Volume 11, Issue 12 December, 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

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.

A milestone by UK Researchers of 100,000 Genomes Project was finally completed.

A novelty to better diagnose and treat an ever expanding number of diseases has finally been completed by the UK researchers and now they are considered as an acknowledged leader in population genomics, which is truly a global effort. Genomics England Researchers in partnership with NHS England have successfully completed sequencing 100,000 Human Genomes which is a noteworthy attainment. The first diagnosis that sequenced around 2,000 whole genomes was in February 2015 and the first diagnosis of children was made in January 2016. Jessica was one of the first children to receive a diagnosis from the 100,000 Genomes Project in which it was discovered that her condition was caused by errors in the SLC2A1 gene that cause 'Glut1 deficiency syndrome', which only affects around 500 people worldwide. The Project has proven the concept of genomic medicine at scale and built the infrastructure that gave the base to NHS Genomics Medicine Service. Genomics and the GMS are metamorphic which balanced the future in healthcare. Via this project of beyond 100,000 genome achievement that brought trust in generic drugs and treatments to medical history which brings hope to thousands of families to receive an effective diagnosis for the first time. It brings hope to large no. of cancer patients with a contingency for a clinical trial or a targeted therapy. Genomics England and NHS have expressed their enough gratitude to 85,000 participants, 1,500 NHS staff and over 3,000 researchers which were a part of the project. Also, the National Institute for Health Research and the UK Government played a very important role in the successful completion of this project. The entire infrastructure was set up from scratch facing numerous hurdles in the initial stages including the absence of technology to deliver genomic medicine, lack of effective bioinformatics pipeline to ensure processing at a large scale. Presently, UK has become the leading nation in the world, to use whole-genome sequencing at a large scale in direct healthcare. As an outcome of the project, NHS Genomic Medicine Service has been established which will provide access to genomic testing to patients across NHS from 2019. Also, it is not the ending here. In oct 2018, the U.K. Govt announced new plans of sequencing 1 million genomes over the next five years.

Source: Genomics England, Dept. Of Health and Social

Pathogen-Host Analysis Tool (PHAT): an Integrative Platform to Analyze Pathogen-Host Relationships in Next-Generation Sequencing Data

The study reveals that the Pathogen-Host Analysis Tool (PHAT) is an application for processing and analyzing next generation sequencing (NGS) data as it relates to relationships between pathogen and host organisms. Unlike custom scripts and tedious pipeline programming, it provides an unifying platform including raw and aligned sequence and reference file input, quality control (QC) reporting, alignment and variant calling, linear and circular alignment viewing, and graphical and tabular output. This unique tool aims to be user-friendly for life scientists studying diverse pathogen-host relationships.

Accessing next-generation sequencing (NGS) data has grown significantly but still some obstacles exist in its processing and analysis. The availability of fast and user-friendly tools has become the limiting factor. There are excellent tools which perform one or several discrete functions in the same domain, e.g., Bowtie2 and SAMtools. Centralizing multi-tool platforms such as Comparative Genomics (CoGe), VirBase ,Pathogen-Host Interaction Data Integration and Analysis System (PHIDIAS) and Unipro UGENE exist, all are server or cloud-based, or maintained by relatively small communities. To combat different related issues finally lead to the development of the Pathogen-Host Analysis Tool (PHAT) by presenting an easy-to-setup and easy-to-use platform for life scientists conducting pathogenhost NGS analysis on common desktop computing hardware (e.g., Windows).

Features:

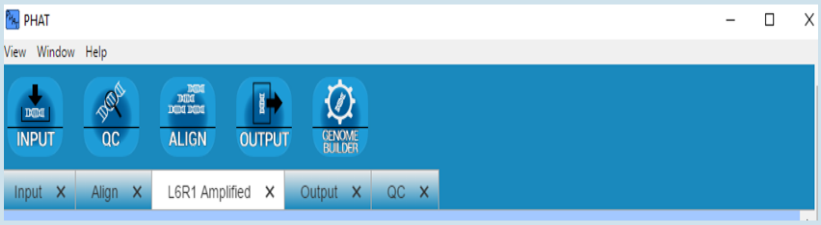
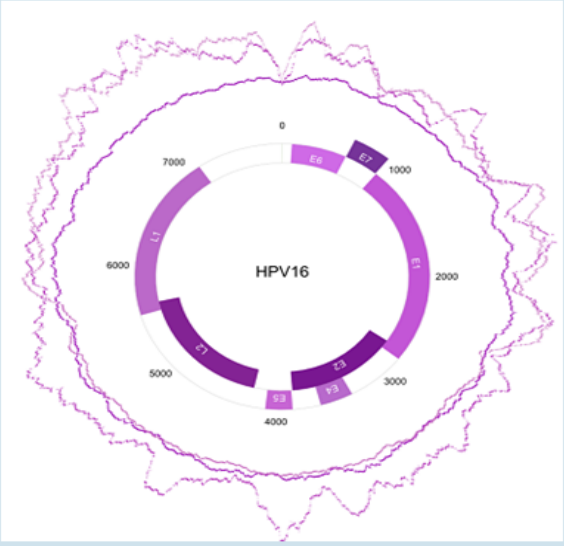


Figure1: Homepage of PHAT

High-throughput sequencing output files commence the Pathogen-host NGS analysis which contains experimentally suited nucleic acid read information. Such a platform for analyzing these output data are provided by PHAT, with a focus on pathogen-host relationships (Figure 1).

DNA read files are input into PHAT as FASTQ files comprised of sequence reads with per base nucleotide identities and quality scores, or pre-aligned SAM/BAM files generated via cloud-based tools such as Galaxy .Quality control analysis can be performed on individual files, with graphical and numeric quality control reports generated. Reference genomes input as FASTA files must be indexed before they can be visualized or used for analysis. Once a pair of forward and reverse reads, paired FASTQ files, and a reference genome have been input, alignment of the paired reads against the reference can occur. PHAT also supports unpaired alignment and visualization of pre-aligned sequences. The core functions of the PHAT platform as well as FASTQ quality control, sequence alignment, and its automated analyses are performed through well-known, established implementations. Namely, quality control scoring for high-throughput sequencing data is performed by FastQC sequence alignment by Bowtie2 alignment visualization by pileup.js circular genome visualization based on our enhancements to Angular Plasmid i.e available as a new project called ngPlasmid and automated variant calling by VarScan 2. The graphical user interface of the application, based on GitHub's Electron project operates in a client-server based architecture.



The development of PHAT, bring simple-to-use, cross-platform NGS analysis to off-the-shelf hardware for life scientists studying pathogen-host relationships. The study was related to human papillomavirus type 16 (HPV16) variants and their tumourigenicity in human skin using NGS, but PHAT can be applied to a wide-variety of pathogen-host relationships (e.g., genotyping of microbes such as viruses, bacteria, and protozoans from host NGS samples).

Source: Christopher M. Gibb *et al. Oxford Bioinformatics*

Bioinfo. Animation (Genomics England – NHS **)**

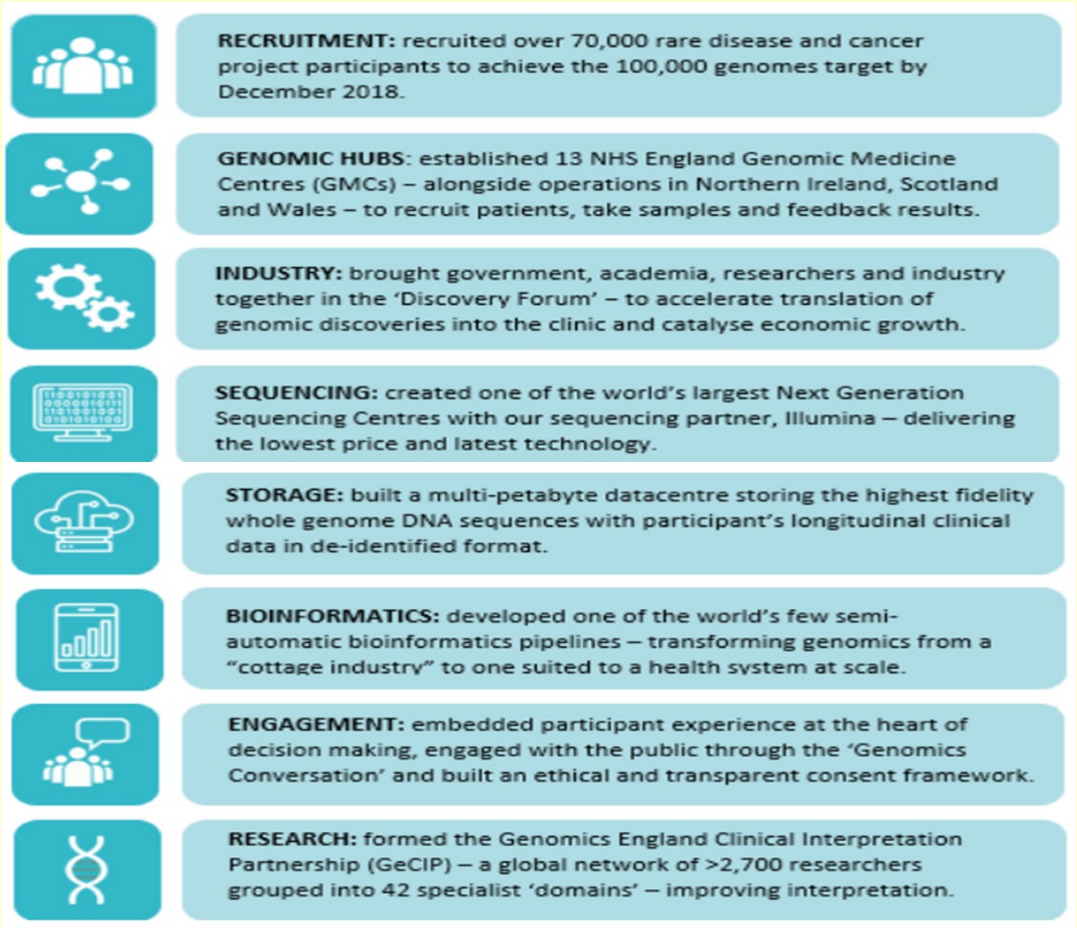
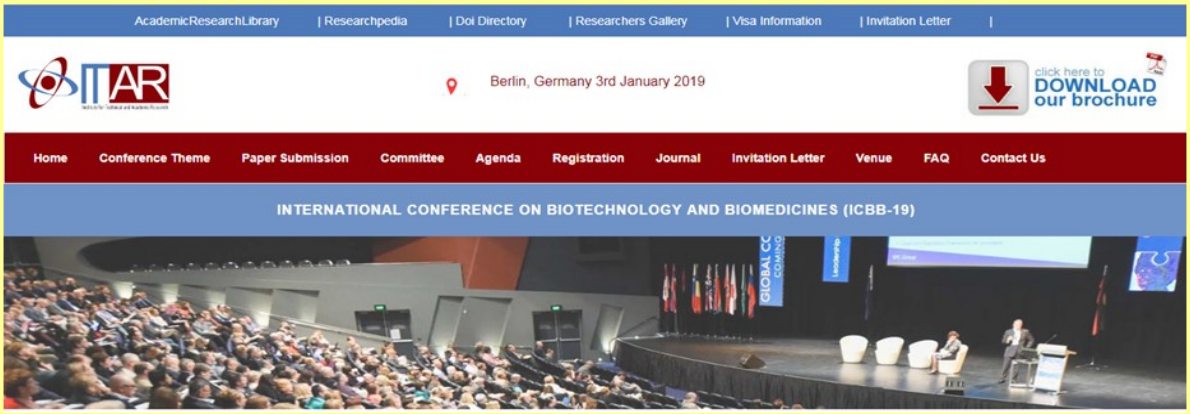


Fig: *Building a genomic medicine infrastructure*

Upcoming event



1. http://itar.in/conf/call.php?id=861480

2. https://enzymology.insightconferences.com/

5ucw, a product of directed evolution

Selection of the fitted individuals over many generations leads to the evolution of finest characteristics for the combination of two different processes of Darwin's conception of biological evolution. Biological evolution is being harnessed in the lab to create new enzymes. The use of evolution is developed to create enzymes with entirely new functions instigated by the winner of the 2018 Nobel Prize in Chemistry, Frances Arnold. The whole process begins by finding a responsive "parent" enzyme, often one that weakly performs the desired reaction or a similar reaction. Then, the parent enzyme is mutated, often at sites that are informed by the 3D structures of the protein, to create many random variants, and they are tested for the desired activity. The best performers are then selected, and subjected to an additional rounds of mutation and testing. After many cycles, an evolved enzyme with the desired new function wins out.

The enzyme from PDB entry 5ucw, performs a complicated reaction that is not performed by any natural enzymes, and for industrial settings which requires precious metal catalysts (such as rhodium, ruthenium or iridium). The assisted evolution began with a cytochrome P450 enzyme. A new serine was added to coordinate with the iron atom. An environmentally-friendly enzyme that doesn't require metal catalysts, and is highly selective for the desired reactants is the final consequence.

Building of the Drugs includes developing an enzyme with a directed evolution was used to create a complex intermediate molecule which is needed to make the diabetes drug sitagliptin and to replace the traditional chemical synthesis method, which requires high-pressure hydrogenation using a special rhodium catalyst. The new enzyme was evolved in 11 steps, first tuning the reaction, then tuning its ability to function under industrial conditions, with organic solvents and at higher temperature. The enzyme creates a very pure product, requiring fewer purification steps and using a much more environmentally-friendly process.

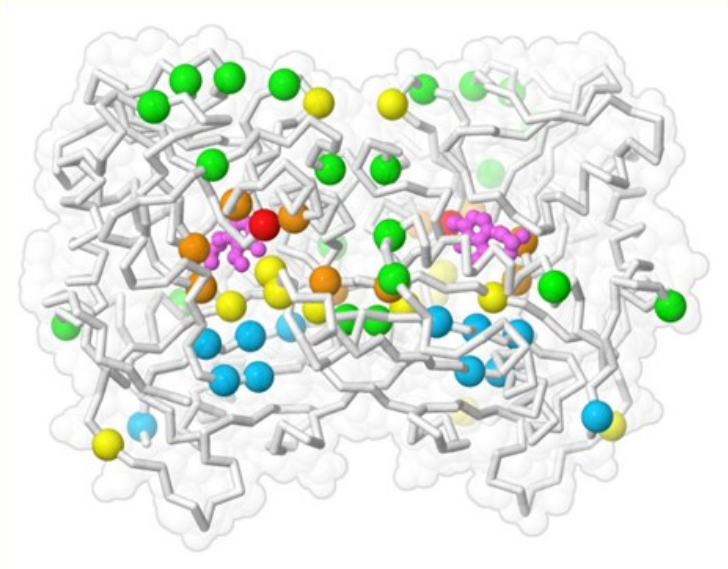


Figure: *An evolved sitagliptin-producing transaminase enzyme, with sites of mutation colored similarly to the illustration above.*

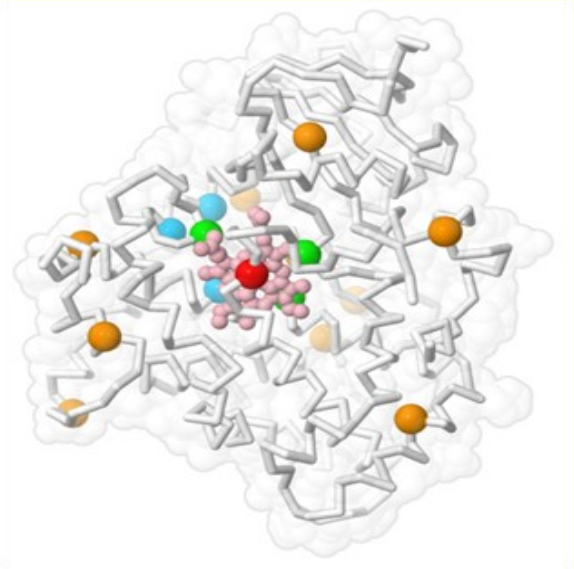


Figure:*An evolved P411 enzyme, with sites of mutation shown with colored spheres.*

Source: http://pdb101.rcsb.org/motm/228

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*