BIOINFOLUCENT

A BIOINFORMATICS MAGAZINE

2022

VOL 1

> *The scientific discovery appears first as the hypothesis of an analogy ; and science tends to become independent of the hypothesis.*
> *~ William Clifford*

INTERVIEW

From the professors desk
I believe that no bioscience research is complete without bioinformatics.

BIOINFOBUZZ

News on new patent laws ,comprehensive DNA sequence analysis on smartphone and a lot more

PRELUDE

The combination of biology and computers has resulted in an extensively powerful tool that is capable of changing the very essence of existence if utilized properly.
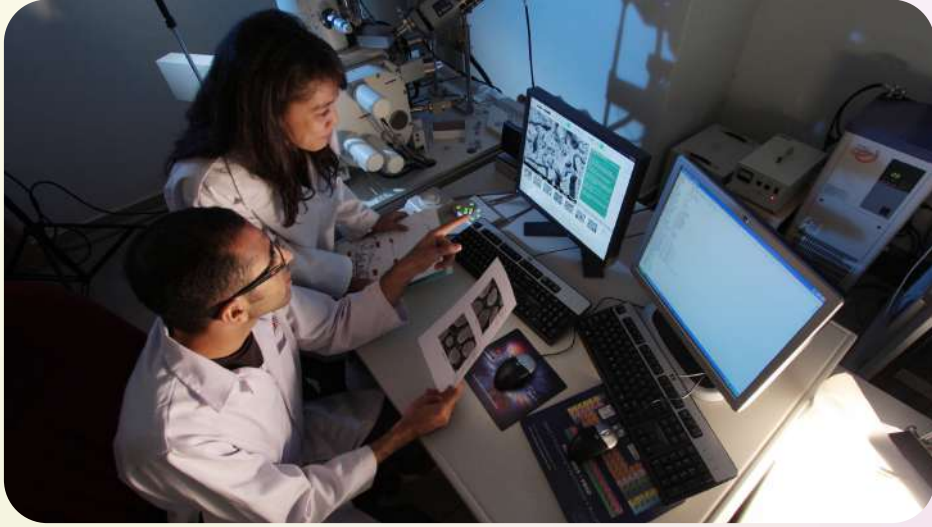
A whole new beginning

FOREWORD

Crescent Bioinformatics Association strives to elicit the different disciplines of bioinformatics among the students interested in the field of biotechnology. Being a student initiated association, their fervor has never dropped the crown. Utilizing the fullest of their online presence for the last two years, the club has heightened their interaction and has successfully conducted various Intra college and inter-college events. Through

Dr. S.Hemalatha
Professor & Dean
School of Life Sciences

ebb and flow, Crescent Bioinformatics Association continues to get the better of each and every hurdle on their way and gain ground as they march to greater heights of laurel.The journey started with a fistful of students, the exuberance of the club and its members have enormously contributed to this very level to publish their first ever magazine "BioInfoLucent". This edition has brought in the roots of bioinformatics through its articles, facts, and news, with a blend of creativity through its memes and games. I extend my heartfelt appreciation to the editorial team and the contributors for their tremendous efforts in this creation. With immense pleasure, I would like to record my heartfelt wishes for their accomplishments and their future endeavors.

● ● ●

TABLE OF CONTENTS

A letter from the Editor

It gives us immense pleasure to present to you, the Crescent Bioinformatics Magazine of the year 2022- **"BioInfoLucent "**. The name BioInfoLucent came from our desire to shed light (knowledge) on the field of bioinformatics in this magazine. The first step is usually the most challenging, and well, it was true. It took us a lot of sweat, time, pain, and joy to craft this very first magazine. We were ambiguous for days on how to begin the notion of establishing this insightful magazine, but as days went by, we got a clear sense of transparency. This creation was crafted by the curious minds of enthusiastic members of Crescent Bioinformatics Association , who put all their efforts towards this creation. The experiences and mistakes we made together while working for the magazine will live on in our hearts for the rest of our lives. I would like to thank everyone who helped in making this magazine a reality. Hopefully, this magazine will enlighten you as much as it did to us while crafting it. The introduction of this promising first-ever CBA magazine, "BioInfoLucent 2022," will encourage many more editions to follow in the coming years. I wish all the readers the best of luck and happy reading.

Thank you very much !

Tasneem Juzer
CO-PRESIDENT

A peek into the Club

The Crescent Bioinformatics Association, established in the year 2020 by the students and functioning for the students, aims to encourage interest among the students in the field of Bioinformatics- an emerging multidisciplinary field that encompasses biology, computer science, mathematics, and physical science. CBA is a fast-growing forum that has achieved several goals in a short period of time. We have conducted an introductory webinar and several group discussions, journal club (online and offline), blog and Instagram posts, coding club, online debate competition and online quiz events. Our future objectives would be to conduct workshops, seminars and conferences, quizzes, journal club, coding club and increased social media engagement.

The club and its members are truly grateful to Dr.Hemalatha S, Dean of School of Life Sciences and Dr. Subhamoy Banerjee, Faculty Coordinator of CBA for their enduring support and guidance in all our endeavours. We at CBA, work as a team to accomplish our target and we strive hard for excellence.

The club organized its first inter-college debate event "DEBioinfoATE" on the 9th of January,2022 with the intention of creating awareness about bioinformatics not only among life science students but also in other disciplines. It is also aimed at spreading knowledge about the existence of bioinformatics and how it could benefit society. Besides the participants from biotechnology, students from other colleges and departments such as Computer Science Engineering, Electrical, and Electronics Engineering also took part, which indeed fulfilled the sole purpose of the event.

CBA has its own social media handles on Instagram, Twitter and WordPress where content related to bioinformatics is posted regularly. It serves as a medium to interact with various professionals in this field and instill the roots of bioinformatics to the common audience.

LINA FATIMA M & THAMARAI G R

INTRODUCTION TO BIOINFORMATICS

Bioinformatics is the use of computational and analytical techniques to capture and understand biological data, utilizing physics, statistics, computer science, and mathematics.

The term bioinformatics was coined in 1970 by Paulien Hogeweg and Ben Hesper who defined it as "The study of information processes in biotic systems"

Margaret Oakley Dayhoff is considered the mother of bioinformatics, she designed a program in the early 1960s that decoded the correct sequence of a small protein in a matter of minutes, a task that would have taken humans months to accomplish! She further went on to create the first protein database using the combination of computers and biochemistry, thus giving rise to a new field- Bioinformatics.

1951- Pauling and Corey

Proposed the alpha helix and beta pleated structure of DNA

1953- Watson and Crick

Gave the double helix structure of DNA

1955- Frederick Sanger

Determined the structure of insulin molecule and how the amino acids are linked

1965- Margaret O. Dayhoff

Created the first database of protein sequences.

Bioinformatics is a sub-discipline that deals with gathering, storing, analyzing, and disseminating biological information, there are further databases created that help in easy organizing and extraction of this data for future requirements.

Bioinformatics has a variety of interesting applications, some of which are :

1. Designing of drugs for various diseases:
 Helps accelerate target identification, find drug resistance, and characterize side effects. This is more like a boost to classical drugs and medicines that helps enhance healthcare and reduce recovery times.

2. Creation of precision medicine also known as personalized medicine:
 A recent approach to cure a patient is to create custom drugs, it's a relatively expensive route but with technological development, it will be highly accessible and might also be the solution to hereditary diseases. This method requires the patient's genomic data though to make these specific drugs.

3. Decode the evolutionary line:
A very interesting application that not only helps us find the roots of our existence but also helps answer whether the egg came first or the hen! Bioinformatics collects genomic data and compares it with the help of data sharing to classify complex evolutionary pathways.

4. In the making of biofuels:
In times when fossil fuels are drying up and we are not yet technologically ready to harness sustainable energy to its full extent, biofuels play an important part to power our energy needs. Bioinformatics is used to enhance the production of ethanol and enhance the optimization of biofuel production.

5. This gave rise to many more subgroups that help in the specialized decoding of selected molecules:
•Transcriptomics- Study of RNA transcripts
•Proteomics- Study of proteins
•Phenomics- A study that interconnects the phenotype to the data present in the DNA
•Chemoinformatics- Computational analysis of chemical data
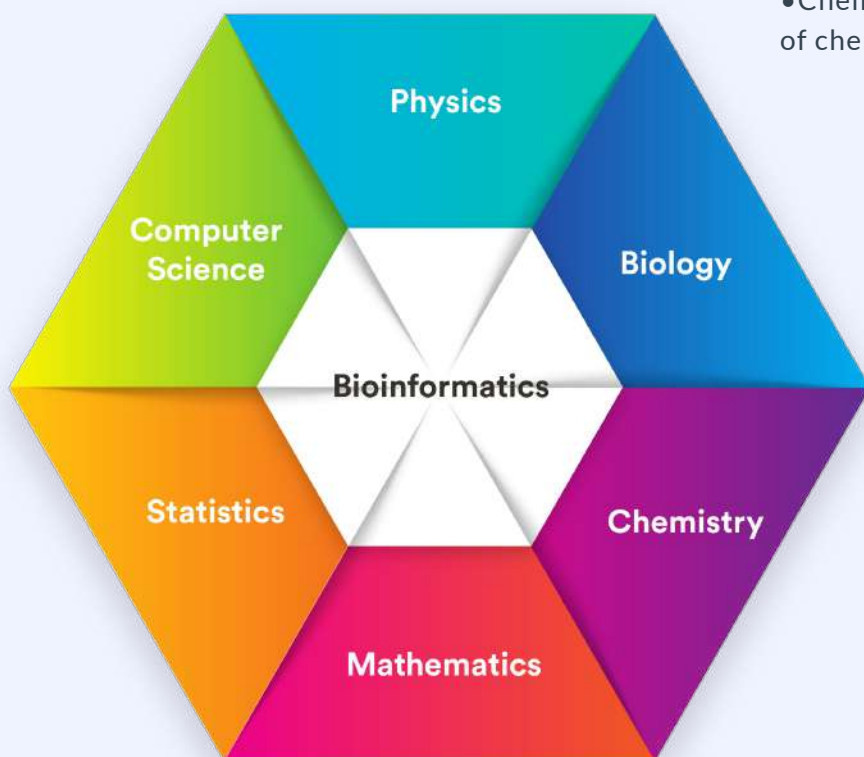
Bioinformatics is a very young field whose applications and fields of research are still being explored, but despite being a novice in the game of science, it has proven to be an extremely useful asset in developing technologies and our way of life as a result. The combination of biology and computers has resulted in an extensively powerful tool that is capable of changing the very essence of existence if utilized properly.

BY MOHAMED ADNAN

GAMES

IDEATED BY ARIVOLI ARAVINDHAN AND PAVAN SAI B
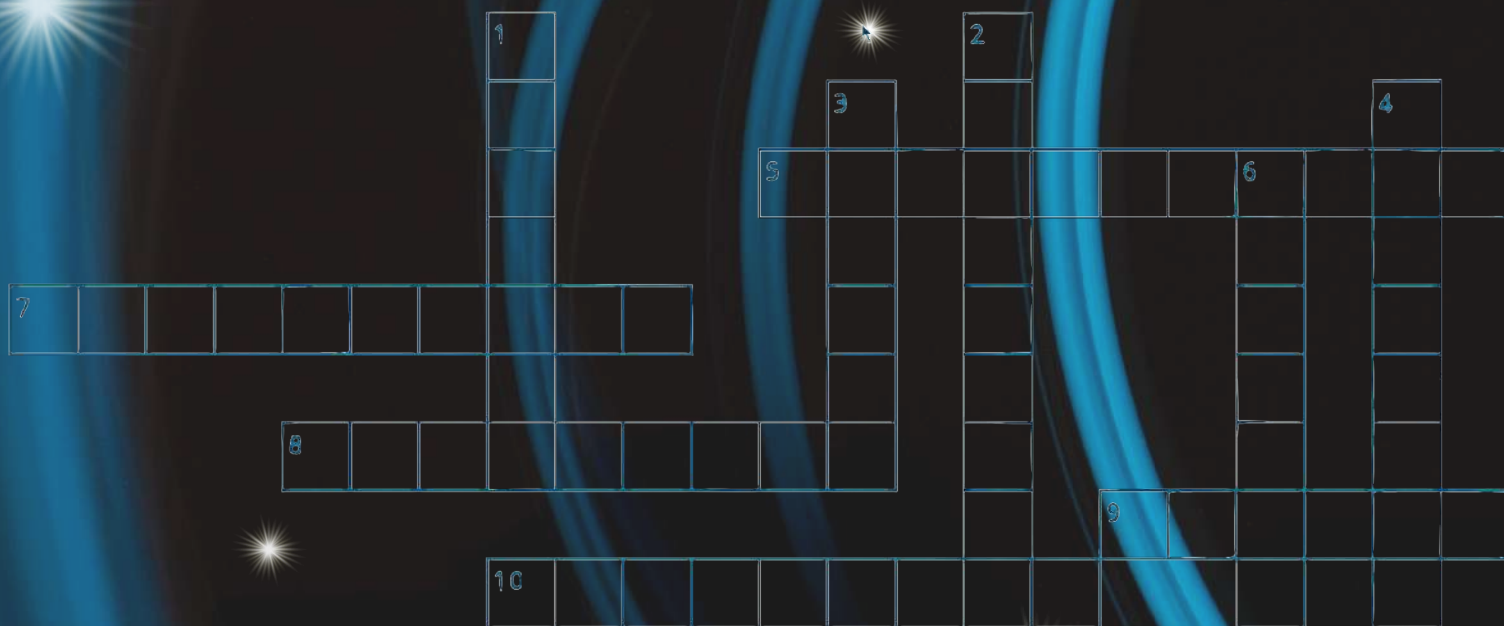
ACROSS

5. Adjective of "genetic recombination"
7. The lesson some of us liked in school level maths
8. The word which states the strategy and solution
9. Founder of 2
10. History of 2

DOWN

1. A well known scene name in the film "interstellar"
2. Growth but for millions of years
3. Add 5th letter of English lexicon in gnome
4. Examination but a cooler word
6. CompLex system/ arrangement/ matrix

Solve the puzzle

A	C	K	A	I	T	S	E	E	C
O	L	E	N	U	Y	E	R	D	L
P	B	G	N	X	Z	S	V	H	U
E	L	A	O	P	G	A	J	M	S
C	G	S	T	R	M	B	K	X	T
H	N	E	A	O	I	A	N	H	E
S	I	Q	T	T	C	T	O	A	R
E	K	U	I	E	R	A	H	S	A
R	R	E	O	O	O	D	T	M	N
O	Z	N	N	M	A	B	Y	R	A
P	Y	C	S	I	R	I	P	C	L
O	T	I	X	C	R	E	O	E	Y
N	P	N	W	S	A	U	I	I	S
A	D	G	S	G	Y	T	B	A	I
N	G	E	N	O	M	E	M	N	S

Solutions on page 18 & page 24

RIDDLES

1. What do you call a cab that provides drug therapy ?

2. How is a dog and a marine biologist alike ?

3. What do you call the leader of the biology gang ?

4. What did one cell say to his sister cell when she stepped on his toe ?

5. What do you call Morgan Freeman when he becomes sentimental ?

6. What do you call a six legged insect that cannot speak ?

7. Why are tertiary structures selfish ?

8. What kind of clothes does a DNA wear ?

9. What do you call a goat which has more than one fur colour ?

10. Why did the recessive gene decide to attend therapy ?

FOOD AND BIOINFORMATICS

Arivoli Aravindhan

"We're changing the world with technology"

Technology arose only a couple of centuries ago but the rate of development has gone from nothing to a very high result in the current world. Education and the will to handle power have made people work harder than ever before. This has been a boon for the current generation where they are having a huge number of advantages compared to their ancestors.

With all of this development happening, the rate of food wastage and population count has been increasing at an all-time high. Concerns of food shortage in the developed countries could be a possibility shortly with the current rate of urbanization happening all around the world (56% in India and 83% in the USA). To start stability of food sources and reduce food wastage, a layout food database should be placed to calculate the food rate functions happening in the real world.

Applications of food databases are

1) Estimating/comparing the nutrient content of foods.
2) Identifying sources of particular nutrients
3) Consumer information and education
4) Recipe and menu development and analysis
5) Completing missing values in a dataset
6) Implementation and monitoring of food legislation
7) Analysing individuals' diets
8) Analysing dietary survey data
9) Assessing how dietary intake affects health and disease outcomes

Such food databases are present below :

FooDB

FooDB is currently the world's largest food database present. It consists of data on food constituents, structural chemistry, and biological facts of each food available. Their website also included data on flavor, color, taste, texture, and aroma. The Alberta Food Metabolome Project, coordinated by David S. Wishart, has merged information with experimentally produced data measured on thousands of chemicals from more than 40 typical food products.

IFCT 2017

Also known as Indian Food Composition Tables. It is more of a manual rather than a digital database where it consisted of most the Indian foods and their nutritional information. Developed by the National Institute of Nutrition Currently, it is posted as a digital manual on their website. It consists of numerical data on individual dietary fiber and nutrients listed.

Food Wiki

Consists of semantic search, match, and inference techniques are used to establish FoodWiki. Customers will be able to match sensitivities and ingredient concepts online using an interface as a web service provided by the system. Before the individual consumes the product, the system assigns an intolerance score based on the implicated substances and food additives in the product on the market shelves.

Foodomics database

The foodomics database enables hypothesis-driven and discovery-based investigations that combine food molecular profiles with data from other high-resolution phenotyping datasets to investigate the many effects of diet on health in a variety of nutritional and non-nutritional perturbations. These techniques will pave the way for nutrition to become an important aspect of precision medicine, allowing the proper food molecules to be given in the right amount, at the right time, and to the right individual.


EuroFIR-BASIS

The EuroFIR eBASIS (Bioactive Substances in Food Information Systems) database for plant-based bioactive substances with potential health benefits is a one-of-a-kind information-based food composition and biological effects database. Over 300 main European plant foods are mentioned, with data taken from peer-reviewed literature, including information on 17 chemical classes (e.g. phytosterols, polyphenols, glucosinolates, and lignans) encompassing various bioactive compound classes and plant foods.
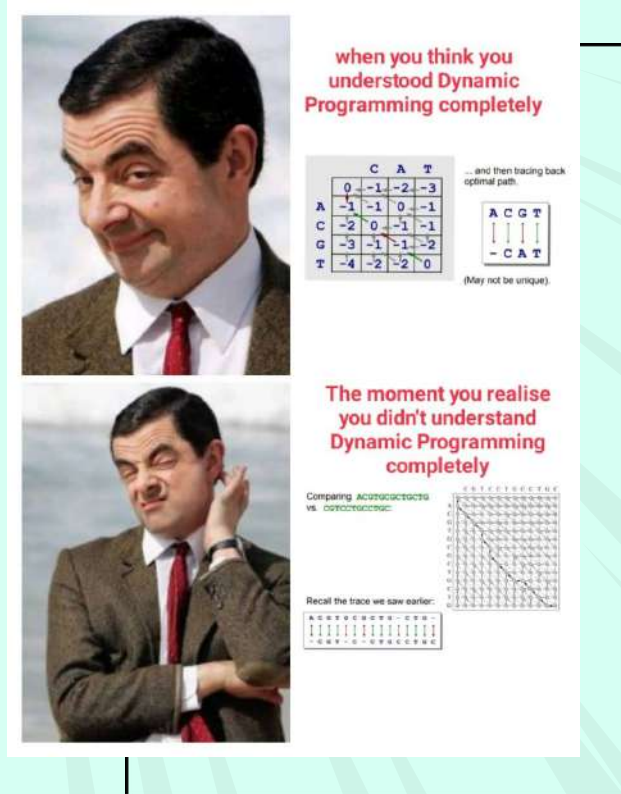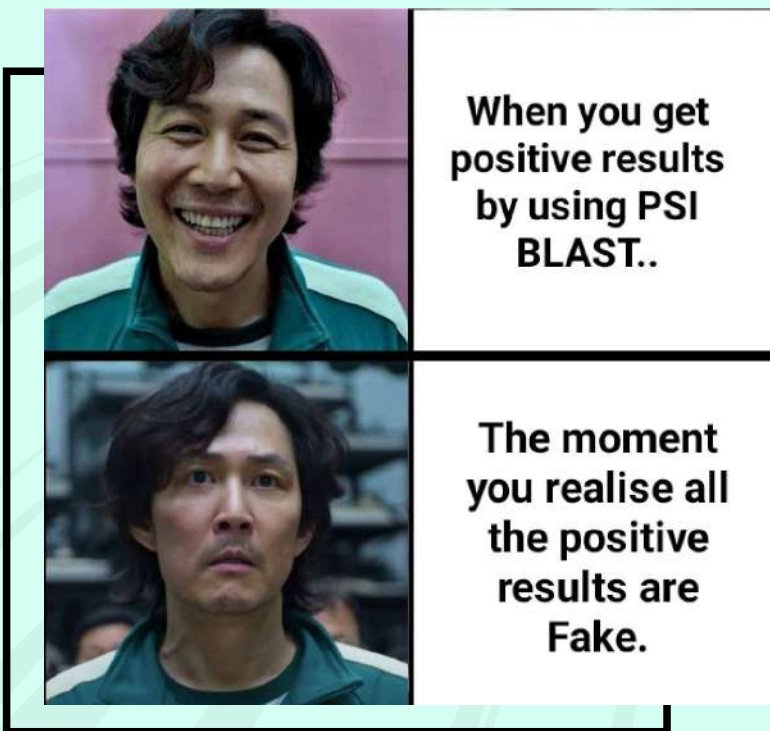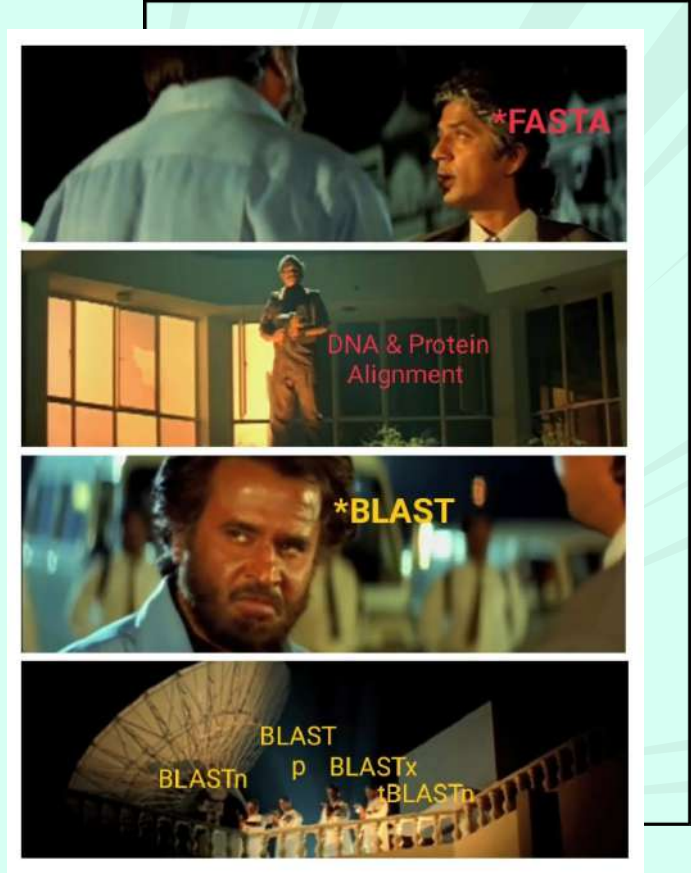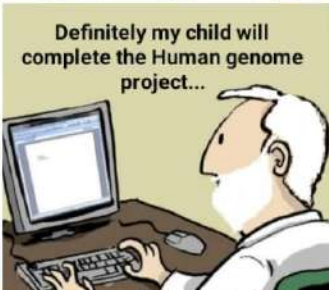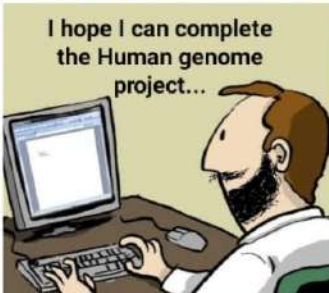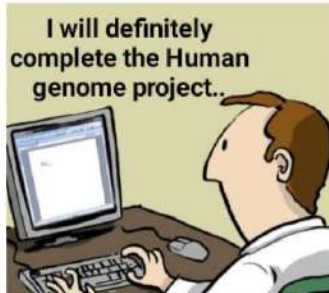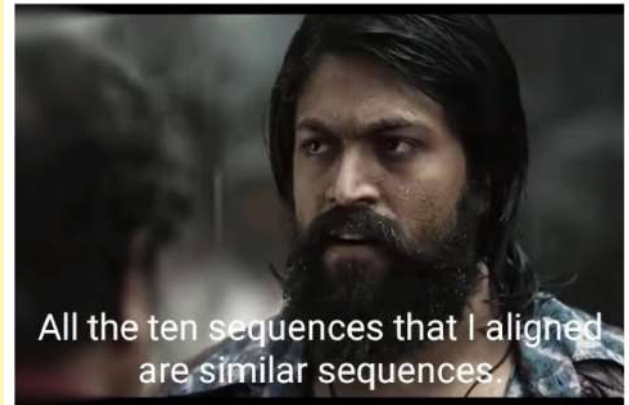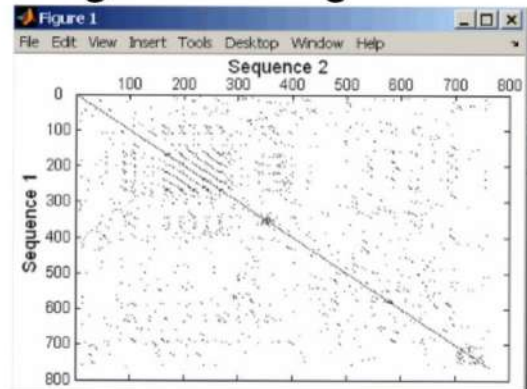
IDEATED BY
SAKTHIVEL & FARAAZ

NEUROINFORMATICS

LINA FATIMA. M

AND

THAMARAI G.R

NEUROINFORMATICS is a supremely transdisciplinary field developed to understand the nervous system in its normal and abnormal state. This domain of bioinformatics broadly encompasses neuromorphic engineering and computational neuroscience including approaches from information systems, computer science, and integrative biology. It seeks to build and preserve web-accessible databases related to experimental and computational neuroscience data by employing ingenious software tools.

Bioinformatics can be comprehensively elucidated as an amalgamation of life sciences and informatics. The escalation of activity in bioinformatics has surpassed the origin of bioinformatics, consequently citing the term *"traditional bioinformatics"*. Traditional bioinformatics is a field that embodies comparing and databasing the DNA, RNA, and proteins as well as modeling the conformation and function of existing and newly designed proteins. The basic data type in traditional bioinformatics is the sequence, thereby making it comparatively simpler than the latest extensions to bioinformatics. Abstractly, traditional bioinformatics consists of sequence-oriented databases together with tools to explore across databases, differentiate either online databases or downloadable software, and progressively model the molecules associated with the sequences. On the other hand, neuroscience data is manifold and heterogeneous. The experimental databases of neuroinformatics house the role and anatomy of genes in the nervous system, images acquired with structural and functional resonance imaging, tissue staining at dimensional scales ranging from subcellular electron microscope images to centimeters of slices through brains of monkeys, optical recordings of voltage and chemically activated dyes.
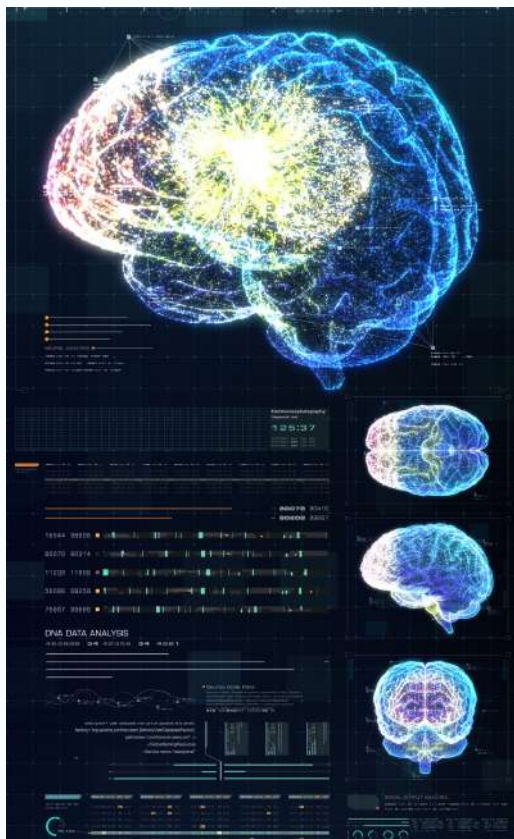
Neuroinformatics acts as an interface between computer science and experimental neuroscience. Computational neuroscience studies the development, composition, physiology, information processing, and cognitive capacity of the nervous system while the archival, retrieval, and consolidation of the substantial amount of generated data through the scientific literature, clinical records, and distinctive databases are achieved by neuroinformatics.

The human brain is composed of 86.1 ± 8.1 billion neurons (Azevedo et al. 2009) and almost an equal number of non-neuronal glial cells; the neurons being interconnected by about a 100 trillion inhibitory and excitatory synapses as well as above-threshold and subthreshold synapses, which in turn constitute large neuronal networks (Goldental et al. 2014). That being cited, computational neuroscience for the last four decades solely aims to investigate the process of information representation and manipulation in the brain employing electrical and chemical signals.

Here are a few examples and research programs are done in neuroinformatics.

In Representative Human Brain Projects (HBP), the group is developing XML protocols that may be used to make data self-descriptive, which means that when data is stored in these XML files, the data is accompanied by a description of the data. Gardner's group has contributed significantly to the conceptual foundation for categorizing all neurophysiology data and is now working on the BrainML transfer standard. The prototype database's main variables are electrical recordings and behavioral event time series; other variables include recording locations, associated publications, brain area and recording sites within that region, kind of neuron, animal species, and so on. They first chose to save data that might be useful to an electrophysiologist colleague who collected electrical recordings from pyramidal cells in somatosensory and associated cortical regions and compared them with videotaped behavioral events during a macaque reaching task. This resulted in a working prototype system that showed how to store annotated physiological data. Their group has lately expanded the extent of web-accessible neurophysiology data. The new information includes the marmoset lateral geniculate nucleus, the macaque primary visual and inferior temporal cortex, and the electrical recordings from retinal ganglion cells.

With 8000 processors, The Blue Brain Project will be operated on an IBM Blue Gene/L prototype supercomputer. One of the most amazing technologies used here is the Blue Gene computer system, the world's fastest supercomputer. The Blue Brain Project's goal is to use comprehensive simulations to better understand mammalian brain function and malfunction. The Blue Brain Project will soon welcome researchers to use Blue Brain Software to create their models of different brain areas in different animals and at various degrees of detail for simulation on Blue Gene. These models will be stored in an Internet database, from which Blue Brain software will extract and link models to create brain regions, allowing the first entire brain simulations to begin!

The main purpose of this initiative is to make it easier to communicate neuroscience data, data-analysis tools, and modeling software. The OECD Working Group on Neuroinformatics has provided its full support to the site. The German Ministry of Science and Education is funding the Neuroinformatics Portal Pilot.

The Neuronal Time Series Analysis(NTSA) Workbench provides a collection of tools, methodologies, and standards for neuroscientists dealing with neural time series data. By creating an information system that will make it easier to save, organize, retrieve, analyze, and share experimental and simulated neuronal data, the purpose of this project is fulfilled. Long-term goals include creating a collection of tools, procedures, and standards to meet the demands of neuroscientists working with neuronal data.

In Japan National Neuroinformatic resource, users can access mathematical models, experimental data, analytic libraries, and other resources through the Neuroinformatics Search Service, or Vision Platform.

The CARMEN project is a multi-site research initiative (11 universities in the UK) that aims to employ GRID computing to allow experimental neuroscientists to preserve their datasets in a framework and make them publically available for future study and usage by modelers and algorithm developers.

Thus, Neuroinformatics is a prime illustration of the rapidly expanding and fascinating increase in bioinformatics knowledge and data. Neuroinformatics databases and tools include a wide range of geographic scales and data kinds that may be stored and examined. Neuroinformatics databases, together with cutting-edge software tools, are critical for studying the nervous system in normal operation, neurological diseases, and therapeutic applications. The short-represented neuroinformatics initiatives provided the larger bioinformatics community with a fascinating sampling of recent advancements.
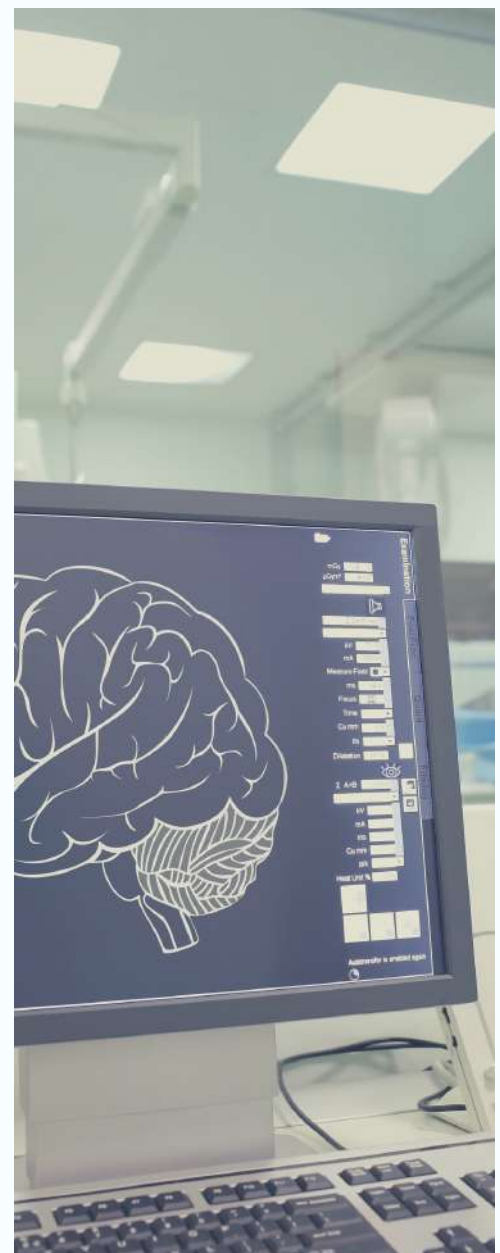
BIOINFO BUZZ

CASSANDRA RIFFLIN C R & SUVETHA P

A PATENT VIEW – BIOINFORMATICS

New patent laws have been introduced to protect biological inventions and changes in bioinformatics. The patent applicants now seek to protect their inventions which include uses of information, software methods of doing business, etc These approaches which are also applicable for traditional biotechnological products are useful in resolving difficulties now faced in prosecuting patent applications.

COMPREHENSIVE DNA SEQUENCE ANALYSIS ON YOUR SMARTPHONE

The algorithm called, Autoglass, is an AI algorithm based on a special deep neural network and steps up from existing statistical methods. AutoClass doesn't make any distributional assumption, hence can effectively correct a wide range of noises or technical variations, identify and remove noises by learning their patterns, and is used to retain the signal by learning their patterns. **The AI teaches itself to differentiate signal vs noise in the data by seeing enough data.** Usually, the more data it sees, the better it performs.

DEEPMP DETECTS DNA BASE MODIFICATIONS ON NANOPORE SEQUENCING DATA

Deep MP, is a conversational neural network-based model that takes information from Nanopore signals and base-calling errors to check the methylation status of a given motif. It has also introduced a threshold-free position modification calling model sensitive to sites methylated at low frequency across cells.

BIG-SLICE MAPS THE DIVERSITY OF 1.2 MILLION BIOSYNTHETIC GENE CLUSTERS

Big slice has opened up new possibilities in Gene mining for **biosynthetic gene clusters (BGCs)** to accelerate natural product discovery. It is also a milestone in constructing a global and searchable interconnected network of BGCs as more genomes are sequenced more information can be gathered to highlight their potentially novel chemistry.

BIOMERCATOR V3: AN UPGRADE OF GENETIC MAP COMPILATION AND QUANTITATIVE TRAIT LOCI META-ANALYSIS ALGORITHMS

An operator is the most advanced QTL analysis software available, offering advanced features such as the QCA tool that can be used to perform a quantitative comparison of alternate breeding schemes, advanced mapping and analysis tasks such as genome-wide association studies, gene, and SNP discovery and mapping, and genome-wide mapping of genetic marker and marker gene associations. The software is now widely used by agronomists, breeders, researchers, and companies to identify genes underlying their quantitative traits. The software has been translated into more than a dozen languages, enabling users to access and use the software in their language.
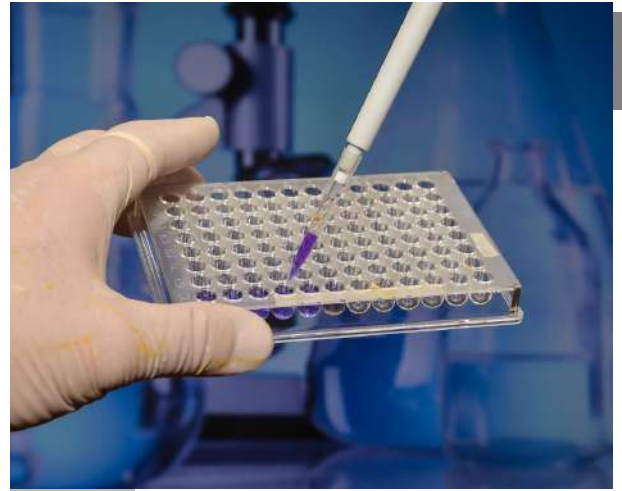

DE NOVO ASSEMBLY OF THE CATTLE REFERENCE GENOME WITH SINGLE-MOLECULE SEQUENCING

ARS-UCD the new reference genome for cattle which is based on the same animal as the original is used for easy transfer and interpretation of results. A combination of modern technology is used to increase continuity, accuracy, and completeness. The increased assembly accuracy will benefit future research on this species.

ACQUIRING REAL-TIME, ACCURATE, AND REPRODUCIBLE 4D BIOINFORMATICS DATA WITH PASER (PARALLEL SEARCH ENGINE IN REAL-TIME)

PASER is available to the market today and is powering applications across the life sciences, pharmaceutical, healthcare, and consumer sectors. The PASER technology will enhance the capabilities of the Tim stop mass spectrometer and the higher level of approach also enables other high-level functions like
de-duplication, which further reduces the amount of data the user needs to shift through.

NEW BIOINFORMATICS PLATFORM PREDICTS OPTIMAL TREATMENT COMBINATIONS FOR CANCER PATIENTS
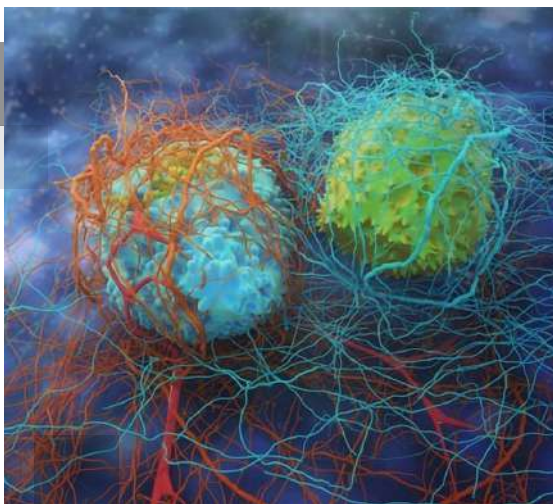
The Recurrent Features Leveraged for Combination Therapy (REFLECT) algorithm can be used to identify other immune response markers that could be targeted by combination therapy and also identified alterations that commonly coincide with response to traditional cancer drugs, such as defects in the regulation of gene expression, changes in the levels of proteins and cancer-related genetic mutations. This could enable the design of combination therapies that are more potent than single agents, with reduced side effects. The findings could also be used to improve the design of single-agent therapies. The REFLECT platform has the potential to improve the efficacy of combination therapy for cancer patients.

Written by
Jaswanthika S
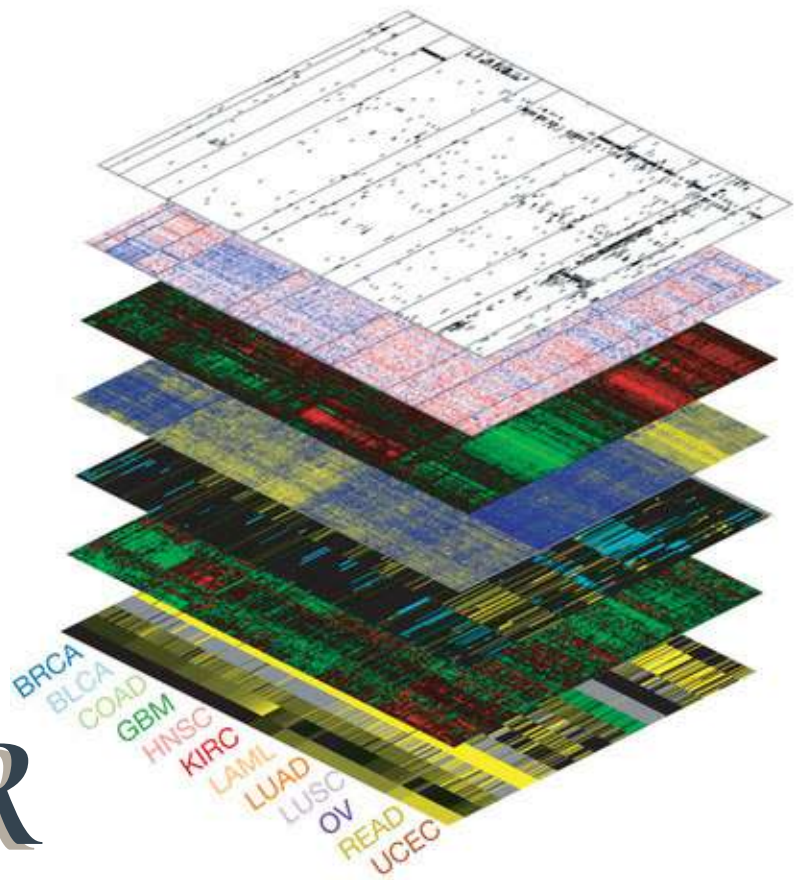
CANCER BIOINFORMATICS

Cancer is on the list of diseases through which people lose their life.

Poor diagnosis, therapeutics, and prognosis are due to differences in duration, complications, localization, susceptibility, cellular differentiation, and resistance to the drug. The new notion in systemic clinical medicine is consolidated with systems biology, clinical science, omics-based technology, bioinformatics, and computational biology to improve disease prognosis, therapeutics, and diagnosis. **Cancer bioinformatics** is a combination of clinical computer technology, computational biology, information technology, arithmetic, and omics technology.

Cancer Bioinformatics allows us to use bioinformatics tools for signaling, communication, and dissemination. To find answers to questions about cancer, we need to develop a progressive bioinformatics tool to find the answer to our queries.

The multivariate assay is a process that can distinguish the errors that occur in the results. miRTrail is a tool to interpret the interactions between genes and miRNAs.

We can study the morphology of the cancer masses with the help of bioinformatics tools. This is done by selecting any desired extract and amplifying this by contrast modification.

Biomarkers-Strategies :

Cancer bioinformatics is used to identify new markers in treating cancer. That could help in its detection in the early stages, monitor a patient's development, and track its response. System clinical medicine is used in the process of generating biomarkers and it is a consolidation of high throughput technology, bioinformatics, and computational science in improving the diagnosis, therapies, and prognosis of the disease. Network biomarkers are used with the knowledge of protein annotation, association, and signaling pathways. Dynamic network biomarkers are used in correlating the patient's information. It is essential to know the association between clinical informatics and bioinformatics.

*Prediction and monitoring of precised medicine :*
System cancer medicine is now recently used in the strategy of P4 ways – i.e predictive, preventive, personal.

Disease perturbed molecular network indicate the abnormality and functioning which has to be carried out by P4 in a cancer drug. Ren and his teammates found an algorithm known as the **Optimization Tool for Clustering and Classification**. This is used for various types of measurements like proteomics and next-generation sequencing data type.

Breast Cancer:

Among the various types of cancer, breast cancer is the one that causes a lot of death. This digit is now depleting remarkably due to the development of therapeutic drugs. Breast carcinoma is divided into three subtypes based on the enrichment of three genes namely – luminal, human epidermal growth factor receptor 2- positive, and ER-negative and basal subtypes.

Tucatinib, trastuzumab, neratinib, and trastuzumab emtansine are some of the drugs used as molecular targeted drugs. For the anti-angiogenic factor, we are employing the vascular endothelial growth factor. Androgen receptor-targeted therapies are showing assuring results in the treatment of breast cancer. Because of the tumor heterogeneity, there is an increasing demand to find a biomarker for the treatment of breast carcinoma.
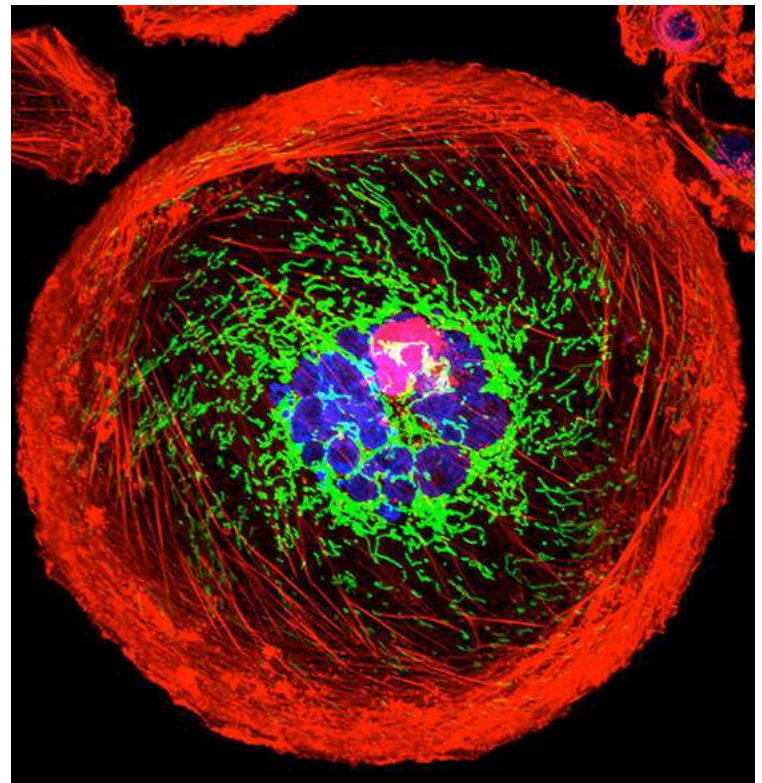
A polyploid giant cancer cell (PGCC) from triple-negative breast cancer.
Credit: National Cancer Institute

Methods :

- **Subject and gene information:**

The GEO database contains genetic information. Platform GPL17586 contains 74 breast carcinoma samples and 12 non-carcinoma samples. Platform GPL570 contains 29 breast carcinoma samples and 4 non-carcinoma samples. GSE42568 contains 104 breast carcinoma samples and 17
 non-carcinoma samples.

- **Data analysis:**

We detect DEG in cancer tissue by GEO2R screening criteria – adj. P<0.05, log2 FC >1.5.

- **Gene ontology and KEGG pathway analysis:**

Gene ontology is used to annotate and analyze the biology process in which they are involved. KEGG database is used to study the relevant signaling pathway.

- **PPI network visualization:**

STRING, an examining tool, is used to construct a PPI network for identified DEGs. Cytoscape software is used to establish the interaction network map and molecular interaction network.

- **Kaplan-Meier survival and redundancy analysis of DEGs:**

Kaplan-Meier plot is developed to calculate the effects of transcripts on cancer patients. To calculate this, we will collect data from the following databases – GEO, European Genome-phenome archive, and the cancer genome atlas. From the result, we can get the survival analysis of the patient and the False Discovery rate was calculated.

Cancer bioinformatics is used to track and estimate the efficiency and effectiveness of precision medicine. This is used for providing a safe and précised medicine for cancer treatment. With the growing trends and evolution in cancer bioinformatics, we can also develop a new drug for cancer treatment. To continue this, we have to integrate the wide data amount for the betterment of human life.

Crossword

ANSWER KEY

ACROSS

5. recombinant
7. statistics
8. programme
9. darwin
10. phlogeny

DOWN

1. docking
2. evolution
3. genome
4. analysis
6. network

Facts

The first protein library was created by a woman named Margaret Oakley Dayhoff who was an American physical chemist and a pioneer in the field of bioinformatics.

The heuristic algorithm of BLAST locates all the similar three-letter words between the sequence of interest.

Errors in databases can be due to sequencing error, Cloning vector contamination, redundancy of data, human error, etc.

In the early 1960s computers were one the tools employed to decode protein sequences.

The coronavirus pandemic in 2020 shows that rapid data analysis and interpretation is a powerful tool to help control the spread when particular data is shared quickly and widely.

Scientists found it hard to find certain genes in bird genome because of a varying GC: AT ratio

Biologists can construct family trees or phylogenies.

Myths

Bioinformatics is a subject that is completely concerned with prediction.

Basically, bioinformatics is practically free of cost.

DNA of a person can be used at crime scenes and to make clones.

Your specific code can be used to make a better software

CRISPR is a tool that can be used to revive dinosaurs

Chromosomal Microarray analysis gives confusing results.

Bioinformatics is completely in the domain of computers and is all about pushing buttons.

FROM THE PROFESSORS DESK

. . . .
. . . . .
. . . . .
. . . .

Q: If we could use your experience in bioinformatics as a bridge to better comprehend it.

Bioinformatics help to understand and comprehend biological data. For example, if you wish to compare the current Covid genome in circulation with previous strains and even other related viruses, you can't do the entire thing in wet lab, but, separately wet lab data has been submitted in the database and you can access it. Use your knowledge and available software to conclude your project.

Q: How crucial is it for a bioinformatician to be familiar with wet lab skills/techniques in order to conduct an effective experiment?

Any in silico work requires wet lab data. When you are collecting large amount of wet lab data on a specific experiment and wish to analyze it. Every experiment has many parameters. So, unless you have a clear concept of that particular experiment, you cannot analyze the obtained dataset to embark on a relevant conclusion. So, my suggestion is, learn as many wet lab techniques as possible during your course so that you develop a clear idea. It may help you to design bioinformatics problems also.

Q: Many students are perplexed by the difference between bioinformatics and computational biology; Could you please provide a brief explanation, sir ?

Bioinformatics is an interdisciplinary field that encompasses biology, information technology and sometimes big data. Primarily, it's an application of information technology to biology to collect, curate, analyze and publish biologically relevant output. Computational biology is a little different and I prefer to call it a subset of bioinformatics. CB deals with a specific set of problems with the help of mathematical and statistical knowledge. Few things are overlapping and few things are distinct. For example, database creation is a bioinformatics project, whereas simulation or docking can be termed as a CB project.

Q: Based on your job experience, how would you define bioinformatics?

I have already defined it. If you ask about the career aspect, as many companies are moving towards biological data analysis and data management and often it's big data, so, definitely bioinformatics has a good prospect. But learning any relevant computer language properly will strengthen your position in the market.

Q: What are some of the skills and knowledge required to become a bioinformatician?

Subject knowledge is very important and different bioinformatics projects demand different subject knowledge, e.g., transcription factor identification project requires molecular biology knowledge, whereas docking requires good knowledge of chemistry. As I said earlier, learning at least one computer language is very crucial.

Q: What are some future scopes of bioinformatics and their significance sir ?

In a data-driven market, the future scope of bioinformatics is definitely high. But, it requires a few additional skills like learning Java, R or Python properly is something biotechnology students often ignore. I don't think only the knowledge of database access will help the students to find a satisfactory job in the market.

Q: What are some simple projects you would recommend for beginners to start with?

During COVID, nearly 90% of biotechnology graduates worked in bioinformatics. So, it is hard to find beginner's projects. But, if you read the theory carefully, you may ask some questions to yourself and as there are zillions of specialized websites and tools present, you can develop a small project. In bioinformatics, concluding your result and finding out biological relevance is also very important.

Q: Any books or courses you would recommend to have a deeper understanding of the subject?

There are many books in the market. And, in order to keep the novelty, the writer includes a few changes in the book. So choosing the book is entirely your and your faculty's choice. However, I prefer books written by *Baxevanis or David Mountor Xiong.*

In an interview with Dr. Sangeetha,

Q: Many students are perplexed by the difference between bioinformatics and computational biology; can you please provide a brief explanation, ma'am?

Bioinformatics and computational biology both involve the best use of computers for biology. Computational biology largely focuses on the mathematical and computational models and methods to resolve the mysteries of life science, whereas bioinformatics is an umbrella covering biological databases to analyse and provide a meaningful interpretation of the biological data.

Q: If we could leverage your experience in bioinformatics as a bridge to better understand it, that would be great.

I believe that no bioscience research is complete without bioinformatics.

Q: Based on your job experience, how would you define bioinformatics?

Bioinformatics is "the information that you can gain from biological data collection, and analysis through mathematics, computers, physical sciences, etc. "

Q: What are some of the skills and knowledge required to become a bioinformatician?

Start by learning R and Python to excel in the field of bioinformatics.

Q: What are some simple projects you would recommend for beginners to start with?

There are several projects to choose from, ranging from sequence analysis to molecular modeling.
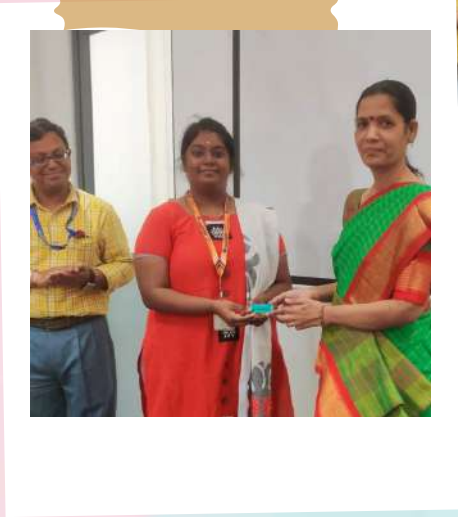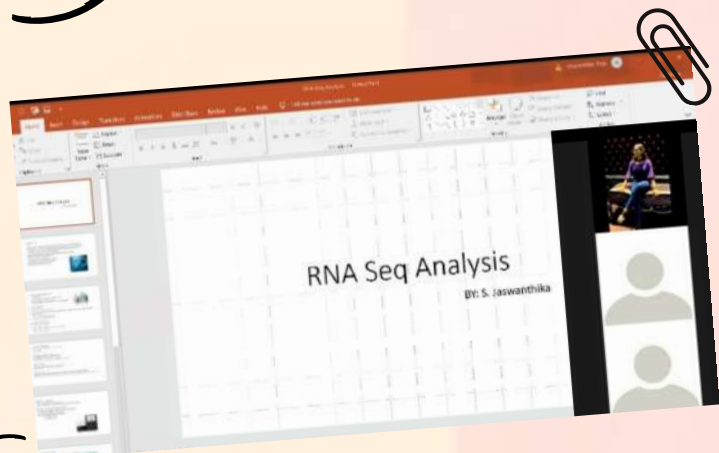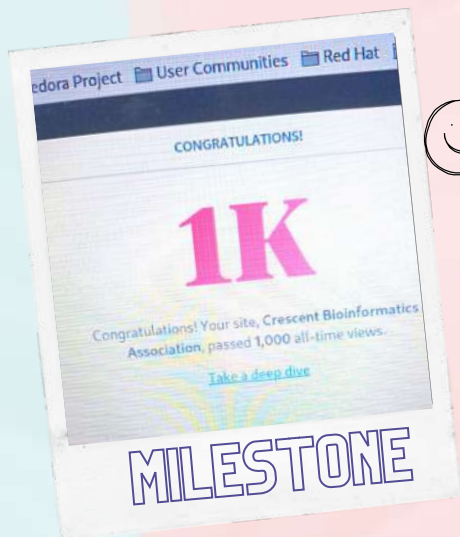
By
Tasneem Juzer

PHOTO GALLERY

Jan 9, 2022

RNA Seq Analysis
BY: S. Jaswanthika

Journal club

MILESTONE

CBA OFFICE BEARERS

ANSWER KEY

Riddles

1. Chemotaxi
2. One wags a tail other tags a whale
3. Nucleus
4. Mitosis (my toe sis)
5. Centimorgan
6. A mutant (mute ant)
7. Because they are wrapped up in themselves
8. Genes
9. Heterozygote (Heterozygoat)
10. It wanted to know how to express itself

Puzzle

ANNOTATIONS
SEQUENCING
NANOPORE
GENOME
BIOPYTHON
DATABASES
CLUSTER ANALYSIS
PROTEOMICS
MICROARRAY

MEET OUR TEAM

EDITOR IN CHIEF
Tasneem Juzer

COMPILER & EDITOR
Lina Fatima M

DESIGNERS

Cassandra Rifflin C.R

Ronisha S

Thamarai G.R

Under the Guidance of
Dr. Hemalatha S (Dean School of Life Sciences)
Dr. A.Baskaran (Faculty Coordinator)

OFFICE BEARERS 2021-2022

Co-Presidents: Tasneem Juzer,Teena Ra

Student Co-ordinator: Mohamed Juvad

Junior Student Co-ordinator: Lina Fatima M

Operations Head: Rifat Aara

Technical Head: Thamarai G.R

Operations Executives:

Annish Shabhiya

Chandhini

Gayathri

Noorliza Fatimah

Sandhya Devi

Prathima Lakshmi

CRESCENT BIOINFORMATICS ASSOCIATION