Title

Characterization of a reverse point mutation in the ATP-binding pocket of φEL, a virally-encoded Group III chaperonin

Authors

Ricardo Avila,[1] Humberto Rojo,[2] and Ricardo Bernal[1,2,3]

[1]Bioinformatics Program, [2]Department of Chemistry, and [3]Border Biomedical Research Center, The University of Texas at El Paso, El Paso, TX

Abstract

Chaperonins are large multimeric protein complexes whose function is to facilitate the folding of nascent or stress-denatured proteins. Their ubiquity across all domains of life is testimony to the vital role that they play in the cell. Until recently, chaperonins had been classified into two groups: I and II. Bacteriophage EL, a large virus that infects the gram-negative bacterium *Pseudomonas aeruginosa*, is source to the first reported virally-encoded chaperonin. This chaperonin, φEL, shares an ATP binding motif along the equatorial domain with GroEL, a well-studied group II chaperonin found in *Escherichia coli*. However, structural characterization of φEL has revealed a drastically different protein folding mechanism to that of group I and group II chaperonins, leading to its classification in a new group III. A single amino acid difference between the two motifs is hypothesized to play an important role in the novel mechanism of φEL, which features a ring separation step. This study uses a colorimetric assay to measure ATP hydrolysis, and structural characterization via Cryo-electron microscopy to analyze the effects of reversing this mutation. Results show that the mutant chaperonin has negligible ATPase activity, and a 20 angstrom Cryo-EM electron density map reveals that the mutant chaperonin is locked in its ATP-bound double-ring conformation, suggesting that the mutated residue plays an important role in the ring-separation step, without which the chaperonin cannot function.

Title

Artocarpus heterophyllus (jackfruit) seed and its peel: Potential functional food?

Authors

Kabita Baral,[1,2] Aashis Gurung,[2] Barsha Gurung,[2] Dristi Adhikari,[2] Rajan Gurung,[2] Sabeer Kandel,[2] and Atisammodavardhana Kaundinnyayana[2]

[1]Bioinformatics Program, The University of Texas at El Paso, El Paso, TX, and [2]School of Health and Allied Sciences, Pokhara University, Kaski, Nepal

Abstract

The aim of this study was to conduct ethnomedicinal survey and in vitro studies of methanolic extract of *Artocarpus heterophyllus* (jackfruit) seed and its peel in order to study its potential as function food to aid in minimizing world hunger and food waste. Questionnaire regarding consumption, health benefits, utilization of seed and seed's taste was prepared and 240 respondents of the selected wards of Lekhnath municipality were interviewed. Further, sample was screened for phytochemical evaluation and general tests; DPPH activity, anti-bacterial activity and total flavonoid content were also measured. Also, the potential of its use as functional food is highlighted. It was found that most people have been consuming jackfruit as fruit and vegetable for its taste rather than health benefits; consuming seed but discarding the peel. Phytochemical analysis revealed the presence of flavonoid in both seed and peel whereas alkaloid and saponin were detected only in peel while general tests revealed that acid value, saponification value, ester value and moisture content were low. Further, DPPH scavenging assay showed that peel has better antioxidant activity than seed which could be correlated to total flavonoid content. Anti-bacterial activity of seed was detected against *Proteus mirabilis* and *Salmonella enterica typhi* only while that of peel was not detected. Various researches support the fact that seed and peel possess various micro and macro nutrients making it promising functional food to be brought forth in market as confectionary, snack and powder.

Title

In silico drug design approach to insomnia therapeutics

Authors

Madan Baral,[1] Ricardo Avila,[1] Haidar Ahmed,[2] Jonathon E. Mohl,[1,3,4] Ming-Ying Leung,[1,3,4,5] and Rachid Skouta[2,4,6]

[1]Bioinformatics Program, [2]Department of Biological Sciences, [3]Computational Science Program, [4]Border Biomedical Research Center, [5]Department of Mathematical Sciences, and [6]Department of Chemistry, The University of Texas at El Paso, El Paso, TX

Abstract

G protein-coupled receptors (GPCRs) are a ubiquitous family of seven-transmembrane proteins that play important roles in signal transduction pathways making them potential targets of modern drugs. Though a vast array of *in silico* drug design and discovery tools to comprehend ligand-protein interactions has been introduced, it is desirable to develop a workflow for drug discovery, using the best available tools to develop novel GPCR targeted drugs. The purpose of this study is to establish a general workflow of *in silico* drug discovery taking the orexin receptor (a GPCR) system which is associated with insomnia and generate a pharmacophore model to further work on ligand based virtual screening to design insomnia therapeutics. We obtained the three available orexin receptors' 3D structures 4S0V, 4ZJ8 and 4ZJC from Protein Databank (PDB) and a total of 139 3D orexin receptor inhibitor ligands were generated using LigPrep tools and performed molecular docking using AutoDock Vina, AutoDock Vina XB, Glide standard precision (SP), Glide extra precision (XP) and Induced Fit Docking (IFD). Pharmacophore modeling was performed using Phase. Comparative docking study using Glide and Vina inferred that dual orexin receptor antagonist (DORA) class of ligands had the best binding affinity when compared to 1-specific orexin receptor antagonists (1-SORA) and 2-specific orexin receptor antagonists (2-SORA). Pharmacophore models on DORA generated 6 pharmacophores which is consistent with the pharmacophores of co-crystallized Suvorexant, a FDA-approved drug for the treatment of insomnia. For future study, we will perform ligand based virtual screening to obtain some potential dual orexin receptor inhibitors.

Title

Computational methods for prediction and classification of G protein-coupled receptors

Authors

Khodeza Begum[1] and Ming-Ying Leung[1,2,3]

[1]Computational Science Program, [2]Bioinformatics Program, and [3]Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, TX

Abstract

G protein-coupled receptors (GPCRs) constitute the largest group of membrane receptor proteins in eukaryotes. Due to their significant roles in many physiological processes such as vision, smell, and inflammation, GPCRs are targets of many prescription drugs. However, the functional and structural diversity of GPCRs has kept their prediction and classification based on amino acid sequence data as a challenging bioinformatics problem. As existing computational methods to predict and classify GPCRs are focused on mammalian (mostly human) data, the ultimate goal of our project is to establish an ensemble approach and implement a web-based software that can be used reliably on a wider range of organisms. As a first step, we have constructed a searchable MySQL database with experimentally confirmed GPCRs and non-GPCRs along with protein features for distinguishing them. This database currently contains 2887 GPCR and 1614 non-GPCR sequences collected from the UniProtKB/Swiss-Prot protein database, covering over 300 species including arthropods, fungi, nematode, etc. Each protein in the database is assigned a unique identification number and linked to information about its source organism, sequence lengths, and other features including amino acid and dipeptide composition. For the GPCRs, family classifications according to the popular GRAFS and IUPHAR systems are also included. This database will provide the training and testing data for subsequent steps in our ongoing work to evaluate existing computational tools, incorporate them into our ensemble, and apply them to identify potential GPCRs in several fly, mosquito, and tick species that are of biomedical or agricultural importance.

Title

Deconvoluting the expression dynamics of m^6A post-transcriptional enzymes and their crosstalk with RNA binding proteins

Authors

Daniel Castaneda-Mogollon[1], Arun Kumar Boddapati,[2] and Sarath Chandra Janga[2,3,4]

[1]Bioinformatics Program, The University of Texas at El Paso, El Paso, TX, [2]Department of Biohealth Informatics, School of informatics and Computing, Indiana University-Purdue University, [3]Center for Computational Biology and Bioinformatics, and [4]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN

Abstract

The m^6A post-transcriptional methylation is a reversible process present in eukaryotes that occurs mostly in mRNA, but also in tRNA, rRNA, snRNA, and long non-coding RNA. This methylation event has been linked to translational output increase, mRNA decay, splicing regulation, export, cancer association, and among many others. We have identified nine important enzymes that have an active participation in this process, and they have been classified into three groups: writers, readers and erasers. It is unclear whether these enzymes act independently or if they perform their enzymatic activities in a more dynamic manner. To test this, we used Cytoscape 3.5.1 to create a network that displays significant coexpression correlation values between m^6A post-transcriptional enzymes and RBPs (RNA binding proteins). Our results show that m^6A erasers seem to act in a more independent manner than writers and readers. Furthermore, m^6A readers have the highest number of significant interactions with other RBPs. Moreover, we found that some m^5C enzymes appear to have a strong co-expression and/or physical interaction with m^6A enzymes. ELAVL1 (HuR) was found to be a potential m^6A reader, as it interacts with the three types of m^6A post-transcriptional enzymes, and its binding behavior is similar to other m^6A readers. In order to determine which cancer types are associated with m^6A enzymes, twenty-eight types of cancers gene expression data were analyzed from FireBrowse. The obtained results indicate that writers and erasers exhibit a higher number of significant differences in gene expression when compared to the control group.

Title

Analysis of prey-predator data by using a stochastic system

Authors

Sumi Dey[1] and Maria C. Mariani[1,2]

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

Abstract

In this work, we propose a system of stochastic differential equations to analyze the prey-predator data. The model is based on the independent Ornstein-Uhlenbeck processes driven by a Gamma process. The independent Ornstein-Uhlenbeck processes is analytically flexible and in addition provides a class of continuous time processes that exhibits long memory behavior. In order to perform our analysis, we first estimate the model parameters and then simulate data using our stochastic system. The simulated data mimic the original prey-predator data, which is observed from the estimates of the root mean square error.

Title

MindBack for brain therapy

Authors

Godínez-Macías, Karla P.,[1,2] Perez-Mata Fernando,[2] Meza-Kubo, Victoria,[2] Flores-Salas, Juan Carlo,[2] and Hernández-Saijas, Manuel[3]

[1]Bioinformatics Program, The University of Texas at El Paso, El Paso, TX, [2]Universidad Autónoma de Baja California, Mexicali, B.C., Mexico, and [3]Centro de Rehabilitación Integral de Ensenada, Ensenada, B.C., Mexico

Abstract

MindBack is a brain computer interface software intended to be used by specialists as a non-invasive tool during cognitive therapy to their patients with cerebrovascular disease (CVA) that cause deficiencies in language, memory, and attention. This software was developed to be used with Mindwave from Neurosky®, a headset with biosensors to collect, digitize, and amplify electrical signals from brain activities and translate them into inputs. To avoid stress for the patient, the therapist can also use the mouse as input device when needed. This new technology provides a new technique for CVA therapy using human-computer interactions to implement the traditional exercises that had led to total or partial recovery of the patient. Following the System Usability Scale protocol, a "within subjects" usability analysis was performed on 36 adults with age between 56 and 76 and no prior CVA. We sampled subjects from this population sector because of their propensity to CVA, and conducted the evaluation without affecting CVA patients. The majority of subjects reported using the MindBack software frequently, as they found it easy and comfortable to use. They also thought the activities were well designed but had trouble concentrating during the activities. We therefore concluded the key for full acceptance of the software is the adjustment of concentration levels so that the electrical signals would be strong enough for the subjects to do the activity. Future work will include evaluating patients with CVA to determine their response to MindBack, and adjust the software according to their special needs.

Title

Optimization of micro-pillar wick structured cooler by using an exa-scale pore network simulator

Authors

Shaikh Tanveer Hossain,[1] Arturo Schafino,[2] Ashesh Chattopadhyay,[3] Murali Kotteda,[2] and Vinod Kumar[1,2]

[1]Computational Science Program and [2]Department of Mechanical Engineering, The University of Texas at El Paso, El Paso, TX, and [3]Department of Mechanical Engineering, Rice University, Houston, TX

Abstract

Continuous increment of compactness of electronic devices throws a great challenge to the heat management process. Thin-film evaporation in wick structures is simple, self-powered, auto-maintained, and reliable with high heat extraction rate utilizing the latent heat of vaporization. Cylindrical shapes are most efficient structures for the repeating element in an array but they are good only for hypothetical array-type structures of same repeating geometric shapes with flow analysis of that small particular section only. Such structures might fail for scattered-type arrangement that could be more efficient if engineered properly. Therefore, room for more optimization of heat management still exists. We have developed a code based on “network and resistance” theory, where a porous medium is compared and simplified to a network of flow paths. The code uses trilinos package (an algebraic solver package developed by Sandia National Lab) and shows great scalability in parallel computing environment. This enables us to analyze structures up to 10,000 million pores for greater understanding of the problem. The resistance of the network paths in this case is dependent on the shape factors of the solid geometry and the shape and height of the meniscus. To determine the resistance, we have implemented and trained a separate neural network code to generate a databank using a large number of Ansys simulations with varieties in the controlling parameters. With this combined process, we are able to have a wider look on the problem and thus we can find the best comprehensive structure for a particular case.

Title

Volatility forecasting of financial time series by using Ornstein-Uhlenbeck type models

Authors

Maria C. Mariani,[1,2] Md Al Masum Bhuiyan,[2] and Osei K. Tweneboah[2]

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

Abstract

In this work, we discuss a statistical technique to estimate the volatility of a financial time series by using an Ornstein-Uhlenbeck type models. Using the daily closing prices from developed and emergent stock markets, we conclude that the incorporation of stochastic volatility into the time-varying parameter estimation improves the forecasting performance via Maximum Likelihood Estimation (MLE). A class of stochastic differential equation arising on the superposition of two independent Gamma Ornstein-Uhlenbeck processes is used to simulate the time series data in a special case where the MLE does not fit the original data. The simulated data mimics the original financial time series, which is observed from the estimates of the root mean square error. Furthermore, the stochastic model exhibits the physical and long memory behavior of the data. We also conclude that the Ornstein-Uhlenbeck type models used in this study guarantees the convergence of the MLE technique, which makes the estimation algorithm feasible with large data sets and facilitate prediction.

Title

Application of SIC-DFT on some closed-shell atoms and molecules

Authors

Kamal Nyaupane[1] and Tunna Baruah[1,2]

[1]Computational Science Program and [2]Department of Physics, The University of Texas at El Paso, El Paso, TX

Abstract

DFT (density functional theory) calculations employ approximations to the exchange-correlation functional, which are not free from self-interaction. The self-interaction error in DFT appears from the fact that the residual self-interaction in the Coulomb part and that in the exchange part do not cancel each other exactly. We examine the recently developed SIC (self-interaction correction) scheme based on Fermi-Lowdin orbitals on closed-shell atoms. The final motivation is to apply the scheme to different transition metal based clusters. The self-interaction leads to delocalized d-orbitals and can lead to incorrect spin states in some systems. We present an ongoing study of the relative energies of various molecular FeO_4 isomers in various oxidation states of both Fe and O: Fe(II)-$3d^6$ $(O_2^{1-})_2$, Fe(III)-$3d^5$ (O_2^{2-}) (O_2^{1-}), Fe(IV)-$3d^4$ $(O_2^{2-})_2$, Fe(V)-$3d^3$ $(O^{2-})_2$ (O_2^{1-}), Fe(VI)-$3d^2$ $(O^{2-})_2$ $(O^{1-})_2$, Fe(VII)-$3d^1$ $(O^{2-})_3$ (O^{1-}) and Fe(VIII)-$3d^0$ $(O^{2-})_4$ through the fermi-orbital based scheme.

Title

Phytochemical screening, cytotoxic and antimicrobial activity of selected Nepalese medicinal plants around Kavrepalanchowk

Authors

Sumod Pokharel,[1,2] Saurav Adhikari,[2] Sumit Gautam,[2] Prashanna Guragain,[2] Nikunj Pradhan,[2] Sabin Sijapati,[2] Rajendra Gyawali,[2] and Tirtha M. Shrestha[2]

[1]Bioinformatics Program, The University of Texas at El Paso, El Paso, TX, and [2]Department of Pharmacy, Kathmandu University, Dhulikhel, Nepal

Abstract

Medicinal plants are gaining recognition worldwide in the treatment and control of diseases. Antibacterial properties of Nepalese plant species - *Myrica esculenta, Mahonia nepaulensis, Madhuca longifolia,* and *Schima wallichii* were evaluated on human pathogenic microorganisms *Salmonella typhi, Staphylococcus aureus,* and *Escherichia coli*. The plants were found to possess phytochemicals saponins, glucosides, flavonoids, tannins, and alkaloids. Methanolic extract from these plants showed antimicrobial activity against tested organisms. The antimicrobial test was carried out for various plant methanolic extract by disc diffusion method using the microorganisms. The zone of inhibition obtained for each plant extract in three different concentrations 3%, 6%, and 9% in the antimicrobial test was compared with that obtained from the standard antibiotic Ciprofloxacin to visualize the significant antimicrobial effect. In addition, cytotoxic test was carried out by "Brine Shrimp Lethality test" in which plant methanolic extract were subjected for their lethality to the matured Brine shrimp. Methanolic extracts showed the LC_{50} (Lethal Concentration 50%) values of 15.4 ppm, 136.4 ppm, 76.9 ppm, and 76.0 ppm for *M. esculenta, M. nepaulensis, M. longifolia,* and *S. wallichii* respectively. When these LC_{50} values were compared to previously established classification of plant toxicity levels, *M. esculenta* was considered moderately toxic, *M. longifolia* and *S. wallichii* mildly toxic, and *M. nepaulensis* nontoxic, suggesting that these plants would be safe to human consumption for drug purpose. So, attention to these medicinal plants can be drawn for herbo-therapeutic treatment or antimicrobial plant principles in the case of infection by selected microorganisms.

Title

Evolutionary studies of furin serine protease enzyme in *Giardia lamblia*

Authors

Roshani Rajapaksha[1] and Siddhartha Das[1,2]

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

Abstract

Giardia lamblia is one of the early-divergent eukaryotes and the organization of its endomembrane systems appears to be substantially different from higher eukaryotes. Instead of highly organized Golgi complexes, *Giardia* contains encystation specific vesicles (ESV) for protein trafficking. Although ESVs share no morphologic similarities typical to Golgi complexes, several known Golgi markers are found to be located in ESVs. Since furin-serine protease (FSP) enzyme is involved in Golgi maturation and a FSP homologue has been annotated in the *Giardia* here, we report the use of bioinformatics tools to make functional and evolutionary inferences about FSP, which may shed light on the evolution of Golgi maturation process. A protein BLAST was performed against UniProt database and homologous sequences of giardial FSP (gFSP) were used to construct a protein network using EFI-EST tool, which resulted in three main clusters. To provide a detailed view of the evolutionary relationships across the clusters, a phylogenetic tree was constructed using MEGA7 software. To examine the conservation of primary sequence of gFSP, a multiple sequence alignment was performed using ClustalW which resulted in several amino acids to be highly conserved across different groups. The three-dimensional (3-D) structure of gFSP was constructed using I-TASSER server. Network and phylogenetic analysis indicated that gFSP and its homologues are present in bacteria and mammals. We will further investigate the conservation of residues using the 3-D structure of gFSP. Our long-term goal is to overexpress and knockdown gFSP and evaluate its function in encysting *G. lamblia* by molecular and biochemical analysis.

Title

Review Presentation: "Identification of cancer risk lncRNAs and cancer risk pathways regulated by cancer risk lncRNAs based on genome sequencing data in human cancers"

Presenter

Marcela Aguilera-Flores, Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

Abstract

This poster is a review of the paper "Identification of cancer risk lncRNAs and cancer risk pathways regulated by cancer risk lncRNAs based on genome sequencing data in human cancers" by Li *et al*. (2016) published in the journal *Nature Scientific Reports*. The abstract as it appears in the original publication is as follows:

“Cancer is a group of diseases involving abnormal cell growth with the potential to invade or spread to other parts of the body. The complexity of cancer can be reduced to a small number of underlying principles like cancer hallmarks which could govern the transformation of normal cells to cancer. Besides, the growth and metastasis of cancer often relate to combined effects of long non-coding RNAs (lncRNAs). Here, we performed comprehensive analysis for lncRNA expression pro les and clinical data of six types of human cancer patients from The Cancer Genome Atlas (TCGA), and identified six risk pathways and twenty three lncRNAs. In addition, twenty three cancer risk lncRNAs which were closely related to the occurrence or development of cancer had a good classification performance for samples of testing datasets of six cancer datasets. More important, these lncRNAs were able to separate samples in the entire cancer dataset into high-risk group and low-risk group with significantly different overall survival (OS), which was further validated in ten validation datasets. In our study, the robust and effective cancer biomarkers were obtained from cancer datasets which had information of normal- tumor samples. Overall, our research can provide a new perspective for the further study of clinical diagnosis and treatment of cancer.”

Title

Review Presentation: “Retinal transcriptome profiling by directional next-generation sequencing using 100 ng of total RNA”

Presenter

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

Abstract

This poster is a review of the paper “Retinal Transcriptome Profiling by Directional Next-Generation Sequencing Using 100 ng of Total RNA” by Brooks *et al.* (2012) published in *Retinal Development: Methods and Protocols*, S.-Z. Wang (ed.), Methods in Molecular Biology, Volume 884, Chapter 13. The abstract as it appears in the original publication is as follows:

“RNA expression profiles produced by next-generation sequencing (NGS) technology (RNA-seq) allow comprehensive investigation of transcribed sequences within a cell or tissue. RNA-seq is rapidly becoming more cost-effective for transcriptome profiling. However, its usage will expand dramatically if one starts with low amount of RNA and obtains transcript directionality during the analysis. Here, we describe a detailed protocol for the creation of a directional RNA-seq library from 100 ng of starting total RNA.”

Title

Review Presentation: “Molecular phylogenetics of ponerine ants (Hymenoptera: Formicidae: Ponerinae)”

Presenter

Abdulmeneem Joma, Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

Abstract

This poster is a review of the paper “Molecular phylogenetics of ponerine ants (Hymenoptera: Formicidae: Ponerinae)” by Schmidt (2013) published in the journal *Zootaxa*. The abstract as it appears in the original publication is as follows:

“Recent molecular phylogenetic studies of ants (Hymenoptera: Formicidae) have revolutionized our understanding of how these ecologically dominant organisms diversified, but detailed phylogenies are lacking for most major ant subfamilies. I report the results of the first detailed phylogenetic study of the ant subfamily Ponerinae, a diverse cosmopolitan lineage whose properties make it an attractive model system for investigating social and ecological evolution in ants. Molecular sequence data were obtained from four nuclear genes (*wingless*, *long-wavelength rhodopsin*, *rudimentary* [CAD], 28S rDNA; total of ~3.3 kb) for 86 ponerine taxa, representing all three ponerine tribes, 22 of the 28 currently recognized genera, and 14 of the 18 informal subgenera of *Pachycondyla*, a heterogeneous grouping whose monophyly is doubtful on morphological grounds. Phylogenetic reconstructions using maximum likelihood and Bayesian inference support the monophyly of Ponerinae and tribe Platythyreini, but fail to support the monophyly of the large tribe Ponerini due to its inclusion of the unusual genus *Thaumatomyrmex*. *Pachycondyla* is inferred to be broadly non-monophyletic. Numerous novel generic and suprageneric relationships are inferred within Ponerini, which was found to consist of four major multi-generic clades (the *Ponera*, *Pachycondyla*, *Plectroctena* and *Odontomachus* genus groups) plus the single genera *Hypoponera* and *Harpegnathos*. Uncertainty remains in some regions of the phylogeny, including at the base of Ponerini, possibly reflecting rapid radiation. Divergence dating using a Bayesian relaxed clock method estimates an origin for stem Ponerinae in the upper Cretaceous, a major burst of diversification near the K/T boundary, and a rich and continual history of diversification during the Cenozoic. These results fail to support the predictions of the “dynastic-succession hypothesis” previously developed to explain the high species diversity of Ponerinae. Though model-based reconstructions of historical biogeography and trait evolution were not attempted in this study, the phylogeny suggests that ponerine evolution was marked by regionalized radiations and frequent faunal exchange between major biogeographic provinces. The reported results also imply multiple origins of cryptobiotic foraging, mass raiding behavior, and gamergate reproduction within Ponerinae, highlighting the value of the subfamily as a model for studying the incipient evolution of these and other ecological and behavioral traits.”

Title

Review Presentation: "The *Drosophila* Gene Expression Tool (DGET) for expression analysis"

Presenter

Regina Umarova, Bioinformatics Program, The University of Texas at El Paso, El Paso, TX

Abstract

This poster is a review of the paper "The *Drosophila* Gene Expression Tool (DGET) for expression analysis" by Hu *et al.* (2017) published in the journal *BMC Bioinformatics*. The abstract as it appears in the original publication is as follows:

"**Background:** Next-generation sequencing technologies have greatly increased our ability to identify gene expression levels, including at specific developmental stages and in specific tissues. Gene expression data can help researchers understand the diverse functions of genes and gene networks, as well as help in the design of specific and efficient functional studies, such as by helping researchers choose the most appropriate tissue for a study of a group of genes, or conversely, by limiting a long list of gene candidates to the subset that are normally expressed at a given stage or in a given tissue.

Results: We report DGET, a Drosophila Gene Expression Tool (www.flyrnai.org/tools/dget/web), which stores and facilitates search of RNA-Seq based expression profiles available from the modENCODE consortium and other public data sets. Using DGET, researchers are able to look up gene expression profiles, filter results based on threshold expression values, and compare expression data across different developmental stages, tissues and treatments. In addition, at DGET a researcher can analyze tissue or stage-specific enrichment for an inputted list of genes (e.g., 'hits' from a screen) and search for additional genes with similar expression patterns. We performed a number of analyses to demonstrate the quality and robustness of the resource. In particular, we show that evolutionary conserved genes expressed at high or moderate levels in both fly and human tend to be expressed in similar tissues. Using DGET, we compared whole tissue profile and subregion/cell-type specific datasets and estimated a potential source of false positives in one dataset. We also demonstrated the usefulness of DGET for synexpression studies by querying genes with expression profile similar to the mesodermal master regulator Twist.

Conclusion: Altogether, DGET provides a flexible tool for expression data retrieval and analysis with short or long lists of Drosophila genes, which can help scientists to design stage- or tissue-specific in vivo studies and do other subsequent analyses."