7TH ANNUAL LA CONFERENCE ON

COMPUTATIONAL BIOLOGY & BIOINFORMATICS

April 5-6, 2019 • LSU Digital Media Center • Baton Rouge, LA

- Cancer Informatics
- Microbiome and Metagenomics
- Health Informatics, Big Data, and Computing
- Translational Bioinformatics and Data Visualization
- Evolutionary Genomics and Phylogenetics
- Virology and Infectious Diseases

LSU Center for Computation & Technology

LSU Office of Research & Economic Development

LBRN

iSCB International Society for Computational Biology

LSU Center for Experimental Infectious Disease Research

Elodie Ghedin, PhD
Director of Center for Genomics and System Biology – New York University

Devin Absher, PhD
Faculty Investigator - Hudson Alpha Institute for Biotechnology

Byoung-Do (BD) Kim, PhD
Director of Research Computing - University of Virginia School of Medicine

Michael Robeson, PhD
Assistant Professor of Biomedical Informatics - University of Arkansas for Medical Sciences

Ying Xu, PhD
Professor of Bioinformatics and Computational Biology - University of Georgia

Jeremy M. Brown
Associate Professor at Biological Science - LSU

Isidore Rigoutsos, PhD
Director of the Computational Medicine Center, Professor, Thomas Jefferson University

Hanoch Kaphzan, MD PhD
Principal Investigator, Faculty member, University of Haifa

Louisiana State University | Digital Media Center | Baton Rouge, LA
Friday, April 5 12:50 p.m. - 9:00 p.m.
Saturday, April 6 9:00 a.m. - 4:30 p.m.

Table of Contents

Day 1: Friday, April 5, 2019

Welcome and Opening Remarks

12:50 - 1:00 pm ..... Gus Kousoulas, Associate Vice President (LSU ORED) / LBRN PI

1:00 - 1:05 pm ..... Stacia Haynie, Executive Vice President & Provost (LSU)

1:05 - 1:25 pm ..... Alfred Tauber, Chairman of the Board of Governors of the University of Haifa Zoltan Kohn Professor Emeritus of Medicine, Boston University

Session I: Microbiome Data Science and Protein Interactions

1:25 – 2:25 pm..... Michael Robeson, University of Arkansas for Medical Sciences
Cultivating a community that enables scalable, extensible, and reproducible microbiome data science

2:25 – 2:45 pm..... Seetharama Jois, University of Louisiana Monroe
Conformation convergence: NMR and MD simulations of SFTI grafted peptides that inhibit protein-protein interaction

Session II: Epigenomics and Complex Human Traits

2:45 – 3:45 pm..... Devin Absher, Hudson Alpha Institute for Biotechnology
Bioinformatic approaches to untangling epigenetic complexity

3:45 – 4:05 pm..... Manisha Panta, University of New Orleans
Prediction of hierarchical classification of transposable elements using machine learning techniques

4:05 – 4:20 pm..... Break

Session III: Computational Evolutionary Biology and Sequence Variation

4:20 – 5:20 pm..... Jeremy Brown, Louisiana State University
The genomic landscape of phylogenetic variation

5:20 – 5:40 pm..... Joshua Vandenbrink, Louisiana Tech University
Transcriptomic analysis of arabidopsis thaliana exhibiting novel blue-light phototropism on board the international space station

Session IV: Health Disparities and Clinical Informatics

5:40 – 6:40 pm..... Isidore Rigoutsos, Thomas Jefferson University
Novel nuclear and mitochondrial RNAs that are linked to key pathways and depend on sex, population origin, and race

6:40 – 7:00 pm..... Loren Gragert, Tulane University
Clinical informatics tools for precision histocompatibility assessments in solid organ transplantation

7:00 pm – 9:00 pm Poster Session and Dinner Provided: Emerging Technology Center

Day 2: Saturday, April 6, 2019

Opening Remarks

9:00 - 9:10 am ..... Gus Kousoulas, Associate Vice President (LSU ORED) / PI LBRN

Session V: Psychiatric Disorders and Viral Genome Evolution

9:10 – 10:10 am..... Hanoch Kaphzan, University of Haifa
The use of bioinformatics tools in the study of Angelman syndrome

10:10 – 10:40 am..... Leonid Brodsky, University of Haifa
Virus genome micro-evolution: computational study of mutations in virus genome sequencing data

10:40 – 10:50 am..... Break

Session VI: Cancer Bioinformatics

10:50 – 11:50 am..... Ying Xu, University of Georgia
Maintaining intracellular acid-base homeostasis is probably at the core of cancer formation, progression and metastasis

11:50 – 12:10 pm..... Chindo Hicks, Louisiana State University Health Sciences Center - New Orleans
A novel integrative bioinformatics approach to modeling pathway crosstalk between germline and somatic genomes in cancer

12:10 pm – 1:10 pm Lunch Provided: DMC Building Lobby

Session VII: Genomics of Infectious Diseases and Respiratory Infections

1:10 – 2:10 pm..... Elodie Ghedin, New York University
Microbial networking in respiratory infections

2:10 – 2:30 pm..... Weishan Huang, Louisiana State University
Heterogeneity of regulatory T cells in the mouse airway during influenza infection

2:30 – 2:40 pm..... Break

Session VIII: High Performance Computing, Cloud Computing and Big Data

2:40 – 3:40 pm..... Byoung-Do Kim, University of Virginia
Advanced computing solutions for biomedical research

3:40 – 4:10 pm..... Ruzhu Chen, International Business Machines
Genomics Application Performance on Power

4:10 - 4:30 pm..... Meeting Wrap-Up and Awards Announcement

Oral Presentation Abstracts

OA-01 Bioinformatic approaches to untangling epigenetic complexity

Devin Absher
HudsonAlpha Institute for Biotechnology

The study of epigenetic links to human disease is complicated by the dynamic nature of the epigenome, the cell-type specificity of epigenetic states, and the difficulty in distinguishing cause and effect. Careful modeling of epigenetic data is required to overcome this complexity, and multiple statistical and informatic approaches are often needed to study the epigenome effectively. Analytical approaches to epigenetic studies will be discussed, in the context of both common human traits as well as autoimmune diseases like lupus.

OA-02 Virus genome micro-evolution: computational study of mutations in the virus genome sequencing data

Leonid Brodsky
University of Haifa

The NGS sequencing of virus genomes across passages of the cell culture allows to monitor virus genome microevolution that exhibits a process of adaptation to the host cell response. Generation of genomes with new mutation variants is the major mechanism of virus adaptation: frequency of the beneficial genome mutation variants is increasing across passages and decreasing for the detrimental variants. A computational analysis of mutations is based on dynamics of their frequency profiles across passages. This analyses includes issues as follows: •How trustable is the frequency of a particular mutation variant under variable coverage of the virus genome positions by the NGS reads? •How significantly beneficial or detrimental is a particular mutation variant. •What patterns of the mutation dynamics across passages appear in micro-evolution of the RNA(+) viruses. •Structural local similarity screening as a way to link a mutation with biological function of the corresponding viral protein. •Modeling of the virus evolution dynamics

OA-03 The Genomic Landscape of Phylogenetic Variation

Jeremy Brown
Louisiana State University-Baton Rouge

Phylogenetic thinking is one of the most important pillars of modern biology. The ability to frame biological questions in the context of common ancestry and shared inheritance underpins not just evolution, but also fields as wide-ranging as forensics, epidemiology, conservation, and medicine. In recent years, genome-wide sequencing has presented both extraordinary new opportunities and new challenges for phylogenetic inference. By radically increasing the amount of data that can be collected quickly and cheaply, new sequencing technologies continue to unlock a wealth of information that can be brought to bear on questions about phylogenetics and molecular evolution. However, these massive datasets can be messy, complicated, and difficult to interpret. My research focuses on how to make the most of the newly unearthed information, while avoiding new pitfalls. In this talk, I will discuss several related efforts I have undertaken with students and collaborators to understand how and why phylogenetic information varies across genomes, how this variation relates to the underlying processes of molecular and genomic evolution, and how a phylogenetic perspective can be used to address questions of both fundamental and applied importance.

OA-04 Genomics Application Performance on Power

Ruzhu Chen
IBM Systems

A highly available, scalable, easy-to-use high-performance data analytics solution is enabled on Power9 systems and ESS storage through IBM cloud private (ICP) for genomics applications. Genomics workloads from sequence alignment, variant detection and quality filtering to deep learning analysis are orchestrated in ICP cluster

deployed with optimized applications such as GATK4 pipeline, PowerAI, H2O and PowerAI Vision. Tuning and performance optimization of these applications are presented.

OA-05 Microbial Networking in Respiratory Infections

Elodie Ghedin
New York University

Few microbes (viruses, bacteria, fungi) live in isolation, or exclusively with members of their own kingdom or domain. Affinities or aversions between microbial members influence the community structure, but these interactions can be re-organized with the arrival of disruptors. In respiratory infections, for example, infectious agents-be they viral or bacterial-are entering an environment within the host where they can impact existing ecological relationships amongst local residents. Disrupting these 'social' networks has ecological and physiological consequences. As we begin to discover the importance of microbial associations in understanding host-pathogen interactions, we need innovative ways to capture direct and indirect effects between viruses, fungi, bacteria, and the host. I will discuss some of our work on respiratory tract infections, such as influenza, and how we tackle the complex host-pathogen interplay by inferring networks of interactions amongst microbes and with their hosts.

OA-06 Clinical Informatics Tools for Precision Histocompatibility Assessments in Solid Organ Transplantation

Loren Gragert
Tulane University

Differences in immune system genes called human leukocyte antigens (HLA) compose an immunological barrier to kidney transplantation. Antibodies can recognize mismatched HLA molecules as foreign. If there are preexisting donor-specific anti-HLA antibodies, the donor organ is incompatible as it would likely be lost to hyperacute rejection. Histocompatibility testing uses data from human leukocyte antigen (HLA) genotyping and solid phase assays of HLA antibodies. While HLA typing is performed by molecular methods, the solid organ allocation system, managed by the United Network for Organ Sharing (UNOS), does not utilize molecular typing data for matching. To decide if organ offers from the UNOS match run should be accepted, molecular HLA typing data is manually analyzed by transplant centers. We developed a virtual crossmatch informatics tool to streamline this process. We found that the current match run will make offers of organs that are clearly incompatible. These unnecessary offers increase cold ischemia time, which reduces the quality of organs. We also found that utilizing molecular HLA typing data can reveal additional compatible donors. To improve equity, patients with many antibodies are given waiting list priority based on their calculated panel-reactive antibody (CPRA) values. CPRA is the percent of the donor pool that is incompatible based on the patient's anti-HLA antibody specificities and population HLA frequencies. The current UNOS CPRA measure is also limited by lack of molecular HLA typing data. Some categories of HLA antibodies don't count towards CPRA. During validation of a new CPRA panel based on molecular typing of >10 million bone marrow registry donors, we also found surprising errors in UNOS CPRA. The kidney allocation system would benefit from implementation of HLA datasets and informatics tools based on molecular typing data.

OA-07 A Novel Integrative Bioinformatics Approach to Modeling Pathway Crosstalk Between Germline and Somatic Genomes in Cancer

Chindo Hicks
LSUHSC New Orleans

Chindo Hicks, Jiande Wu and Tarun Karthik Kumar Mamidi Department of Genetics, Louisiana State University Health Sciences Center, School of Medicine, New Orleans Abstract Advances in high-throughput genotyping and next generation sequencing of cancer genomes have enabled discovery of germline and acquired somatic

mutations. Germline mutations, contained within the heritable genome, and somatic mutations, acquired de novo by cancer cells in the tumor genome, have largely been considered as separate endeavors in cancer research. Emerging evidence from our research and others suggests that germline and tumor genomes interact to drive tumorigenesis. Germline mutations increase the likelihood that an individual will develop cancer in his or her lifetime. Therefore, understanding the molecular mechanisms underlying oncogenic interactions between the germline and tumor genomes to drive tumorigenesis provides a rational basis for the development of novel precision prevention strategies. Here we report a novel integrative bioinformatics approach for modeling network and pathway crosstalk between the germline and somatic genomes in breast and prostate cancer using genotype data on germline genomes from genome-wide association studies (GWAS) and next generation sequence data on cancer genomes from The Cancer Genome Atlas (TCGA). Our analysis has revealed oncogenic interactions and cooperation between the germline and the somatic genomes during tumorigenesis. Most notably, the investigation has revealed crosstalk between the pathways regulated by the germline genome and the pathways regulated by the somatic genome. We conclude that oncogenic interactions and cooperation between germline and somatic mutations occurs through network and pathway crosstalk between the two genomes.

OA-08 Heterogeneity of regulatory T cells in the mouse airway during influenza infection

Weishan Huang
Louisiana State University-Baton Rouge

Influenza (flu) infection is the leading cause of respiratory illness worldwide. While flu vaccines are effective at reducing morbidity and mortality, viral clearance relies on the development of a strong immune response. Virus-specific antibodies and T cells are required for killing of flu-infected cells and clearance of the virus. Although strong T cell responses are desired, the cytotoxic effector function of the innate and adaptive immune responses, can lead to the development of pulmonary immunopathology, which together with recurrent bacterial infections is a major cause of death during flu pandemic. While T cells produce inflammatory cytokines during flu infection, they can also produce the immunosuppressive cytokine IL-10, which is critical for limiting the immunopathology caused by the excessive immune responses. The population composition and molecular signatures of the IL-10-producing regulatory T cells in the airway during flu infection are unclear. Using mouse models that report the production of IL-10 by GFP and expression of conventional regulatory T cell marker Foxp3 by RFP, a mouse adaptive model of influenza infection, transcriptomic analyses in the population and single cell levels, and transgenic mouse models that are impaired in T cell-derived IL-10 production, we found that IL-10+ regulatory T cells in the mouse airway during flu infection is mainly comprised of CD4+ Foxp3+, CD4+ Foxp3- and CD8+ subsets, and they differ from IL-10-producing T cells under pneumonitis or asthmatic conditions. Within each subset, IL-10-producing T cells exhibit significant heterogeneity. Information gained from this dataset provides insights into the T cell subset heterogeneity and signature markers, and shed light on the strategic designs of therapeutic development utilizing the immunomodulatory features of T cells for the treatment of pulmonary immunopathology.

OA-09 Conformational convergence: NMR and MD simulations of SFTI grafted peptides that inhibit protein-protein interaction.

Seetharama Jois
University of Louisiana at Monroe

Seetharama Jois, Pravin Parajuli, Achyut Dahal, Sitanshu Singh. College of Pharmacy, University of Louisiana Monroe, Monroe LA USA 71201 Multicyclic plant-derived peptides with disulfide bonds, are resistant to thermal, chemical, and enzymatic degradation, and are orally bioavailable. As a model system to inhibit protein-protein interaction (PPI) by these peptides, we will use CD2-CD58 molecular pairs. CD58 is a cell adhesion molecule with only one known ligand, CD2. It plays a critical role in the facilitation of antigen-specific

recognition through interaction with CD2 on T lymphocytes and natural killer cells. Blocking of adhesion/co-stimulatory molecules results in blocking PPI and preventing the generation of the immune response. We have generated multicyclic peptides from plant sources for grafting biologically active epitopes from CD2 protein to to inhibit PPI of CD2-CD58. However, the designed sunflower trypsin inhibitor (SFTI) grafted peptides exhibited multiple conformers in solution. To stabilize the peptide conformation to single conformation, we have designed analogs of SFTI with dibenzofuran (DBF) moiety. Using two dimensional NMR, we have shown that SFTI exhibits at least three possible conformations in solution as shown by multiple sets of NMR resonances for each amino acid residue in the peptide. By introducing the DBF moiety at a crucial position in the peptide, we were able to constrain the peptide conformation in solution to one major conformation as shown by 2D NMR. MD simulations were carried out to evaluate the conformational stability of SFTI grafted peptides. Detailed analysis indicated that Pro-Pro presence in the peptide sequence might have an effect on the existence of multiple conformations in solution.

OA-10 Utilizing bioinformatics tools in studying Angelman syndrome

Hanoch Kaphzan
University of Haifa

Angelman syndrome (AS) is a human neurodevelopmental disorder associated with autism, intellectual disability, motor dysfunction, epilepsy and lack of speech. AS is caused by the lack of expression of the E3-ligase UBE3A, usually due to a deletion of the UBE3A gene on the maternal allele. The paternal UBE3A gene is imprinted after birth and is not expressed in the mature brain. Interestingly, in the brain, UBE3A effects are dosage sensitive, and UBE3A duplication results in autism. In order to reveal the physiological effects and the underlying molecular mechanisms affected by UBE3A, we utilize mice models and mice embryonic fibroblasts (MEFs) with Ube3a deletion. The mouse model of AS exhibit an array of abnormalities that correlate with the neurological deficits observed in AS human patients. In addition to physiological and behavioral studies, we utilize bioinformatics tools for studying the proteomic, transcriptomic, and epigenetic changes underlying AS.

OA-11 Advanced Computing Solutions for Biomedical Research

Byoung-Do Kim
University of Virginia School of Medicine

Biomedical research has become data-intensive, and the computational biology and informatics community is now producing an unprecedented scale of big data on a daily basis. Traditional high-performance computing platform still serves well for large-scale genomics or structural biology computation, but there is a wide variety of computational requirements in other biomedical fields. Emerging technologies such as Cloud Computing and Linux Container provide researchers with more options for their specific research needs, and the easily accessible and customizable environment of the new technology will attract more researchers. In this talk, we will look into the available advanced computing solutions for biomedical research with a couple of use cases developed by the School of Medicine Research Computing team at The University of Virginia.

OA-12 Prediction of Hierarchical Classification of Transposable Elements using Machine Learning Techniques

Manisha Panta
University of New Orleans

Motivation: Transposable Elements (TEs) or jumping genes are the DNA sequences that have intrinsic capability to move within a host genome from one genomic location to another - genomic location can be either same or different chromosomes. Studies show that TEs have a role in genome function and evolution as their presence can modify the functionality of genes and increase the size of the genome. Thus, proper classification of the identified jumping genes is important to understand their role in germline and somatic evolution. While there

are computational methods that perform either binary classification or multi-label classification of TEs, few studies have focused on the hierarchical classification of the elements. The state-of-the-art method has used a neural network for hierarchically classifying TEs. In this regard, we explored different machine learning approaches for hierarchical classification of transposable elements. Results: Firstly, we compared state-of-the-art machine learning models, based on the Multi-Layer Perceptron (MLP) neural network algorithm, for two hierarchical classification strategies: non-Leaf Local Classifier per Parent Node (nLLCPN) and Local Classifier Per Parent Node and Branch (LCPNB) using a variety of machine learning (ML) models. We experimentally demonstrated that LCPNB classification strategy performs statistically better with most of the models. The comparative results with different learning algorithms showed that optimized Support Vector Machine (SVM), with a radial basis function (RBF) kernel, has higher balanced-accuracy for both classification strategies. We found that the balanced-accuracy (using the hF measure) for the LCPNB strategy with the SVM model on benchmark hierarchical datasets of TEs from the Plant Genome and Systems Biology (PGSB) was 0.905; for TEs from the Repbase database, hF was 0.886, and when TEs from both databases were combined, hF was 0.881. We then proposed a new method, ClassifyTE, based on an augmented stacking approach to predict the hierarchical class of a transposable element using the LCPNB strategy. Our proposed model outperforms existing state-of-the-art Multi-Layered Perceptron method. Furthermore, ClassifyTE outperforms all the existing machine learning approaches for hierarchical classification of TEs. Using this approach, the balanced-accuracy (hF) for PGSB, Repbase, and the mixed datasets were found to be 0.915, 0.882 and 0.892 respectively, higher than our previous results using the SVM model.

OA-13 Novel Nuclear and Mitochondrial RNAs that are Linked to Key Pathways and Depend on Sex, Population Origin, and Race

Isidore Rigoutsos
Thomas Jefferson University

Ever since the discovery of microRNA (miRNA), the community has largely adhered to a paradigm according to which a single stretch of genomic DNA is transcribed into a miRNA precursor molecule that is then processed to produce a single (regulatory) miRNA product, approximately 22 nucleotides (nts) in length. Similarly, for more than 50 years, the community viewed transfer RNA (tRNA) as ancillary molecules that participated in the translation of codons to amino acids. As next generation sequencing (NGS) became widespread, scientists observed that each miRNA precursor gives rise to multiple miRNA products with slightly different endpoints and different abundances. These 'variants' were called 'isomiRs.' In the case of tRNA, NGS revealed that both precursor and mature tRNA give rise to multiple 'tRNA-derived fragments' or 'tRFs' with different endpoints and abundances. The initial reaction was to dismiss isomiRs and tRFs as inconsequential bystanders. By analyzing datasets from thousands of healthy individuals and patients, we were first to show that human isomiRs and tRFs are produced constitutively and that the identities and abundance levels of isomiRs and tRFs depend on a person's sex, population origin, and race. We also showed that isomiRs and tRFs additionally depend on the tissue at hand, the tissue's state (health vs. disease), and the disease type/subtype. From a functional standpoint, we showed experimentally that distinct isomiRs from the same miRNA precursor arm can target different groups of messenger RNA (mRNA). For tRFs, we showed that different tRFs from the same parental tRNA are anti-correlated with different mRNA. Particularly intriguing is our finding that the mitochondrially-encoded tRNA are a very rich source of tRFs. In fact, sequence for sequence, the 22 mitochondrial tRNA produce 15x as many tRFs as their nuclearly-encoded counterparts.

OA-14 Cultivating a community that enables scalable, extensible, and reproducible microbiome data science.

Michael Robeson
University of Arkansas for Medical Sciences

The advent of high-throughput sequencing and other '-omics technologies has provided microbial ecologists unprecedented access to better characterize the microbial world. Improvements in biotechnological and computational approaches have become more robust and economically feasible over the last several years. Resulting in the ability to not only survey entire sets of microbial communities, but also elucidate the functional roles of microbes and their interactions with one another and their hosts. However, there is a need for easy-to-use tools and approaches that can effectively integrate a variety microbiome associated multi-omics data. Just as important, these tools should be accessible to beginners and also cultivate a highly invested user and developer community willing to share and discuss ideas. In this presentation, I will use QIIME 2 (https://qiime2.org/) as an example of how an inclusive community of scientists, software engineers, statisticians, educators, students, and others invested in microbiome research, can meaningfully contribute to development of microbiome data science platforms with an emphasis on reproducibility and transparency.

OA-15 Transcriptomic analysis of Arabidopsis thaliana exhibiting novel blue-light phototropism on board the International Space Station

Joshua Vandenbrink
Louisiana Tech University

Authors: Joshua P. Vandenbrink 1,2, Raul Herranz3, William Poehlman4, F. Alex Feltus4, Alicia Villacampa3, Malgorzata Ciska3, F. Javier Medina3, and John Z. Kiss 2 1 School of Biological Sciences, Louisiana Tech University, Ruston, LA 71272 USA 2 Department of Biology, University of North Carolina at Greensboro, Greensboro, NC 27402, USA 3 Centro de Investigaciones Biológicas (CSIC), Madrid, E28040, Spain 4 Department of Genetics and Biochemistry, Clemson University, Clemson, SC 29634, USA Growth-mediated movements ('tropisms') are essential for plants to respond to changing environmental conditions. These tropisms also play a direct role in determining the architecture and final form of the plant. Water (hydrotropism), light (phototropism) and gravity (gravitropism) are a few of the environmental stimuli that cause tropistic movement. Currently, little is known about the interaction between tropistic movement. Here, we utilize conditions of microgravity present on the International Space Station to identify genes that are differentially expressed between µg and 1-g when plants are grown under unidirectional blue light. Gene expression of these groups was then analyzed via RNA sequencing. A total of 296 genes were found to be significantly differentially expressed between microgravity when compared to 1-g controls ($p < 0.05$). In addition, Pathway Analysis of significantly differentially expressed genes (µg vs 1-g) identified 8 molecular pathways that are significantly affected by reduced gravity conditions. These findings show that light-associated pathways (e.g. photosynthesis-antenna proteins, photosynthesis, porphyrin and chlorophyll metabolism) show significant downregulation while ribosome biosynthesis and oxidative phosphorylation are upregulated in conditions of microgravity.

OA-16 Maintaining Intracellular Acid-Base Homeostasis is Probably at the Core of Cancer Formation, Progression and Metastasis

Ying Xu
University of Georgia

In this talk, I will present some of our recent discoveries made through mining and modeling omic data of tens of thousands of samples of cancer tissues and inflammatory diseases. We have demonstrated: cancer cells of at least 16 of the most prevalent cancer types all harbor Fenton Reactions: $Fe2+ + H2O2 \rightarrow Fe3+ + \cdot OH + OH-$ in their cytosols and mitochondria. A consequence of the cytosolic Fenton reactions is that they continuously

produce OH- at rates that can gradually overwhelm the intracellular pH buffer, which will drive the pH up and ultimately kill the cells if not neutralized. As response, all cancer cells increasingly alter their metabolism, commonly referred to as metabolic reprogramming. We have studied 30+ such reprogrammed metabolisms and found that they have one thing in common: they each produce more H+ than their original metabolisms, presumably to neutralize the OH-. Among them, the most significant is the Warburg effect, the basis for cancer detection using PET/CT. I will present data to show that the induced Warburg effect can drive cell division. In addition, I will present preliminary data to show that two lesser known reprogrammed metabolic pathways; over-production of sialic acids and gangliosides of specific types produce more protons. Their deployment in plasma membranes create increasingly stronger cell-cell repulsion because of their negative charges, which may gradually result in substantial deformation of affected cells, leading to the activation of a mechanical stress response program, which coordinates a series of counterbalancing activities to the cell deformation, including cell protrusion and contraction, ultimately giving rise to cell migration.

Poster Abstracts

PA-01 ***Predictor Model for HIV Patient Pregnancy for Newborn***

Bhupendra Acharya
University of New Orleans

The majority of children that are living with HIV are infected via mother during pregnancy, childbirth or breastfeeding. Children that were infected during pregnancy with HIV are at risk of mental impairment. The mental impairment impacts the children ability to learn and perform in daily day to day activities. The result showed that impairment in cognitive functioning is associated with HIV disease advancement during the growth of the newborn. Our result showed that children that are between 3-7 are comparatively functioning with poor performance than healthy children. The children that were exposed to HIV had a harder time in learning and grasping the environment than compared to uninfected.

PA-02 ***Accurate Identification of Brain Hemorrhage using Genetic Algorithm based Feature Selection and Stacking***

Duaa Alawad
University of New Orleans

Abstract: Brain hemorrhage is a type of stroke. It can be caused by an artery in the brain bursting and causing localized bleeding in the surrounding tissues. Computerized tomography (CT) scan of the brain is one of a variety of imaging examinations, which enables the accurate detection and diagnosis of the brain hemorrhage. In this work, we developed an automated method to detect the existence and the type of brain hemorrhage in a CT scans image of the brain. The proposed method comprises of several steps such as image preprocessing, image segmentation, feature extraction, feature selection and the design of an advanced classification framework. The image preprocessing and the segmentation steps involve the removal of the skull region from the image and the identification of the region of interest (ROI) using Otsu's method, respectively. Subsequently, feature extraction involved the collection of a comprehensive set of features from the ROI, such as the size of the ROI, centroid of the ROI, perimeter of the ROI, the distance between the ROI, and the skull. Moreover, for improved performance, an Evolutionary Algorithm (EA) based feature selection algorithm is used to select relevant features. These features are then used to train the Stacking-based machine learning framework to predict different types of brain hemorrhage. Finally, the evaluation results show that the proposed predictor achieves a 10-fold cross-validation (CV) accuracy (ACC) of 99.17% on the benchmark dataset. Therefore, the proposed approach significantly outperforms existing brain hemorrhage classification approaches and can be useful for effective prediction of brain hemorrhage types.

PA-03 ***Interplay of Folate Receptors and Glucocorticoids in the Human Prostate Cancer***

Bashir Atteia
Southern University at New Orleans

Background: Prostate cancer (CaP) is the most prevalent non-skin and least understood of all human malignancies in males in the world. There is a significant controversy about the role of treatment for prostate cancer patients. Practice guidelines recommend that these patients can be surgically operated and treated with radiation, based on their tumor characteristics, prostate-specific antigen level, age, comorbidities, and preferences. Another potential treatment modality is androgen deprivation therapy (ADT) as monotherapy for localized prostate cancer. Although ADT has a well-defined role in patients with metastatic disease or with high-risk localized disease undergoing radiotherapy, its role as monotherapy in patients with localized disease has not been established in clinical trials and its use is not uncommon. Although, the adjuvant glucocorticoids therapy such as dexamethasone is provided to hormone-refractory prostate cancer patients and to ameliorate the pain cancer therapy, the recurrence of prostate cancer is high. Goal: The aim of this study is to investigate

the correlation between glucocorticoids and folate receptors. Materials and Methods: Human prostate cancer cells C4-2B and PC3 were sed. Results: treatment of prostate cells with dexamethasone or folic acid elevates the cell proliferation rates. However, co-treatment of both dexamethasone and folic acid reduced the proliferation rate of cancer cells. Furthermore, methotrexate induced apoptosis to cancer cells. Conclusions: Dexamethasone may be implicated in the enhancement of recurrence of prostate cancer in low folate levels, where dexamethasone may interact with folate receptors and enhance the proliferations of prostate cancer cells. Co-Treatment with folate and dexamethasone reduced the prostate cancer cell proliferations which should be considered in the treatment protocol for prostate cancer patients. Methotrexate can be used as a chemotherapeutic agent for treatment of prostate cancer.

PA-04 ***Characterization of Proteins Properties of Breast Cancer Proteins using Support Vector Machine***

Babu Baniya
Grambling State University

In the healthcare system, different clinical examinations were performed to identify the particular cause and corresponding cure. Each cause has distinct characteristics and needs special treatment of it. Mammography is one of the imaging technique that is used to detect breast cancer. However, approximately 20% of screening mammogram results dragonize as a false negative. It is due to high breast tissue density. This can make it hard to see small tumors in or around the dense tissue. In the next stage, patients require an invasive biopsy or surgical procedure to acquire tumor tissue for further medical assessment. Therefore, there is an urgent need to develop a novel, minimally invasive diagnostic strategy for the early diagnosis of breast cancer. We collected the different proteins from patients, analyzed them, and calculated the different statistics. These variables presented into two dimensions space how they contributed to improve the class label separability. The goal is to minimize the distance within a class and maximize the distance among the classes. We got impressive classification accuracy using a support vector machine (SVM).

PA-05 ***Continuous Monitoring of Animals Using IR Cameras - Disease Staging and Modeling***

Elia Brodsky
Pine Biotech

Advances in video technology have opened an opportunity to collect detailed datasets on the movement, fine-scale motion, social interactions, vocalizations and physiological responses of individual animals to environmental and pathological changes. In many areas of animal experimentation, improvements in our ability to collect such extensive and detailed data sets are outstripping our ability to analyze them. These diverse, complex and often high-dimensional data sets exhibit nonlinear dependencies and unknown interactions across multiple variables, and may fail to conform to the assumptions of many classical statistical methods. The field of machine learning provides methodologies that are ideally suited to the task of extracting knowledge from these data. The assessment of animal welfare and the emotional states that may reveal it can be highly subjective, and poor welfare is often only indicated by multiple interacting factors. Machine Learning and mathematical modeling can assist in monitoring such behaviors and extrapolating studied variation to explore important factors related to disease progression. Such methods can be extended to provide a diagnostic tool for mechanistic study of health and disease as well as the study of efficacy and safety for psychopharmacological drugs. Here we present TRKIR.com - a web application based on IR cameras and remote sensors for continuous monitoring of animal models of human disease. Digital readouts are extracted from the video using computer vision algorithms to provide specific behavioral and physiological measures that provide objective data throughout the course of studies. The data is then used to model disease progression and identify change points in the animal condition using Hidden Markov Models. The ability to

obtain continuous data non-invasively can help improve the analysis of disease progression and treatment and thus improve decision making during drug development.

PA-06 ***Changes of microbial community in the soybean plots treated with biochar and poultry litter***

Rosalie Calderon
Louisiana State University-Baton Rouge

Soil microbial communities play a significant role in nutrient cycling and soil conditioning that promotes plant growth and health. Four different soil amendments, which include untreated control, biochar at 11.2 Mg ha-1, poultry manure at 0.8 11.2 Mg ha-1 dry weight, and a combination of biochar and poultry manure, were applied annually for three consecutive years (2016, 2017, and 2018) in soybean plots in Red River Research Station, Bossier City, Louisiana. DNA were directly extracted from the soil samples in each treatment. Results showed that the composition and structure of bacterial communities varies from phylum to genus. In the phylum level, the top five dominant phyla were Proteobacteria, Acidobacteria, Bacteroidetes, Verrucomicrobia, and Gemmatimonadetes. The ten most abundant culturable genera throughout all the treatments were Blastocatella, Sphingomonas, Nitrospira, Desulfurellales, Nitrososphaera, Gemmatomonas, Bryobacter, Ramlibacter, Bacillus, and Paenibacillus. Interestingly, nitrifying bacteria of the phylum Nitrospirae known to promote plant growth and antimicrobial peptide-producing bacteria of the phylum Firmicutes were among the dominant phyla in the biochar and the poultry treated soils, respectively. Noteworthy, the microbial distribution in soil tends to become more even by repeated treatments of biochar for three consecutive years. This research provides the information about the significant changes of microbial community from the treatment of biochar, which could have an impact on soil-borne diseases of soybean.

PA-07 ***Characterization of HER-2 Targeted Modified Peptidomimetic***

Achyut Dahal
University of Louisiana at Monroe

Overexpression of HER2 is mostly associated with development of more aggressive and poorer prognosis in breast cancer and non-small cell lung cancer. HER2 has become an important target for cancer treatment therapy. We have designed several peptidomimetics that targets domain IV of HER2 and inhibit HER2 mediated dimerization. Among them one of the peptidomimetics, Compound 18 that exhibited antiproliferative activity in nanomolar range in HER2 overexpressing cancer cell line and showed the in-vivo anti-tumor activity in mice. Compound 18 has limited aqueous solubility and to enhance the solubility of compound 18, hydroxy moiety was added to the compound at proline residue of peptide sequence. Compound 18 with hydroxy moiety exhibited significantly increased aqueous solubility compared to compound 18. Molecular docking was used to evaluate the binding of the hydroxy-compound 18 with the extracellular domain of HER2. Surface Plasmon Resonance study also confirmed that the modified peptidomimetic binds to the extracellular domain of HER2 receptor. Modified peptidomimetic retained the antiproliferative activity towards HER2 positive cancer cell line and IC50 in BT474 cell lines and A549 cells line were found to be 54.19 ±1.0382 nM and 519 ±1.56 nM respectively. The thermal stability was evaluated using circular dichroism which showed that modified hydroxy compound 18 was stable at higher temperature. In summary, the modified hydroxy compound 18 showed better anti-proliferative activity on HER2 positive cancer cell line and increased water solubility compared to compound 18. The research reported in this presentation was supported by NCI of NIH under the grant number 1R15CA188225-01A1

Poster Abstracts

PA-08 ***Louisiana Biomedical Research Network Bioinformatics Training Program Authors: Justin Dore, Sahil Sethi, Julia Panov, Avi Titievsky, Elia Brodsky***

Justin Dore
Louisiana State University-Baton Rouge

Louisiana Biomedical Research Network Bioinformatics Training Program Authors: Justin Dore, Sahil Sethi, Julia Panov, Avi Titievsky, Elia Brodsky Starting from June 2018, the LBRN Bioinformatics Training Program was made available to over 200 participants on campuses throughout Louisiana and abroad. The online training program was supported by regular workshops conducted in New Orleans, Baton Rouge, Ruston, and Monroe. The program was launched from a small pilot at Loyola University of New Orleans and expanded to a summer program open to a small number of students and faculty at the LBRN Summer Research program. After the renewal in October 2018, an extended pilot was started at the Center for Computation and Technology at LSU. Small scale pilots were started at Grambling State University and Louisiana Tech University. Here we present our analysis of participation data, describing a growing interest in bioinformatics from students as well as new and established research labs. We also compare this data to our database of over 1000 participants that took online courses through the edu.t-bio.info platform. The majority of active participants are already involved in basic or translational research or plan to get involved in the near future. More than half of those surveyed have a non-engineering degree in the life-sciences in biology, biotechnology or biochemistry. Participants have explored informatics or similar courses as electives or external courses. Most hold a B.Sc. or higher. Similar to online participants, over half of the participants have over one year of experience in a life-sciences-related field and one-third have an interest in biomedical translational research. The LBRN training program continues to expand providing students, faculty and researchers with new resources and perspectives on the Bioinformatics field. Users remain receptive and engaged and the Pilot Program continues to demonstrate the applicability and utility of the Program content.

PA-09 ***An Integrative Genomics Approach for Associating Genetic Susceptibility with the Tumor Immune Microenvironment in Triple Negative Breast Cancer***

Jacob Elnaggar
LSUHSC New Orleans

Triple-negative breast cancer (TNBC) is the most aggressive form of breast cancer. It is a heterogeneous disease with poor prognosis. Contributing to the worse prognosis in TNBC is the higher rates of relapse and rapid progression to metastatic disease which is often lethal. With the exception of cytotoxic chemotherapy, there is currently no effective targeted therapies. Immunotherapy such as vaccines offer new opportunities for treatment of TNBC. But realizing the potential of immunotherapy and vaccination may require understanding the association between the tumor immune microenvironment and genetic susceptibility to TNBC. The objective of this exploratory study was to investigate the potential association between genetic susceptibility and tumor immune microenvironment in TNBC. We integrated information on genetic variants and genes associated with an increased risk of developing breast cancer with gene expression data from the Caucasian women diagnosed with the basal-like immune activated (N=54) and basal- like immune suppressed (N=60) subtypes of TNBC to discover and characterize immune modulated gene signatures, molecular networks and biological pathways enriched for genetic susceptibility variants. The investigation revealed immune modulated gene signatures, molecular networks and biological pathways enriched for genetic susceptibility variants. The discovered pathways included the role of BRCA1 in DNA damage response, hereditary breast cancer, aryl hydrocarbon receptor and molecular mechanisms of cancer signaling pathways. The investigation suggests the link between genetic susceptibility and the tumor immune microenvironment in TNBC and establishes putative functional bridges between genetic predisposition and immune modulated pathways.

Poster Abstracts

PA-10 ***Prediction of Protein-Carbohydrate Binding Residue using Stacking.***

Suraj Gattani
University of New Orleans

Carbohydrate-binding proteins play vital roles in many vital biological processes. Study of these protein-carbohydrate interactions, at the residue level, is useful in treating many critical diseases. Analyzing the local sequential environments of the binding and non-binding regions to predict the protein-carbohydrate binding sites is one of the challenging problems in molecular and computational biology. Existing experimental methods for identifying protein-carbohydrate binding sites are laborious and expensive. Thus, prediction of such binding sites, directly from sequences, using computational methods, can be useful to fast annotate the binding sites and guide the experimental process. Because the number of carbohydrate-binding residues is significantly lower than the number of non-carbohydrate-binding residues, most of the methods developed for the prediction of protein-carbohydrate binding sites are biased towards over predicting the negative class (or non-carbohydrate-binding). Here, we propose a balanced predictor, called StackCBPred, which utilizes features, extracted from evolution-driven sequence profile, called the position-specific scoring matrix (PSSM) and several predicted structural properties of amino acids to effectively train a Stacking-based machine learning method for the accurate prediction of protein-carbohydrate binding sites. We predict the binding at the residual level and the robustness of the proposed method has been confirmed through proper validation and testing.

PA-11 ***Machine Learning to Analyze the impact of Various Drug Treatments on Artificial Peripheral Nerve Tissues***

Hamid Karimian
Tulane University

The peripheral nervous system lacks a protective barrier separating sensitive nerve tissue from potentially toxic compounds circulating in the blood. As a result, drug-induced peripheral neuropathy (DIPN) is a common side effect that compromises treatment efficacy and prevents newly developed drugs from reaching market. Therefore, early detection of DIPN using in-vitro models is a valuable asset for both the pharmaceutical industry and the healthcare system. Our microphysiological 'nerve-on-a-chip' model [1] enables both high-resolution imaging and electrophysiological analysis of explanted peripheral nerve tissue grown in an in-vitro hydrogel environment. We attempt to segment the high-resolution images taken from these tissues to automatically quantify the morphological properties of myelin, axons, and Schwann cells using deep neural networks [2]. In this work, we perform learning representations for classification of electrophysiological waveforms generated by tissues to understand the influence of various drug treatments on their properties. Our data is electrophysiological signals taken from both tissue treated with known neurotoxic compounds and healthy controls. We design a novel deep learning architecture that learns interpretable representations of the signals and classifies them jointly. The proposed architecture includes an encoder part that can be trained unsupervised. In fact, the unsupervised part is a denoising Auto-encoder [3], which is adapted with LSTM layers in both encoder and decoder parts. We design a attention mechanism in the encoder to highlight the main features to obtain a better latent representation of the signals. In the supervised setting, the neuropathy labels of those tissues are provided during training. The trained model will make predictions on the following five treatments, Cisplatin, Bortezomib, Forskolin, Paclitaxel and Vincristine. The auto-encoder LSTM provides a new smooth representation of the signals that highlights the important features that help in recognizing the influence of neurotoxic drug treatments on electrical conduction in artificial nerve tissues. For the evaluation of the obtained representation we apply the traditional statistical analysis of the signals on this intermediate representation. We then quantify the absolute magnitude, signal onset velocity, peak velocity distributions, peak amplitude distributions, and integrated trace area of both the traditional and the

intermediate representations of the electrophysiological waveforms generated by cultures in each of these treatment groups. The results show that the trained representations are more expressive and informative for the traditional statistical analysis. This approach has the potential to alleviate the need for performing high-resolution imaging of actual animal nerve tissue, which is very labor-intensive, costly and time-consuming compared to automatic analysis of electrophysiological signals produced by artificial nerve tissue.

PA-12 ***Unifying Interpretation of Molecular HLA Typing Data Between the Solid Organ Allocation Match Run and Virtual Crossmatch***

Navchetan Kaur
Tulane University

Virtual crossmatch compares a transplant candidate's HLA antibody assay results to the HLA typing of the donor before an organ offer is accepted. While HLA typing is now performed by molecular methods, the United Network for Organ Sharing (UNOS) UNet system's match run that determines organ offers allows for entry of only two donor HLA antigens per locus and does not capture precise information about molecular HLA typing ambiguity. We have developed a web application to compute virtual crossmatch and the histocompatibility component of match run using molecular HLA typing data in IPD-IMGT/HLA nomenclature, using genotype lists or multiple allele codes. The application computes the probability of each potential conflicting HLA specificity, both at two-field allele and antigen level, based on high resolution allele frequencies from 26 US populations obtained from the National Marrow Donor Program. Our analysis of simulated donor-candidate compatibility assessments reveals the risks of conflicts for alternative less probable HLA antigens and alleles of the donor that cannot be listed in the UNet system. We detail cases where organ offers from the match run would be turned down upon examination of the molecular HLA typing data. We show how analysis of molecular HLA typing can be aided by a computational decision support system. The VIrtual CrossmaTch for mOleculaR HLA typing (VICTOR) application can be accessed at http://www.transplanttoolbox.org/victor. HLA typing ambiguity has been an unappreciated source of immunological risk in transplantation. This application automates translation of HLA data between antigen-based and molecular nomenclature systems to streamline virtual crossmatch in the short time available to accept organ offers. A standardized interpretation of molecular HLA typing data could be integrated in the match run along with epitope-based interpretations to conduct more comprehensive risk assessments.

PA-13 ***Accurate RNA Binding Protein Prediction using Advanced Machine Learning Techniques.***

Reecha Khanal
University of New Orleans

Identification of RNA Binding Proteins from sequence information alone is one of the challenging problems in computational biology. RBPs play crucial roles in several fundamental biological functions including transcriptional regulation of RNAs and RNA metabolism splicing. Existing Experimental techniques are time-consuming and costly. Thus, efficient computational identification of RBPs directly from the sequence can be useful to annotate RBP and assist the experimental design. Here, we put forward AIRBP, a computational sequence-based method, which utilizes features extracted from evolutionary information, physiochemical properties, and disordered properties to train a machine learning method designed using stacking, an advanced machine learning technique, to effectively predict RBPs. Moreover, to develop a robust classifier AIRBP is trained on the useful feature-subset identified by the evolutionary algorithm (EA). AIRBP attains Accuracy (ACC), F1-score, and MCC of 95.38%, 0.917, and 0.885, respectively, on the benchmark dataset, using a 10-fold cross-validation and ACC, F1-score, and MCC of 93.04%, 0.943, and 0.855, for Human test set; 91.60%, 0.942 and 0.789 for S. cerevisiae test set; and 91.67%, 0.953 and 0.594 for A. thaliana test set, respectively. Furthermore, AIRBP outperforms the state-of-the-art method, RBPPred, by 3.26%, 6.75% and

9.53% in terms of ACC, F1-score, and MCC, respectively, based on the benchmark dataset. Further, using independent test datasets, performance comparison based on average percentage of improvement, calculated over three different independent test sets, in terms of accuracy, F1-score, and MCC, AIRBP outperforms RBPPred by 4.76%, 3.58%, and 9.11%, respectively. Therefore AIRBP, can be applied to the prediction of RBPs directly from the sequence and provide valuable insight to treat critical diseases.

PA-14 ***Computational Modeling of Blood Metabolomics Biomarkers Addressing Racial Disparity***

Phillip Kilgore
Louisiana State University-Shreveport

Alcohol use disorders are among the most devastating disorders contributing to the global burden of disease (139 million disability-adjusted life-years). In the United States, over 8% of the population meets the criteria for alcohol dependence with alcohol-related problem cost exceeding 223 billion dollars, 88,000 deaths, and nearly 10,000 driving fatalities per year. Furthermore, alcohol contributes to over 200 diseases such as alcohol dependence, liver cirrhosis, cancers, and injuries. We hypothesize that a metabolite in the blood of alcohol dependent patients could associate with anxiety based heavy drinking patterns and/or race, which may play an important role in explaining Acamprosate efficacy in subpopulations of alcohol dependent patients. We developed a biomarker paradigm to apply metabolomics and clinical measurement data using computational modeling. We used these results to establish a prediction algorithm between drinking patterns and depression in Acamprosate efficacy. We are currently working to test whether racial disparity between Caucasian and African-American alcohol dependent populations influences the drinking pattern. The results from this study may suggest a prediction paradigm for the identification of a sub-population in alcohol dependence.

PA-15 ***Mechanistic Modeling of Evolutionary Dynamics for Phylogenetic Inference***

Christina Kolbmann
Southeastern Louisiana University

Fossilized birth-death (FBD) models enable researchers to co-estimate phylogenies, divergence times, and diversification parameters. These models can be time homogeneous, in which the model and parameters are expected not to change with time, or time heterogeneous, in which parameter values may be expected to vary. Many organisms have a discontinuous fossil record. In this talk, we will discuss modeling the fossil record of ants, which is known to be highly dependent on amber deposits, and the presence of sap-bearing plants. Amber deposits create discontinuous sampling, with intervals of presence and of absence. FBD models include extinct and extant data in the same process of diversification and sampling, allowing all data to be modeled as one process. Using a combined molecular and morphological dataset from the Formicidae, we will discuss using time-heterogeneous FBD models to estimate phylogenies, and how these models impact the estimated rates of speciation, extinction, and fossil sampling.

PA-16 ***Genomic Analysis for Virulence Determinants in Feline Herpes Virus Type-1 Isolates***

Andrew Lewin
Louisiana State University-Baton Rouge

Feline Herpes Virus Type-1 is a significant cause of ocular and respiratory disease in cats. Ocular lesions include painful ulceration and the possibility of permanent blindness. Twenty-five viral isolates were obtained from shelter-housed cats in seven separate locations in the USA and were fully sequenced using Illumina MiSeq. All host animals from which viral isolates were obtained were assigned a clinical disease severity score according to a defined standardized scheme, with ocular and respiratory clinical signs assessed separately. Ancestral character state reconstruction was performed in an attempt to infer the ancestral virulence trait associated

with each terminal sample in the phylogenetic tree. This analysis uses likelihood reconstruction by maximizing the probability the observed states of virulence would arise under a stochastic model of evolution. In an attempt to identify regions of the viral genome that may be associated with varying degrees virulence in the host, a genomic variant analysis was performed. A genome wide survey was used to determine if the allele frequencies of any single nucleotide polymorphism vary systematically as a function of virulence severity scores with the goal that identification of disease severity associated variants may subsequently reveal insights into the mechanisms underlying the variation in disease severity.

PA-17 ***G-Dash: A Genome Dashboard Integrating Modeling and Informatics***

Zilong Li
Louisiana Tech University

Genomics is a sequence based informatics science and a structure based molecular material science. There are few tools available that unite these approaches in a scientifically robust manner. Here we describe G_Dash, a web based prototype of a genomics dashboard, specifically designed to integrate informatics and 3D material studies of chromatin. G_Dash unites our Interactive Chromatin Modeling(ICM) tools with the Biodalliance genome browser and the JSMol molecular viewer to rapidly fold any DNA sequence into atomic or coarse_grained models of DNA, nucleosomes, or chromatin. As a chromatin modeling tool, G_Dash enables users to specify nucleosome positions from various experimental or theoretical sources, interactively manipulate nucleosomes, and assign different conformational states to each nucleosome. As an informatics tool, data associated with 3D structures are displayed as tracks in a genome browser. The exchange of data between informatics and structure is bi-directional so any informatics track can inform a molecular structure (e.g. color by function) and structure features can be displayed as informatics tracks in a genome browser(e.g. Roll, Slide, or Twist). As a sample application, models of the CHA1 promoter based on experimentally determined nucleosome positions are explored with G_Dash. Steric clashes and DNA knotting are observed but can be resolved with G_Dash's minimal coarse_grained model without significant variation in structure. Results raise questions about the interpretation of nucleosome positioning data and promoter structures. In this regard, G_Dash is a novel tool for investigating structure_function relationships for regions of the genome ranging from base pairs to chromosomes and for generating, validating and testing mechanistic hypotheses. http://dna.engr.latech.edu/~gdash/GDash-landing-page

PA-18 ***Mapping the Genomic and Epigenomic Interaction Landscape for the Discovery of Molecular Markers and potential Targets in Aggressive Prostate cancer***

Tarun karthik kumar Mamidi
LSUHSC New Orleans

Although prostate cancer (PCa) is the second most diagnosed and second leading cause of cancer related death among men in the US, not all men will die from the disease. The key challenge faced by clinicians is identifying men at high risk of developing aggressive PCa and discovery of molecular makers to guide treatment options. Alterations in the germline, somatic and epi-genomes play roles in PCa. However, to date, germline, somatic and epigenetic mutation datasets have largely been analyzed as separate endeavors in PCa research. The objective of this study was to investigate the possible oncogenic interactions and cooperation among genes containing germline, somatic and epigenetic mutations, and to discover molecular networks and biological pathways enriched for these genetic and epigenetic alterations in aggressive PCa. We addressed this knowledge gap by integrating germline mutation information from genome-wide association studies (GWAS) with somatic and epigenetic mutation information from The Cancer Genome Atlas (TCGA) using gene expression data on aggressive PCa as the intermediate phenotype. Preliminary results revealed that genes containing germline mutations also harbor somatic and epigenetic mutations and that these genes are

functionally related. Additionally, the analysis revealed molecular networks and biological pathways enriched for germline, somatic and epigenetic mutations demonstrating that integrative analysis is powerful approach for biomarker discovery and holds the promise of linking genetics with environmental perturbations in aggressive PCa.

PA-19 ***AMLGA: A Genetic Algorithm with Adaptive and Memory-Assisted Local Operators***

Yashwanth Karthik Kumar Mamidi
University of New Orleans

Genetic algorithm (GA) is a population-based random search method that mimics the action of evolution. With encoded chromosomes, like employing a binary alphabet, GA operates by discovering, prioritizing and recombining fitter building blocks or schemata that may end in a fitter chromosome. Crossover operation produces healthier offspring by rearranging the parent chromosomes and is treated as the core of GA. We present a novel mechanism of applying it at gene level locally to improve the performance of crossover. We store fitter gene schema by employing memory and emphasize on it at the time of crossover. We implement this memory-assisted native crossover in an adaptive manner by accessing the memory versus the section of the current parent to obtain superior offspring. Further, we add a recently proposed adaptive gene replacement operator that enhances the elites globally with their homolog fitter gene schema. We name the new hybrid Genetic Algorithm with Adaptive and Memory-assisted Local operators as AMLGA. On a set of 40 complex mathematical benchmark test problems, AMLGA exhibits competitive performance compared to the present state-of-the-art evolutionary algorithms. Being effective on scalable issues, we believe AMLGA will be effective in the era of massive information.

PA-20 ***Continuous Monitoring of Animals Using IR Cameras - Disease Staging and Modeling***

Joshua McKendall
Pine Biotech

Advances in video technology have opened an opportunity to collect detailed datasets on the movement, fine-scale motion, social interactions, vocalizations and physiological responses of individual animals to environmental and pathological changes. In many areas of animal experimentation, improvements in our ability to collect such extensive and detailed data sets are outstripping our ability to analyze them. These diverse, complex and often high-dimensional data sets exhibit nonlinear dependencies and unknown interactions across multiple variables, and may fail to conform to the assumptions of many classical statistical methods. The field of machine learning provides methodologies that are ideally suited to the task of extracting knowledge from these data. The assessment of animal welfare and the emotional states that may reveal it can be highly subjective, and poor welfare is often only indicated by multiple interacting factors. Machine Learning and mathematical modeling can assist in monitoring such behaviors and extrapolating studied variation to explore important factors related to disease progression. Such methods can be extended to provide a diagnostic tool for mechanistic study of health and disease as well as the study of efficacy and safety for psychopharmacological drugs. Here we present TRKIR.com - a web application based on IR cameras and remote sensors for continuous monitoring of animal models of human disease. Digital readouts are extracted from the video using computer vision algorithms to provide specific behavioral and physiological measures that provide objective data throughout the course of studies. The data is then used to model disease progression and identify change points in the animal condition using Hidden Markov Models. The ability to obtain continuous data non-invasively can help improve the analysis of disease progression and treatment and thus improve decision making during drug development.

Poster Abstracts

PA-21 ***Analysis of a high-throughput transcription factor knockdown RNA-seq dataset using LSU's HPC system.***

John McKowen
Louisiana State University-Baton Rouge

The project is to analyze the RNA-seq data from a high-throughput gene knockdown experiment. In this experiment, 483 transcription factors were knocked down using RNAi in a particular Drosophila melanogaster cell type, S2 cells. Each instance of knockdown will produce a gene expression dataset (2-10Gb), with replicates and controls this amounts to 1900 RNA-seq datasets. The HPC system is used to download and process each dataset into normalized alignment signal (~5Mb). At this stage, the data is transferred to a local machine for analysis. This data was then clustered and analyzed to answer a number of biological questions related to gene expression and transcription factor regulatory networks. The two main goals of this anaylsis are to 1: Cluster TFs into groups with similar gene regulation behavior. 2: Identify TFs which perturb specific genes.

PA-22 ***Three-Dimensional Ideal Gas Reference State based Energy Function for Flexible Proteins***

Avdesh Mishra
University of New Orleans

Flexible proteins (FPs) lack single stable 3D shapes. Because FPs structure constantly fluctuates between several possible conformations, they have challenged the classical structure-to-function paradigm. In order to achieve a comprehensive understanding of the behavior of FPs, prediction of single structural conformation is often insufficient and therefore, requires an ensemble of predicted structures. Development of an effective conformational ensemble generator relies heavily on two key components, energy function, and conformational sampling algorithms. In this study, we have developed an energy function for flexible proteins (flexEgy), which utilizes knowledge obtained from multiple models of NMR structures and three-dimensional (3D) ideal gas reference state formulated from three distinct hydrophobic-hydrophilic interactions of amino acids. The three distinct interactions including hydrophobic-hydrophilic, hydrophobic-hydrophobic and hydrophilic-hydrophilic are controlled using 3D optimized values of alpha. Likewise, the contributions of each of the interaction group are controlled by the optimized values of beta. Furthermore, we establish the optimization and independent test decoy set for flexible proteins which are obtained using single point angular mutation of torsion angle guided by Ramachandran distribution. The values of alpha and beta are optimized through an evolutionary algorithm (EA) on the optimization dataset. The test on independent test dataset indicates that flexEgy provides outstanding performance in selecting native structures and in z-score which, shows how well the native is separated from the decoy structures. Therefore, the proposed energy function can be useful in effective near-native conformation selection and consequently can be applied in conformational ensemble generation of flexible proteins.

PA-23 ***Optimizing the Alu Detection Program PolyDetect***

Grayce Mores
Louisiana State University-Baton Rouge

Retrotransposons are repetitive fragments of genetic information that can change locations within a genome by copying and pasting themselves at novel insertion sites. The short-interspersed element Alu is a currently active retrotransposon within the primate species. Alu insertions that are identical by descent can be used in phylogenetic research to show common ancestry and inheritance among individual primates. Vallmer Jordan created the program polyDetect to identify novel Alu insertions within primate genomes using unassembled short read data. This project focuses on optimizing the polyDetect program to be shorter, easily readable, and user friendly. The original nine program files that make up polyDetect were condensed and organized into

three program files, removing over 2,000 lines of unnecessary code. A Graphical User Interface (GUI) was created to allow for simplified user input and to create a user-friendly experience. This new program, titled polyDetectV2, ran successfully and produced the same output file as the original polyDetect. Future work includes further testing of polyDetectV2 using several organisms' whole genome sequences and comparing the results to the results from Jordan's Papio research using the original polyDetect.

PA-24 ***The Use of Next-Generation Sequencing Techniques to Determine Biomarkers for Developmental Toxicity - A Case Study using Retinoic Acid***

Matthew Overturf
University of Louisiana at Monroe

Adverse outcome pathways (AOPs) have become a popular discussion with the push to implement the 3Rs (reduce, refinement, and replacement of animal tests) in environmental toxicology. The central theme of AOPs is to link molecular initiating events to population-level responses. The fish early-life stage (FELS) test, an environmental risk assessment tool, is labor and resource intensive with apical endpoints of survival and growth. To be able to predict these apical endpoints from embryo exposures would greatly enhance the implementation of AOPs for such analyses. Key molecular events of embryonic development must be fully understood. Therefore, the retinoic acid signaling pathway, responsible for regulation of neurogenesis, cardiogenesis, body axis extension, and development of the foregut and eye, may provide biomarkers for use in predicting survival and growth endpoints of FELS. RNA-seq technology was utilized to determine differentially expressed genes of Japanese medaka (Oryzias latipes) following exposure to 0.1 μM (30.0 μg/L) retinoic acid. The number of differentially expressed genes was determined to be 442 using the differential expression tool DESeq2. Many of those genes are involved in embryonic, heart, brain, and bone development, angiogenesis, and hatching. The abnormal expression of these genes during embryonic development may be useful in predicting adverse outcomes for future developmental time periods such as that of the FELS.

PA-25 ***Transcriptomic and epigenetic aberrations in Angelman syndrome, a multi-omics approach***

Julia Panov
University of Haifa

Ubiquitin-protein ligase E3A (UBE3A) is one of the E3-ligases in the ubiquitin-proteasome system. Alterations in Ube3a levels, either deletion or overexpression, culminate in severe neurodevelopmental disorders such as Angelman syndrome or Autism spectrum disorder. Using bioinformatics tools we analyzed publicly available transcriptomic and epigenetic data together with transcriptomic data generated in our lab to elucidate transcriptional and epigenetic mechanisms governed by UBE3A.

PA-26 ***Prediction of Hierarchical Classification of Transposable Elements using Machine Learning Techniques***

Manisha Panta
University of New Orleans

Motivation: Transposable Elements (TEs) or jumping genes are the DNA sequences that have intrinsic capability to move within a host genome from one genomic location to another - genomic location can be either same or different chromosomes. Studies show that TEs have a role in genome function and evolution as their presence can modify the functionality of genes and increase the size of the genome. Thus, proper classification of the identified jumping genes is important to understand their role in germline and somatic evolution. While there are computational methods that perform either binary classification or multi-label classification of TEs, few studies have focused on the hierarchical classification of the elements. The state-of-the-art method has used a neural network for hierarchically classifying TEs. In this regard, we explored

different machine learning approaches for hierarchical classification of transposable elements. Results: Firstly, we compared state-of-the-art machine learning models, based on the Multi-Layer Perceptron (MLP) neural network algorithm, for two hierarchical classification strategies: non-Leaf Local Classifier per Parent Node (nLLCPN) and Local Classifier Per Parent Node and Branch (LCPNB) using a variety of machine learning (ML) models. We experimentally demonstrated that LCPNB classification strategy performs statistically better with most of the models. The comparative results with different learning algorithms showed that optimized Support Vector Machine (SVM), with a radial basis function (RBF) kernel, has higher balanced-accuracy for both classification strategies. We found that the balanced-accuracy (using the hF measure) for the LCPNB strategy with the SVM model on benchmark hierarchical datasets of TEs from the Plant Genome and Systems Biology (PGSB) was 0.905; for TEs from the Repbase database, hF was 0.886, and when TEs from both databases were combined, hF was 0.881. We then proposed a new method, ClassifyTE, based on an augmented stacking approach to predict the hierarchical class of a transposable element using the LCPNB strategy. Our proposed model outperforms existing state-of-the-art Multi-Layered Perceptron method. Furthermore, ClassifyTE outperforms all the existing machine learning approaches for hierarchical classification of TEs. Using this approach, the balanced-accuracy (hF) for PGSB, Repbase, and the mixed datasets were found to be 0.915, 0.882 and 0.892 respectively, higher than our previous results using the SVM model.

PA-27 ***Design of a Grafted Multicyclic Peptide for Immunomodulation***

Pravin Parajuli
University of Louisiana at Monroe

T-cell activation by an antigen presenting cell (APC) requires binding of APC and T-cells. Protein-protein interaction (PPI) of co-stimulatory molecules in APC and T-cells helps in attachment of T-cells to APC and amplify the signaling for T-cell activation. PPI such as CD2(T-cells)-CD58(APC) interaction is important for T-cell activation and generation of immune response. Expression of CD58 in antigen presenting cells (APC) is found to be increased in autoimmune diseases. Inhibition of CD2-CD58 protein-protein interaction (PPI) is known to hinder activation of T-cells and have promising role in slowing the progression of autoimmune diseases. From the alanine scanning studies of previously designed peptides, we obtained a potent CD2-CD58 PPI inhibitor peptide, AS1, having IC50 of 3 nM in lymphocyte-epithelial cell adhesion assay. In this study, we are evaluating the potency and stability of a modified peptide using multicyclic framework grafting strategy. The important amino acid residues in peptide AS1 were grafted on to a multicyclic peptide sunflower trypsin inhibitor (SFTI) framework. SFTI-AS1 inhibited adhesion between OVCAR-3:T-cells and HFLS-RA:T-cells with IC50 of 23.34 ±4.62 nM and 37.23 ±1.02 nM respectively. Flow cytometry, molecular docking, and SPR analysis data reveled that SFTI-AS1 binds to CD58 protein in CD2 binding region. From NMR studies, we found that the peptide might have multiple conformation. Molecular modeling was done to study the possibility of multiple conformation of peptide. SFTI-AS1 exhibited exceptional thermal, chemical and enzymatic stability. In summary, muticyclic grafting of peptides increased the stability of peptide and such grafted peptides can be used as therapeutic agents to modulate autoimmune diseases. Keywords: Sunflower Trypsin Inhibitor, Multicyclic Grafted Peptide, Protein-Protein Interaction

PA-28 ***Prediction of DNA binding residues from sequences using machine learning methods.***

Pujan Pokhrel
University of New Orleans

Thousands of proteins are known to bind to DNA; for most of them the mechanism of action and the residues that bind to DNA, i.e. the binding sites, are yet unknown. Experimental identification of binding sites requires expensive and laborious methods such as mutagenesis and binding essays. Hence, such studies are not

applicable on a large scale. This work presents a highly reliable computational technique for predicting DNA-binding function at the level of protein-DNA complex structures, rather than low-resolution two-state prediction of DNA-binding as most existing techniques do. In this work, a novel algorithm for feature selection based on Genetic Algorithm has been investigated. Likewise, various machine learning methods like Support Vector Machines, Logistic Regression, k Nearest Neighbors, Neural Networks have been used to predict the DNA binding residues and their performance is measured. Finally, the methods are combined using a novel stacking-based framework and the results are compared to the other state-of-the-art approaches.

PA-29 ***A Machine Learning Approach towards Functional Morphology Study***

Pooja Pun
University of New Orleans

Functional morphology involves the study of the relationship between the physical structures and the functions of the various parts of an organism, and has been a topic of interest since the 1950s. Morphology-to-performance relationships are well explained for individual performance traits, particularly in specifies that are highly specialized for conducting the specific biological task. However, an organism's morphology and physiology are altered due to multiple conflicting environmental factors. We currently lack a proper understanding of the multivariate morphology-to-performance relationships, at least in part because of logical challenges in measuring multiple performances on the same individuals. In this study, we developed a machine learning method to study the functional morphology of lizard species. First, the missing values in the dataset were replaced using a two-step k-Nearest Neighbor (kNN) based approach. Then, the relevant features were selected using an evolutionary algorithm. Different types of classifiers were used to classify over 30 lizard species based on their relevant functional and morphological data. Then, the best performing classifiers were combined, in a stacking manner, to obtain a more robust classification framework. In a subsequent step, Stacking-based regression framework was established to produce novel predictions of unmeasured performance traits. The comparison of stacked prediction models and their predictive-performance based on cross-validation indicates that this approach can be successfully applied to advance the study of functional morphology.

PA-30 ***Investigation of Intertidal Metazoan Biodiversity within Previously Oiled Sheared and Intact Coastal Margins in Barataria Bay***

Patrick Rayle
Louisiana State University-Baton Rouge

The effects of margin shearing of marsh islands on the biodiversity of intertidal metazoan communities was investigated using a metabarcoding approach. Many of these intertidal metazoans are meiofauna which have short range dispersal, are largely sedentary, and spend their entire lives in the marsh. Therefore, these animals are good indicators of changes in marsh health; one limitation of using meiofauna as bioindicators is that they are difficult to identify using traditional methods. Following the immediate devastation of Louisiana coastal marshes during the 2010 Deepwater Horizon Oil Spill, the physical structure of certain marshes was weakened during the long-term due to a reduction in belowground biomass. During subsequent hurricanes, the weakened marsh edges were sheared away, leaving sites which erode faster than the surrounding marsh. In 2017, soil was collected from selected sheared sites as well as nearby intact sites. DNA was extracted from these soil samples and then amplified using Earth Microbiome Project 18S small ribosomal subunit primers. Resultant amplicons were then sequenced on the Illumina Hiseq platform. Biodiversity was then analyzed using the SILVA database in QIIME 2. Sheared sites had significantly reduced biodiversity and different community structures compared to intact sites.

Poster Abstracts

PA-31 ***OmicsLogic: Practical Bioinformatics Modules for Online Training***

Sahil Sethi
Pine Biotech

Over the past couple of years, we conducted a number of educational programs at university campuses around the world, working with faculty, researchers and students to build a collection of 6 modules that cover the most important data types, methods and applications in bioinformatics. These include: Introduction to bioinformatics, Genomics, Transcriptomics, Epigenomics, Metagenomics and an Introduction to Machine Learning. The modules are available through our online portal and help establish a good foundation that one can build upon with practical assignments and projects. In this poster, we present the developed materials and an innovative platform that helps link the theoretical aspect of learning with a practical, hands-on analytical platform developed at University of Haifa, Israel. The modules rely on public-domain datasets acquired through online databases like NCBI. Project materials are designed to help transition online learning into applied research by following structured tutorials and getting to understand the logic of analytical methods. After the logic has been applied in the context of a tutorial and enough practice has been given each topic, the user transitions to a next level account that allows new datasets to be analyzed on the specially-designed HPC infrastructure.

PA-32 ***Impact of Intermittent Palatable Solution Access on Alcohol Drinking and Associated Global Protein Expression Changes in the Rat Brain Reward Circuitry***

Krishna Shah
Xavier University of Louisiana

Previous studies in our lab have shown that intermittent access to palatable diet (High-Fat Diet) reduces alcohol drinking in rats, which has potential therapeutic implications in the management of alcohol use disorder. However, the underlying mechanisms are unknown. It is also unclear if the reduction in alcohol drinking could be attributed to the caloric density and/or palatability of the high-fat diet. In the present study, we investigated the impact of a non-caloric sweetener on alcohol drinking (two-bottle choice drinking) and global protein expression (proteomics and bioinformatics) in the rat brain reward circuitry. Male Long Evans rats, matched for body weight, water, and food intake, received intermittent (24 hrs twice a week on Tue and Thru) access to the saccharine solution or normal water (controls) for two-weeks. Chow and water were always provided ad libitum. Following initial two-weeks of intermittent saccharine solution exposure, rats were allowed to drink ethanol (20% v/v) on the water-only access days (Mon, Wed, Fri) while intermittent saccharine solution cycling was continued. A significant reduction ($p<0.05$) in alcohol drinking was observed in the rats that were exposed to the intermittent saccharine solution. Brains were isolated at the end of the study and quantitative proteomic analysis (tandem mass tag labelling with triple labeling for each sample) of brain homogenates was performed to assess global protein expression changes in the brain reward circuitry regions (striatum). Over 1700 proteins were quantified; however, no consistent significant changes in global protein expression was observed in the striatum. Further studies assessing other brain regions within the brain reward circuitry are needed to further understand the possible neurobiological mechanism regulating alcohol drinking as a result of palatable diets exposure.

Poster Abstracts

PA-33 ***Stability Assessment, Pharmacokinetic Evaluation, and Anti-tumor Studies of a HER-2 Targeted Peptidomimetic***

Leeza Shrestha
University of Louisiana at Monroe

HER2 protein overexpression has been observed in 3-38% of non-small cell lung cancer (NSCLC) while HER2 gene amplification is found in 10-20% of NSCLC. Co-expression of EGFR and HER2 results in poor survival rate in NSCLC patients Therefore, HER2 is not only a biomarker but also an important therapeutic target for cancer treatment. We have designed several peptidomimetics that targets domain IV of HER2 and inhibit HER2 mediated dimerization. One of the peptidomimetics, Compound 18 exhibited antiproliferative activity in nanomolar range in HER2 overexpressing lung cancer cell lines. Molecular docking studies showed the binding of compound 18 to domain IV of HER2 receptor. Stability of compound 18 was evaluated at higher temperature and various pH. In addition, metabolic stability was assessed by incubating compound 18 with human liver microsomes in vitro. Pharmacokinetics of compound 18 were studied in mice with single dose of compound 18 injection via tail vein and the result showed the half-life to be around 27 hours. Based on its high stability and longer biological half-life, the pharmacokinetic results support the therapeutic potential of compound 18 for cancer treatment. In vivo activity of compound 18 on pulmonary tumor were evaluated using an experimental metastasis lung cancer mouse model. The result showed significant difference in tumor growth rate between control and treated groups. In summary, the designed peptidomimetic exhibited potent in vitro as well as in vivo anticancer activity. The funding for this project was from grant 1R15CA188225-01A1 from NCI/NIH.

PA-34 ***Respiratory Syncytial Virus Disease is Associated with Altered Gut Microbiota***

David Siefker
Louisiana State University-Baton Rouge

Rationale Respiratory syncytial virus (RSV) is the number one cause of lower respiratory tract infections in infants. There are still no vaccines or specific antiviral therapies against RSV, mainly due to the inadequate understanding of RSV pathogenesis. Recent studies suggest a role for gut microbiota in determining RSV disease severity. Objectives To determine if gut microbiota is associated with RSV disease severity. Methods We enrolled 61 infants hospitalized with RSV. Patients that were admitted to the pediatric intensive care unit (PICU) were classified as having severe disease, while patients admitted to the general ward were considered to have moderate disease. In addition, healthy controls without respiratory or gastrointestinal tract diseases were enrolled. We evaluated the composition of gut microbiota within 24 hours of enrollment using 16S metagenomic sequencing of stool samples. Measurements and Main Results Patients with severe RSV disease had significantly lower alpha diversity of the gut microbiota compared to patients with moderate disease or healthy controls. Beta diversity was significantly different between all RSV patients (severe and moderate) compared to controls. Relative abundance analysis showed that S24-7 was significantly enriched in infants with severe RSV disease. Conclusions These data indicate that a unique gut microbial profile is associated with severe RSV disease. More studies are needed to determine if gut microbiota could be used as a biomarker to identify patients at risk for developing severe RSV disease and whether the differences we observed in gut microbiota are the cause or consequence of RSV disease.

Poster Abstracts

PA-35 ***Molecular Docking Studies of Cyclic Peptides for Targeting HER2 Protein-Protein Interaction in Lung Cancer***

Sitanshu Singh
University of Louisiana at Monroe

Non-small-cell lung cancer (NSCLC) is the most common and is a leading cause of cancer deaths worldwide. Approximately 18-33% of NSCLC tumors show human epidermal growth factor receptor 2 (HER2) overexpression, indicating the importance of HER2 in NSCLC. The inhibition of protein-protein interactions of HER2 with other EGFRs ultimately leads to control and inhibition of tumor growth in lung cancer patients. To improve the stability of the peptide, cyclic peptidomimetics were grafted into the cyclic peptide frameworks of sunflower trypsin inhibitor (SFTI). The structure-activity relationship of different peptides was studied by NMR and molecular modeling. To optimize the grafting of the peptide in SFTI, computational docking method was used. Using the crystal structure of HER2 with 620 amino acid residues of the extracellular domain that represents domain IV of HER2, different peptides were docked using AUTODOCK software to the region around domain IV. The designed multicyclic peptides were evaluated for their ability to inhibit HER-2 dimerization, with one candidate, showing antiproliferative activity in the nanomolar range (IC50 ~ 280 nM) in various HER2+ cancer cells. Surface plasmon resonance studies suggested that compound binds to the HER2 extracellular domain and in particular to domain IV. Proximity ligation assay and western blot analysis suggested that compound is able to block HER2: HER3 interaction and inhibit phosphorylation of kinase domain of HER2 respectively. Our future studies will address the validation of the results in HER2+ over-expressive lung cancer animal models. The research is funded by the National Institutes of Health, grant number R15CA188225.

PA-36 ***STRING Data Mining of GWAS Data in Canine Hereditary Pigment-Associated Deafness***

George Strain
Louisiana State University-Baton Rouge

STRING Data Mining of GWAS Data in Canine Hereditary Pigment-Associated Deafness Maria Kelly-Smith and George M. Strain Comparative Biomedical Sciences, School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA 70803 Most canine deafness is linked to white pigmentation caused by the piebald locus, shown to be the gene MITF (melanogenesis associated transcription factor), but numerous studies have failed to identify a cause. The coding regions of MITF are not mutated deaf dogs, leading us to pursue genes acting on or controlled by MITF. We have genotyped 501 DNA samples from deaf and hearing Australian cattle dogs, Dalmatians, and English setters, breeds with high deafness prevalence. Genome-wide significance was not attained in any of our analyses, but we did identify several suggestive associations. Genome-wide association studies (GWAS) in complex hereditary disorders frequently fail to identify causative gene variants, so advanced data mining techniques are needed to extract information to guide future studies. STRING diagrams (https://string-db.org) are graphical representations of known and predicted protein-protein interactions identifying documented relationships between gene proteins based on the scientific literature, to identify functional gene groupings to pursue for further scrutiny. The STRING program predicts associations at a preset confidence level and suggests biological functions based on the identified genes. Starting with (1) genes within 500 kb of GWAS-suggested SNPs, (2) known pigmentation genes, (3) known human deafness genes, and (4) genes identified from proteomic analysis of the cochlea, we generated STRING diagrams that included these genes, plus associated genes predicted by the program. We then reduced the number of genes by excluding genes with no relationship to auditory function, pigmentation, or relevant structures, and identified genes that warrant further investigation.

Poster Abstracts

PA-37 ***Through the Logrus: Bioinformatics Internship at NCGR***

Joseph Straub
Louisiana Tech University

NCGR, the National Center for Genome Resources, is a non-profit that conducts next-generation sequencing (NGS) and bioinformatics research. I had the opportunity to be one of 18 students selected from across the country to participate in their 2018 summer internship funded by NM-INBRE. Before the 6-week program began, participants were required to take a short edX course on Linux Bash. While at NCGR, our wranglers-NCGR scientists-taught the fundamentals of NGS, genome assembly, variant calling, RNA-seq, transcriptome analysis, metagenomics, and about careers in bioinformatics. This internship gave participants a chance to learn hands-on in terminal emulators working with real data hosted on NCGR's data center. Participants were able to interact with career bioinformaticians and build their understanding of the field beyond just the rote learning of commands and workflows. At the conclusion of the program, participants were required to present before the class and the wranglers about what they learned from the program, and if they brought their own data, what findings they gathered from that data. Personally, the NCGR Summer Internship gave me a chance to learn how to combine my love for computers with my interest in biology, opened me up to new experiences in the Land of Enchantment, and left me with memories that will last a lifetime.

PA-38 ***TMB-iBIOMES: A Database of All Atom Simulation and Analysis for Nucleosomes***

Ran Sun
Louisiana Tech University

Nucleosomes are the fundamental building blocks of chromatin, the biomaterial that houses the genome in all higher organisms. A nucleosome consist of 145-147 base pairs of double stranded DNA wrapped 1.7 times around eight histones. There are approximately 100 atomic resolution structures of the nucleosome available from the protein data bank. Collectively they explore histone mutations, species variations, binding of drugs and ionic effects but only a few sequences of DNA. Given a four-letter code (A, C, G, T) there are on the order of 4^147 ~ 10^88 possible sequences of DNA that can form a nucleosome. Exhaustive studies are not possible. Here we introduce a database containing over 20 microseconds of all atom molecular dynamics simulations for over 500 different realizations of the nucleosome. Our iBIOMES-Lite based database serves as a reference for comparative studies and future on-demand simulations of nucleosomes. For every simulation we present the Amber formatted topology and coordinate input files, NAMD formatted output, log and trajectory files, and RMSD and DNA helical parameter analysis data. Closely related simulations are grouped together. A summary analysis is provided for each group, and a meta analysis of all simulations is provided. All data can be easily navigated in a file browser or iBIOMES Lite web browser format or downloaded with command line tools. Results indicate that the workflow and simulation protocol developed are robust, that DNA sequence can affect the structure and dynamics of nucleosomal DNA at some locations but not others, and that sequence differences can be observed in 10's of nanoseconds.

PA-39 ***Role of different macrophage phenotypes for survival of Chlamydia trachomatis***

Illya Tietzel
Southern University at New Orleans

Chlamydia trachomatis is bacterial pathogen causing the highest numbers sexually transmitted diseases in the USA. Albeit infection with Chlamydia is treatable with antibiotics, many sequelae such as inflammatory tissue damage, scarring, pelvic inflammatory disease, ectopic pregnancy or infertility occur. Different phenotypes of macrophages such as classically activated, alternatively activated, and resident resting macrophages have been described recently. Classically activated macrophages, with their anti-microbial effector mechanisms,

are known to be involved in acute inflammatory processes during the course of infection. In contrast, alternatively activated macrophages contribute to tissue repair at sites of wound healing, and have reduced bactericidal functions. Their role in pathogenesis of Chlamydia is not understood. It is hypothesized that they differ in their ability to support growth of Chlamdyia. To address this question, macrophages derived from CD14-positive monocytes isolated from peripheral blood mononuclear cells were treated with interferon-_ or interleukin-4 to produce classically (CA m_) or alternatively activated macrophages (AA m_), respectively. Confocal microscopy of chlamydial inclusions and quantification of infectious yields revealed better pathogen growth and development in AA m_ than CA m_, which correlated with the reduced expression of indoleamine 2,3-dioxygenase, a known anti-chlamydial effector of the host. Furthermore, AA m_ secreted higher amounts of anti-inflammatory interleukin-10 compared to CA m_, characteristics that indicate its suitability as host to Chlamydia trachomatis. The data suggest that IL-10 produced by infected AA m_ attenuated the anti-chlamydial function of CA m_ with growth recovery observed in infected CA m_ in the presence of infected, but not mock-infected AA m_. Treatment of infected CA m_ with IL-10 treatment promoted better chlamydial growth.

PA-40 ***Changes in Water Microbiota Structure Along the Rio Piedras River Due to Anthropogenic Impacts***

Eduardo Tosado Rodriguez
Univeristy of Puerto Rico

The Rio Piedras river is the most important watershed in the urban area of San Juan, Puerto Rico. Throughout years, this watershed has been anthropogenically impacted by urban development and contaminants. Currently, only culture dependent methods are used to evaluate water microbial diversity and quality. We hypothesized that the structure and diversity of the water bacterial communities changed according to the anthropogenic impact levels. To test our hypothesis, we aimed to perform a molecular 16S rDNA profiling to a watershed gradient using next-generation sequencing. Water samples were collected on the river in four sites over an urbanization gradient. Water was filtered, gDNA was extracted from the filters and 16S rDNA V4 region genes were amplified and sequenced by Illumina MiSeq. Data analyses were done using QIIME1 with the SILVA database as reference. A total of 29 phyla and 213 genera were found, in varying complexity along river sites. The less impacted site showed only 5 phyla with a dominance of Proteobacteria and Actinobacteria the medium impact sites had 15 phyla and a dominance of Chlorobi, Actinobacteria and Bacteroidetes while the most impacted site had a higher diversity and number of phyla (26) with a dominance of Proteobacteria, Firmicutes (Faecalibacterium) and Bacteroidetes (Bacteroides) - fecal matter derived taxa. Richness was significantly higher in the most impacted site. Also, we found 4 Enterobacteriaceae genera including Enterobacter and Yersinia, in the high impact sites. With a 16S rDNA approach, we were able to define the varying and dynamic microbiota of this urban watershed to a much more detailed depth, revealing association of bacterial diversity to anthropogenic impact. We found an unprecedented diversity of microbial communities in water samples directly linked to an increase in human impact of overall environmental and enteric bacteria.

PA-41 ***Unifying Phenotypic and Molecular Data for Phylogenetic Estimation***

Tyler Tran
Southeastern Louisiana University

A phylogenetic tree is used to represent the evolutionary relationships between different species as they diversify overtime. Over 12,000 described ant species are scattered across the globe. Due to their ecological importance and relevance, much work has been performed on the evolution of this group. There are many molecular datasets, and fossil collections which can. provide substantial data that can be used for phylogenetic methods. Using these extant and extinct records of ants will provide powerful insight on the evolutionary

forces that affect and have affected this species as well as the many other species interacting with the ants. Previous studies have demonstrated conflict between molecular and morphological data, in terms of the phylogeny of ants, and the true origin time of ants. By combining both types of data into a single analysis, we may provide a more exact origin time of the species. To estimate a phylogeny requires estimation of values for unknown parameters that we believe best model the process of evolution. The Fossilized Birth-Death Model is a Bayesian 'total evidence' phylogenetic analysis that jointly models extant and fossils species, combining morphological and molecular data as well as stratigraphic ranges. The FBD model will run using RevBayes and is calibrated based a constant diversification rate as well as a continuous stratified time scale. This analysis will be one of the largest FBD analyses ever performed, and will include over 700 ant species.

PA-42 ***The Regulatory Roles of ncRNAs in the Innate Immune Response to Pneumovirus Infection***

Ifeanyi Uche
Louisiana State University-Baton Rouge

Background: Human metapneumovirus (HMPV) is a clinically important single stranded RNA pneumovirus that is one of the major causes of lower respiratory infections in babies and infants worldwide. Small noncoding RNAs, including micro RNAs (miRNAs) represent a large family known to play a key post-transcriptional role in gene expression. Despite their importance in infectious diseases, the role of miRNAs in respiratory viruses remains largely understudied. Rationale: We have recently identified that HMPV induces different types of miRNAs in infected cells. However, the molecular mechanisms regulating the antiviral response to HMPV are largely unknown. Given the extensive impact on human health, studies are needed to identify the mechanisms responsible for the HMPV-induced disease. Methods: We used an in vitro model of pneumovirus infection in human epithelial cells. Quantification of gene expression was determined by qRT-PCR assays. Results: We observed that the expression of miRNA suppressed the antiviral response against HMPV by targeting the pattern recognition receptors in the host infected cells. Decreased viral gene expression was also observed. Significance: These studies contribute to a better understanding of the molecular mechanisms by which miRNAs modulate the antiviral immune response against respiratory viral infections, which is critical to design future therapeutic approaches against this highly relevant viral infection.

PA-43 ***The probiotic microbiota of the Caribbean Mauby, a fermented tropical wood beverage***

Brayan Vilanova
Univeristy of Puerto Rico

Microbial species derived from fermented foods, are able to colonize the gastrointestinal tract. As antibiotics and processed foods have a detrimental effect on the human microbiome, it is important to identify new fermented food sources and beverages that could benefit us. Mauby and Tepache are tropical fermented beverages made from wood and pineapple respectively. We hypothesized they these beverages likely possess probiotic bacteria. To test our hypothesis we prepared both tropical beverages in a controlled environment and produced a robust microbiota study comparing them. We prepared two mauby beverages, one was made with a tree wood (C. elliptica) from Puerto Rico (PR), and another was made with a vine (G. lupuloides) from the Dominican Republic (DR), using two different wood sources (cortex and medulla). The Tepache drink was prepared using pineapple rind (Ananas comosus). The preparations were done in a mini-fermenter at 28°C for ~5-6 days. Genomic DNA extractions were followed by amplification of 16S V4 region and sequenced with Illumina MiSeq. Data analyses was done in QIITA and QIIME using the SILVA database as reference. We found a variety of at least 26 different Lactobacilli species. We found a significant dominance of Oenococcus oeni, Lactobacillus mali in PR mauby and a dominance of L. ghanensis, L. casei, Lactococcus sp and L. uvarum in the cortex-derived DR mauby. Mauby made with medulla did not yield a specially probiotic rich content. Tepache was enriched with L. Paracasei and L. fermentum. Our preliminary data suggests that mauby and tepache are

probiotic-rich beverages, containing a variety of Lactobacillaceae OTUs as well as other non-conventional lactic acid bacteria, beneficial for human consumption.

PA-44 ***Integrative Analysis of the Genomic and Epigenomic Interaction Landscape in Triple Negative Breast Cancer***

Jiande Wu
LSUHSC New Orleans

Triple negative breast cancer (TNBC) is the most aggressive and lethal form of breast cancer. Currently, no molecularly targeted therapeutic agents have been clinically approved for TNBC. Cytotoxic chemotherapy remains the only effective therapeutic modality. Therefore, there is an urgent need for the discovery of novel clinically actionable molecular markers and targets for the development of novel therapeutics. Advances in High-Throughput genotyping, genomic and epigenomic profiling and the recent surge of next generation sequencing technologies have enabled discovery of germline, epigenetic and somatic mutations. However, these information have not been integrated to define the genomic and epigenomic interaction landscape in TNBC. The objective of this study was to investigate the possible oncogenic interactions among and between genes containing germline, epigenetic and somatic mutations in TNBC and to identify molecular networks and biological pathways enriched for these genetic and epigenetic alterations shaping the TNBC phenotypes. We addressed this objective by integrating germline mutation information from genome-wide association studies (GWAS) with somatic mutation and epi-mutations from The Cancer Genome Atlas (TCGA) on TNBC, using gene expression data on TNBC from the TCGA as the intermediate genotype. Preliminary results have revealed that genes containing germline mutations also contain somatic and epi mutations and that these genes are functionally related and interact in molecular networks. Additionally, the analysis revealed biological pathways enriched for germline, somatic and epi mutations. These results demonstrate that integrative analysis combing genetic, genomic and epigenomic information is a powerful approach for biomarker discovery in TNBC.

PA-45 ***CYCS human gene analysis insilico***

Mohammed Elmoataz Yusif

This paper study CYCS gene importance main function and direct disease cause and associated disease then suspected disease. Estimated from 13 software Result beside comparing results to estimate the significant mutation in human. The must frequented mutation and highest score are K26M rs11548798. Previous study in yeast Reported Y48H.and spot light molecular biological effects on Protein characteristics due to this mutation. Affiliation: MOHAMMED ALMOATAZ W.YUSIF SUDAN OMDURMAN ALMORADA 4-4 282 E-MAIL: biology.logy16@gmail.com Tel:00249919790092

PA-46 ***Tumor grade, patient age and geographic area related racial disparity patterns in prostate cancer mortality: epidemiological implications and the underlying molecular biology***

Kun Zhang
Xavier University of Louisiana

A reported major racial disparity in prostate cancer (PCa) is that African American (AA) patients have a higher mortality rate than their European American (EA) counterparts. We perform an integrative analysis of the SEER data and molecular biology data to further clarify such a disparity. In this study, we filtered the SEER 2009-2011 records and divided them into 4 groups, i.e. EA-HG, EA-LG, AA-HG and AA-LG, where HG and LG respectively represent the high- and low- grade tumors. On such a partition, we performed a series of analyses using standard statistical methods. Molecular evidence for a primary result of the epidemiological analysis is

obtained on several public gene expression datasets. Based on the registry-specific measures, a significant linear regression of total mortality rate on the percentage of high-grade cancers (PHG) is demonstrated in EAs ($p < 0.01$) but not in AAs. PHG and its racial disparity are differentiated across ages and the groups defined by patient outcomes. For patients with cancers in the same grade category, i.e. LG or HG, the survival stratification between races is not significant in most geographic areas. The genes with different expression levels between AAs' and EAs' tumors of the same grade category are rare. Our study suggests that the perception that prostate tumors are more evasive in AAs than in EAs is true when AAs have a higher PHG than EAs. However, this perception is questionable when the comparison is focused on cases within the same grade category.

PA-47 ***The druggable role of E-Cadherin (CDH1) and Cadherin 11 (CDH11) in claudin-low breast cancer cell lines compared to normal-like breast cancer cell lines revealed by Transcriptomic profiling***

Rica Zinsky

Claudin-low is a molecular subtype of breast cancer discovered in gene-expression studies and is associated with an unfavourable outcome. To reveal new targets specific for the therapy of claudin-low breast cancer, differential gene expression approaches for analysing RNA-seq datasets were applied and compared claudin-low and normal-like breast cancer to identify stand-alone-marker for claudin-low breast cancer subtype. We used RNA-seq datasets of Daemen et al. for our project. Among the 64 cell lines analysed in this reference project, six cell lines are claudin-low and five normal-like. Those eleven cell lines were used for our analyses with the T-Bio-info platform. To analyse the differential gene expression DESeq2, EdgeR, Random Forest, and the T-test were performed. The differentially expressed genes were annotated with the human GAGE pipeline and with the DAVID annotation tool to found out dysregulated pathways. Among the differentially expressed genes, at least for two a targeted therapy is available. The overexpressed CDH11 (cadherin 11) is also a therapeutic target in rheumatoid arthritis and treated with Celecoxib, a small molecule that showed an inhibitory effect in cancer cells. And a lack of CDH1 (cadherin 1) expression in breast cancer cell lines, like shown in the claudin-low breast cancer cell lines, was reported to be sensitive to Crizotinib treatment, a small molecule that is very effective in the treatment of ROS1 or ALK positive non-small cell lung cancer. Conclusion: Both drugs show a potential to inhibit especially claudin-low breast cancer and are already used in other diseases which could accelerate clinical testing of Celecoxib and Crizotinib for claudin-low breast cancer patients.

Oral Presentation Index by Author

Poster Abstracts Index by Author

Poster Abstracts Index by Author

Poster Abstracts Index by Author

Organizing Committee

Gabrielle Broussard
LBRN Grant Coordinator
Louisiana State University-Baton Rouge

Gus Kousoulas
LBRN Principal Investigator
Louisiana State University-Baton Rouge

Ramesh Subramanian
LBRN Program Coordinator
Louisiana State University-Baton Rouge

Lyndon Coghill
LBRN Bioinformatician
Louisiana State University-Baton Rouge

John Quebedeaux, Jr.
LBRN Computer Manager
Louisiana State University-Baton Rouge

Christopher Taylor
LBRN BBC Core Director
Louisiana State University Health Sciences Center-New Orleans

Nayong Kim
LBRN BBC Assistant Director
Louisiana State University-Baton Rouge

Ram Ramanujam
LSU CCT Director
Louisiana State University-Baton Rouge

Alexis White
LBRN Grant Administrator
Louisiana State University-Baton Rouge

Sponsors

Louisiana Biomedical Research Network

LSU – Center for Computation and Technology

International Society for Computational Biology

LSU – Office Research and Economic Development

LSU – Center for Experimental Infectious Disease Research

Funded by LBRN (5P20GM103424-17) and CEIDR (5P30GM110760-05) grants from the National Institute of General Medical Sciences of the National Institutes of Health.