College of Science

Bioinformatics

The University of Texas at El Paso

Annual Poster Session 2010
Titles and Presenters

1) **Venkata Alluri**
   Title: Computational Identification and RNA Secondary Structure Prediction for siRNA Target Sites

2) **Karla Garcia**
   Title: Phosphogenomics of Rotifera Exposed to Heavy Metals

3) **Matthew Hartman**
   Title: Dinucleotide Composition for Intronic and Non-Intronic miRNA

4) **Hector Lopez**
   Title: Applying a Mini-Array Technique to Increase Autoantibody Detection in Cancer

5) **Javier Ordonez**
   Title: Uncovering Cryptic Diversity in Zooplankton through DNA Sequencing

6) **Daniel Rodarte**
   Title: Sex in *Giardia lamblia*: Analysis of Meiotic Genes

7) **Marco Rodriguez**
   Title: Using Mutual Information to Search for Co-Evolving Networks of Amino Acids in Aldehyde Dehydrogenase (ALDH)

8) **Rahul Vegesna**
   Title: Protein Secondary Structure Prediction Servers and Their Prediction Accuracies

9) **Erika Villa**
   Title: Stable Isotope Analysis of Gladius to Investigate Trophic Patterns

Bell Hall 130A
Friday, November 19, 2010, 10:30 AM – 12:30 PM

For more information, please contact the Program Coordinator, Ms. Oliana Alikaj, at oalikaj2@utep.edu or 915.747.8484.

#1 Computational Identification and RNA Secondary Structure Prediction for siRNA Target Sites

Suman Alluri[1,2] and Ruchi Yadav[2]
[1]Bioinformatics Program, The University of Texas at El Paso
and [2]Sathyabama University, India

The discovery of RNA interference (RNAi) is one of the most revolutionary advances in history of biology. RNAi can suppress gene expressions in a highly specific and potent manner. Small interfering RNA (siRNA) is a class of double-stranded RNA molecules involved in the RNA interference pathway. Prediction of effective siRNA to target Nuclear Shuttle Protein (NSP) gene helps in silencing this gene via RNA interference phenomenon to prevent translation and production of proteins. SVM RNAi 3.6 and other online software such as Ambion's, siRNA target finder, Duplex Finder, GenScript, are used to identify siRNA. The effective siRNA was selected and searched for the position in the Adenovirus gene, it was found in the 1076th position of the gene. Secondary structure design for silencing of NSP protein of Adenovirus is predicted using CLC workbench.

#2 Phosphogenomics of Rotifera Exposed to Heavy Metals

Karla Garcia[1] and Elizabeth Walsh[1,2]
[1]Bioinformatics Program and [2]Department of Biological Sciences
The University of Texas at El Paso

Phosphorylation plays a significant role in a wide range of cellular processes. The exposure to heavy metals is of great concern since it may have an effect on phosphorylation. In this study, we will determine the effects of heavy metals on the phosphorylation of tyrosine using rotifers as the model organism. Additionally, we will investigate the effects of heavy metals on reproductive mode of the rotifer species studied. The four different species of rotifers selected are two bdelloids and two monogononts. Bdelloid species include (*Philodina megalotrocha*, *Adineta vaga*) and monogononts include (*Epiphanes chihuahuaensis*, *Brachionus plicatilis*). Rotifer species were chosen based on their ability to culture easily, having fast generation times and their distinct life cycles. Preliminary studies conducted on *E. hawaiiensis* and *E. dilatata* demonstrated protein expression patterns that appear to change with metal exposure. We also expect to see a similar change for additional rotifer species.

#3 Dinucleotide Composition for Intronic and Non-Intronic miRNA

Matt Hartman,[1] Sameera Viswakula,[2] Leo Saldivar,[1,2] and Ming-Ying Leung[1,2]
[1]Bioinformatics Program and [2]Department of Mathematical Sciences
The University of Texas at El Paso

Emerging research has shown that microRNAs play an important role in genetic regulation and are linked to the development of cancer. Machine learning techniques have been successful in categorizing true microRNA precursors from null hairpin sequences. Issues in the identification of miRNA precursors include finding appropriate features, data sets and null sequences. A significant portion of microRNA precursors are found in sequences annotated as intronic. Additionally, intronic microRNA may be processed differently from intergenic precursors. In this study, we investigate the statistical characteristics of the known microRNA dataset available at mirbase.org. Expected versus actual dinucleotide compositions based on mononucleotide frequencies was analyzed for all known microRNA precursors. Data was grouped on species, precursor structural components (stem, loop, and entire precursor), and whether the sequences were annotated as intronic or intergenic. This study is part of an effort to identify features and sequence composition for a microRNA prediction algorithm based on machine learning.

#4 Applying a Mini-Array Technique to Increase Autoantibody Detection in Cancer

Hector Lopez[1] and Jianying Zhang[1,2]
[1]Bioinformatics Program and [2]Department of Biological Sciences
The University of Texas at El Paso

Studies on autoimmune diseases have provided abundant evidence suggesting that antigens might play a role in cancer disease process. An antigen is a molecule recognized by the immune system. Originally, the term came from antibody generator and was a molecule that binds specifically to an antibody, but the term now also refers to any molecule or molecular fragment that can be bound by a major histocompatibility complex (MHC) and presented to a T-cell receptor. In addition, many studies demonstrated that cancer antibodies react with a unique group of autologous cellular antigens called tumor-associated antigens (TAA). Moreover, many researchers have been interested in the use of antibodies as serological markers for cancer diagnosis. Also, because of the general absence of antibodies in humans and in non-cancer conditions, there is low sensitivity regarding to the approach of autoantibody detection in cancer.

#5 Uncovering Cryptic Diversity in Zooplankton through DNA Sequencing

Javier Ordonez[1] and Elizabeth Walsh[1,2]
[1]Bioinformatics Program and [2]Department of Biological Sciences
The University of Texas at El Paso

Over the years, taxonomists have generated extensive knowledge in biodiversity for macroscopic animals, but in the case of microscopic animals, biodiversity patterns are less well known due to the lack of taxonomic expertise and effort, and the lack of easily identifiable morphological characteristics. My expectation is that the amount of biodiversity in aquatic microorganisms (zooplankton) still to be uncovered may be very high because of the following reasons. The small size of the organisms, which implies more available niches for populations. Their ability of producing dormant stages is associated with long-distance dispersal and the capacity to survive unfavorable periods. Our research objective is to uncover the hypothetical diversity of some of these aquatic microscopic invertebrates, specifically members of the Genus *Alona*, within the Chihuahuan Desert region using molecular genetics tools.

#6 Sex in *Giardia*: Analysis of Meiotic Genes

Daniel Rodarte,[1] Trevor Duarte,[2] Tavis Mendez,[2] and Siddhartha Das[1,2]
[1]Bioinformatics Program and [2]Department of Biological Sciences
The University of Texas at El Paso

Giardia lamblia is a flagellated protozoan and the causative agent of giardiasis. Once the parasites attach to and colonize the upper small intestine of their host, the vegetative trophozoites reproduce by binary fission. It is therefore thought that the reproductive cycle of *Giardia* is entirely asexual and that its population is clonal. In addition, previous studies have shown that the sequenced genome presents low levels of allelic heterozygosity along with ploidy variation during differentiation; a characteristic that is evocative of meiosis. The genome of *Giardia* contains genes specific for meiotic recombination, suggesting that sexual reproduction may occur in this organism. In this work, a phylogenetic analysis of the *Giardia* meiotic-like genes Dmc1b, Spo11, Mnd1, Hop1, and Rad51, was conducted for the comparison of these meiotic genes with homologs found in other species.

#7

Using Mutual Information to Search for Co-Evolving Networks of Amino Acids in Aldehyde Dehydrogenase (ALDH)

Marco Rodriguez
Bioinformatics Program, The University of Texas at El Paso

Aldehyde dehydrogenases (ALDHs) are enzymes that aid in detoxification and glycolysis among other functions in living organisms. Mutual information (MI) allows for an estimate of mutual co-evolutionary relationships between two positions in a protein sequence. It is hypothesized that with the use of MI programs one can predict which co-evolving networks of residues will include allosteric networks and allosteric signaling networks within the subclasses of the ALDH. We will analyze the subclasses methylmalonate semialdehyde (MMSA) and glutamate-gamma-semialdehyde (GGSA). The networks constructed are visualized using the Visual Molecular Dynamics program and a comparative analysis model is used to validate the results. The predicted networks created with the MI programs identified allosteric network similarities between the subclasses. These results will be assessed for accuracy in regards to signaling network predictions between subclasses. This work is in progress, but it will be beneficial for the understanding of how subclasses in the ALDH are related.

#8

Protein Secondary Structure Prediction Servers and Their Prediction Accuracies

Rahul Vegesna,[1] Marco Rodriguez,[1] Anurag Gautam,[1] and Ming-Ying Leung[1,2]
[1]Bioinformatics Program and [2]Department of Mathematical Sciences
The University of Texas at El Paso

Protein secondary structures (PSS) are important in predicting the 3D protein structure and in understanding how proteins fold and interact with other molecules. Most known PSS and their 3D structures are available at Protein Data Bank (PDB). For those proteins with amino acid sequence data only, their secondary structures can be predicted computationally. There are many publicly available PSS prediction servers (PSSPS) for predicting PSS from their amino acid sequences but with varying degrees of prediction accuracies. Our main goal is to identify those PSSPS with higher degrees of accuracy in their predictions. In this study, we selected 100 proteins having different protein structural domains from PDB and submitted them to different PSSPS. The overall comparison of PSS predictions with the known PSS from PDB helped us identify PSSPS with better prediction accuracies.

#9 Stable Isotope Analysis of Gladius to Investigate Trophic Patterns

Erika Villa
Bioinformatics Program, The University of Texas at El Paso

Stable isotopes have been useful when investigating the feeding ecology of cephalopods since they reflect assimilated diet. The gladius of the squid was analyzed for reconstructing squid trophic activity by using stable isotope analysis of carbon and nitrogen. With the growth of the squid and its gladius, both isotopic ratios increase due to consumption of larger prey. Jumbo squids (*Dosidicus gigas*) were collected along the Eastern Pacific and ten gladius were selected. Isotopic ratios were determined along the gladius and a high positive correlation between length and both carbon and nitrogen isotopic signatures was observed. The Δcarbon and Δnitrogen values increased as the squid increased in size, suggesting an ontogenic change in diet. Isotopic pattern differences were viewed amongst sites, suggesting that diet varied as a function of prey diversity and abundance. The gladius has been proposed to represent a lifespan trophic record since increments are deposited as it grows.