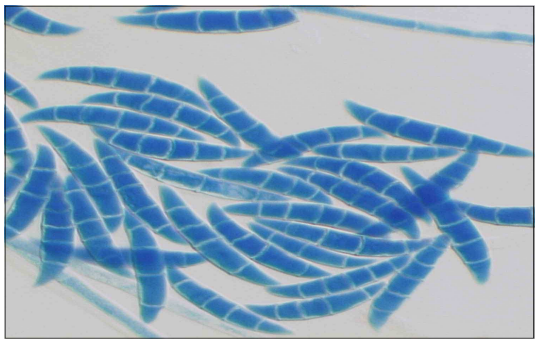
Alternative splicing and the transcription are the most familiar processes amongst the biological processes. This leads to the production of different kind of proteins from the same gene with different forms of mRNA which are known as "transcript variants", or "splice variants" or "isoforms"

The proteins produced after the alternative splicing are affected in different ways. As these transcript variants encode for different proteins having different amino acid sequence and hence produce different functions . Addressing the above problems, Mall et al., (2016) has developed a new software known as "ProtAnnot" as a plug-in in the IGB (Integrated Genome Browser). IGB is a user- friendly genome browser which helps the user to analyze the genomic data and the RNA-seq data. ProtAnnot provides a deep insight into how the transcription and alternative splicing affects the protein and its function.

ProtAnnot provides a fast and efficient way to visualize the impact of alternative transcribed proteins and display linked blocks which represent transcript structures and the thickness of the block represents the translated region.

Draft Genome Sequence of a UK Strain (UK99) of Fusarium culmorum

Fusarium culmorum is a soilborne fungal plant pathogen that causes foot and root rot and *Fusarium* head blight on small-grain cereals, in particular on wheat and barley. We report herein the draft genome sequence of a 1998 field strain called FcUK99 adapted to the temperate climate found in England. The ascomycete fungus *Fusarium culmorum* is one of two commonly found pathogenic species identified on flowering wheat spikes that cause yield losses and mycotoxin contamination of grain in the United Kingdom.



The draft assembly of UK99 is 41,928,875 bp in length with 1,062,015 gaps, although this includes chromosome 5 and chromosome 6 as pseudomolecules. Removal of all the unknown bases results in an assembly of 39,005,997 bp, which is 1,429,176 bp greater than the CS7071 assembly. Strain CS7071 lacks a gene annotation, unlike UK99. The UK99 genome annotation can be publicly accessed in a genome browser at http://pre.fungi.ensembl.org prior to its incorporation into the full Ensembl Fungi release. Ensembl Fungi also provides a community curation tool for the UK99 annotation available at http://cap.ensemblgenomes.org/Fculmorum.

Raw data and the assembled sequences have been submitted to the European Nucleotide Archive (ENA). The study accession number is PRJEB12835. Accession numbers for the assembled chromosomes are LT598659to LT598662, FJUU01000001, and FJUU01000002.

[Source: Genome Announc. 2016 Sep 15;4(5). pii: e00771-16. doi: 10.1128/genomeA.00771-16]

Roar: A software to detect alternative polyadenylation

Identifies preferential usage of alternative polyadenylation (APA) sites, comparing two biological conditions, starting from known alternative sites and alignments obtained from standard RNA-seq experiments. The ROAR approach is based on Fisher test to detect disequilibriums in the number of reads falling over the 3'UTRs when comparing two biological conditions. Counts and fragments lengths are used to calculate the prevalence of the short isoform over the long one in both conditions, therefore the ratio of these ratios represents the relative "shortening" (or lengthening) in one condition with respect to the other.

The algorithm is implemented in a Bioconductor package to facilitate its broad usage in the scientific community. The ability of this approach to detect shortening from libraries with a number of reads comparable to that needed for differential expression analyses makes it useful for investigating if alternative polyadenylation is relevant in a certain biological process without requiring specific experimental assays.

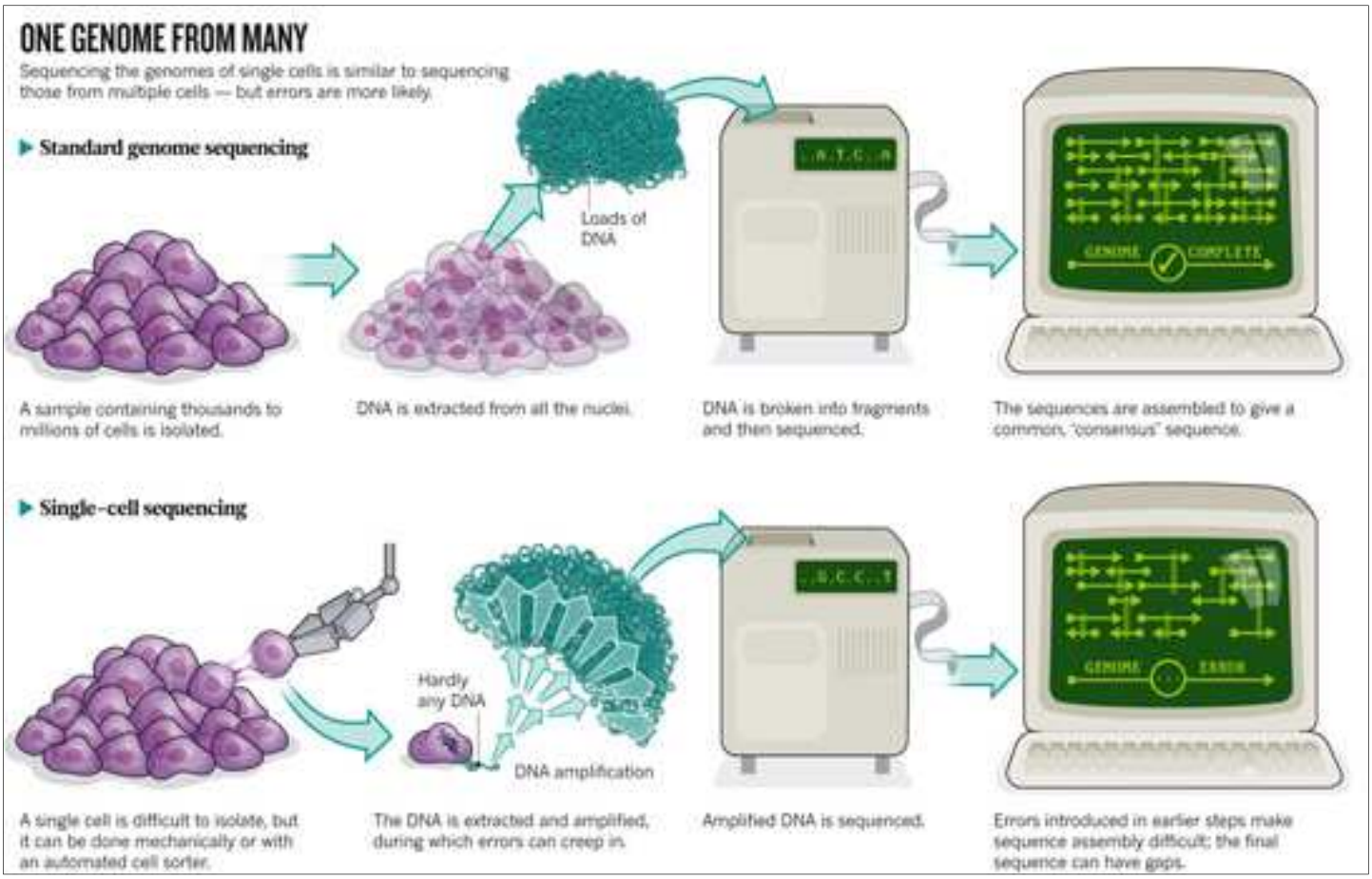
[SOURCE: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-016-1254-8]

IRNdb: A database of immunologically relevant non-coding RNAs

IRNdb is a database that combines *microRNA*, *PIWI-interacting RNAs*, and *long non-coding RNA*information with immunologically relevant target genes. The database is intended to advance research on the influence of ncRNAs on immunological processes. Currently, IRNdb contains information for *mouse* as it is often used as a model organism for immunological research purposes. The current version of IRNdb documents 12 930 experimentally supported miRNA-target interactions between 724 miRNAs and 2427 immune-related mouse targets. IRNdb is a comprehensive searchable data repository which will be of help in studying the role of ncRNAs in the immune system. IRNdb collects information on ncRNAs targeting immunologically relevant protein-coding genes and biological annotation data. To access the content in IRNdb, a web-interface is developed (accessible at http://irndb.org), which is divided into four main sections 'miRNAs', 'piRNAs', 'lncRNAs' and 'Target genes'.

[SOURCE: Database (2016) 2016 : baw138doi: 10.1093/database/baw138]

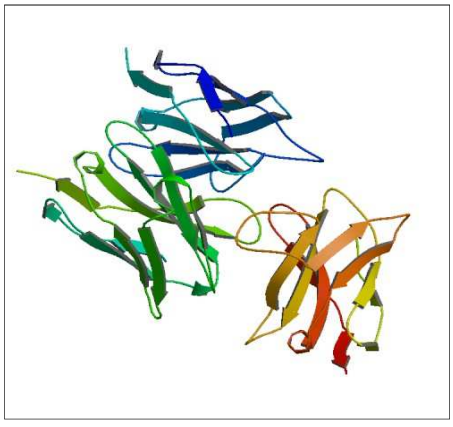
Bioinfo. Animation



Molecule of the month

Pembrolizumab

Pembrolizumab is an FDA-approved therapeutic antibody that targets the programmed cell death-1 (PD-1) to block the immune checkpoint pathway for the treatment of various types of cancer. It receives remarkable attention due to the high degree of efficacy. Very recently, the crystal structure of the Fab fragment of pembrolizumab (PemFab) in complex with the extracellular domain of human PD-1 (PD-1_{ECD}) was reported at a resolution of 2.9 Å. However, this relatively low-resolution structural data fails to provide sufficient information on interfacial water molecules at the binding interface that substantially contribute to affinity and specificity between the therapeutic antibody and target. Herein presenting the independently determined crystal structure of the Fv fragment of pembrolizumab (PemFv) in complex with the PD-1_{ECD} at a resolution of 2.15 Å. This high-resolution structure allows the accurate mapping of the interaction including water-mediated hydrogen bonds and provides, for the first time, a coherent explanation of PD-1 antagonism by pembrolizumab. The structural data provides new insights into the rational design of improved anti-PD-1 therapeutics. .



Experimental Data Snapshot

- Method: X-RAY DIFFRACTION
- Resolution: 2.15 Å
- R-Value Free: 0.226
- R-Value Work: 0.184
- Deposited: 2016-06-14
- Expression System: Brevibacillus choshinensis

Upcoming events

Patents

Bioinformatics computation using a map reduce-configured computing system

US20080133474A1

Inventor: Ruey-Lung Hsiao, Ali Dasdan, Hung-Chih Yang

Abstract

A MapReduce architecture may be utilized for sequence alignment algorithm processing (such as BLAST or BLAST-like algorithms). In addition, a MapReduce architecture may be extended such that memory of the computing devices of a MapReduce-configured system may be shared between different jobs of sequence alignment and/or other bioinformatics algorithm processing, thereby reducing overhead associated with executing such jobs using the MapReduce-configured system.

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