PRATYARTH'18

Symposium on Data Analytics in Bioinformatics

Organized by

Computational Biology Group @ SASTRA,

SASTRA Deemed University

February 3, 2018

SCHEDULE

09.30-09.45	**Inauguration**
09.45-10.30	**Compilation, curation and exploration of the phytochemical space of Indian medicinal plants** ***Dr Areejit Samal*** *The Institute of Mathematical Sciences, Chennai, India*
10.30-11.15	**Learning and Predicting Novel Metabolic Pathways through Subgraph Mining** ***Dr Karthik Raman*** *Department of Biotechnology,* *Indian Institute of Technology Madras, India*
11.15-11.30	**Tea Break**
11.30-12.15	**Global recombination landscape in human: Insights from an in-depth analysis of the 1000 genomes project** ***Dr Thiyagarajan S*** *Institute of Bioinformatics and Applied Biotechnology, Bangalore*
12.15-13.30	**Lunch Break**
13.30-14.30	**Poster presentation**
14.30-15.15	**Computational Techniques in Data Integration and Big Data Analytics for Bioinformatics** ***Dr Subramaniya Swami V*** *School of Computing,SASTRA Deemed University*
15.15-16.15	**Oral presentation**
16.15-16.30	**Awards Presentation & Valedictory Function**

INDEX

PP 2	Structure Based Design of Novel Therapeutic Agents for Breast Cancer Targets	Mathangi S SivakamaMeenakshi P ThamaraiSelvi S Saraboji K	Poster	16
PP 3	Anti-NHSL property of 13-octadecenal: a study on their effects in diabetic foot ulcers	Rajivee J Sabithadevi N Gowdhaman D	Poster	17
PP 4	Studies on molecular genetic analysis of HLA predisposition to HPV infection	Aarthi S Varsha R Arunkumar G	Poster	18
PP 5	In silico analysis of promoter nucleosome features targeted by SWI/SNF	AlaguSankareswaran S ArunGoutham R N Suma Mohan S	Poster	19
PP 6	Study of conformational flexibility of loops in protein-ligand complexes	Faiz.S Vignesh.P Vigneshwar R	Poster	20
PP 7	Drug resistance mechanism of Pyrazinamidase in Mycobacterium tuberculosis: Insights from molecular dynamics	Janani S M Dharani N Thamotharan S	Poster	21
PP 8	Analysis of binding sites and prediction of glycation sites in HAS	Harini M Saraswathi N T	Poster	22
PP 9	Cross-linking guided molecular modeling of SWI/SNF core sub-unit assembly	Nivedha T Swetha M Suma Mohan S	Poster	23
PP 10	**Investigation oF (S)-ß-(carbobenzoxyamino)-γ-butyrolactone as a novel quorum sensing quencher of uro pathogenic *escherichia coli***	Mahesh V G ShyamSundar S Karthi Shanmugam	Poster	24
PP 11	Analysis of genetic and epigenetic changes associated with cardiomyopathy	MuthuMeenakshi M Hemallatha Y Alex Stanley B	Poster	25
PP 12	Analysis of gene expression pathways in Wheat seed development	Nivedhithaa J G Vishnupriya S S Sathya A	Poster	26

PP 13	Targeting drug resistance mechanisms in leishmaniasisntd	Dhakshayini J Logachandar R Ashok Palaniappan	Poster	27
PP 14	Dynameomic approach to study protein-DNA interactions	Supraja V Mohan Yashwanth Kumar K Udaykumar M	Poster	28
PP 15	Molecular visualization tool for dimer generation using crystallographic information file (CIF)	Abishek M Swetha G Udaykumar M	Poster	29
PP 16	Stage specific analysis of Liver cancer gene expression	Vignesh R Varsha K Ashok Palaniappan	Poster	30
PP 17	Modelling studies on Cu/Zn SOD1 from Wuchereriabancrofti and structure based inhibitor designing against Filariasis	Abinaya Lakshmi G Saraboji K	Poster	31
PP 18	Conformational Analysis of Some *para*-Substituted Xanthene analogs	Kavitha M Hemarajalakshmi Thamotharan S	Poster	32
PP 19	Analysis and prediction of glycation sites of human hemoglobin	PriyaDharshini G Surya R Saraswathi N T	Poster	33
PP 20	Genetic and epigenetic changes in liver disease	Priyatharsshini S P Shweta S Vigneshwar R	Poster	34
PP 21	Compositional vector based distance calculation for two protein sequence	Mohana Krishnan Sankara Subramanian M Arunachalam J	Poster	35
PP 22	Insilico structural studies on redox enzymes of Wuchereriabancrofti	Rakshandha V Saraboji K	Poster	36
PP 23	Automated genomic signal processing for genes associated with cancer	Ramya.K Uday Kumar M	Poster	37

PP 24	Identification of potential inhibitor for bromodomain of brahma related gene 1 (BRG1)	Sindhuja A Sivaramakrishnan	Poster	38
PP 25	An Exploratory Analysis of Gene Expression Data in Human Viral Hepatitis	Sureshan M Arunachalam J	Poster	39
PP 26	Conformational flexibility of the Cu/Zn SOD1 induced by the point mutation at glu100	BhalekarShraddha LekhaSundararajan Saraboji K	Poster	40

Invited Lectures

IL 1

COMPILATION, CURATION AND EXPLORATION OF THE PHYTOCHEMICAL SPACE OF INDIAN MEDICINAL PLANTS

Areejit Samal

The Institute of Mathematical Sciences (IMSc), Chennai, India.

E-mail:asamal@imsc.res.in

Phytochemical constituents of medicinal plants encompass a diverse chemical space for drug discovery. India is rich with a flora of indigenous medicinal plants that have been used for centuries in traditional Indian medicine to treat human maladies. A comprehensive online resource on the phytochemistry of Indian medicinal plants will enable computational approaches towards natural product-based drug discovery. In this talk, I will present, IMPPAT, the most-comprehensive, freely available, curated, digital database of 1742 Indian Medicinal Plants, 9596 Phytochemicals, And 1124 Therapeutic uses spanning 27074 plant-phytochemical associations and 11514 plant-therapeutic associations. Notably, our curation effort led to a non-redundant in silico library of 9596 phytochemicals with standard chemical identifiers and structure information. Using cheminformatic approaches, we have computed the physicochemical, ADMET and drug-likeliness properties of the IMPPAT phytochemicals. We show that the structural complexity of IMPPAT phytochemicals differ from libraries of commercial compounds or diversity-oriented synthesis compounds while being similar to other libraries of natural products. Within IMPPAT, we have filtered a subset of 960 potential druggable phytochemicals, of which majority have no significant similarity to existing FDA approved drugs, and thus, rendering them as good candidates for prospective drugs.

IL 2

LEARNING AND PREDICTING NOVEL METABOLIC PATHWAYS THROUGH SUBGRAPH MINING

Karthik Raman

Department of Biotechnology, Indian Institute of Technology Madras,India

E-mail: kraman@iitm·ac·in

The ability to predict pathways for biosynthesis of metabolites is very important in metabolic engineering. It is possible to mine the repertoire of biochemical transformations from reaction databases, and apply the knowledge to predict reactions to synthesise new molecules. However, this usually involves a careful understanding of the mechanism and the knowledge of the exact bonds being created and broken. There is clearly a need for a method to rapidly predict reactions for synthesising new molecules, which relies only on the structures of the molecules, without demanding additional information such as thermodynamics or hand-curated reactant mapping, which are often hard to obtain accurately. In this talk, I will describe a robust method based on graph mining, to predict a series of biochemical transformations, which can convert between two (even previously unseen) molecules. We mine the reaction database and store reaction centres and signatures in a reaction rule network. Such a novel representation enables us to rapidly predict pathways. We also propose a heuristic that predominantly recovers natural biosynthetic pathways from amongst hundreds of alternatives, through a directed search of the reaction rule network, enabling us to provide a reliable ranking of pathways. Our approach scales well, even to databases with>100,000 reactions.

IL 3

GLOBAL RECOMBINATION LANDSCAPE IN HUMAN: INSIGHTS FROM AN IN-DEPTH ANALYSIS OF THE 1000 GENOMES PROJECT

Thiyagarajan S

Institute of Bioinformatics and Applied Biotechnology

E-mail:sthiyaga@ibab.ac.in

Meiotic recombination shuffles different alleles along the chromosomes generating the genetic diversity in the offspring. This process is vital for the adaptation and evolution of a species. Study of the plasticity of recombination rates and presence of hotspots of recombination in the human genome has been of interest lately. An in-depth analysis of the 1000 Genomes Project (1KGP) has been made to study the fine-scale variation in the autosomal recombination rates across 20 human populations to uncover the global recombination landscape in human. The results showed more than 80K hotspots and coldspots which are subsequently validated using a high-resolution pedigree based genetic map. Our analysis yielded clusters of continental groups, reflecting their shared ancestry and genetic similarities in the recombination rates that are linked to their migratory and evolutionary histories. Genomic locations and strengths of hotspots and coldspots across all the populations studied have also been generated, the details of which will be presented.

IL 4

COMPUTATIONAL TECHNIQUES IN DATA INTEGRATION AND BIG DATA ANALYTICS FOR BIOINFORMATICS

Subramaniya Swami V

School of Computing, SASTRA Deemed University, Thanjavur-613401

Email:swamy@cse.sastra.edus

In the era of precision medicine, enormous amounts of data are being generated from disparate sources, including omics, imaging, sensing and beyond. Today, computational scientists need to develop better tools to manage, integrate and share data to make it clinically actionable. Biomedical data has grown too large for most research groups to host and analyze the data from large projects themselves. In the recent years, parallel, incremental, and multi-view machine learning algorithms have been developed to address the research challenges in bioinformatics. Similarly, graph-based architectures and in-memory big data tools have been developed to minimize I/O cost and optimize iterative processing. Data commons provide an alternative by co-locating data, storage and computing resources with commonly used software services, applications and tools for analyzing, harmonizing and sharing data to create an interoperable resource for the research community.

Oral Presentations

OP 1

UNDERSTANDING THE MOLECULAR MECHANISM OF FREE RADICAL SCAVENGING BY FISETIN-A COMPUTATIONAL STUDY

Aathithya D*, Karthi S

Department of Bioinformatics, SASTRA Deemed University, Thanjavur-613401

Email: aathithyadiaz@gmail.com

Oxidative stress results from an imbalance between the production of reactive oxygen species (ROS) and free radical scavenging. Antioxidants are substances that transfer its proton to the ROS to circumvent the effects of free radicals. Natural and synthetic polyphenols are known for its antioxidant activity and are used in the prevention of cancer and cardiovascular diseases. Fisetin is a natural polyphenol that possesses antioxidant and free radical scavenging activity. Many studies reported that the free radical scavenging potential of polyphenolic compounds were well correlated with the quantum mechanical descriptors (QMD). However the molecular mechanism of free radical scavenging activity of fisetin and its correlation with QMD were not well studied. In this present work we performed quantum mechanical calculations using Gaussian 09W to understand the free radical scavenging activity of fisetin. We performed density functional theory (DFT) calculations to determine the QMD associated with the scavenging activity. Our major findings are (i) Planarity of the molecule influences the antioxidant property fisetin (ii) Bond dissociation energy descriptor determines the site of hydrogen abstraction. Our findings will be helpful for the medicinal chemist for further lead optimization processes.

OP 2

IN SILICO SCREENING OF CANCER-ASSOCIATED MUTATIONS IN THE HSA DOMAIN OF BRG1 AND ITS ROLE IN AFFECTING THE ARP-HSA SUBCOMPLEX OF SWI/SNF

Arun Goutham R. N*, Alagu Sankareswaran S*, Suma Mohan S

*Department of Bioinformatics, SASTRA Deemed University,Thanjavu*r - 613401

Email: sankareswaransubramanian@gmail.com, gouthamarun162@gmail.com, sumamohan@scbt.sastra.edu.

The family of Chromatin remodeling complexes facilitates the access of condensed genomic DNA during transcription, replication, and repair, by altering the histone-DNA contacts in the nucleosome structures. SWI/SNF (SWItch/Sucrose Non-Fermentable) family of ATP dependent chromatin remodeling complexes have been documented for their tumour suppressor function. Based on the recent cancer genome sequencing studies, 20% of the cancer causing mutations were identified in this family of remodeling complex which is close to the mutational frequency of p53. The functional characterization of mutations in the SWI/SNF subunits is important for understanding their role in tumorigenesis and identifying potential therapeutic strategies. In an effort to study the human actin related complex and the cancer causing mutations associated with it, we modelled the structure of theß-actin-BAF53A-HSA complex based on the yeast Arp-HSA complex (PDB id: 4I6M). The COSMIC mutations found in the HSA domain of BRG1 were subjected to analysis using 'predict SNP' program. The cancer-associated mutations in the HSA domain, R513Q, K502N, R521P, Y489C, R466H were subjected to MD simulation studies using DESMOND program. Analysis of the trajectory indicates destabilizing effect of the mutations in the ARP-HSA sub-complex which would have an impact in the SWI/SNF complex assembly and remodeling function.

OP 3

MOLECULAR PHYLOGENY USING COMPOSITION VECTORS

Lakshmi Seshadri*,Varsha R, Arunachalam J

*Department of Bioinformatics, SASTRA Deemed University,Thanjavu*r - 613401

Email: varsharaman8@gmail.com /lakshsesh97@gmail.com

The current methods used for reconstructing a molecular phylogeny tree are Neighbour joining, UPGMA, maximum parsimony & Maximum Likelihood Methods. These methods quite often produce varied results depending on the data and the tree building parameters used. Composition vector method is an alignment free method that can be used to overcome the disadvantages faced with alignment based methods while constructing phylogeny: such as the need for homologs, the difficulty in aligning when the length of the sequences are not the same and it is also computationally expensive especially for a big dataset. Using this method, the phylogentic tree for a large number of sequences can be constructed in a time efficient manner. By adjusting the parameters like the weights of the vectors, the k-mer length etc.,the tree reconstruction procedure can be further optimized.

OP 4

CAN AGGREGATION RESISTANCE PATTERNS DESIGN THERMALLY STABLE PROTEINS?

Sowmya R K*and Saraboji K

Department of Bioinformatics, SASTRA DeemedUniversity, Thanjavur-613401

Email:krish.krishnan67@gmail.com

Thermophilic organisms can produce proteins with extreme thermal stability and the stability of thermophilic proteins has been elucidated with various experiments as well as using computational methods. It has been reported that increase in number of ion pairs and side chain–side chain packing, hydrogen bond forming capability, aromatic clusters, electrostatic interactions and hydrophobic residues at protein surface, and main-chain hydrophobic free energy in thermophilic proteins can enhance their thermodynamic stability.Interestingly, based on thermal stability studies, we know that the increased thermodynamic stability of thermophilic proteins may not necessarily equip them with lower susceptibility toward aggregation. So here we aimed to explore the structural conservation of aggregation sites and discriminate the strategies used by thermophiles and mesophiles to resist aggregation. In this present study, we compare the aggregation-resistance mechanism adapted by the thermophilic proteins and its mesophilic pairs using a dataset of 373 protein families. We have used Aggrescan3D (A3D) software to compute aggregation propensity using a score calculated by combining the prediction of 11 algorithms based on the Structure of thermophilic proteins. A3D uses a structure-based approach, allows the specific detection of those spatially-adjacent aggregation-prone amino acids that are relevant to protein aggregation from folded states. Our results show that mesophiles have most of the APRs in Exposed regions and hence, can interact with the polar environment leading to Aggregation. These observations suggest that, thermophiles stow away their APRs in the core of the protein and gain thermostability by protecting them from exposure to solvent. Our results show that, though thermophiles and mesophiles have similar number of APRs, when the size of APRs were analyzed, longer APRs of 8-10 residues were found in greater numbers in mesophiles than thermophiles indicating that, thermophiles gain further thermostability by reducing longer APRs. The results of the study and future perspectives will be discussed during the presentation.

OP 5

STRUCTURAL BIOINFORMATICS STUDIES OF *WOLBACHIA* SURFACE PROTEIN *AND* FERROCHELATASE OF *BRUGIA MALAYI* – PROMISING ANTI-FILARIAL DRUG TARGETS

Reethu Anand, Harini R, Sampath N*

School of Chemical and Biotechnology, SASTRA Deemed University, Thanjavur-613401

Email: sams76@gmail.com; sampath@scbt.sastra.edu

Wolbachia is a common and the obligate intracellular α-proteobacteria, which are found in insects, mites, and arthropods species. The parasitic nematodes cause lymphatic filarial disease in human. The current treatment approach for lymphatic filariasis is chemotherapy with high dosage. The existing drugs are efficient only for early stage worms. *Wolbachia* has endosymbiont activity with nematodes which is essential for its survival. So, *wolbachia* may be the promising drug target to develop antifilarialdrugs. The *wolbachia* genome suggested many drug targets, which are playing an indispensable role in nematode survival and reproduction. *Wolbachia* surface protein (*Wsp*) is one among them, which is observed in the outer membrane surface of the cell and helps in nematodes survival through porin activity. *Brugiamalayi* is one of the infectious filarial nematodes which is most prevalent parasite observed in India. So, our interest is the structural study of *wolbachia* surface protein (*Wsp*) from *B. malayi*, a promising drug target to eradicate the filarial disease by stopping the porin activity. We have chosen Wsp gene from *B. malayi* (*wBm-Wsp*) for structural investigation through modeling study. In addition to *wBm-Wsp*, heme bio-synthetic pathway enzymes are also found to be another promising target to develop drugs against nematodes due to lack of ferrochelatase (FC) enzyme in *B. malayi* that present in all *wolbachia*endosymbionts and as well as in human. So, it is believed that nematodes might get heme from *wolbachia*. In this view we also have modeled wBm-FC (hemH) structure and further structural analysis is in progress.

Poster Presentations

PP 1

AN ANDROID APPLICATION FOR COUNTING AND CLASSIFICATION OF CELLS

Sairam G*, Nanda Kumar M R*, Swetha R* and Arunachalam J

Department of Bioinformatics, SASTRA Deemed University,Thanjavur-613401

Email: arunachalam@bioinfo.sastra.edu

The traditional methods of cell analysis involve the manual counting of cells observed under the microscope from the smear. This method poses large dependency on the skills of the laboratory technician and may cause errors. The paper proposes an image processing technique to count and classify the cells based on their morphology using machine learning concepts in an android application .This method produces accurate results and do not requires any trained expert to operate this android application. This provides a low cost and portable solution for counting and classification of cells using an image processing and machine learning algorithm that works on the images captured by an android application from the microscope with considerable accuracy. This android application will be useful for pathological tests in remote or isolated geographical locations. So far the android application is created for counting the cells using image processing algorithm. In future the complete android application for classifying the cells for diagnosing diseases (malaria, sickle cell anemia, etc.) using machine learning concepts will be created.

PP 2

STRUCTURE BASED DESIGN OF NOVEL THERAPEUTIC AGENTS FOR BREAST CANCER TARGETS

Mathangi S, Sivakama Meenakshi P*, Thamarai Selvi S and Saraboji K

Department of Bioinformatics,SASTRA DeemedUniversity, Thanjavur-613401

Email: shivathenu@gmail.com

Breast cancer is a type of cancer that involves the uncontrollable division of cells in the breast tissues, common in women. There are various risk factors such as age, family history, food habits causing breast cancer, among which 5-10% of cases are due to genes inherited from the parents. Breast cancer is due to the over-expression of **HER2** (human epidermal growth factor receptor 2) and mutations in **BRCA1** (Breast cancer 1) gene.These proteins control the cell division and repair mechanism of breast cells, ensures DNA stability, act as an important biomarker and target to approximately 30% of breast cancer. Several studies have demonstrated that ligands can be used to inhibit the over expression of these proteins so that the breast cancer can be treated. The small molecule derivative of indol-1-ones and benzenesulfonamide were already identified for their anti-cancer property in the *in vivo* studies and is now targeted to HER2 and BRCA1. This study will help in designing suitable ligand to target HER2 and BRCA1 that would inhibit the tumorigenic activity of the same, serving as a treatment for the breast cancer. The aim of this project is to find the binding mode and inhibition mechanism of HER2 and BRCA1. Based on the multiple structure alignment, the common binding sites of these targets have been identified. The molecular docking was performed using GOLD suit and the results will be presented in detail.

PP 3

ANTI-NHSL PROPERTY OF 13-OCTADECENAL: A STUDY ON THEIR EFFECTS IN DIABETIC FOOT ULCERS

Rajivee J, Sabithadevi N and Gowdhaman D

Department of Bioinformatics SASTRA Deemed University, Thanjavur-613401

Email: jayarajivee@gmail.com

Diabetic foot ulcer is one of the major factors responsible for lower limb amputations in the world. Most of the diabetic foot infections are mono or poly microbial in nature and these microbial species tend to produce biofilms which hinder the normal wound healing process. These microbial species are drug resistant and there is a constant need for development of novel anti-bacterial agents to fight these infections in the case of diabetic foot ulcers. Studies reveal that most of the Gram positive bacteria which are devoid of NHSL fail to display/establish infections and hence NHSL can be an effective target to fight bacterial infection in diabetic foot ulcers.Our objective of the study is to explore the possible active compounds from *Cassia auriculata,*a potential folklore medicinal plant used for Ayurveda and siddha systems of medicine. Some parts of plant like flower, seeds help in curing chronic and other diabetic related diseases against pathogenic bacteria which possess NHSL- surface protein anchoring enzyme in microbial isolates of diabetic foot ulcers.

PP 4

STUDIES ON MOLECULAR GENETIC ANALYSIS OF HLA PREDISPOSITION TO HPV INFECTION

Aarthi S, Varsha R and Arunkumar G

Department of Bioinformatics, SASTRA Deemed University, Thanjavur -613401

*Email***:** *aarthi02saro20@gmail.com, varsha1697@gmail.com.*

Human HLA exhibit enormous polymorphism to combat pathogen diversity and HPV mutates its epitopes to escape human immune response. Both show geographic variation. Identifying the HLA-HPV epitope interaction shall shed light on the molecular basis of disease susceptibility/resistance.Over 60% of Human Papilloma virus (HPV) infection cause cervical cancer. Inhuman, MHC gene cluster codes for Human Leukocyte Antigen (HLA) that are expressed on antigen presenting cells and recognise foreign antigens. The objective of the present study is to identify the binding affinity and molecular interactions between positively selected HPV epitopes and HLA-B supertypes.HLA-B*2704 protein was taken from PDB and viewed it using PYMOL, which is also used to find the binding pocket. The HPV-L1 protein epitopes was taken from epitope prediction server “Immune Epitope Database Analysis”. The structure of various epitope were built using Argus lab. The binding affinity between HPV epitopes and HLA-B supertypes were identified using Autodock4.A total of eight number of epitopes were docked onto HLA B*2704**.** A range of binding energy values (-2.76 to 4.28) were obtained for the various epitopes Variation in amino acid property as a consequence of positive selection of HPV creates low affinity for binding with HLA thus demonstrating the molecular basis of HPV evolution and disease susceptibility.

PP 5

IN SILICO SCREENING OF CANCER-ASSOCIATED MUTATIONS IN THE HSA DOMAIN OF BRG1 AND ITS ROLE IN AFFECTING THE ARP-HSA SUBCOMPLEX OF SWI/SNF

ArunGoutham R. N*, Alagu Sankareswaran S, Suma Mohan S.

Department of Bioinformatics, SASTRA Deemed University, Thanjavur - 613401

Email: gouthamarun162@gmail.com

The family of Chromatin remodeling complexes facilitates the access of condensed genomic DNA during transcription, replication, and repair, by altering the histone-DNA contacts in the nucleosome structures. SWI/SNF (SWItch/Sucrose Non-Fermentable) family of ATP dependent chromatin remodeling complexes have been documented for their tumour suppressor function. Based on the recent cancer genome sequencing studies, 20% of the cancer causing mutations were identified in this family of remodeling complex which is close to the mutational frequency of p53. The functional characterization of mutations in the SWI/SNF subunits is important for understanding their role in tumorigenesis and identifying potential therapeutic strategies. In an effort to study the human actin related complex and the cancer causing mutations associated with it, we modelled the structure of theß-actin-BAF53A-HSA complex based on the yeast Arp-HSA complex (PDB id: 4I6M). The COSMIC mutations found in the HSA domain of BRG1 were subjected to analysis using ‘predict SNP’ program. The cancer-associated mutations in the HSA domain, R513Q, K502N, R521P, Y489C, and R466H were subjected to MD simulation studies using DESMOND program. Analysis of the trajectory indicates destabilizing effect of the mutations in the ARP-HSA sub-complex which would have an impact in the SWI/SNF complex assembly and remodeling function.

PP 6

STUDY OF CONFORMATIONAL FLEXIBILITY OF LOOPS IN PROTEIN-LIGAND COMPLEXES

Faiz.S, Vignesh P and Vigneshwar R

Department of Bioinformatics, SASTRA Deemed University, Thanjavur– 613401

Email:S.faiz_pilot@yahoo.com, vignesh1oct@gmail.com

The amino acid sequence forms the primary structure of the protein and their secondary structures are classified as alpha-helix, beta-sheet and loops. The loops are considered as irregular secondary structure and it consists of random coils, turns and bends. Analyse the residue in protein loops in the protein-ligand complex and their contribution in the ligand binding. It is not that all ligands bind to the loop region.so our aim is to analyse the loop that can interact with ligand molecule.Background study of our project is to understand the flexibility of protein loops and their significance in the protein-ligand interaction criteria for filtration of protein-ligand complexes are select the protein structures crystallized in atomic resolution (<1.5Å),proteins should be monomer and sequence similarity 30% only and structure which has ions as ligand are not taken into account.The non-bonded interaction involves close contacts, ionic and hydrophobic interactions are analysed for protein-ligand complexes using Chimera. Secondary Structure details are identified using DSSP. Chimera is a highly extensible program for interactive visualization and analysis of molecular structures and related data that includes density maps, supramolecular,sequence alignments, docking results, trajectories, and conformational ensembles.DSSP is a database of secondary structure assignments for all protein entries in the Protein Data Bank.DSSP is also the program that calculates Stride and Secondary Structure loop conformation. Hence, finally analysing the hydrophobic, ionic interaction and close contact. We are able to understand the residues responsible for the binding of ligands and their structural complexity.

PP 7

MOLECULAR DYNAMICS SIMULATION OF *MYCOBACTERIUM TUBERCULOSIS* PYRAZINAMIDASE W68R MUTANT: A STRUCTURAL PERSPECTIVE TO UNDERSTAND THE DRUG RESISTANCE MECHANISM

Janani S M*, Dharani N and Thamotharan S

Department of Bioinformatics, SASTRA DeemedUniversity, Thanjavur-613 401

E-Mail: jananimohan76@gmail.com

Tuberculosis (TB) remains world's leading cause of death and illness. The emergence of drug resistance bacterial strains is one of the main causes for the current spread of TB. Multidrug resistance tuberculosis is caused when the resistance is developed for two or more first-line drugs. It is of interest to note that the pyrazinamide (PZA) is an effective first- and second-line drug for the treatment of TB. It is a prodrug that is converted into its active form pyrazinoicacid by the bacterial enzyme pyrazinamidase(PZase). In *Mycobacterium tuberculosis*, this enzyme is encoded by the *pncA* gene and it consists of 185 amino acids. Hundreds of mutations that include missense, insertions and deletions were observed from the clinical isolates. It is to be noted these mutant were scattered throughout this gene. The single mutant (W68R) is located in the vicinity of the PZA binding site. Mutation of a bulky hydrophobic amino acid (W) by a positively charged residue (R) on the PZase attracts greater importance. Owing to this reason, we have performed molecular dynamics simulation for four systems (WT, W68R, WT-PZA and W68R-PZA) and each of these systems was carried out for 100 ns simulation time. Various parameterssuch as root mean squared deviation (RMSD), root mean squared fluctuation (RMSF) and the hydrogen bonding (HB) interaction between protein and ligand were analyzed to understand the structural changes caused by the mutant. The results are compared with the corresponding WT and WT-PZA systems. The detailed results will be presented during the symposium.

PP 8

ANALYSIS OF BINDING SITES AND PREDICTION OF GLYCATION SITES IN HUMAN SERUM ALBUMIN

Harini M *, Saraswathi N T

Department of Bioinformatics, SASTRA Deemed University,Thanjavur-613 401

Email: harini.murali27@gmail.com

Glycation (non-enzymatic glycosylation) is the result of the covalent bonding of a sugar molecule, such as glucose or fructose, to a protein or lipid molecule non enzymatically.Human serum albumin is the serum protein found in human blood.Albumin acts as a transporter for hormones, fatty acids, metal ions and other compounds.Human Serum Albumin was downloaded from Protein Data Bank and the binding sites were predicted using various servers like CASTp, PROBis, LIGSITE, and COACH. The glycation sites of human serum albumin for the downloaded pdb were predicted by calculating the distance between Lysine-Lysine and Lysine-Arginine residues within a range of 5-7Å (based on AGE such as GOLD, MOLD, DOGDIC, MODIC). The altered functional properties are such as reduced antioxidant property and reduced binding affinity when compared to non-glycated albumin. Some of the glycation residues identified through the Mass spectrometry studies are Lys 525, Arg 410, Cys 34, Arg 114, Lys 351, Lys 475, Arg 117, Arg 160, Arg 186, Arg 218, Arg 428. Here we would like to predict the residues which are responsible for the formation of advanced glycation end products such as GOLD, MOLD,GODIC and DOGDIC. These AGE molecules are crosslinked products between Lysine-Lysine or Lysine-Arginine of HSA. Thus the above residues were responsible for the formation of AGE molecules at binding site which probably disturb the binding property of the protein.

PP 9

CROSS-LINKING GUIDED MOLECULAR MODELING OF SWI/SNF CORE SUBUNIT ASSEMBLY

Nivedha T, Swetha M*, Suma Mohan S

Department of Bioinformatics, SASTRA Deemed University,Thanjavur-613 401

Email:bratzcoolgirl1056@gmail.com

SWI/SNF Chromatin remodeling complex is a ~2 MDa complex consists of 8 to 12 subunits conserved from yeast to human. This complex is recruited to the target promoters and can perform nucleosome remodeling in an ATP dependent manner. The core complex comprises of subunits, snf2(BRG1/BRM), Snf5(BAF47), and two copies of SWI3(BAF155/BAF170) which can form the minimized complex capable of doing the *in vitro* remodeling activity. Subsequent to the functional assembly of the core complex, the accessory subunits which help in targeting or regulation of the complex and the subcomplex specific signature subunits are recruited for target specific remodeling activity. Proper functional assembly of SWI/SNF is required for the transcriptional regulation of a large set of genes. The recent cancer genome sequencing studies highlight the core subunits, BRG1, BRM and BAF47 because of the high frequency of cancer causing mutations associated with it. Despite its functional importance, the structure or the subunit order for this large remodeling complex is not solved so far. Here, we propose a molecular model for multi-subunit functional assembly of yeast SWI/SNF core complex based on the cross-linking mass spectrometry data available from Sen P et al. The proposed molecular model of SWI/SNF core complex will aid in understanding the role of core subunits in nucleosome remodeling mechanism mediated by SWI/SNF.

PP 10

INVESTIGATION OF (S)-ß-(CARBOBENZOXYAMINO)-Γ-BUTYROLACTONE AS A NOVEL QUORUM SENSING QUENCHER OF URO PATHOGENIC *ESCHERICHIA COLI*

V.G Mahesh, S.Shyam Sundar and Karthi Shanmugam

Department of Bioinformatics, Sastra Deemed University,Thanjavur-613401

Email:mahesh21296.ms@gmail.com, shyamsundarsrini08@gmail.com

The uropathogenic Escherichia coli (UPEC) pathogenecity is affected by quorum sensing transcriptional regulator SdiA. In this study, we virtually screened Asinex small molecular database (~1,50,000 compounds) to discover novel quorum sensing inhibitors. Using Schrodinger drug discovery suite the database compounds were initially filtered for REOS, PAINS, and Ligprop for their drug likeliness. Using high throughput virtual screening the filtered compounds were screened against SdiA of UPEC. HTVS revealed that few database compounds with unique structural scaffold have higher affinity than the natural autoinducer molecule acyl homoserine lactone (AHL). These novel compounds can be further studied for its quorum quenching activity using in vitro and in vivo studies.

PP 11

ANALYSIS OF GENETIC AND EPIGENETIC CHANGES ASSOCIATED WITH CARDIOMYOPATHY

Muthu Meenakshi M, Hemallatha Y and Alex Stanley

Department of Bioinformatics, SASTRA Deemed University, Thanjavur -613401

Email: m.muthumeenakshi24@gmail.com, hemallathayuvaraj@gmail.com

Cardiomyopathy is the disease associated with heart muscle. It is defined as a morphologically abnormal heart where the muscles become enlarged, thick or rigid. This condition makes it hard for the heart to deliver blood to the body and leads to heart failure. It is more common in the present generations and there is a greater risk for the upcoming generations. At the genetic level, the disease is triggered by the mutation of the protein Titin. There are various epigenetic changes leading to the cardiac dysfunction. The core epigenetic change is the miRNAs which are involved in gene silencing. They bind to 3' end of the untranslated region and repress protein production by destabilizing mRNA and translational silencing. The raw data that were obtained from the ArrayExpress repository were normalized and differentially expressed using R packages. This differential expression leads to the ontological classification of the gene, which belongs to the "**Molecular Level**". The specific function that is associated to the epigenetic changes and the binding process responsible for the ischemic changes was interpreted as "**Protein Binding**". At the transcriptional level, there was reduction in the transcription speed at the exon 12, 13, 14 and 15 transcripts. Epigenetic repression of the glucocorticoid receptor (GR) and promotorhypermethylation leads to the downregulation of the genes which thus leads to the heart failure. The CpG methylation occurred at CREs and Sp1 binding sites and suppressed transcription. Thus, the epigenetics played a greater role in the developing hearts causing the repression of the glucocorticoid receptor which is involved majorly in the protein binding and hence, induces ischemic-sensitive phenotypic changes in the developing hearts.

PP 12

ANALYSIS OF GENE EXPRESION PATHWAYS IN WHEAT SEED DEVELOPMENT

Nivedhithaa J G*, Vishnupriya S S, .Sathya A

Department of Bioinformatics, SASTRA DeemedUniversity, Thanjavur -613401

Email: nivedhi1996@gmail.com

Wheat (*Triticumaestivum*) seed development is an important physiological process of seed maturation. Starch and protein are the principle storage reserves in the Wheat seed. In this project, we performed microarray data analysis using the Affymetrix Gene Chip to reveal the gene expression profiles in the phases of Wheat cultivar. Grain morphology observations showed that the period of 11 to 15 Days Post Anthesis (DPA) was a key stage for synthesis and accumulation of seed starch. The gene expression profiles varies during different stages of development in all organisms. In the present study, the differential gene expression profile has been visualized as Gene Expression network using Cytoscape during the seed development in Wheat. The steps to identify differential gene expression:The literatures has been collected from PubMed database. The genes has been retrieved for early phases. The gene expression data has been retrieved from GEO DataSets. The raw data has been normalized. The raw data has been annotated using R script. The expression network has been constructed using STRING software. Network development for expressed genes in early stage of Wheat seed development using Cytoscape. The network has been analysed using Network Analyzer. The analysis of the network helps us to identify the differentially expressed genes in early stage of Wheat seed development in Wheat. The network shows the enzyme of peroxidase family are upregulated in 11-20 DPA and the enzymes of transferases family are upregulated in 11-15 DPA.

PP 13

TARGETING DRUG RESISTANCE MECHANISMS IN LEISHMANIASIS NTD

J. Dhakshayini, R.Logachandar, AshokPalaniappan*

Department of Bioinformatics, SASTRA Deemed University, Thanjavur-613401

E-mail: kousijayaraman246@gmail.com,

One of the neglected tropical diseases was chosen in this study: LEISHMANIASIS. Caused by parasite *Leishmania major*. We are interested in drug resistance mechanisms mediated by *Leishmania*P-glycoprotein. Leishmaniasis is transmitted by the bite of infected female phlebotomine sandflies which can transmit the protozoa *Leishmania.Leishmaniasis* occurs in 88 tropical and subtropical countries and ug resistance is a major problem. To validate by MD simulations the inhibitor potential of Docked ugswith the leishmaniasis modelled p-glycoprotein structure.The first step was to determine the sequence identity of *Leishmania*Pglycoprotein with the human/mouse P-glycoprotein (hPGP) (UniProt ID: P08183). The sequence identity was determined by running a PSI-BLAST. The molecular docking was carried out using the AutoDock suite of tools. Interaction of known and novel ugs with P-glycoprotein[*Leishmania major*] was studied using the *in silico* tools. Cladosporin, the best (least) free energy of binding. The best pose of each docked complex was viewed using RasMol 2.1, and all interacting residues within a radius of 4.5 Å of the ligand were identified and analyzed. Molecular dynamics simulation was done using GROMACS. The drug topology was prepared using PRG2 server. The structures were solvated and 9 CL ions added to neutralise the charge on the system. Energy minimisation was carried out and it converged. We are doing molecular dynamics simulation of this large protein-drug system. The studies were carried out to identify a suitable Pgp inhibitor in drug resistant Leishmania worms. So far, we have identified cladosporin as a potential candidate and need to further ascertain the molecular dynamics stability of the docked complex of cladosporin with Leishmania P-glycoprotein.

PP 14

DYNAMEOMIC APPROACH TO STUDY PROTEIN-DNA INTERACTIONS

Supraja .V. Mohan *, Yashwanth Kumar K, Udaykumar M

Department of Bioinformatics, SASTRA Deemed University, Thanjavur – 613401

E-mail: suprajavasanthy@gmail.com

Protein-DNA interactions arefundamental to the cellular responses in our body. Factors that affect the protein-DNA interactions are: spatial arrangement of hydrogen bond donors and acceptors, protein dynamics, water, sequence-dependent DNA flexibility etc. Understanding the interplay of these factors in contributing to protein-DNA interactions requires a systematic analysis of various protein-DNA complexes. The dynameomic approach can provide a comprehensive understanding of protein-DNA interactions. There are currently, two databases that deal with dynameomic information: Dynameomics and BIGNASim. The aim of this project wasto create an open repository of MD simulated data of protein-DNA complexes and provide tools for analysis of the dynameomics data. A MySQL database was created with relational schema to store the data from molecular dynamics simulations. The data from simulation of the EcoRI-DNA complex was imported into the database. As a first step towards creating a database of simulations, in this miniproject, a single dataset consisting of 8222 atoms was inserted into the database. The querying speed was optimized using indexing which reduced the time taken for a single query from 20 minutes 0.2 seconds. However, a simple analysis such as distance calculation between two atoms (one of the fundamental analyses), took about 10 minutes. Options to optimize this speed is being explored. A database of protein-DNA simulations has been created. Data entry and data querying has been optimized.

PP 15

MOLECULAR VISUALIZATION TOOL FOR DIMER GENERATION USING CRYSTALLOGRAPHIC INFORMATION FILE (CIF)

Abishek M*, Swetha G and Udaykumar M

Department of Bioinformatics, SASTRA Deemed University, Thanjavur-613401

E-mail:abishekms7@gmail.com

The X-Ray crystal structure analysis are very important for structural elucidation of compounds. X-Ray diffractionis the most powerful technique to study the structure of the small and macromolecules. It is used to obtain structural information about crystalline solids and also useful to solve the three dimensional structure of biomolecules. Crystals are repeating arrays of molecules packed in an ordered way. Unit cells are the building blocks of the crystal structure. Unit cells are built up by repetition of blocks in a three dimensional manner. CIF file format is a standard file format for representing the crystallographic information. CIF file contains information about the small molecular crystal structure and it can be validated using checkCIF, ADIT validation server. CIF files can be visualized using Mercury, crystal explorer.The objective of the present work is to develop a computational tool for dimer generation using the crystallographic information file (.cif).The application Molvis provides excellent user interface for crystallographers for dimer generation by uploading CIF file and .MLC files. The software is developed using the Java core and advanced libraries.User interaction with Molvis starts by uploading a CIF file which contains Symmetry information, unit cell information and the coordinate information of the given small molecule.The toolkit also uses another file format (.mlc) which is generated from PIXEL program for the uploaded small molecule.The analytic modules implemented in Molvis increases the strength for analyzing the dimers and identify the stable dimers for further calculations.

PP 16

STAGE SPECIFIC ANALYSIS OFLIVER CANCER GENE EXPRESSION

Vignesh R *, Varsha K and Ashok Palaniappan

Department of Bioinformatics, SASTRA Deemed University, Thanjavur - 613401

E-mail: *vignesh.rajasekar.1994@gmail.com*

Liver Hepatocellular carcinoma (LIHC) is a major health problem worldwide. It is the fifth most common cancer in the world, with more than half million new cases yearly. The incidence of LIHC rose in the last decade. Cancer staging should serve to select the appropriate primary and adjuvant therapy, to estimate the prognosis, and also to assist in the evaluation of the results of treatment, and to exchange information. Even though many genes have been associated with the risk of LIHC, the genetic differences across different stages of LIHC are not known which can be studied through this study. The purpose of this study is to explore the gene expression differences in the different stages of liver cancer using the clinical and experimental datasets retrieved from TCGA. The expression and clinical data are merged into a single file based on the 32 Aliquot Barcodes since the genes in the files corresponds to these barcodes. Using the Unique command the duplicates (the repeated columns and rows) are removed. Grouping by various factors such as age, gender, ethnicity, stage of the cancer etc. Group-wise analysis of the subsets of data generated.Differential Expression analysis and PCA analysis is performed on the grouped subsets. The expression (20532x423) & clinical data (1748x377) are merged based on the search for common aliquot identifiers. The aliquot identifiers are spread out in many columns in the clinical data. After removing the duplicates the final data has the dimensions (22250x423). The histogram plots obtained for each of the specific factors explain us the gene expression corresponding to liver cancer with respect to that respective factor. After which the PCA analysis and Gene enrichment analysis will be performed on the grouped subsets.

PP 17

MODELLING STUDIES ON CU/ZN SOD1 FROM *WUCHERERIA BANCROFTI* ANDSTRUCTURE BASED INHIBITOR DESIGNING AGAINST FILARIASIS

Abinaya Lakshmi G* and Saraboji K

Department of Bioinformatics, SASTRA Deemed University, Thanjavur-613401

Email: abiganesh28@gmail.com

Lymphatic Filariasis, commonly known as elephantiasis, is a human disease caused by infection with parasites classified as nematodes (roundworms) of the family Filariodidea parasitic worms, of which *WuchereriaBancrofti* is responsible for 90% of the class. Preventive treatments that stop the spread of this infection is being given to patients through Mass drug Administration (MDA) strategy.The aim of the present study is to perform modelling studies on one of the redox enzyme Cu/Zn Superoxide dismutase of *Wuchereriabancrofti* and explore its structure-function relationships. This study focuses to find a suitable inhibitor to block the Filariasis activity caused by*W.bancrofti* Superoxide dismutase. This dimeric enzyme catalyses the dismutation of the superoxide (O2−) radical into the molecular oxygen (O_2) and hydrogen peroxide (H_2O_2). The three-dimensional structure prediction of protein was performed using Modeller 9.17 package. Structure visualization and analysis was done using PyMol and sequence Alignment was performed between the redox enzyme of Human and filarial worm using EMBOSS-Needle pairwise alignment program. Validation of the generated models was done using Ramachandran plots. The RMS deviation between the human and wb SOD1 was found to be 0.315 Å. Interestingly wb SOD1 displaces quite a handful of residues that are similar to human pathogenic mutants. Further we have used DeepSite, CASTp to predict the possible binding sites present in wbSOD1 and the screening of small molecule ligands, which can inhibit the filarial activity, are underway.

PP 18

CONFORMATIONAL ANALYSIS OF SOME *PARA*-SUBSTITUTED XANTHENE ANALOGS

Kavitha M, Hemarajalakshmi V

Department of Bioinformatics, SASTRA DeemedUniversity, Thanjavur-613401

Email: kavitham1196@gmail.com

Xanthene is a chemical compound acts as an antioxidant, food preservatives, food colouring agents, flavouring agents, anti-infective agents and excipients. Oral fluorescein, a derivative of xanthene, is used in angiography, brain tumor therapy and diagnosing vascular disorders. The chemical formula for xanthene is $C_{13}H_{10}O$. This project aims to find out the conformation with minimum energy amongst the modelled structure of xanthene derivatives. Crystal structures of ten (4-Br, 4-Cl, 4-H,4-F, 4-Me, 4-NMe2, 4-OH, 4-OMe, 4-NO2, 4-O-CH2-COOH) derivatives of xanthene were obtained from Cambridge Structural Database (CSD) and each derivative was modelled into four different conformations using GaussView. The modelled structures were optimized without any geometrical constraint. The vibrational frequencies were computed to assess the minimum energy conformer on the potential energy surface.The modelled were optimized using Gaussian-09 program using M05-2X/6-31+G(d) level of theory. The potential energy was calculated for all the 10 derivatives. The minimum energy conformation was found to be UP-UP conformation. In the minimum energy conformer, the dimethyl substituted carbon atom is oriented above the mean plane of the ring. The results suggest that 4-F, 4-NO2, 4-OMe, 4-OH compound crystallized in the minimum energy conformer, while othercompounds crystallized in the transition state geometry.

PP 19

ANALYSIS OF BINDING SITES AND PREDICTION OF GLYCATION SITES IN HUMAN HEMOGLOBIN

Surya R*, PriyaDharshini G, .Saraswathi N T

Department of Bioinformatics, SASTRA Deemed University, Thanjavur-613401

Email: rsurya816@gmail.com

Glycation is a non-enzymatic process in which free amino and thiol groups of the protein reacts with the glucose. Glycation-induced structural and functional modifications of hemoglobin may thus have deleterious effects on the protein leading to oxidative stress and pathological complications of diabetes mellitus. The objective of this project is to predict the glycation sites and analyse with binding sites of the Hemoglobin. Initially the binding sites of hemoglobin were identified using online servers such as CASTp, Ligsite and Coach and the glycation sites of Hemoglobin were identified by literature studies and by the construction of distance matrix. The distance matrix was constructed by calculating the distance between the Lys-Lys and Lys-Arg within the distance of 5-7Å. Thus the residues involved in the binding and glycation sites of hemoglobin were found out. The residues involved in the binding sites are LYS-99(A), LYS-127(A), LYS-7(C), LYS-127(C), ARG-141(A), ARG-141(C) and the residues involved in the glycation sites are LYS-7(C), LYS-11(C), LYS-127(A), ARG-141(C), LYS-127(C), ARG-141(A). The residues that are common in both binding and glycation sites may tend to undergo glycation. From the analysis of binding and glycation sites, the following residues were present in scommon they are Chain A - k(99), k(127), R(141) and Chain C – K(7), K(127), R(141). These residues were highly prone for the formation of crosslinking AGE's such as MOLD, GOLD, MODIC and DOGDIC.

PP 20

GENETIC AND EPIGENETIC CHANGES IN LIVER DISEASE

Priyatharsshini.S.P*, Shweta.S, Vigneshwar R

Department of Bioinformatics, SASTRA Deemed University, Thanjavur-613401,

E-mail:tharsshinikumar96@gmail.com

The basic idea behind cancer and cancer causing factors were identified.Epigenetics and the factors causing epigenetics were studied which includes study aboutepigenetic modifications such as methylation, histone modification, acetylation, etc. Sochoose to study Hepatocellular carcinoma. The patient data were collected from ArrayExpress and PubMed database .The basic idea was to explore the highly expressed genes in hepatocellular carcinoma and to find the genetic, epigenetic causes of the disease. Raw data of hepatocellular disease were collected and classified based on the age and gender.Data were extracted and by using R software data were normalized, and the highly expressedgenes were listed. Based on the ontological study, the data were classified as Molecular Factor(MF),Biological Process(BP) and Cellular Component(CC).As most of them werebecause of Biological process, the epigenetic significance was noted. From the highlyexpressed gene list, most predominant genes were found and by doing literature study,epigenetic role of such genes were found and methylation of the highly expressed gene was found to be one such cause for hepatocellular carcinoma. Hence highly causing genes forthis particular disease was identified.

PP 21

COMPOSITIONAL VECTOR BASED DISTANCE CALCULATION FOR TWO PROTEIN SEQUENCE

Mohana Krishnan P*, Sankara Subramanian M and Arunachalam J

Department of Bioinformatics, SASTRA DeemedUniversity, Thanjavur-613401,

E-mail: mohanakrishnan31296@gmail.com

Till today we are using alignment based method for calculating the distance between the protein sequences. The success of alignment based methods generally depends on the degree of homology between the sequences studied. Moreover when the size of the data increases as in Nextgen sequencing experiment data, creating alignments become an intensive task both in terms of computational power and memory. These aw backs can be overcome by developing efficient alignment free methods.Alignment-free methods can broadly be classified into four categories: a) methods based on *k*-mer / word frequency, b) methods based on substrings, c) methods based on information theory and d) methods based on graphical representation. Alignment-free approaches have been used in sequence similarity searches, clustering and classification of sequences, and more recently in phylogenetics .In this we are going to use compositional vector method.Objective of the project is to calculate the distance between two proteins based on resultant vector and design a web-based tool. Amino acid composition for the given sequence is calculated and multiplied with the values of 10 property amino acid table. A matrix is calculated for the obtained value. With the value of matrix resultant value is calculated and finally the web-based tool is designed .The programming language used is client based programming (php, javascript) and when the calculate button is clicked, it will display the distance between two proteins.A web-based tool is designed with the facility to give two protein sequence. The protein sequence is given by uploading a FASTA file.The corresponding distance for the two protein sequence was found.

PP 22

*In-silico*STRUCTURAL STUDIES ON REDOX ENZYMES OF *WuchereriaBancrofti*

Rakshandha V*, Saraboji K

Department of Bioinformatics, SASTRA Deemed University, Thanjavur-613401

Email: raksha.velumani @ gmail.com

Lymphatic Filariasis – tropical disease caused by the species of nematodes like *Wuchereriabancrofti* (*Wb*). Mosquito transmits the third larval stage of nematode into blood, matures into adult worm on reaching lymph gland. Current treatments includes mass drug administration of Ivermectin, Albandazoland diethylcarbamazine citrate. It's effective in eliminating the microfilaria but not adult worm. These drugs have no common inhibitory mechanism. The defence mechanism of the host is by Reactive oxygen species (ROS). The ROS is neutralized by the redox enzymes of*Wb*. We can devise ways to inhibit the function of redox enzymes by understanding their Structure - function relationship thereby killing the parasite. This study aims to analyse the structure and dynamics of redox enzymes of *Wb*such as Protein disulphideisomerase (PDI), an antioxidant with four domains. It catalyses the formation and breakage of disulphide bonds when proteins fold. The oxidized and reduced states of *Wb* PDI were modelled using the template, 'The crystal structure of reduced Hpdi (abb'xa') from Homo sapiens (PDB ID: 4EKZ_A)' which shared 32% sequence identity to target sequence. The active site region (CGHC) is conserved in the A and A' domains of PDI. MD simulation for oxidised form of *Wb* PDI for 10 ns using OPLS-AA force field is performed. In the MD simulation output, OPDI shifts to the RPDI state. The conformation at 7 ns is similar to the reduced form of PDI with RMSD of 6.364Å. Mode vector analysis gave the collective motion of the complete protein. The domain A is found to have higher flexible motion towards A'. B and B' domains are seen to accommodate for change in A and A' domain.The solvent accessible surface area (SASA) of PDI is comparable with Yeast-PDI and Human-PDI.

PP 23

AUTOMATED GENOMIC SIGNAL PROCESSING FOR GENES ASSOCIATED WITH CANCER

Ramya K*, Udhaykumar M

Department of Bioinformatics, SASTRA Deemed University, Thanjavur-613401

Email: *kumarramya96@gmail.com*

Genomic signal processing (GSP) is the technique used for analyzing and processing of genomic signals to gain biological knowledge, and translation of it, to systems-based applications can be used for disease gene prediction and to classify between diseased and non diseased gene sequences. Genomic signal processing involves obtaining gene sequences, especially cancer gene sequences. These sequences can be further processed using Digital Signal Processing (DSP) techniques to obtain the statistical parameters (mean, median, standard deviation, range, mean average density, median average density, standard deviation and entropy) for the cancerous and non cancerous gene sequences. The values for the above statistical parameters can be used to classify the sequences into cancerous and non cancerous sequences using machine learning techniques.

.

PP 24

IDENTIFICATION OF POTENTIAL INHIBITOR FOR BROMODOMAIN OF BRAHMA RELATED GENE 1 (BRG1)

Sindhuja A*, Sivaramakrishnan V

Department of Bioinformatics, SASTRA Deemed University, Thanjavur-613401

Email:sindhuja18021996@gmail.com

Transcription factor STAT3 is inappropriately activated in multiple hematopoietic and solid malignancies, in which it drives the expression of genes involved in cell proliferation, differentiation, survival and angiogenesis. The inappropriate activity of STAT3 is due to its interaction with other proteins. One of them is BRG1 (ATP-dependent helicase SMARCA4) protein encoded by the SMARCA4 gene. This gene is a member of the SWI/SNF (switch/sucrose-non fermenting) family of proteins. BRG1 regulates transcription by remodeling the chromatin structure via utilization of ATP and thereby disrupting the DNA–histone interactions at the nucleosomes. Therefore, BRG1 may represent a promising strategy for targeting inappropriately activated STAT3 in cancer. The small molecule compounds required for screening was downloaded from ZINC biogenic database and screened against bromodomain of human transcription activator BRG1 by using Autodock4. Ligands are ranked based on binding energy. Interactions between the ligands to bromodomain of human transcription activator BRG1 was also analyzed using Ligplot. Through computational studies we conclude on the lead compound 2-Naphthylhydrazine which shows better interaction to bromodomain of human transcription activator BRG1 than benchmark compound. We believe that the lead compound can further be optimized with additional future studies for improving its drug likeliness against bromodomainof Brahma Related Gene1.

PP 25

AN EXPLORATORY ANALYSIS OF GENE EXPRESSION DATA IN HUMAN VIRAL HEPATITIS

Sureshan M*, Arunachalam J

Department of Bioinformatics, SASTRA Deemed University, Thanjavur 613 401

Email: sureshan.tnj@gmail.com

Liver is a major organ which is responsible for several important functions like metabolism, detoxification, immunity towards infection, storage & synthesis of important cellular molecules etc.,The main viral infection that affects the liver is hepatitis. Our objective is to identify the patterns (similarities, dissimilarities or correlation) of gene expression profiles of four hepatitis viral types A, B, C and E whose microarray gene expression profiles were retrieved from gene expression omnibus database. Initially Gene set enrichment analysis was performed to identify the enrichment pathways for the viral types A, B, C and E whose gene set along with their pathways were subjected to two ordination techniques DCA & CIA. DCA (Detrended correspondence analysis) which identified 23 genes that are differentially expressed among six domains of KEGG ontologies for four viral types and CIA (Co-inertia analysis) which identified that 38 genes to be highly correlated across all six domain for the four viral types. Comprehensive analysis of four human viral hepatitis types shows twelve genes which primarily cause the inflammation.

PP 26

CONFORMATIONAL FLEXIBILITY OF THE Cu/Zn SOD1 INDUCED BY THE POINT MUTATION AT GLU100

Bhalekar Shraddha*, Lekha Sundararajan and Saraboji K

Department of Bioinformatics, SASTRA Deemed University, Thanjavur-613401

Email: shraddha.bhalekar976@gmail.com

Superoxide Dismutase1 (SOD1) is an enzyme that catalyzes the conversion of superoxide (O_2^-) to oxygen (O_2) and hydrogen peroxide (H_2O_2) which is then transformed into water and oxygen by the enzyme catalase. It is a homodimericmetalloprotein, with 153 residues, Cu^{2+} for catalytic activity and Zn^{2+} for structural stability of the protein, in each monomer. The secondary structure of SOD1 is an eight stranded Greek-key β-barrel with two α-helices. More than 100 different mutations in the SOD1 gene cause amyotrophic lateral sclerosis (ALS): a fatal neurodegenerative disease in which aggregation of the SOD1 protein is the primary mode of pathogenesis. Several results show that these mutations have remarkably diverse and unexpected effects on the structure, activity and stability of SOD1, but several mutations seem to have no effect on the properties of SOD1, except for altering the net charge of the protein.The aim of present study is to examine the dynamics of the far positioned mutation sites (6Å away from the metal site, interface-region and disulphide-bridge) and analyze it. Of the various SOD1 mutants, we have focused on mutations occurring in E100 (E100G, E100K), as studies speculate it has less negative charge than wtSOD1. MD Simulation was performed for E100G and E100K mutants using GROMACS-5.1.2 package utilizing Gromos43a1 force-field, steepest-descent Integrator, SPC water-model. All the MD simulations were performed for 50ns. Solvent-accessibility and deviation from the wtSOD1 were analyzed using NACCESS program and GROMACS respectively. Structure visualization and analysis was done using PyMol and COOT. Interestingly the dynamics results show that the point mutation E100G lead to the loss of β-barrel stability, which alters the shape of β6-strand towards the hydrophobic core of the monomeric unit. The MD simulation clearly displays the induced conformational flexibility due to the loss of negative charge in the β-sheet, whereas the E100K mutation retain the enzyme more similar to the wtSOD1. The detailed results will be presented.