Title

Electronic structure and charge transfer excited states of endohedral fullerene-porphyrin/phthalocyanine complexes utilized in organic photovoltaics

Authors

Fatemeh Amerikheirabadi,[1] Luis Basurto,[2] Rajendra Zope,[2] and Tunna Baruah[1,2]

[1]Computational Science Program, UTEP

[2]Department of Physics, UTEP

Abstract

Organic Donor-Acceptor complexes form the main component of the organic photovoltaic devices (OPVs). The open circuit voltage of OPVs is directly related to the charge transfer excited state energies of these complexes. Currently a large number of different molecular complexes are being tested for their efficiency in photovoltaic devices. In this work, density functional theory as implemented in the NRLMOL code is used to investigate the electronic structure and related properties of these donor-acceptor complexes. The charge transfer excitation energies are calculated using the perturbative delta self-consistent field method recently developed in our group as the standard time dependent density functional approaches fail to accurately provide them. The model photovoltaics systems analyzed are as follows: Sc3N@C80_ZnTPP, Y3N@C80_ZnTPP and Sc3N@C80_ZnPc. In addition, a thorough analysis of the isolated donor and acceptor molecules is also provided. The studied acceptors are chosen from a class of fullerenes named trimetallic nitride endohedral fullerenes. These molecules have shown to possess advantages as acceptors such as long lifetimes of the charge-separated states.

Title
Estimation of the coefficients of a linear differential operator

Authors
Maria Barraza-Ríos and Joan G. Staniswalis
Computational Science Program and Department of Mathematical Sciences, UTEP

Abstract
Principal Differential Analysis (PDA) is used to study functional data, that is, data where each observation is represented by a curve or a function.

PDA methodology seeks to identify a linear differential operator (LDO)

$$L = \omega_0 I + \omega_1 D + \cdots + \omega_{m-1} D^{m-1} + \omega_m D^m$$

that satisfies (as closely as possible) $Lx_i = 0$ for each functional observation $x_i(t)$ $(i = 1, \cdots, n)$. A theorem from analysis and results in nonparametric statistics prove that the coefficients $\omega_0, \cdots, \omega_m$ are in the Sobolev space, and thus can be approximated by B-splines.

Current PDA software used to estimate the LDO assumes that $\omega_m = 1$. We present a method that eliminates this restriction, and thus ensures that the coefficients $\omega_0, \cdots, \omega_m$ are in the Sobolev space and that their approximation by B-splines is mathematically valid. The new method is inspired on results in linear regression that show that the weighted average of pair-wise slopes between data points is equivalent to the least-squares estimator for the estimate of the regression line slope.

Title
Numerical methods for coupled fluid flow and geomechanics

Authors
Maranda Bean,
Computational Science Program, UTEP, and
Son-Young Yi,
Department of Mathematical Sciences and Computational Science Program, UTEP

Abstract
We consider the time-dependent coupling between fluid flow and solid deformation in an elastic porous material using Biot's model. A mixed finite element method is used for both the flow and mechanics subproblems. Four iteratively coupled schemes, drained split, undrained split, fixed-strain split, and fixed-stress split, are presented. In each of these numerical algorithms, the flow and mechanics subproblems are split based on some physical assumptions and solved in a staggered fashion until convergence. Numerical results are explored.

Title

Electronic structure of planar aggregates of boron clusters

Author

Carlos Diaz, Computational Science Program, UTEP

Abstract

Using simulated annealing, random searches and basin hopping, we have searched for low lying isomers of boron cluster containing 27 atoms in its neutral and charged states. For those searches, the inter-atomic potential between boron atoms is described using density functional theory at the generalized gradient approximation level. The structures of low lying isomers found in our searches are predominantly quasi-planar. Several of these structures are seen as growth of smaller size boron clusters. The population of low energy neutral boron cluster isomers was used to study the influence that charge has on the structural pattern of charged boron clusters. The calculations indicate that the boron clusters also tend to prefer quasi-planar geometries. These are the largest planar boron cluster aggregates reported so far. Electronic properties such as ionization energies, vertical detachment energies and electron affinities and also the infrared and Raman spectra of neutral clusters will be presented.

Title

DNA motif screening for genes expressed in the mushroom body brain structure of *Drosophila melanogaster* using Fisher's exact test and the Chi-Square test of homogeneity

Authors

Raymond A. Ford,[1] Carolina Guerra,[2] Ming-Ying Leung,[1,2] and Kyung-An Han[3]

[1]Department of Mathematical Sciences, UTEP

[2]Bioinformatics Program, UTEP

[3]Department of Biological Sciences, UTEP

Abstract

Several existing bioinformatics software for identifying DNA motifs are combined with a novel tool to develop a new approach for identifying potential enhancer elements for a set of genes related to learning and memory in *Drosophila*. Using six different DNA motif discovery software tools, we identified 160 conserved motifs among the DNA sequences driving mushroom body expression. We took two different approaches to identify DNA motifs among 28 lines known to display γ-lobe expression: the first based on the length of individual line sequences, and the second based upon stochastic line selection. The FASTA files for all Janelia Farm lines were computationally searched to determine if any of the discovered motifs were present in these sequence data. Eight candidate DNA motifs were identified. The frequency of each motif's occurrence was noted among the 28 lines displaying γ-lobe expression against 28 randomly selected lines known to not display γ-lobe expression. The Fisher's exact test and Chi-Square test of homogeneity were performed for each motif and the *p*-values were compared to a level of significance established with the use of a Bonferroni correction. This information can be used to identify potential enhancer elements for the genes expressed in the mushroom body γ neurons. With these results, we can further identify additional genes expressed in the γ neurons responsible for dopamine and octopamine signals. This study would ultimately help understand the underlying pathological mechanisms of ADHD, autism, schizophrenia, Parkinson's disease, and drug abuse/addiction.

Title

Molecular evidence of cryptic speciation in the rotifer *Brachionus plicatilis* across the Chihuahuan Desert

Authors

Kayla I. Galindo[1] and Elizabeth J. Walsh[1,2]

[1] Bioinformatics Program, [2] Biological Sciences Department, The University of Texas at El Paso.

Abstract

The Chihuahuan Desert is rich in biodiversity with freshwater and inland saline water habitats. *Brachionus plicatilis,* a monogonant rotifer that lives in brackish habitats worldwide, was once thought to be a cosmopolitan species. However through molecular methods, it has been revealed as a cryptic species complex. Cryptic speciation is defined as a biological process whose end result is a group of distinct but morphologically similar species. This project focuses on investigating cryptic speciation of *B. plicatilis* in the Chihuahuan Desert region. Sequences for the mitochondrial cytochrome c oxidase subunit I (COI) gene and the 16S ribosomal ribonucleic acid (rRNA) region were obtained from 6 populations in the Chihuahuan Desert. Preliminary analyses of the 16SrRNA gene region shows little to no genetic distance (.004-.029) among 28 lineages from two populations, Ascarate Lake and Keystone Heritage Park. Preliminary analysis for COI also shows little to no genetic distance (0.0-.006) among 18 lineages from the same two populations.

Title

A bioinformatics approach to identify potential enhancer elements for genes expressed in the mushroom body neurons in *Drosophila*

Authors

Carolina Guerra,[1] Raymond A. Ford,[2] Ming-Ying Leung,[1,2] and Kyung-An Han[3]

[1]Bioinformatics Program, UTEP

[2]Department of Mathematical Sciences, UTEP

[3]Department of Biological Sciences, UTEP

Abstract

Bioinformatics is an interdisciplinary science that focuses on the management and interpretation of data obtained from complex biological phenomena using mathematical models, computational algorithms, and statistical methodologies. In this project, we combine several existing bioinformatics software for identifying DNA motifs with other computational tools to develop a new approach for finding potential enhancer elements of a set of genes related to learning and memory in Drosophila. Through the use of web scraping scripts we collected data of interest from the HHMI Janelia Farm Research Campus and FlyBase to construct 6,931 FASTA files containing the DNA sequence data for D. melanogaster. Using six different DNA motif discovery software, 18 conserved motifs were identified among the DNA sequences driving mushroom body expression. The FASTA files for all Janelia Farm lines were computationally searched to determine if any of the discovered motifs were present in these sequence data. This information can be used to identify potential enhancer elements for the genes expressed in the mushroom body gamma neurons, which are responsible for learning and memory. With these results, we can further identify additional genes expressed in the gamma neurons, which mediate dopamine and octopamine signals that are involved in neuropsychiatric disorders including ADHD, autism, schizophrenia, Parkinsons disease, and drug abuse/addiction. This study would ultimately help understand the underlying pathological mechanisms.

Title

Investigating gene expression profiles and identifying biomarkers for papillary thyroid cancer

Author

Akshita Gurram, Bioinformatics Program, The University of Texas at El Paso.

Abstract

Papillary thyroid cancer (PTC) is the most common malignant thyroid tumor which has increasing prevalence all over the world. According to National Cancer Institute at NIH; the estimated new cases and deaths in United States in 2014 are 62,980 and 1890 respectively. The changes in gene expressions results from high DNA copy number abnormalities which are the contributors of thyroid oncogenesis.There may exist genes which are amplified in thyroid cancer and may not have been revealed yet. Hence, there is an urge to identify novel prognostic biomarkers that would help in earlier detection of the PTC. Main aim of study is to find marker genes that help in identifying tumors in the patients who are more likely to experience aggressive progression of PTC. Present work is done by analyzing microarray dataset GSE3678 using Multiple Experiment Viewer software which generates informative and interrelated displays of expression and annotation data from single or multiple experiments. The dataset had 7 tumor samples and 7 paired normal tissue samples. First the data was normalized and then T-test between the two groups of data was done to identify the differentially expressed genes. Next the entire lists of genes were uploaded into DAVID database to identify the significant genes. This database helps to identify enriched biological themes, visualize genes on KEGG pathway, list interacting proteins and link gene-disease associations. Lastly the significant genes were uploaded into Kyoto enclyopedia of genes and genomes database which deals with genomes, biological pathways, diseases, drugs. Based on the study, there are two genes which might be possible biomarkers such as PAX8 which is responsible for thyroid follicular growth and IP3R responsible for cell division, proliferation and apoptosis control.

Title

Validated solutions for ordinary differential equations via affine arithmetic

Authors

Luis C. Gutierrez, Computational Science Program, UTEP and
Martine Ceberio, Computer Science Department and Computational Science Program, UTEP

Abstract

In many scientific applications, ordinary differential equations (ODE) are used as mathematical models to describe physical phenomena. When safety or reliability of a model is required, guaranteed solutions of the model are necessary. Common numerical methods to solve ODEs produce fast and accurate results, but in some cases the results are not guaranteed to be correct. Moreover, when initial values of an ODE are uncertain (given as an interval), traditional numerical methods cannot handle such parameters. To address such issues, validated methods compute enclosures of the exact solution and can handle uncertain parameters by using interval arithmetic. Although using interval arithmetic can guarantee an enclosure, the exact enclosure might be over-estimated because of dependencies in the functions. This problem is known as the wrapping effect. Affine arithmetic was developed to overcome the wrapping effect by using first order polynomials to enclose the exact solution. In this work, we explore the use of affine arithmetic to compute solutions of Initial Value Problems for ODE's. In particular, we use Taylor series combined with affine arithmetic to improve the over-estimation of interval methods.

Title

Tumor suppressor PTEN: Potential cancerous genes interactors with PTEN in *Homo sapiens*.

Authors

Amal Hassan[1] and Zhang[1,2].

[1]Bioinformatics Program, The University of Texas at El Paso.

[2] Department of Biological Science, The University of Texas at El Paso.

Abstract

PTEN gene that stands for "phosphatase and tension homolog" is a human gene, located at chromosome 10q23.3 and the protein made by the gene is found in most of the body tissues. It acts as a tumor suppressor and any mutation in this gene leads to development of many cancers such as colorectal, lung, breast and gastric cancers. PTEN major function, as tumor suppressor, is to slow down cell division, repair DNA mistakes, and to tell cells when to die in a process called apoptosis (Berger 2011). PTEN is an enzyme that belongs to the family of PTP (protein tyrosine phosphatase) that help to remove a phosphate group from amino acid tyrosine and send signals from outside the cell to cell nucleus. These signals guide the cell to either grow or divide or to mature or even to die. The objective of the research is to search for protein coding genes that can potentially interact with PTEN using protein-protein interaction database such as BioGrid, IntAct, SPIKE, MINT, as well as predicted protein-protein interaction database such as STRING. The results suggested that several cancer related protein-coding genes appeared to interact with PTEN. SLC9A3R1/NHERF1 was the common PTEN interactor gene in all databases. NHERF1 plays the role of cytoplasmic adaptor protein that recruits various signaling proteins to the plasma membrane of epithelia and other cell types. As a conclusion, PTEN activity may have direct relation to the over expression of NHERF1 in cancer patients and might be possible biomarker candidate for cancer detection.

Title

Review Presentation: “Genetic diversity and linkage disequilibrium estimation among the maize breeding germplasm for association mapping”

Presenter

Anastasia Kellogg, Bioinformatics Program, The University of Texas at El Paso.

Abstract

This poster is a review of the paper “Genetic diversity and linkage disequilibrium estimation among the maize breeding germplasm for association mapping” by Liu et al. (2014) published in the journal *International Journal of Agriculture & Biology.* The abstract, as originally published by the authors is as follows:

“Analyzing the genetic basis and linkage disequilibrium (LD) of maize inbred lines is important for maize breeding and marker-trait association. In this study, a total of 201 SSR markers were used to assay the genetic diversity, population structure, and LD of a maize association mapping population consisting of 290 inbred lines, which mainly represented temperate Chinese and US germplasm. Results of genetic diversity analysis showed that this population presented relatively abundant genetic variation and a high level of gene diversity. According to population structure analysis, breeding lines could be clustered into 5 sub-groups, which corresponded well with known pedigree records. Compared with the other 4 sub-groups, the Lvdahonggu (LRC) sub-group showed higher genetic diversity. LD evaluation results showed significant LD levels in pair-wise SSR markers, with up to 40.6% of pairs linked with chromosomes, ranging from 38.1 to 80.1%. The results of the present studies will provide useful information to perform genome-wide association study to improve the efficiency of maize breeding in maize growing areas represented by the panel.”

Title

*Ab-initio c*omputations of electronic, transport, and bulk property of cubic zinc sulphide (zb-ZnS)

Authors

B. Khamala[1,2], L. Franklin[2], E. C. Ekuma[4], Y. Malozovsky[2], A. Stewart[2,3] and D. Bagayoko[2]

[1]Computational Science Program, The University of Texas, El Paso, TX, 79968
[2]Department of Physics, Southern University and A&M College, Baton Rouge, LA 70813
[3]Division of Mathematics and Natural Sciences, Allen University, Columbia, SC, 29204
[4]Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA 70803

Abstract

We present the results from *ab initio,* self-consistent, local density approximation (LDA) calculations of electronic and related properties of zinc-blende zinc sulphide (zb-ZnS). We employed the Ceperley and Alder LDA potential and the linear combination of atomic orbital (LCAO) formalism. Our calculations are non-relativistic. The implementation of the LCAO formalism followed the Bagayoko, Zhao, and Williams method as enhanced by Ekuma and Franklin (BZW-EF). The BZW-EF method includes a methodical search for the optimal basis set that yields the minima of the occupied energies. This search entails increasing the size of the basis set and related modifications of angular symmetry and of radial features. Our calculated, direct band gap of 3.725 eV, at the Γ point, is in excellent agreement with experiment. We have also calculated the total (DOS) and partial (pDOS) densities of states, electron and hole effective masses and the bulk modulus that agree very well with available, corresponding experimental results.

Title

Review Presentation: “Structural biology and bioinformatics in drug design: opportunities and challenges for target identification and lead discovery”

Presenter

Kishore Reddy Anekalla, Bioinformatics Program, The University of Texas at El Paso

Abstract

This poster is a review of the paper “Structural biology and bioinformatics in drug design: opportunities and challenges for target identification and lead discovery” by Blundell TL *et al.* (2006) published in the journal *Philosophical Transactions of The Royal B Society*. The abstract, as originally published by the authors is as follows:

“Impressive progress in genome sequencing, protein expression and high-throughput crystallography and NMR has radically transformed the opportunities to use protein three-dimensional structures to accelerate drug discovery, but the quantity and complexity of the data have ensured a central place for informatics. Structural biology and bioinformatics have assisted in lead optimization and target identification where they have well established roles; they can now contribute to lead discovery, exploiting high-throughput methods of structure determination that provide powerful approaches to screening of fragment binding.”

Title

Predictive functional profiling of missense mutations identified in the kinome and phosphatome of high-risk leukemia

Authors

J Knapka,[1] DM Oaxaca,[2] HL Wilson,[5] B Carcamo,[5] JA Ross,[2,3] G Rodriguez,[2] and MY Leung[1,3,4]

[1]Bioinformatics Program, UTEP

[2]Department of Biological Sciences, UTEP

[3]Border Biomedical Research Center

[4]Department of Mathematical Sciences, UTEP

[5]El Paso Children's Hospital

Abstract

There is an immediate medical need for new targeted strategies for the detection and treatment of leukemia, especially in Hispanic populations who are disproportionately affected by this type of cancer. Dysregulation of protein phosphorylation has been shown to play key roles in cancer cell growth, but the mutational burden within the kinome and phosphatome has not been fully elucidated in high-risk lymphoid leukemias. Therefore, we sought to explore the mutational landscape within the kinome and phosphatome of relapse and treatment refractory acute lymphoblastic leukemia (ALL). Whole-exome sequencing and mutational analysis of the kinome and phosphatome was performed on genomic DNA isolated from 20 primary ALL (new onset and relapsed disease) patient samples and 3 normal control samples selected from an established regional tumor bank. We present an integrative function-based bioinformatics analysis that defines and prioritizes kinase and phosphatase mutations identified in the target tumor cell population. For each identified missense mutation, the affected gene was translated in all predicted splice variants based on NCBI's human genome reference assembly (build 37). JCVI's PROVEAN protein-variant scoring tool was used to predict the functional consequences of the identified mutations. Variants with PROVEAN scores < -2.5 were selected as candidates for further investigation. A PubMed search for previously-reported clinical investigations of each SNP was used to validate the PROVEAN threshold. This work provides significant insight into novel aberrant cell signaling pathways driving leukemia tumor cell growth and thereby brings new diagnostic and therapeutic strategies into focus for further investigation.

Title

Genetic variation and genealogical structure among *Eriophorum vaginatum* ecotypes using next generation sequencing.

Authors

Armando Lerma[1], Jon Mohl[2,3], Ned Fetcher[5], Ming-Ying Leung[2,3,4], Michael L. Moody[1]

[1]Department of Biological Sciences, UTEP

[2] Bioinformatics Program, UTEP

[3]Computational Science Program, UTEP

[4]Department of Mathematical Sciences, UTEP

[5]The Institute for Environmental Science and Sustainability, Wilkes University,

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. 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. The RAD sequence data is filtered through the FASTX sequence preprocessing tools and then loci are assembled using the STACKS pipeline followed by the identification of SNPs. Finally, a model based clustering program (Structure) and a tree building program (Geneious) are used to evaluate genetic variation, structure and relatedness between ecotypes. Distinguishable genetic structure between ecotypes from the north and south end of the range is expected with genetic distance between populations correlating with geographic distance. A high number of ecotype specific SNPs are expected to be discovered for use in large-scale population genetic studies of the group in the Arctic.

Title

ISOGlyP: Protein O-glycosylation prediction using results from random peptide studies

Authors

Jonathon Mohl,[1,2,3] Gerardo Cardenas,[1,2,3] Thomas Gerken,[5] and Ming-Ying Leung[1,2,3,4]

[1]Border Biomedical Research Center, UTEP

[2]Bioinformatics Program, UTEP

[3]Computational Science Program, UTEP

[4]Department of Mathematical Sciences, UTEP

[5]Departments of Pediatrics, Biochemistry, and Chemistry, Case Western Reserve University

Abstract

O-glycosylation has many biological purposes within humans and changes in O-glycosylation patterns have been associated with a range of diseases, such as gastric and renal cancers and diabetes.. There are existing O-glycosylation prediction algorithms, based on neural networks (NetOGlyc 3.1) and support vector machines (NetOGlyc 4.0 and CKSAA-OGlycsite), that look at sequences from a compiled database of experimentally confirmed sites of O-glycosylation within a set of proteins (O-GlycBase) to develop the predictive features of their programs. But within humans, there are at least 20 isoforms of GalNAc-transferases that could have glycosylated any of those proteins. ISOGlyP (Isoform Specific O-Glycosylation Prediction) is a web-based O-glycosylation prediction program, which is currently the only one that takes into account the differences in biochemical properties of different isoforms of GalNAc-transferase. The current prediction is based on data collected from 10 different polypeptide GalNAc-transferases using small random peptides to determine the site-specific amino acid preference of the polypeptide GalNAc-transferase. These experimental values are then used to return a calculated value of the expected increase (or decrease) in glycosylation of a Threonine or Serine within a protein and are specific for the human polypeptide GalNAc-transferases. A working prototype is available for public use at ISOGlyP.utep.edu. This presentation describes our analysis of the human O-GlycBase entries using ISOGlyP and several enhancements to the website.

Title

Towards the scalability and hybrid parallelization of a spatially variant lattice algorithm

Author

Henry R. Moncada, Computational Science Program, UTEP

Advisors

Drs. Shirley V. Moore and Raymond C. Rumpf

Abstract

We use the spatially variant algorithm to synthesize a spatially variant lattice for a periodic electromagnetic structure. The spatially variant algorithm has the ability to spatially vary the unit cell orientation and exploit its directional dependencies. The technique produces a lattice that is smooth and continuous. The lattice spacing will remain strikingly uniform when the unit cell orientation is spatially varied. This is important for maintaining consistent properties throughout the lattice.

Periodic structure like a photonic crystal or metamaterial devices can be mimicked using the spatial variant algorithm. Properties of these periodic structures include unit cell orientation, lattice spacing, fill fraction, and more. The algorithm is not a coordinate transformation technique so it can easily produce a more complicated and arbitrary spatial variance. This is important because scattering and field concentrations at discontinuities can degrade device performance.

Our current effort is to develop high performance, portable spatially variant codes for parallel architectures. For the scalable code development, the first phase consists of the development of an optimized serial C code as a basis for a high performance Message Passing Interface (MPI) implementation for distributed memory to improve its performance. In the second phase, the MPI version of the spatially variant algorithm will be produced for portability and evaluation of the benefits. In the third phase, combining the MPI parallelization with OpenMP shared memory multiprocessing not only benefits us in load balancing but also in cutting down the memory overhead. In the fourth phase, we push the limits to speed up our code performance using the advantage of the graphics processing units (GPU) and CUDA (Compute Unified Device Architecture). In the fifth phase, we expect to develop other scalable computational electromagnetics algorithms as part of this effort.

Title

A GPCR-focused investigation of the *R. microplus* transcriptome

Authors

Sergio Muñoz,[1] Alexandria Ogrey,[1] Felix G. Guerrero,[2] and Ming-Ying Leung[1,3]

[1]Bioinformatics Program UTEP

[2]USDA-ARS Knipling-Bushland US Livestock Insects Research Laboratory

[3]Department of Mathematical Sciences, UTEP

Abstract

Rhipicephalus microplus, also known as the southern cattle tick, has been found in tropical and subtropical regions all over the world, including Mexico. It is a vector for parasites responsible for cattle diseases that can lead to decreased weight, anemia, loss of milk/meat production, and death. The cattle tick was eradicated from the United States in the early 20th century; however, the ease for tick-carrying animals to cross the border, cattle importation from Mexico, and the cattle tick's growing resistance to acaricides are becoming causes for concern that reinfestation of the United States might occur.

The objective of the project is to develop a bioinformatics pipeline for identifying possible G-protein coupled receptors (GPCRs), utilizing sequence length and number of helices. A series of Python scripts were written to analyze the cDNA sequences reverse-transcribed from the transcriptome of the synganglion of the cattle tick. Using these scripts, possible protein coding regions are obtained from the cDNA, and are input into the Transmembrane Hidden Markov Model web server. The protein sequences identified by the scripts as potential GPCRs were run through a standalone version of Pfam, which classified the submitted proteins by comparing them to the Pfam library of Hidden Markov Models. This output, resulted in 28 probable GPCRs. All protein coding regions for the synganglion were run through Pfam for comparison (72 GPCRs were identified) and for identifying the encoded proteins.

Title

Using bioinformatics tools to help determine if heat shock protein 70 (HSP70) is a potential biomarker for hepatocellular carcinoma

Authors

Janeth Rodriguez[1], Ming-Ying Leung[1,2], Jianying Zhang[1,3] and Hugh Nicholas[4],

[1]Bioinformatics Program, The University of Texas at El Paso.
[2]Department of Mathematical Sciences, The University of Texas at El Paso.
[3]Biological Sciences, The University of Texas at El Paso.
[4]Biomedical Applications, Pittsburgh Supercomputing Center.

Abstract

Hepatocellular Carcinoma (HCC) is the most common type of liver cancer in the world and patients generally will die within a year of its detection due to the lack of early diagnostic methods. As Heat Shock Protein 70 (HSP70) is known to be overexpressed in HCC, it is important to determine if HSP70 can be a potential biomarker for the early detection of HCC. 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 HCC. Once a correlation between HSP70 and HCC is found, then a homology analysis using proteins from the heat shock protein family is done in order to verify that there are conserved regions across the protein family. An optimal biomarker has regions that are 90-100% conserved, and can be also targeted by an antibody in the patient's sera. If a protein does not have highly conserved regions then it cannot serve as a good biomarker for cancer because there will be multiple mutations in the protein sequence that the antibody will not recognize. The results in this analysis will be used to help design wet-lab experiments for testing HSP70 as a possible biomarker and is expected to save valuable time and lab resources.

Title

A computational prediction of T cell epitopes for avian flu H5N1 virus

Authors

Shivangi Sharma, Bioinformatics Program, The University of Texas at El Paso.

Mingtao Zeng, Center of Excellence in Infectious Diseases, Texas Tech University Health Sciences Center, El Paso.

Abstract

Avian flu (influenza A subtype virus H5N1) is caused in humans due to direct contact with the poultry. Despite severe epidemic outbreaks on several occasions, not much progress has been made with regard to an epitope-based vaccine designed for H5N1 virus. In this study, 75 envelope proteins were collected from UniprotKB and analyzed with an *in silico* tool, Vaxijen to identify the most immunogenic protein. The protein was then verified through several parameters viz hydrophobicity, stability and binding affinity to predict the T-cell epitopes. T-cell immunity was assessed using secondary structure prediction tools SOPMA and grand average of hydropathicity (GRAVY) to determine that the protein is stable and can induce cell-mediated immunity. The peptide sequence FINVPEWSY was found as the most potential T-cell epitope, respectively. Furthermore, as an RNA virus, one important thing was to establish the epitope as a conserved one.Epitope conservancy for individual peptides was predicted using the online tool from the IEDB analysis resource, showing 63.51% conservancy. The epitope was further tested for binding against the HLA molecule, (which is the molecule present on the cell surface binding to which brings about the immune response) by computational docking techniques (AutoDock) to verify the epitope interaction. However, this is a preliminary study of designing an epitope-based peptide vaccine against H5N1 virus; the results await validation by *in vitro* and *in vivo* experiments.

Title

Review Presentation: “A comprehensive and high-resolution genome-wide response of p53 to Stress”

Presenter

Sivakumar Kareti, Bioinformatics Program, The University of Texas at El Paso

Abstract

This poster is a review of the paper “A comprehensive and high-resolution genome-wide response of p53 to Stress" by Gue Su Chang et al. (2014) published in *Cell Reports* 8, 514-527, July 24th 2014.

P53 is a tumor suppressor that regulates the transcription of stress-response genes; unfortunately many p53 target genes are still unknown. Using ChIP-exo (chromatin immunoprecipitation followed by exonuclease treatment) 2,183 recognition elements (REs) were identified to be associated with p53. Res are positional constrained with other Res and other regulatory elements. Stress resulted in increased occupancy of transcription factor IIB (TFIIB) and RNA polymerase II near REs. By analyzing genome wide response to stress, allowed us to identify 151 high confidence p53 regulated genes which composed a large portion of a predefined DNA-damage stress response network. Therefore p53 plays an important role in regulating the stress response network.

Title

Performance analysis of *in-situ* visualization of large-scale codes

Authors

Umayanganie Munipala, Computational Science Program, UTEP

Shirley V. Moore, Department of Computer Science and Computational Science Program, UTEP

Abstract

As the computing power and capacity of super computers grow, complexity of the attempted problems on these machines has become detailed with high accuracy, thus the data generated by the massive scientific simulations have reached the scale of Petabytes and are still increasing in size. Generated data have become too large to analyze with conventional post processing. In-Situ Visualization techniques enable visualizations to be coupled with high performance computing simulation codes. Visualizations are generated while the simulation is running, enabling the user to monitor the simulation prior to getting the end result and enabling study of the current status of the visualization. This reduces the data storage required and transferring over networks. Integration of the visualization with the simulation has not yet been evaluated at large scale and could introduce unexpected complications and bottlenecks. We propose a performance and scaling analysis of in situ visualization coupled with the large-scale MPAS-Ocean simulation code, which is a project for developing ocean simulation components for use in climate, regional climate and weather studies.

Title

A novel *in-silico* drug designing approach for identification of natural compounds for treatment of hypothyroid

Authors

Pankhuri Wanjari[1] and R.M Jayadeepa[2]

[1]Department of Bioinformatics, National Institute of Technology Bhopal INDIA

[2]Bioinformatics program, Institute of Computational Biology, Bangalore INDIA

Abstract

Hypothyroidism is an endocrine disorder characterized by abnormally low thyroid hormone production, resulting in the deficiency of thyroid hormone. Hypothyroid can be treated by giving synthetic drug intake of L-thyroxin (T4) and L-triiodothyronine (T3) as a replacement or supplementary therapy in patients. Some of the drugs used for this purpose are Liotrix, Thyrolar and many more. The main objective of this research work is to find natural compounds that can help to stimulate thyroid hormone production. The molecules were screened based on the Lipinski's rule of 5. A total of 115 molecules were selected. 1NAV, a thyroid receptor alpha in complex with an agonist selective for thyroid receptor beta, was selected as the target protein. The molecules were subjected to docking analysis with 1NAV. The most effective compounds were isolated from Ashvagandha (*Withania Somnifera*), Astragalus (*Huang Qi*), Gotukola (*Centella Asiatica*), Triphala, Punarnava, Bauhinia Purpurea, Watercress and Dandelions (*Taraxacum Officinale*). The compounds were subjected to toxicity analysis and those that passed the toxicity tests were analyzed for binding site. From the active site analysis it was found that amino acids like Asp, Gly, Lys, Val, Asn, Gly, and Arg; which were present on site 7, showed the most prominent site. In silico docking results showed that 5 compounds viz Withaferin A, Calycosin-7-O-Beta-D-Glucoside, Withanolide D, Rotenone and Quercetin were the lead compound for the disease. Amongst these 5, Withaferin A and Calycosin-7-O-Beta-D-Glucoside gave the best results.

Title

Inverse problem optimization of 2-dimensional Bouguer gravity anomalies

Authors

Azucena Zamora, Computational Science Program, UTEP
Adrian E. Gutierrez, Department of Geological Sciences, UTEP
Aaron Velasco, Computational Science Program and Department of Geological Sciences, UTEP

Abstract

Bouguer gravity anomalies measured on the Earth's surface contain information about natural resources (e.g., hydrothermal vents, hydrocarbon deposits, ore bodies, salt domes, etc.) critical for their prospecting. Tracking changes in the gravity of the Earth can determine the density distributions of buried masses and constitutes an inexpensive, non-invasive, and non-destructive technique for the imaging of complex shallow Earth structures. However, inverse modeling of gravity data presents a very ill-posed mathematical problem, given that solutions are non-unique.

Usually, geophysicists and geologists implement "trial and error" forward model schemes to obtain density distributions of the Earth's shallow substructure. Implementing primal-dual interior point methods for constrained optimization we can achieve an optimal density model that minimizes the differences between calculated and observed gravity anomalies through the use of known density constraints for transitional areas. We reduce the solution space and overcome the ambiguity associated with the inversion of Bouguer gravity anomalies by focusing only on feasible models as constrained by other geophysical data (e.g. seismic and borehole data). Non-linear constrained optimization reduces the computer time required to obtain a relatively good anomaly fit and the amount of "manual" forward model-interpreter interaction.

Three synthetic tests and gravitational data obtained from a Porphyry Cu-Mo deposit formation (Cooper Flat Mine, Hillsboro at Sierra County, NM) are included. The final analysis reveals an improved 2-D density model for the Copper Flat mine and shows the juxtaposed density contrasts surrounding the ore body and the relationship among the geological features that shaped this copper porphyry deposit.