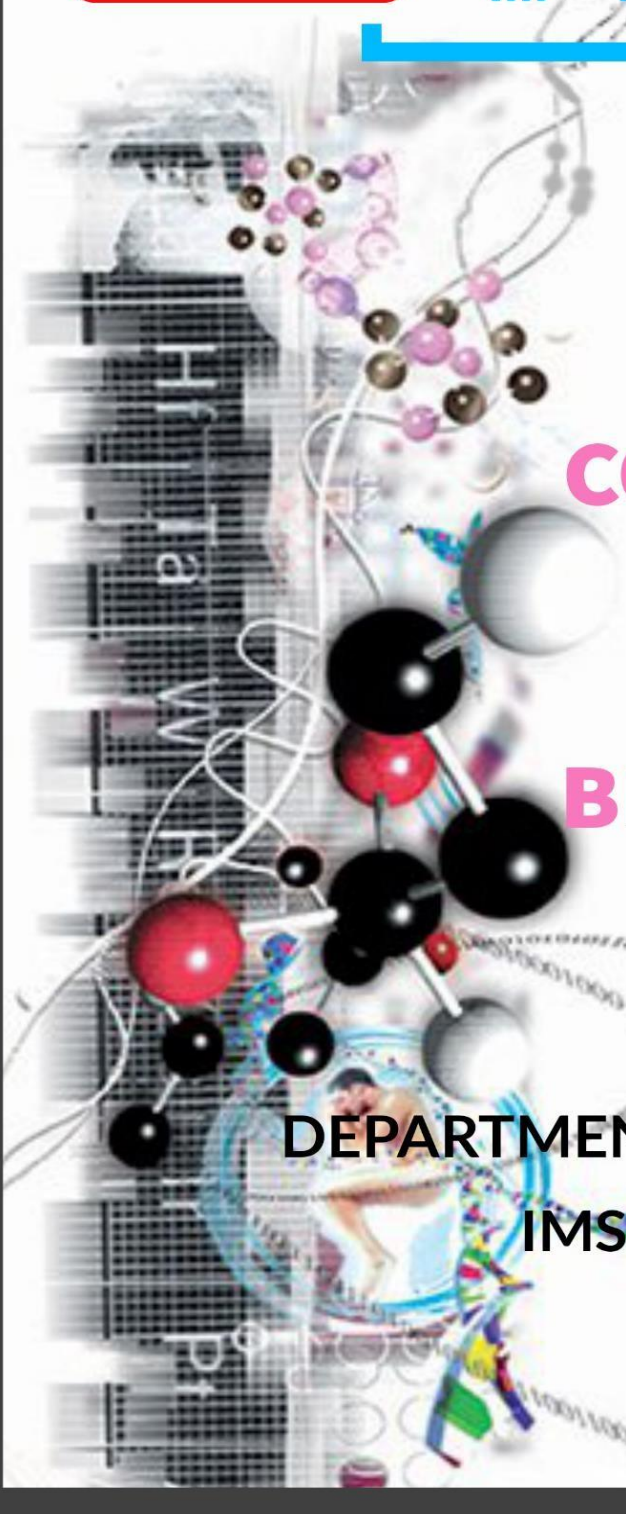
VOLUME 2• ISSUE 2• SEPTEMBER 2021

GEMS BIO VISION



ISSUE ON

COMPUTATIONAL BIOLOGY & BIOINFORMATICS

DEPARTMENT OF BIOTECHNOLOGY
IMS ENGINEERING COLLEGE
GHAZIABAD

Shri Naresh Agarwal
Visionary Guide, IMS Society

Shri Sanjay Agarwal
Chief Mentor & Guide

Prof. Prabal K. Chakravorty
Director,
IMS Engineering College, Ghaziabad

Dr. Santosh K. Mishra
Editor in Chief

Editors

Dr Prabir Kumar Paul
Mr. Priya Ranjan Kumar

Associate Editors

Mr. Aditya Raj
Mr. Prashant Singh

Gems Editorial Team

Mr. Ashutosh Khaswal
Ms. Priyanshi Garg
Mr. Keshav Saxena
Mr. Prashant Tiwari
Mr. Garvit Gupta
Ms. Rashi Tyagi
Ms. Pragya Tiwari
Ms. Taneya Gupta

Editorial Advisory

Dr. Vivek Kumar
Dr. Indu Bhatt
Dr. Shomini Parashar
Dr. Abhidha Kohli

Creative team

Mr Priya Ranjan Kumar
Dr. Meghna Singh

Publishing and Social Media Team

Dr. Indu Bhatt
Dr. Vivek Kumar
Mr. Aditya Raj

Cover page Design

Mr. Aditya Raj
Mr. Prashant Singh

FROM THE EDITOR

DR. SANTOSH K MISHRA
EDITOR-IN-CHIEF

Last two years was very adverse for human health due to Covid-19 pandemic. This problem severely affects us our educational system, healthcare facilities and human behaviour etc. This pandemic insisted us to rethink about our developmental model. Now whole word is trying to overcome from these situations. Our country also leading in the vaccine development program and we are running a largest vaccination drive. In the development of vaccine, drug molecule and data analysis computational techniques play a significant role. From last several decade human-computer interactions need an affective dimension to create realistic and believable scenarios. The recent developments in Artificial intelligence, robotics, data mining and data analysis techniques led to new and challenging applications of Computing and Human-Computer interactions. Moreover, dedicated affective mediation technologies can be effectively integrated into assistive tools to solve various problems of human especially in the area of biosciences and bioengineering. Computational thinking and techniques are so central to the quest of understanding life that toady all biology is computational biology. Computational biology brings order into our understanding of life and it makes biological concepts rigorous and testable. Biological knowledge today is defined, organized and accessed through with the help of computational biology.

In this issue of Bio vision our young biotechnologist are trying to give an insight of computational techniques in the area of biotechnology. I hope that this issue will enrich the knowledge of our readers.

INDEX

ABOUT BIOTECHNOLOGY DEPARTMENT

The Department of Biotechnology was established in the year 2002 with a clear vision of educating students with latest technology in the growing field of Biotechnology as an undergraduate course. Now the department is NBA accredited and also offers postgraduate program (M.Tech). Since its inception, the Department has continuously grown and taken initiatives to impart quality education and inculcate research aptitude in Biotechnology students. The department is actively engaged in research activities in various areas of Biotechnology and related fields. It is also an authorized research s for Ph.D. program through AKTU, Lucknow. The department has consciously taken a decision to strengthen research activity in various areas of Biotechnology with a view to develop practical solutions to problems faced by industries. The department is unique in having established expertise across a broad range of scientific disciplines, thereby encouraging innovative approaches to teaching and research. Our expert faculties in various disciplines also provide research consultancy in Environmental Biotechnology with a view to develop practical solutions to environmental problems faced by industries and the municipalities. Besides this, other areas such as animal biotechnology, plant biotechnology, microbiology, biochemistry, immunology etc are also covered for consultancy and technical support.

Government Funded Projects:-

1) DST SERB-TARE (Government of India) funded research project entitled “PfSEA1 antigen characterization as potential vaccine candidate against human malaria parasite Plasmodium falciparum” is ongoing in the Biotechnology Department.

2) A consultancy project of UP pollution control board for monitoring of ambient air quality of Hapur city is also ongoing in the department.

Major Departmental Highlights include:-

- Highly qualified faculties with extensive experience in industry, research and teaching

- 10 State-of-the art laboratories including Genetic Engineering, Fermentation Technology, Bioprocess and Computational Biology
- International internship at University of Louisville, KY, USA
- Advanced R&D in areas of Cancer Biology, Recombinant DNA Technology, Drug Discovery and Herbal Product Formulation
- Average 20 research papers published annually in peer-review journals by Faculty and Students
- Consistent 90% placement record in Core Biotechnology companies
- The students from Biotechnology department are well-placed in different MNC's and reputed government/non-government organizations.
- Many students started their own start-ups in the area of biotechnology

VISION OF DEPARTMENT

To be a Centre of Excellence in field of Biotechnology education, research, training and entrepreneurship guided by sound scientific principles, quality teaching and thrust for improvement.

MISSION OF DEPARTMENT

1. **To develop a strong Biotechnology Engineering program based on quality, education, research and training.**
2. **To impart quality education to the students and enhance their skills which willmake them globally competitive.**
3. **To develop trained biotechnology professionals who can contribute to the continuous improvement of biotechnological services and products.**
4. **To develop scientific and/or technical resources as per biotechnology industry demands.**

ABOUT GEMS SOCIETY

GEMS (Genetic Engineers and Molecular Scientists), the professional society of Biotechnology department was established in the year 2008. The aim of this society is to encourage students for various professional as well as social activities. These activities of this society provide students a platform where they can excel their talent in the area of science and technology with better understanding of their professional & social responsibilities. It also helps students to show leadership skills well as team work culture in among students Faculties of department constantly providing guidance necessary support students for organizing various activities of societies. GEMS Society is constantly working to organize various techno-cultural events for overall development of students. Every semester this society organizes events like Guest lecture from experts, Seminars, Innovative idea presentation, Biofiesta (Annual techno-cultural event), Plantation of herbal plants etc. Expert from industry and research institute like CDRI, BIBCOL, DABUR Research foundation, Sun Pharmaceutical Industries, Codon Biotech, Envirotech etc. delivered invited talk in the various activities organized by GEMS society. The structure of this society is as follows:-

President : Dr. Prabir Kumar Paul (HoD, BT)

Coordinator : Dr. Santosh Kr. Mishra (Assistant Professor, BT)
Vice President : Mr. Aditya Raj (Student of BT 4th Year)
Secretary : Ms. Rashi Tyagi (Student of BT 3rd Year)
Joint Secretary : Ms. Manisha (Student of BT 2nd Year)
Treasurer : Dr. Ashwini Kumar (Assistant Professor, BT)
Members of GEMS : All students and Faculty members of Biotechnology Department

ABOUT THE FACULTY

Prof. Prabir K Paul
Ph.D. (University of Delhi)

Ms. Shomini Parashar
M.Sc. (Gurukul Kangri University, Uttrakhand)

Dr. Santosh K Mishra
Ph.D. (AKTU, Lucknow)

Dr. Vivek Kumar
Ph.D. (Delhi University)

Dr. Meghna Singh
Ph.D (JNU,New Delhi)

Dr. Indu Bhatt
Ph.D. (Banasthali Vidyapith, Rajasthan)

Mr. Priya Ranjan Kumar
M.Tech (IIIT, Hyderabad)

Dr. Abhidha Kohli
Ph.D. (Graduate Institute of Biomedical Sciences, Chang Gung University, Taiwan)

Dr. Ashwini Kumar
Ph.D. (NIT, Raipur)

Dr. Uday Kumar Gupta
Ph.D. (IIT, BHU)

Dr. Saswat Kumar Pradhan
Ph.D. (BITS, Pilani)

Mr. Umesh Bhati
M.Tech (IIT, BHU)

Dr. Ashutosh Kumar Pandey
Ph.D. (MNNIT, Allahabad)

Dr. Namrata Yadav
Ph.D. (IIT, BHU)

Dr. Anupama Singh
Ph.D. (IIT, Patna)

Mr. Neeraj Agarwal

MACHINE LEARNING APPROACH IN BIOINFORMATICS

GARVIT GUPTA

DEPARTMENT OF BIOTECHNOLOGY

IMS ENGINEERING COLLEGE, GHAZIABAD

The origin of Bioinformatics can be from the Mendel's discovery of genetic inheritance in 1865. Since 1953 big revolution achievements took place by James Watson and FrancisCrick as they determined the structure of DNA. Bioinformatics is an interdisciplinary field that develops methods and software tools for understanding biological data. Machine learning is the adaptive process that makes computers improve from experience, by example, and by analogy. Machine learning includes the learning speed, the guarantee of convergence, and how the datacan be learned incrementally. We usuallyrefer to methods like Artificial Neural Networks (ANNs), Genetic algorithms (GAs), and Fuzzy systems along with hybrid methods including a combination of some of these methods.

While the first wave of computational analysis did focus on sequence analysis, where many highly important unsolved problems still remain, the current and futureneeds will in particular concern sophisticated integration of extremely diverse sets of data. These novel types of data originate from a variety of experimental techniques of which many are capable of data production at the levels of entire cells, organs, organisms, or even populations. The main driving forcebehind the changes has been the advent of new, efficient experimental techniques, primarily DNA sequencing, that have led to an exponential growth of linear descriptions of protein, DNA and RNA molecules. Other newdata producing work as massively parallel versions of traditional

experimental methodologies. Genome-wide gene expression measurements using DNA microrarrays is, in essence, a realization of tens of thousands of Northern blots. The large amounts of data create a critical need for theoretical, algorithmic, and software advances in storing, retrieving, networking, processing, analyzing, navigating, and visualizing biological information. In turn, biological systems have inspired computer science advances with new concepts, including genetic algorithms, artificial neural networks, computer viruses and synthetic immune systems, DNA computing, artificial life, andhybrid VLSI-DNA gene chips .

Computational tools for classifying sequences, detecting weak similarities, separating protein coding regions from non-coding regions in DNA sequences, predicting molecular structure, post-translational modification and function, and reconstructing the underlying evolutionary history have become an essential component of the research process. This is essential to our understanding of life and evolution, as well as to the discovery of new drugs and therapies.

Bioinformatics has emerged as a strategic discipline at the frontier between biology and computer science, impacting medicine, biotechnology, and society in many ways.

Large databases of biological information create both challenging datamining problems and opportunities, each requiring new ideas. In this regard, conventional computer science algorithms have been useful, but are increasingly unable to address many of the most interesting sequence analysis problems.

Machine-learning approaches (e.g. neural networks, hidden Markov models, vector support machines, belief networks),on the other hand, are ideally suited for domains characterized by the presence of large amounts of data, “noisy” patterns, and the absence of general theories. The fundamental idea behind these approaches is to learn the theory automatically from the data, through a process of inference, model fitting, or learning from examples. Thus they form a viable complementary approach to conventional methods. The aim of this book is to present a a

broad overview of bioinformatics from machine learing perspective.

Machine-learning methods are computationally intensive and benefit greatly from progress in computer speed. It is remarkable that both computer speed and sequence volume have been growing at roughly the same rate since the late 1980s, doubling every 16 months or so. More recently, with the completion of the first draft of the Human Genome Project and the advent of high-throughput technologies such as DNA microarrays, biological data has been growing even faster, doubling about every 6 to 8 months, and further increasing the pressure towards bioinformatics. In our minds, in fact, there is little difference between machine learning and Bayesian modeling and inference, except for the emphasis on computers and number crunching implicit in the first term. It is the confluence of all three factors—data, computers, and theoretical probabilistic framework—that is fueling the machine-learning expansion, in bioinformatics and elsewhere.

and it is fair to say that bioinformatics and machine learning methods have started to have a significant impact in biology & medicine.

REFERENCES:-

[1] C. YR and K. M, “Interpretable machine learning in bioinformatics,” *Methods*, vol. 179, pp. 1–2, Jul. 2020, doi:10.1016/J.YMETH.2020.05.024.

[2] D. Hanauer, D. Rhodes, C. Sinha-Kumar, and A. Chinnaiyan, “Bioinformatics Approaches in the Study of Cancer,” *Curr. Mol. Med.*, vol. 7, no. 1, pp. 133–141, Feb. 2007, doi: 10.2174/156652407779940431.

[3] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, “Deep learning for healthcare: review, opportunities and challenges,” doi: 10.1093/bib/bbx044.

NEXT GENERATION SEQUENCING

SHWETA SINGH & SRASHTI JADAUN

DEPARTMENT OF BIOTECHNOLOGY

IMS ENGINEERING COLLEGE, GHAZIABAD

Nucleic acid Sequencing can also be a way for determining the precise order of nucleotides present in a given DNA or RNA molecule. The chain termination method (Sanger or di-deoxy sequencing method), published in 1977, has remained the foremost commonly used DNA Sequencing technique. The Human Genome Project, led by the International Human Genome sequencing Consortium and Celera Genomics, was accomplished with first-generation Sanger Sequencing. After the completion of the Human Genome project, demand for inexpensive and faster sequencing methods has increased greatly. The demand has been driven by the development of second-generation sequencing methods (next-generation Sequencing).

Next-generation sequencing platforms perform massively parallel sequencing during which many fragments of DNA from one sample are sequenced in unison. Massively parallel sequencing, technology facilitates high throughput screening, which allows an entire genome to be sequenced instantly.

WORKING:-

The process of next-generation sequencing is as follows, consider a single genomic DNA sample.

1) Library Preparation

Libraries are created by fragmenting sample DNA and adding specific adapters to both ends. Fragments can then be amplified and purified. During the method of adapter ligation, distinctive index sequences, or “barcodes,” are

added to each library. The barcodes are used to distinguish between the libraries.

2) Cluster Amplification

Clusters are generated through bridge amplification. DNA polymerase moves throughout the length of the DNA strand, generating its complementary strand. The initial strand is washed away, leaving only the reverse strand.

3) Sequencing

During the sequencing step, libraries are loaded onto a flow cell and put down on the sequencer. The newly identified strings of bases, called reads, are then reassembled employing a known reference genome as a scaffold (resequencing), or in the absence of a reference genome (De novo sequencing).

4) Data Analysis

The data analysis software determines nucleotides (a process called base calling) and provided the accuracy of the base calls. Today, you may use intuitive data analysis apps to study NGS data to produce sequence alignment, variant calling, data visualization, or interpretation.

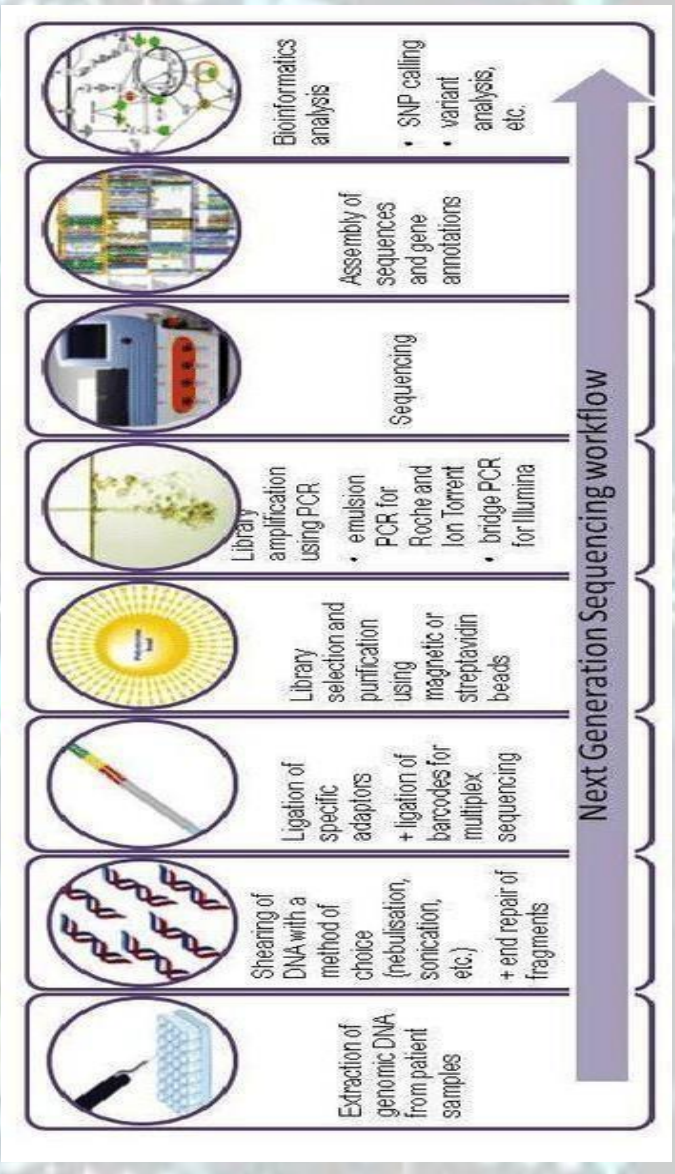


Figure 1: Flow Chart of the workflow of NGS.

TYPES:-

1) Illumina (Solexa) Technology

This technique is also called "bridge amplification" and works on sequencing by synthesis(SBS) technology where fragments of the genome to be sequenced are immobilized in a flow cell then, amplification synthesis reaction on a solid support (glass slide) that contains oligonucleotide sequences complementary to a ligated adapter, resulting clonal "clusters" consisting of about 1000 copies of each oligonucleotide fragment. Each glass slide can support many parallel cluster reactions. During the synthesis reactions, modified nucleotides, matching to each of the four fluorescently labeled bases, are incorporated and detected. These nucleotides also act as terminators of synthesis for every reaction, which is unblocked after detection for the successive round of synthesis. These reactions are iterated for 300 or more cycles.

2) Ion Torrent sequencing

Ion Torrent technology directly converts the data of nucleotide sequence into digital information on a semiconductor chip. This process exploits the fact that the incorporation of a dNTP into a growing DNA strand and involves the formation of a bond and thus the discharge of pyrophosphate and a positively charged hydrogen ion. The release of hydrogen ions results in changes within the pH of the solution, which is detected by a detector. The primary step during this approach includes library construction which involves DNA fragmentation and adaptor ligation. The fragments are clonally amplified on the small beads by emulsion PCR. Beads are primed for sequencing of nucleotides by annealing a sequencing primer, then placed into the wells of a silicon Ion chip to observe pH changes within single wells of the sequencer because the reaction proceeds stepwise. The unattached dNTP molecules are washed out before consequent cycle when a particular dNTP is introduced in it. If the introduced dNTP isn't complementary there is no incorporation and no biochemical reaction.

3) 454 pyrosequencing

In this technique, pyrophosphate is detected, which is the byproduct of nucleotide

incorporation when a base was incorporated during a growing DNA chain. Individual DNA fragments are ligated to adapters, amplified by PCR in an emulsion "bead" (emPCR) reaction. After DNA synthesis the chemical detection of reactions occurs during a picoliter-sized chamber where pyrophosphate release is measured utilizing a light-generating reaction.

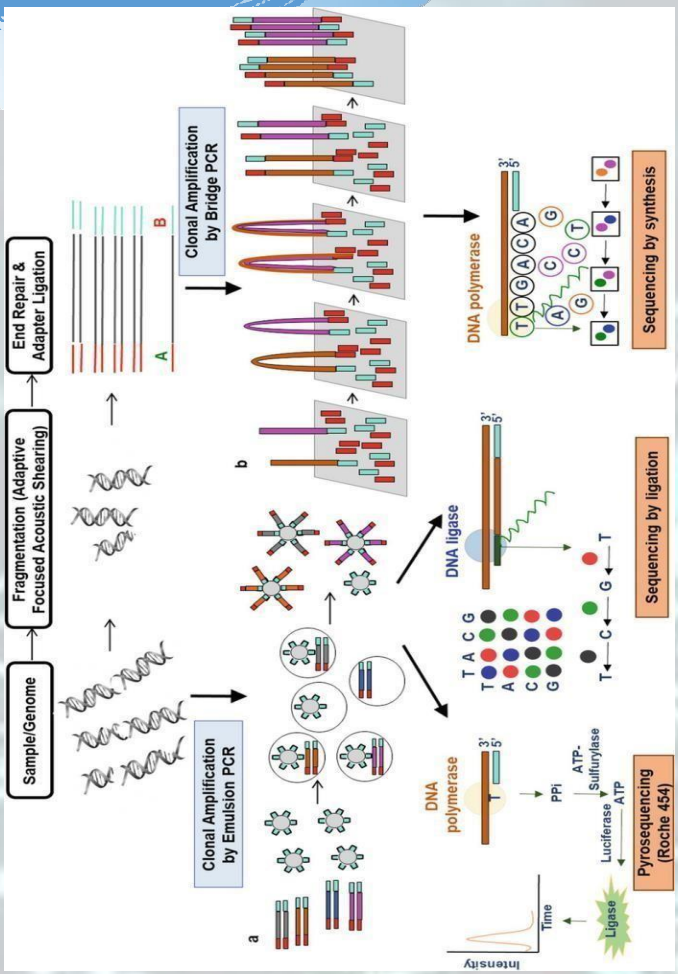


Figure 2: Flow Chart of the workflow of NGS.

APPLICATIONS:-

- Detection of unknown diseases caused by viral pathogens and discovery of Novel Viruses.
- Detection of Tumor Viruses.
- Characterization of the human virome.
- Study of Antiviral Drug Resistance.
- Quality Control and potency of Live-Attenuated Viral Vaccines.
- Epidemiology of viral Infections and evolution of virus

REFERENCES:-

[1] M. WR, M. JD, and M. ER, "Next-Generation Sequencing Technologies," *Cold Spring Harb. Perspect. Med.*, vol. 9, no. 11, Nov. 2019,doi:10.1101/CSHPERSPECT.A036798.

[2] L. Barzon, E. Lavezzo, V. Militello, S. Toppo, and G. Palù, "Applications of Next-Generation Sequencing Technologies to Diagnostic Virology," *Int. J. Mol. Sci.*, vol. 12, no. 11, p. 7861, Nov. 2011, doi: 10.3390/IJMS12117861.

BIOINFORMATICS AND SARS-COV-2 VIRUS

NEHA CHATURVEDI & ASHUTOSH KHASWAL

DEPARTMENT OF BIOTECHNOLOGY

IMS ENGINEERING COLLEGE, GHAZIABAD

CoVid-19 infection caused by positive sense RNA virus is a severe acute respiratory syndrome (SARS-CoV-2) that becomes the cause for pandemic across the world in late 2019 till date. The acuteness of this pandemic and its spread all over the world produces an unprecedented effort of communities like medicine, biology, health, bioinformatics and computer science researchers leading to the rapid development of several vaccines. Group of researchers observed that SARS-COV-2 virus controls mitochondria indirectly when enters into the host (human body) results in Manipulating or regulating mitochondrial functions just by changing open reading frames like ORF-9B. Once virus controls mitochondria of cell, It stops the immune function of host cells and promote viral replication, causing COVID-19 disease.

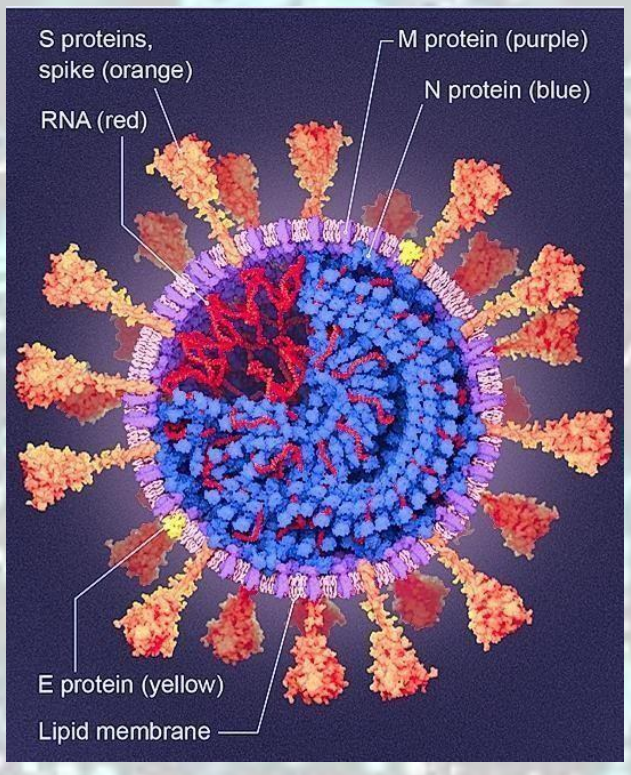


Figure 1: Describes the cross section presentation of SARS-CoV-2 virus particle

isolated from nose of infected person. It was noticed by the scientists that corona virus canbe visible in electron microscope due to its nanoscale size i.e. around 100 nm. S protein from spikes region of Corona virus represented by orange colour in the diagramis hundreds of times larger that extends fromthe surface and grab onto a human cell just by slipping inside. The structural proteins N,M and E represented by blue, purple and yellow colour in the diagram helps in formation of new virions form by movinginside the cell.

To control this virus, an urgent need is required to study and analyse the complete genome of SARS-COV-2 virus. It was suggested that Bioinformatics plays a major role in CoVid-19 drug discovery using different computational approaches with the help of various types of software's and tools. Bioinformatics software tools and databases for analysing and storing virus interactions helps in various ways. SARS-CoV-2 researches have several ideas and themes, including next-generation sequencing for genome detection, metagenomics and database storing various genome and variants.

NEXT-GENERATION SEQUENCING FOR SARS-COV-2 VIRUS:-

Next-generation sequencing [NGS] is the massive parallel sequencing tool majorly used in bioinformatics approaches that provides ultra-high throughput, speed and scalability to determine nucleotides order of targeted region of RNA/DNA or whole genome sequencing of any living organisms. Recently it was majorly used to characterize cancers at the transcriptomic, epigenetic and genomic levels. For detection of SARS-COV-2, Next-generation sequencing is the dominant technology that provides us the fundamental data about SARS-COV-2 virus like its origin and intermediate hosts. A group of scientists analyse Phylogenetic network of 160 complete SARS-CoV-2 genome sequences collected from around the world which revealed that there are three distinct but closely related variants of SARS-CoV-2. To perform Next generation sequencing, some steps are involved that are discussed below. (1) RNA extraction and fragmentation of input genomes. (2) Attaching adaptors for barcoding and preparation of a sequencing library. (3) Millions of fragments are simultaneously and

independently sequenced. (4) Human related DNA sequence reads are removed. (5) Coting's of long DNA stretches are assembled and aligned to a reference database for taxonomic classification.

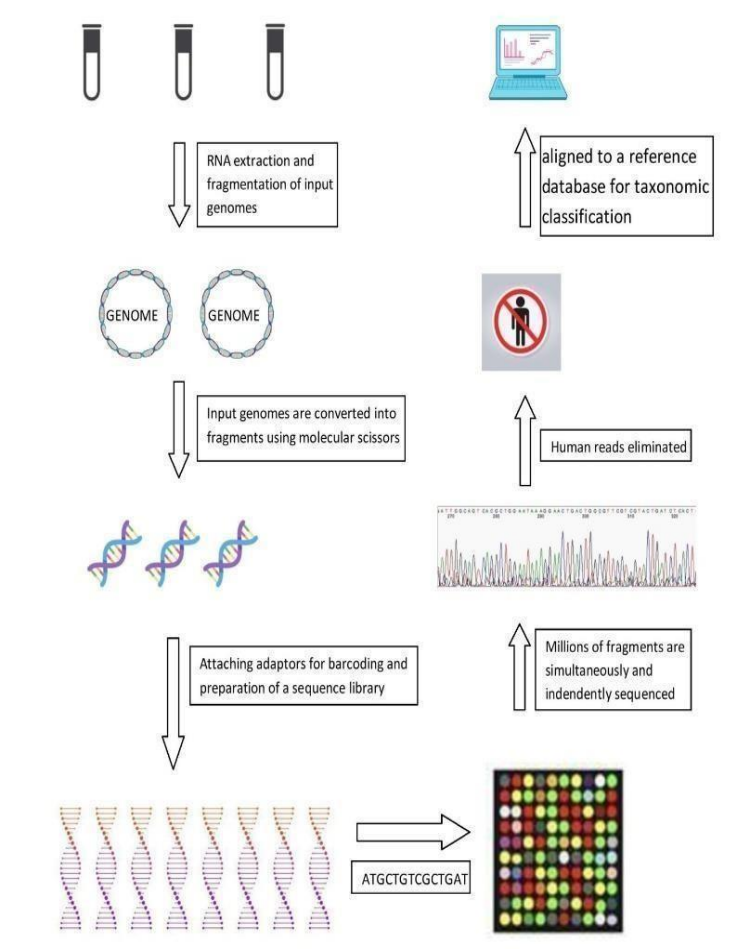


Figure 2: Flowchart for simple demonstration of Next generation sequencing steps

It was also observed that Shotgun metagenomics sequencing is an extremely powerful tool for the identification of previously uncharacterised pathogens like SARS-CoV-2 virus. It is a culture-independent technique that can interrogate all of the DNA in a sample, allowing the characterization of complex communities of microorganisms, without any prior knowledge of their genome sequences. Hence, shows its potency in SARS-CoV-2 virus detection and analysis.

Next-Generation sequencing are publicly available for researchers to study the origin of SARS-CoV-2. Also, hundreds of corona viruses and SARS-CoV-2 genomes were determined by using Next-Generation sequencing method.

CONCLUSION:-

In this COVID-19 pandemic, It was clear that ample amount of biotech researches are needed globally to prevent these kind of pandemics in near future. Present Bioinformatics tools are used widely for many vital functions like sequence data analysis, SARS-CoV-2 detection,

COVID-19 evolution, containment and tracking of COVID-19, discovery of potential drug targets and related therapeutic strategies. In this

article, It was pointed out that the next generation sequencing method is a potential tool that is widely used in SARS-CoV-2 virus genome analysis and could be helpful in future use for COVID-19 researches.

REFERENCES:-

[1] M. Cannataro and A. Harrison, "Bioinformatics helping to mitigate the impact of COVID-19-Editorial," *Brief. Bioinform.*, vol. 22, no. 2, pp. 613–615, Mar. 2021, doi: 10.1093/BIB/BBAB063.

[2] X. Chen *et al.*, "Next-Generation Sequencing Reveals the Progression of COVID-19," *Front. Cell. Infect. Microbiol.*, vol. 0, p. 142, Mar. 2021, doi: 10.3389/FCIMB.2021.632490.

[3] F. G and M. A, "COVID-19 vaccines: where we stand and challenges ahead," *Cell Death Differ.*, vol. 28, no. 2, pp. 626–639, Feb.2021, doi: 10.1038/S41418-020-00720-9.

[4] Q. C, W. AW, S. JT, L. NJ, and S. N, "Shotgun metagenomics, from sampling to analysis," *Nat. Biotechnol.*, vol. 35, no. 9, pp. 833–844, Sep. 2017, doi: 10.1038/NBT.3935.

[5] M. F *et al.*, "A complete protocol for whole-genome sequencing of virus from clinical samples: Application to coronavirus OC43," *Virology*, vol. 531, pp. 141–148, May 2019, doi: 10.1016/J.VIROL.2019.03.006.

[6] B. Hu, H. Guo, P. Zhou, and Z.-L. Shi, "Characteristics of SARS-CoV-2 and COVID-19," *Nat. Rev. Microbiol. 2020 193*, vol. 19, no. 3, pp. 141–154, Oct. 2020, doi: 10.1038/s41579-020-00459-7.

[7] R. Pereira, J. Oliveira, and M. Sousa, "Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics," *J. Clin. Med. 2020, Vol. 9, Page 132*, vol. 9, no. 1, p. 132, Jan. 2020, doi: 10.3390/JCM9010132.

[8] R. M, S. MN, S. S, and H. V, "Essential interpretations of bioinformatics in COVID-19 pandemic," *Meta gene*, vol. 27, Feb. 2021, doi: 10.1016/J.MGENE.2020.100844.

[9] K. K. Singh, G. Chaubey, J. Y. Chen, and P. Suravajhala, "Decoding SARS-CoV-2 hijacking of host mitochondria in COVID-19 pathogenesis," *https://doi.org/10.1152/ajpcell.00224.2020*, vol. 319, no. 2, pp. C258–C267, Aug. 2020, doi: 10.1152/AJPCELL.00224.2020.

[10] G. S. Watts and B. L. Hurwitz, "Metagenomic Next-Generation Sequencing in Clinical Microbiology," *Clin. Microbiol. Newsl.*, vol. 42, no. 7, pp. 53–59, Apr. 2020, doi: 10.1016/J.CLINMICNEWS.2020.03.004.

[11] P. C. Y. Woo, Y. Huang, S. K. P. Lau, and K.-Y. Yuen, "Coronavirus Genomics and Bioinformatics Analysis," *Viruses*, vol. 2, no. 8, p. 1804, 2010, doi: 10.3390/V2081803.

[12] J. S. Reis-Filho, "Next-generation sequencing," *Breast Cancer Res. 2009 113*, vol. 11, no. 3, pp. 1–7, Dec. 2009, doi: 10.1186/BCR2431.

[13] M. F. Konig *et al.*, "Preventing cytokine storm syndrome in COVID-19 using α-1 adrenergic receptor antagonists," *J. Clin. Invest.*, vol. 130, no. 7, pp. 3345–3347, Jul. 2020, doi: 10.1172/JCI139642.

PROTEIN STRUCTURE PREDICTION & VISUALIZATION

SHUBHAM TIWARI & SHUBHAM SINGH

DEPARTMENT OF BIOTECHNOLOGY

IMS ENGINEERING COLLEGE, GHAZIABAD

Mid-20th century, a diverse research took place in bio-related (Bioinformatics) programs; protein structure, visualization, and sequencing of amino acids. Protein is ubiquitous and convoluted macromolecules within life organisms; differ from organism to organism. A different spatial shape and function toward molecules. A precocious pragmatic technique developed in the past several decades, leads to exponential growth. UniProt and Protein data Bank (PDB) helps to achieve much sequence directly. X-ray crystallography and NMR spectroscopy are currently empirical techniques for protein structure prediction. This information reveals a detail of protein structure prediction through virtual functioning to obtain a 3D dimensional protein model. Ultimate goal to obtain a desired structure based on sequence alignment. You may hear the name of Homology Modelling; it is a most accurate way to predict the protein structure. In case of no similarity found in PDB, can find out, and predictstructural similarity of target protein. Both homology modelling and threading methods are used to predict a sequence. On other hand, ab initio method has also been used. In this article a detailed prediction method is discussed, including conformation initialization, conformation search, structure selection and visualization of proteins.

STEPS IN PROTEIN STRUCTURE PREDICTION:-

1) Conformation initialization

The input of protein structure prediction is in 1D and output in 3D structure. The structural

template of homologues to target protein can be identified from the PDB database by sequence alignment. In contrast, the conformation copied from the template is close to the optimal one; in case of homology modelling. The two proteins without any sequential similarity fold into similar folds. Sometimes Conformational initialization is obtained from structural templates which are vastly better than any ones built from the scratch; it dramatically curtails the process of subsequent conformational search. The template free method is the best choice for hard target proteins of those non satisfactory templates to be identified.

2) Conformational Search

Next, run a simulation with a guide of a certain force field to know the near-native conformation one by one. Almost all protein structures assemble the simulation method and do conformational search. The protein conformational energy landscape is needed in the force field to depict the conformational search. Various machine learning methods like the hidden Markov model, artificial neural network have been used for deriving energy functions. Once the energy function is determined, look for the lowest energy conformation of the target protein. Other hand in searching, molecular dynamics simulation is commonly used.

3) Structure Selection

Following the conformational search, a bulk number of target structure protein is generated. The key point of structure selection is the assessment method for different native-like structures from non-native ones. CASP helps to assess the specific structural quality. The force field itself can filter structure, but what we need to be careful of is to design a force field between the accuracy and speed contradiction. The energy function can either be physics based or knowledge based. Another way is to select structure based on clustering in all structural similarity.

4) Reconstruction and Structure Refinement

Most prediction methods adopt a protein representation for conformational search, rather than what could be obtained from one reduced model. Through reduced models all-atom be reconstructed. The backbone and side chain are

reconstructed; methods used for side chains such as Scwrl, SCATD, RASP and so on. Structural refinement selection can be done through clustering methods. The structural issues in the reduced model can affect the quality of the final all-atom structure. In fact, they refine the reduced model and structure it according to the reconstruction schedule. Structure refinement also acquires a force field to conduct molecular dynamics and Monte Carlo, but the main purpose is to improve the quality of all-atoms.

5) Visualization

Plotting the coordinate of a protein structure in 2D plane (Secondary structure) or 3D plane (tertiary structure). Visualization of any protein molecule is necessary to study the interaction between ligand and the target, active site, body, angle rotation etc. Computational technology makes it very easy to visualize any structure and it saves the cost, manpower and time. In silico visualization provides you various online and offline tools to visualize, construct or edit any protein/chemical molecule.

The importance of in silico protein structure visualization is interactive, which gives us permission to manipulate the structure of protein molecules or any other molecule through graphical user interface according to your study or work purpose. Just at a click of a mouse you rotate, zoom in and out, and moves the structure in real time. You can study and analyse a segment of the molecule in great detail. It provides various shapes and colours for your structure; manipulation of protein structure may involve adding or removing charges, adding polar hydrogens, minimizing energy and identification of active sites for docking purposes.

In-silico visualization provides various visualization styles as wire frame, ball and stick, space filling, solid spheres and ribbons. Wire frame is considered as the simplest form for structure visualization. Iti is just a line draw which represents the bond between atoms. It is a useful or skeletal view of the molecule where Cα of each residue are connected. Ball and stick model can be imagined as a sphere connected with a rod. The sphere represents the atom and the rod is the bond between atoms. In the space filling model, atoms are visualized as large solid

spheres according to the Vander waals radii of particular atoms. Ribbon diagrams are the most interactive form of visualization. Cylindrical or spiral ribbons indicate α-helices and β-sheets are represented by a flat arrow. This form allows easy identification and a clear view of topological areas of structure. Different visualization forms can be chosen as per convenience.

Here are list of some structure visualization tools: -

Rasmol *http://www.openrasmol.org/*

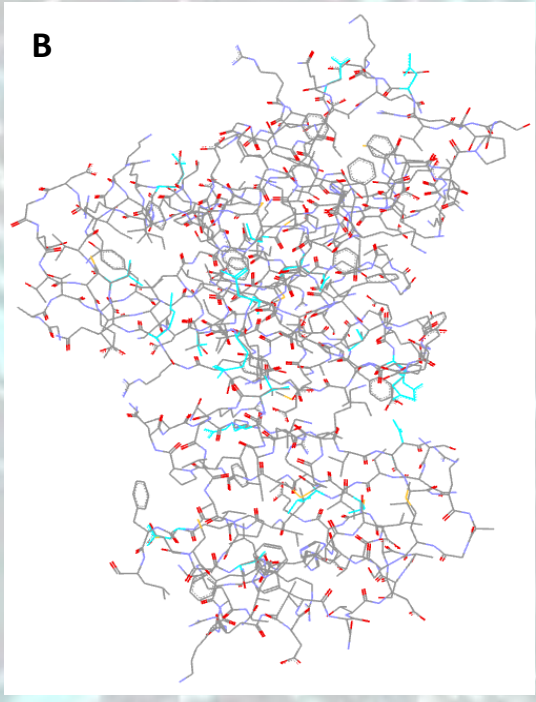
Swiss-PDB Viewer *https://spdbv.vital-it.ch/*

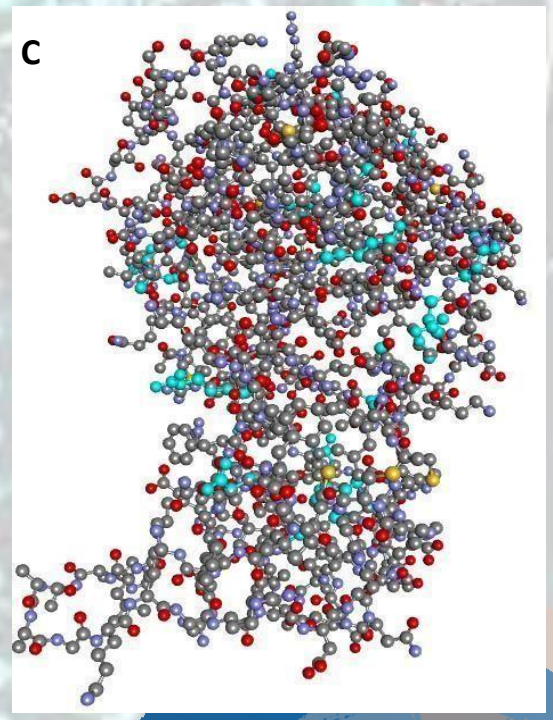
Mol-script *www.avatar.se/molscript/*

Chime *www.mdlchime.com/chime/*

Chimera *https://www.cgl.ucsf.edu/chimera/*

Pymol *https://pymol.org/2/*





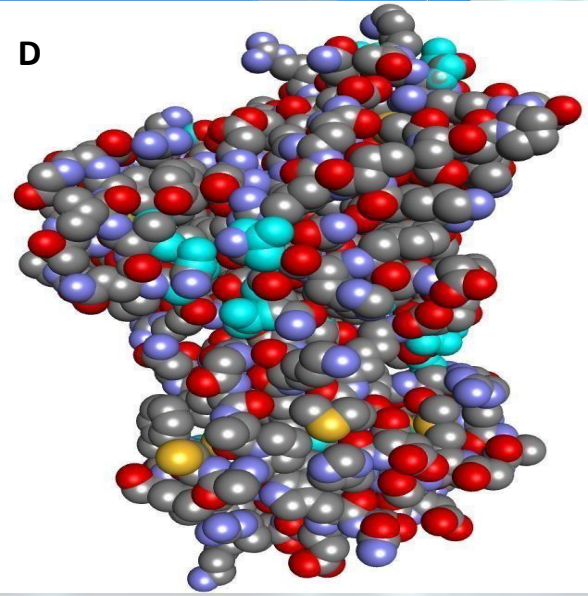


Figure 1: Examples of different visualization forms of protein molecules. (A) Ribbons (B) Wireframe (C) Ball and Stick (D) Space-filling Spheres

CONCLUSION:-

Protein structure prediction and visualization is an important step in study of gene expression and drug designing for a particular disease. Bioinformatics and computational biology really made it very easy to predict the structure of a particular protein molecule and its visualization. Various tools and applications are available to perform these steps as described above. Visualization gives a whole idea about the topology, molecular interactions, active sites and the geometry of the molecule. It helps for identifying the best position for drug interaction. Structure prediction and visualization makes your calculation for precise and accurate for docking which will result in time and money saving.

REFERENCES:-

[1] H. Deng, Y. Jia, and Y. Zhang, "Protein structure prediction," *Int. J. Mod. physics. B*, vol.32,no.18,Jul.2018,doi:10.1142/S021797921840009X.

[2] B. Y. Khor, G. J. Tye, T. S. Lim, and Y. S. Choong, "General overview on structure prediction of twilight-zone proteins," *Theor. Biol. Med. Model.*, vol. 12, no. 1, Dec. 2015, doi: 10.1186/S12976-015-0014-1.

[3] J. Moult, "Predicting protein three-dimensional structure," *Curr. Opin. Biotechnol.*, vol. 10, no. 6, pp. 583–588, Dec. 1999.

[4] T. Hameduh, Y. Haddad, V. Adam, and Z. Heger, "Homology modeling in the time of collective and artificial intelligence," *Comput. Struct. Biotechnol. J.*, vol. 18, pp. 3494–3506, Jan.2020, doi: 10.1016/J.CSBJ.2020.11.007.

[5] D. H, J. Y, and Z. Y, "Protein structure prediction," *Int. J. Mod. physics. B*, vol. 32, no. 18,Jul.2018,doi: 10.1142/S021797921840009X.

[6] X. Martinez, M. Chavent, and M. Baaden, "Visualizing protein structures — tools and trends," *Biochem. Soc. Trans.*, vol. 48, no. 2, pp. 499–506, Apr. 2020, doi: 10.1042/BST20190621.

BIOINFORMATICS IN THE ROLE OF DISEASE DIAGNOSIS & PROTEIN TARGETING

SHRADDHA PANWAR & SHUBHAM TIWARI

DEPARTMENT OF BIOTECHNOLOGY

IMS ENGINEERING COLLEGE, GHAZIABAD

Bioinformatics is a combination of biological science, computer, mathematics and statics which stores, analyses, and computes various biological data dealing with RNA, DNA and proteins. Various bioinformatics tools such as proteomics, population genetics, genomic, transcriptomic and molecular phylogenetic is extremely useful in verifying the disease symptoms, identifying drug targets and achieving results. By using bioinformatics, we can relate the origin basis of disease by sequencing of gene method furthermore analyzing the gene expressions.

ROLE IN DIAGNOSIS AND TARGETING:-

Bioinformatics provide us with various tools and databases through which targeting and diagnosis process get easy. Databases are the collection of subsets of raw data stored in an arranged manner in server. At the genetic level of disease identification is first step is to retrieve the sequence of the organism (or gene coding for affective protein) this can be done by using numerous databases such as Genebank, EMBL, Uniport, GeneCard etc. NCBI as the major database can avail the tools for comparison and proper origin of the disease that is to be diagnosed. Bioinformatics makes the structure designing and visualization very smooth. Sequencing analysing uses DNA sequencing and genome annotation which uses genomics that deals with genome editing, evolution, structure and functions. Which is helpful in spotting genetic disorders caused by mutations. Single nucleotide polymorphism are used as

detection methods for point mutation. Second methodology is analysis of cellular organism which indeed use microscopy and image analysis, protein localization nuclear organization of chromatin. A lot of gene and protein expressions are also being used in analysis of gene expressions, protein structure etc. and molecular interaction networks are extensively used.

The diagnosing of a disease, visualization of structure of affected protein molecule will help to make precise decision and effective drug molecule development. Visualization can be done by available bioinformatics tools such as RasMol, PyMol, MolMol, Discovery studio, Chimera. These visualization tools aid in identifying the specific position for ligand binding on the target molecule. The process chain which is being followed in identification of disease, is predicting gene relevant to the disease constructing gene networks further connecting the protein-protein interaction network hence if the role of protein is correctly understood then the drug which could be used for cure is identified.

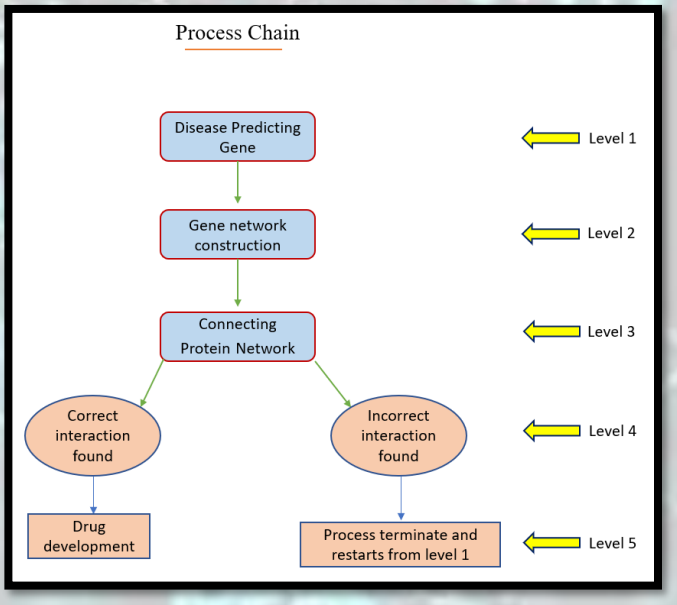


Figure 1: Flow Chart of process. This flow chart indicates the various steps at different levels in the drug designing with correct identification of the target.

For the identified disease, drug designing is a major time taking process. It takes around 14 years to avail the drug for public use. In-Silico drug designing minimize the time consumed which help in bioinformatics to allow target identification, target visualization, ligand designing, ligand validation and docking. In-silico drug design allows minimum human work time and is highly cost effective. It increases the efficiency of disease diagnostic many folds as

most appropriate match of target and disease is predicted by help of various bioinformatics tools. Ligand is a chemical molecule which can be treated as drug component after successful chemical trials. Molinspiration, Expassy, PreADMET are online available tools of ligand designing. These tools allow you to calculate the bond angles no. of hydrogen, lipophilic properties. Apart from designing ligands databases are also available that can be used a source of chemical molecules.

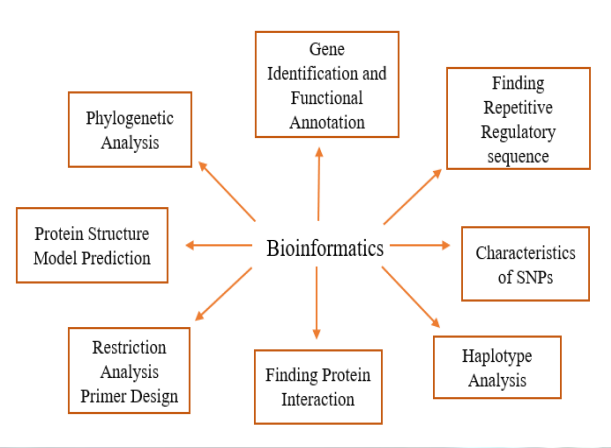


Figure 2: Various application of Bioinformatics. It can be used for interaction identification analyzing the phylogenetic tree, modelling the protein structure from sequences. Bioinformatics made primer designing very easy.

NGS has been reported as the advance method for sequencing it helps in mapping the genetic contribution of complex diseases such as cancer, Alzheimer. Cutting edge sequencing (NGS) is a hugely equal sequencing innovation that offers super high throughput, versatility, and speed. The innovation is utilized to decide the request for nucleotides in whole genomes or designated districts of DNA or RNA. NGS has reformed the organic sciences, permitting labs to play out a wide assortment of utilizations and study natural frameworks at a level at no other time conceivable. NGS can make easy and flexible classification of various type of cancer and allows to track the patient's genetic activity

CONCLUSION:-

Bioinformatics is extremely useful in analysing and controlling infection and even with the help of various tools, the pathogen which has caused the infection can also be identified. The bioinformatics is even used to make customized medicines which is specially designed for a particular patient according to his/her genetic makeup. Bioinformatics is an enlarging branch of science in every aspect more and more

applications are recognized in various fields such as drug discovery, target identification, ligand designing, target and ligand binding etc. genome studies and proteomics is extremely useful for cure of genetic disorders which are not curable, if the comparison of genetic makeup would not be done. Hence making bioinformatics valuable and useful in disease targeting and diagnostics.

REFERENCES:-

[1] A. M. Lesk, "Bioinformatics," *Encycl. Br.*, Jul. 2013, Accessed: Jun. 30, 2021. [Online].Available:https://www.britannica.com/science/bioinformatics.

[2] Z. Jiang and Y. Zhou, "Using bioinformatics for drug target identification from the genome," *American Journal of PharmacoGenomics*, vol. 5, no. 6. Am J Pharmacogenomics, pp. 387–396, 2005, doi: 10.2165/00129785-200505060-00005.

[3] R. Chaudhari and A. Joshi, "Targeted drug delivery for personalized cure," in *Advances in Personalized Nanotherapeutics*, Springer International Publishing, 2017, pp. 97–115.

[4] L. Suneetha, P. Marsakatla, R. S. Suneetha, and S. Suneetha, "Bioinformatics as a Tool to Identify Infectious Disease Pathogen Peptide Sequences as Targets for Antibody Engineering," in *Antibody Engineering*, InTech, 2018.

[5] V. Majhi, S. Paul, and R. Jain, "Bioinformatics for Healthcare Applications," in *Proceedings - 2019 Amity International Conference on Artificial Intelligence, AICAI 2019*, Apr. 2019, pp. 204–207, doi: 10.1109/AICAI.2019.8701277.

[6] X. Xia, "Position Weight Matrix, Gibbs Sampler, and the Associated Significance Tests in Motif Characterization and Prediction," *Scientifica (Cairo).*, vol. 2012, pp. 1–15, 2012, doi: 10.6064/2012/917540.

[7] A. T. M. Saeb, "Current Bioinformatics resources in combating infectious diseases," *Bioinformation*, vol. 14, no. 01, pp. 031–035, Jan. 2018, doi: 10.6026/97320630014031.

BIOINFORMATICS IN CANCER RESEARCH

SHIVANGI VERMA

DEPARTMENT OF BIOTECHNOLOGY

IMS ENGINEERING COLLEGE, GHAZIABAD

Cancer is one of leading cause of deaths in the world. No matter what your age or gender is "Bioinformatics" is one of the emerging fields, which has not only provided certain tools, software to store a large databases of a particular dieses like cancer but also helping in prediction of structure of various proteins etc, which helped researchers to discover a new drugs or methods to treat number of diseases in comparatively less time.

Few years back, scientists has adopted genital basis in order to study about the cancer. They did not centre on a specific gene; instead they are looking forward towards the substantial components of expressed genome. Bioinformatics has emerged as an essential part in not only tools or techniques but it has set a new bench mark in the area of research. In this short review we will particularly discuss about the contribution of bioinformatics in cancer researches and about the advances made in this filed which can save only masses which are suffering from this disease.

NEW TECHNOLOGIES IN BIOINFORMATICS THAT HELPS US IN CANCER STUDEIES:-

1) DNA Microarray

DNA microarray is a propitious method. This helps in analysing the any differences in the genomic DNA of a rouge (cancer) cells. Thus, by measuring MRNA transcripts in fragments in a given sample, it provides a view of genomic expression of a cancer. Microarray technology is a new and efficient approach to extract data of biomedical relevance for a wide range of applications. In cancer research, it will provide

high-throughput and valuable insights into differences in an individual's tumor as compared with constitutional DNA, mRNA expression, and protein expression and activity.

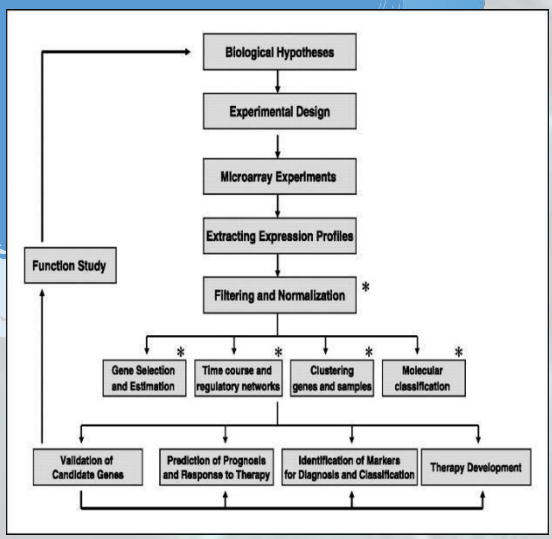


Figure 1: Flowchart to explain the method of DNA microarray

2) Role of SLAM in Cancer Research

SLAM stands for "Signalling lymphocytic activation molecule", it is one of the banderol methods which are used in cancer researches. They are generally been used in cancer researches. It gives us information clinical results and also about the cancer therapies. It also used in cancer (breast cancer mainly) diagnosis. These new findings may prompt SLAM to be considered as new tumour markers, diagnostic tools or potential therapeutic targets for controlling the tumour progression.

SLAM family	Pathology	Treatment type	Treatment name	Study phase	References
SLAMF1	ALL	Oncolytic MV vaccine strains	CAM-70	Preclinical	[28]
	Non-Hodgkin's lymphoma		$MV^{VACT}NIS$	Preclinical	[29]
	CLL		WM-63	Phase 1 clinical trial	[35]
SLAMF2	B lymphoma	anti-SLAMF2 Ab	HuLy-m3	Preclinical	[36]
	MM		1B4	Preclinical	[37]
SLAMF4	Leukemia	2B4 chimeric receptor	2B4-ζ	Preclinical	[57]
SLAMF6	CLL and lymphoma	anti-SLAMF6 Ab	994.1 and 480.12	Preclinical	[67]
	CLL		αSLAMF6 and Ibrutinib	Preclinical	[68]
		CS1-specific peptide	$CS1_{229-247}$	Preclinical	[75]
			HuLuc63	Preclinical	[70, 72]
	MM	anti-SLAMF7 Ab	Elotuzumab combined to Bortezomib	Preclinical	[71]
			Elotuzumab combined to Bortezomib	Phase 1 clinical trial	[77]
			Elotuzumab combined to Lenalidomide and Dexamethasone	Phase 1 clinical trial	[79]
			Elotuzumab combined to Lenalidomide and Dexamethasone	Phase 3 clinical trial	(ELOQUENT-1, NCT01335399)
SLAMF7	MM with renal impairment	anti-SLAMF7 Ab	10 mg/Kg Elotuzumab, 5-25 mg Lenalidomide and 40 mg Dexamethasone	Phase 1b clinical trial	[81]
			Elotuzumab	Phase 1 clinical trial	[76]
	Refractory or relapsed MM	anti-SLAMF7 Ab	10 or 20 mg Elotuzumab, 25 mg Lenalidomide and 40 mg Dexamethasone	Phase1b-2 clinical trial	[80]
			10 mg/Kg Elotuzumab, 25 mg Lenalidomide and 40 mg Dexamethasone	Phase 3 clinical trial	[82]
			10 mg/Kg Elotuzumab, 1,3 mg Bortezomib and 20 mg Dexamethasone	Phase 2 clinical trial	[78]
			10 mg/Kg Elotuzumab, 200 mg Thalidomide and 40 mg Dexamethasone	Phase 2 clinical trial	[83]
Elotuzumab in combination with Lenalidomide and Dexamethasone obtained FDA in November 2015					

Abbreviations: MM: Multiple Myeloma; CLL: Chronic Lymphocytic Leukemia; ALL: Acute Lymphocytic Leukemia.

Figure 2: Table regarding SLAM family

3) Cancer Outlier Profile analysis (COPA)

It is the oldest techniques that help in finding of cancer gene expressions. They target only those profiles which are variants of rouge cells.

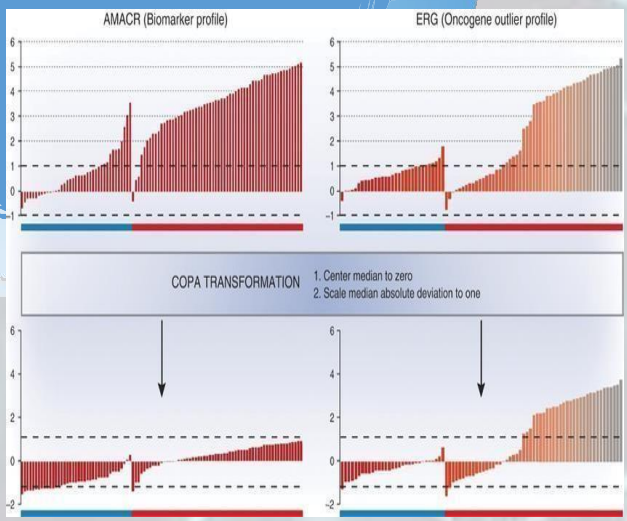


Figure 2: biomarker profile on upper left, *oncogene profile on upper right, * and COPA transformation (centre median to zero& scale median absolute deviation to one).

CONCLUSION:-

Bioinformatics has played wide roles in cancer researches. In future more new techniques and tools will be invented which will help in providing in depth knowledge about cancer. Few months back Indian scientists T-ZAR LABhave devise a non-invasive tool for cancer diagnosis. Researchers are now working to find a link between the development of cancer cells and foetus.

REFERENCES:-

[1] A. Beg and R. Parveen, "Role of Bioinformatics in cancer research and drug development," *Transl. Bioinforma. Healthc. Med.*, pp. 141–148, 2021, doi: 10.1016/B978-0-323-89824-9.00011-2.

[2] D. Hanauer, D. Rhodes, C. Sinha-Kumar, and A. Chinnaiyan, "Bioinformatics Approaches in the Study of Cancer," *Curr. Mol. Med.*, vol. 7, no. 1, pp. 133–141, Feb. 2007, doi: 10.2174/156652407779940431.

[3] D. Wu, C. M. Rice, and X. Wang, "Cancer bioinformatics: A new approach to systems clinical medicine," *BMC Bioinformatics*, vol. 13, no. 1, pp. 13–16, 2012, doi: 10.1186/1471-2105-13-71.

[4] C. Brenner, "Applications of Bioinformatics in Cancer," *Cancers (Basel).*, vol. 11, no. 11, Nov. 2019, doi: 10.3390/CANCERS11111630.

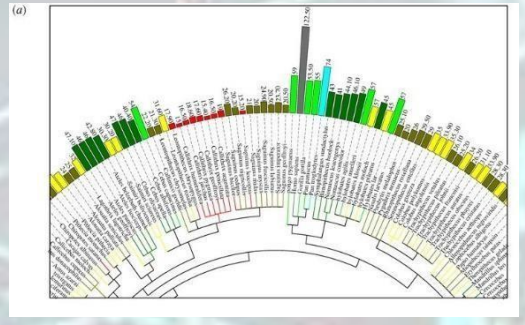
COMPUTATIONAL BIOLOGY FOR AGING

ABHINAV ROHILA & RIYA DUBEY

DEPARTMENT OF BIOTECHNOLOGY

IMS ENGINEERING COLLEGE, GHAZIABAD

Ageing in organisms is a process in which there is progressive decline in organism's health and can ultimately lead to death. The speed of ageing in different organisms is different, even when they are closely related. However, it is not clear yet, "What is the process of ageing?", but definitely genes play some role.



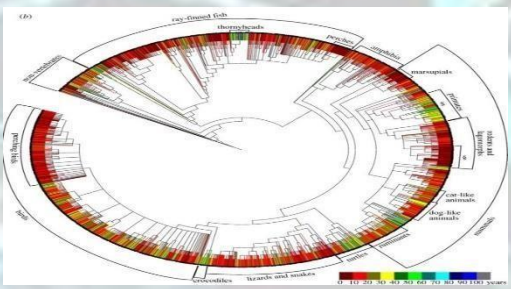


Figure 1(a, b): Bioinformatics analysis of phylogenetic clustering of longevity records

For finding the cause of ageing different computational biology comes into action. Computational biology comprises with computational tools which are available for researchers.

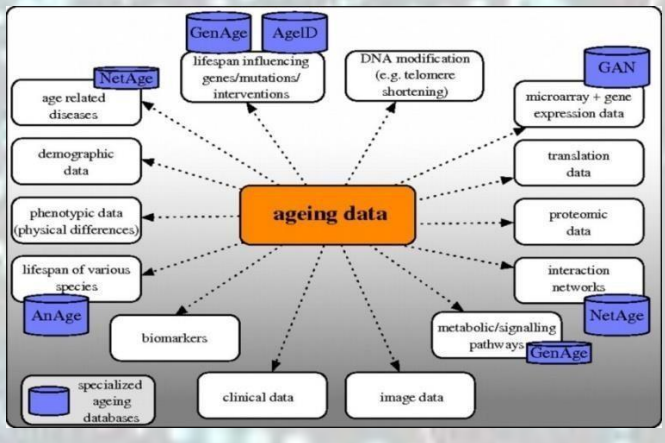


Figure 2: Bioinformatics analysis of phylogenetic clustering of longevity records

RESOURCES, DATABASES & TOOLS:-

Data collection in ageing researches are not easy to collect as there are different researches on different platforms around the world. A variety of data could be seen in the form of genomics, proteomics, metabolomic and transcriptomic. HUMAN AGEING GENOMIC RESOURCES (HAGR) database have following components: GenAge and AnAge .

1) GenAge

GenAge is a database related to ageing and genes. GenAge identifies gene pathways related to ageing & studies wrt to a database containing a list of human genes analysed for their connection with human lives.

2) AnAge

AnAge is an informative database related to ageing & life history of animals including previous records. AnAge was majorly developed for comparative biological studies for the ageing & life related aspect.

3) Gene Aging Nexus

Gene Aging Nexus is a web database and data search platform for microarray data on aging that is freely accessible to query, analyse & visualizing cross-species data on aging.

database	description	URL
GenAge (HAGR)	a database of genes related to longevity or ageing	http://genomics.senescene.info/genes
AnAge (HAGR)	a database of longevity and ageing in animal species	http://genomics.senescene.info/species
Gene Aging Nexus	a data mining platform for the biogerontological–geriatric research community	http://gan.usc.edu
AgeID	ageing genes and interventions database	http://uwaging.org/genesdb
NetAge	a database containing miRNA-regulated PPI networks for longevity, ARDs and ageing-associated processes	http://netage-project.org/

Figure 3: Databases currently available for searching and/or downloading data related to ageing

DATA PROCEDURAL WAYS:-

1) Gene-Expression Analysis

In this way of computational aging, most of the gene expression data originates from microarrays. For experiment at least 2 different conditional genes are taken & are hybridized in micro-arrays for the detailed bio-informatics

analysis. These -omic techniques such as (microarrays proteomics and metabolomics) requires specific computational methods. Combining the results of the different techniques is a general challenge in biological research which needs to be addressed in order to get a more complete understanding.

2) Proteomics & Metabolomics

As gene expression uses microarrays or Next Gen Sequencing, Proteomics works on protein profiling. Different proteomics techniques have different focuses like protein concentration, modification, complexions etc. The no of direct targeting proteomics studies is less. Protein changes are observed which could lead to cellular maintenance & several processes using transcriptional analysis.

Metabolomics is the examination of the metabolic processes in the cells or tissues throughout. Eg. Glucose or Oxygen species are important for ageing. The comparison of long lived mutant, if profiled by combining microarray, proteomics & metabolomics would be an interesting study.

3) Pathway Analysis

One of the main challenges in interpreting gene expression data is trying to understand the biological consequences of gene-expression changes. To tackle out the challenges in interpreting gene expression data, the pathway analysis is involved, be in signalling or regulatory pathway.

We have 3 approaches in pathway analysis to ageing:

i. Analysis of protein-protein interaction
ii. Gene-regulatory network analysis
iii. Modelling quantitative properties of reaction.

4) Image analysis

Various laboratories have developed tools for the expression & automated analysis of microarray images. As a second approach, image-textured analysis is done with two characteristics reflecting grey-level co-occurrence & directionaling.

This methods, fitted to estimate the age of normal living wormsfrom images, can be used to estimate how much younger a long-lived worm looks when compared with its chronological age.

COMPARATIVE ANALYSIS OF SPECIES & TISSUES:-

By comparison of different species can help for analysis of ageing and causes of ageing. For the comparison of data information on ageing in phenotypes, experiment details and other's knowledge about the organism's genomics.

Lifespan analysis of organism can also be done for the studies of comparative analysis. For the experimenting about ageing generally in yeast and invertebrates with worms, which are mostly used. These organisms have short lifespan and could be used as model organism for research purpose.

1) **Within Species (Comparing genotypes)**

By comparing genotype within species, certain information could be gained regarding ageing. Example: A study in human beings stated similarity and difference in expression changes in the medulla and cortex of kidney with age.

2) **Within Species (Tissue Specificity)**

Ageing could also be associated with changes in tissue specificity. Example: Decrease of genes which makes the mitochondrial electron transport chain with increasing age.

3) **Within Species (Population Studies)**

In all types of comparative studies the most powerful study is population based. Example: Datasets have been collected from different families which shows longitivity exceptionally. The long-lived have been established as 'model' populations in study of longitivit.

CONCLUSION:-

Computational systems modelling is a novel integrated approach that provides a powerful foundation for gaining an in-depth understanding of how human metabolism is perturbed by aging. It also highlighted the rationale for using computational systems models. The steps involved in the model

building process were also outlined, and a wide variety of models from cellular to whole body were discussed that emphasized the utility of modelling to aging research. It is highly probable that in future years computation systems modelling will be further embedded within systems biology.

Even the data is collected and organized from various ageing researches & studies for their examination, they have one thing which is common i.e. the techniques are dependent on Bioinformatics which is omnipresent in the way that it is used extensively in all the fields of ageing studies from genomics to proteomics.

The majority of the research in ageing involve proteomics, which results in making data sets with data from high-expression combining the resulting datasets could lead to more useful information results at the mechanism of ageing. As we are gaining the knowledge and information, there is a need for the betterment & development of the tools for more critical computational biology

REFERENCES:-

[1] M. T. M. Auley and K. M. Mooney, "Computational Systems Biology for Aging Research," *Interdiscip. Top. Gerontol.*, vol. 40, pp. 35–48, 2015, doi: 10.1159/000364928.

[2] R. A. Avelar *et al.*, "A multidimensional systems biology analysis of cellular senescence in aging and disease," *Genome Biol. 2020 211*, vol. 21, no. 1, pp. 1–22, Apr. 2020, doi: 10.1186/S13059-020-01990-9.

[3] D. Wieser, I. Papatheodorou, M. Ziehm, and J. M. Thornton, "Computational biology for ageing," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 366, no. 1561, p. 51, Jan. 2011, C

[4] R. Song, E. A. Sarnoski, and M. Acar, "The Systems Biology of Single-Cell Aging," *iScience*, vol. 7, pp. 154–169, 2018, doi: 10.1016/j.isci.2018.08.023.

TEST YOURSELF

1. The method used for prediction of three dimensional structure of a protein from known structure(s) of one or more related proteins is_____

(A) MSA | (B) Homology Modelling | (C) Docking | (D) Phylogeny

2. Bioinformatics cannot analyse____

(A) Mathematical analysis | (B) Statistical analysis | (C) Biomedical analysis | (D) Chemical analysis

3. Bioinformatics comprise of all except___

(A) DNA microarrays | (B) Medical informatics | (C) DNA microarrays | (D) Agarose gel Electrophoresis

4. The categories in which Machine learning approaches can be traditionally categorized

(A) Supervised learning | (B) Unsupervised learning | (C) Reinforcement learning | (D) All of the above

5. You need to use a first generation sequencing method for de novo sequencing, which template should give optimum results for this project?

(A) Genomic DNA | (B) PCR product | (C) Plasmid DNA | (D) BAC

6. To avoid ligation of separate DNA fragments, which of the enzyme is used?

(A) Phosphate | (B) Kinase | (C) Ligase | (D) Endonuclease

7. Which of these projects would be best suited for Next Generation Sequencing?

(A)To determine if tumour sample contains a common missense mutation | (B) To find the Transcriptome of a tumour sample | (C) To genotype ten genomic DNA samples for a known SNP | (D) All of the above

8. Which of the following is an anti-apoptotic protein?

(A) NOXA | (B) BCL-Xs | (C) Bim | (D) Bfl-1

9. The common sequence at the beginning of each gene in coronavirus is _____

(A) CUUAACAA | (B) CCCUUUAA | (C) UUCCAAUU | (D) UCUAUAUA

10. For prediction of three dimensional structure of protein
(P) Homology modelling tries many possible alignments
(Q) Threading first identifies homologous
(R) Threading evaluates many rough models
(S) Homology modelling optimizes one model

(A) Q,S | (B) P,Q | (C) R,S | (D) Q,R

RECENT NEWS

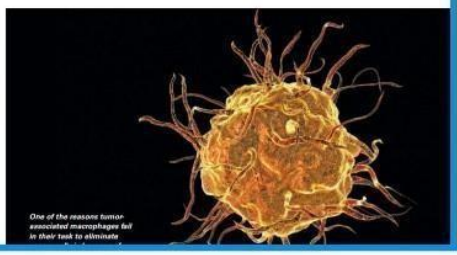
Tools against tumors

Grant from Innovate UK to Pathios seeks to advance cancer immunotherapy

BY ILENE SCHNEIDER

OXFORD, UK—Pathios Therapeutics Ltd., an early-stage drug discovery company focused on developing macrophage conditioning to unlock the power of the innate immune system against cancer, has received a $475,000 Smart Grant from Innovate UK to accelerate its cancer immunotherapy program targeting the innate immune checkpoint known as GPR65. The company is targeting a genetically validated innate immune checkpoint as an approach to unleashing the power of macrophages in the fight against cancer.

Pathios, which was launched in 2017, will use the grant to collaborate with researchers from the Department of Oncology at The University of Oxford to develop tools to expedite the translation of small-molecule GPR65 inhibitors for use in cancer immunotherapy. The program will include the development of early clinical target engagement biomarkers as well as using bioinformatics techniques to identify patients



One of the reasons tumor-associated macrophages fail in their task to eliminate

Novel SARS-CoV-2 neutralizing antibodies

Therapeutic candidates retain binding to major COVID-19-causing "South African" variant

BY DDN STAFF

CAMBRIDGE, UK; PORTO, Portugal & LONDON—IONTAS Ltd., which focuses on the discovery and optimization of fully human antibodies, has announced the identification of neutralizing antibodies that bind to multiple emerging SARS-CoV-2 variants.

The results, generated from a combination of phage display technology and B cell receptor repertoire sequencing of hospitalized COVID-19 patients, have identified what the company called "potent neutralizing antibodies with distinct mechanisms of action." Based on this approach and its work with biologics-focused contract research organization FairJourney Biologics S.A., IONTAS has developed a panel of therapeutic candidates.

SARS-CoV-2 virus enters the body as a result of the viral spike protein binding to the ACE2 protein on the surface of long cells. This interaction is driven by a small region of the spike protein and the majority of antibodies currently undergoing clinical trial bind in this site. However, this site

Exosomes to the rescue in cancer

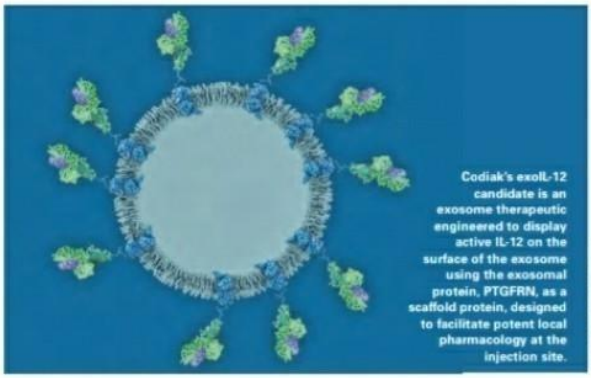
Codiak demonstrates potential of engineered exosomes to stimulate targeted, antitumor immunity

BY LORI LESKO

CAMBRIDGE, Mass.—Believed to be the first-ever company to have led preclinical trials of engineered exosomes, clinical-stage biopharmaceutical Codiak Biosciences Inc. recently unveiled the results from multiple studies demonstrating the potential of Codiak's precision exosomes to direct pharmacological payloads to specific cells—and achieve enhanced immune mediated antitumor activity with an expanded safety margin.

Codiak unveiled its results in November at the 35th Annual Meeting of the Society for Immunotherapy of Cancer (SITC), describing the exosomes' action as extracellular vesicles that act like the body's FedEx delivery trucks to carry messages between cells to stimulate targeted, integrated anti-tumor immunity in IL-12 and STING—well-validated, yet historically elusive, immuno-oncology pathways.

Codiak's engEx platform can engineer exosomes with distinct properties, load them with various types of therapeutic molecules, and alter tropism so they reach specific cellular targets such as cancer and infectious diseases. The company believes that exosomes may serve as a valuable therapeutic because of their



Codiak's exoIL-12 candidate is an exosome therapeutic engineered to display active IL-12 on the surface of the exosome using the exosomal protein, PTGFRN, as a scaffold protein, designed to facilitate potent local pharmacology at the injection site.

CREDIT: CODIAK BIOSCIENCES INC

approaches," says Douglas E. Williams, Codiak's president and CEO.

At the time he spoke with *DDN*, he expected to see results from the healthy volunteer portion of the company's exoIL-12 study by the end of 2020 and safety, biomarker and preliminary

therapeutic index with exoIL-12, delivering a more robust tumor response, dose control and an improved safety profile.

Codiak intends to focus development of exoIL-12 on tumors that have, in previous clinical testing, shown clinical responses to IL-12

retention and increased tumor growth inhibition across multiple mouse models. Additionally, these data demonstrate significant remodeling of the tumor microenvironment and confirm tissue-retained pharmacology in non-human primate models, thereby widening the therapeutic window for this potent cytokine.

Among the data presented at SITC 2020, Codiak reported that exoIL-12 was ~100 fold more potent in tumor growth inhibition than recombinant IL-12 (rIL-12), with complete responses observed in 63 percent of mice treated with exoIL-12 compared to 0 percent in mice treated with an equivalent dose of rIL-12. In addition, exoIL-12 showed dramatic change in the tumor microenvironment, triggering a ~8-fold increase in cytotoxic T-cell infiltration and ~150-fold increase in M1 macrophage recruitment. In non-human primates, exoIL-12 showed tissue-localized pharmacology, local induction of IFNg and a lack of systemic exposure.

In multiple in-vitro and in-vivo studies, the engEx Platform was found to affect cellular tropism and boost payload delivery, with Codiak noting in a press release that exosomes

WHAT'S INSIDE

WELCOMING THE FOURTH WAVE OF DRUG DISCOVERY

Amgen executive forecasts multispecific drugs as the next key movement

BY KRISTEN SMITH

THOUSAND OAKS, Calif.—Dr. Ray Deshaies, senior vice president of global research at Amgen, believes he has identified the "fourth wave" of drug discovery: the advent of multispecific drugs that induce proximity to enable targeted cell degradation. His assessments in this concept will radically transform how scientists understand how drugs work and how new methodology can be applied to targeting disease. In a paper and an accompanying perspective piece recently published in *Nature*, he explored the potential of this trajectory and what it could mean for the future of drug development.

Induced proximity is already ingrained in the body, as ubiquitin-tagged proteins can be absorbed by spontaneous and targeted. Some viruses have been able to override a cell's natural systems by linking to those proteins and tricking the cell into tracking its own defenses. Proteolysis targeting chimera (PROTAC) molecules use the same playbook, initiating the degradation process by utilizing two classes of multispecific

AMGEN CONTINUED ON PAGE 6

Here, a PROTAC (small pink molecule, lower center) works by bringing a target protein (shown in yellow) into proximity with enzymes known as ubiquitin ligases (purple). Once in proximity, the enzymes attaches a small protein called ubiquitin (green) to the target protein, which marks the target for degradation.

Deep learning for drug design

Metabolite Translator assesses safety of novel compounds

BY ILENE SCHNEIDER

HOUSTON—A deep learning-based technique created at Rice University's Brown School of Engineering is designed to tell pharmaceutical researchers how drugs in development will perform in the human body. Computer scientist Lydia Kavraki—the Noah Harding Professor of Computer Science; a professor of bioengineering, mechanical engineering and electrical and computer engineering; and director of Rice's Ken Kennedy Institute—and her team have introduced Metabolite Translator. This computational tool predicts metabolites, the products of interactions between small molecules like drugs and enzymes, and helps to determine the safety of drug candidates.

Department of Biotechnology
IMS Engineering College, Ghaziabad
NH-24, Adhyatmik Nagar, Near Dasna, Distt: Ghaziabad Uttar Pradesh-201015
Phone: (91)120-4940000 Mobile: 9821396582/83
Website: http://www.imsec.ac.in/
E-mail: enquiry@imsec.ac.in, hodbt@imsec.ac.in

http://www.imsec.ac.in

https://www.facebook.com/IMSEC143/

https://www.facebook.com/biotechnologyimsec/

https://twitter.com/IMSECGhaziabad

https://www.linkedin.com/in/imsecghaziabad/

https://www.youtube.com/channel/UCYqo6uXH4BOp4juzGZhh-zg