Volume 12, Issue 12 December, 2019

Bioinformatics up to Date

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

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.

Using deep learning to associate human genes with age-related diseases:

The only way to identify genes possibly related with ageing is to build a classification model (from the machine learning field) that is capable of classifying the genes as related with multiple age-related diseases. To build this model they have use a pre-compiled list of human genes related with age-related diseases and they apply a novel Deep Neural Network (DNN) method to find relation between gene descriptors (e.g. Gene Ontology terms, protein-protein interaction data, biological pathway information) and other age-related diseases.

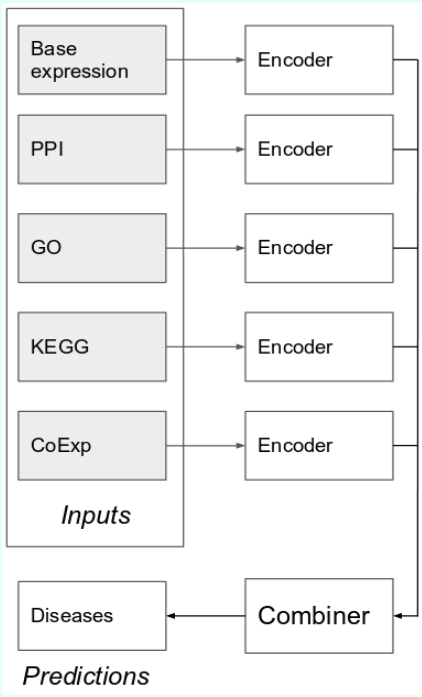

Fig: showing the predection method

The novelty of their new DNN method is its modular architecture, which have the capability of combining several sources of biological data to predict ageing-related diseases a gene may associated with it. Their DNN method have a better predictive performance than a standard DNN approaches, that is a Gradient Boosted Tree classifier (a strong baseline method) and a Logistic Regression (LR) classifier. Given the DNN model produced by their method, they use two approaches to identify human genes that are not known to be associated with age-related diseases according to the dataset. First, they investigate genes that are close to other disease-associated genes in a complex multi-dimensional feature space known by the DNN algorithm. Second, by using the class label probabilities output by their DNN approach, they identify genes with a high probability of being related with age-related diseases according to the model. They provide evidence of these putative relations retrieved from the DNN model with literature support. An increasing number of researchers are also focusing on solving the 'ageing problem',which is, trying to delay ageing in humans. This goal seems to be more and more possible in the not so distant future: biologists can already considerably extend the lifespan of several animal species such as the fruit fly and the mouse.

Source: Fabris et al. 2019, Bioinformatics, (OXFORD ACADEMICS).

Structure of H3K36-methylated nucleosome–PWWP complex reveals multivalent cross-gyre binding:

Idenfying of histone-modified nucleosomes by some specific reader domains underlies the regulation of chromatin-associated processes. On the other hand structural studies revealed how reader domains can bind to modified histone peptides, it is very unclear how reader domains interact with the modified nucleosomes. Here, they report the cryo-electron microscopy structure of the PWWP reader domain of human transcriptional coactivator LEDGF in complex with an H3K36-methylated nucleosome at 3.2–Å resolution. The structure thus reveals multivalent binding of the reader domain to the methylated histone tail and to both gyres of nucleosomal DNA, explaining the known cooperative interactions. The observed cross-gyre binding may contribute to the nucleosome integrity during transcription. The structure also explains how human PWWP domain-containing proteins are recruited to H3K36-methylated regions of the genome for transcription, histone acetylation and methylation, and for DNA methylation and repair.

Covalent modifications of nucleosomes control the chromatin-based processes such as DNA transcription, replication and its repair. Many modifications that occur on the accessible histone tails that are produced from the nucleosome core particle. These modifications include acetylation and methylation of lysine residues and are recognized by the 'reader' domains that recruit several proteins. The molecular basis for how reader domains identify histone modifications has been provided by structural studies of reader domain–histone peptide complexes. Some reader domains not only bind to modified histone tails, but can also additionally bind to DNA. Of over 20 types of reader domains, at least four (PWWP, Tudor, chromo and bromo) are known to bind both the modified histone and DNA. These reader domains are predicted to identify the histone modification in the context of nucleosomal DNA, with both types of interactions contributing to the affinity of the domain for modified nucleosomes. However, such multivalent binding of a reader domain was thus far not observed directly.

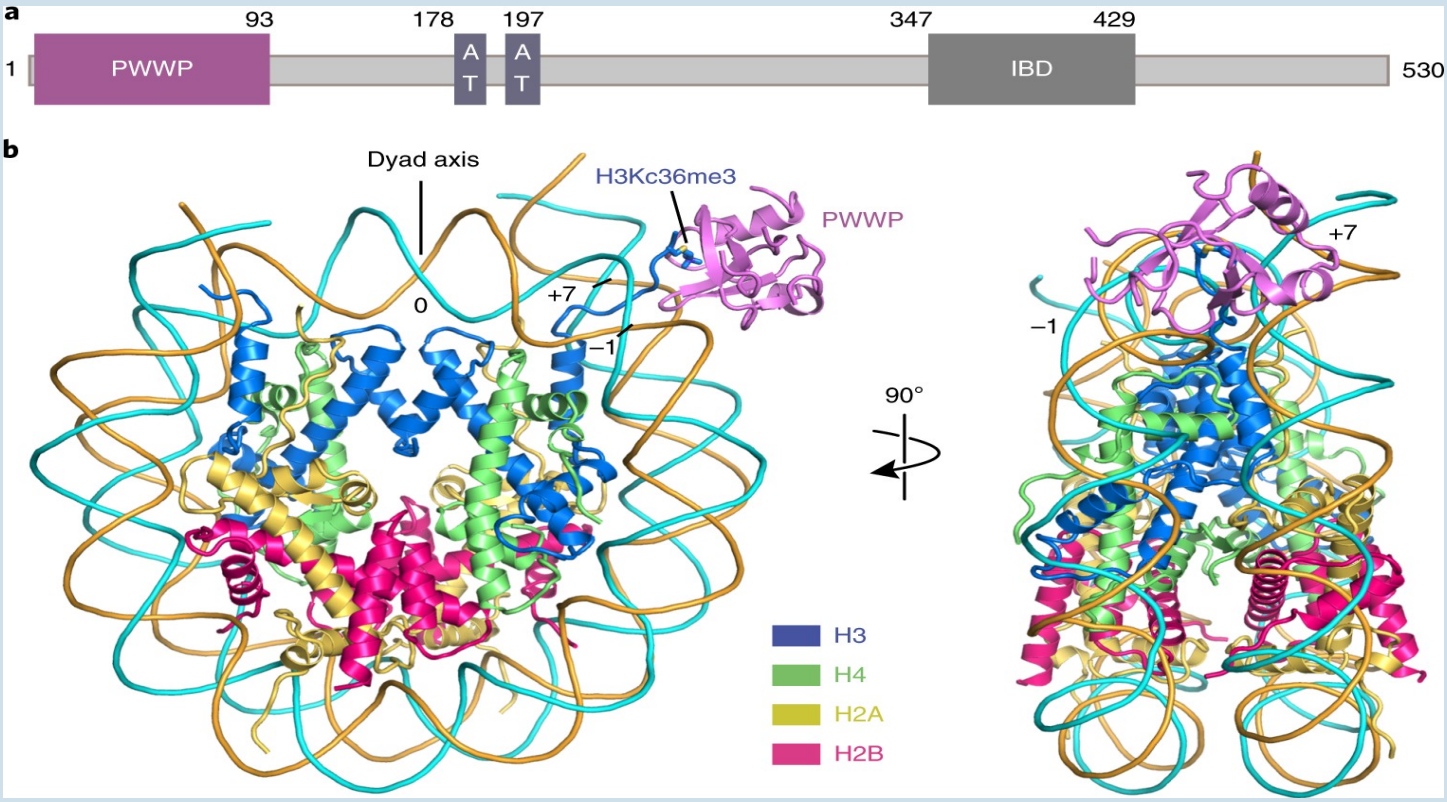

Fig: Structure of the H3KC36me3-modified nucleosome bound to the PWWP domain of LEDGF.

They collected all recombinant nucleosomes with H3K36me3 mimicked by a methyl-lysine analog. The lysine is mutated to cysteine and alkylated to form a thioether mimicking methylated lysine (H3KC36me3). The Purified recombinant full-length LEDGF bound to the H3KC36me3-modified nucleosome, as seen in an electrophoretic mobility shift assay (EMSA) (Extended Data). The complex which was assembled by mixing LEDGF and an H3KC36me3-modified nucleosome with 145 bp DNA with the canonical Wisdom 601 sequence at a molar ratio of 2:1. The complex was then isolated by size-exclusion chromatography, cross-linked with glutaraldehyde and used for the preparation of cryo-EM grids. Cryo-EM data were collected on a Titan Krios microscope (FEI) with a K2 detector (Gatan) sis of the cryo-EM data revealed weak density for LEDGF. For stabilizing LEDGF on the nucleosome, They prepared H3KC36me3-modified nucleosomes with an additional 20 bp of extranucleosomal linker DNA at the exit site. Thus the, LEDGF bound more tightly to the extended nucleosome containing 165 bp of DNA Cryo-EM analysis revealed a defined additional density for the PWWP domain of LEDGF. So they therefore subjected the modified nucleosome–LEDGF complex with the 165 bp DNA for structure determination.

Source: Wang et al. 2019, Nature Structural & molecular biology.

Bioinformatics Animation

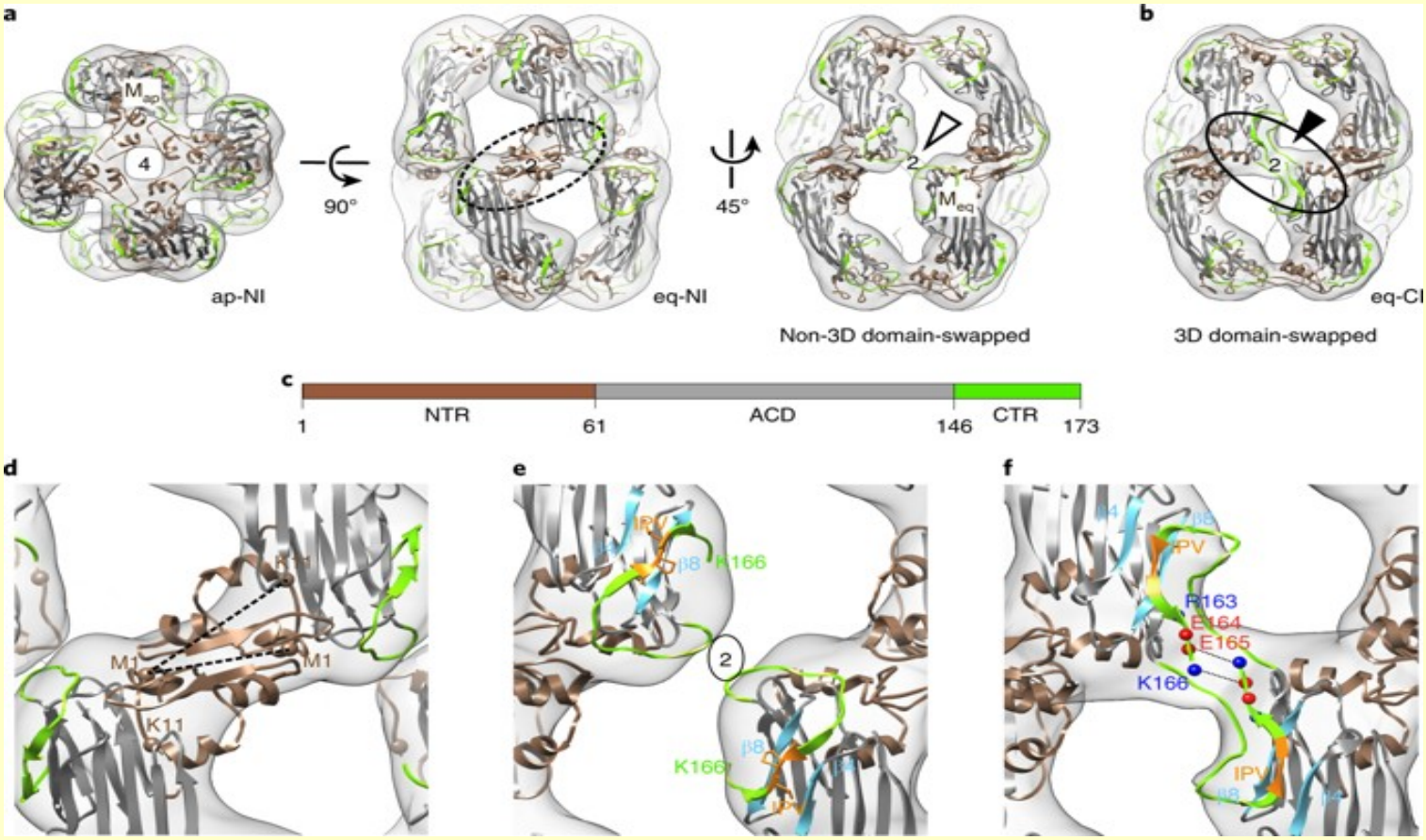

Fig: Pseudoatomic models of the human αA-crystallin (reduced) 16-mer.

Source: Kaiser *et al*. 2019, Nature Structural & molecular biology.

Upcoming Events

Medical Genomics & Metagenomics — Hands on Training in Medical Genomics & Metagenomics 14 Jan 2020 - 17 Jan 2020 • Cochin, India

Please contact +918129425785 or mail to adelbertinnovationresearchlab9@gmail.com for clarifications.

Kindly visit www.adelbertinnovationresearch.com for brochure.

Contact: Organizing Secretary; Phone: [+91 8129425785]; Email: adelbertinnovationresearchlab9@gmail.com Event

listing ID: 1305808

Event website: http://www.adelbertinnovationresearch.com

IWOMB 2020 — 3rd International Workshop on Mathematical Biology 06 Jan 2020 - 08 Jan 2020 • University of the Philippines, Los Baños, Laguna,

Event listing ID: 1265795

Related subject(s): Applied Mathematics (in general), Systems Biology and Computational Biology

Event website: https://iwomb2020.weebly.com/

1. https://conference-service.com/conferences/bioinformatics.html
2 . https://conference-service.com/conferences/bioinformatics.html

Molecule of the month

Hypoxia-Inducible Factors

Oxygen is very important with out it, our cells will rapidly die. Because of this, today we have evolved a dedicated system that monitors the amount of oxygen and controlls all the responses when levels get low (termed hypoxia). Oxygen-starved cells send out signals that make our body to create more red blood cells and build more blood vessels. The oxygen-starved cells reprogram their metabolism to shift energy production towards several pathways that do not need so much oxygen, for example, by decreasing pyruvate dehydrogenaseand increasing lactate dehydrogenase. The Nobel Prize for Physiology or Medicine was awarded this year to three researchers who discovered the molecular details of this central oxygen-sensing process, termed as the HIF (Hypoxia-Inducible Factor) system.

Hypoxia-inducible factor α (HIF-α) is the central switch that makes cells to respond to limiting oxygen. It is a protein that is about 800 amino acids in length, with several functional elements. The structure shown here (PDB entries 1lqb and 1lm8) includes a small portion of its central region, which has two key proline residues (one is shown here). When there is plentiful oxygen, these prolines are hydroxylated by PHD enzymes (HIF prolyl hydroxylases). Then, the hydroxyproline is recognized by a complex including pVHL (von Hippel-Lindau disease tumor suppressor) that targets HIF-α for ubiquitination and degradation by proteasomes. So, at normal oxygen levels, HIF-α is continuously degraded and cells carry on its usual function.

PHD enzymes helps in sensing the oxygen levels. They attach oxygen atoms to two key prolines in HIF-α using a metal ion and the cosubstrate α-ketoglutarate. When oxygen is scarce, PHD catalysis is slowed and the prolines are did not get modified. An additional enzyme, termed FIH (factor inhibiting HIF), performs a second type of hydroxylation reaction, targeting an asparagine in HIF-α and modifying the way it interacts with the transcription machinery (PDB entry 1h2n, not shown).

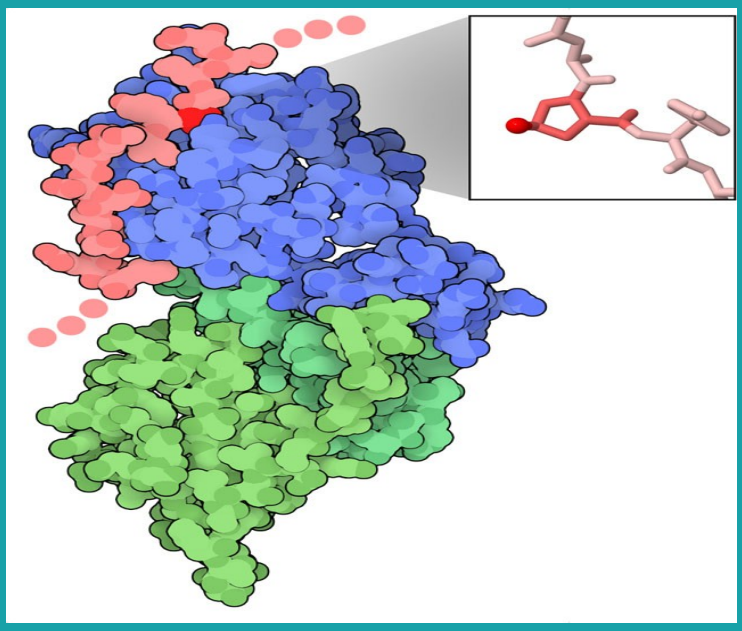

Fig:Complex of a peptide from HIF-α (pink, with proline in red), pVHL (blue), and two elongins (green). The inset shows a close-up of the hydroxylated proline.

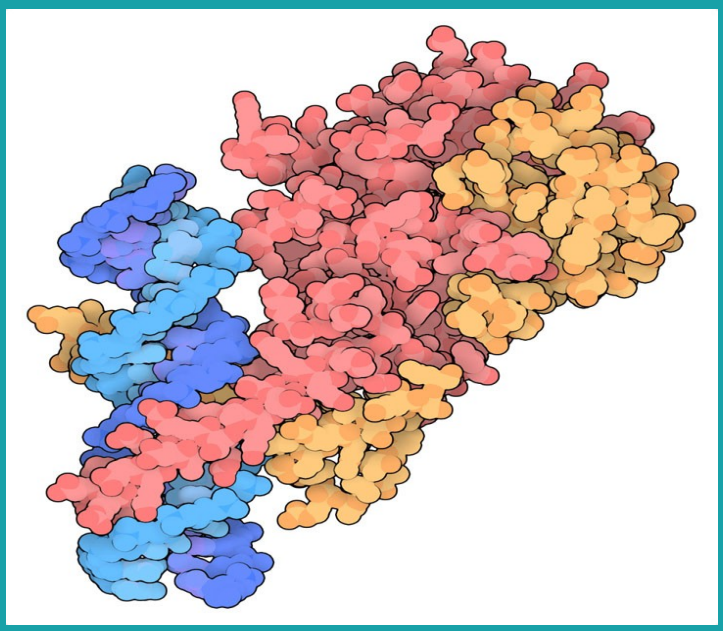

Fig: Complex of HIF-α (pink), HIF-β (yellow), and DNA (blue).

When oxygen is very low, HIF-α is not hydroxylated and are not degraded by proteasomes, so it goes into action. It moves towards the nucleus and associates with a protein, called HIF-β. Together, they may bind to many sites in the genome and promote transcription of genes involved in low-oxygen metabolism and remodeling the circulatory system to improve the oxygen delivery. The structure shown here (PDB entry 4zpr) includes the DNA-binding portion of the complex bound to a short piece of DNA.

Source: http://pdb101.rcsb.org/motm/240

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
Advanced Computation and Data Sciences Division, ACDSD
CSIR-North East Institute of Science and Technology, Jorhat, Assam
E-mail: *bio.sillayumnam@gmail.com*