SIXTH ANNUAL
Arthropod Genomics
S Y M P O S I U M and
i5k Community Workshop

May 30 - June 2, 2012 • Kansas City

POSTER ABSTRACTS

*Indicates name of person presenting poster

♦♦♦Cross-reference list on last page lists first authors alphabetically.♦♦♦

SESSION: EMERGING GENOMES

EG-1
Improving the Honey Bee Consensus Gene Set
*Bennett, Anna K.; Elsik Christine G.
Georgetown University, Department of Biology, Washington, DC 20057

We produced an improved official gene set (OGSv3.1) for honey bee (*Apis mellifera*) to facilitate comparative analyses and manual annotation. Challenges encountered in generating the first honey bee official gene set (OGSv1.0), published in 2006, included the AT richness of the genome, highly heterogeneous GC composition, limited EST/cDNA data and the large evolutionary distance between honey bee and other sequenced genomes. As a part of the Bee Genome Sequencing and Analysis Consortium, the Baylor College of Medicine Human Genome Sequencing Center sequenced the genomes of two closely related species, dwarf honey bee (*A. florea*) and buff-tailed bumble bee (*Bombus terrestris*); added sequence and improved the draft *A. mellifera* genome; and deep sequenced multiple *A. mellifera* tissue transcriptomes. These data were used to detect genes believed to be missing from OGSv1.0 and improve the gene set.

Genes were predicted using a variety of methods (NCBI RefSeq and Gnomon, Augustus, SGP2, GeneID, FgenesH++ and N-SCAN), each leveraging the new transcriptome data. The Augustus analysis also incorporated available peptide data, and the N-SCAN analysis leveraged sequence conservation between the *A. mellifera* genome and the other bee genomes. We used GLEAN to create several consensus gene sets with different combinations of the input gene prediction sets, and evaluated the consensus sets by comparing them to high quality annotations, cDNA sequences and peptides, before deeming one of the sets to be the official gene set.

OGSv3.1 represents a significant improvement as it includes 16,038 genes, 5,881 genes more than OGSv1.0. To determine whether genes that were not detected in OGSv1.0 have common characteristics that make them more challenging to predict, we compared them to the OGSv1.0 genes. We evaluated features such as tissue expression specificity, coding feature length, GC content and existence of arthropod homologs. Previously unannotated genes were found to have shorter average coding lengths and are more likely to be expressed in a tissue specific manner.

The improved honey bee gene set will be invaluable to the honey bee research community in efforts to elucidate the mechanisms behind fundamental biological processes, such as evolution of insect eusociality, as well as agricultural issues, such as pollinator health and immunity. Furthermore, understanding the reasons genes were not predicted in OGSv1.0 will lead to more effective gene prediction strategies for new genome projects.

EG-2

Sequencing the genome of the spruce budworm, *Choristoneura fumiferana*

*Cusson, M.; Maaroufi, H.; Nisole, A.; Boyle, B.; Laroche, J.; Béliveau, C.; Kukavica-Ibrulj, I.; Doucet, D.; Lucarotti, C.; Sperling, F.; Lumley, L.; Brunet, B.; Bird, H.; Nealis, V.; Levesque, R.C.
(MC, CB, LL) Natural Resources Canada (NRCan), Laurentian Forestry Centre, Québec, QC, G1V 4C7; (DD) NRCan, Great-Lakes Forestry Centre, Sault Ste. Marie, ON, P6A 2E5; (CL) NRCan, Atlantic Forestry Centre, Fredericton, NB, E3B 5P7; (VN) NRCan, Pacific Forestry Centre, Victoria, BC, V8Z 1M5; (HM, AN, BBo, JL, KII, RCL) Institut de biologie intégrative et des systèmes, Université Laval, Québec, QC G1V 0A6; (FS, BBr, HB) Dept. of Biological Sciences, University of Alberta, Edmonton, AB, T6G 2E9.

The spruce budworm, *Choristoneura fumiferana*, is the most devastating insect pest of spruce-fir forests in North America. Populations undergo cyclical changes in density and reach outbreak levels every 30-40 years, causing significant tree mortality. Current management approaches are dominated by the application of *Bacillus thuringiensis* to protect trees where budworm populations have reached outbreak levels. As a first step towards developing alternative management strategies, we recently undertook the sequencing and assembly of the budworm genome. We report on our strategy and progress made in genome assembly and annotation, and provide an overview of ongoing companion studies.

Our sequencing strategy combines several approaches and NGS technologies including: (i) Roche 454 GS-FLX and (more recently) GS-FLX+ shotgun sequencing, using DNA extracted from a single male pupa, (ii) 454 GS-FLX sequencing of 6 kb and 20 kb paired-end libraries using DNA extracted from a small pool of male pupae, (iii) GAIIx Illumina sequencing of a 2 kb paired-end library, using the same DNA employed for the 454 shotgun sequencing, and (iv) 454 GS-FLX+ sequencing of pools of 50 BAC clones (average insert size of 140 kb; BAC library constructed using DNA extracted from ~5 g of larvae). For identification of expressed ORFs during the annotation process, we are also conducting 454 GS-FLX+ sequencing of a normalized cDNA library, developed from a pool of RNAs extracted from all budworm life stages. These transcriptomics resources will complement existing EST libraries developed by our group.

To date, 454 GS-FLX shotgun and 6 kb paired-end sequencing has generated 12.9 million reads with an average length of 376 nts and a modal peak of 500 nts, achieving 11x coverage of the estimated 450 Mb genome, distributed over 30 chromosomes. By comparison, Illumina sequencing has generated 512 million reads, representing 100x genome coverage. Genome assembly using Newbler, with all available 454 sequences and a portion of the Illumina sequences as input, has generated 28,000 scaffolds, representing ¾ of the chromosomal genome, while assembly of the mitochondrial genome is complete and contained within a single 15.53 kb scaffold. Additional sequencing is currently underway (454 GS-FLX+ shotgun, 454 GS-FLX 20 kb paired-end and BAC sequencing) and is expected to substantially reduce the number of scaffolds. Initial estimates of the proportion of repetitive DNA indicate it is low (7.2%) compared to *Bombyx mori* (38.4%). Mapping of the available scaffolds onto the 28 *B. mori* chromosomes suggests extensive synteny and gene conservation between the two species. Through this comparative process, a preliminary annotation of the budworm genome has been generated.

Several companion studies are currently underway, including the Illumina sequencing of the genomes of other *Choristoneura* species within the "*C. fumiferana* species complex" (*C. pinus*, *C. biennis*, *C. occidentalis*). Although these species can interbreed, reproductive isolation is maintained through differences in life history and behavioural traits, such as differences in host associations and/or sex pheromone blends. Nonetheless, their genomes are expected to be very similar. We thus plan to map Illumina reads onto the assembled *C. fumiferana* genome. In addition, two RNA-seq transcriptomics studies have been initiated: one comparing transcripts from diapausing and non-diapausing *C. fumiferana* strains (eggs, 1st and 2nd instars) to identify genes involved in the induction of diapause, and another one comparing pheromone biosynthesis and perception-related transcripts in relevant tissues (pheromone gland, antennae) of *C. fumiferana* and *C. pinus* to identify adaptive genes involved in species delimitation. Finally, a genotype-by-sequencing approach is being used to identify SNP markers specific to regional populations of the budworm within its geographical distribution. These markers will be used to monitor budworm moth migration. A similar approach is also being applied to hybrid *Choristoneura* families to identify SNPs for the construction of the first genetic linkage maps for spruce budworm species.

EG-3

Discovery and characterization of the first mosquito Y chromosome genes

*Hall, A. Brantley; Criscione, Frank; Qi, Yumin; Tu, Zhijian (Jake)
Virginia Tech, Biochemistry, Blacksburg, VA, 24061

Y chromosome genes are serially underrepresented in published genomes, and have only been characterized in a few species. Using Identify-Y, a novel algorithm to discover Y chromosome sequences, we report the identification and verification of the first three mosquito Y chromosome genes.

Identify-Y analyzes the number of alignments to a query sequence from sex-specific next-generation sequencing databases to gauge differences in relative abundance between males and females. To detect differences in relative abundance, Identify-Y calculates a value called the chromosome quotient, the number of female alignments divided by the number of male alignments. Autosomal sequences are present in the same number in males and females, and subsequently have a chromosome quotient near one. X chromosome sequences are twice as numerous in females as in males, and subsequently have a chromosome quotient near two. Sequences unique to the Y chromosome are only present in males, and subsequently have a chromosome quotient of zero. Y chromosome sequences that have autosomal or X chromosome homologues have chromosome quotients greater than zero, but often less than the chromosome quotients of autosome or X chromosome sequences. We discovered a range of chromosome quotients near zero, distinct from the chromosome quotients of autosome and X chromosome sequences, which is highly enriched for Y chromosome sequences. In the malaria mosquito *Anopheles stephensi*, we report the identification of 462 Y chromosome sequences in this range.

After comparing the Y chromosome sequences to RNA-seq databases, we identified three novel Y chromosome genes and confirmed their Y status by either digital PCR, or chromosomal *in situ* hybridization. The first two genes, which we named GUY1 and GUY2, were unique to the Y and have no homology to any known genes. GUY1 is expressed exclusively in the early embryo and GUY2 is expressed throughout various time points. The third gene, which we named GOY1, is homologous to a transcription factor critical to embryonic development, and is expressed in the *A. stephensi* early embryo. GOY1 is the result of a duplication event from chromosome 2L, and retains high sequence similarity with its paralog. We are investigating the functions of these genes by RNA interference in hopes of finding the canonical male-determining gene, or other functionally important genes on the Y.

EG-4

Transcriptomics of the invasive brown marmorated stink bug (*Halyomorpha halys*)

Ioannidis, Panagiotis[1]; Lu, Yong[2]; Kumar, Nikhil[1]; Creasy, Todd[1]; Ott, Sandra[1]; Tallon, Luke J[1]; Pick, Leslie[2], Dunning Hotopp, Julie C.[1,*]
[1]Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD 21201
[2]Department of Entomology, University of Maryland, College Park, MD 20742
*Corresponding author

The brown marmorated stink bug (BMSB), *Halyomorpha halys* (Hemiptera: Pentatomidae), is an invasive species in the mid-Atlantic region of the US that is native to Asia. They were first found in the US in Allentown, PA and since then they have spread to many neighboring states. Their name is due to the foul odor they produce by specialized glands in their abdomens when threatened. While quite well known to the general public because they invade residences in the fall in search of overwintering sites, the biggest threat is to agriculture. The effects on agriculture are devastating as they cause damage in various crops such as soybeans, apples, peaches and raspberries. In 2010 alone, BMSB were estimated to cause $37 million in damages. Most importantly, a much greater number of crops are at risk of BMSB infestation because early this year they were detected as south as Florida. Controlling BMSB populations becomes a very difficult task because natural predators for this introduced pest are lacking in the US, and BMSB are resistant to most of the reduced-risk insecticides. Application of other, less specific insecticides could have long-term negative effects on the ecosystem and also on beneficial insects. Given that they have only recently become a significant agricultural pest to be studied, only ~100 bp of sequence data was available in the public databases for the BMSB. Therefore we undertook a transcriptome-based sequencing project in order to provide sequence data to the community trying to develop control strategies. RNA was isolated from more than ten different samples including: embryos; 1st, 2nd, 3rd, 4th, 5th instar nymphs; active adult males and females; and males and females in diapauses. PolyA-selected RNA from two pools of RNA were normalized with Dsn and sequencing on an Illumina HiSeq generating 366 million reads and 37 Gbp of data. The transcriptome was assembled with Trinity to generate 202,598 sequences including sequence and splice variants. The length of transcripts varied from 201 – 26,984-bp, with a mean length of 766 ± 1,221-bp, a median of 368-bp and a mode of 202-bp. Annotation of the longest splice variant of each gene took place using the IGS Annotation Engine. Both basic and applied research is underway to further characterize key BMSB genes and their potential as targets for pest control.

EG-5

De novo* assembly of three symbiont genomes of the Whitefly *Bemisia tabaci

Jiang, Zi-Feng *; Ghanim, Murad*; Xia, Fang-Fang; Stevens, Rick; White, Kevin P.
(ZFJ, KPW) Department of Human Genetics, The University of Chicago, Chicago, IL 60637; (FFX, RS) Argonne National Laboratory, Argonne, IL 60439; (MG) The Agricultural Research Organization (ARO), Department of Entomology, Volcani Center, Israel 50250

The sweet potato whitefly *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) is one of the most globally damaging insect pests in open fields and protected agricultural crops, causing annual losses estimated by 1-2$ billion, and is one of the top 100 invasive species worldwide (http:// **www.issg.org/database**). *B. tabaci* harbors the primary symbiont *Porteria* as well as six different secondary symbionts belonging to the genera *Arsenophonus, Cardinium, Fritschea, Hamiltonella, Rickettsia,* and *Wolbachia.* Bacteriocytes, the specialized cells that harbor these bacteria, as well as whole insects from two *B. tabaci* biotypes known as B and Q, were used for sequencing the whole genomes of three symbionts: *Porteria, Hamiltonella*, and *Rickettsia.* These results will advance our understanding of the host-bacteria interactions in whiteflies, for developing novel methods for whitefly control.

EG-6

The development of the genomic database of the diamondback moth (*Plutella xylostella*) and genomic analysis using it

*Jouraku, Akiya; Yamamoto, Kimiko; Kuwazaki, Seigo; Urio, Masahiro; Suetsugu, Yoshitaka; Narukawa, Junko; Miyamoto, Kazuhisa; Kurita, Kanako; Kanamori, Hiroyuki; Katayose, Yuichi; Matsumoto, Takashi; Noda, Hiroaki
National Institute of Agrobiological Sciences 1-2 Owashi, Tsukuba, Ibaraki, Japan

The diamondback moth (DBM), *Plutella xylostella*, is one of the most harmful insect pests against crucifer crops due to its high resistance to conventional insecticides such as pyrethroids and *Bacillus thuringiensis* (BT). In order to contribute to the analysis of the insecticides resistance-related genes and the development of the efficient insect pest control methods, we have been developing a genomic database of the DBM as one of the first steps for the development of integrated databases for broad agricultural pests. The database currently provides the following sequence data of the DBM: (1) WGS contigs (3.8x coverage, 454 GS-FLX Titanium), (2) predicted genes CDS generated from the WGS contigs, (3) unigene sequences constructed from RNA-seq sequences (20Gbp HiSeq2000) and 22,034 EST/mRNA sequences (extracted from cDNA libraries of three tissues and public databases), and (4) a putative gene set generated by clustering the predicted genes CDS and unigenes. The sequences were annotated with the result of (1) the BLAST search against non-redundant protein database (NCBI-nr), gene sets of model insects (*Bombyx mori*, *Tribolium castaneum*, etc.) and EST databases of lepidoptera insect pests (*Spodoptera frugiperda*, etc.), (2) the domain search against the Pfam database by HMMER, (3) the assignment of Gene Ontology (GO) term by BLAST2GO. The database provides Rich GUI based web interfaces for efficiently browsing, searching and downloading the generated sequences and their annotation data, by which genes of interest can be easily investigated by using a web browser. The feature of the database and the result of genomic analysis using the database (analysis of the feature of the transcriptome data, comparative analysis with lepidoptera insects, etc.) will be presented.
This work is supported by Grant-in-Aid for Scientific Research (B) (22380041).

EG-7

WebApollo: A Web-based Sequence Annotation Editor for Distributed Community Annotation

*Lee, Ed; Helt, Gregg; Harris, Nomi; Buels, Robert; Childers, Christopher; Reese, Justin; Muñoz-Torres, Mónica C; Elsik, Christine G; Holmes, Ian; Lewis, Suzanna E.
(EL, GH, NH, SEL) Berkeley Bioinformatics Open-source Projects, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA; (CC, JR, MCM, CGE) Georgetown University, Washington DC, 20057, USA; (RB, IH) Department of Bioengineering, University of California at Berkeley, Berkeley, CA, 94720, USA

As technical advances make sequencing faster and cheaper, genomic annotation efforts must adapt to keep pace. The upward trend in the number of genome sequencing projects means there will be a larger reliance on contributions from domain specialists. Thus the curation environment is shifting from a traditional centralized model, in which all curators for a given genome project share the same physical location, to a geographically dispersed community annotation model—which requires new tools to support community annotation efforts. WebApollo was designed to provide an easy to use, web-based environment that allows multiple distributed users to edit and share sequence annotations.

An extensive group of curators and investigators from the bee genome research community, spread over academic institutions worldwide, are currently beta-testing WebApollo. Their curation efforts, findings and interactions will dramatically upgrade the quality of the annotation data for the genomes of honeybee (Apis mellifera), and two bumble bees (Bombus impatiens and Bombus terrestris), which will lead to a better understanding of the biology of these social insects.

WebApollo is comprised of three components: a web-based client, a server-side annotation editing engine, and a server-side service that provides the client with data from different sources, including databases at the University of California at Santa Cruz, Ensembl, and Chado.

The web-based client is designed as an extension to JBrowse, a JavaScript-based genome browser that provides a fast, highly interactive interface for the visualization of genomic data. This JBrowse extension provides the gestures needed for editing annotations, such as dragging and dropping features to create new annotations of genes, transcripts and other genomic elements, dragging to change exon boundaries of existing annotations, and using context-specific menus to modify features. It has support for deep sequencing visualization (e.g., BAM data). The extension also connects to the annotation-editing service and the data-providing services.

The server-side annotation-editing engine is written in Java. It handles all the necessary logic for editing and deals with the complexities of modifications in a biological context, where a single change can have multiple cascading effects (e.g., when splitting or merging transcripts). Edits are stored persistently in the server, allowing users to quickly recover their data in the event of unexpected browser or server crashes. The server provides synchronized updates over multiple browser instances, so that every edit is immediately visible to all users who are viewing or editing the same region. It offers multiple levels of user accessibility, allowing project owners to decide with whom to share their work, and whether to allow read-only or both read and write access.

The server-side service that provides data to the client is built on top of Trellis, a Distributed Annotation System (DAS) server framework. It sends JBrowse-supported JavaScript Object Notation (JSON) data, rather than the more verbose DAS XML. We developed Trellis plugins to access data from the UCSC MySQL genome database, Ensembl DAS services, and Chado databases. All three components are open source and provided under the BSD License.

WebApollo will be publicly released in fall 2012.
Public demo: http://icebox.lbl.gov:8080/ApolloWebDemo
Project web page: http://gmod.org/wiki/WebApollo

EG-8

Developing a community annotation system for arthropod genomes

*Lawson, D.; Hughes, DST.; Megy, K.; VectorBase.
European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, CB10 1SD

VectorBase is a NIAID-funded Bioinformatic Resource Center that provides genomic resources for the vector community for both established vector species as well as newly sequenced ones. Changes to our genome annotation pipeline have highlighted the need to improve our ability to capture and present community annotations, with an emphasis on the intiial annotation effort but also as part of the ongoing curation for our species.

Development of this system has been driven by the desire to involve the research community for species undergoing genome annotation to ensure that important loci are appraised prior to submission to INSDC archival databases and publication. This pre-loading of community led assignment of gene names and descriptions should increase the utility of the genome to all researchers upon publication.

Essentially, users download the genomic region (in Fasta format), annotation and evidence tiers (in GFF3 format) from a web interface. They can then edit and refine gene models using existing tools such as Apollo or Artemis and submit modified annotation back to VectorBase.

A number of quality assurance/control procedures are ran on submitted data to identify potential errors, which are communicated via email back to the submitter. Annotations that pass these procedures are stored in a SQLlite database and can be integrated into the canonical gene set on an appropriate schedule.

This system has been used in house for the *Glossina morsitan*s and *Rhodnius prolixu*s annotation efforts and has been deployed for non-vector communities such as *Heliconius melpomene* and *Strigamia maritima*. We will present an overview of the system, ongoing developments to capture associated metadata and discuss the suitability of deploying this system to the wider arthropod species (especially those slated for sequencing as part of the i5k project).

EG-9

From The Genome to Population Dynamics: The Glanville fritillary butterfly as a model system

*Lehtonen R.; Somervuo P.; Salmela L.; Paulin L.; Ahola V.; Koskinen P.; Saastamoinen M.; Duplouy A.; de Jong M.: Taipale M.; Mäkelä N.; Rastas P.; Wong S.; Kvist J.; Lawson D.; Taipale J.; Mäkelä V.; Ukkonen E.; Frilander M.; Holm L.; Auvinen P.; Hanski I.
University of Helsinki, (LR, SP, SV, SM, DA, JM, RP, WS,HI) Metapopulation Research Group, Department of Biological and Environmental Sciences; (PL, KP, KJ, FM, HL, AP) Institute of Biotechnology; (SL, MN, MV, UE) Department of Computer Science , Helsinki, Finland; (TM, TJ) Karolinska Institutet, Department of Biosciences and Nutrition, Huddinge, Sweden; (LD) VectorBase/Ensembl Metazoa, European Bioinformatics Institute, Hinxton, UK.

Our main research interests are the effects of spatial population structure on natural selection and demographic and evolutionary dynamics using the Glanville fritillary butterfly (*Melitaea cinxia*) as a model system. Long-term research conducted on the Glanville fritillary metapopulation in the Åland islands has yielded examples of landscape structure maintaining nonsynonymous polymorphisms in candidate genes, genetic effects on population dynamics as well as putative examples of coupling between demographic and evolutionary dynamics.

The Glanville fritillary genome which have been sequenced using the newest Illumina, SOLiD and Roche 454 platforms, will be completed during this spring. Through resequencing of pooled population samples we have been able to identify 2 million SNP sites which information is the bases of our large-scale metapopulation-wide variation analysis including up to 30,000 samples. In parallel with the assembly of the genome, we are reconstructing a genetic map by next-generation RADtag sequencing and targeted genotyping of the parents and the offspring from carefully selected crosses between distant populations.

We have conducted pioneering transcriptome-wide analyses of pooled population samples which were used to identify loci under selection using cross-population SNP allele frequency comparison and Gene Ontology enrichment analysis.

We have designed genotyping assays for Illumina GoldenGate and Sequenom MassArray platforms and have genotyped up to 2,000 population samples with 800 markers so far. We have been able to find significant genetic associations to phenotypic and life-history traits, such as larval growth, flight metabolism and total reproductive output, using the carefully selected SNP panels. Additionally, we have already identified interesting phenotypic and genotypic differences between old vs. newly established populations, fragmented vs. continuous landscape and an isolated single population vs. a large metapopulation. The study in the metapopulations originate from two highly fragmented landscapes in Finland and Sweden, two extensive continuous landscapes in Estonia and Sweden, and one extremely isolated small population in Russia.

We are now extending the variation analysis to the whole transcriptome level by ultra-deep Illumina RNA sequencing in 5 wild butterfly populations (~40 samples/population) living in different kinds of landscapes. These experiments have become possible because of rapid evolution in high-through-put sequencing technologies and development of cost-effective in-house RNA library construction protocols. We are aiming to gain knowledge about local adaptation, micro-evolutionary effects and gene regulatory pathways using SNP allele, alternative splicing and gene expression variation data.

This project will yield rich molecular tools and material to study adaptation and evolution in contrasting environments at the genomic level as well as candidate genes for organismal studies.

EG-10

New World screw-worm Genomics: a First Step Towards Understanding the Evolution of Parasitism in Calliphoridae

Leotti, Fabricio C.; Cardoso, Gisele A.; Azeredo-Espin, Ana Maria, A. M. L.; *Torres, Tatiana, T.
(FCL, TTT) Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo (USP), São Paulo, Brazil; (GAC, AMLAE) Departamento de Genética, Evolução e Bioagentes, Instituto de Biologia, Universidade Estadual de Campinas (Unicamp), Campinas, Brazil

Calliphoridae flies are characterized by the ability of their larvae to develop in animal flesh. Species in the Calliphoridae family can be divided generally into three groups based on their larval feeding habits: saprophagy, facultative ectoparasitism, and obligate parasitism. It has been proposed that this functional division may reflect the progressive evolution of parasitism in the Calliphoridae family. The range of life-history strategies and the appearance of obligate parasitism independently in at least three separate occasions in Calliphoridae make it an ideal model system for the evolution of parasitism in a phylogenetic context.

Among the obligate parasites in the Calliphoridae family is the New World screw-worm (NWS), *Cochliomyia hominivorax*. This species is one the most important parasitic insect pests in South and Central America. Larvae of this fly infest tissues of warm-blooded vertebrates causing significant reductions in the quality of leather and in the production of milk and meat. This insect pest also represents a serious public health problem in the Caribbean region, where screw-worm infestations in humans are frequently reported.

Despite its medical and veterinary importance and its negative economic impact on the livestock sector, only limited genetic information is available for this species. Until very recently, the database of genomic sequences for insect species has been restricted to model species. Now, with the recent advances in DNA sequencing technology, the generation of sequence data has increased at an unprecedented rate. As new sequencing technologies became less expensive, it is possible to generate genomic information from non-model species such as *C. hominivorax*. Massively parallel sequencing has been demonstrated as an excellent tool to generate genome-wide sequence information as well as levels of gene expression. Hence, as a part of an effort to use comparative functional genomics to study the evolution of feeding behavior in Calliphoridae, we sampled NWS transcriptome by deep sequencing of poly-A transcripts and generated low-coverage genome survey sequences using the 454 sequencing technology.

Deep sequencing on the 454 platform of three normalized libraries (larval, adult male and adult female) generated a total of 548,940 reads that were assembled into 37,432 unigenes. About 44% of the NWS unigenes were annotated using a database of protein-coding transcripts from 17 insect species. We provided the largest database of NWS expressed sequences. Several candidate genes involved

in feedind behavior in larvae and adult flies were selected for the comparison among different Calliphoridae species, and at least two of them show an expression level related to the feeding habit.

Sequencing of NWS genome using shotgun and paired-end sequences, provided an estimated depth of coverage of 1x. With the low-coverage sequencing, we discovered highly repetitive elements as retroviruses, pseudogenes and transposable elements, conserved in closely related species, *Chrysomya bezziana* and *Lucilia cuprina*. We also assembled the complete sequence of the NWS mitochondrial genome with high accuracy. The sequencing and characterization of *C. hominivorax* genome is still underway and it will be an important resource for the future comparative studies in Calliphoridae.

EG-11

Annotating genomes and training the annotators: insights and lessons from Hymenoptera Genome Database.

*Munoz-Torres, Monica C.; Reese, Justin T.; Childers, Christopher P.; Bennett, Anna K.; Hagen, Darren E; Elsik, Christine G.
Georgetown University, Department of Biology, Washington, DC 20057 USA

Hymenoptera Genome Database (HGD, hymenopteragenome.org) is an informatics resource in support of genomic research in insect species of the order Hymenoptera; its main objective is to enhance our understanding of the biology of agriculturally, ecologically and evolutionarily relevant Hymenoptera species through genomics. HGD covers approximately 200 MY in the phylogeny of Hymenoptera, and includes an outgroup for social behavior, all of which allow the leverage of genetic data, genome sequences, gene expression data, as well as the biological knowledge of related model organisms. Our relational database implements open-source software and components that provide access to curated data continuously contributed by an extensive, active research community. The true wealth of knowledge behind each predicted gene model of each sequenced genome is mostly an unknown treasure, only uncovered for a few chosen models of particular interest; our only certainty is that finding a way to access all details behind each and every protein of each genome (both in terms of functional information and sequence characterization) is not only necessary to the expansion of our basic biological and evolutionary knowledge of a species, but also vital to bringing the genome to life. Our group has acquired extensive experience in developing and implementing tools to aid in the annotation and curation of insect genomes, both for established research models such as the honey bee (*Apis mellifera)* and the parasitoid jewel wasp (*Nasonia vitripennis*), as well as for emergent genomes such as those of the bumble bees (*Bombus terrestris* and *B. impatiens*). As the cost of completing genome sequencing projects drops, a larger number of insect researchers working on species of Hymenoptera have ventured into obtaining their own genome sequences. In the last year we have especially focused our attention in developing special tutorials for bumble bee researchers, a special community with special kinds of data. New Generation Sequencing (NGS) techniques have brought an onset of new challenges to annotators, ranging from managing the introduction of frameshift errors into an assembly built using pyrosequencing data (454), to the reoccurring annotation of gene models dispersed across two assembled scaffolds. HGD continues to capture, curate and distribute research contributions from a constantly growing community of nearly 80 institutions in 14 countries, constituting some of the largest dispersed manual annotation efforts reported to date. Here we detail our efforts in training dispersed communities of scientists and how they can be emulated and improved to become part of a model for dissemination and implementation of annotation and curation tools by other members of the genomics community, such as the research teams included in the i5K initiative.

EG-12

First Draft Genome Sequence of Coffee Berry Borer: the most invasive insect pest of coffee crops

Nuñez, Jonathan.; Hernández, Erick.; Giraldo, William.; Navarro, Lucio.; Gongora, Carmenza.; Cristancho Marco A.; Gaitán, Álvaro L.; *Benavides Pablo.
(LN, EMH, WG, LN, CG, MAC, ALG, PB) National Coffee Research Center - CENICAFE, Kilometro 4 via Antigua a Manizales, Manizales-Colombia; (LN) Purdue University, Department of Entomology, West Lafayette, IN 97907-2089

Coffee Berry Borer (CBB), *Hypothenemus hampei*, is a curculionidae beetle with a reproductive system that ensures high degree of inbreeding and fitness, facilitating a quick colonization of coffee crops worldwide. This insect has a great impact on coffee producer countries, which is aggravated since it has the ability to speed up its life cycle and reproduction while temperature rises. Novel strategies to control this insect-pest are now required to face climate change and it is feasible to do so based on insect molecular data. We have sequenced the genome of both males and females, using a combination of WGS and 8-20kb paired end libraries by mean of 454-FLX pyrosequencing. As a result, we obtained 55,350 contigs and 11,187 scaffolds (N50: 461 Kb) and an estimated CBB genome size of 170 – 180Mb. This information has allowed to identified 20,570 gene models and the mapping of 20,653 unigenes obtained from a previously assembled global transcriptome. The identification of ubiquitous ultra-conserved and single-copy core genes was conducted, estimating a genome completeness of 95%. Genes involved in CBB metabolism are related to enzyme families involved with arabinoxylans and other metabolism polysaccharides, which possibly had the same evolutive dynamic as mannanases, an enzyme recently described in CBB as a key compound in the quick worldwide dispersal and strong specialization of CBB over *Coffea* berries. No orthologous genes have yet been reported in insects such as *Tribolium castaneum* and neither in *Hypothenemus obscurus*, a close relative with different host plants, which we sequenced its global transcriptome. Transcriptional profiles are being identifying with the objective of quantifying their levels of gene expression during CBB attack to coffee berries. We are addressing studies related to insect-plant interaction using RNA-Seq and have been identified expression profiles associated with the response of the insect under entomopathogenic infection with *Beauveria bassiana*, obtaining differential expression of genes such as thaumatin, hypothetical antimicrobials and high Cytochromo C activity. We are performing a Focused Genome Profiling sequencing strategy from a BAC library with 9,226 clones (110kb mean insert size) in order to improve our assembly and genome analysis and therefore addressing studies related to CBB genetics. We Will discuss the implications of this research on studies of CBB diversity, genetics, behaviour and pest management in Colombia.

EG-13

Practical Challenges of Organizing Community Annotation Resources

*Reese, Justin T.; Childers, Chris P.; Muñoz-Torres, Mónica C.; Bennett, Anna K.; Elsik, Christine G.
Georgetown University, Department of Biology, Washington, DC 20057

Manual annotation greatly increases the ability to gain biological insight from newly sequenced genomes. While the time and expense of sequencing genomes has fallen rapidly, annotation remains expensive and time consuming. Community annotation is an effective strategy for reducing the time and expense involved in manually annotating newly sequenced genomes. Free, open-source software exists for most tasks involved in coordinating community annotation efforts, but many technical challenges remain. Over the past several years, our effort to develop community annotation resources for various genome projects at the Hymenoptera Genome Database has led to the constant evolution of our standard operating procedures. Every new project poses new challenges, especially as new sequencing technologies emerge. In hopes of providing guidance to investigators of new insect genome projects, we share some of our approaches to developing community annotation resources. We discuss practical issues involved in collecting data from sequencing centers and other bioinformatics groups, managing data within our research group, presenting data to the research community, and collecting feedback from the community.

EG-14

A Draft genome assembly for the Western corn rootworm, *Diabrotica virgifera virgifera*

Robertson, Hugh M.; *Coates, Brad S.; Walden, Kimberly KO.; French, B.W; Miller, Nicholas M.; Yandell, Mark; Tagu, Denis; Siegfried, Blair D., Sappington, Thomas W.
HMR, KKOW) University of Illinois at Champaigm/Urbana, Champaign, IL51801; (BSC, BWF, TWS) United States Department of Agriculture, Agricultural Research Service, Ames, IA 50011; (NMM, BDS) University of Nebraska, Lincoln, NE 68583; (DT) Institut National de la Recherche Agronomique (INRA), France; (MY) University of Utah, Salt Lake City, UT 84112.

The Western corn rootworm, *Diabrotica virgifera virgifera* (Coleoptera: Chyrsomelidae) is a major insect pest of cultivated corn and cellulosic biofuel crops in the United States, and has recently become invasive in corn producing regions of Europe. Extensive crop damage is caused by larval feeding on plant roots and adult consumption of plant reproductive tissue (mainly the silk of corn ears). Effective field control of *D. v. virgifera* has remained difficult due in part to evolution of resistance to multiple management tactics. A whole genome sequence project has been initiated, which is hoped will lead to development of novel control tactics. The initial draft genome has been assembled from > 183 million paired end Illumina HiSeq 2000 reads from 500-bp insert libraries constructed from DNA extracted from an inbred strain, and are estimated to comprise ~30-fold coverage of the 2.58 Gbp *D. v. virgifera* genome. Over 190,000 sequence scaffolds were generated using SOAP *de novo* v.1.05 and GapCloser v. 1.12. Mapping the position of known genes indicated that single exon sequences are present within contigs, but full genes are disjoined in the current draft through interruption by intervening intron sequences. Additional Illumina HiSeq 2000 data is being generated from 1.5- and 5.0-kbp libraries, which will be integrated into a combined assembly. This should increase overall scaffold lengths and result in greater contiguous coverage of full gene sequences. Although preliminary, these results indicate that assembly and analysis of the *D. v. virgifera* exome will eventually be possible, opening the door to development of novel tactics for effective management of this serious insect pest.

EG-15

Transcriptome Analysis of *Lymantria dispar* Ld652Y Cells Suggests Substantial Viral-Related Expression and Activity.

*Sparks, Michael E.; Gundersen-Rindal, Dawn E.
USDA-ARS Invasive Insects Biocontrol and Behavior Laboratory, Plant Sciences Institute, Beltsville Agricultural Research Center, Beltsville, MD 20705

The gypsy moth (Lymantria dispar)-derived IPLB-Ld652Y cell line is a popular *in vitro* system for studying virus-related phenomena in the Lepidoptera. A transcriptomic library of 14,368 putatively unique transcripts (PUTs) comprising 8,476,050 high-quality, informative bases was generated using both single-pass EST sequencing and 454-based pyrosequencing. The NCBI nonredundant protein database was leveraged to assess gene content of the IPLB-Ld652Y transcriptome, and the Swiss-Prot subset of UniProtKB provided for inference of more detailed functional annotation, where possible. A diverse range of virus-associated transcripts was identified in the dataset, which suggests substantial levels of viral expression and activity in Ld652Y cells. These sequence resources should provide a high level of utility to insect virologists for developing testable experimental hypotheses.

EG-16

The Genome of the Hessian Fly Reveals a Massive Expansion of Secreted Salivary Gland Proteins Necessary for Gall Formation and Larval Survival

Stuart, Jeffery J.*; Chen, Ming-Shun; Brown, Susan J. Richards S. ; representing the Hessian Fly Genome Consortium.
(JJS) Purdue University, Entomology, West Lafayette, IN 47907; (MSC) USDA-ARS, Hard Winter Wheat Genetics Research Unit, Manhattan, KS 66505; (SJB) Kansas State University, Biology, Manhattan, KS 66506; (SR) Baylor College of Medicine, Human Genome Sequencing Center, Houston, TX 77030

The Hessian fly is an obligate parasite and important pest of wheat a crop that provides more nourishment for people than any other food source. First instar larvae of this gall midge secrete salivary compounds that induce the wheat host to form a nutritive tissue feeding the larvae. Genome sequencing revealed approximately 1,500 Short Secreted Salivary Gland Protein Genes (SSSGPs) in multiple families with high conservation of cis-regulatory, UTR, intron and signal peptide sequences, but not other protein coding sequence. EST and proteomic data shows SSSGPs are transcribed in the salivary gland and proteins make it to the plant. Mapping of virulence traits identified 2 SSSGPs providing genetic evidence that SSSGPs are necessary for gall formation and larval survival. We postulate this massive family of genes records the evolution of this insect-plant co-evolutionary battle.

EG-17

Progress in genome sequencing of the brown planthopper, *Nilaparvata lugens*

*Suetsugu, Yoshitaka; Daimon, Takaaki; Jouraku, Akiya; Kobayashi, Tetsuya; Nakamura, Yuki; Kuwazaki, Seigo; Kayukawa, Takumi; Katayose, Yuichi; Kanamori, Hiroyuki; Kurita, Kayoko; Sanada-Morimura, Sachiyo; Matsumura, Masaya; Tanaka, Yoshiaki; Tanaka, Hiromitsu; Shiotsuki, Takahiro; Shinoda, Tetsuro; Yamamoto, Kimiko; Noda, Hiroaki
(YS, AJ, YN, SK, YK, HK, KK, KY) National Institute of Agrobiological Sciences, Agrogenomics Research Center, Tsukuba, Ibaraki 305-8634, Japan; (TD, TK, TK, YT, HT, TS, TS, HN) National Institute of Agrobiological Sciences, Division of Insect Sciences, Tsukuba, Ibaraki 305-8634 Japan; (MM, SS) NARO Agricultural Research Center for Kyushu Okinawa Region, Koshi, Kumamoto 861-1192, Japan

The brown planthopper (BPH), *Nilaparvata lugens (Stål)*, is one of the most serious destructive insects of rice in Asia. Controlling *N. lugens* has been depending on chemical insecticides. However, the long-term application of chemical insecticides results in development of resistance against them. Development of biotypes being able to infest resistant rice varieties has also been a serious problem to be solved. Accumulation of knowledge of *N. lugens* in genomic level will provide us clues to comprehensive understanding of development of insecticide resistance and new biotypes, which leads to sustainable pest management technologies. BPH ESTs derived from various tissues or developmental stages has been accumulated through Sanger techniques. Thanks to recent emergence of next-generation high-throughput DNA sequencing techniques, increase of BPH transcriptome data has been getting larger and larger. On the other hand, genomic information of BPH has been unavailable yet. Recently, our group has started genome sequencing of BPH, aiming at publication of draft genomic sequence. The availability of draft BPH genomic sequence will play an important role not only in the development of novel and sophisticated technology for pest control but also in other genomic studies. In this report, we introduce the progress of genome sequencing and relevant ongoing sequencing projects. The BPH genome sequencing is conducted in combination of pyrosequencing (Roche/454 Life Sciences GS FLX) and Illumina Sequencing technology. So far, total length of 454 reads reached about 7.2Gbp, corresponding to genome coverage of 6.2X. To obtain a global view of transcriptome in the whole BPH life cycle, high-throughput RNA-seq has been also performed using Illumina sequencing technology. With the construction of transcriptome data of a standard strain, identification of genes responsible for resistance against pesticides and for virulence to plant resistance will be accelerated. All the data will be available through the database specializing in BPH genomic data.

SESSION: COMPARATIVE GENOMICS

CG-18

Whole Genome Evolutionary Analysis of the Monarch Butterfly Genome

*Boore J.L.; Fuerstenberg S.I.
(JLB) Genome Project Solutions, Inc., 1024 Promenade St., Hercules, CA 94547 USA and University of California Berkeley, 3060 Valley Life Sciences Building, Berkeley, CA 94720 USA; (SIF) Genome Project Solutions, Inc., 1024 Promenade St., Hercules, CA 94547 USA

A description of the complete 273 Mb genome of the monarch butterfly, *Danaus plexippus*, was recently published (Shuai Z, Merlin C, Boore JL, Reppert SM, 2011. *Cell* 147:1171-1185). We have since employed our “PHRINGE” (Phylogenetic Resources for the Interpretation of Genomes) system to do evolutionary analysis of its compete gene set in comparison to those of the arthropods *Daphnia pulex*, *Ixodes scapularis*, *Pediculus humanus*, *Tribolium castaneum*, *Pogonomyrmex barbatus*, *Apis mellifera*, *Drosophila melanogaster*, *Anopheles gambiae*, and *Bombyx mori*, along with the outgroups of *Homo sapiens* and *Caenorhabditis elegans*. Briefly described, PHRINGE creates a graph with all inferred protein sequences as nodes, with edges formed from distance scores calculated from their full length alignments, then clusters these into gene families using a method that considers the evolutionary relationships among the organisms, then performs a true, thorough phylogenetic analysis for each. This allows assignment of orthologous (related by lineage splitting) and paralogous (related by gene duplication) relationships, creates reconstructions of gene duplications and losses, and provides the most accurate possible inference of gene function. This is in contrast with non-phylogenetic methods in common usage that rely on simple similarity matching and that are known to make errors, including the incorrect association of the pairs of more slowly evolving paralogs and lack of annotation for those more rapidly evolving. Every gene of each animal has a "Detail Page" with information such as its position, length, and intron-exon structure, and the most similar sequences, with links. Every cluster of homologous genes has a "Cluster and Compare Page" with the multiple sequence alignment, comparison of intron-exon structures, and the phylogeny of the member genes. The associated Synteny Viewer allows the user to see the relative physical locations of sets of orthologs in sets of user-specified genomes. This is presented at: http://danaus.genomeprojectsolutions-databases.com/.

CG-19

Honey Bee Queen Mandibular Pheromone Inhibits Ovary Development and Fecundity in a Fruit Fly

*Camiletti, Alison L.; Percival-Smith, Anthony; Thompson, Graham J.
Western University, Biology, London Ont., Canada N6A 3K7

A key feature of eusocial insects is their reproductive division of labour. The queen signals her fecundity to her potentially reproductive daughters via a pheromone, which in effect renders them sterile. In contrast, solitary insects lack division in reproductive labour and, as such, there is no need for ovary regulating pheromones. Nonetheless, females from both non-social and eusocial lineages regulate their ovaries to maximize their inclusive reproduction, and it is possible that the same underlying networks are involved. In this study we provide some evidence from fruit flies that solitary insects share a conserved ovary regulating pathway with eusocial honey bees. Specifically, we demonstrate that a honey bee pheromone suppresses fly female ovaries in much the same way as it suppresses worker ovaries within the beehive. Flies exposed to bee pheromone showed significant reduction in ovary size, egg number, and number of viable offspring as measured through ovary dissection and reproductive assays. Flies therefore respond to interspecific social cues that they are not behaviourally adapted to. This conspicuous observation suggests that flies and bees, though distantly related, retain a functionally conserved network that can regulate female reproduction.

CG-20

Hymenoptera Genome Database: Resources for Comparing Hymenopteran Genomes

*Childers, Christopher P.; Reese, Justin T.; Munoz-Torres, Monica C.; Bennett, Anna K., Elsik; Christine G.
Georgetown University, Biology, Washington, DC 20057

The Hymenoptera Genome Database (HGD; http://HymenopteraGenome.org) is a resource that supports genome informatics of hymenoptera species. HGD incorporates data from honey bee (*Apis mellifera*), two bumble bee species (*Bombus terrestris* and *B. impatiens*), the parasitoid wasp *Nasonia vitripennis*, and seven ant species (*Acromyrmex echinatior*, *Atta cephalotes*, *Camponotus floridanus*, *Harpegnathos saltator*, *Linepithema humile*, *Pogonomyrmex barbatus*, and *Solenopsis invicta*). The eleven species currently represented in HGD will soon be joined by two additional Apis species (*A. florea* and *A. dorsata*). The ongoing addition of new species improves the power of comparative analyses, advancing our understanding of biological processes relevant to agriculturally important hymenopterans and providing additional insights into social insect biology. Users may access the data via genome browsers (Gbrowse), web-based BLAST searches and direct file downloads. Datasets include genome assemblies, computed and manually-annotated genes, protein homologs, cDNA sequences, non-coding RNA sequences, RNA-Seq data and genetic markers. All way comparisons between the official gene sets (OGS) for each species at HGD, and the FlyBase gene set for *Drosophila melanogaster* were used to compute ortholog relationships between species. The HGD genome browsers leverage these pre-computed ortholog information for the official gene sets to link between the species, allowing users to navigate across the genomes at HGD. Database cross-references directly connect features to external resources. BLAST queries to a reference assembly are linked to the genome browsers, allowing visualization of BLAST hits in the context of the genome.

CG-21

Molecular clock reveals extremely rapid diversification of a communication system

K. H. Frederick-Hudson*, Corey Hudson, J. Schul
(KHFH, JS) University of Missouri, Division of Biological Sciences, Columbia MO, 65211; (CMH) University of Missouri, Informatics Institute, Columbia MO, 65211

Evolution of communication systems is inherently complex, requiring coevolution between sender and receiver mechanisms. This study uses comparative methods and molecular clock analyses to investigate the diversity and evolutionary history of call diversity in communication of *Neoconocephalus* katydids. Ancestral state reconstructions revealed three derived call traits each with multiple independent origins. Additionally, ancestral state reconstructions found that the temperate life history (with egg diapause) is derived. Seven of ten temperate *Neoconocephalus* species are found in a monophyletic clade. This temperate clade includes species with each of the four call phenotypes (ancestral & 3 derived). The genetic diversity found in this temperate clade is less than that among populations of its sibling clade. A molecular clock analysis dates the speciation within the temperate clade to less than 50,000 years ago. This supports the hypothesis that speciation and diversification of the temperate *Neoconocephalus* communication system has taken place since the last glaciation (100,000-15,000 years ago). Our results highlight how rapidly acoustic communication systems may diversify, which has important implications for our understanding of the underlying evolutionary mechanisms.

CG-22

Perfect Arthropod Genes Constructed from Gigabases of RNA

Gilbert, Don
Indiana University, Biology, Bloomington, IN 47405; gilbertd@indiana.edu

Gene construction, not prediction.
The decade of gene prediction is over; gene construction from transcriptome sequence now surpasses predictions for biological validity. To paraphrase broch
cent paper: ".. over half the computational gene predictions were imperfect, with missing exons, false exon predictions, wrong intron boundaries, fused and fragmented genes". Gene construction from RNA is not without similar and other problems, but these can be reliably resolved with attention to detail and a river of sequence data.

Is that a honey bee gene in your wasp genome?
Species expressed genes often differ from those mapped from other species**.** Transcript evidence can give truer answers for phylogeny, gene function, etc., than when one protein is fit to many species. Alternatives are everywhere and expression trumps protein mapping.

Too much data or not enough?
Complexity and conflicts: transcript assemblies can be more accurate than prediction guesses, but harder and more time consuming to resolve conflicts. RNA data quality sets limits, and software struggles at both ends of the data river. Sensible data reduction is a major gene construction task, where 10^9 RNA reads are assembled to 10^6 of competing transcripts, and those filtered with multiple criteria for the closest approach to 10^4.5 biological genes.

Genes without genomes?
Is a genome assembly needed? Is that transcriptome real? Assembling a transcriptome can be more costly than genome assembly, and less useful without one. Alternates, paralogs and bad guesses can be resolved with a genome. Measuring expression of imperfect genes has pitfalls.

Best practices for perfect genes.
Gene construction software and methods continue to improve, but are imperfect. A current best strategy uses several methods, extracting the best of their many results. Rough edges need smoothing: predictor models and transcript assemblies each have qualities the other lacks, for coding sequences and sequence signals, gene holes and mash-ups. Multiple lines of gene evidence can score the quality of competing gene constructions to select a best, if not yet perfect, gene set.

These lessons are illustrated with new genes constructed for *Daphnia*, pea aphid, a jewel of a wasp, and arthropods hiding in the Sequence Read Archive.

CG-23
Analysis of small RNA data in Hymenoptera reveal lineage-specific microRNAs
*Hagen, Darren E.; Elsik, Christine G.
(DEH, CGE) Department of Biology, Georgetown University, Washington DC 20057

MicroRNAs (miRNAs) are small RNA molecules, which are involved in post-transcriptional regulation of gene expression by targeting messenger RNA (mRNA) causing mRNA destabilization or repression of translation. Insect miRNAs help regulate expression of proteins involved in development, metabolism, and many other important molecular mechanisms and cellular processes. Identification and analysis of lineage and species-specific miRNA will advance the understanding of the evolution of biological traits. Utilizing Illumina sequence data generated by our group and data mined from publicly available databases, we characterize of a set of putative miRNA in a range of species from the Order Hymenoptera. Our analysis identifies a set of miRNA that appear to be unique to Hymenoptera. Additional analysis reveals miRNA unique to members of Aculeata and Formicidea. Ongoing miRNA target analysis will ultimately lead to a greater understanding of the role miRNA play in the biologically unique development and reproduction of Hymenoptera.

CG-24
Genome size diversity in Arthropods.
*Hanrahan, Shawn J.; Johnston, J. Spencer.
Texas A&M University, Entomology, College Station, TX 77843.

Arthropod genome size diversity remains poorly sampled, with far less than one percent of species surveyed. Herein are presented over 200 estimates that allow a first glimpse into the extent of genome size variation. Within these estimates are 8 new orders, and over 38 families that previously had no recorded genome size estimate. To date, the genome size for these species has ranged from 104 Mbp for the body louse all the way up to 18,000 Mbp for a grasshopper. The upper range has been exclusively found within the Orthoptera, with many species surpassing 10,000 Mbp. Most other orders contain a wide range of estimates. Of the species estimated, the Coleoptera range from 154 Mbp to 2,600 Mbp. The full range of genome size has yet to be discovered. We have expanded the known range by over 1,500 Mbp by sampling additional species. It is likely this range will change as additional sampling is performed.

Knowledge of the genome size of a species is paramount before sequencing. Genome size will directly correlate with the size and cost of any sequencing project. The data presented here allows one to answer some key questions: What species have currently genome size estimates? Can I predict the genome size of my species if I know the estimate of a relative? And, how strong is the phylogenetic signal found in genome size?

These data show very clearly that genome size changes are greatest at lower taxon levels. Yet, at the same time, even closely related species may differ very significantly in genome size. Knowledge of the genome size of a relative will often be a poor predictor of the genome size of a given species. We therefore recommend direct estimation of genome size be made for several individuals of the same strain and sex as those to be sequenced. Our lab routinely conducts arthropod genome size estimates, and is more than willing to determine genome size for any fresh or frozen arthropod that has been correctly identified to the species level.

CG-25
tRNA expression and post-transcriptional modifications in reduced-genome endosymbionts (*Buchnera aphidicola*) of divergent aphid species
*Hansen, Allison K.; Moran, Nancy A.
Dept. of Ecology and Evolutionary Biology, West Campus, Yale University, PO Box 27388, West Haven, CT 06516-7388

Most sap-feeding insects feed on a diet that is low in nitrogen, especially in the form of essential amino acids, and require obligate endosymbionts to synthesize their essential amino acids. *Buchnera aphidicola*, is an obligate unculturable bacterial endosymbiont of aphids. *Buchnera* has coevolved with its aphid hosts for 200-250 MY, during which its genome shrunk to only 416-652 kbp depending on the lineage. This genome reduction and accelerated sequence evolution has resulted in changes that appear to lower the efficiency and accuracy of transcription and translation as compared to free-living relatives.

Translational efficiency and accuracy is controlled by tRNAs and other genome-encoded mechanisms. In organelles, translational processes are dramatically altered, due to genome shrinkage and to the incorporation of host-encoded gene products used in translation. The influence of genome reduction on translation in endosymbionts is largely unknown. In this study, we investigate whether divergent lineages of *Buchnera aphidicola*, the reduced-genome bacterial endosymbiont of aphids, possess less optimal translation compared to their free-living relatives, such as *Escherichia coli*. Based comparative genomics and directional RNAseq data, we found substantial evidence that translation is less optimal in *Buchnera* than in *E. coli*. Broadly, we observed a specific, convergent, pattern of tRNA loss in *Buchnera* and other endosymbionts that have undergone genome shrinkage. Furthermore, many modified nucleoside pathways that are important for *E. coli* translation are lost in *Buchnera*. Additionally, *Buchnera*'s base compositional bias favouring A+T has resulted in reduced tRNA thermostability, and may have altered aminoacyl-tRNA synthetase recognition sites. *Buchnera* tRNA genes are significantly shorter than those of *E. coli*, as the majority no longer have a genome-encoded 3' CCA; however, all of the expressed, shortened tRNAs undergo 3' CCA maturation. Moreover, expression of tRNAs with particular anticodons was not correlated with usage of the corresponding codons. Overall, our data suggest that endosymbiont genome evolution results in less optimal translational efficiency compared to their free-living ancestor. Consequently, the translational efficiency and fidelity evident in *Buchnera* appear to be in an intermediate state between free-living bacteria and organelles. Analogous to organelles, if further tRNA erosion occurs in *Buchnera* importation of aphid host encoded tRNAs and translation related enzymes may be required to maintain *Buchnera*'s translational processes.

CG-26

Evolution at Two Levels in Fire Ants: The Relationship between Gene Expression and Protein Evolution

*Hunt, Brendan G.; Ometto, Lino; Keller, Laurent; Goodisman, Michael A. D.
(BGH, MADG) Georgia Institute of Technology, School of Biology, Atlanta, GA 30332; (LO) Fondazione Edmund Mach, Department of Biodiversity and Molecular Ecology, San Michele all'Adige, Italy; (LK) University of Lausanne, Department of Ecology and Evolution, Lausanne, Switzerland

Variation in protein sequence and gene expression each contribute to phenotypic diversity, and thus may be subject to similar selective pressures. Ants and other eusocial insects are particularly useful for investigating the evolutionary link between protein sequence and gene expression because expression variation plays a central role in determining differences between their sexes and castes. We investigated the relationship between protein sequence evolution and gene expression evolution in the fire ants *Solenopsis invicta*, *S. richteri*, and their hybrids in order to gain a greater understanding of how selection jointly operates on expression patterns and coding sequence. We found that genes with high expression variability within castes and sexes were frequently differentially expressed between castes and sexes, as well as between species and hybrids. These data indicate that genes showing high variation in expression in one context tend to show high variation in expression in other contexts as well. Our analyses further revealed that gene expression variability in *Solenopsis*, arising in multiple contexts, was positively associated with rate of protein sequence evolution. These results indicate that gene expression variability is negatively associated with the strength of selective constraint operating on a protein. Overall, our study provides one of the strongest demonstrations that selective constraint mediates both protein sequence evolution and gene expression variability across different biological contexts and timescales.

CG-27

Estimating Arthropod Genome Size: Pitfalls and Practical Advice.

*Johnston, J. Spencer; Hanrahan, Shawn J. Texas A&M University, Entomology, College Station, TX 77843.

An accurate estimate of genome size is important for anyone interested in sequencing an Arthropod species, and is an essential first step into the I5K Project. Sufficient papers have been written on genome size estimation that it should be possible for any lab with access to a flow cytometer to estimate genome size accurately. However, when one looks at the literature one finds numerous examples where genome size estimates vary widely between labs. We give a number of examples of that variation and list pitfalls that may lead to erroneous estimates. The first problem is that there are methods which appear attractive, but which consistently produce incorrect results. These approaches should not be used without strong support from other methodologies. The improper use of a DNA standard is a second, and very common, problem. This problem is easily overcome in most cases; the exception is for genomes in the range 1000 to 2000 mbp, where a suitable standard is still being sought. Endoreduplication is very common in arthropods. Failure to correctly identify the level of endorduplication in the recorded data leads to estimates that are in error by 0.5X, 2X, 4X, or more. Proper choice of tissue, and an awareness of endoreduplication is the only way around these errors. Finally endogenous DNAase and stain inhibitors can lead to genome size estimates that are 30% of the correct value. A simple test reveals these problems; proper choice of tissue seems the only way to avoid them. Our lab routinely conducts arthropod genome size estimates, and is more than willing to determine genome size for any fresh or frozen arthropod that has been correctly identified to the species level. Alternatively, we will be pleased to provide additional advice for those who decide to try those estimates on their own.

CG-28

Genomic Analysis of the Lateral Gene Transfer from a *Wolbachia* Endosymbiont to *Drosophila ananassae* through DNA Re-sequencing and Its Implications for Invertebrate Genomics

Kumar, Nikhil; Sieber, Karsten; Bromley, Robin; Flowers, Melissa; Ott, Sandra H.; Tallon, Luke J.; *Dunning Hotopp, Julie C.
Institute for Genome Sciences, University of Maryland School of Maryland, Baltimore, MD 21201

Lateral (or horizontal) gene transfer is the transfer of DNA between organisms in the absence of sex. It can allow organisms to acquire novel traits that are unique from those inherited. Often such transfers go overlooked in invertebrate animal genome projects. Yet, extensive LGT has now been described between endosymbionts and their invertebrate hosts. The most extensive transfers are between *Wolbachia* endosymbionts and its invertebrate hosts; we call such LGTs "nuwts" using the established nomenclature for numts. Nuwts have now been identified in the genomes of diverse invertebrate taxa. We show through genome re-sequencing with qPCR validation that the nuwt in *Drosophila ananassae* is the largest nuwt known with multiple integrations of the 1.5 Mbp *Wolbachia* genome in the insect chromosome. As many as eight copies of large portions of the *Wolbachia* genome exist per insect chromosome in some lines, with as many as sixteen copies per nucleus. This extensively duplicated transfer can be found in multiple lines of *Drosophila ananassae* from Asia and the Pacific indicating that it is widely distributed. An additional line was identified that has a different duplication pattern. We will present our analysis of this multi-genome re-sequencing project. This includes re-sequencing and qPCR data that shows that nuwts can become under-replicated in the genome under specific conditions. This suggests that this LGT may move between the euchromatic and heterochromatic portions of the genome, which is under further investigation. We will also present our data from across numerous invertebrate genome projects to highlight the challenges associated with examining nuwts in invertebrate genomes.

CG-29

Novel mechanisms of rhodopsin management in the mosquitoes *Aedes aegypti* and *Anopheles gambiae*

*Leming, Matthew; Metoxin, Alex; Moon, Young; Whaley, Michelle; O'Tousa, Joe
University of Notre Dame, Biology, Notre Dame, IN 46556

The dipteran compound eye is composed of ommatidial units, each containing an organized set of photoreceptor cells. A microvillar structure in these cells, known as the rhabdomere, is a large plasma membrane surface containing rhodopsin and other proteins responsible for phototransduction. The rhodopsins detect light photons and thereby initiate both the process of vision and the setting of the circadian clock. *Drosophila melanogaster* is the best-studied invertebrate model for visual functions. *Drosophila* shares a common

ancestor with the mosquito species that existed approximately 250 million years ago. Both *Drosophila* and mosquitoes possess multiple rhodopsins, and each of these is expressed in specific photoreceptors. The *Drosophila* rhodopsin found in the major class of photoreceptors is Rh1. By examining the pattern of rhodopsin expression in both *Aedes aegypti* and *Anopheles gambiae*, we show that both mosquito species also express long wavelength rhodopsins (Aaop1 and Agop1, respectively) in the major class of photoreceptors. The research has revealed two key differences in the management of the mosquito rhodopsins than found in the *Drosophila* model. In *Drosophila,* Rh1 remain localized to the rhabdomere and the steady state levels of Rh1 are constant through a 24 hour cycle. In contrast, Aaop1 opsin is localized to vesicles in the cytoplasm during daylight and no Aaop1 is found in the rhabdomere. The levels gradually increase during the morning period but then decline during the afternoon and evening periods. At dusk, Aaop1 moves from the cytoplasmic vesicles and is localized within the rhabdomeres. We examined the influence of circadian rhythms on Aaop1 behavior by exposing the animals to either sustained light, or a precocious dusk. These experiments show light is directly responsible for the movement of Aaop1 but that a circadian rhythm may influence the daily change in Aaop1 levels. Agop1 also exhibits a daily change in steady state levels, but the decline occurs during the morning hours and then Agop1 begins to accumulate later in the day. Similar to Aaop1, Agop1 also exhibits a light-triggered relocation from the rhabdomere to the cytoplasm. This relocation mechanism, however, does result in total loss of Agop1 from the rhabdomeric compartment. These data show these two mosquito species possess unique rhodopsin control mechanisms not anticipated from research on *Drosophila*. Furthermore, the results show that the visual systems of the two mosquitoes are further specialized, likely to support the different requirements for visual function in the nocturnal *An. gambiae* and the diurnal *Ae. aegypti* lifestyles.

CG-30

De novo transcriptome assembly and SNP discovery in _Euphydryas gillettii_ to test for balancing selection

McCoy, Rajiv C.*; Boggs, Carol L.; Petrov, Dmitri A.
Stanford University, Biology, Stanford, CA 94305

The dominant view today (based on both theoretical studies and genomic scans) is that balancing selection only marginally contributes to overall patterns of genomic variation. Motivated by recent theoretical work and using a novel test for balancing selection, our study refocuses attention on the relative importance of balancing selection versus other evolutionary forces in maintaining consequential genetic variation in natural populations. We investigate this question in *Euphydryas gillettii*, a univoltine montane checkerspot butterfly that was intentionally introduced to Gothic, Colorado in 1977. The population established and subsequently experienced wild fluctuations including extreme bottlenecks of fewer than 25 adult individuals. This demographic scenario is ideal to identify candidates for balancing selection, as balanced SNPs should be refractory to frequency change while neutral and deleterious SNPs will be disproportionately lost or fixed due to strong drift during a bottleneck.

We prepared and sequenced barcoded cDNA libraries from 8 whole larvae from the introduced Colorado population and 8 whole larvae from its ancestral Wyoming population on the Illumina HiSeq2000 platform. These data were used to assemble the *E. gillettii* transcriptome *de novo* using Trinity, resulting in 75180 contigs of mean length 732 bp, which were annotated using Blast2GO. We then mapped data back to this reference transcriptome using BWA, performing variant discovery and genotyping with GATK. We filtered out SNPs on contigs that represented potential contamination as well as putative false SNPs representing likely artifacts of our assembly, mapping, and variant detection pipeline. The remaining 16415 SNPs were annotated as synonymous, nonsynonymous, or untranslated using OrfPredictor and BlastX. Our pipeline identified 1675 nonsynonymous SNPs that were maintained through the bottleneck, representing candidates for future investigation of balancing selection.

CG-31

Genomic Trends in Aquatic Arthropods: Comparative Transcriptomics of the Calanoid Copepod _Eurytemora affinis_ and the Water Flea _Daphnia pulex_

*Munro, James B.; Orvis, Joshua; Posavi, Marijan; Borntrager, Martin; Lee, Carol E.; Silva, Joana C.
(JBM, JO, JCS) Department of Microbiology and Immunology and Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD 21201, USA; (MP, MB, CEL) Department of Zoology, University of Wisconsin Madison, Madison, WI 53706.

Copepods form the largest biomass of metazoans in the world's oceans, and typically dominate zooplankton assemblages in nearshore environments. They play a predominant role in global food webs, both as algal grazers and as food sources for some of the world's most important fisheries. Copepods are primary vectors of waterborne disease agents, including those that cause Typhoid fever, shigellosis, and numerous diarrheal diseases such as cholera, as well as of a wide range of non-pathogenic taxa. Despite their tremendous impact on global ecosystems, relatively little is known about the genome composition and structure of copepods, or of aquatic arthropods in general, when compared to the wealth of genomic information already gathered for terrestrial arthropods.

The copepod *Eurytemora affinis* has several features that make it an ideal model system for copepod genomics and as a waterborn disease vector. It is geographically and ecologically widespread and has an enormous biomass (10^4-10^5/m^3) in many coastal systems worldwide, including the Gulf of Mexico and the major bays and estuaries in North America and Northern Europe. This species hosts several human bacterial pathogens, most notably *Vibrio cholera*, as well as a variety of other bacterial taxa (see poster by Lee and collaborators). Most importantly, *E. affinis* has one of the smallest genomes among copepods, with 1C=294 Mb, and several genetically distinct inbred lines are actively maintained.

In order to initiate the genomic characterization of *E. affinis*, we are analyzing the transcriptome of this species. Our study is based on the *E. affinis* inbred line VA30, started with animals obtained from L'Verte, in the St. Lawrence estuary, Canada. We have generated the transcriptome of three samples: adult males and females treated a panel of antibiotics, TM-1 (treated males) and TF-1 (treated females), and a pool of adult males and females that were not antibiotic-treated, UMA-2 (untreated mixed adults). Strand-specific RNAseq data was obtained for the three samples, and assembled with the software Trinity. The assembly process resulted in 154K, 120K and 134K transcripts for TM-1, TF-1 and UMA-2, respectively. Preliminary ORF analyses suggest that the number of expressed genes in females is slightly lower than in males. Analyses of these transcripts will be conducted to cluster transcripts by orthology

group, confirm taxonomic origin, assign coding potential and function, and identify metabolic pathways. These data will provide the first insight into the core transcriptome of *E. affinis* from a saline environment, and give us the ability to investigate sex-based differential expression and the impact of the copepod microbiome on the host's transcription patterns.

With the intent to help define patterns characteristic of aquatic arthropods, we are analyzing the *E. affinis* data in the context of the data available for the crustacean *Daphnia pulex*. *D. pulex* is a model organism for biomedical research, and the only aquatic arthropod for which substantive genomic resources are available. Two striking features of the *D. pulex* genome are the large number of species-specific genes and the expansion of several gene families, in particular those that encode for genes relevant to the aquatic lifestyle. Interestingly, our preliminary analyses reveal that several hundred genes so far unique to *Daphnia* have significant homology to ORFs in *E. affinis*. Also, as was observed in other aquatic organisms including *Daphnia*, *E. affinis* transcripts reveal a large number of opsin and cryptochrome ORFs. Detailed comparative transcriptome analyses of *E. affinis* and *D. pulex* will highlight traits characteristic of aquatic arthropods and help us understand the evolution of aquatic specialization.

CG-32

Identification of hymenopteran non-protein-coding RNA candidates by comparative genomics

*Nunes, Francis M. F.; Paschoal, Alexandre R.; Cristino, Alexandre S.; Freitas, Flávia C. P.; Simões, Zilá L. P.; Durham, Alan M.
(FMFN, FCPF) Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto, SP 14049-900, Brazil; (ARP, AMD) Universidade de São Paulo, Instituto de Matemática e Estatística, SP 05508-090, Brazil; (ASC) The University of Queensland, The Queensland Brain Institute, Qld 4072, Australia; (ZLPS) Universidade de São Paulo, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, SP 14049-901, Brazil; (FMFN, ARP) Equally contributing authors.

A growing body of evidences indicates that eukaryotic genomes are pervasively transcribed into a wide diversity of RNA classes. While protein-coding sequences (mRNA) are encoded by a small part of a genome, a great variety of non-protein-coding sequences (ncRNA) covers a large genomic percentage. Different types of ncRNA are critically involved in a number of regulatory pathways controlling gene expression. Although some insect genomes were sequenced, only very little is known about the *'entomos'* ncRNA repertory. In particular, genomic data of two hymenopteran species (the honey bee, *Apis mellifera*, and the jewel wasp, *Nasonia vitripennis*) lay unexploited in terms of their non-coding layer. In the current study we developed an automated pipeline to predict and annotate conserved ncRNA elements from the aforementioned hymenopteran genomes using the fruit fly *Drosophila melanogaster* as reference species. Toward this end, we used the eQRNA algorithm that performs a comparative genomics analysis based on probabilistic inference of genomic alignments. To avoid false positive results, AT content from each genome and minimum identity percentage across them were calculated and used to adjust alignment parameters (BLAST). A total of 5,508 insect-conserved ncRNA were predicted in the honey bee genome, ranging from 83 to 7,834 nucleotides (nt) in length. In the wasp genome, we found 2,540 putative ncRNA which vary from 69 to 14,416 nt. Approximately 48% of *Apis* candidates and 44% of *Nasonia* candidates are located in genomic regions that are not yet organized according to the chromosomal mapping. On the other hand, chromosome 1 of both species presented the highest number of ncRNA -- 400 candidates in honey bee and 310 candidates in wasp. These results suggested that loci for both ncRNA and mRNA sequences are placed in microsynteny blocks shared among hymenopteran genomes. Our *in silico* analysis indicated that ~54% (in wasp) and ~79% (in honey bee) of ncRNA have similar structural properties with known ncRNA. Most candidates were annotated as miRNA or snoRNA. Non-coding candidates were also searched against available expression data such as Expressed Sequence Tags and RNA-Seq libraries, and 1,241 *Apis-Nasonia* ncRNA were validated. Our study integrated data from dry and wet experiments resulting in a novel and robust strategy that predicted high-confidence ncRNA from Hymenoptera genomes, opening new avenues for insect RNomics.
Financial support: FAPESP (2011/03171-5, 2010/08332-4), CNPq (AMD, 161917/2011-9, 481000/2009-7), CAPES (ARP)

CG-33

Improved Prediction of Ultraconserved Regions in Insect Genomes

Rhodes, Adelaide.
Harte Research Institute for Gulf of Mexico Studies, Texas A&M University-Corpus Christi 78412.

The increase in available genomes for arthropods allows the testing of general hypotheses on conservation in genome structure. In vertebrates, ultraconserved elements in noncoding regions of human, mouse and rat genomes are a useful tool for predicting potential gene regulators as well as to determine amniote ontogeny. Ultraconservative elements have been identified in insects as well. This poster will retest the conservation of previously identified ultraconservative elements among *Drosophila melanogaster*, *Drosophila pseudoobscura* and *Anopheles gambiae*. Previously, the two longest ultraconservative region identified by Glazov et al. (2005) among these three species were the intron-exon junction of the homothorax gene and the snRNA:U6 gene cluster. A more recent test of homology found that the ultraconserved region within the snRNA:U6 gene cluster was maintained across several additional *Drosophila* species, but not in *Anopheles mellifera* or the red flour beetle *Tribolium castaneum*. Ultraconservation was not retained across the exon-intron region of the homothorax gene among additional *Drosophila* species. The addition of insect, crustacean and other arthropod genomes to the existing sequences will allow for the discovery of additional ultraconserved elements as well as the verification of proposed ultraconserved elements in insect genomes.

CG-34

Identification of Potential Chemosensory Gene Families in Funnel Web Spiders using NGS

*Sánchez-Gracia, Alejandro; Frias-López, Cristina; Guirao-Rico, Sara; Almeida, Francisca C.; Librado, Pablo; Vieira, Filipe G.; Arnedo, Miquel A.; Rozas, Julio
(AS-G, CF-L, SG-R, FCA, PL, FGV, JR) Universitat de Barcelona and Institut de Recerca de la Biodiversitat (IRBIO), Departament de Genètica, Barcelona 08028, Spain; (CF-L, MAA) Universitat de Barcelona and Institut de Recerca de la Biodiversitat (IRBIO), Departament de Biología Animal, Barcelona 08028, Spain; (SG-R) University of Edinburgh, Institute of Evolutionary Biology, Edinbugh EH9 3JT, UK; (FGV) University of California, Berkeley, Department of Integrative Biology, Berkeley CA 94720-3140, USA

The study of the origin and evolution of arthropod chemosensory system offers a perfect framework to investigate the role of the evolutionary forces in determining the structure, organization and evolution of genomes. Using the four major chemosensory families of arthropods as a model, we have performed the first rigorous and comprehensive comparative analysis of the dynamics of gene gains and losses and the impact of natural selection in shaping variation at the coding sequences of new duplicates. For some of these analyses we have developed and applied new probabilistic models under the maximum likelihood framework; these methods have been implemented in the software *BadiRate*.

Our analysis of the genomic sequence of *Ixodes scapularis* (Chelicerata) points to the absence of the typical insect olfactory gene families in this species. In order to gain insights into the origin and evolution of the chemosensory system in chelicerates, we have obtained subtractive cDNA libraries from palps and the first two legs (the putative chemosensory organs) in the funnel web spider *Macrothele calpeiana* (Hexathelidae). Using the Roche 454 GS-FLX Titanium technology, we obtained over 50,000 reads for each library, with an average read length of 352 bp. De novo assemblies yielded 1300 and 1680 potential genes for legs and palps, respectively. Identification of gene families was carried out using reciprocal BLAST searches within each library. We used the Blast2GO software to conduct a prospective functional annotation of transcripts to identify candidate receptor and binding protein genes that could play a role in the chemosensory system of *M. calpeiana*. The results obtained from this analysis will allow us to better understand the molecular strategies developed by different groups of arthropods in its independent adaptation to the terrestrial environment.

CG-35

A Genome-Wide Survey of RNA Interference Genes in a Galling Insect

*Shreve, Jacob T.; Subramanyam, Shubha; Johnson, Alisha J.; Williams, Christie E.; Shukle, Richard H.
(SJT) Department of Entomology, Purdue University, West Lafayette, IN 47907; (SS) Department of Agronomy, Purdue University, West Lafayette, IN 47907; (JAJ) USDA-ARS Crop Production and Pest Control Research Unit, Department of Entomology, Purdue University, West Lafayette, IN 47907; (WCE) USDA-ARS Crop Production and Pest Control Research Unit, Department of Entomology, Purdue University, West Lafayette, IN 47907; (SRH) USDA-ARS Crop Production and Pest Control Research Unit, Department of Entomology, Purdue University, West Lafayette, IN 47907

The Hessian fly, *Mayetiola destructor*, is a gall midge in the dipteran family Cecidomyiidae and is becoming a tractable model species for galling insects. Many gall midges are significant pests in agro-ecosystems and among these is the Hessian fly, which is a pest of wheat and frequently causes significant yield losses in many wheat production areas worldwide. On seedling wheat, larvae feed within the leaf sheath near the crown and all damage to wheat is from larval feeding. Secreted salivary gland proteins (SSGPs) produced by larvae are hypothesized to be the effectors that elicit reprograming of susceptible plants and interact with *R* gene products to elicit the defense response of resistant plants. While the transcripts encoding putative SSGP effectors in Hessian fly have been extensively characterized, their functional role in eliciting the reprogramming and altered physiological pathways in susceptible plants is unknown. One approach to discover the function of a gene is perturbing its function by reverse genetic methods such as RNA interference (RNAi). *Caenorhabditis elegans* and *Tribolium castaneum* exhibit robust systemic RNAi responses; however, *Drosophila melanogaster* does not show a robust systemic RNAi response. This is proposed to be due to a lack of core component genes for systemic RNAi in the *Drosophila* genome. To date, our efforts to perturb the function of genes encoding SSGPs in Hessian fly larvae through direct injection, feeding, and soaking with dsRNA have met with little or no success. Hessian fly is a primitive dipteran (suborder Nematocera) in the infraorder Bibionomorpha that is considered to have a close relationship with the high dipteras in the suborder Brachycera that include *Drosophila*. We hypothesize the poor results to elicit a systemic RNAi response and RNAi mediated gene silencing in Hessian fly is due to a lack of some core component genes for a robust systemic RNAi response as is the case with *Drosophila*. To test our hypothesis we have utilized the recently available Hessian fly genome sequence to mine out homologs of genes involved in RNAi-mediated gene silencing and systemic RNAi response and compared these results with the core component genes in *Tribolium*, *Drosophila*, and *C. elegans*. These data support our hypothesis that there are differences in the number of RNAi core component genes compared to *Tribolium*, with Hessian fly being similar in its compliment of RNAi genes to *Drosophila*. Additionally, we inventoried core component genes for systemic RNAi in the genome of *Anopheles gambiae* and compared these results with those of Hessian fly. These data further documented differences between Hessian fly and *Anopheles* as well as between *Anopheles* and *Tribolium*, *Drosophila*, and *C. elegans*. Results are discussed from the perspective of the robust RNAi response observed in *Tribolium* with that obtained in Hessian fly and why some dipterans may be less amenable to systemic RNAi and RNAi mediated gene silencing.

CG-36

Comparative Analyses to Explore the Evolutionary Signal of Gene Partitions in the Insect Tree of Life

*Simon, S.; Narechania, A.; DeSalle, R.; Hadrys, H.
(SS, AN, RD) Sackler Institute for Comparative Genomics, American Museum of Natural History, New York, NY 10024; (HH) Institute for Ecology and Evolution, University of Veterinary Medicine Hannover, D-30559 Hannover, Germany

With the increase of available transcriptomes and whole genomes we have entered a new era – that of phylogenomics. In the phylogenomic era, great progress has been made in reconstructing the Tree of Life by inferring the phylogenetic relationships using genome-scale data. However, it has also been shown that gathering a large amount of data is not necessarily sufficient to produce reliable trees. Recent discussion in phylogenomic studies centers on the data quality and inference methods, both known to reinforce systematic bias resulting in a highly supported but incorrect tree. One important factor contributing to the data quality is how the evolutionary history of the genes, which compose the data sets, is of relevance for the resulting topology. Here, the phylogenetic signal of a gene is likely to be related to its evolutionary constraint, and it has been suggested that a polytomy can be resolved by using genes that evolve at the optimal rate in the relevant timescale. Consequently, the major question is which genes should be concatenated and could we identify genes with congruent phylogenetic signal present in all organism under investigation allowing the separation of true phylogenetic signals from random noise.

In our new approach we have started to analyze the impact of the selected genes on the resulting topology using the wealth of published sequence information (transcriptome- and genome-projects) for insects in addition to newly generated transriptomic data for

three polyneopteran orders (Dermaptera, Plecoptera and Zoraptera). In addition to total evidence analyses (all genes concatenated) we selected individual proteins according to their biological function with the assumption that they harbour the same evolutionary history along the branches of the organismal phylogeny. These comparative analyses show that within the Insect Tree a large number of genes (>300) are required to recover some deep nodes, e.g. Neoptera, Eumetabola or the sister-group relationship of Plecoptera and Dermaptera. In addition, based on the analyses using functional groups of genes, these nodes are in disagreement with the total evidence analyses. In contrast, only a few genes are required to recover holometabolan relationships and the functional categories exhibit strong agreement with the total evidence analysis. In sum, this study shows that within the Insect Tree phylogenetic ambiguity remains for several deep branches with high incongruences among independent data sets. In addition, the extent to which the biological function of the genes is related to its phylogenetic signal requires more detailed analyses. However, new genome and transcriptome projects encompassing all recognized insect orders might guide us to a deeper understanding of the evolutionary signal of different underlying gene partitions and ultimately of insect evolution.

CG-37

Single nucleotide polymorphic analysis reveals sub-structuring within California *Culex pipiens* complex populations and phenotypic associations

Trout Fryxell, Rebecca*; Lanzaro, Gregory, Cornel, Anthony
(RTF) Department of Entomology and Plant Pathology, University of Tennessee Knoxville, TN 37996-4560; (GL) Department of Population, Microbiology and Immunology, School of Veterinary Medicine, University of California- Davis, Davis, CA 95616; (AC) Department of Entomology, University of California – Davis, Davis, CA 95616

The genetic relationships differentiating the various members of the ubiquitous and phenotypically variable *Culex pipiens* complex of mosquitoes have long been a focus of controversy. In some settings local members of the complex are genetically isolated whereas in a different setting the complex can freely mate; a complexity that confounds aspects of their systematics, vector pathogen epidemiology and their control. Our objective was to identify patterns of gene flow and population structuring within above and below ground California *Cx. pipiens* populations using single nucleotide polymorphisms (SNPs) and to determine if specific phenotypes were associated with observed genotypes. Population structuring was analyzed using classical F-statistics, Bayesian statistics (Structure), and principle coordinates analysis (PCoA) of genetic distance among individuals.

Bayesian analyses indicated there was a significant amount of sub-structuring present, comprising two “groups", the northern *Cx. pipiens* and southern *Cx. quinquefasciatus,* which actively intermix creating a hybrid zone within which backcrosses frequently occur that create appearances of further sub structuring. The two “groups” were identified as *Cx. pipiens* and *Cx. quinquefasciatus* based on accepted population specific diagnostic assays. Within species analyses identified additional populations, at least three *Cx. pipiens* and four *Cx. quinquefasciatus* populations that were also present in the hybrid zone. According to F_{ST} tests, some phenotypic traits were better associated than others, but geography was the best indicator of population structuring.

This SNP analysis identified *Cx. pipiens* and *Cx. quinquefasciatus* with significant sub-structuring present within California and this may explain the high degree of phenotypic variation observed within the complex. Use of the underground phenotype did not concretely associate with a distinct *Cx. molestus* population, rather mosquitoes collected from underground sites is most likely to be a mixture of the complex members in California. The possibility that in addition to the genetic and spatial structuring of *Cx. pipiens* populations in California there may also be structuring in phenotypic behaviors such as autogeny and blood meal preferences that may be associated with some of these distinct populations**.** It is critical that future genetic studies use a genotype-phenotype approach to improve the vector community’s understanding of the complex behavioral differences among *Cx. pipiens s.l.* populations.

CG-38

Insights into intron evolution from the genome of the *Ixodes* tick.

*Waterhouse, Robert M.; Zdobnov, Evgeny M.; Tick Genome Consortium.
Department of Genetic Medicine and Development, University of Geneva Medical School, 1 rue Michel-Servet, 1211 Geneva, Switzerland; Swiss Institute of Bioinformatics, 1 rue Michel-Servet, 1211 Geneva, Switzerland.

The deer tick, *Ixodes scapularis*, is a vector of human and animal pathogens including bacteria that cause Lyme disease and anaplasmosis, as well as babesiosis-causing protazoan parasites. During their two-year life cycle from larval, to nymph, and then adult stages, they must take a blood meal between each stage from a host where they can feed for several days, and where pathogen transmission may occur. *I. scapularis* is amongst the first sequenced representatives of the chelicerates (horseshoe crabs, scorpions, spiders and mites), it improves the genomic sampling of non-insect arthropods and bridges the gap between insects and other metazoans. Taking advantage of the basal position of ticks in the arthropod phylogeny, we examined patterns of intron sharing amongst orthologous genes from species representing the major arthropod lineages and several outgroup species. Using single-copy metazoan orthologues from OrthoDB (www.orthodb.org) we mapped intron positions onto protein sequence alignments to identify shared and unique intron positions amongst the selected species. Strikingly, this quantification revealed a dramatic difference to the crustacean, *Daphnia pulex*, where the tick shares greater than ten times more intron positions exclusively with non-arthropods and about four times fewer intron positions exclusively with insects. The intron presence/absence data provides sufficient phylogenetic signal to successfully reconstruct the species tree and it consistently reveals that *I. scapularis* intron positions are more similar to vertebrates and the sea anemone, than to the pancrustaceans.

ResearcherID: http://www.researcherid.com/rid/A-1858-2010
Acknowledgements: Cate Hill & other members of the Tick Genome Consortium, the Computational Evolutionary Genomics Group (http://cegg.unige.ch) & the OrthoDB team (www.orthodb.org).

CG-39

OrthoDB: The Catalog of Arthropod Orthologs

Evgeny M. Zdobnov, Robert M. Waterhouse, and Evgenia V. Kriventseva
Department of Genetic Medicine and Development, University of Geneva Medical School, 1 rue Michel-Servet, 1211 Geneva, Switzerland; Swiss Institute of Bioinformatics, 1 rue Michel-Servet, 1211 Geneva, Switzerland

The OrthoDB resource presents a catalog of eukaryotic orthologs delineated in an explicitly hierarchical manner at each radiation of the species phylogeny. The current release spans 33 arthropod species, as well as 48 vertebrates, 73 fungi and 12 basal metazoan species. Dipterans currently represent the largest sampled insect order with 16 species, followed by hymenopterans with 10 species, with the *Ixodes* tick and the *Daphnia* water flea presenting clear outgroup references for the 31 insect species. The database features include searchable functional annotations from UniProt gene descriptors as well as Gene Ontology and InterPro attributes. These annotations, and the mapping of detailed *D. melanogaster* phenotype data, allow putative functional inferences for the vast majority of arthropod genes. Additional annotations for the metazoan orthologous groups describe links to Human diseases from Online Mendelian Inheritance in Man (OMIM), high-level mammalian phenotype terms from Mouse Genome Informatics (MGI), and phenotypic data from WormBase. The groups of orthologues are also characterised by evolutionary features including the quantification of protein sequence divergence amongst group members and homology to other orthologous groups. The interface features extended phyletic profile querying, enhanced text-based searches, query combination options, and a BLAST sequence search. The ever increasing sampling of sequenced genomes brings a clearer account of the majority of gene genealogies that will facilitate informed hypotheses of gene function in newly sequenced genomes, as well as providing essential data for uncovering and quantifying long-term trends of gene and genome evolution. OrthoDB has so far attracted over 50 citations, many from studies on arthropods, a testament to it utility and popularity in genomics research. OrthoDB is referenced with link-outs from a number of major international resources including FlyBase, NCBI and UniProt, and is freely accessible from www.orthodb.org.

References: Kriventseva *et al*, NAR, 2008; Waterhouse *et al*, NAR, 2011.

SESSION: GENOME REGULATION AND EDITING

GRE-40

Regulation of transposable elements in the *Aedes aegypti* mosquito genome by piRNAs

Arensburger, P. *; R. H. Hice, J. A. Wright, N. L. Craig, P. W. Atkinson
(PA, RHH, JAW, PWA) University of California Riverside, Center for Disease Vector Research, Riverside, CA 92507;
(NC) Department of Molecular Biology & Genetics and Howard Hughes Medical Institute, Johns Hopkins School of Medicine, Baltimore, MD 20742,USA.;

Uncontrolled transposable element (TE) movement is highly detrimental to genome integrity. To protect themselves host genomes have evolved pathways to limit the movement of endogenous TE sequences by inactivation of transcription via methylation of TE specific regions, chromatin effects, and others mechanisms. Recently, the central role of PIWI interacting small RNAs (piRNAs) in many of these TE inactivating pathways has been elucidated in *D. melanogaster* (see Siomi et al. 2011 for a recent review) . These studies show that in *D. melanogaster* nearly half of all identified piRNA sequences have sequences matching known TEs. When associated with PIWI proteins these can form piRNA-induced silencing complexes (piRISC) that act to silence specific TE transcripts. We have extended this analysis to the *Aedes aegypti* mosquito genome.

Ae. aegypti has a larger genome size than *D. melanogaster* (1,380 Mb. vs. 120 Mb.) and a much greater TE sequence load (47% of sequenced genome in *Ae. aegypti* vs. 3% of euchromatic genome in *D. melanogaster*). We found that while the mosquito genome contains roughly an equal diversity piRNA sequences as the fruitfly only 19% of *Ae. aegypti* piRNAs matched known TE sequences (compared to 51% for *D. melanogaster*). Furthermore, while TE matching piRNAs in *D. melanogaster* predominantly originate from a few “master” loci this model does not match our observations in *Ae. aegypti*. We discuss the possible origins of TE matching piRNAs in *Ae. aegypti* and the possible targeting of genes of viral origin.

Reference: Siomi, M. C, K. Sato, D. Pezic, and A. A Aravin. “PIWI-interacting Small RNAs: The Vanguard of Genome Defence.” *Nature Reviews Molecular Cell Biology* 12, no. 4 (2011): 246–258.

GRE-41

Characterization and functional analysis of the *Knk* gene family in the red flour beetle, *Tribolium castaneum*

*Chaudhari, Sujata S; Arakane, Yasuyuki; Specht, Charles A; Moussian, Bernard; Kramer, Karl J; Beeman, Richard W and Muthukrishnan, Subbaratnam
(SSC, KJK, SM) Department of Biochemistry, Kansas State University, Manhattan, KS, USA; (YA) Division of Plant Biotechnology, Chonnam National University, Korea; (CAS) University of Massachusetts, Worcester, MA, USA; (BM) Univesity of Tubingen, Tübingen, Germany; (KJK, RWB) USDA-ARS-CGAHR, Manhattan, KS 66502, USA

The insect cuticle is mainly composed of chitin and proteins, and forms a protective barrier that shields insects against biological and mechanical stresses. Proteins associated with chitin metabolism and/or cuticle assembly are attractive targets for biopesticide development. Recent studies of *Drosophila melanogaster* (*Dm*) have identified a gene, *knickkopf* (*knk* = broken head), whose expression is important for tracheal tube expansion and cuticular organization (Moussian et al., 2006). Using a bioinformatics search of insect genomes, we have identified three *Knk*-like genes in *Tribolium castaneum, TcKnk1, TcKnk2* and *TcKnk3*, for which orthologs are present in the *D. melanogaster* genome (*DmKnk1, DmKnk2* and *DmKnk3*, respectively). All three genes are differentially expressed during different developmental stages of the beetle, suggestive of distinct roles for each TcKnk in *Tribolium* development. All three genes are expressed in the carcass (whole body without gut) but not in gut tissue. RNA interference (RNAi) of *TcKnk1* results in lethal phenotypes at larval-larval, larval-pupal and pupal-adult molts, whereas RNAi of *TcKnk2* and *TcKnk3* leads to developmental arrest only at the pupal-adult molt with ~55% and 100% mortality, respectively. Interestingly, a lethal phenotype is observed only with dsRNAs specific for the C-terminal region of *TcKnk3* but not its N-terminal region, suggestive of the presence of alternatively spliced variants of *TcKnk3*. Functional analysis has revealed an important role for Tcknk1 in the organization and protection of procuticular chitin from chitinases (Chaudhari et al., 2011), whereas the other two genes appear to influence the laminar organization of chitin. Collectively, our

results suggest that Knks play crucial and distinctive roles in the maintenance and organization of chitin in the cuticles of the red flour beetle and other species of insects.

This work was supported by a grant from the national Science Foundation (IOS-1022227)

References

Moussian B, *et al.* (2006) *Drosophila* Knickkopf and Retroactive are needed for epithelial tube growth and cuticle differentiation through their specific requirement for chitin filament organization. *Development* 133:163-171.
Chaudhari, S. S., *et al.* (2011) The Knickkopf protein protects and organizes chitin in the newly synthesized insect exoskeleton. Proc Natl Acad Sci USA. 108:17028-17033.

GRE-42

Annotation of serine protease superfamily and transcriptional response to a plant protease inhibitor in *Helicoverpa armigera*

Kuwar, *Suyog S.; Pauchet, Yannick; Vogel, Heiko; Heckel, David G.
Department of Entomology, Max Planck Institute for Chemical Ecology, Jena, Germany D-07745

Protease inhibitors (PI) are major direct defenses evolved by plants to deter herbivorous insects. Insects exposed to PI grow slowly, show retarded development and increase in mortality, due to the inhibition of digestive proteases. However, during the course of evolution insects have overcome the effect of protease inhibitors in their diet, by increasing production of their proteases and/or evolving PI-insensitive proteases. We are interested in understanding the regulation of PI-sensitive and PI-insensitive proteases of *Helicoverpa armigera,* in response to SKTI, a Kunitz-type trypsin inhibitor from soybean. First we annotated a total of 113 serine protease genes and determined their intron/exon structure. Five major groups are inferred by phylogenetic analysis; one group of trypsin is azurocidine-like while two groups of chymotrypsins have different pro-peptide length. Gene sequences are highly conserved at their catalytic residues H_{57} D_{102}, S_{195} and N-terminal signature sequence. Transcriptional response in midgut tissue of 4^{th} instar *H. armigera* to SKTI in artificial diet shows up regulation of several trypsins at first, but their expression decrease gradually, while expression of up regulated chymotrypsins and their number increases over time. Larval performance also shows initial slow growth on SKTI diet but timely adaptation and attaining same weight as of control diet. Differential expression of trypsins and chymotrypsins genes over time, suggest candidate sensitive and insensitive proteases. These candidate genes will be further validated using real time qPCR and proteomics experiments.

GRE-43

Patterns of Conversation of Exceptional Recombination Rates in Honey Bees

Rueppell, Olav
Department of Biology, Univ. North Carolina at Greensboro

The exceptional recombination rates of social insects provide a strong argument for adaptive evolution of recombination. With about 20cM/Mb, honey bees exhibit recombination rates that are 5x the expected rates based on genome size and chromosome number. However, the ultimate and proximate causes for this phenomenon are unclear. We are also only beginning to understand the fine scale patterns of recombination and its evolutionary conservation in honey bees. Here, I will explore how variable local recombination rates are in honey bees. The results will be discussed in the light of ultimate and proximate hypotheses to explain the exceptional recombination rates in honey bees.

GRE-44

Full-length cDNA Database of a Lepidopteran Model Insect, *Bombyx mori*, Based on over 11,000 cDNA Clones

Suetsugu, Y.: Futahashi, R.; Kanamori, H.; Kadono-Okuda, K.; Sasanuma, S.; Narukawa, J.; Ajimura, M.; Jouraku, A.; Namiki, N.; Shimomura, M.; Sezutsu, H.; Osanai-Futahashi, M.; Suzuki, M-G.; Daimon, T.; Shinoda, T.; Taniai, K.; Asaoka, K.; Niwa, R.; Kawaoka, S.; Katsuma6, T. Tamura1, H. Noda1, M. Kasahara7, S. Sugano8, Y. Suzuki8, H. Fujiwara4, H. Kataoka, H.; Arunkumar, KP.; Nagaraju, J.; Goldsmith, MR.; Feng, Q.; Xia, Q.; Yamamoto, K.; Shimada, T.; *K. Mita
(YS; HK, KK-O, SS, JN, MA, AJ, HS, MO-F, TD, TS, KT, KA, TT, HN, KY, KM) National Institute of Agrobiological Sciences, Tsukuba, Japan; (RF) National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan; (NN, MS) Mitsubishi Space Software Co., Ltd., Tokyo, Japan; (M-GS) The Univ. of Tokyo, Graduate School of Frontier Sciences, Kashiwa, Japan; (RN) University of Tsukuba, Life and Environmental Sciences, Tsukuba 305-8572, Japan; (SK, SK, TS) The University of Tokyo, Graduate School of Agricultural and Life Sciences, Tokyo, Japan; (MK) The Univ. of Tokyo, Computational Biology, Graduate School of Frontier Sciences, Kashiwa 277-0882, Japan; (SS, YS) The Univ. of Tokyo, Human Genome Center, Institute of Medical Science, Tokyo 108-8639, Japan; (KPA, JN) Centre for DNA Fingerprinting and Diagonostics, Hyderabad, India; (MRG) The University of Rhode Island, Kingstone, USA; (QF) South China Normal University, Guangdong Provincial Key Laboratory of Biotechnology for Plant Development, School of Life Sciences, Guangzhou 510631, China; (QX) Southwest University, State Key Laboratory of Silkworm Genome Biology, Chongqing 400716, China

A new *Bombyx* genome assembly was well established in 2008, which ushered in functional genomics of Lepidoptera as well as silkworm. The goal of the current genome analysis is a complete annotation of the genome, for which we have prepared new full-length cDNA sequences and a web-based curation system. In this symposium, we will present full-length cDNA information. The average length of these FL-cDNAs was 1,813bp, and 10,409 of 10,666 FL-cDNAs that could be aligned on the silkworm genome assembly had open reading frames. The average intron size was 1,904 bp, possibly resulting from a high accumulation of transposons. We determined the frequency of identical ESTs in a deep EST database (>408,000 ESTs) to estimate their levels of transcription. The lengths of 3' UTRs correlated inversely with transcription levels, suggesting that long 3' UTRs may contain regulatory sequences depressing transcription. Using gene models predicted by GLEAN and published mRNAs, we identified 16,823 gene loci on the silkworm genome assembly, of which 27% were specific to silkworm, whereas 53% were common with six other model insects. Classification with GO-terms showed an enrichment of DNA-related and response-to-stress genes in silkworm, potentially reflecting its distinctive chromosome structure and many species-specific antibacterial genes, whereas genes common to all six insects were abundant in other protein-related classes. Among 1,367 tissue-specific genes, analysis of testis-/ovary-specific genes revealed distinctive features of sexual dimorphism including depletion of female-specific genes on the Z chromosome in contrast to an

enrichment of testis-specific genes. More than 40% of genes expressed in specific tissues mapped in tissue-specific chromosomal clusters, potentially corresponding to active chromatin domains enabling efficient gene expression. The newly obtained cDNA sequences enabled us to annotate the genome of this lepidopteran model insect more accurately, enhancing genomic and functional studies of Lepidoptera and comparative analyses with other insect orders, and yielding new insights into gene evolution and lepidopteran-specific genes.

GRE

Genome-Wide Approaches for Identifying Genes Controlling Sexual and Asexual Reproduction in the Pea Aphid

TAGU, D.*; LEGEAI, F.; JAQUIERY, J.; MIEUZET, L.; MAHEO, F. ; LETERME, N. ; BONHOMME, J. ; NOUHAUD, P.; RISPE, C.; LAROSE, C. GAGGIOTTI, O.; STOECKEL, S.; SIMON, J-C.
(TD, LF, JJ, ML, MF, LN, JB, NP, RC, LC, SS, SJC) INRA, UMR1349 IGEPP, F-35653 Le Rheu, France; (OG) LECA UMR CNRS 5553, Université Joseph Fourier BP 53 38041 Grenoble, France

One key issue for the success of pest management is the understanding of mechanisms involved in pest adaptations to environmental pressures. Aphids are among the main insect pests in countries of temperate and continental climates. They feed from phloem sap and provoke damage on plants by hijacking nutriments and wounding stems and leaf veins. By feeding from plant to plant, aphids also transmit numerous plant viruses circulating in the phloem and responsible for drop downs in crop yields. The success of aphids as pests is related to their peculiar life history traits, in particular their reproductive mode alternating asexual parthenogenesis and sexual reproduction. An additional complexity is the presence of lines or populations that have become entirely asexual. This work aims at integrating population genomics, quantitative genetics and transcriptomics to get comprehensive insights into these variations of reproductive mode in aphids, by linking phenotypic plasticity (molecular bases of clonal and sexual phases within a given genotype) and polymorphism (co-existence of sexual and asexual lines within a given species). This program combines three different genome-wide approaches.

We first identified loci of the pea aphid genome involved in differences in reproductive mode by genome scanning of multiple sexual and asexual populations. Since the variation of reproductive mode is shaped by climate factors, we sampled sexual populations in regions that have a cold winter (3 sites in Eastern France) and asexual populations in regions that have a mild winter (3 sites in Western and Southern France). To identify the genomic regions linked to reproductive phenotypes, we genotyped 124 individuals (72 asexuals and 52 sexuals) at 378 microsatellite markers chosen to cover different scaffolds of the referenced annotated genome. We detected loci under divergent selection between asexual and sexual phenotypes in populations. We found 5 genomic regions under divergent selection. The most divergent genomic region supported two genes (Farnesyl Diphosphate Synthase 2, HMG Co-reductace) of the juvenile hormone (JH) biosynthesis pathway known to be involved in the seasonal plastic switch of reproductive mode in aphids. We genotyped another microsatellite located at 30kb from this region which was again detected as an outlier, strengthening the hypothesis of its involvement in the control of reproductive mode variation.

The second step was to identify quantitative trait loci (QTLs) for the reproductive mode in the pea aphid. We i) generated 6 F1 and 253 F2 individuals from F0 that present contrasted phenotype for the reproductive mode, ii) genotyped these individuals at 443 microsatellite markers, iii) assessed the phenotype (i.e. reproductive mode), and iv) constructed a genetic linkage map to statistically identify QTLs. We pointed out 1 microsatellite marker associated with the reproductive phenotype: all individuals bearing a specific combination of two alleles at this locus express the asexual phenotype, while all other combinations of alleles resulted in sexual phenotypes. Interestingly, this locus corresponds to the locus identified from the genome scan approach. This suggests that some of the genes located at this locus and annotated as related to the JH pathway might be involved in the determinism of the reproductive mode.

To complete this genome wide analysis, we performed RNA-Seq transcriptomes of 10 sexual and 10 asexual populations. Statistical analyses to identify differentially transcribed mRNAs are in progress with a special focus on GO terms associated with the JH pathway. In conclusion, we found at least 2 independent pieces of evidence that the polymorphism of reproductive mode is related to loci involved in JH biosynthetic pathways.

GRE-45

Computational Identification of Potential A-to-I RNA Editing Sites in *Apis mellifera*

*Tao, Shu; Elsik, Christine G.
Georgetown University, Biology, Washington, DC 20057

RNA editing is a post-transcriptional process that leads to addition, substitution and deletion of certain bases of both protein-coding and non-protein-coding RNA molecules, and accordingly alters the coding capacity of mRNAs and biological property of non-protein-coding RNAs. While alternative splicing is a well-known mechanism that increases the number of different gene products compared to the number of genes encoded in a eukaryotic genome, A-to-I RNA editing may also play an important role in proteome complexity.

The computational identification of A-to-I RNA editing is a relatively new area of research that has become feasible with the availability of next generation sequencing technologies. Here we present the results of computational identification of potential A-to-I RNA editing sites in honey bee by virtue of its latest genome assembly (Amel_4.5) and Illumina-generated RNA-Seq data from ten honey bee individuals (five nurses and five foragers, 2-3 replicates per individual). We will use this data to investigate the effect of A-to-I RNA editing on protein structural components (e.g. motifs, domains, signal peptides). Our efforts to understand the potential role of A-to-I RNA editing in the honey bee brain will increase our understanding of potential discrepancies between the *Apis mellifera* genome and transcript sequences, as well as their potential effects on the functions of certain genes.

GRE-46

Identification of Conserved and Novel MicroRNAs in *Manduca sexta* and Their Possible Roles in the Regulation of Immunity-related Gene Expression

*Zhang, Xiufeng; Zheng, Yun; Jagadeeswaran, Guru; Sunkar, Ramanjulu; Jiang, Haobo
(XZ, HJ) Oklahoma State University, Entomology and Plant Pathology, Stillwater, OK 74078; (YZ) Fudan University, Institute of Development Biology and Molecular Medicine, Shanghai 200433, China; (GJ, RS) Oklahoma State University, Biochemistry and Molecular Biology, Stillwater, OK 74078

The tobacco hornworm *Manduca sexta* has served as a model for insect biochemical and physiological research. However, the knowledge on posttranscriptional regulation of gene expression by microRNAs is still limited in this species. Our previous study identified 163 conserved and 13 novel microRNAs in *M. sexta*, most of which are present at low levels in pupae. To identify more *M. sexta* microRNAs and examine their possible roles in the expression regulation of immunity-related genes, we constructed four small RNA libraries using fat body and hemocytes from fifth instar naïve or bacteria-injected larvae and obtained 32.9 million reads of 18-31 nucleotides by Illumina sequencing. In the dataset, there were two new conserved microRNAs and nine conserved variants. We also found seven novel microRNAs and one cluster of those. mse-miR-1b, predicted as the anti-sense microRNA of mse-miR-1, and putative mse-miR-929 from the previous study, were confirmed in the tissue small RNA samples. mse-miR-281-star, mse-miR-965-star, mse-miR-31-star, mse-miR-9a-star, mse-miR-9b-star, and mse-miR-2a-star were maintained at a high level compared to respective mature strand. Abundance changes of microRNAs were observed before and after the immune challenge. Based on the *M. sexta* CUFF transcripts and fat body-hemocyte transcriptome data, we suggest that certain microRNAs and microRNA*s regulate gene expression of some pattern recognition receptors, prophenoloxidase activation cascade components, serpins, and members of the putative intracellular signaling pathways (Toll, IMD, JAK-STAT, and MAPK-JNK-p38). In summary, this work enriches our knowledge on *M. sexta* microRNAs and how some of them may regulate immunity-related gene expression.

SESSION: METAGENOMICS

M

Genome-wide diversity of the European tick *Ixodes ricinus* and its associated microbiome.

*Carpi, Giovanna; Miller, Webb; Ratan, Aakrosh; Sarofeen, Christian M.; Harris, Robert S.; Drautz, Daniela I.; Rizzoli, Annapaola; Schuster, Stephan C.
(GC, AR) Fondazione Edmund Mach, Department of Biodiversity and Molecular Ecology, 38010 San Michele all'Adige (TN), Italy; (WM, AR, CMS, RSH, SCS) Pennsylvania State University, Center for Comparative Genomics and Bioinformatics, University Park, PA 16802, USA; (SCS, DID) Nanyang Technical University, Singapore Centre on Environ. Life Sciences Engineering, Singapore, 637551.

Ixodes ricinus is recognized as the most common hard-tick species and the primary vector of Lyme disease and of other newly emerging zoonotic pathogens in Europe. Despite *I. ricinus* being of increasing public health relevance, the extent of intraspecies genetic variability and the tick-associated microbiome (pathogen-load and commensal microbes) in natural field populations remain largely unexplored and limited to gene-by-gene approach by using low-throughput molecular techniques.

In this study, we employed a next-generation sequencing strategy (whole-genome sequencing) to obtain, as a first report, the complete mitochondrial genome and a draft assembly of the nuclear genome of *I. ricinus*. Field-collected *I. ricinus* from a tick-borne disease hot-spot in Northern Italy (local scale) were individually analyzed for this purpose. Since genomic resources are only available from distantly related species, a custom pipeline was developed to assemble the sequence data and to identify putative nucleotide polymorphisms (SNPs). We have obtained the complete mitochondrial genomes that resulted in 14,570 bases in length, for seven *I. ricinus* ticks. Preliminary results indicate a mean pairwise diversity of 0.0047 and, 106 informative polymorphic sites in the mitochondrial genome at a local scale. These findings suggest a higher degree of genetic diversity of *I. ricinus* at the mitochondrial level compared to previous studies evaluating a few mitochondrial genes on a larger geographic scale.

In addition, we have detected thousands of genome-wide SNPs among the samples at the nuclear level (draft assembly of the nuclear genome of 400.6 Mbp in length) that will facilitate future genotyping studies aimed to understand *I. ricinus* population structure at a higher resolution.

By using the same study design and sequencing approach, we have analyzed the microbiome harbored by each tick individual. The metagenomic analysis was performed with reference- and protein homology-based methods to taxonomically profile non-host genetic DNA-derived sequences and to assess the relative abundance of bacterial taxa. The broad bacterial diversity detected confirms pathogenic bacteria of known presence (*Borrelia burgdorferi* s.l., *Anaplasma phagocytophilum*), bacteria species with previously described tick-associations (i.e. *Candidatus* Midichloria mitochondrii), and indicates other bacterial taxa yet to be associated with ticks and of undetermined function. The wide range of bacteria diversity among individuals may have correlation to the *I. ricinus'* habitat and tick-host interactions.

This research creates the basis for elucidating the microbial communities residing in *I. ricinus* tick, directed towards a better understanding of the composition and the structure of bacteria in a medically important arthropod vector.

M-47

The evolution of eusociality (i.e., reproductive and non-reproductive colony members) represents a major transition in evolution, and is thus likely associated with other evolutionary changes occurring at additional levels of biological organization. The link between host-symbiont dynamics and host social biology is a potentially important, yet heretofore unstudied, aspect of coevolution.

Bees are an ideal group to study microbial symbionts within an evolutionary framework because diets in each phase of their life cycle are based upon just two ingredients: nectar and pollen from flowering plants. However, access to and use of these resources varies as a function of social caste. One of the salient features of bee biology most directly associated with gut microbial composition is variation in nutrient sequestration within colony members of social species. Dominant individuals specializing in reproduction (queens) have a larger and more protein-enriched diet than subordinate individuals (workers), which specialize in foraging and brood care. This link between food and caste is an important foundation for social evolution, as caste determination is highly influenced by nutrition during development in the highly eusocial stingless bees and honey bees. Moreover, individual task allocation within social bee colonies is regulated by stimuli associated with food intake, as well as through evolutionary modifications of conserved nutrient-related molecular pathways. Gut microbes may mediate the relationship between bees and their floral resources. In turn, the host intestinal environment created by variation in social status may influence the microbiome. Symbiotic relationships between bees and their microbiomes may also enhance food conversion, alter host reproductive potential, influence host behavior, or provide a protective barrier against pathogens.

We are investigating the effects of bee social organization on gut microbial symbiont dynamics, and reciprocally, the effects of community composition of gut microbes (“microbiome”) on bee social dynamics. In our first experiment, we tested the hypothesis that colony members of social species have caste-specific, functionally disparate microbiomes. Although previous research suggested the gut microbiome of honey bees is limited in diversity and is consistent across populations, these analyses focused primarily on workers. We characterized the hindgut microbiomes from queens, nurses, foragers, and drones from three colonies of the Western honey bee (*Apis mellifera*) using next-generation sequencing techniques. We extracted DNA from the hindgut of each individual; PCR amplified the V1-V3 hypervariable region of the16s rRNA gene; and pyrosequenced these fragments using 454 technology with GS FLX Titanium series reagents. Our presentation will report the results of this study within the framework of the coevolutionary reciprocity between social behavior and host-symbiont dynamics.

M-48

Exploring the microbial metagenomic diversity of the copepod host during habitat invasions

*Lee, Carol Eunmi; Gelembiuk, Greg W.; Silva, Joana C.
(CEL, GWG) Center of Rapid Evolution, University of Wisconsin Madison, Madison, WI 53706; (JCS) Department of Microbiology and Immunology and Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD 21201, USA

While many human diseases are known to be waterborne (such as cholera, diphtheria), the prevalence and diversity of pathogens associated with aquatic vectors have never been previously explored. The copepod *Eurytemora affinis* is a dominant coastal species, such that its microbial community assemblage could directly impact human health. Within the past few decades, the copepod *E. affinis* has invaded freshwater habitats multiple times independently throughout the world (Lee 1999). As *E. affinis* is an invasive species, moving frequently from coastal habitats into inland waters, harboring pathogens could have serious implications for disease transmission. We are currently exploring the microbial composition and metagenome of marine copepod invaders, uncovering an enormous diversity that had been largely uncharacterized.

We performed high-throughput 454 pyrosequencing of a segment of the 16S rDNA gene from microbial assemblages associated with saline and freshwater copepod populations. We sequenced paired samples from ancestral saline source and freshwater derived populations, representing four independent invasions from two genetically distinct clades. Levels of diversity in our samples were high, comparable to those found in microbiome studies of free-living marine bacterioplankton. The microbial assemblage within the copepod differed sharply from that of the surrounding water. **Most remarkable was the presence of a large number of unknown taxa, including undescribed genera and families**. For example, the single most common taxon in our freshwater samples (comprising >1/4 of sequences) appears to belong to an undescribed genus. Moreover, our sequencing uncovered a wide variety of potentially pathogenic taxa, including *Vibrio cholerae*.

When we compared microbial composition across paired samples of saline and freshwater populations, representing multiple independent invasions, **we found dramatic shifts in the copepod microbiome and parallel shifts across independent invasions**. However, despite the shifts in microbial community composition with habitat invasions, **a set of core microbial taxa appeared to persist in all copepod populations across all environments**. These core taxa include species of *Klebsiella*, *Streptomyces*, and *Pseudonocardia*, which are known to be mutualists in ants and other insects. Such microbial constituents might have important functions for host fitness during invasions (e.g. by producing antibiotics or providing nutritional benefits). We are currently performing whole-genome shotgun sequencing of the copepod microflora in order to capture the repertoire of functional genes in the copepod microbiome. Our goal is to gain insights into the metabolic and biogeochemical processes associated with this undiscovered component of the ecosystem.

SESSION: FUNCTIONAL GENOMICS

FG-49

Transposon-based germ-line transformation of the Coffee Berry Borer, *Hypothenemus hampei*: Opportunities for functional genomics.

Acevedo, Flor E.; Barrera, Carlos; Aluvihare, Channa; Harrell, Robert A.; *O'Brochta, David A.: Benavides, Pablo.
(FEA) Pennsylvania State University, Department of Entomology, University Park, PA 16802; (CB, PB) Cenicafé. Department of Entomology, Km. 4 Vía Antigua Chinchiná - Manizales, Colombia; (CA, RAH, DAO'B) University of Maryland, Institute for Bioscience and Biotechnology Research, Rockville, MD 20850; (DAO'B) University of Maryland, Dept of Entomology, College Park, MD 20850

Coffee is one of the most traded commodities in the world and supports the livelihoods of 20-25 million families. The most important insect pest of coffee is *Hypothenemus hampei* (Coffee Berry Borer; Coleoptera: Curculionidae), a small beetle that infests and destroys coffee berries and the beans within. Originating in Africa and now found in all coffee-producing nations including the United States (Hawaii), the Coffee Berry Borer severely impacts the quantity and quality of coffee production. Various efforts are underway to find ways to control this costly pest and a leader in those efforts is Colombia's Centro Nacional de Investigaciones de Café - Cenicafé (National Coffee Research Center). Cenicafé has more than 70 years of experience in investigating the biology, genetics and breeding of coffee and the pests that attack it. Ongoing research at Cenicafé includes the investigation of *H. hampei* with the intent of understanding the biological basis for this insect's unique life history characteristics and habits, and to use this knowledge to devise new strategies to control or eradicate these insects. To this end, Cenicafé has a genomics and functional genomics research program that has recently determined the genome sequence of *H. hampei* and a number of expressed sequence tag (EST) libraries from various tissues and developmental stages. In collaboration with the University of Maryland, Cenicafé has successfully developed transgenic technologies for *H. hampei* and has demonstrated that the transposons *Minos* and *piggyBac* are effective transgene vectors in this species. Here we describe our successful efforts to genetically modify the Coffee Berry Borer using transposon-based gene vectors. We undertook a photomicroscopic analysis of embryonic development of *H. hampei* in order to determine the timing of major embryological events and the location of presumptive germ cells and other embryological landmarks. Female *H. hampei* lay approximately 3 eggs per day after they have taken up residence within a coffee berry and these conditions can be duplicated on an artificial diet in the laboratory. Eggs of *H. hampei* are large, relative to the size of adult females, and are approximately 500 µm long and 100 µm wide. *Hypothenemus hampei* eggs have no visible chorion and are covered by a transparent vitellin membrane. Embryogenesis of *H. hampei* is typical of long germ-band insects such as *Drosophila melanogaster* although the timing is species-specific. At 27ºC embryogenesis takes approximately 144 hrs and begins with a series of nuclear divisions resulting in a syncitium. The syncitial blastoderm stage occurs during the first ten hours of development and at approximately 8.5 hours after egg laying (AEL) pole cells are visible at the posterior end of the embryo. Gastrulation begins at approximately 36 hrs AEL with maximum germ band extension occurring at 60 hrs AEL. Dorsal closure begins at 96 hrs AEL and sclerotization of the larval mandibles begins at 120 hrs AEL followed by eclosion of the first instar larvae approximately 24 hrs later. The embryos of *H. hampei* are readily manipulated and injected with transposon-based gene vectors. Two transgenic experiments were completed. In the first experiment we injected a mixture of two vectors and two transposase expressing plasmids (pPB (*piggyBac*) 3xP3ECFP + PBhsΔSST & pMos (*Mos*) 3xP3DsRed + PKhsp82Mos) and recovered 251 G_0 adults. One transgenic line containing pPB 3xP3ECFP was recovered. In the second experiment a different mixture of vectors and transposase-expressing plasmids was injected (pMinos (*Minos*) 3xP3DsRed + PHss6hs1LMi20 & pHermes (*Hermes*) Actin5CEGFP + pBCHsHH) resulting in 144 G_0 adults. Three of these adults produced transgenic progeny containing pMinos 3xP3DsRed. We will now integrate three critical Coffee Berry Borer technology platforms (genome sequence, EST profiles and transgenic technology) to advance the functional genomics efforts at Cenicafé by creating and using transposon-based forward genetics technologies (gene- and enhancer-traps). These technologies will be used initially to investigate the physiological genetics of feeding, digestion and excretion of *H. hampei* larvae and adults.

FG-50

Investigating the Spread and the Mechanisms of DDT Resistance in M-form *Anopheles gambiae* from BENIN

Fiacre Agossa[1*]., Christopher Jones[2]., Hilary Ranson[2]., Vincent Corbel[1]., Martin Akogbeto[1].
[2]Vector Group, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA
[1]Centre de Recherche Entomologique de Cotonou (CREC), 06 BP 2604, Cotonou, Bénin

Insecticide resistance in *Anopheles* mosquitoes is threatening the success of malaria control programs. A time series of collections showed that the L1014F kdr allele has increased in frequency from 6% to 47% in 5 years and as a result there has been a drastic reduction in susceptibility to DDT and Pyrethroids in the north part of Benin (Malanville). In this study, we investigate the contribution of other potential resistance mechanisms. RT-qPCR was used to analyse *CYP6M2, CYP6P3, GSTe2, GSTD3* and *chymotrypsin* (candidate genes associated lately with DDT resistance following microarray analysis) transcription profile from DDT resistant, *An. gambiae* M-form field collections from Houéyiho and Malanville (vegetable growing areas) against the laboratory susceptible M-form strain from N'Goussou (Cameroon). Several detoxifying enzymes were up regulated, despite of the presence of kdr (0.81 in Malanville and 0.9 in Houéyiho), *GSTe2* looks like the strongest candidate gene for DDT resistance in Benin. We looked also for the origin of the first event of L1014S kdr mutation in *An. gambiae* s.l. in Malanville, and the upstream inton 1 sequencing analysis revealed a possible migration event from Burkina-Faso. In addition we have reported for the first time the N1575Y mutation which also is associated with L1014F mutation in Malanville.

FG-51

Genomic organization and evolution of the vertebrate-specific venom neurotoxin α-latrotoxin from the Western black widow spider *Latrodectus hesperus*

Bhere, Kanaka V.; Ayoub, Nadia A.; Hayashi, Cheryl Y.; *Garb, Jessica E.
(KVB, JEG) University of Massachusetts Lowell, Department of Biological Sciences, Lowell, MA 01854; (NAA) Washington and Lee University, Department of Biology, Lexington, VA 24450; (CYH) University of California Riverside, Department of Biology, Riverside CA 92521

The Western black widow spider (*Latrodectus hesperus*) has been selected for whole genome sequencing by the i5K initiative. Widow spiders (genus *Latrodectus*) are widely known for their highly potent neurotoxic venom, which contains the 120 kD protein α-latrotoxin (ALTX). ALTX is a neurotoxin that elicits massive neurotransmitter release from vertebrate pre-synaptic nerve terminals, and is responsible for the extreme pain caused by black widow spider bites. The toxin is encoded by a 4200 bp gene, and is one of four latrotoxin paralogs that are exclusively known from black widows. Although black widow spiders and the ALTX toxin have important medical significance, the genomic organization and evolution of the ALTX gene remains largely unstudied. We have sequenced a ~21 kilobase fosmid clone positive for the ALTX gene, which was obtained from a genomic library constructed from *Latrodectus hesperus*. This was accomplished by randomly sequencing overlapping fragments cloned from multiple restriction digest libraries of the fosmid, followed by primer walking. Our results show the presence of a previously undescribed latrotoxin paralog found in close proximity to the ALTX gene in the *L. hesperus* genome, indicating a recent gene duplication event during the evolution of this gene family. Consistent with previous reports, both latrotoxin paralogs are single-exon genes. We characterized the genomic regions surrounding the adjacent paralogs to identify the transcriptional and repetitive elements associated with the latrotoxin genes and to better understand their regulation and evolution. Phylogenetic analyses of our genomic latrotoxins, along with published members of this family, also confirm the newly discovered paralog is the result of a recent duplication of the ALTX gene.

FG-52

Development of RNAi-based Dominant Markers for use in Genetic Pest Management

*Chu, Fu-Chyun; Lorenzen, Marcé D.
North Carolina State University, Department of Entomology, Raleigh, NC 27695

The ability to confer a dominant phenotype is required for distinguishing transgenic insects from their non-transgenic siblings. Early studies used wild-type genes to rescue a mutant phenotype in a host strain, while current studies rely heavily on fluorescent proteins. Such strategies however require mutant strains and/or specialized equipment to detect fluorescence. To overcome these problems we designed an RNA interference (RNAi)-based marker system for use in the red flour beetle, *Tribolium castaneum*, which is expected to obviate the need for specialized equipment and produce an easily observed phenotype in wild strains. The *T. castaneum aspartate 1-decarboxylase* (*TcADC*) gene was selected as the RNAi target because a lack of TcADC protein is known to change the body color of these rust-red beetles to jet-black. Moreover ADC mutant beetles appear to be perfectly healthy. We constructed a *piggyBac*-based transformation vector that uses the native *TcADC* promoter to drive expression of an mRNA that will fold back upon itself producing a 571-bp *TcADC*-specific hairpin. This dsRNA hairpin is expected to evoke the RNAi pathway, degrade the endogenous *TcADC* transcripts, and reduce (or eliminate) the production of TcADC protein. We are currently screening offspring from two sets of G_0 beetles. The first set is the result of *piggyBac*-mediated transformation in a *vermilion*^{white} (v^w) background, while the second is the result of *piggyBac*-mediated transformation in a *piggyBac* helper (M26) background. Results of these experiments will be presented. If successful, RNAi-based marker systems could be designed for use in any insect for which the sequence of a suitable gene is known.

FG-53

A Galaxy-integrated framework using network graphs to exploit multidimensional data from disparate "Omics" sources for creating/exploiting functional genomics gene sets across multiple species.

*Dunn, W. Augustine; James, Anthony; Xie, Xiaohui
(WAD,AJ) University of California: Irvine, Molecular Biology and Biochemistry, Irvine, CA 92697
(XX) University of California: Irvine, Department of Computer Science, Irvine, CA 92697

Here we develop a framework for importing, formatting, integrating, and analyzing a wide range of "Omics" data types with graph-based relationships using preexisting Galaxy tools whenever possible and writing new ones when needed. We then use it to integrate and analyze RNA-seq data (midgut-specific time-course: 0, 2, 4, 6, 8, 10 hr post bloodmeal in *Anopheles gambiae* & *Culex quinquefasciatus*); orthology/parology relationships; predicted ecdysone (20E) responsive transcription factor binding site locations, and promoter sequence conservation. Using this tool, we compile gene sets that are simultaneously correlated by way of functionality (e.g. putatively 20E-induced), evolutionary history, and maintenance of mRNA abundance-patterns across species-divergence times as large as 150 million years.

Functional Genomics frequently requires construction of such gene sets for the purposes of classification, discovery of novel gene interactions, predicting active regulatory elements, etc. The rise of the Omics Era presents the challenge of integrating data from multiple LARGE scale input streams in an informative way. Network Graphs offer a natural framework for storing and analyzing highly dimensional relationships and have provided insight when applied to areas such as gene interaction relationships as well as human relationships.

The Galaxy analysis paradigm has provided arguably the single greatest advance for bench-top biologists with regard to improved access to and usability of cutting edge bioinformatics analysis programs that usually exist only as command line applications. Local Galaxy installations are easily modified and virtually any command line application can be fitted with a friendly user interface and strung together with existing tools to form powerful and complex analysis pipelines called workflows. We have written python scripts that allow the user to incorporate many common genomics/transcriptomics data types into graph-based representations that can be used to arrive at gene sets after considering the combined relationships that result from the integration of multiple data streams (genomics, transcriptomics, etc).

FG-54
TARGETED MUTAGENESIS IN *Aedes aegypti* AS A MEANS TO ELUCIDATE THE MODE OF ACTION OF BACTERIAL TOXINS.
*Gill, Sarjeet S.; Evans, Amy M; Qureshi, Nadia; Lee, Subum
University of California, Cell Biology and Neuroscience, Riverside, CA 92521

Aedes aegypti, an important vector of human diseases, is controlled by a variety of approaches including the use of bacterial insecticides, such as *Bacillus thuringiensis israelensis*. The high insecticidal activity and the low toxicity to other organisms have resulted in the rapid use of *B. thuringiensis* as an environmentally friendly method for use in controlling of mosquito populations. To elucidate the mechanism of action of these toxins we previously demonstrated that a cadherin-like receptor in *Ae aegypti* larvae binds the Cry11Aa toxin with high affinity. We also identified aminopeptidases, alkaline phosphatases and another cadherin bind the Cry11Aa toxin with high affinity. However, it has not been established if all these different proteins are actually involved in the in vivo toxicity of the Cry11A toxin.

To address this issue we specifically targeted mutagenesis of the cadherin using zinc finger nucleases. A six base sequence was chosen that if mutagenized would lead to formation of a truncated protein after cadherin repeat 4. Mosquito embryos were injected with a pair of zinc finger nucleases mRNA. G0 adults were individually mated with wild-type, and G1 eggs obtained. The G1 progeny were analyzed using DNA fragment analysis and DNA sequencing to identify mutant alleles. Of about 20 different families derived from single G0 adults, we identified three different mutant alleles. These alleles were selected to analyze for the response to toxins. Details of the methods used for targeted mutagenesis and the response of the cadherin gene knockout to individual toxins in *Bacillus thuringiensis israelensis* will be discussed.

The ease by which specific genes can be target facilitates investigation of gene function in insects in which other genomics tools are not readily available.

FG-55
DNA methylation signatures in termites
*Glastad, Karl M.; Hunt, Brendan G.; Goodisman, Michael A. D.
Georgia Institute of Technology, Biology, Atlanta, GA 30318

DNA methylation has been implicated in many important processes such as gene expression regulation, suppression of transposable elements, and epigenetic crosstalk. DNA methylation has been identified in several insects. Furthermore, DNA methylation occurs in one social taxon where it is believed to mediate phenotypically plastic developmental outcomes. Thus DNA methylation may play a role in mediating phenotypic variation in insects. Here, we demonstrate that expressed genes of another social insect, the termite *Reticulitermes flavipes*, exhibit a strong signature of DNA methylation. Interestingly, this signature is limited to only a subset of such genes, an observation that has been made in other insects with genic DNA methylation. Gene ontology analysis further reveals that putatively methylated and unmethylated genes differ in their function, as well as their expression bias. We further explore the relationship between the signal of DNA methylation in termites and in other social insects.

FG-56
***De novo* transcriptome sequencing for unraveling the effects of insecticide selection in the German cockroach**
*Gondhalekar, Ameya D.; Scharf, Michael E.
Purdue University, Entomology, West Lafayette, IN 47907

The German cockroach, *Blattella germanica L.* is an important urban pest that contributes significantly to childhood asthma and enteric pathogen transmission. Gel bait formulations of insecticides like indoxacarb, fipronil, dinotefuran etc. are been currently used for cockroach control in the United States. Although the above mentioned insecticides are providing satisfactory control of this pest, historical data indicates that German cockroaches are capable of developing resistance to any group of insecticide that has been used for its management for a prolonged period of time. As a proactive approach toward understanding resistance mechanisms and resistance management, we selected a field strain of the German cockroach for six generations in the laboratory with a commonly used insecticide, indoxacarb. Indoxacarb tolerance in the three selected lines (replicates) varied between 4- to 7-fold and 20- to 23-fold for feeding and surface contact bioassay methods, respectively. Preliminary investigations into mechanisms of indoxacarb tolerance in the selected lines implicated cytochrome P450 monooxygenase and carboxylesterase gene families. However, a large number of studies on insecticide resistance in different insect species have shown that multiple gene families are usually involved in resistance. In the present study, to determine the effects of indoxacarb selection at the whole transcriptome-level, we performed Illumina RNA sequencing with the indoxacarb-selected and unselected lines noted above. The RNA-seq data was first used for *de novo* transcriptome assembly, and then using this data, gene expression analysis was conducted to identify the differentially expressed genes. Such information is crucial for future resistance management programs and for preserving the efficacy of insecticides like indoxacarb for German cockroach control. Moreover, we hope that the whole transcriptome data presented here will be a useful resource for gene prediction in the German cockroach genome sequencing project.

FG-57
The University of Maryland Insect Transformation Facility: A Resource for the Study of Insect Functional Genomics.
*Harrell, Robert A.; Aluvihare, C.; O'Brochta, David
University of Maryland, Institute for Bioscience and Biotechnology Research, Rockville, MD 20850

The University of Maryland's Insect Transformation Facility (UM-ITF) is a valuable resource for scientists interested in insect functional genomics. Integral to the success of functional genomics is the ability to genetically manipulate organism being studied. Although the technology for the creation of transgenic insects has been available for many years now, it is still not widely used, mainly because the technology is demanding and requires specialized equipment. Six years ago, in an effort to make this technology more accessible to

researchers, the University of Maryland set up a “State-of-the-Art” Insect Transformation Facility. The facility operates on a cost recovery model allowing researchers around the world to economically access insect transformation technology. The facility has become the “go to” resource for insect transformation, training, and facility use for many, allowing researchers to focus on the science not the technology. During the six year history of the facility the staff has built a proven track record of success. The majority of projects completed by the UM-ITF have been fee for service microinjections of the mosquito species *Aedes aegypti* and *Anopheles stephensi* but it has continued to expand its services to include other insect species, such as *Tribolium castaneum, Lucillia sericata* and various species of Drosophila other then melanogaster. The UM-ITF provides species specific service based on the biology of the species and the maturity of the transgenic technology involved. For example, the facility provides either “full service transgensis” or “microinjection only” service for *Aedes aegypti*. “Full service transgensis” includes client vector confirmation, microinjection of enough embryos to produce a contract specified number of Go adults, Go and G1 rearing and mating, G1 screening and transgenic confirmation by PCR, all for a reasonable cost. “Microinjection only” includes client vector confirmation and microinjection of a contract specified number of embryos; the injected embryos are then shipped to the client for rearing and screening. All injections include a quality assurance protocol that guarantees injections were of a quality to produce transgenic lines.

The UM-ITF staff also provides training in all aspects of insect transgensis technology. There are many reasons why it may be better for the researcher to learn the techniques of insect transgensis then for the facility to make transgenic insects for them. For example when the insect species is a regulated species where regulations governing its possession or transport make it impractical to work directly with the organism. Alternatively some researchers have the resources required to undertake insect transformation but need their staff to be trained. In these cases the UM-ITF can provide “hands on” training and post training consultation.

The UM-ITF also can provide researchers access to the facility. After an initial orientation session researchers may utilize the UM-ITF facilities at a reasonable per hour charge. The ability of researchers to use the UM-ITF equipment allows them access to “State-of-the-Art” equipment that they may not be able to afford or for which they only have limited needs.

Over the six year period that the UM-ITF has been in existence it has build a solid reputation with clients based on a successful track record. In this time the facility has a 97% success rate for producing transgenics for *Aedes aegypti*, *Anopheles stephensi* , and *Tribolium castaneum* (n=96).

FG-58

A Viral Histone H4 Interacts with Nucleosome Components and Alters Target Insect Transcriptome Revealed by an Illumina RNASeq

Hepat, Rahul [1]; Song, Hojoon [2]; [*]Kim, Yonggyun[1]
(RH, YK) Department of Bio-Sciences, Andong National University, Andong 760-749, Korea; (HS) Department of Biological Sciences, KAIST, Daejon 305-701, Korea

A viral histone H4 (= CpBV-H4) is encoded in a polydnavirus, *Cotesia plutellae* bracovirus. Its predicted amino acid sequence is highly homologous to host insect histone H4 except an prolonged N-terminal tail containing 38 amino acids with nine lysine residues. Its expression induces immunosuppression of target insects by suppressing immune-associated genes presumably through an epigenetic control. This study addressed its molecular interaction with host nucleosomes and its target genes to be altered in their transcription by an Illumina RNASeq. Reassociation of nucleosome components with a recombinant CpBV-H4 produced octamers containing each monomer of H4 and CpBV-H4 plus dimers of H2A, H2B, and H3, which were separated by a size-exclusion chromatography and analyzed by SDS-PAGE. Transient expression of CpBV-H4 in a model insect, *Tribolium castaneum*, was performed by microinjection of a recombinant expression vector and confirmed by RT-PCR and immunoblotting. Under this transient expression condition, total RNAs were extracted and read with an Illumina HiSeq 2000. Annotated transcripts were classified into different GO categories and compared with those of control insects injected with a nonrecombinant expression vector. Genes showing significant differences ($P < 10^{-5}$) in corrected transcript levels were determined to be targets manipulated by expression of CpBV-H4. When the target genes were physically mapped, they were scattered on entire chromosomes of *T. castaneum*. Moreover, the target genes were not significantly different with a typical transcriptome at the developmental stage. These suggest that the viral histone H4 alters host gene expression by a direct molecular interaction with target insect nucleosomes without clear specificity.

FG-59

Functional Genomics of Two Families of Genes Encoding Cuticular Proteins Analogous to Peritrophins in *Tribolium castaneum*

Jasrapuria, Sinu; Li, Beibei; Chaudhari, Sujata S; Arakane, Yasuyuki; Moussian, Bernard; Mezendorfer, Hans; Specht, Charles A; Beeman, Richard W; Kramer, Karl J and Muthukrishnan, Subbaratnam
(SJ, BL, SSC, KJK, SM) Department of Biochemistry, Kansas State University, Manhattan, KS, (YA) Chonnam National University, Gwangju, South Korea, (BM) Univesity of Tubingen, Tubingen Germany, (HM) University of Osnabrueck, Osnabrueck, Germany, (CAS) University of Massachusetts Medical School, Boston, MA and (RWB, KJK) USDA-ARS-CGAHR, Manhattan, KS

This study is focused on the functional characterization of several *Tribolium castaneum* genes that encode Cuticular Proteins Analogous to Peritrophins (CPAP). These proteins contain one or three copies of the chitin-binding domain, ChtBD2, with its six characteristically spaced cysteine residues. *CPAP* genes are expressed exclusively in cuticle-forming tissues and have been classified into two families, *CPAP1* and *CPAP3*. The *CPAP1* family has 10 members, each with one ChtBD2 domain, while the *CPAP3* family has eight members, each with three ChtBD2 domains. Individual members of the *CPAP1* and *CPAP3* gene families have distinct developmental patterns of expression. Many of these genes are essential for development, molting, cuticle integrity or proper locomotion and fecundity. RNA interference (RNAi) targeting *TcCPAP1-C, TcCPAP1-H, TcCPAP1-J* or *TcCPAP3-C* transcripts results in death at the pharate adult stage of development. RNAi for other *CPAP3* genes results in different developmental defects, including abnormal elytra or hindwings, abnormal gait, and/or adult/embryonic mortality. Scanning electron microscopic analysis of pharate adults following RNAi for the *TcCPAP3-D* gene reveals that elytral cuticle morphology is affected. These results provide experimental support

for specialization in the functions of several CPAP proteins in *T. castaneum* and provide a biological rationale for the conservation of orthologs within these CPAP families of proteins in insects of different orders. Many of these proteins serve essential and non-redundant functions in maintaining the structural integrity of the cuticle. This is the first comprehensive RNAi study of any family of cuticular proteins in a single insect species.

This work was supported by a grant from the national Science Foundation (IOS-1022227)

FG-60

A Comprehensive Study on the Effects of the Insect Growth Regulator Diflubenzuron in the Model Beetle Species *Tribolium castaneum*

Kelkenberg, Marco; Kim, Hee Shin; Chaudhari, Sujata S.; Kumari, Meera; Specht, Charles A.; Brown, Susan J.; Muthukrishnan, Subbaratnam and *Merzendorfer, Hans
(MK, HM) Department of Biology/Chemistry, University of Osnabrück, 49069 Osnabrück, Germany; (HSK, SJB) Division of Biology, Kansas State University, Manhattan, KS 66506, USA; (SSC, MK, SM, HM) Department of Biochemistry, Kansas State University, Manhattan, KS 66506, USA; (CAS) Department of Medicine, University of Massachusetts, Worcester, MA 01605, USA

Benzoylurea-derived insecticides are in wide use to control various insect pests. Although this class of insect growth regulators is known to disrupt molting and to affect chitin content, their precise mode of action is still not understood. To gain a broader insight into the mechanism underlying the insecticidal effects of benzoylurea compounds, we conducted a comprehensive study with the model beetle species and stored product pest *Tribolium castaneum* (red flour beetle) utilizing genomic and proteomic approaches. DFB was added to a wheat flour-based diet at various concentrations and fed to larvae and adults. We observed abortive molting, hatching defects and reduced chitin amounts in the larval cuticle and eggs. Electron microscopic examination of the larval cuticle revealed major structural changes and a loss of lamellate structure of the procuticle. We used a genomic tiling array for determining relative expression levels of about 11,000 genes predicted by the GLEAN algorithm. Genes encoding enzymes involved in chitin metabolism were unexpectedly unaffected, but several genes encoding cuticle proteins were affected. In addition, genes presumably involved in detoxification pathways were up-regulated. Interestingly, genes encoding peritrophic matrix proteins were largely unaffected, although the chitin content of the peritrophic matrix was reduced by about 50%. Therefore, we tested whether diflubenzuron-treatment affects the permeability of the peritrophic matrix. Feeding FITC-dextrans of defined sizes allowed us to monitor their distribution in the midgut of control and diflubenzuron-treated larvae. The results of this study demonstrate that diflubenzuron-treatment disrupts the barrier function of the peritrophic matrix in the *T. castaneum* midgut. In summary, the red flour beetle turned out to be a good model organism for investigating the global effects of bioactive materials such as insect growth regulators and other insecticides.

FG-61

Functional Genomics of dsRNA-Mediated Gene Silencing in Transfected Silkmoth-derived Bm5 Cells

Kolliopoulou, Anna; Iatrou, Kostas; *Swevers, Luc
Insect Molecular Genetics, Institute of Biology, NCSR “Demokritos”, Athens, Greece

While the intracellular RNAi machinery has been well characterized in *Drosophila* and *Drosophila*-derived S2 cells, no systematic investigation was reported so far regarding the genetic requirements for successful RNAi in other insect species. In this study, the involvement in RNAi of different Dicer enzymes, Argonaute effectors and dsRNA-binding proteins was evaluated in transfected Bm5 cells, which are derived from the silkmoth *Bombyx mori*.

Efficient RNAi is observed in Bm5 cells after co-transfection of a luciferase reporter construct with dsRNA targeting the luciferase ORF (dsLuc). To investigate the mechanism of RNAi, initially an “RNAi-of-the-RNAi” approach was used in which dsRNAs that target the core factors of the RNAi machinery were co-administered. Factors of all three small RNA pathways were evaluated (miRNA, siRNA and piRNA). Results indicate that knock-down of Ago1 (miRNA pathway), Dcr2 and Ago2 (siRNA pathway) as well as Ago3 (piRNA pathway) could affect the efficiency of dsRNA-mediated gene silencing. On the other hand, dsRNAs that targeted Dcr1, Loqs, R2D2, Translin, Trax-B or Aubergine proved to be non-effective. These results indicate the direct involvement of both siRNA and piRNA pathways in intracellular RNAi in Bm5 cells, while the effects of Ago1 (miRNA pathway) may be indirect.

To carry out the reciprocal set of experiments, i.e. to evaluate whether increased expression of RNAi factors can result in increased gene silencing, ORFs for silkmoth Dcr2, R2D2 and Ago2 (siRNA pathway) were cloned in expression vectors for tagged (Flag or Myc) expression in Bm5 cells. Besides functional assays of dsRNA-mediated gene silencing, the expression constructs will be used for intracellular co-localization studies for the core factors of the siRNA pathway.

Our studies aim at a better understanding of the mechanism of RNAi in the silkmoth, which in many cases appears to be refractory to silencing following injection of dsRNA into the body cavity, and that could serve as a basis to improve the technique of RNAi in lepidopteran insects.

FG-62

A model system for the study of the mode of action of the chitin inhibitor, diflubenzuron

*Kumari, Meera; Merzendorfer, Hans; Specht, Charles A.; Beeman, Richard W.; Kramer, Karl J.; Muthukrishnan, Subbaratnam
(MK, KJK, SM) Department of Biochemistry, Kansas State University, Manhattan, KS 66506; (HM) University of Osnabrück, Department of Biology/Chemistry, 49069 Osnabrück, Germany; (CAS) Department of Medicine, University of Massachusetts, Worcester, MA 01605, USA; (RWB, KJK) Center for Grain and Animal Health Research, ARS-USDA, Manhattan, KS 66502

We have focused on the elytron of the red flour beetle, *Tribolium castaneum*, as the model tissue for studying the mode of action of the “chitin inhibitor”, diflubenzuron (DFB). In addition to the ease of isolation and manipulation, this specialized wing tissue is essentially made-up of epidermal cells and is amenable to immunolocalization and transmission electron microscopic (TEM) studies. We have standardized the protocol for topical administration of DFB on precisely timed prepupae to achieve the desired level of mortality (50-

90%) on day 5 of the pharate adult stage when the elytra (and other cuticle forming tissues) can be analyzed for chitin and immunolocalization of proteins involved in chitin metabolism and cuticle assembly. Exposure of prepupae to DFB at 1000 ppm results in near complete loss of chitin in the newly forming procuticle of the elytron and body wall. However, RT-PCR analyses indicate no significant changes in transcript levels for chitin synthase-A (CHS-A) in these insects compared to mock-treated controls. Immunolocalization of CHS-A also did not detect any significant change in the level or localization of this enzyme inside the cell following DFB-treatment. However, there were changes in the distribution of other cuticular proteins in DFB-treated insects. These results will be presented.

TEM analysis of larvae fed different concentrations of DFB showed a progressive loss of chitin-protein laminae in the newly forming procuticle (1). At high concentrations, the procuticle closest to the epidermal cell was amorphous presumably because of the reduction in chitin. Adults females fed a diet containing 100 ppm DFB were viable and fecund, but the eggs failed to hatch. Fully grown larvae could be visualized inside the egg case, but they failed to hatch. Analysis of chitin content of eggs collected at different time points during embryonic development indicated substantial reduction in chitin content of eggs compared to control consistent with the reduction of chitin in the newly forming larval cuticle.

This work was supported by the NSF grant IOS-1022227.

FG-63

Multicopper Oxidase-1 Plays a Role in Iron Metabolism in *Drosophila melanogaster*

Lang, Minglin*; Gorman, Maureen J.; Kanost, Michael R.
(ML, MRK, MJG) Kansas State University, Department of Biochemistry, Manhattan, KS 66502, USA; (ML) Agricultural University of Hebei, College of Life Science, Baoding, China 071001

Iron is an essential metal involved in many fundamental biochemical processes within a living cell or organism. In mammals, ceruloplasmin and hephaestin, two enzymes known as multicopper oxidases, have been well characterized as having iron oxidase (ferroxidase) activity and playing critical roles in iron metabolism. Surprisingly, multicopper ferroxidases from invertebrates have not been identified, and the mechanisms of iron metabolism in invertebrates (including model organisms such as *Drosophila melanogaster*) are not well understood. A candidate insect ferroxidase is multicopper oxidase-1. Multicopper oxidase-1 is present in all insect genomes analyzed to date; however, its physiological role in insects is unknown. In this study, a set of experiments was designed to characterize the function of *Drosophila* multicopper oxidase-1 (CG3759, DmMCO1). By using immunohistochemistry and confocal laser-scanning microscopy, we localized DmMCO1 to the basal side of cells in the digestive system and Malpighian tubules. Using Prussian blue staining and a ferrozine assay, we tested iron accumulation in the iron cell region of the midgut and in whole bodies, respectively. Knockdown of DmMCO1 markedly reduced iron accumulation in the iron cells of both wandering larvae and adults, as well as in whole wandering larvae and adults. Moreover, knockdown of DmMCO1 increased the longevity of flies cultured on food supplemented with a toxic concentration of iron. qRT-PCR results showed that knockdown of DmMCO1 resulted in a significant reduction of DMT1 mRNA; because DMT1 transports iron from the food into the midgut epithelial cells, this reduction in DMT1 could explain the reduction in iron in knockdown insects. Finally, we purified recombinant DmMCO1 and verified that it has ferroxidase activity. Taken together, our results indicate that multicopper oxidase-1 is a ferroxidase that participates in iron metabolism in insects.

FG-64

A "Genome-to-Lead" Approach for Insecticide Discovery: Chemical Library Screening Reveals Antagonists of Vector Arthropod Dopamine Receptors with *In Vivo* Toxicity

*Meyer, Jason M.; Ejendal, Karin F.K.; Conley, Jason M.; Avramova, Larisa V.; Garland-Kuntz, Elisabeth E.; Giraldo-Calderón, Gloria I.; Brust, Tarsis F.; Watts, Val J.; Hill, Catherine A.
(JMM, GIG-C, CAH) Purdue University, Department of Entomology, West Lafayette, IN 47907; (KFKE, JMC, EEG-K, TFB, VJW) Purdue University, Department of Medicinal Chemistry and Molecular Pharmacology, West Lafayette, IN 47907; (LVA) Purdue University, Integrated Screening Technologies, Bindley Bioscience Center, Discovery Park, West Lafayette, IN 47907

Research is needed to develop new mode-of-action chemistries to control vector arthropods. The biogenic amine-binding G protein-coupled receptors (GPCRs) are attractive targets for insecticide discovery research because of their integral role in diverse and essential neurobiological processes. The Purdue Invertebrate Receptor Group (PIRG) is mining vector genomes for GPCRs suitable for functional characterization and evaluation as novel chemical control targets. Recent studies of the PIRG involve molecular cloning and pharmacological analysis of the D_1-like dopamine receptors in the yellow fever mosquito, *Aedes aegypti,* and Lyme disease tick, *Ixodes scapularis*. In heterologous expression studies, the comparative pharmacology of these receptors was investigated in response to dopamine and other biogenic amines, as well as to several agonists and antagonists. A cell-based assay was implemented to enable screening of the $LOPAC_{1280}$ chemical library to identify lead antagonistic chemistries active at the mosquito and tick DOP2 receptor. The toxicity of selected antagonists identified in the screen was confirmed *in vivo* with mosquito larval bioassays, and their relative potency paralleled that observed in the cell-based system; this discovery provided a proof-of-concept for our "genome-to-lead" approach for insecticide discovery. Ongoing work involves structure-activity relationship studies to explore the activity of chemistries related to lead molecules at the vector dopamine receptors *in vitro* and to investigate their effect on mosquitoes and other key arthropod pests *in vivo*. Our research may change the paradigm of drug discovery to combat neglected arthropod-borne diseases by which genomic data are exploited to facilitate a target-specific method for vector control.

FG-65

MicroRNA-mediate host-pathogen interaction in *Bombyx mori*

*Nagaraju, J.; Singh, J.; Singh, CP.; Badrinarayan, D.
Centre for DNA Fingerprinting and Diagnostics (CDFD), Laboratory of Molecular Genetics, Hyderabad, India.
(jnagaraju@cdfd.org.in)

MicroRNAs (miRNAs) are evolutionally conserved class of small (~22 nucleotides) non-coding RNAs that post-transcriptionally regulate the expression of mRNAs bearing complementary target sequence. Since the discovery of miRNA, a plethora of studies has shown the important role of miRNAs in almost all the biological processes in eukaryotes. Recently, miRNAs have emerged as critical players in the intricate host-viral interactions. Host utilizes its cellular miRNAs for not only regulating its own immune genes, but also includes modulation of various viral genes required for viral entry and its establishment inside the host. Apart from eukaryotes, viruses have also evolved a highly sophisticated gene-silencing mechanism mediated by miRNAs with the purpose of enhancing production of progeny viruses by combating the host-immune response and abrogating the molecular arsenals mounted by the host cells.

We have carried out a comprehensive study using both computational and experimental methods for identification and characterization of many novel miRNAs and their targets from *B. mori* and its natural pathogen *Bombyx mori nucleopolyhedrosis virus* (BmNPV), which play an important role in the intricate host-pathogen interactions. As a result, we have discovered four miRNAs encoded by BmNPV, which turned out to have a number of interesting cellular targets that are involved in different antiviral host defense mechanisms. One of these targets was GTP-binding nuclear protein Ran, an important component of Exportin-5-mediated nucleocytoplasmic transport machinery mainly involved in non-coding RNA trafficking. The suppression of Ran by the viral miRNA leads to impaired Ran-mediated small RNA export from nucleus to cytoplasm in the host, which results in enhanced BmNPV proliferation. These results give an insight into one of the evasion strategies used by the virus to counter the host small RNA-mediated defense for its effective proliferation, and has relevance in the development of insect virus control strategies.

FG

Transposon-based forward and reverse genetics in Anopheles mosquitoes.

*O'Brochta, David A.; Pilitt, Kristina L.; Harrell, Robert A.; Aluvihare, Channa; Alford, Robert T.
(DAO'B, KLP, RAH, CA, RTA) University of Maryland, Institute for Bioscience and Biotechnology Research, Rockville, MD 20850;
(DAO'B) University of Maryland, Entomology, College Park, MD 20850

Transposons can be configured to carry specific sensors that permit the detection of various genetic elements such as enhancers, promoters and genes. Transposon-based enhancer detection is not only an effective way to sense the presence of enhancers but when coupled to robust binary transcription regulatory circuits such as the *Gal4* system, the enhancer sensing elements can be used to drive the expression of any transgene under the regulatory control of a promoter containing the *Gal4* upstream activation sequence (*UAS*) without having to physically isolate and characterize the enhancer. This can greatly increase the options available for expressing transgenes in specific temporal and spatial patterns without increasing the need to generate new transgenic lines with the transgene of interest. Transposons configured somewhat differently and after they have inserted into introns or exons can be used to sense the presence of genes. These so-called gene traps not only enable genes to be detected based on patterns of transgene expression from reporter genes within the sensing element but in many cases transposon integration results in disabling the target gene. The resulting recessive hypomorphic or null mutations can be of great use in trying to determine a gene's function. Transposon-based transgenic technologies have been available for mosquitoes for over a decade but they are utilized only occasionally because the creation of primary transgenic mosquitoes can be technically challenging and because some integrated transposons have shown little or no remobilization activity in *Aedes aegypti* and, in some cases, *Anopheles stephensi*, severely limiting the utility of the transposons as functional genomics tools. The *piggyBac* transposon does not suffer from these limitations in *Anopheles stephensi*. Integrated *piggyBac*-based gene vectors in *An. stephensi* readily remobilize in the presence of functional *piggyBac* transposase and can efficiently detect enhancers. Functional *Gal4*-based enhancer- and splice acceptor site-based gene-trap systems were assembled and introduced into *An. stephensi*. The detection strategies and performance of these systems are described as well as descriptions of collections of resources arising from enhancer- and gene-trap screens. The availability of these powerful functional genomics technologies in conjunction with currently available structural genomic resources will enhance efforts to investigate the genetic basis of important aspects of mosquito biology. These technologies also provide the vector biology community with new genetic resources.

FG-66

Identification and expression analysis of putative genes related to innate immune pathways in the Asian honeybee, *Apis cerana*

Park, D.; Kim, W.; Hong, K.; Jung, J. W.; Kim, J.; Kwon, H. W.*
Seoul National University, WCU Biomodulation major, Department of Agricultural Biotechnology, Seoul, Korea 151-742.

Genes related with behavioral and physiological resistance mechanisms between European honeybee and Asian honeybee seem be somewhat apart each other based on recent transcriptome analysis by NGS. To study the genes related to the innate immune pathways of the Asian honeybee, *Apis cerana*, total RNA samples were extracted from the 500 worker bees's brains, antennae, and hypopharyngeal glands. The total RNA was sequenced with 454-pyrosequencing technique to read 271,728 sequences, of which about 70% of the sequences were subsequently made contiguous sequences to construct a transcriptome database with 20,171 isogroups. Based on this transcriptome database, 50 gene sequences of *Apis mellifera*, which are homologous to innate immune related genes in cell signaling pathways such as Toll, Imd, and JNK in the fruit fly, *Drosophila melanogaster*, were subjected to BLAST search to identify homologues in *Apis cerana*. Coupled with injection experiments with honeybee pathogens to *A. cerana* and *A. mellifera*. We examined the expression levels of the genes involved in immune responses. This study will pave the way to better understand molecular and cellular mechanisms underlying immune signaling and induction on various endemic/exotic pathogens in two different honeybee species.

FG-67

Analyses of RNAi gene targets for the control of western corn rootworm

Ramaseshadri, Partha*; Bolognesi, R.; Zhao, S. ; Wiggins, E. ; Clinton, W.; Heck, G.; Flannagan, R.; Segers, G.
Monsanto Company, 700 Chesterfield Pkwy W, Chesterfield, MO 63017

Monsanto has developed an *in planta* delivery method of insect control utilizing the dsRNA mediated gene suppression (RNAi) pathway, a first of its kind for crop pest management. The method is mediated by expression of a corn rootworm specific double stranded RNA (dsRNA) in plant tissues that, when ingested by corn rootworms, triggers the RNAi mechanisms in the insects, conferring protection against the rootworm complex (*Diabrotica spp.*) in corn. Here we report the RNAi mechanisms in Western corn rootworm (WCRW, *Diabrotica virgifera virgifera*) with respect to target gene specificity using both *in silico* and functional approaches. Small RNA profiles derived from the region encompassing the RNAi target sequence were obtained from WCRW insects fed with dsRNA correspondent to the target gene region. The results revealed no transitivity of siRNA. We also evaluated whether shorter transcripts embedded into longer carriers are active against the insects. *In planta* assay with a target sequence of 27 bp embedded within a 150 bp carrier showed comparable activity against WCRW in relation to results obtained with the longer transcripts (≥150 bp). In addition, the importance of RNAi target specificity was validated: 1) by *in silico* analysis of the target gene variation among rootworm species from different geographic locations; and 2) by comparing the activities of two different targets with variation in the SNP frequencies between related rootworm species. These studies provide insights on the fine tuning of RNAi target selection that could be employed to control insect pests through *in planta* delivery method.

FG-68

The Termite Gut Bioreactor: Transcriptional and Translational Analysis of Lignin Degradation

Sethi, Amit*; Slack, Jeffrey; Kovaleva, Elena S.; Buchman, George W.; Scharf, Michael E.
(AS, MES) Department of Entomology, Purdue University, West Lafayette IN, USA; (JS, ESK, GWB) Chesapeake-PERL Inc., Savage MD, USA

Production of biofuels from lignocellulosic biomass, the woody part of plants, is very attractive and sustainable as it is the most abundant renewable resource available on Earth. However, the major limiting factor in lignocellulose processing for biofuel production is the recalcitrance of lignin. Lignin is a complex phenylpropanoid polymer found in almost all terrestrial plants. This robust resin is highly resistant to both chemical and biochemical processing. However, termites have evolved various physiological mechanisms in their gastrointestinal tracts to overcome the lignin obstacle in wood to release fermentable simple sugars. Thus, knowledge on gained on termite digestive mechanisms can be exploited in the design and operation of industrial bioreactors for efficient enzyme-based biofuel production from plant biomass. For this reason, most recent termite metagenomics studies have been focused on analyzing cellulose digestion capabilities of termites; however, while lignin is degraded in termite gut within hours, the mechanisms by which termites digest lignin are not yet clear. Here, we fed *Reticulitermes flavipes* termites three diets containing differing degrees of lignin complexity and used high throughput Roche 454-titanium pyrosequencing, MS-MS proteomics, and recombinant enzymology approaches to ask the question: what enzyme systems do termites use to deal with their lignin-rich lignocellulose diets? Our findings reveal over 9,500 distinct host and symbiont transcripts that are differentially expressed in response to diets with varying degrees of lignin complexity, including over 300 responsive cellulase, hemicellulase and candidate lignase transcripts. Parallel proteomic investigations and functional digestive studies with recombinant lignocellulases provided strong evidence of congruence at the transcriptional and translational levels. Our findings provide insights into key lignocellulases that termites use to cope with a lignin-rich diet and reveal novel, more efficient bioconversion strategies for use in industrial biofuel production from lignocellulosic biomass.

FG-69

IDENTIFICATION OF IMMUNE SYSTEM GENE SILENCING TARGETS IN A *DE NOVO* ASSEMBLY OF THE TRANSCRIPTOME OF THE SQUASH BUG, *ANASA TRISTIS* (DE GEER) (HETEROPTERA: COREIDAE)

Shelby, Kent S., USDA-ARS-Biological Control of Insects Research Laboratory, 1503 S. Providence Rd., Columbia, MO 65201, USA.

The Squash bug is a major piercing/sucking pest of cucurbits causing extensive damage to plants and fruits, and transmitting phytopathogens. There are few effective biological control agents or cultural practices for controlling this highly destructive pest. A promising new approach for control is insect pest specific RNAi targeted to critical physiological systems *via* the plant phloem. Thus we have initiated studies to determine the feasibility of delivering immunosuppressive RNAi to phloem feeding insects such as the squash bug. First steps towards this goal include construction of the first *de novo* exome, identification of candidate gene silencing targets, and laboratory based *per os* delivery methods for gene silencing of immunity. RNA was extracted from insects challenged with bacterial and fungal immunoelicitors, insects fed on different cucurbit species, and insects from all life stages from egg to adult. All treatments and replicates were separately barcoded for subsequent analyses, then pooled for sequencing in a single lane using the Illumina HiSeq2000 platform. Over 211 million 100-base tags generated in this manner were trimmed, filtered, and cleaned, then assembled into a *de novo* reference transcriptome using the Broad Institute Trinity assembly algorithm. The assembly was annotated using NCBIx NR, BLAST2GO, KEGG and other databases. Of the >130,000 total assemblies 37,327 were annotated identifying the sequences of candidate gene silencing targets from immune, and other physiological systems. To accomplish expression profiling the 100-base tags from each treatment, controls and replicates were independently aligned to the annotated assembly of the *A. tristis* reference transcriptome. Immune system components upregulated by infection will be presented. From this first reference transcriptome a comprehensive RNA-seq expression profiling of squash bug immune defense has been completed, and candidate gene silencing targets have been identified to accomplish RNAi-based immunosuppression.

FG-70
Gene Expression in a *Tribolium castaneum* cell line
Silver, Kristopher*; Beeman, Richard; and Park, Yoonseong
(KSS, RB) USDA-ARS-Center for Grain and Animal Health, Manhattan, KS 66502; (YP) Kansas State University, Entomology, Manhattan, KS 66506

Beetles are important agricultural pests that account for a large amount of crop loss annually, both pre- and post-harvest. Currently, *Tribolium castaneum* is the only beetle with a fully-sequenced genome, and offers a sophisticated toolkit for analysis of gene function, including RNA interference (RNAi) for knockout of potential biopesticide target genes. Recently, a cell line was successfully established using a mixture of *Tribolum castaneum* adults and pupae, offering a unique opportunity to investigate beetle gene expression and basic cell biology *in vitro* with the added aid of a complete genome sequence. Accordingly, we have used RNA from a modified version of this cell line as a template for next-generation transcriptome sequencing. The gene expression patterns of the cultured cells will be compared to those of tissues and stages of development at the organismal level, and will be used to guide further studies into unique aspects of cell biology of beetles.

FG-71
Identification of direct FOXO targets controlling diapause of the mosquito *Culex pipiens* by chromatin immunoprecipitation sequencing (ChIP-Seq).
Sim, Cheolho ; Bai, Xiaodong; Denlinger, David L.
(CS) Department of Biology, Baylor University, Waco, TX 76798, USA; (DLD, XB) Department of Evolution, Ecology, and Organismal Biology and Department of Entomology, Ohio State University, 318 West 12th Avenue, Columbus, OH 43210, USA; (XB) Center for RNA Molecular Biology, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106, USA

FOXO, a forkhead transcription factor, is a key regulator of adult diapause in *Culex pipiens*. The precise mechanism by which FOXO regulates multiple functions, however, is poorly understood. Here, we used chromatin immunoprecipitation sequencing (ChIP-Seq) to identify direct targets of FOXO. Our initial analysis centered on the sequence reads from ChIP-Seq, which yielded 4642 biding sites in a cut-off level between 12 to 1,202 reads supporting each binding site. Then, we only selected the binding sites, which have consensus binding motif, GTAAACA (A/T) of invertebrate FOXO. The nearest one gene in a 10 kb region surrounding the predicted binding site was extracted for each binding site, resulting in a dataset containing the genes that are potentially regulated by FOXO. The genes in the dataset were annotated for sequence similarity by using blastp. By selecting the candidate genes based on their functional relevance to diapause syndrome, we recapitulated four categories, such as cell cycle & growth regulation, metabolic pathways, stress tolerance and lifespan extension. Thus, we could prioritize 13 targets for further analysis. Among the candidate sequences, 12 targets were validated by a quantitative PCR method. Our results show that the ChIP-Seq based targeting strategy leads to greater improvement for FOXO target genes than other screening strategies.

FG-72
Gene Class Specificity of Transcription Elongation Factors in Drosophila
Smith ER; Lin C; *Garrett AS; Thornton J; Mohaghegh N; Hu D; Jackson J; Saraf A; Swanson SK; Seidel C; Florens L; Washburn MP; Eissenberg JC; Shilatifard A
(ERS,CL,*ASG,JT,NM,DH,AS,SKS,CS,LF,WMP,AS) Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO 64110, USA; (JJ,JCE) Edward A. Doisy Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, Saint Louis, MO 63104; (MPW) Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, KS 66160, USA

Eleven-nineteen lysine-rich leukemia (ELL) participates in the super elongation complex (SEC) with the RNA polymerase II (Pol II) CTD kinase P-TEFb. SEC is a key regulator in the expression of HOX genes in mixed lineage leukemia (MLL)-based hematological malignancies, in the control of induced gene expression early in development, and in immediate early gene transcription. Here, we identify an SEC-like complex in Drosophila, as well as a distinct ELL-containing complex that lacks P-TEFb and other components of SEC named the “little elongation complex” (LEC). LEC subunits are highly enriched at RNA Pol II-transcribed small nuclear RNA (snRNA) genes, and the loss of LEC results in decreased snRNA expression in flies and mammals. The specialization of the SEC and LEC complexes for mRNA and snRNA-containing genes, respectively, suggests the presence of specific classes of elongation factors for each class of genes transcribed by RNA polymerase II.

FG-73
Transcriptome analyses of larvae and pupae stages of *Anastrepha fraterculus* (Diptera: Tephritidae) investigated by Illumina RNA-seq technology.
Sobrinho Jr. Iderval S.; Campanini, Emeline B.; Andrade, Paulo H.M.; Rezende, Victor R.; Zuffo, Livia; Lima, Andre L.A.; Ribeiro, Daniel, T.; Nakamura, Aline M.; Chahad-Ehlers, Samira ; de Brito, Reinaldo A.*
Dept. Genetics and Evolution, Universidade Federal de Sao Carlos, Via Washington Luis Km 235, CP 676. 13565-905 Sao Carlos, SP. Brazil.

Species of the *fraterculus* group of *Anastrepha* (Diptera: Tephritidae), especially the South-American Fruit Fly (*Anastrepha fraterculus*) are some of the species of greatest economic importance because of their wide distribution and large range of fruits they attack. Several species of the group *fraterculus* are generally identified by subtle differences in morphological traits, which is particularly troublesome because of the inherent plasticity in these markers and potential existence of cryptic species in the group. This problem has not been placated by the use of some nuclear and mitochondrial molecular markers, which have, in general, shown high levels of intraspecific polymorphism but limited phylogenetic resolution among species. In our lab, we have screened female and male reproductive cDNA libraries in order to identify possible targets of positive selection that might show species specific differences among species of the group in order to investigate patterns of differentiation and speciation in this group, but so far have failed to identify genes

potentially involved with species differences in the group fraterculus. Here we turn to next generation sequencing in order to more thoroughly canvass the genes expressed in larval and pupal stages. In order to do so, we extracted total RNA from 200 stage I larvae, 20 stage II larvae and 20 stage III larvae which were equimolarly combined into two different pools of larval RNA. We also extracted 10 stage I pupae and 10 stage II pupae which were combined into two pools of pupae RNA. These four pools were sequenced using Illumina GAIIx next generation sequencing platform to generate more than 100 million mate-paired 100bp reads. More than 45000 contigs were generated using the package Trinity. These contigs have been processed using different strategies and the more expressed contigs and more relevant gene families have been subject to a more detailed investigation of patterns of expression. We will show data on this species' larval and pupae transcriptome as well as a functional investigation of the differential expression patterns at these stages, focusing on genes with high evolutionary rates that may be important in species differentiation, such as serin proteases, lectins and odorant binding proteins.

Financial Support: FAPESP 2010/02455-4 (RAB), CAPES (ISSJ)

FG-74

Lessons from Illumina Data with *Anopheles gambiae*: Identification of pri-miRNA and Discovery of Substantial Differences with Microarray Results

*Vannini, Laura; Dunn, W. Augustine; Willis, Judith H.
(LV, JHW) University of Georgia, Cellular Biology, Athens, GA 30602; (WAD) University of California, Molecular Biology and Biochemistry, Irvine, CA 92697

MicroRNAs (miRNAs) are an evolutionarily ancient component of genetic regulation well conserved in eukaryotic organisms. By binding a complementary sequence on a target mRNA, miRNAs (20-22 nucleotides) usually lead to translational repression or mRNA degradation. The miRBase website (http://www.mirbase.org) reported 66 miRNAs in *Anopheles gambiae*; all identified according their sequence homology with other miRNAs and the secondary structure of their precursors (pre-miRNAs). A miRNA can be transcribed as a pri-miRNA (more than 1 Kb long) and later trimmed into a hairpin-shaped pre-miRNA (~ 90 nucleotides long). So far, little is known about the composition of the initial miRNA transcripts and no pri-miRNAs have been identified in *A. gambiae*.

By comparing genomic sequences with transcriptome data from Illumina (RNAseq) we have identified a 1444 bp genomic fragment that does not appear to contain a functional open reading frame and is transcribed without splicing events. The aga-miR-308 (22 nucleotides) and its pre-miRNA (100 nucleotides) are the only annotated transcripts in this genomic region. This strongly suggests that the identified genomic fragment is transcribed as the pri-mir-308 in *A. gambiae* and later processed as aga-miR-308. The secondary structure of the deduced pri-mir-308 has been predicted by using the RNAfold software (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi). The obtained minimum free energy model shows a highly complex secondary structure that consists of several arms each with stems and loops.

Our original experiment was designed to compare mRNA levels 3 h after blood feeding *vs.* non–blood fed mosquitoes. Surprisingly, we found substantial differences between our Illumina data and published microarray data collected under similar physiological conditions (Marinotti *et al*., 2006 IMB 15:1). Preliminary qRT-PCR analyses confirm our Illumina data.

FG-75

Expression Analysis of Multiple Libraries of Unreplicated RNA-seq Data

*Walk, Thomas C.; Geib, Scott M.
USDA Agricultural Research Service, Pacific Basin Agricultural Research Center, 64 Nowelo St, Hilo, HI 96720

Considerations in performing expression analysis with RNA-seq data are the large number of read counts, the typically wide dispersions of background noise or sampling error, and the fact that a common experimental design includes large numbers of unreplicated samples. These factors preclude the use of standard parametric statistical tests, and make pairwise testing impractical. This research compared 3 methods for appropriate analysis using simulated data with read count distribution based upon the empirically observed distribution in RNA-seq data collected from 6 developmental libraries of the oriental fruit fly, *Bactrocera dorsalis*. The simulations consisted of 1000 transcripts of true positive differentially expressed genes and 9000 true negatives following a hyperbolic decline in number of transcripts with read count, with nondifferentially expressed read counts varying from 1 to 80000. Simulated differential expression ranged from 2 to 32 fold, and background noise was fit to negative binomial distributions with dispersions ranging from 0.00 to 0.95. Three methods tested for identifying differential expression were probabilities generated from the negative binomial cumulative distribution function (CDF test), a log likelihood ratio test developed for EST data (R test), and a Bayesian approach published for RNA-seq data (baySeq). All tests produced similar results at the extremes of differential expression and dispersion, with low specificity at high dispersion and low differential expression, and high specificity and sensitivity at low dispersion and high differential expression. Under intermediate conditions, the CDF test and baySeq produced similar results while the R test lagged behind. Performance improved with filtering out of low read counts for the CDF and R test, and with log transformation with the R test. CDF test performance declined with increased inequality in library sizes, while the R test and baySeq performance remained constant. We conclude that each test may be used for RNA-seq data. The CDF test is suitable when library sizes vary by less than a factor of 2. When dispersion is low, or with log transformation, the R test yields accurate results. The baySeq method requires input of a statistical model, and so it is appropriate if there is prior knowledge of how expression might vary among libraries. The baySeq test has been made available by previous authors as a package in the R statistical programming language. The CDF test and R test for RNA-seq data are available from the current authors in a GUI and command line based program called RNAseqR (sourceforge.net/projects/rnaseqr/)

FG-76

Pipe-lines for design, quality control and analysis of SNP genotype data in non-model organisms

*Wong, Swee Chong; Lehtonen, Rainer
University of Helsinki, Faculty of Biological and Environmental Science, 00014 Helsinki

The long-term Glanville fritillary (*Melitaea cinxia*) butterfly study system is one of the few internationally recognized metapopulation models in population and evolution biology. This information allows Glanville fritillary to be an excellent model system for integrated biology, which combines molecular, individual, population and ecological data. As many other non-model organisms, Glanville fritillary lacks the extensive genomic and genetic information. Information from the genome sequences is needed to further develop Glanville fritillary into a comprehensive model system in population and evolutionary biology. In this study, we created Bioinformatics pipe-lines for large-scale SNP genotyping, as well as genotype data quality control, as part of the effort to enable genome-wide variation detection in the Glanville fritillary. Custom-designed SNP chips are required for targeted large scale genotyping in non-model organisms. Our custom SNP chips assay designing pipe-line identifies SNP sites at genomic scaffolds, and creates suitable input files for SNP assay designer programs, such as Sequenom MassARRAY Assay Designer (Sequenom, Inc., CA) and Illumina Prem. assay design tool, ADT (Illumina, Inc., CA). This pipe-line can easily process large-scale SNP genotyping projects, where custom chips with > 50,000 SNPs will be designed.

The quality control pipe-line was created to analyze raw genotype data from Sequenom and Illumina for data visualization, filtering and pre-processing. Each SNP is plotted in separate signal intensity scatter plots to analyze the integrity and quality of the SNP data. The SNP which does not portrait the expected clustering pattern from the scatter plot is excluded from further analysis. The remaining SNPs will be analyzed for Mendelian errors, Hardy-Weinberg Equilibrium, and genotype probability checking. Subsequent call rates for each SNP and individual are calculated, where SNPs which exceed the preset evaluation threshold values are used in further statistical analysis, such as linkage and association tests in family-based and population-wise settings. Both assay design and genotype quality control pipe-lines were written with Python and R.

FG-77

Mass spectrometry-based semi-quantitative analysis reveals dynamic changes in the proteome of *Manduca sexta* larval plasma after immune challenge

*Zhang, Shuguang; Hartson, Steve; Rogers, Janet; He, Yan; Jiang, Haobo
(SZ, YH, HJ) Oklahoma State University, Entomology and Plant Pathology, Stillwater, OK 74078; (SH, JR) Oklahoma State University, Biochemistry and Molecular Biology, Stillwater, OK 74078.

Due to its large size, the tobacco hornworm *Manduca sexta* has been widely used as a model to study insect biochemistry and physiology. With the development of next-generation sequencing technology, we have extensively investigated its immunotranscriptome while the *Manduca* genome project rapidly moves forward. In contrast, there is only one published study on *M. sexta* hemolymph proteome and, limited by method, instrument and database, Furusawa et al. (2008) merely identified 58 non-redundant proteins in the larval plasma. In this work, we used state-of-the-art nano-liquid chromatography and tandem mass spectrometry (nanoLC-MS/MS) to identify, quantify, and compare proteins in hemolymph samples from bacteria- and buffer-injected larvae. We injected a mixture of *Escherichia coli*, *Micrococcus luteus*, and curdlan into day 1, 5th instar *M. sexta* larvae to trigger immune responses and, as a negative control, injected sterile buffer to another set of insects. The hemolymph samples were collected 24 h later, plasma proteins were separated by SDS-PAGE on a 4-15% gradient gel, and each lane was cut into nine slices based on the Coomassie Blue staining pattern. Proteins in different sections were extracted from the gel, digested by trypsin, and analyzed by nanoLC-MS/MS, respectively. To better investigate changes in small proteins, we also treated the plasma samples with acetonitrile to crush large proteins out of aqueous solution and, after centrifugation, the supernatants were directly analyzed by the LTQ Orbitrap XL mass spectrometer. We computationally translated all known *M. sexta* cDNA sequences and coding regions from the draft genome to generate a database that contains all the theoretical proteins in this species. Taking advantage of the transcriptomic and genomic sequence information, we have identified about 1000 proteins, many exhibiting abundance changes after the bacterial challenge as judged by significant differences in spectrum counts between the control and treatment groups. The immunity-related proteins we identified include pattern recognition receptors, signal transducers and modulators, and effectors such as antimicrobial peptides. Together with the results from our transcriptome analyses, these proteomic data further support the roles of plasma factors in defense against invading pathogens.

FG-78

Discovery of Avirulence Genes in the Hessian Fly

*Zhao, Chaoyang; Navarro, Lucio; Katz, Andrew M.; Shukle, Richard H.; Richards, Stephen; Stuart, Jeffrey J.
(CZ, LN, JJS) Purdue University, Entomology, West Lafayette, IN 47907; (AMK) Purdue University, Agronomy, West Lafayette, IN 47907; (RHS) Purdue University, USDA-ARS and Dept. of Entomology, West Lafayette, IN 47907; (SR) Baylor College of Medicine, Human Genome Sequencing Center, Houston, TX 77030

The Hessian fly (HF, *Mayetiola destructor*) is believed to interact with wheat (*Triticum spp.*), in a gene-for-gene fashion. According to this hypothesis, virulence to each individual HF resistance (*R*) gene in wheat (named *H1* through *H32* and *Hdic*) is caused by mutations in cognate HF avirulence (*Avr*) genes in the HF. To test this hypothesis, we examined whether virulence to HF *R* genes *H5, H24*, and *Hdic* is caused by mutations in different HF avirulence (*Avr*) genes. We developed mapping populations by crossing individual flies that were avirulent and virulent to each *R* gene. To map the mutations that confer virulence to each *R* gene, we developed PCR-based genetic markers associated with simple sequence repeats (SSR) present in the HF genome sequence. *H5*-virulence (*vH5*) mutations were mapped between two DNA-sequenced scaffolds (A2.5 and A2Random.1) near the centromere on the short arm of HF autosome A2. In contrast, *H24*-virulence (*vH24*) and *Hdic*-virulence (*vHdic*) mutations were mapped on X chromosomes. v*H24* was positioned within a 230-kb DNA segment in scaffold X1Random.8, which is located near the telomere on the long arm of chromosome X1. *vHdic* was positioned between scaffolds X2.6 and X2.8, which are located near the centromere on the long arm of chromosome X2. Gene prediction and RT-PCR analyses identified a candidate *vH24* gene that encodes a small protein (266 amino acids) containing an N-

terminal secretory signal peptide and an RXLR-like motif. Consistent with expectations, expression of this gene is limited to avirulent first-instar larvae. A candidate *vHdic* gene was also identified. This gene encodes another small secreted salivary gland protein (193 amino acids) whose mature peptide also includes an RXLR-like motif in the N-termini. These results bring the total number of *Avr* genes mapped in the HF genome to seven. Because resistance in *H5*-, *H24*- and *Hdic*-wheat lines is clearly elicited by different HF mutations, these results support the gene-for-gene hypothesis.

SESSION: DEVELOPMENTAL AND BEHAVIORAL

DB-80

Development of Next Generation Sequencing Resources to Study Behavior in Social Wasps: Applications to Caste Determination and Facial Recognition

*Berens, Ali J.; Toth, Amy L.
(AJB, ALT) Iowa State University, Program in Bioinformatics and Computational Biology, Ames, IA 50011; (AJB, ALT) Iowa State University, Department of Ecology, Evolution, and Organismal Biology, Ames, IA 50011; (ALT) Iowa State University, Department of Entomology, Ames, IA 50011

Genomic resources are being developed for a group of social insects, paper wasps in the genus *Polistes*, which are an important model for studying the evolution of social behavior. These wasps live in small societies consisting of a queen and her daughter workers. Although they cooperate to form a “eusocial” colony, they are considered to be “primitively eusocial” because workers have the ability become queens, and there is a substantial amount of conflict and aggression among females for opportunities to reproduce. This poster will describe two examples of the application of genomic tools to this emerging model system. 1) As in many other social insects, queen caste determination in *Polistes* is associated with higher nutritional state during larval stages. Using next-generation RNA-sequencing, we are characterizing transcriptomic signatures of queen and worker caste development and investigating whether a nutritional manipulation affects the expression of caste-related genes. 2) Some species of *Polistes* have been discovered to possess the remarkable ability to recognize faces of conspecifics. Species with facial recognition, such as *Polistes fuscatus*, can learn faces much more readily than other visual patterns compared to species without facial recognition such as *Polistes metricus*. A candidate gene expression study using qRT-PCR indicates that these learning differences are associated with changes in the expression of genes related to eye development, mushroom body development, and memory function. Additional genomic resources for *Polistes* are now under development, including de novo whole genome sequencing of the 300 Mb *P. dominulus* genome, currently being assembled and annotated.

DB-81

Correlating gene expression with gastrulation behavior in *Parhyale hawaiensis*

*Chaw, R. Crystal; Patel, Nipam H.
(RCC) University of California at Riverside, Biology, Riverside, CA 92521; (NHP) University of California at Berkeley, IB and MCB, 94720

Little is known about the molecular regulation of early morphogenesis in the amphipod crustacean *Parhyale hawaiensis.* Previous research established a *Parhyale* fate map at the 8-cell stage through lineage tracing and ablation experiments. During gastrulation, the progeny of specific blastomeres become distinct cell populations that interact to establish a multi-layered embryo. We are interested in correlating lineage-specific gastrulation behavior and gene expression. We propose combining a microarray technique that was first described in the nematode *C. elegans* with techniques available in *Parhyale.* In *C. elegans,* it is possible to fluorescently tag mRNAs in cells of interest with an epitope-tagged poly-A binding protein (FLAG-PAB) driven by cell- specific promoters. These transcripts are then detected during microarray analysis. In *Parhyale,* microinjection of mRNA and DNA constructs into specific blastomeres at the 8-cell stage is a well-established technique. We will create a Parhyale-specific mRNA construct (FLAG-ph-PAB) to tag transcripts in a given lineage. At specific gastrulation stages, tagged transcripts can then be separated and analyzed using RNA-seq or a microarray chip created from *Parhyale* EST sequences. Here, we describe early *Parhyale* development and cellular interactions during gastrulation, outline our experimental design, and show that we have successfully cloned *Parhyale*- poly-a binding protein (ph-PAB).

DB-82

Differential gene expression in the evolutionary development of the weevil rostrum

*Davis, S.R.; Cartwright, P.; Engel, M.S.
(SRD, MSE) University of Kansas, Division of Entomology, Natural History Museum & Department of Ecology & Evolutionary Biology, Lawrence, KS 66049; (PC) University of Kansas, Department of Ecology & Evolutionary Biology, Lawrence, KS 66045.

Weevils (Coleoptera: Curculionoidea) are a diverse group of extant organisms with approximately 50,000 described species. They are of great agricultural significance because they are associated with all major groups of plants and plant tissues. The weevil rostrum, for example, is a key evolutionary innovation that has enabled this group to feed on and oviposit in nearly all plant tissues, giving rise to diverse life histories and tremendous diversity in rostrum form. Insights into comparative development of the rostrum will provide insight into the evolution of this key innovation that may be responsible for the explosive radiation of the lineage. Although weevils are an enormous group and countless species are significant agricultural pests, no weevil species have been utilized in developmental studies. In order to better understand the formation and evolution of this structure, transcriptomes from the developing head tissue of 4 weevil species, representing disparate clades and divergent rostral forms, and 1 outgroup (non-weevil species) are being produced in *de novo* assemblies. While there are difficulties in assessing differences among transcriptomes from divergent taxa, tests for differential expression patterns of transcripts are being performed to characterize differences in the gene networks that are producing the profound phenotypic diversity observed in the rostrum and to better understand the genetic framework that permitted the diversification of such an immense lineage as the weevils.

DB-83

Development of Genomic Resources for Studying Acoustic Communication in the Lesser Wax Moth, *Achroia grisella*

*Gleason, Jennifer M.; Zhou, Y.; Hackett; Jennifer L.; Harris, Bethany R.; Greenfield, Michael D.
(JMG, YZ, BRH) University of Kansas, Ecology and Evolutionary Biology, Lawrence, KS, 66045; (JLH) University of Kansas, Molecular Biosciences, Lawrence, KS 66045; (MDG) Université François Rabelais de Tours, NSRS UMR 7126, Tours, France, 37200; (YZ) currently University of Vermont, Biology, Burlington, VT 05405.

Males of the Lesser Wax Moth, *Achroia grisella*, produce an ultrasonic signal through tymbal buckling to call females for mating. Extensive work on the behavioral biology of the moth has identified the major song parameters favored by females as well as extensive variation in calling song in this lekking species. To understand the genetics of calling song, we undertook a quantitative trait locus (QTL) study. For our molecular markers, we developed an EST library from head tissue. The majority of the ESTs had homologs within the *Bombyx mori* genome. To map QTL, we crossed two inbred lines from Kanas and Florida for which we had identified single nucleotide polymorphisms (SNPs). Our genetic map indicated that synteny is conserved between *A. grisella* and *B. mori.* We identified QTL for two calling song traits, pulse pair rate and pulse amplitude, as well as development time and weight. Preliminary analysis indicated a significant association of a calling song trait with the Z chromosome and a significant correlation between development time and pulse amplitude.

DB-84

Transcriptome profiling of sexual maturation and mating in the Mediterranean fruit fly, *Ceratitis capitata*

*Gomulski, Ludvik M.; Dimopoulos George; Xi, Zhiyong; Scolari, Francesca; Gabrieli, Paolo; Siciliano, Paolo; Clarke, Anthony R.; Malacrida, Anna R.; Gasperi, Giuliano
(LMG, FS, PG, PS, ARM, GG) University of Pavia, Biology and Biotechnology, Pavia, Italy; (GD, ZX) Johns Hopkins University, Molecular Microbiology and Immunology, Baltimore, MD, USA; (ZX) Michigan State University, Entomology, East Lansing, MI, USA; (ARC) Queensland University of Technology, Biogeosciences, Brisbane, Queensland, Australia

Sexual maturation and mating in insects are generally accompanied by major physiological and behavioural changes. Many of these changes are related to the need to locate a mate and subsequently, in the case of females, to switch from mate searching to oviposition behaviour. The prodigious reproductive capacity of the Mediterranean fruit fly, *Ceratitis capitata*, is one of the factors that has led to its success as an invasive pest species. To identify the molecular changes related to maturation and mating status in male and female medfly, a microarray-based gene expression approach was used to compare the head transcriptomes of sexually immature, mature virgin, and mated individuals. Attention was focused on the changes in abundance of transcripts related to reproduction, behaviour, sensory perception of chemical stimulus, and immune system processes. Broad transcriptional changes were recorded during female maturation, while post-mating transcriptional changes in females were, by contrast, modest. In male medfly, transcriptional changes were consistent both during maturation and as a consequence of mating. Of particular note was the lack of the mating-induced immune responses that have been recorded for *Drosophila melanogaster*, that may be due to the different reproductive strategies of these species. This study, in addition to increasing our understanding of the molecular machinery behind maturation and mating in the medfly, has identified important gene targets that might be useful in the future management of this pest.

DB-85

A Provisional Gene-Regulatory Pathway for the Control of Ovary Activation in Worker Honey Bees

*Mullen, Emma K.; Thompson, Graham J.
The University of Western Ontario, Biology, London, ON Canada N6A 3K7

The well-understood theory of kin selection explains how complex social behaviour can evolve at the gene level, yet there is controversy surrounding this theory because it does not identify which genes are involved in the expression of selfishness or altruism. The recent genome sequencing of the honey bee (*Apis mellifera)* and of other highly social organisms is creating unprecedented opportunities to identify genes important to social coordination. Microarray analyses have revealed hundreds of genes involved in reproductive regulation, but there has not yet been an attempt to link these genes into pathways that explain the expression of altruistic sterility in worker bees. I am using a knowledge-based gene network modeling program to reconstruct pathways that describe this form of altruism by means of ovary activation and de-activation. By linking gene lists derived from microarrays with gene-function

information, we have assembled a provisional network – the first of its kind. This hypothetical network is significant because it provides a testable hypothesis that describes the very genes controlling worker sterility, a quintessential trait at the core of social gene theory. We are currently probing this network for genes related to reproduction in social and non-social taxa, such as pheromone receptors, transcription factors and genes related to oogenesis. From this inclusive analysis, we will address one outstanding question in insect sociobiology: did complex social behaviour evolve from co-option of genes that regulate solitary life histories, or are new 'social' genes required?

DB-86

Developmental Time-Course of Expression in *Spodoptera frugiperda*

*Nègre, Nicolas; Gimenez, Sylvie; Legeai, Fabrice; Cousserans, François; d'Alençon, Emmanuelle; Fournier, Philippe
(NN, SG, FC, EdA, PF) Université de Montpellier 2, UMR1333 DGIMI, F-34095 Montpellier, France; (NN, SG, FC, EdA, PF) INRA, UMR1333 DGIMI, F-34095 Montpellier, France; (FL) IRISA, équipe Symbiose, Campus universitaire de Beaulieu, 35042 Rennes, France

The Lepidopteran *Spodoptera frugiperda*, also know as the Fall Army Worm (FAW), is a major corn, rice, sorghum and sugarcane crop eater. Genomic resources are currently being produced in order to understand what makes this polyphagous noctuidae shift into pest behaviour. In particular, the genome of a laboratory strain has been sequenced in 2010 and its assembly is in process. A reference transcriptome has also been generated based on a normalized 454 RNA-seq library from a mixed pool of developmental time-points. To refine the annotation of the genome, these NGS-based resources are combined with more classical data such as BAC and EST libraries. In parallel, we generated 6 developmental time-points (eggs, L2early, L2late, L3early, L3late and L6) transcriptomes by Illumina RNAseq. We used clustering approaches to detect groups of genes specific of each stages. A comparative approach with Drosophila homologs allowed the detection of FAW-specific developmental regulators.

DB-87

Comparative Genetics of Arthropod Salivary Gland Development

*Nguyen, Chilinh; Andrews, Emily; Le, Christy; Clemons, Anthony; Severson, David W.; and Duman-Scheel, Molly.
(CN, CL, AC, DWS, MDS) Eck Institute for Global Health and Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, 46556; (CN, EA, CL, DWS, MDS) Department of Medical and Molecular Genetics, Indiana University School of Medicine, South Bend, IN, 46617.

Understanding mosquito salivary gland development is critical given the importance of this tissue in blood feeding and pathogen transmission. A recent survey of the mosquito genomes suggests that mosquitoes have orthologs of many genes that regulate embryonic salivary gland development in *Drosophila melanogaster,* a well-characterized insect genetic model organism. Expression of a number of these orthologs was investigated in the dengue and yellow fever vector *Aedes aegypti,* an emerging model for vector mosquito development. These studies revealed that mosquito salivary gland development significantly differs from that of *D. melanogaster. Ae. aegypti* salivary glands, unlike their *Drosophila* counterparts, do not arise from invagination of salivary gland placode cells originating in the ventral neuroectoderm. In *Ae. aegypti,* the salivary gland develops as an outpocketing of the foregut. Despite this divergent early embryonic salivary gland development, orthologs of many *Drosophila* salivary gland development genes are expressed in the *Ae. aegypti* embryonic salivary gland. Thus, this investigation has revealed both divergent and conserved mechanisms of salivary gland development in dipteran insects. Future experiments will be directed at functional assessment of the roles of these genes in *Ae. aegypti,* as well as examination of the enhancer elements that regulate gene expression in the *Ae. aegypti* salivary gland. These studies may foster further comparative studies of salivary gland development in additional vector mosquito and insect species.

DB-88

Identifying transcriptional components of diapause preparation in the Asian tiger mosquito, *Aedes albopictus.*

*Poelchau, Monica F.; Reynolds, Julie A.; Denlinger, David L.; Elsik, Christine G.; Armbruster, Peter A.
(MFP, CGE, PAA) Georgetown University, Biology, Washington, DC 20057; (JAR, DLD) The Ohio State University, Entomology and Ecology, Evolutionary and Organismal Biology, Columbus, OH 43210

Many temperate insects survive the harsh conditions of winter by undergoing photoperiodic diapause, a pre-programmed developmental arrest initiated by short day lengths. Despite its well-established ecological significance, the molecular basis of photoperiodic diapause remains largely unresolved. The Asian tiger mosquito, *Aedes albopictus*, represents an outstanding emerging model to investigate the molecular basis of photoperiodic diapause in a well-defined ecological and evolutionary context. *Ae. albopictus* is currently considered the most invasive mosquito in the world, and traits related to diapause appear to be important factors contributing to the rapid spread of this mosquito. To discover transcripts involved in diapause preparation, we sequenced and assembled the transcriptomes of *Ae. albopictus* early- (3 d post-oviposition) and late- (6 d post-oviposition) stage embryos reared under diapause-inducing or non-diapause-inducing conditions. Illumina Hi-Seq yielded ~120 million paired-end reads, which we combined with a previous assembly of 454-GS FLX reads derived from *Ae. albopictus* oocytes. We identified 12,293 ESTs with high-confidence annotations to known Dipteran gene models**.** Analyses of differential expression revealed strong ontogenetic changes in gene expression between early- and late- stage embryos in both diapause-destined and non-diapause destined embryos. While not as prevalent, gene expression differences between diapause-destined and non-diapause-destined embryos were present in both developmental stages, and were more pronounced later in development. Genes and pathways related to morphogenesis, metabolism and stress response were among the most strongly differentially expressed between diapause- and non-diapause-destined embryos in both developmental stages. These results are consistent with studies from other diapausing insects. Our identification of differentially expressed transcripts related to diapause enriches the limited knowledge base for the molecular basis of insect diapause, in particular for the preparatory stage.

DB-89

A New Family of Cuticular Proteins (CPCFC) with Unusual Properties

*Reed, Tyler W.; Wu, Sheng-Cheng; Willis, Judith H.
University of Georgia, Department of Cellular Biology, Athens, GA 30602

BCNCP1, from *Blaberus craniifer* (Jensen *et al.,* 1997 IBMB27:109), an early addition to the now hundreds of cuticular protein sequences, turns out to have orthologs or paralogs in all orders of insects for which there are sequence data except the Hymenoptera. The presence of two paralogs in *Tribolium* and other beetles and some moths, makes it possible to elevate these sequences to family status, named CPCFC because it is a **C**uticular **P**rotein with two or three motifs of CPxxPxC, hence the two **C**ysteines are separated by **F**ive amino acids. Coleoptera and Lepidoptera have only two motifs in each protein; in all other orders, there is a single ortholog with three motifs. Each motif could well serve to hold metals and we have carried out an elemental analysis on cast cuticles (pupal exuvia and larval head capsules) of *Anopheles gambiae* that revealed the unexpected presence of several metals, e. g. Ca, Mg, Zn, Mn, Cu. We have expressed *AgamCPCFC1* in *E. coli*, which resulted in a brown-colored protein that becomes a sticky, glue-like aggregate at low salt concentration. The brown color of AgamCPCFC1 suggests its association with a metal. *In situ* hybridization with histological sections from *An. gambiae* revealed the mRNA is abundant in the epidermis over much of the body in larvae, pupae and pharate adults. Except for the CPAP1 and CPAP3 families (Jasrapuria *et al*., 2010 IBMB 40:214), cysteines are rarely found in cuticular proteins. Hence, we anticipate that the CPCFC family is serving some important and unique function in cuticle. We shall speculate on this function and describe how we plan to learn whether it is correct.

DB-90

Metabolic trajectories and gene expression during post-diapause development between E and Z pheromone strain of the European Corn Borer (*Ostrinia nubilalis)*

Wadsworth, Crista. B.*; Woods, William; Dopman, Erik A.
Tufts University, Biology, Medford, MA 02155

Speciation is the driving force behind evolution, and results in the vast varieties of species seen on this planet today. Using an invasive moth pest, the European Corn Borer (ECB), this study focuses on temporal isolation. Our interest is addressing the genetic mechanisms responsible for this reproductive barrier. In ECB there are two distinct phenotypes for post-diapause development time (PDD), which is the time it takes a diapausing caterpillar to pupate once they have received a specific diapause-breaking cue. For insects in upstate NY, PDD long strain ECB remain in PDD for 44 days and PDD short strain ECB remain in PDD for 14 days. This difference alters the life histories between the strains and results in different voltinism patterns seen in the field, with the PDD long strain being univoltine and the PDD short strain being bivoltine in upstate NY. Differential PDD acts as a prezygotic reproductive barrier and isolates gene flow by 65%. In this study we seek to, 1) use metabolism as a biomarker to isolate days in PDD that the two strains show major differences, 2) ask if the shapes of the metabolic trajectories can tell us about the mechanism controlling differential PDD, 3) use transcriptome sequencing to quantify differential expression of genes between the strains in a time course during PDD. Transcriptome sequencing will provide us with a list of candidate genes for PDD as well as be informative about what types of genes are involved in diapause maintenance and termination. Ultimately we hope to understand what genetic mechanisms are giving rise to differential PDD.

2012 ARTHROPOD GENOMICS SYMPOSIUM POSTERS
SORTED BY LAST NAME OF FIRST AUTHOR

Poster Number	Author(s)	Title
FG-49	Acevedo, Flor E.; Barrera, Carlos; Aluvihare, Channa; Harrell, Robert A.; *O'Brochta, David A.: Benavides, Pablo.	Transposon-based germ-line transformation of the Coffee Berry Borer, *Hypothenemushampei*: Opportunities for functional genomics.
FG-50	Agossa, Fiacre[1*], Christopher Jones[2]., Hilary Ranson[2]., Vincent Corbel[1]., Martin Akogbeto[1].	Investigating the Spread and the Mechanisms of DDT Resistance in M-form *Anopheles gambiae* from BENIN
GRE-40	Arensburger,P.*; R. H. Hice, J. A. Wright, N. L. Craig, P. W. Atkinson	Regulation of transposable elements in the *Aedesaegypti*mosquito genome by piRNAs
EG-1	Bennett, Anna K.; Elsik Christine G.	Improving the Honey Bee Consensus Gene Set
DB-80	Berens, Ali J.; Toth, Amy L.	Development of Next Generation Sequencing Resources to Study Behavior in Social Wasps: Applications to Caste Determination and Facial Recognition
FG-51	Bhere, Kanaka V.; Ayoub, Nadia A.; Hayashi, Cheryl Y.; *Garb, Jessica E.	Genomic organization and evolution of the vertebrate-specific venom neurotoxin α-latrotoxin from the Western black widow spider *Latrodectushesperus*
CG-18	Boore J.L.; Fuerstenberg S.I.	Whole Genome Evolutionary Analysis of the Monarch Butterfly Genome
CG-19	Camiletti, Alison L.; Percival-Smith, Anthony; Thompson, Graham J.	Honey Bee Queen Mandibular Pheromone Inhibits Ovary Development and Fecundity in a Fruit Fly
M	Carpi, Giovanna; Miller, Webb; Ratan, Aakrosh; Sarofeen, Christian M.; Harris, Robert S.; Drautz, Daniela I.; Rizzoli, Annapaola; Schuster, Stephan C.	Genome-wide diversity of the European tick *Ixodesricinus* and its associated microbiome.
GRE-41	Chaudhari, Sujata S; Arakane, Yasuyuki; Specht, Charles A; Moussian, Bernard; Kramer, Karl J; Beeman, Richard W and Muthukrishnan, Subbaratnam	Characterization and functional analysis of the *Knk* gene family in the red flour beetle, *Triboliumcastaneum*
DB-81	Chaw, R. Crystal; Patel, Nipam H.	Correlating gene expression with gastrulation behavior in *Parhyalehawaiensis*
CG-20	Childers, Christopher P.; Reese, Justin T.; Munoz-Torres, Monica C.; Bennett, Anna K., Elsik; Christine G.	Hymenoptera Genome Database: Resources for Comparing Hymenopteran Genomes
FG-52	Chu, Fu-Chyun; Lorenzen, Marcé D.	Development of RNAi-based Dominant Markers for use in Genetic Pest Management
EG-2	Cusson, M.; Maaroufi, H.; Nisole, A.; Boyle, B.; Laroche, J.; Béliveau, C.; Kukavica-Ibrulj, I.; Doucet, D.; Lucarotti, C.; Sperling, F.; Lumley, L.; Brunet, B.; Bird, H.; Nealis, V.; Levesque, R.C.	Sequencing the genome of the spruce budworm, *Choristoneurafumiferana*
DB-82	Davis, S.R.; Cartwright, P.; Engel, M.S.	Differential gene expression in the evolutionary development of the weevil rostrum
FG-53	Dunn, W. Augustine; James, Anthony; Xie, Xiaohui	A Galaxy-integrated framework using network graphs to exploit multidimensional data from disparate "Omics" sources for creating/exploiting functional genomics gene sets across multiple species.
CG-21	Frederick-Hudson, K. H.*, Corey Hudson, J. Schul	Molecular clock reveals extremely rapid diversification of a communication system
CG-22	Gilbert, Don	Perfect Arthropod Genes Constructed from Gigabases of RNA
FG-54	Gill, Sarjeet S.; Evans, Amy M; Qureshi, Nadia; Lee, Subum	TARGETED MUTAGENESIS IN *Aedesaegypti* AS A MEANS TO ELUCIDATE THE MODE OF ACTION OF BACTERIAL TOXINS.
FG-55	Glastad, Karl M.; Hunt, Brendan G.; Goodisman, Michael A. D.	DNA methylation signatures in termites
DB-83	Gleason, Jennifer M.; Zhou, Y.; Hackett; Jennifer L.; Harris, Bethany R.; Greenfield, Michael D.	Development of Genomic Resources for Studying Acoustic Communication in the Lesser Wax Moth, *Achroiagrisella*
DB-84	Gomulski, Ludvik M.; Dimopoulos George; Xi, Zhiyong; Scolari, Francesca; Gabrieli, Paolo; Siciliano, Paolo; Clarke, Anthony R.; Malacrida, Anna R.; Gasperi, Giuliano	Transcriptome profiling of sexual maturation and mating in the Mediterranean fruit fly, *Ceratitiscapitata*
FG-56	Gondhalekar, Ameya D.; Scharf, Michael E.	*De novo*transcriptome sequencing for unraveling the effects of insecticide selection in the German cockroach
CG-23	Hagen, Darren E.; Elsik, Christine G.	Analysis of small RNA data in Hymenoptera reveal lineage-specific microRNAs
EG-3	Hall, A. Brantley; Criscione, Frank; Qi, Yumin; Tu, Zhijian (Jake)	Discovery and characterization of the first mosquito Y chromosome genes
CG-24	Hanrahan, Shawn J.; Johnston, J. Spencer.	Genome size diversity in Arthropods.

Poster Number	Author(s)	Title
CG-25	Hansen, Allison K.; Moran, Nancy A.	tRNA expression and post-transcriptional modifications in reduced-genome endosymbionts (*Buchneraaphidicola*) of divergent aphid species
FG-57	Harrell, Robert A.; Aluvihare, C.; O'Brochta, David	The University of Maryland Insect Transformation Facility: A Resource for the Study of Insect Functional Genomics.
FG-58	Hepat, Rahul [1]; Song, Hojoon[2]; *Kim, Yonggyun[1]	A Viral Histone H4 Interacts with Nucleosome Components and Alters Target Insect Transcriptome Revealed by an IlluminaRNASeq
CG-26	Hunt, Brendan G.; Ometto, Lino; Keller, Laurent; Goodisman, Michael A. D.	Evolution at Two Levels in Fire Ants: The Relationship between Gene Expression and Protein Evolution
EG-4	Ioannidis, Panagiotis[1]; Lu, Yong[2]; Kumar, Nikhil[1]; Creasy, Todd[1]; Ott, Sandra[1]; Tallon, Luke J[1]; Pick, Leslie[2], Dunning Hotopp, Julie C.[1,*]	Transcriptomics of the invasive brown marmorated stink bug (*Halyomorphahalys*)
FG-59	Jasrapuria, Sinu; Li, Beibei; Chaudhari, Sujata S; Arakane, Yasuyuki; Moussian, Bernard; Mezendorfer, Hans; Specht, Charles A; Beeman, Richard W; Kramer, Karl J and Muthukrishnan, Subbaratnam	Functional Genomics of Two Families of Genes Encoding Cuticular Proteins Analogous to Peritrophins in *Triboliumcastaneum*
EG-5	Jiang, Zi-Feng *; Ghanim, Murad*; Xia, Fang-Fang; Stevens, Rick; White, Kevin P.	*De novo* assembly of three symbiont genomes of the Whitefly *Bemisiatabaci*
CG-27	Johnston, J. Spencer; Hanrahan, Shawn J. Texas A&M University, Entomology, College Station, TX 77843.	Estimating Arthropod Genome Size: Pitfalls and Practical Advice.
EG-6	Jouraku, Akiya; Yamamoto, Kimiko; Kuwazaki, Seigo; Urio, Masahiro; Suetsugu, Yoshitaka; Narukawa, Junko; Miyamoto, Kazuhisa; Kurita, Kanako; Kanamori, Hiroyuki; Katayose, Yuichi; Matsumoto, Takashi; Noda, Hiroaki	The development of the genomic database of the diamondback moth (*Plutellaxylostella*) and genomic analysis using it
M-47	Kapheim, Karen M.*; Yeoman, Carl J.; Wilson, Brenda A.; White, Bryan A; Robinson, Gene E.	Genomic Characterization of the Coevolutionary Process: Gut Microbiomes Associated with Honey Bee Social Caste
FG-60	Kelkenberg, Marco; Kim,Hee Shin; Chaudhari, Sujata S.; Kumari, Meera; Specht,Charles A.; Brown, Susan J.; Muthukrishnan,Subbaratnam and *Merzendorfer, Hans	A Comprehensive Study on the Effects of the Insect Growth Regulator Diflubenzuron in the Model Beetle Species *Triboliumcastaneum*
FG-61	Kolliopoulou, Anna; Iatrou, Kostas; *Swevers, Luc	Functional Genomics of dsRNA-Mediated Gene Silencing in Transfected Silkmoth-derived Bm5 Cells
CG-28	Kumar, Nikhil; Sieber, Karsten; Bromley, Robin; Flowers, Melissa; Ott, Sandra H.; Tallon, Luke J.; *Dunning Hotopp, Julie C.	Genomic Analysis of the Lateral Gene Transfer from a *Wolbachia*Endosymbiont to *Drosophila ananassae* through DNA Re-sequencing and Its Implications for Invertebrate Genomics
FG-62	Kumari, Meera; Merzendorfer, Hans; Specht, Charles A.; Beeman, Richard W.; Kramer, Karl J.; Muthukrishnan, Subbaratnam	A model system for the study of the mode of action of the chitin inhibitor, diflubenzuron
GRE-42	Kuwar, *Suyog S.; Pauchet, Yannick; Vogel, Heiko; Heckel, David G.	Annotation of serine protease superfamily and transcriptional response to a plant protease inhibitor in *Helicoverpaarmigera*
FG-63	Lang, Minglin*; Gorman, Maureen J.; Kanost, Michael R.	Multicopper Oxidase-1 Plays a Role in Iron Metabolism in *Drosophila melanogaster*
EG-8	Lawson, D.; Hughes, DST.;Megy, K.; VectorBase.	Developing a community annotation system for arthropod genomes
M-48	Lee, Carol Eunmi; Gelembiuk, Greg W.; Silva, Joana C.	*Exploring the microbial metagenomic diversity of the copepod host during habitat invasions*
EG-7	Lee, Ed; Helt, Gregg; Harris, Nomi; Buels, Robert; Childers, Christopher; Reese, Justin; Muñoz-Torres, Mónica C; Elsik, Christine G; Holmes, Ian; Lewis, Suzanna E.	WebApollo: A Web-based Sequence Annotation Editor for Distributed Community Annotation
EG-9	Lehtonen R.; Somervuo P.; Salmela L.; Paulin L.; Ahola V.; Koskinen P.; Saastamoinen M.; Duplouy A.; de Jong M.: Taipale M.; Mäkelä N.; Rastas P.; Wong S.; Kvist J.; Lawson D.; Taipale J.; Mäkelä V.; Ukkonen E.; Frilander M.; Holm L.; Auvinen P.; Hanski I.	From The Genome to Population Dynamics: The Glanville fritillary butterfly as a model system
CG-29	Leming, Matthew; Metoxin, Alex; Moon, Young; Whaley, Michelle; O'Tousa, Joe	Novel mechanisms of rhodopsin management in the mosquitoes *Aedesaegypti* and *Anopheles gambiae*
EG-10	Leotti, Fabricio C.; Cardoso, Gisele A.; Azeredo-Espin, Ana Maria, A. M. L.; *Torres, Tatiana, T.	New World screw-worm Genomics: a First Step Towards Understanding the Evolution of Parasitism in Calliphoridae
CG-30	McCoy, Rajiv C.*; Boggs, Carol L.; Petrov, Dmitri A.	*De novo* transcriptome assembly and SNP discovery in *Euphydryasgillettii*to test for balancing selection

Poster Number	Author(s)	Title
FG-64	Meyer, Jason M.; Ejendal, Karin F.K.; Conley, Jason M.; Avramova, Larisa V.; Garland-Kuntz, Elisabeth E.; Giraldo-Calderón, Gloria I.; Brust, Tarsis F.; Watts, Val J.; Hill, Catherine A.	A "Genome-to-Lead" Approach for Insecticide Discovery: Chemical Library Screening Reveals Antagonists of Vector Arthropod Dopamine Receptors with *In Vivo* Toxicity
DB-85	Mullen, Emma K.; Thompson, Graham J.	A Provisional Gene-Regulatory Pathway for the Control of Ovary Activation in Worker Honey Bees
EG-11	Munoz-Torres, Monica C.; Reese, Justin T.; Childers, Christopher P.; Bennett, Anna K.; Hagen, Darren E; Elsik, Christine G.	Annotating genomes and training the annotators: insights and lessons from Hymenoptera Genome Database.
CG-31	Munro, James B.; Orvis, Joshua; Posavi, Marijan; Borntrager, Martin; Lee, Carol E.; Silva, Joana C.	Genomic Trends in Aquatic Arthropods: Comparative Transcriptomics of the Calanoid Copepod *Eurytemoraaffinis* and the Water Flea *Daphnia pulex*
FG-65	Nagaraju, J.;Singh, J.; Singh, CP.; Badrinarayan, D.	MicroRNA-mediate host-pathogen interaction in *Bombyxmori*
DB-86	Nègre, Nicolas; Gimenez, Sylvie; Legeai, Fabrice; Cousserans, François; d'Alençon, Emmanuelle; Fournier, Philippe	Developmental Time-Course of Expression in *Spodopterafrugiperda*
DB-87	Nguyen, Chilinh; Andrews, Emily; Le, Christy; Clemons, Anthony; Severson, David W.; and Duman-Scheel, Molly.	Comparative Genetics of Arthropod Salivary Gland Development
CG-32	Nunes, Francis M. F.; Paschoal, Alexandre R.;Cristino, Alexandre S.; Freitas, Flávia C. P.; Simões, Zilá L. P.;Durham, Alan M.	Identification of hymenopteran non-protein-coding RNA candidates by comparative genomics
EG-12	Nuñez, Jonathan.; Hernández, Erick.; Giraldo, William.; Navarro, Lucio.; Gongora, Carmenza.; Cristancho Marco A.; Gaitán, Álvaro L.; *Benavides Pablo.	First Draft Genome Sequence of Coffee Berry Borer: the most invasive insect pest of coffee crops
FG	O'Brochta, David A.; Pilitt, Kristina L.; Harrell, Robert A.; Aluvihare, Channa; Alford, Robert T.	Transposon-based forward and reverse genetics in Anopheles mosquitoes.
FG-66	Park, D.; Kim, W.; Hong, K.; Jung, J. W.; Kim, J.; Kwon, H. W.*	Identification and expression analysis of putative genes related to innate immune pathways in the Asian honeybee, *Apiscerana*
DB-88	Poelchau, Monica F.; Reynolds, Julie A.; Denlinger, David L.; Elsik, Christine G.; Armbruster, Peter A.	Identifying transcriptional components of diapause preparation in the Asian tiger mosquito, *Aedesalbopictus.*
FG-67	Ramaseshadri, Partha*; Bolognesi, R.; Zhao, S. ; Wiggins, E. ; Clinton, W.; Heck, G.; Flannagan, R.; Segers, G.	Analyses of RNAi gene targets for the control of western corn rootworm
DB-89	Reed, Tyler W.; Wu, Sheng-Cheng; Willis, Judith H.	A New Family of Cuticular Proteins (CPCFC) with Unusual Properties
EG-13	Reese, Justin T.; Childers, Chris P.; Muñoz-Torres, Mónica C.; Bennett, Anna K.; Elsik, Christine G.	Practical Challenges of Organizing Community Annotation Resources
CG-33	Rhodes, Adelaide.	Improved Prediction of Ultraconserved Regions in Insect Genomes
EG-14	Robertson, Hugh M.; *Coates, Brad S.; Walden, Kimberly KO.; French, B.W; Miller, Nicholas M.; Yandell, Mark; Tagu, Denis; Siegfried, Blair D., Sappington, Thomas W.	A Draft genome assembly for the Western corn rootworm, *Diabroticavirgiferavirgifera*
GRE-43	Rueppell, Olav	Patterns of Conversation of Exceptional Recombination Rates in Honey Bees
CG-34	Sánchez-Gracia, Alejandro; Frias-López, Cristina;Guirao-Rico, Sara;Almeida, Francisca C.; Librado, Pablo;Vieira, Filipe G.;Arnedo, Miquel A.; Rozas, Julio	Identification of Potential Chemosensory Gene Families in Funnel Web Spiders using NGS
FG-68	Sethi, Amit*; Slack, Jeffrey; Kovaleva, Elena S.; Buchman, George W.; Scharf, Michael E.	The Termite Gut Bioreactor: Transcriptional and Translational Analysis of Lignin Degradation
FG-69	Shelby, Kent S.	IDENTIFICATION OF IMMUNE SYSTEM GENE SILENCING TARGETS IN A *DE NOVO* ASSEMBLY OF THE TRANSCRIPTOME OF THE SQUASH BUG, *ANASA TRISTIS* (DE GEER) (HETEROPTERA: COREIDAE)
CG-35	Shreve, Jacob T.; Subramanyam, Shubha; Johnson, Alisha J.; Williams, Christie E.; Shukle, Richard H.	A Genome-Wide Survey of RNA Interference Genes in a Galling Insect
FG-70	Silver, Kristopher*; Beeman, Richard; and Park, Yoonseong	Gene Expression in a *Triboliumcastaneum* cell line
FG-71	Sim, Cheolho*; Bai, Xiaodong; Denlinger, David L.	Identification of direct FOXO targets controlling diapause of the mosquito *Culexpipiens* by chromatin immunoprecipitation sequencing (ChIP-Seq).

Poster Number	Author(s)	Title
CG-36	Simon, S.; Narechania, A.; DeSalle, R.; Hadrys, H.	Comparative Analyses to Explore the Evolutionary Signal of Gene Partitions in the Insect Tree of Life
FG-72	Smith ER; Lin C; *Garrett AS; Thornton J; Mohaghegh N; Hu D; Jackson J; Saraf A; Swanson SK; Seidel C; Florens L; Washburn MP; Eissenberg JC; Shilatifard A	Gene Class Specificity of Transcription Elongation Factors in Drosophila
FG-73	Sobrinho Jr. Iderval S.; Campanini, Emeline B.; Andrade, Paulo H.M.; Rezende, Victor R.; Zuffo, Livia; Lima, Andre L.A.; Ribeiro, Daniel, T.; Nakamura, Aline M.; Chahad-Ehlers, Samira ; de Brito, Reinaldo A.*	Transcriptome analyses of larvae and pupae stages of *Anastrephafraterculus* (Diptera: Tephritidae) investigated by Illumina RNA-seq technology.
EG-15	Sparks, Michael E.; Gundersen-Rindal, Dawn E.	Transcriptome Analysis of *Lymantriadispar* Ld652Y Cells Suggests Substantial Viral-Related Expression and Activity.
EG-16	Stuart, Jeffery J.*; Chen, Ming-Shun; Brown, Susan J. Richards S. ; representing the Hessian Fly Genome Consortium.	The Genome of the Hessian Fly Reveals a Massive Expansion of Secreted Salivary Gland Proteins Necessary for Gall Formation and Larval Survival
GRE-44	Suetsugu, Y.: Futahashi, R.; Kanamori, H.; Kadono-Okuda, K.; Sasanuma, S.; Narukawa, J.; Ajimura, M.; Jouraku, A.; Namiki, N.; Shimomura, M.; Sezutsu, H.; Osanai-Futahashi, M.; Suzuki, M-G.; Daimon, T.; Shinoda, T.; Taniai, K.; Asaoka, K.; Niwa, R.; Kawaoka, S.; Katsuma6, T. Tamura1, H. Noda1, M. Kasahara7, S. Sugano8, Y. Suzuki8, H. Fujiwara4, H. Kataoka, H.; Arunkumar, KP.; Nagaraju, J.; Goldsmith, MR.; Feng, Q.; Xia, Q.; Yamamoto, K.; Shimada, T.; *K. Mita	Full-length cDNA Database of a Lepidopteran Model Insect, *Bombyxmori*, Based on over 11,000 cDNA Clones
EG-17	Suetsugu, Yoshitaka; Daimon, Takaaki; Jouraku, Akiya; Kobayashi, Tetsuya; Nakamura, Yuki; Kuwazaki, Seigo; Kayukawa, Takumi; Katayose, Yuichi; Kanamori, Hiroyuki; Kurita, Kayoko; Sanada-Morimura, Sachiyo; Matsumura, Masaya; Tanaka, Yoshiaki; Tanaka, Hiromitsu; Shiotsuki, Takahiro; Shinoda, Tetsuro; Yamamoto, Kimiko; Noda, Hiroaki	Progress in genome sequencing of the brown planthopper, *Nilaparvatalugens*
GRE	TAGU, D.*; LEGEAI, F.; JAQUIERY, J.; MIEUZET, L.; MAHEO, F. ; LETERME, N. ; BONHOMME, J. ; NOUHAUD, P.; RISPE, C.; LAROSE, C. GAGGIOTTI, O.; STOECKEL, S.; SIMON, J-C.	Genome-Wide Approaches for Identifying Genes Controlling Sexual and Asexual Reproduction in the Pea Aphid
GRE-45	Tao, Shu; Elsik, Christine G.	Computational Identification of Potential A-to-I RNA Editing Sites in *Apismellifera*
CG-37	Trout Fryxell, Rebecca ;Lanzaro, Gregory, Cornel, Anthony	Single nucleotide polymorphic analysis reveals sub-structuring within California *Culexpipiens*complex populations and phenotypic associations
FG-74	Vannini, Laura; Dunn, W. Augustine; Willis, Judith H.	Lessons from Illumina Data with *Anopheles gambiae*: Identification of pri-miRNA and Discovery of Substantial Differences with Microarray Results
DB-90	Wadsworth, Crista. B.*; Woods, William; Dopman, Erik A.	Metabolic trajectories and gene expression during post-diapause development between E and Z pheromone strain of the European Corn Borer (*Ostrinianubilalis)*
FG-75	Walk, Thomas C.; Geib, Scott M.	Expression Analysis of Multiple Libraries of Unreplicated RNA-seq Data
CG-38	Waterhouse, Robert M.; Zdobnov, Evgeny M.; Tick Genome Consortium.	Insights into intron evolution from the genome of the *Ixodes* tick.
FG-76	Wong, Swee Chong; Lehtonen, Rainer	Pipe-lines for design, quality control and analysis of SNP genotype data in non-model organisms
CG-39	Zdobnov, Evgeny M., Robert M. Waterhouse, and Evgenia V. Kriventseva	OrthoDB: The Catalog of Arthropod Orthologs
FG-77	Zhang, Shuguang; Hartson,Steve; Rogers, Janet; He, Yan; Jiang, Haobo	Mass spectrometry-based semi-quantitative analysis reveals dynamic changes in the proteome of *Manducasexta* larval plasma after immune challenge
GRE-46	Zhang, Xiufeng; Zheng, Yun; Jagadeeswaran, Guru; Sunkar, Ramanjulu; Jiang, Haobo	Identification of Conserved and Novel MicroRNAs in *Manducasexta*and Their Possible Roles in the Regulation of Immunity-related Gene Expression
FG-78	Zhao, Chaoyang; Navarro, Lucio; Katz,Andrew M.; Shukle,Richard H.; Richards, Stephen; Stuart,Jeffrey J.	Discovery of Avirulence Genes in the Hessian Fly