Title

Review Presentation: "Fractal-based analysis of optical coherence tomography data to quantify retinal tissue damage"

Presenter

Abel Alemeshet, Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

Abstract

This poster is a review of the paper "Fractal-based analysis of optical coherence tomography data to quantify retinal tissue damage" by Somfai *et al*. (2014). The abstract is as follows:

"The sensitivity of Optical Coherence Tomography (OCT) images to identify retinal tissue morphology characterized by early neural loss from normal healthy eyes is tested by calculating structural information and fractal dimension. OCT data from 74 healthy eyes and 43 eyes with type 1 diabetes mellitus with mild diabetic retinopathy (MDR) on biomicroscopy was analyzed using a custom-built algorithm (OCTRIMA) to measure locally the intraretinal layer thickness. A power spectrum method was used to calculate the fractal dimension in intraretinal regions of interest identified in the images. ANOVA followed by Newman-Keuls post-hoc analyses were used to test for differences between pathological and normal groups. A modified p value of <0.001 was considered statistically significant. Receiver operating characteristic (ROC) curves were constructed to describe the ability of each parameter to discriminate between eyes of pathological patients and normal healthy eyes."

Title
Transcriptome analysis of alveolar macrophages in aging and influenza infected mice using RNA-Seq approach

Authors
Kishore R. Anekalla,[1] Alexander S. Misharin,[2] Jacob I. Sznajder,[2] and Scott G.R. Budinger[2]
[1]Bioinformatics Program, The University of Texas at El Paso, El Paso, TX, and [2]Pulmonary and Critical Care, Feinberg School of Medicine, Northwestern University, Chicago, IL

Abstract
Proteostasis network (PN) is an intricate system comprising chaperones, trafficking systems, degradation components, and signaling pathways that play an important role in the normal functioning of cellular mechanism. Molecular chaperones and co-chaperones form the core component in PN. Conditions like pathogen infections and aging may cause imbalance in intracellular protein homeostasis (proteostasis). This imbalance may disrupt protein folding pathways. PN in the lung is more frequently challenged due to its constant exposure to stressors. Alveolar macrophages (AM) of lungs are cells present in the highly dynamic microenvironment between the airway epithelium and blood vessels, making them the first cells to get exposed to the stressors. AM can also trigger the immune response and initiate repair mechanisms after tissue damage. Due to continuous stress conditions in the lungs, levels of tissue-resident AM cells may decline and their function may be supported by recruited monocyte-derived macrophages that adapt into tissue-resident AM-like cells. The disturbances in chaperone network of AM need to be understood as they may play a major role in triggering a wide array of tissue responses. Murine models were selected as the experimental system, with an RNA-seq approach to address the problem. Transcriptomic sequencing was done using an Illumina-based HiSeq2000 sequencing machine with single-end reads. These reads were aligned to the mouse genome using Tophat. Counts data were generated using different Bioconductor packages and differential expression analysis was done using edgeR. Chaperone signatures were mapped in the list of genes generated and 24 genes of different chaperone families were found to be differentially expressed.

Title
Dopamine receptors involved in ethanol sensitivity and tolerance

Authors
José A Barragán, Ivan Mercado, and Kyung-An Han
Department of Biological Sciences, Neuromodulation Disorders Cluster, Border Biomedical Research Center, The University of Texas at El Paso, El Paso, TX

Abstract
Ethanol is known to disrupt the dopaminergic system and induces behavioral changes. Increased tolerance and sensitization are examples of altered behaviors that underlie alcohol addiction. In addition, reduced sensitivity to the sedative effect of alcohol is associated with susceptibility to alcohol addiction. Alcohol abuse and addiction can lead to detrimental effects on mental and physical well-being of individuals and also poses additional problems in our society through violence and traffic fatalities. The mesolimbic dopamine pathway is shown to be involved in alcohol addiction; however, the role of dopamine receptors is still ambiguous. In this study we used *Drosophila* as a model system to elucidate the roles of dopamine receptors in sensitivity and tolerance to the sedative effect of ethanol. Flypub was used to deliver ethanol vapor to the freely moving flies and the number of sedated flies was measured every 2 min. Tolerance index was calculated by subtracting the mean sedation time of the first exposure minus that of the second exposure, which is then divided by the mean sedation time of the first exposure. Preliminary data show that the male but not female flies with mutation in D5-dopamine receptor had reduced tolerance compared to the control CS flies. This suggests that D5-dopamine receptor may be important for tolerance development in a gender-dependent manner. Studies are in progress to characterize all dopamine receptor mutants and map the neural substrates of their functions. Overall this research may provide insight into a genetic factor predisposing to alcohol addiction.

Title
Discrimination of seismic signals arising in earthquakes and mining explosions by using Fourier
analysis

Authors
Md Al Masum Bhuiyan,[1] Maria C. Mariani,[1,2] Osei K. Tweneboah,[1] and Hector Gonzalez-Huizar[3]
[1]Computational Science Program, [2]Department of Mathematical Sciences, and [3]Department of
Geology, The University of Texas at El Paso, El Paso, TX

Abstract
The Fourier Transform is by far the most important method used in seismic data processing and
interpretation. Seismic data is recorded into what is termed the time-domain. Using the Fourier
Transform, a time domain signal is transformed to the frequency domain where it is equivalent to
an Amplitude Spectrum and a Phase. Fourier's theory states that a given signal can be synthesized
as a summation of sinusoidal waves of various amplitudes, frequencies and phases. Our study is
to discriminate between natural seismicity and anthropogenic events by using Fast Fourier
transform (FFT). The seismic signals generated by activities like nuclear tests and explosions
employed in mining, are recorded by seismic station and might be confused with natural tectonic
earthquakes. Usually seismologist visually inspect seismograms (time series that record the
displacement of the ground) in order to identify particular phases that can define the source of the
signal, however, incorrect interpretations sometimes lead to an erroneous analysis of seismic
hazard in a particular region.

In our investigation, we collected the seismograms containing the seismic waves generated by the
earthquakes and the explosions recorded by three nearby seismic stations (IU.TUC, US.ANMO,
IU.WUAZ) and observe that they generate different frequency spectra. Our job is to make a model
by using Fourier analysis as a tool to discriminate between natural tectonic earthquakes and
explosions generated in mining, the data corresponds to earthquakes occurred in 2014 in Arizona,
and explosions cataloged with similar magnitude as the earthquakes (Magnitude M=3.0-3.3) and
located in the same region (within a radius of 10km).

Title

Study of volatility structures applied to high frequency data arising in Geophysics and Finance

Authors

Francis Biney and Maria C. Mariani

Department of Mathematical Sciences and Computational Science Program, The University of Texas at El Paso, El Paso, TX

Abstract

This work investigates the underlying volatility processes in earthquake series, high frequency (tick) data, financial indices and explosive series. Furthermore it examines the applicability of a range of GARCH specifications for modeling volatility of these series in order to identify similarities and differences in the volatility structures. The GARCH variants considered include the basic GARCH, IGARCH, ARFIMA (0,d,0)-GARCH and FIGARCH specifications. The major contribution of this work comes in the realm of applications. The methodology is applied to three domains: Geophysics (earthquake data), Finance (high frequency data and indices) and explosives data. In all the applications the methodology provides insight into features of these series volatility.

The results show that the FIGARCH specification is favored in the DJIA, S&P500 and the explosives series volatility but in the BAC and JPM high-frequency data and in the earthquake series the ARFIMA-GARCH specification is preferred as it more reliably describes the volatility of these series. The WMT and IBM high-frequency data volatility were best described using the GARCH model.

Title
Genetic diversity of a widespread tropical mysid, *Mysidium gracile*

Authors
Rosaura J. Chapina,[1, 2] Jennifer Chavez-Ramos,[1] and Elizabeth Walsh[1,2]
[1]Department of Biological Sciences and [2]Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

Abstract
Genetic diversity is essential for organisms to respond to environmental factors such as climate change and ocean acidification. Greater genetic diversity amongst organisms allows a higher probability of adaptation to changing environments. Since seawater temperatures continue to increase and become more acidic, marine organisms must adapt or migrate to new habitats to prevent populations of being decimated. The mysid *Mysidium gracile* is a shrimp-like crustacean that aggregates to swarms near coral reefs. They play a vital role in reef systems by occupying an important intermediate level of marine food webs and transferring energy from planktonic species to reef fishes. Thus, there is concern regarding the tolerance of reef- associated organisms to climatic changes and overall coral reef health. The objective of this study is to determine the level of genetic variation within and among Caribbean *M. gracile* populations. Mysids from 7 islands were collected and preserved for genetic analysis. The CO1 gene was amplified and sequenced for 100 mysids representing 14 swarms. Genetic diversity among populations varied from 0-30% and distances within populations ranged from 0-26%. Haplotype diversity was determined using DnaSP5.0. Twenty-three haplotypes were detected with a haplotype diversity of 0.94, thus indicating a high level of haplotypic variation. Mysids from two different populations shared a haplotype, implying that there is potential gene flow between these populations. Mysids from additional swarms are in process of being characterized. Overall, these results show that there is genetic variation within and among mysid populations that may allow them to adapt and withstand environmental factors.

Title
Predicting G-protein coupled receptors in the *Rhipicephalus microplus* synganglion

Authors
Anastasia Kellogg,[1] Felix Guerrero,[2] and Ming-Ying Leung[1,3]
[1]Bioinformatics Program, The University of Texas at El Paso, El Paso, TX, [2]USDA Knipling-Bushland U.S. Livestock Insects Research Laboratory, Kerrville, TX, and [3]Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, TX

Abstract
The cattle tick, *Rhipicephalus (Boophilus) microplus*, is a pest that causes multiple health complications in cattle, often leading to decreased production and occasionally death. Therefore, the development of successful pest control strategies greatly benefits cattle producers. The prevalence of the G-protein coupled receptor (GPCR) super-family as a drug target draws interests to their candidacy as a pesticide target in the cattle tick, but GPCRs are known to share little sequence similarity, and at this time; there is no well-characterized reference genome available for *R. microplus*. This limits the effectiveness of alignment dependent methods such as BLAST and PFAM. However, GPCRs share a common structure consisting of seven transmembrane helices.

Presented here is a combination of structurally-based and alignment-free methods, which supplement the identification of potential GPCRs from sequence data. TMHMM predicts the number of transmembrane helices present in a protein sequence. GPCRpred is a SVM for predicting GPCRs using the dipeptide composition of the sequence. These tools were applied to our transcriptome assembly of the Texas cattle tick synganglion. Together, BLAST and PFAM identified 64 unique contigs containing GPCR candidates. Collectively, TMHMM and GPCRpred identify a greater number of prospective GPCR candidates than BLAST and PFAM alone. This demonstrates that a combination of structurally-based and alignment-free methods produces a greater collection of prospective GPCRs than common alignment-based methodologies that require similarity to known sequences. The combination of all of these tools presents a pool of likely GPCR candidates for further laboratory analysis as possible pesticide targets.

Title
Numerical simulation of the morphosynthesis of polycrystalline biominerals

Authors
Bethuel Khamala[1] and Natasha Sharma[1,2]
[1]Computational Science Program and [2]Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, TX

Abstract
We consider a multistage process consisting of a polymer-induced liquid-precursor (PILP) phase and the occurrence of spherulites due to surface nucleation process for the synthesis of crystal polymorphs of biogenic substances. This process is modeled by the coupled Cahn-Hilliard Navier-Stokes system which describes the mixing of both miscible and immiscible liquids as well as the PILP phase. We present some preliminary results for the coupled system as well for the fourth order Cahn Hilliard equation. The numerical experiments were implemented within the software package FeniCs.

Title
A Next-Generation Sequencing approach to evaluate genetic variation among *E. vaginatum* populations.

Authors
Armando Lerma,[1] Jon Mohl,[2,3] Ned Fetcher,[4] Ming-Ying Leung,[2,3,5] and Michael L. Moody[1]
[1]Department of Biological Sciences, [2]Bioinformatics Program, and [3]Computational Science Program, The University of Texas at El Paso, El Paso, TX, [4]The Institute for Environmental Science and Sustainability, Wilkes University, Wilkes-Barre, PA, and [5]Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, TX

Abstract
Eriophorum vaginatum is an ecologically important Arctic plant lineage and it serves as a model for understanding home site advantage in light of climate change. Ecophysiological research has recognized distinct ecotypes from the north and south of its range in Alaska; analysis of its genetic structure can help to predict how these populations will respond to climate change. This research employs double digest restriction associated DNA (RAD) sequencing, a next generation sequencing technique, to identify single nucleotide polymorphism (SNP) data for distinguishing the ecotypes at the genetic level and ultimately levels of gene flow among populations. In a preliminary analysis, a script was written in Python computer language to create a workflow that filtered, processed and analyzed the RAD sequence data using different programs (FASTX, STACKS). Subsequently, a tree-building program, Geneious, was used to evaluate genetic variation, structure and relatedness between ecotypes. After completing the analysis, a significant number of SNPs were identified and neighbor-joining trees allowed us to observe an expected north/south split but the resolution of it suggested that inclusion of more loci may be more important than using only loci with high sample coverage for an increased resolution. A VCF (variant call format) file was produced containing genotype information on each sample, by parsing and manipulating this file we will be able to identify and divide the specific SNPs that varied among the different populations for further analysis.

Title
Increasing resolution of Hox evolution with whole-genome sequencing from three echinoderm species

Authors
Kyle Long,[1,2] Carlos Nossa,[3] Mary Sewell,[4] Nicholas Putnam,[3] and Joseph Ryan[1]
[1]Whitney Laboratory for Marine Bioscience, University of Florida, St. Augustine, FL,
[2]Bioinformatics Program, The University of Texas at El Paso, El Paso, TX,
[3]Department of Ecology and Evolutionary Biology, Rice University, Houston, TX, and
[4]University of Auckland, NZ

Abstract
Hox genes are a group of transcription factors known to be important in patterning the primary body axis of most animals during development. Unlike other animals, the echinoderm body plan incorporates a pentaradial adult form built upon a bilateral larval plan. It has long been suspected that changes in the presence/absence, arrangement, and expression of Hox genes in stem ancestors of echinoderms have contributed to their unique adult body plan. Although the genome of Strongylocentrotus pupuratus has been fully sequenced and studied extensively, genomic information from other echinoderms remains scarce. We have sequenced and assembled the genomes of three echinoderm species from three of the remaining four classes: Ophionereis fasciata (brittle star), Patiriella regularis (sea star), and Australostichopus mollis (sea cucumber). In this study we characterize 361 previously unidentified homeobox genes. We report the first evidence of the presence of Hox2 in Asterozoa, which had been previously thought to be absent. The absence of Hox4 within holothuroids was also confirmed. Hox4 had previously been shown to be absent in echinoids, meaning that Hox4 was most likely lost in Echinozoa. We identified a novel family of posterior-class hox genes, Hox11/13d, in each of the three species, suggesting that this might also be present within Crinoids and Echinoids. The increased resolution of echinoderm hox genes will aid further studies of the role of Hox genes in echinoderms and shed light on the role Hox genes in echinoderm body patterning.

Title
Use of satellite data to perform atmospheric studies in the Paso del Norte Region

Authors
Suhail Mahmud,[1] Asad U.H. Gulib,[1] Rosa Fitzgerald,[2] and Scott Williams[3]
[1]Computational Science Program and [2]Department of Physics, The University of Texas at El Paso, El Paso, TX, and [3]Department of Plant and Environmental Sciences, New Mexico State University, Las Cruces, NM

Abstract
Using our research group's 1.2 meter satellite antenna, we extract several types of atmospheric data around the region of Paso del Norte (PdN), such as Normalized Difference Vegetation index (which helps interpret the greenness of any land surface), Land surface temperature (the radiative skin temperature of ground which depends on the albedo, the vegetation cover and the soil moisture) and, in addition, airborne dust. We track different satellites, e.g., NOAA 18 and 19, and using the Advanced Very High Resolution Radiometer (AVHRR) sensor we retrieve environmental parameters using programs from NOAA's Coast watch, Metoffice's ATOVS and AVHRR Pre-processing Package (AAPP) and developed algorithms.

Title
Ab-initio computation of electronic, transport, and related properties of zinc blende boron arsenide (zb-BAs)

Authors
Ifeanyi H. Nwigboji,[1,2] Y. Malozovsky,[2] Bethuel Khamala,[1] and D. Bagayoko[2]
[1]Computational Science Program, The University of Texas at El Paso, El Paso, TX, and
[2]Department of Physics, Southern University and A&M College, Baton Rouge, LA

Abstract
We present results from ab-initio, self-consistent density functional theory calculations of electronic, transport, and related properties of zinc blende boron arsenide (zb-BAs). We used a local density approximation (LDA) potential of Ceperley and Alder, as parameterized by Vosko and his group, the linear combination of Gaussian orbitals (LCGO) formalism, and the Bagayoko, Zhao, and Williams (BZW) method, as enhanced by Ekuma and Franklin (BZW-EF) in carrying out our completely self-consistent calculations. With this method, the results of our calculations have full physical content. Our results include electronic energy bands, and densities of states. Our calculated, indirect band gap of 1.48 eV, from Γ-X, for the room temperature lattice constant of 4.777 Å, is in excellent agreement with experimental value of 1.46 ± 0.02 eV. We explain the agreement between these findings, including the indirect band gap, and available, corresponding, experimental ones. This work confirms the capability of DFT to describe accurately properties of materials, provided the computations adhere strictly to the conditions of validity of DFT as done by BZW-EF method. Acknowledgements: Work funded in part by: The National Science foundation (NSF) and the Louisiana Board of Regents, through LASiGMA [Award Nos. EPS-1003897, NSF(2010-15)-R11-SUBR] and NSF HRD-1002541, The US Department of Energy - National Nuclear Security Administration (NNSA) (Award No. DE-NA0001861) and DE-NA0002630) LaSPACE and LONI-SUBR.
.

Electronic structure of endohedral fullerenes with magnetic endohedral units

Authors
Kamal Nyaupane,[1] Andrew Li,[2] Rajendra Zope,[1,3] and Tunna Baruah[1,3]
[1]Computational Science Program, The University of Texas at El Paso, El Paso, TX, [2]Department of Physics, University of California at Santa Barbara, Santa Barbara, CA, and [3]Department of Physics, The University of Texas at El Paso, El Paso, TX

Abstract
Endohedral fullerenes represent a novel family of carbon nanostructures, which are characterized by a fullerene cage with atoms, ions, or clusters trapped in its interior. The encapsulated molecule or cluster can determine the properties of the fullerenes. Often charge transfer occurs from the endohedral unit to the outer cage leading to high stability. Resulting from the charge transfer from the encaged species to the fullerene cage, endohedral fullerenes hold a lot of fascinating properties inaccessible by the empty fullerenes, and consequently promise potential applications in biomedicine, molecular electronics and photonics etc. In this work we have examined the electronic properties of several transition metal oxide clusters encapsulated within a carbon cage. The chosen transition metal clusters have cubane structure with high spin. The motivation is to examine whether the encapsulation can stabilize transition metal clusters with high magnetic anisotropy energy. We have calculated the electronic and magnetic properties of $Co_4O_4@C_{80}$, $Mn_4O_4@C_{76}$, $Mn_4O_4@C_{70}$, $Mn_4O_4@C_{78}$ using density functional theory. And also we have calculated the electronic and magnetic properties by rotating endohedral unit. The magnetic anisotropy energy of these endohedral fullerenes is calculated using a perturbative approach. Such magnetic clusters can have potential application in magnetic storage devices.

Title

Computational methods for G-protein coupled receptors prediction

Authors

Echezona Okoye,[1] Anastasia Kellogg,[1] and Ming-Ying Leung[1,2]

[1]Bioinformatics Program and [2]Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, TX

Abstract

One of the most commonly used and successful targets for drugs is a specific protein group called G-protein coupled receptors (GPCRs). Approximately forty to fifty percent of all modern medicines interact with this protein group. GPCRs are prime targets of interest due to their biological relevance such as signaling, transporting of molecules and cell-cell communication. We are developing a standalone and web-based GPCR prediction tool, called GPCR Pipeline, capable of identifying likely GPCRs from genomic sequences and pinpointing their transmembrane regions. The GPCR Pipeline is being developed using the Python programming language and incorporating the approaches used in popular bioinformatics software such as Pfam, BLAST, and TMHMM, which are based on sequence alignments, database searches, and hidden Markov models. Aiming at better performance, our tool will be further benchmarked against existing GPCR predictors such as GPCRpred, GPCRHMM, and PCA-GPCR, and then made publicly available to assist scientists with predicting GPCRs.

Title

Review Presentation: "Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: a test case in the euasterid plant clade"

Presenter

Roshani Rajapaksha, Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

Abstract

This poster is a review of the paper "Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: a test case in the euasterid plant clade" by Feinan Wu *et al*. (2006) published in the journal Genetic Society of America. The abstract, as originally published by the authors, is as follows:

"We report herein the application of a set of algorithms to identify a large number (2869) of single-copy orthologs (COSII), which are shared by most, if not all, euasterid plant species as well as the model species Arabidopsis. Alignments of the orthologous sequences across multiple species enabled the design of "universal PCR primers," which can be used to amplify the corresponding orthologs from a broad range of taxa, including those lacking any sequence databases. Functional annotation revealed that these conserved, single-copy orthologs encode a higher-than-expected frequency of proteins transported and utilized in organelles and a paucity of proteins associated with cell walls, protein kinases, transcription factors, and signal transduction. The enabling power of this new ortholog resource was demonstrated in phylogenetic studies, as well as in comparative mapping across the plant families, tomato (family Solanaceae) and coffee (family Rubiaceae). The combined results of these studies provide compelling evidence that (1) the ancestral species that gave rise to the core euasterid families Solanaceae and Rubiaceae had a basic chromosome number of $x = 11$ or 12.2) No whole-genome duplication event (*i.e.*, polyploidization) occurred immediately prior to or after the radiation of either Solanaceae or Rubiaceae as has been recently suggested."

Title
Characterization of Heat Shock Protein 70 in cancer through protein-protein interaction networks and homology analysis

Authors
Janeth Rodriguez,[1] Ming-Ying Leung,[1,2] Jianying Zhang,[1,3] and Hugh Nicholas,[4]
[1]Bioinformatics Program, [2]Department of Mathematical Sciences, and [3]Department of Biological Sciences, The University of Texas at El Paso, El Paso, TX, and [4]Biomedical Applications, Pittsburgh Supercomputing Center, Pittsburgh, PA

Abstract
Purpose: Heat Shock Protein 70 (HSP70) is known to be overexpressed in cancer. It is important to determine if HSP70 can be a potential biomarker for early cancer detection. This analysis will use bioinformatics to characterize HSP70 in relation to other cancer-related proteins and the heat shock protein family.

Design Methods: Protein-protein interaction (PPI) network databases were used to investigate whether HSP70 interacts with other known cancer-related proteins. If so, then there is a high probability that HSP70 is specific to cancer and can be a potential biomarker for cancer detection. Once a correlation between HSP70 and cancer is found, then a homology analysis of the heat shock protein family is done through the construction, optimization and subsequent mining of a high-accuracy multiple sequence alignment. This is done in order to better understand the role of HSP70 in the protein family and to verify that there are conserved regions across the sequences of the protein family to target the biomarker.

Results/Expected Results: Through the PPI analysis it was found that HSP70 has 49 cancer-related protein interactions including Heat Shock Protein 90 (HSP90A). It is also involved in the Mitogen-Activated Protein Kinase (MAPK) signaling pathway that is involved in cancer. Three conserved regions were found in the ATP/ADP binding domain of HSP70 in the homology analysis.

Discussion/Conclusion: Based on this analysis HSP70 is a good candidate for a biomarker in detecting cancer because it interacts with other cancer-related proteins and has the ADP/ATP conserved region. In order for a protein to serve as good biomarker in cancer it is important for it to have highly conserved regions for the antibody to recognize. The results in this analysis will help design wet-lab experiments for testing HSP70 as a possible biomarker and is expected to save valuable time and lab resources.

Title
Review Presentation: "VISA – Vector Integration Site Analysis server: a web-based server to rapidly identify retroviral integration sites from next-generation sequencing"

Presenter
Dewan Shrestha, Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

Abstract
This poster is a review of the paper "VISA – Vector Integration Site Analysis server: a web-based server to rapidly identify retroviral integration sites from next-generation sequencing" by Hocum *et al.* (2015) published in the journal *BMC Bioinformatics*. The abstract, as originally published by the authors, is as follows:

"**Background:** Analyzing the integration profile of retroviral vectors is a vital step in determining their potential genotoxic effects and developing safer vectors for therapeutic use. Identifying retroviral vector integration sites is also important for retroviral mutagenesis screens.

Results: We developed VISA, a vector integration site analysis server, to analyze next-generation sequencing data for retroviral vector integration sites. Sequence reads that contain a provirus are mapped to the human genome, sequence reads that cannot be localized to a unique location in the genome are filtered out, and then unique retroviral vector integration sites are determined based on the alignment scores of the remaining sequence reads.

Conclusions: VISA offers a simple web interface to upload sequence files and results are returned in a concise tabular format to allow rapid analysis of retroviral vector integration sites.

Keywords: Bioinformatics, Gene therapy, Retroviral vectors, Next-generation sequencing, Retroviral mutagenesis".

Title
Structural truss optimization by finite element analysis and genetic algorithm

Authors
Mohammad Tauhiduzzaman[1] and Pavana Prabhakar[1,2]
[1]Computational Science Program and [2]Department of Mechanical Engineering, The University of Texas at El Paso, El Paso, TX

Abstract
Structural optimization is one of the key factors for designing massive structures, where the main concern is not merely to ensure a complex design, but also to identify the limiting factors and resolve the issues by generating optimum values of the main parameters. Algorithms, like the genetic algorithm, that mimic natural processes can be employed in the structural optimization technique. In this paper, both finite element analysis and genetic algorithm are employed to obtain the optimum result of the cross-sectional area for truss structures. Area is the main variable for this optimization technique that can be expressed by the array of binary numbers to carry out the genetic algorithm operation. Subsequently stress analysis is performed using the material properties. Since minimization of the weight is the objective function, decreasing the cross-sectional areas subjected to higher stress of the truss members. Allowable stress operates as a stopping criterion for this iterative process. Finally, stress analysis and genetic algorithm create a possible solution set for areas and weight of the unit cell for the truss structure to be determined. Finite element analysis is conducted by combining FEA (using ABAQUS) and the genetic algorithm implemented in MATLAB.

Title
Review Presentation: "Adaptive evolution and elucidating the potential inhibitor against schizophrenia to target DAOA (G72) isoforms"

Presenter
Kristin Valdez, Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

Abstract
This poster is a review of the paper "Adaptive evolution and elucidating the potential inhibitor against schizophrenia to target DAOA (G72) isoforms" by Sehgal *et al*. (2015) published in the journal *Drug Design, Development and Therapy*. The abstract, as originally published by the authors, is as follows:

"Schizophrenia (SZ), a chronic mental and heritable disorder characterized by neurophysiological impairment and neuropsychological abnormalities, is strongly associated with d-amino acid oxidase activator (*DAOA*, *G72*). Research studies emphasized that overexpression of *DAOA* may be responsible for improper functioning of neurotransmitters, resulting in neurological disorders like SZ. In the present study, a hybrid approach of comparative modeling and molecular docking followed by inhibitor identification and structure modeling was employed. Screening was performed by two-dimensional similarity search against selected inhibitor, keeping in view the physiochemical properties of the inhibitor. Here, we report an inhibitor compound which showed maximum binding affinity against four selected isoforms of DAOA. Docking studies revealed that Glu-53, Thr-54, Lys-58, Val-85, Ser-86, Tyr-87, Leu-88, Glu-90, Leu-95, Val-98, Ser-100, Glu-112, Tyr-116, Lys-120, Asp-121, and Arg-122 are critical residues for receptor–ligand interaction. The C-terminal of selected isoforms is conserved, and binding was observed on the conserved region of isoforms. We propose that selected inhibitor might be more potent on the basis of binding energy values. Further analysis of this inhibitor through site-directed mutagenesis could be helpful for exploring the details of ligand-binding pockets. Overall, the findings of this study may be helpful in designing novel therapeutic targets to cure SZ."

Title
Identification of a dataset and features for ubiquitination prediction

Authors
Mariana Vasquez,[1] Jon Mohl,[2,3,4] and Ming-Ying Leung[1,2,3,4,5]
[1]Undergraduate Participation in Bioinformatics Training Program (UPBiT), [2]Border Biomedical Research Center, [3]Bioinformatics Program, [4]Computational Science Program, and [5]Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, TX

Abstract
Ubiquitination is a post-translational modification that tags a protein for destruction by attaching ubiquitin, commonly referred to "the kiss of death". It regulates various cellular processes such as apoptosis (programmed cell death), the presentation of antigens to the immune system, and protein degradation. Ubiquitin pathway disorders play a role in diseases such as Alzheimer's, lupus, and cancer. Ubiquitination is catalyzed by three enzymes: ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzymes (E2), and ubiquitin ligase (E3). Together they play a role in binding ubiquitin to a protein. Ubiquitination sites are difficult to obtain experimentally. UbiProt is a database specific for ubiquitination, and was used to construct a dataset with confirmed positive and negative ubiquitinated sites. Protein homology and structural features along with chemical features of amino acids within the sequence were compiled. Using this feature data, a machine learning algorithm can now be selected and trained for ubiquitination prediction.

Title
Review Presentation: "Gene network requirements for regulation of metabolic gene expression to a desired state"

Presenter
Bofei Wang, Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

Abstract
The abstract is a review of the paper "Gene network requirements for regulation of metabolic gene expression to a desired state" by Jan Berkhout, Bas Teusink, and Frank J. Bruggeman (2013) published in the journal *Scientific Reports*. The original abstract published by the authors is as follows:

"Gene circuits that control metabolism should restore metabolic functions upon environmental changes. Whether gene networks are capable of steering metabolism to optimal states is an open question. Here we present a method to identify such optimal gene networks. We show that metabolic network optimization over a range of environments results in an input-output relationship for the gene network that guarantees optimal metabolic states. Optimal control is possible if the gene network can achieve this input-output relationship. We illustrate our approach with the best-studied regulatory network in yeast, the galactose network. We find that over the entire range of external galactose concentrations, the regulatory network is able to optimally steer galactose metabolism. Only a few gene network parameters affect this optimal regulation. The other parameters can be tuned independently for optimisation of other functions, such as fast and low-noise gene expression. This study highlights gene network plasticity, evolvability, and modular functionality."

Title
Study of the photolysis rate coefficients for the Paso del Norte (PdN) Region

Authors
Pema Wangchuk,[1] Suhail Mahmud,[1] Rosa Fitzgerald,[1,2] William Stockwell,[3] and Duanjun Lu[4]
[1]Computational Science Program and [2]Department of Physics, The University of Texas at El Paso, El Paso, TX, [3]Department of Chemistry, Howard University, Washington, DC, and [4]Department of Physics, Jackson State University, Jackson, MS

Abstract
Photolysis rate coefficients are calculated from spectrally resolved actinic fluxes that are derived from radiation transfer models. A radiation transfer model calculates a spectrally resolved actinic flux for average and clear sky conditions, although it is strongly affected by the atmospheric distribution of clouds and site-specific meteorological conditions. We measure hemi-spherically integrated spectrally resolved solar photon flux between wavelengths of 300 and 700 nm (actinic flux) for our research. Therefore, the goal is to investigate the use of measured actinic flux to improve photolysis rate coefficients used by air quality models and then to determine if this improvement will allow the air quality models to simulate more accurate ozone concentrations for the Paso del Norte Region.

Title
Review Presentation: "Ancestry, admixture and fitness in Colombian genomes"

Presenter
Pankhuri Wanjari, Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

Abstract
This poster is a review of the paper "Ancestry, admixture and fitness in Colombian genomes" by Rishishwar *et al*. (2015) published in *Scientific Reports, Nature*. The abstract, as originally published by the authors, is as follows:

"The human dimension of the Columbian Exchange entailed substantial genetic admixture between ancestral source populations from Africa, the Americas and Europe, which had evolved separately for many thousands of years. We sought to address the implications of the creation of admixed American genomes, containing novel allelic combinations, for human health and fitness via analysis of an admixed Colombian population from Medellin. Colombian genomes from Medellin show a wide range of three-way admixture contributions from ancestral source populations. The primary ancestry component for the population is European (average = 74.6%, range = 45.0%–96.7%), followed by Native American (average = 18.1%, range = 2.1%–33.3%) and African (average = 7.3%, range = 0.2%–38.6%). Locus-specific patterns of ancestry were evaluated to search for genomic regions that are enriched across the population for particular ancestry contributions. Adaptive and innate immune system related genes and pathways are particularly over-represented among ancestry-enriched segments, including genes (HLA-B and MAPK10) that are involved in defense against endemic pathogens such as malaria. Genes that encode functions related to skin pigmentation (SCL4A5) and cutaneous glands (EDAR) are also found in regions with anomalous ancestry patterns. These results suggest the possibility that ancestry-specific loci were differentially retained in the modern admixed Colombian population based on their utility in the New World environment."