

THIRD ANNUAL

Arthropod Genomics

S Y M P O S I U M

Center for Genomic Studies on Arthropods Affecting Human, Animal and Plant Health • Kansas State University

June 11-14, 2009 • Kansas City

POSTER ABSTRACTS

1

Global Gene Expression in Honey Bees during Intense Flight and Senescence

*Ammons, Andrew D.; Elekonich, Michelle M.; Roberts, Stephen P.

University of Nevada, Las Vegas, School of Life Sciences, Las Vegas, NV 89154

A primary focus of research in ecological and evolutionary biology is the connection between behavioral development and the intensity and frequency of physiological mechanisms that influence aging and senescence. Current oxidative stress theories of aging implicate the accrual of reactive oxygen species (ROS) produced through aerobic respiration to chronic cellular damage. The accumulation of this damage and the breakdown of antioxidant defenses lead to aging. Intense metabolic activity has also been linked to the formation of ROS and eventual oxidative damage. Honey bees are an excellent model to test these theories, as they are a free-living organism whose behavior and flight (the most aerobically demanding activity in the animal kingdom) are amenable to manipulation. Utilizing oligonucleotide microarrays, we are studying patterns of global gene expression in the brains and thorax muscles (tissues known to undergo high oxidative stress) of age-matched bees that vary in their behavior (nursing and foraging) and flight experience. Some preliminary results indicate up-regulation of genes involved in muscle structure (myofilin, myosin alkali light chain), energy metabolism (isocitrate dehydrogenase, trehalose-6-phosphate synthase I), and heat shock (heat shock chaperones 70-4, 70-5) in the thorax muscles of foragers. By comparing RNA expression in tissues from bees at different points in the aging process, we can better understand the molecular networks that influence senescence and death – phenomena that affect all organisms, including humans.

2

Characterization of *Anopheles gambiae* SRPN6, an immune inducible midgut and salivary gland invasion marker

*An, Chunju; Michel, Kristin

Kansas State University, Division of Biology, Manhattan, KS 66506

Serpins play important roles in the outcome of malaria parasite infection in mosquito vectors. *Anopheles gambiae* SRPN6 is an immune inducible midgut and salivary gland invasion marker. SRPN6 knockdown increases the number of melanized parasites and delays parasite lysis. Moreover, SRPN6 knockdown increased the number of parasites present in salivary glands. However, the precise function of SRPN6 remains elusive. The aim of our current study is to biochemically characterize SRPN6 and identify its target protease *in vivo*. Recombinant SRPN6 with six histidine residues at the N-terminus was expressed in *E. coli* and purified using Ni-NTA and ion exchange chromatography. Purified recombinant SRPN6 exists as a monomer and has the tertiary structure with the mixture of alpha-helix and beta-sheet. The specific inhibitory activity assay indicated that recombinant SRPN6 could inhibit the amidase activity of Kallikrein from porcine pancreas *in vitro*. Furthermore, SRPN6 also significantly inhibited PAP1, a prophenoloxidase-activating protease from the lepidopteran *Manduca sexta*. Indeed, the addition of recombinant serpin to *Manduca* plasma, inhibited phenoloxidase activity significantly, further supporting a regulatory role of SRPN6 in the melanization immune response. We are currently taking two complementary approaches to identify the *in vivo* target protease of SRPN6. Firstly, we are testing whether SRPN6 can also inhibit the activity of CLIPB10, the *An. gambiae* ortholog of PAP1, *in vitro*. Secondly, using a polyclonal antibody against the C-terminal part of SRPN6, we will affinity-purify serpin-serine protease complexes from infected tissues and identify the protease components by mass spectrometry.

3

Functional Analysis of Genes Encoding UDP-N-Acetylglucosamine Pyrophosphorylases in the Red Flour Beetle, *Tribolium castaneum*

*Baguion, Marilyn C; Arakane, Yasuyuki; Kramer, Karl J; Beeman, Richard W ; Muthukrishnan, Subbaratnam (MCB) Department of Biology, Kutztown University of Pennsylvania, Kutztown, PA 19530, USA; (YA, KJK, SM) Department of Biochemistry, Kansas State University, 141 Chalmers Hall, Manhattan, KS 66506, USA; (KJK, RWB) Grain Marketing and Production Research Center, ARS-USDA, 1515 College Avenue, Manhattan, KS 66502, USA

UDP-N-acetylglucosamine pyrophosphorylase (Uap) catalyzes the formation of UDP-N-acetylglucosamine (UDP-GlcNAc). UDP-GlcNAc is a key intermediate for the synthesis of chitin, the glycosylation of proteins, and the formation of glycosylphosphatidylinositol (GPI)-linkers, which anchor proteins to the cell membrane. In the fruitfly, *Drosophila melanogaster*, the gene that encodes Uap has been called *mummy* (*mmy*) or *cystic* (*cyst*). Previous studies had shown that in *Drosophila*, the *mmy/cyst* gene is required for cuticle formation, and in morphogenesis in the central nervous system, as well as in tracheal tubulogenesis. In this study, we searched the genome of *Tribolium* and have identified two genes that share high sequence identity with the *Drosophila mmy/cyst* gene. We designated these two genes as *TcUap1* and *TcUap2*. Their specific functions in *Tribolium* have not yet been investigated. To determine their roles in *Tribolium* development, we studied the expression profiles of *TcUap1* and *TcUap2* using RT-PCR and carried out double-stranded RNA (dsRNA)-mediated RNA interference (RNAi)-based experiments. Expression profiles indicated that both genes are expressed in all developmental stages. However, the transcript level of *TcUap1* is significantly lower in the embryonic stage compared to that of *TcUap2*. When young larvae were injected with dsRNA for *TcUap1* (*dsTcUap1*), their development was arrested in the larval stage and they gradually shrank in size, became dehydrated and eventually died. When last instar larvae were injected with *dsTcUap1*, their development was arrested at the pharate adult stage with 100% mortality. When a mixture of young and last instar larvae were injected with *dsTcUap2*, both larval-larval and larval-pupal molts were inhibited and molting defects resulted in 100% mortality as well. Injection of *dsTcUap2* into pharate pupae or 0-1 day old pupae prevented pupal-adult eclosion. Analysis of transcript levels in the insects three days after dsRNA injection confirmed that the mRNA level of only the targeted gene was suppressed. Our results, therefore, strongly suggest that *TcUap1* is required for larval development as well as for pupal-adult eclosion, but not for larval-pupal molt. On the other hand, our data clearly indicate that *TcUap2* is critical for all molt types during development.

4

Functional analysis of Polycomb and Trithorax group chromatin factors in *Drosophila* eye development

*Baig, Jawaid Ahmad¹; Françoise Chanut²; Tom B. Kornberg³; Ansgar Klebes^{1,3}

1 Freie Universität Berlin, Institut für Biologie – Genetik Arnimallee 7, 14195 Berlin, Germany;

2 George Williams Hooper Foundation, UC San Francisco, CA 94143-0552; present address: Office of Career and Professional Development, UC San Francisco, CA 94143-0376

3 University of California, San Francisco, Dept. of Biophysics and Biochemistry, CA 94143, USA

Polycomb group (PcG) and Trithorax group (TrxG) genes are required for the stable maintenance of the expression domains of homeotic and other developmental genes. PcG factors are believed to epigenetically silence gene expression, while TrxG function is required for active transcription of target genes. In addition to their crucial function in early development PcG and TrxG genes also play essential roles in photoreceptor differentiation and eye development.

We are investigating the function of two related PcG genes, *Suppressor of zeste 2* (*Su(z)2*) and *Posterior sex combs* (*Psc*) and the TrxG member *osa* in eye development. Simultaneous removal of the partially redundant *Su(z)2* and *Psc* functions by mitotic recombination results in disruption of eye development and de-regulation of *wingless*. GAL4-mediated over-expression of *Su(z)2* generates a small eye phenotype that is characterized by retardation of the morphogenetic furrow and a decrease in the number of differentiating photoreceptor cells.

The over-expression of *Osa* also causes a small eye phenotype that appears to arise from a retardation in cell cycle progression. This phenotype can be modified by mutant *CyclinE* alleles and the two proteins can be co-immunoprecipitated. *Osa* has been characterized as a component of the Brahma protein complex, a *Drosophila* SWI/SNF chromatin remodeling complex that functions as a transcriptional co-activator. We discuss a dual function of *Osa* in transcriptional regulation and interaction with *CyclinE* to control cell cycle progression.

5

Identification of Genes Encoding for Substances Interfering with Blood Clotting in the Midgut of *Rhodnius prolixus*

Bandarra, R.¹; Tremonte, H.¹; Vieira, L.R.¹; Rego, C.¹; Algarve, L.¹; Neves, B.C.¹; Tanaka, A.S.²; *Braz, G. R. C.¹

Federal University of Rio de Janeiro, Biochemistry Department, Chemistry Institute, Rio de Janeiro, Brazil; (AST) Federal University of São Paulo, Biochemistry Department, Brazil

Rhodnius prolixus is an obligate hematophagous triatomine and a Chagas disease vector. ESTs from a cDNA library of *R. prolixus* midgut were analyzed and 17 sequences were found that have Kazal type serine protease inhibitors domains, which prevent blood clotting in this organism. An attempt to cluster these sequences using CAP3 returned 5 different contigs, regardless the criteria used for clusterization. Upon the release of *R. prolixus* genomic trace sequences we identified 19 putative exons, some of which presented in more than one contig. Some contigs found could be either the product of alternative splicing of one or more genes or the transcription product of different genes sharing local similarities, as it would be expected as the result of gene duplication followed by divergent evolution. Two approaches to clarify this hypothesis were used: bioinformatics and PCR cDNA amplification/sequencing. GenSeed was used as local assembler. This software uses blast to identify in a formatted database the genomic trace sequences the ones that match a selected exon, clusters them in a contig using CAP3 software and blasts the genomic contig obtained against the same data base. GenSeed performs several cycles of assembly/extension and eventually stops. Using this method two non overlapping genomic sequences covering our longer EST were found. In these genomic sequences our putative exons occur in the predicted order, but in two separated segments. None of this gene fragments contains exon 5. Unfortunately, using this approach, we could not confirm the genomic linkage between the initial and the final part of this contig. On the other hand, no alternative genomic sequence bearing the identified exons was found, suggesting that the four contigs that share local similarities are the product of alternative splicing, although gene duplication cannot be ruled out yet. Contig

5 does not share sequence similarities with the other four contigs and seems to be the processed transcript of a totally different gene. Using PCR and specific oligonucleotides we amplified several DNA fragments which sequences will confirm the findings above described. In other triatomines as *Triatoma infestans* and *Triatoma brasiliensis*, only one transcript having Kazal type domains was found. These data show that either *R. prolixus* has a more complex set of anticoagulants as compared to other bugs or the useful potential of these related insects as sources of anticoagulant drugs of medical interest has not been fully exploited yet.

6

Sex Determination in the red flour beetle *Tribolium castaneum*

Blattner, Ariane; Meier, Nicole; Held, Beat; Quinter, Christoph; Lorenzen, Marcé; *Bopp, Daniel (AB, NM, BH, CQ, DB) University of Zurich, Zoological Institute, Zurich 8057, Switzerland (ML) USDA, Agricultural Research Service, Manhattan, KS 66502

Studies in different Dipteran and Hymenopteran insects suggest that the *transformer* (*tra*) and *doublesex* (*dsx*) genes act as a conserved regulatory module in the pathway that determines the sexual fate of the zygote. Our research group extended this analysis to a Coleopteran insect, *Tribolium castaneum*, and uncovered several candidate genes in this beetle's genome using BLAST analysis. Gene silencing studies and analysis of insertional mutations provide evidence for a conserved sex-determining function of the *Tribolium dsx* homologue (*Tcdsx*). In genotypically male (XY) beetles homozygous for an insertional mutation in *Tcdsx* (*Tcdsx*⁶⁹), gonads are composed of a mix of testicular and ovarian tissues (ovotestes). In addition, as adults homozygous *Tcdsx*⁶⁹ "males" display partial or complete loss of male-specific structures such as the aedeagus and sex patch (a male-specific gland on the femur of the first leg pair).

Furthermore, homozygous *Tcdsx*⁶⁹ XY pupae differentiate female-like papillae structures in the genital region and produce basal levels of vitellogenin (typically only detected in females). No gross morphological abnormalities were found in genotypically female (XX) *Tcdsx*⁶⁹ homozygotes, however, the presence of skleroterized structures in the outer genitalia suggest a weak masculinizing effect of the insertion allele. In addition, such females have an egg-laying defect and are sterile. Silencing of *Tcdsx* by dsRNA injection during larval stages causes intersexual phenotypes similar to those observed in *Tcdsx* mutants. Based on these findings we propose that *Tcdsx* acts as a double switch that controls proper sexual dimorphic development in *Tribolium*. The *Tcdsx* gene expresses a small set of sex-specific splice variants which are homologous in sequence to the male and female variants of *dsx* in *Drosophila*. Sex-specific regulation of *Tcdsx* is achieved at the splicing level. BLAST analysis identified the *tra* homologue, *Tctra*, and the *transformer2* homologue, *Tctra2*, both of which act as upstream regulators of *Tcdsx* splicing. While *Tctra2* products are detected in XX and XY individuals, *Tctra* expresses functional products only in genotypically female animals (XX). Silencing of *Tctra* in XX animals leads to a shift from the female to the male mode of *Tcdsx* splicing and results in masculinization at the phenotypic level. In XY animals, *Tctra* silencing has no effects on *Tcdsx* splicing and no phenotypic consequences. These results are consistent with *Tctra* being the upstream component in the pathway acting as an ON/OFF switch that determines the sexual fate of the zygote by controlling the activity of the double switch *Tcdsx*. We propose that the *tra-dsx* module also operates in Coleopterans and hence represents an ancestral mechanism that transduces and implements the selected sexual fate in insects.

7

THE TRIA PROJECT: Mountain Pine Beetle System Genomics

Project co-leaders: Bohlmann, J. and Cooke, J.

Co-investigators: Aukema, B.; Breuil, C.; Colman, D.; Cooke, B.; Erbilgin, N.; Evenden, M.; Hamelin, R.¹; Holt, R.; Huber, D.; Jones, S.; Keeling, C.; Marra, M.; Murray, B.; *Sperling, F.

(JC, DC, NE, ME, FS) Univ. Alberta, Edmonton, AB; (BA, BC, RH¹) Canadian Forest Service, Natural Resources Canada; (RH², SJ, MM) Genome Sciences Centre, Vancouver, BC; (JB, CB, CK) Univ. British Columbia, Vancouver, BC; (DH, BM) Univ. Northern British Columbia, Prince George, BC;

Introduction: The Mountain Pine Beetle (MPB) has devastated over 15 million hectares of forest in western North America and there is potential for continued spread. At least three interacting biological components are involved — the MPB, beetle-associated fungi, and their pine hosts. MPB and fungal-associate attack overwhelms tree defenses, interrupts water and nutrient transport, and can lead to tree mortality. Genomic information is vital to understanding the eruption and spread of MPB as molecular interactions between these organisms play an important role.

Objectives: We are providing new, state of the art knowledge about MPB, their fungal associates, host pine species, interactions between these organisms, and how these interactions impact MPB population dynamics. We will use this information to equip forest managers and policy makers with better tools and management options to guide more effective management decisions.

Approach: This is a multidisciplinary, highly collaborative project involving researchers in genomics, population genetics, molecular biology, ecology, systematics, and mathematics. Our approach is to combine high-throughput expression profiling and genotyping with ecological assays.

Building Genomics Resources: Genomic resources provide the foundation for investigating molecular-scale interactions between organisms, and include whole genome and/or expressed gene sequences as well as molecular markers from MPB, fungal associates, and pine host species.

Population Genomics: Understanding of population structures within the existing outbreak provides the basis to examine both historical and current spread. An extensive, integrated hierarchical sampling of landscapes in Alberta and British Columbia was

conducted in 2007 and 2008, generating over 10,000 MPB, fungal-associate, and pine species samples. SNP and microsatellite markers are being developed for each of these taxa and used to genotype these samples. Analyses of these data will allow us to assess genetic substructure, relatedness of populations, species composition and distribution, adaptive variation and movement.

Functional Genomics: We are characterizing MPB genes that play putative roles in processes important for olfaction, overwintering and overcoming tree defenses, fungal associate genes that are putatively important in pathogenicity, and pine genes that may be important in the defense arsenal. Environmental conditions are hypothesized to influence pine defense against MPB/fungal-associate attack, colonization and subsequent reproductive success: to this end, we are examining pine genomic responses to water limitation and determining whether these genomic responses may affect MPB reproductive and behavioral traits.

Maintaining Healthy Forests: The MPB epidemic represents a serious economic and ecological threat to Canadian forests. Forest management strategies rely on the ability to predict MPB spread. This project is generating detailed information on the complex biological interactions between organisms and between organisms and their environment that occur in an MPB epidemic. These data will be used to generate genomic-enhanced risk models that can be used by foresters and policy makers to make informed forest management strategies.

8

Ectopic Activation of Wnt Signaling Represses Anterior Development

*Bolognesi, Renata; Posnien, Nico; Fischer, Tamara D.; Bucher, Gregor and Brown, Susan J.

(RB, TDF, SJB) Kansas State University, Division of Biology, Manhattan, KS, USA

(NP, GB) Johann-Friedrich-Blumenbach Institut, Department for Developmental Biology, Göttingen, Germany

Proper *Wnt* signalling is essential to normal development of virtually all metazoans. In short-germ arthropods, including the red flour beetle *Tribolium castaneum*, *Wnt* signalling is implicated in posterior patterning and germband elongation; depletion of positive pathway regulators produces embryos lacking abdominal segments (Bolognesi et al., 2008), a phenotype conserved with other arthropods and vertebrates, but not *Drosophila*. We used the *Tribolium* genome sequence to identify orthologs of *Wnt* pathway inhibitors. *Tc-axin*, a conserved negative pathway regulator, is expressed in the anterior region of blastoderm embryos. We then investigated the effects of *Wnt* ectopic activation by performing RNAi with *Tc-axin*. Anterior segments (labrum to T3) in the resulting embryos were highly reduced or absent. Domains of posterior determinants (*Tc-caudal* and *Tc-millipatte*) and *hox* genes (*Tc-AbdA* and *Tc-Dfd*) were shifted anteriorly in the blastoderm embryo, while the domains of *Wnt* gene expression expanded in the head. *Tc-axin* appears to be required to keep *Wnt* targets repressed to allow normal anterior development. This function is also conserved with vertebrates, indicating that proper *Wnt* signalling is an ancient requirement for normal development.

9

Evolutionary Analysis as the Basis for Interpreting, Comparing, and Presenting Genomes: The GATOR and PHRINGE System

*Boore, Jeffrey L.; Fuerstenberg, Susan I.

(JLB, SIF) Genome Project Solutions, Hercules, CA 94547 (JLB) University of California, Department of Integrative Biology, Berkeley, CA 94720

We anticipate a great acceleration in whole genome sequencing over the next few years. Current tools for interpreting, comparing, and presenting these data cannot handle the expected pace, lack integration, require extensive IT support and computational expertise, and do too little to facilitate biological discovery. In particular, the standard “browser” format is anachronistic, with the genome assembly, rather than the biological information, being the organizing principle. It requires great manual effort to identify any particular gene or biochemical pathway. Fortunately, two new developments are enabling a better approach. First, next generation sequencing technology allows very deep sequence coverage of the set of expressed genes. For example, 200-fold mean sequence coverage can be obtained on a typical transcriptome for about \$20,000. This means that genes can be modeled with much greater accuracy, so even the early stages of analysis can focus on biological discovery instead of manual gene curation. Second, we have developed an effective tool (“PHRINGE”, for Phylogenetic Resources for Interpreting Genomes) for assigning orthologous and paralogous relationships among genes based on phylogenetic analysis of complete gene sets. In the absence of biochemical characterization, the best inference of gene function is by inferring that orthologous genes retain the same function. This is incorporated into the “GATOR” (Genome Analysis Tools and Online Resources) system under development, a “gene-centric”, user-friendly, streamlined approach to genome interpretation, comparison, and presentation. The entry point is the gene catalog itself, sortable by many categories, including domain content, intracellular location, SNP content, biochemical pathway, protein characteristics, number of members in any gene family, and many others. Users can view evolutionary trees, gene colinearity maps, and links to protein structures for all genes in multiple sequenced genomes.

10

A candidate golden gene (slc24a5) from a lepidopteran, Bombyx mori by comparative genomics.

*Brown, Thomas M. Geneectar Com LLC, Whitefish, MT 59937

After the zebrafish golden gene was elucidated, it was reported that a SNP in SLC24A5, the human homologue, was associated with pigmentation of the human skin. The SNP was nearly fixed in Caucasians, while being nearly absent among Asians and Africans. Beginning with a similar sequence from *Apis mellifera*, a Unigene automatic BLAST search returned 100 sequences including many from arthropods; however, only Anaplura, Coleoptera, Diptera and Hymenoptera were represented among orders of insects. Searching public databases, a lepidopteran sequence was found and assembled from short contigs. Conserved Domain

searching revealed that this sequence is very similar to SLC24A5 in possessing two domains of pfam01699, Na_Ca_ex, Sodium/calcium exchanger protein. This assembled lepidopteran sequence will be presented as aligned to known and inferred slc24a5 sequences.

11

Comparative Genomics Software Packages at GMOD

*Cain, Scott; Clements, David; and the GMOD Consortium

(SC) Ontario Institute for Cancer Research, 101 College St, Toronto, Ontario, Canada; (DC) National Evolutionary Synthesis Center, 2024 W. Main Street, Suite A200, Durham, NC 27705-4667 USA

The study of comparative genomics has been increasing recently due to the explosion of inexpensive genome sequences, and the visualization software packages at the Generic Model Organism Database (GMOD) project reflects that interest in comparative genomics. There are at least five software packages for visualizing synteny at GMOD. Here we present each of those packages, along with suggestions for use cases as well as caveats for their use. Software packages presented here include CMap, GBrowse-syn, SynBrowse, SynView, and Sybil. All of these packages are free and open source. The GMOD project is supported by grants from the National Institutes of Health and the USDA Agricultural Research Service.

12

Functional Analysis of the *knickkopf* Gene Family in Organization of *Tribolium castaneum* Cuticle

*Chaudhari, Sujata S; Arakane, Yasuyuki; Boyle, Daniel L; Kramer, Karl J; Beeman, Richard W; Muthukrishnan, Subbaratnam (SC, YA, KJK, SM) Department of Biochemistry, Kansas State University, 141 Chalmers Hall, Manhattan, KS 66506, USA; (DB) Department of Biology, Kansas State University, 116 Akert Hall, Manhattan, KS 66506, USA; (KJK, RWB) Grain Marketing and Production Research Center, ARS-USDA, 1515 College Avenue, Manhattan, KS 66502, USA

The insect cuticle is composed mainly of chitin and proteins and is a protective barrier that shields insects against biological and mechanical stresses. Proteins associated with chitin metabolism and cuticle assembly are attractive targets for biopesticide development. Recent studies in *Drosophila (Dm)* have identified a gene, *knickkopf (knk)*, whose expression is important for tracheal tube expansion and cuticle organization. In the present study, we have identified three Knk-like genes in *Tribolium*, *TcKnk1*, *TcKnk2* and *TcKnk3* which are orthologs of *DmKnk1*, *DmKnk2* and *DmKnk3*, respectively. All of these genes were differentially expressed throughout the different developmental stages of the beetle, suggestive of crucial roles for TcKnk's in *Tribolium* development. All three genes are expressed in the carcass (whole body without gut) but not in gut tissue. RNAi of *TcKnk1* resulted in lethal phenotypes preventing larval-larval, larval-pupal and pupal-adult molts, whereas RNAi of *TcKnk2* and *TcKnk3* resulted in developmental arrest only at the pupal-adult molt with ~55% and 100% mortality, respectively. Interestingly, lethal phenotype was observed only with dsRNAs specific for the C-terminal region of *TcKnk3* but not its N-terminal region, suggestive of alternative sliced variants of *TcKnk3*. TEM analysis confirmed that *TcKnk1*-specific RNAi lethality is associated with defects in cuticle organization. Collectively, our results indicate that each Knk plays a different role in the synthesis or organization of chitin in the cuticle of *Tribolium*.

13

Transcriptome Response to a Soybean Cystatin in Cowpea Bruchid Midgut

Chi, Yong-Hun; Salzman, Ron A; Balfe, Susan; Ahn, Ji-Eun; Sun, Weilin; Pittendrigh, Barry; Murdock, Larry L; *Zhu-Salzman, Keyan

(YC, RAS, JA, KZS) Department of Entomology, Texas A&M University, College Station, TX 77845; (SB, WS, BP, LLM) Department of Entomology, Purdue University, W. Lafayette, IN 47907.

The insect digestive system is the first line of defense protecting cells and tissues of the body from a broad spectrum of toxins and anti-nutritional factors in its food. To gain insight into the nature and breadth of genes involved in adaptation to dietary challenge, a collection of 20,352 cDNAs were prepared from the midgut tissue of cowpea bruchid larvae (*Callosobruchus maculatus*) fed on regular diet and diets containing anti-nutritional compounds. Transcript responses of the larvae to dietary soybean cystatin (scN) were analysed using cDNA microarrays, followed by quantitative RT-PCR confirmation with selected genes. The midgut transcript profile of insects fed a sustained sublethal scN dose over the larval life was compared with that of insects treated with an acute high dose of scN for 24 hrs. A total of 1,756 scN-responsive cDNAs were sequenced; these clustered into 967 contigs, of which 653 were singletons. Many contigs (451) did not show homology with known genes, or had homology only with genes of unknown function in a BLAST search. The identified differentially-regulated sequences encoded proteins presumptively involved in metabolism, structure, development, signaling, defense and stress response. Expression patterns of some scN-responsive genes were consistent in each larval stage, while others exhibited developmental stage-specificity. Acute (24 hr), high level exposure to dietary scN caused altered expression of a set of genes partially overlapping with the transcript profile seen under chronic lower level exposure. Protein and carbohydrate hydrolases were generally up-regulated by scN while structural, defense and stress-related genes were largely down-regulated. These results show that insects actively mobilize genomic resources in the alimentary tract to mitigate the impact of a digestive protease inhibitor. The enhanced or restored digestibility that may result is possibly crucial for insect survival, yet may be bought at the cost of weakened response to other stresses.

14

Sex-specific Splicing of the *doublesex* Genes in Homometabolous Insects Reveals 300 Million Years of Evolution at the Bottom of the Insect Sex-determination Pathway

*Cho, Soochin; Brichacek, Angela C.; Zhang, Jianzhi

(SC and ACB) Creighton University, Department of Biology, Omaha, NE 68178; (JZ) University of Michigan, Department of Ecology and Evolutionary Biology, Ann Arbor, MI 48105

Sex-determination mechanisms vary greatly among taxa. It has been proposed that genetic sex-determination pathways evolve in reverse order from the final step in the pathway up to the first. Consistent with this hypothesis, *doublesex* (*dsx*), the most downstream gene in the *Drosophila* sex-determination cascade that determines most sexual phenotypes, also determines sex in other dipterans and the silk moth, while the upstream genes vary among these species. However, it is unknown when *dsx* was recruited to the sex-determination pathway during insect evolution. Furthermore, sex-specific splicing of *dsx*, by which *dsx* determines sex, is different in pattern and mechanism between the moth and fly, raising an interesting question of how these insects have kept the executor of sex determination while allowing flexibility in the means of execution. To address these questions, we here study the *dsx* gene of the honey bee *Apis mellifera*, a member of the most basal lineage of holometabolous insects. We report that honey bee *dsx* is sex-specifically spliced and that it produces both the fly-type and moth-type splicing forms, indicating that the use of different splicing forms of Dsx in controlling sexual differentiation was present in the common ancestor of holometabolous insects. Our data suggest that in ancestral holometabolous insects, the female Dsx form is the default and the male form is generated by suppressing the splicing of the female form. Thus, it is likely that the *dsx* splicing activator system in flies, where the male form is the default, arose during early dipteran evolution. We corroborate this idea by characterizing the sex-specific splicing patterns of *dsx* in the blackfly (*Simulium vittatum*), a basal dipteran species, and comparing it with other dipteran insects.

15

Single nucleotide polymorphism (SNP) markers for genetic and genomic analysis of Lepidoptera

*Coates, Brad S; Sumerford, Douglas V; Hellmich, Richard L; Bayles, Darrell O; Wanner, Kevin W; Robertson, Hugh M; Sappington, Thomas W

(BSC, DVS, RLH, TWS) USDA-ARS, Corn Insects & Crop Genetics Research Unit, Ames, IA, 50011; (DOB) USDA-ARS, Bacterial Diseases of Livestock, Ames, IA, 50010; (KWW) Department of Plant Sciences & Plant Pathology, Montana State University, Bozeman, MT, 59717; (HMR) Department of Entomology, University of Illinois, Champaign, IL, 61801

Species from the insect order Lepidoptera are pests of major agronomic crops and also include species with endangered survival due to diminishing habitat ranges. Single nucleotide polymorphisms (SNPs) are biallelic nucleotide changes for which locus-specific molecular genetic markers can be developed to genotype individuals within a species. Although only two alleles are present per locus, SNPs tend to be at high density within the genomes of Lepidoptera, and are readily predicted from expressed sequence tag (EST) resources. Herein we show that putative SNPs predicted from the alignment of individual EST reads from *Ostrinia nubilalis* midgut and antennal EST libraries are a source for the development of molecular genetic markers. Validation of putative SNPs was carried out by development of multiplex single base extension assays followed by polymorphism detection on the SEQUINOM MassARRAY®. Individual PCR multiplex reactions successfully detected changes at ³ 34 loci, and show a near 70% validation rate among SNPs that were tested. Application of validated SNP markers to *O. nubilalis* population genetic analyses and linkage mapping indicate that “null” allele frequencies are low compared to validated microsatellite markers. These experiments describe a pipeline for the development and application of SNP markers, and demonstrate their use in population and genomic analyses of *O. nubilalis*. Our prediction of putative SNPs within other lepidopteran EST libraries suggests that intraspecific polymorphism data is readily available within existing sequence resources, and indicate that analogous marker development efforts are possible.

16

Moving towards a *Diabrotica virgifera virgifera* genome sequence*Coates, Brad S; Sumerford, Douglas V; Miller, Nicholas J; Siegfried Blair D; Sappington, Thomas W
USDA-ARS, Corn Insects & Crop Genetics Research Unit, Ames, IA, 50011

The corn rootworm, *Diabrotica virgifera virgifera*, is a major crop pest in North America that recently has been introduced to Europe, and despite varied control efforts remains a casual factor of corn yield reductions. The species is of great interest due to observed adaptation to transgenic *Bacillus thuringiensis* toxin- and rotation-based control measures. The *Diabrotica* Genetics Consortium has initiated an effort to obtain a whole genome sequence assembly, and in preparation has directed the collection and analyses of genetic and genomic tools. These efforts are funded by a USDA-CREES, National Research Initiative (NRI), Arthropod and Nematode Genomic Tools Development grant “Genomic Resources to Accelerate Western Corn Rootworm Research”. Herein we describe experiments to estimate the haploid (1N) genome size, perform bacterial artificial chromosome (BAC) library construction, screening, and end sequencing, and development of microsatellite- and single nucleotide polymorphism (SNP)-based markers for the construction of a high density linkage map. Currently, ongoing sequence assemblies for full BAC full inserts will give insight into the difficulty repetitive elements will impose on genome assembly using current and future sequencing technologies.

17

The association of mobile elements with microsatellites in the genomes of Lepidoptera

*Coates, Brad S.; Hellmich, Richard L.; Sumerford, Douglas V;
USDA-ARS, Corn Insects & Crop Genetics Research Unit, Ames, IA, 50011

The insect order Lepidoptera contains numerous species that are pests of major agronomic crops in addition to endangered species that are suffering from diminishing habitat ranges. Despite their economic and ecological importance, species of Lepidoptera are not well characterized genetically due to difficulties encountered during the development of one of the most common molecular tools available, microsatellite or simple sequence repeat (SSR) markers. Two phenomena often arise during analyses of single locus microsatellite markers; 1) PCR co-amplification of > 1 locus this is observed in the presence of > 2 fragments following size fractionation, or 2) the absence of PCR-generated fragments due to non-amplification of “null” alleles that results in significant reductions in heterozygosity compared to Hardy-Weinberg expectations. In the following, we show that (TTTG)_n and (GTCY)_n repeats are propagated within the genomes of Lepidoptera by hitchhiking within transposable elements (TEs). Additionally, we predict that (AAAG)_n microsatellites are a genome target site for the insertion of TEs which introduces haplotype variation that can give rise to “null” alleles during subsequent PCR amplifications. The described TEs are short interspersed nuclear element (SINE) or *Helitron*-like. The LSCS3 SINE contains a putative RNA polymerase III promoter, a tRNA-like domain, and an internal (TTTG)_n repeat. *Helitron*-like TEs all show characteristic 5'-TC and 3'-CTRR termini, and insert between AT or TT dinucleotides without target site duplication. Conserved *Helitron* secondary structures [subterminal inverted repeats (SIRs), 5' inverted repeats (IRs), and 3' stem-loop (IR)] and internal (GTCY)_n repeats suggest common descent with the *Drosophila* intersperse nuclear element (DINE-1) class of *Helitrons* (*Helitron1_Dm*). These *Helitrons* are common within genomes of Lepidoptera and are at high copy number in the *Bombyx mori* genome assembly, which suggests ubiquitous presence within lepidopteran genomes will render associated microsatellite unusable for genetic analyses and **may** have implications in full genome assemblies that rely solely on short read technologies.

18

BEEWORM: Gene Expression in a Host-Parasite Interaction

*Colgan, Thomas J.; Blaxter, Mark L.; Sumner, Seirian S.; Brown, Mark J.F.
(TJC) Department of Zoology, School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland; (MLB) Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, EH9 3JT, UK; (SSS) Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK; (MJFB) School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, TW20 0EX, UK.

The host-parasite interaction represents one of the most dynamic systems in the biological world. Using both parasitological and molecular techniques, this project will examine the natural host-parasite interaction of the bumblebee, *Bombus terrestris* L. (Apidae), and the nematode, *Sphaerularia bombi* Dufour (Tylenchida), at the gene expression level. *S. bombi* is unique in that it infects overwintering queens, with one of its main pathological effects being the castration of the host. In addition, the parasite manipulates the immune system and behaviour of the host. Using molecular and transcriptomic techniques, the main aim of this study is to identify genes, both in the host and parasite that have an altered gene expression during a host-parasite interaction. The project will identify genes in the nematode that allow the parasite to avoid the host's immune system, castrate the host and manipulate the host's behaviour. In addition, the study will identify genes targeted in the host by the parasite. The results of the study will provide a novel insight into the interaction at the genetic level of both host and parasite.

19

Genomics of *Nosema ceranae*, an emergent pathogen of honey bee (*Apis mellifera*): clarifying microsporidian evolution and host interactions.

*Comman, R. Scott; Chen, Y.; Pettis, Jeffrey S.; Evans, Jay. D.
USDA-ARS Bee Research Laboratory, Beltsville, MD 20705.

Recent steep declines in honey bee health have severely impacted the beekeeping industry, presenting new risks for agricultural commodities that depend on insect pollination. These declines could reflect increased pressures from parasites and pathogens. The incidence of infection by the microsporidian *Nosema ceranae* has increased significantly in the past decade and is a possible cause of bee declines. Microsporidia are obligate intracellular parasites related to fungi that have highly reduced genomes, cell ultrastructure, and metabolic pathways. Here we describe a draft assembly (~7.9 MB) of the *N. ceranae* genome derived from pyrosequence data. The assembly was complicated by a strong AT bias (74%), frequent homopolymer runs, and a diversity of repetitive elements. Of 2,614 predicted protein-coding sequences, we estimate that 1,366 have homologs in the human pathogen *Encephalitozoon cuniculi*, the most closely related organism for which a genome sequence is available. Allowing for transposable element proteins and false gene predictions, the two species are very comparable in terms of total gene number (1,996 in *E. cuniculi*), synteny, and GO annotations. Nonetheless, a substantial fraction of *N. ceranae* genes lack clear homologs in *E. cuniculi* and vice versa, and we identified few conserved genes that are potentially specific to the phylum. This suggests that, despite their small genomes, gene turnover contributes substantially to the divergence of microsporidian lineages. Evolution is also rapid at the amino acid level. For example, two of the three major protein components of the polar tube □ a large coiled structure of mature spores that is essential for initiating infection and the major synapomorphy of the taxon □ align poorly across species. Comparative analyses also identified candidate regulatory elements associated with genes. For *N. ceranae* gene models with well-supported 5' annotations, a highly significant sense-strand motif occurred within 15 bases upstream of most start codons. We detected a version of this motif in *E. cuniculi* as well despite a very different background nucleotide composition (53% AT). These analyses provide a foundation for future functional and population genomic studies of this enigmatic pathogen.

20

Diuretic hormones in the red flour beetle *Tribolium castaneum* (Herbst)

*Cosme, Luciano V; Begum, Khurshida; Park, Yoonseong

Kansas State University, Entomology, 123 W. Waters Hall, Manhattan, KS 66506

Neuropeptides are diffusible signal molecules mediating vital physiological processes in metazoans. We have been interested in a group of neuropeptides in osmoregulatory neuroendocrine system which has been suggested as a possible target for development of new biopesticides. Since the genome sequence of the *T. castaneum* has recently been completed, we were able to identify three diuretic hormones on its genome, one calcitonin like (CT)-DH31 and two corticotropin releasing factor like (CRF) – DH37 and DH47. The CRF-like DHs (DH37 and DH47) are encoded by one gene which undergoes alternative splicing. We have studied expression patterns of the three transcripts encoding DHs by using semi-quantitative reverse transcription PCR (RT-PCR), *in situ* hybridization, and immunohistochemistry. The RT-PCRs for different developmental stages (egg, larvae, pupa, adult) and dissected tissue (central nervous system, gut, Malpighian tubules and carcass) revealed that DH31 is constitutively expressed during the development, except low expression in the embryo. The DH31 is expressed in all tissue examined (CNS, MT, HG, carcass and gut) and highly transcribed in the central nervous system (CNS). Agreement with RT-PCR results *in situ* hybridization of DH31 specific probe recognized cellular signal in CNS, mid gut and hindgut. The gene encoding CRF-like peptides DH37 and DH47 produces four different transcripts by using different transcription initiation sites and an alternative splicing. These transcripts found to be present throughout developmental stages except early embryo where lower expression was evident and in all tissues examined (CNS, MT, HG, carcass and gut) at least one isoform was detected. *In situ* hybridization and immunohistochemistry identified strong DH37 expressing cells in the brain, thoracic ganglia and abdominal ganglia as well as in gut. *In situ* hybridization of DH47 revealed strong reaction in six pairs of small anterior-lateral neurons in the larval and adult brain. Biological functions of these hormones were further investigated by using RNA interference (RNAi). Injections of dsRNA were made at the last larval instar and a numbers of different biological parameters were measured until 30 days after adult emergence. Significant defects in oviposition and egg development were observed in the RNAi of the genes encoding DHs.

21

What is the genome structure of the *Wolbachia* wPcurA2 strain infecting figitid parasitoids of fruit flies?

*Davies, Andrew P.; Shanker, Savita; Handler, Alfred M.; Shoemaker, DeWayne; Schetelig, Marc F.; Xavier, Nirmala; Sivinski, John

(APD, AMH, DDS, MFS, JS) United States Department of Agriculture – Agricultural Research Service, Center for Medical, Agricultural, and Veterinary Entomology, Gainesville, FL 32608; (SS) University of Florida, Interdisciplinary Center for Biotechnology Research – Genomics, Gainesville, FL 32610; (NX) University of Florida, Entomology and Nematology Department, Gainesville, FL 32611

The maternally inherited intracellular bacterium, *Wolbachia pipientis*, induces reproductive alterations in host insect populations such as male-killing, feminization, parthenogenesis induction (PI, or reproduction without males) and cytoplasmic (reproductive) incompatibility (CI). Phylogenomic comparisons between *Wolbachia* strains inducing different reproductive alternations in their hosts reveal many insights into the biology and evolution of *Wolbachia* in general. Analyses of three sequenced *Wolbachia* genomes including wBm mutualistic symbiont strain infecting the nematode *Brugia malayi* and two strains that cause CI in *Drosophila* hosts (wMel and wRi) show that *Wolbachia* are among the most highly recombining intracellular bacterial genomes studied to date. Current evidence suggests lateral gene transfer between CI *Wolbachia* and hosts has not occurred despite their close associations and detection of gene transfer between A- and B-supergroup *Wolbachia* strains that share the same host. In comparison to mechanisms by which *Wolbachia* induce CI, those which drive PI in some parasitoid hosts are relatively well understood, yet the genomes of PI *Wolbachia* strains remain largely unassembled. We intend to sequence the genome of the *Wolbachia* wPcurA2 strain (A-supergroup) that causes PI in figitid parasitoids currently being examined for biological control of tephritid fruit flies. Our goal is to help elucidate biological mechanisms that drive *Wolbachia* PI in hosts as well as *Wolbachia*'s obligate intracellular lifestyle in general.

22

Expressed Sequence Tags from the Reproductive Tract of *Anastrepha fraterculus*

*de Brito, RA.; Resende, AL; Kaminski, AC; Paula, FPP; da Silva, FH

Dept Genética e Evolução; Universidade Federal de São Carlos, São Carlos, Brazil. 13565-905.

We are celebrating 200 years of Darwin's birth, and the speciation process has continued to be one of the central puzzles to be unraveled in biology. One of the most important questions in evolutionary biology that still remains elusive relates to which forces and genes are involved in the speciation process. Albeit a great deal of knowledge has accrued in recent years due to more robust quantitative analytical methods, faster computers and a plethora of methods to identify and screen new genetic markers, we still know surprisingly little of which evolutionary changes lead to species differentiation. It has been suggested that most speciation processes occur as a by-product of the independent evolution in allopatric populations. Because it is difficult to distinguish between traits that directed the speciation process and others that differentiated after the split, the identification of such genes has been pretty elusive. Many comparative studies have indicated that sex-related molecular and morphological traits evolve faster than others, in part due to sexual selection. Studies on such genes have shown that many are subject to positive selection that has driven the large differences observed, but most of these studies have focused on male-expressed proteins. In our study we investigate a group of cryptic species of tephritid flies of the group *Anastrepha fraterculus* seeking genes that may be involved in these species differentiation. Many nuclear and morphological markers have failed to separate the species of this group, which leads to look for sexually selected genes, that may be evolving faster. In order to identify such genes, we created a cDNA library of adult female reproductive tissues of one of the species of this group, *Anastrepha fraterculus*, seeking genes that may be

preferentially expressed in females and may be also subject to positive selection and the counter-response to male proteins. We extracted total RNA from the reproductive tissue of a pool of 50 sexually mature females (10 days old) of *A. fraterculus*, and after quantification, we isolated mRNA using the PolyATract mRNA Isolation System. We prepared the cDNA and cloned the expressed RNA copies into a library using the CloneMiner cDNA Library Construction kit. This library has been screened by colony PCR and, PCR products were PEG 8000 purified and sequenced in a MegaBACE 1000. Over 500 clones have been sequenced to identify expressed genes in the female reproductive trait of *Anastrepha*. Sequences were edited for quality and sequences longer than 140 bases with Phred >20 were checked for Open Reading Frames (ORFs) in all six frames using the software ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf.html>). We also checked for the presence of a signal peptide with the software SignalP 3.0. Functional domains were identified using the CD-search online tool and the Conserved Domain Database and Search Service (CDD) v2.16 from NCBI. This strategy has allowed us to identify many proteins at least to a protein family either using the DNA sequences or the derived ORF aminoacid sequences using BLASTn or BLASTp, respectively. We aligned these identified products to homologues from other insects such as species of *Drosophila* and *Aedes* and *Apis*. Some of these aligned sequences, particularly some we believe might be targets for sexual selection, such as proteases, were analyzed in the software K-estimator to calculate dN/dS evolutionary rates, which indicates patterns of synonymous to non-synonymous substitutions in these regions. We have identified many standard house-keeping genes such as Tubulin, Actin, ATP synthase 6, mitochondrial Cytochrome Oxidase and at least 10 different ribosomal proteins. It is not surprising that ribosomal proteins account for almost 30 % of the clones sequenced so far, considering the amount of protein expression that takes place in the ovary. More importantly, we have also identified many tissue specific genes, such as chorion, and some trypsin-like proteases. Overall, we have related 27% of the proteins expressed to signal transduction, 16% to transport, 10% to binding, 7% to proteolysis, 6% to constitutive. About 24% of the proteins did not have any known function and about 10% are unknown. For 30 of the isolated proteins we investigated the ratio of dN/dS to investigate patterns of selection these proteins may be subject to by contrasting their sequence to sequences of these genes in *Drosophila melanogaster*. 28 of these showed a dN/dS below 1 (for most, well below 1), indicating a pattern of purifying selection. One (*vitellin Ab*) showed dN/dS close to 1, indicating a pattern similar to neutral evolution and one (CG6503) had a dN/dS of 2.5 which is indicative of positive selection occurring on this gene. We hope that by continuing to screen these genes we should be able to identify important genes that affect species differentiation in this group of flies.

Financial Support: FAPESP 2007/00500-2 to RAB, CAPES Pró-Equipamentos to RAB and FHS.

23

Study of Gene Expression in *Nicotiana tabacum* Induced by *Helicoverpa zea* Herbivory

DesRochers, Brittany L.*; Thomas, Vickie ; and Musser, Richard O.

Western Illinois University, Biological Sciences Department, Macomb, IL 61455

The relationship between plants and insects is incredibly intricate, involving complicated cascades of biochemical pathways that have a goal of resisting herbivory. Nicotine is an alkaloid synthesized in the roots of tobacco plants in response to insect damage and functions as a chemical defense compound. From the roots the nicotine is transported to the aerial parts of the plant through the xylem and finally accumulates in the vacuoles. Cross-talk between many hormone signaling pathways in plants has been shown to be important in signal transduction. In previous studies from our lab nicotine induction has been shown to be suppressed as a result of glucose oxidase in *Helicoverpa zea*, caterpillar saliva. In this experiment microarrays were utilized to study the gene expression in *Nicotiana tabacum* in relation to caterpillar saliva and herbivory. The tobacco plants were grown from seeds in an environmentally controlled growth chamber until 8 weeks. The 6th instar caterpillars, *Helicoverpa zea*, were surgically wounded for two different treatments: an ablated treatment, where the salivary glands were removed, or a mock treatment where the surgery was mimicked, but the salivary glands remained intact. Then the caterpillars were caged onto individual tobacco leaves and allowed to feed for 24 hours, periodically repositioning the cage to ensure optimal herbivory. The leaves were then harvested and the RNA was purified from each sample. Amplified RNA was formed from the template RNA strand and labeled with cyanine 3 or cyanine 5 dyes, which then was hybridized onto the Agilent tobacco oligo microarrays. The analysis of the arrays produced hundreds of significantly altered genes from the non-wounded control treatment in both the mock and ablated treatments. Significantly altered genes ($P < 0.05$) showed major trends in the data that corresponded to different plant physiological pathways such as defense, photosynthesis, metabolism, and stress. In particular, my results indicate that many widely known pathogenesis-related genes such as basic chitinase and beta-1,3 glucanase are activated 5 fold higher in the mock compared to the ablated treatment. These results suggest that the caterpillar oral elicitors impact the plant defense response.

24

Oocyte patterning in non-model insects: creating transcriptomes of the ovaries and embryos of two insect species using 454 sequencing

Ewen-Campen, Ben*; Alwes, Frederike; Wahl, Mary; Extavour, Cassandra G. M.

Harvard University, Department of Organismic and Evolutionary Biology, Cambridge, MA 02138

The body axes of many animals are initially patterned by the asymmetric distribution of maternal mRNAs and proteins in the oocyte. In *Drosophila*, these maternal factors are synthesized by the 15 nurse cells that connect to each oocyte within the ovary. However, insect ovaries fall into three anatomical categories, of which *Drosophila* represents only one. The telotrophic meroistic ovaries of the true bugs (Hemiptera), for example, possess a single group of syncytial nurse cells that connect to all oocytes simultaneously via extended nutritive tubes. The panoistic ovaries of basal insects including crickets (Orthoptera) do not possess nurse cells. Little is known about the production or localization of maternal factors within the developing oocytes of insects possessing these different ovary types.

To address this issue, we have begun to characterize the ovarian and embryonic transcriptomes of a Hemipteran (the milkweed bug, *Oncopeltus fasciatus*) and an Orthopteran (the cricket *Gryllus bimaculatus*) using 454 pyrosequencing. Both of these species have rich histories as laboratory species, and both possess large ovaries that facilitate live imaging and microinjection. Our

sequence data will allow us to identify orthologs of candidate genes involved in a variety of embryonic patterning processes, and will lay the groundwork for future studies on the development and evolution of insect body plans. Additionally, these sequences will augment the relatively few genomic resources available for hemimetabolous insects, thereby contributing to our understanding of genomic evolution during the 330 million years since the divergence of Hemimetabola and Holometabola from their last common ancestor.

25

Will Your Favorite Genome be Sequenced?

*Fuerstenberg, Susan I; Boore, Jeffrey L.

(SIF, JLB) Genome Project Solutions, Hercules, CA 94547 (JLB) University of California, Department of Integrative Biology, Berkeley, CA 94720

The past decade has seen a fantastic acceleration in our ability to sequence complete genomes. We anticipate that this exponential increase will continue over the next few years. Will your favorite genome be among those sequenced? We review the several technologies that are state-of-the-art for genome sequencing, describe their capabilities, provide an interpretation of the pros and cons of each, and argue for the most appropriate mix of techniques for whole genome sequencing. We describe the factors important in the choice of genomes to target, the various mechanisms of funding and coordinating such a large project, and the options for follow-on analyses that will maximize biological discovery and the utility of the work to the broader scientific community.

26

The power of multiple genomes in *Nasonia* – Genes for species differences and speciation

*Gadau, Jürgen; Niehuis, Oliver; Gibson, Joshua D.; Joachim Ruther; Thomas Schmitt; Jan Buellesbach, Juergen Liebig

(JDG, JG, ON, JL) Arizona State University, School of Life Sciences, Tempe, AZ 85287;

(TS,JB) Universitaet Freiburg, Biology I, Department of Evolutionary Biology and Animal Ecology, University of Freiburg, Hauptstr.

1, 79104 Freiburg, Germany; (JR) Institute of Biology, Applied Zoology/Animal Ecology, Free University of Berlin, Haderslebener Str. 9, 12163 Berlin, Germany

Three closely related species of the parasitoid wasp genus *Nasonia* have been sequenced. All three species can be interbred once they are cured of their endosymbionts (*Wolbachia*), which normally cause bidirectional incompatibility. Virgin hybrid females will produce haploid males that allow efficient mapping and positional cloning of genes for traits that differ between species and/or cause pre- or postzygotic isolation ("speciation genes"). Female traits can be studied with equal efficiency using clonal sibships, i.e. females (sisters) generated by backcrossing F2 hybrid males to either parental line.

We present a comparison of the genetic architecture underlying species differences in morphological (wing size; body size), physiological (cuticular hydrocarbons, male sex pheromones) and behavioral (male courtship behavior) traits. We also demonstrate how we quickly identified the gene for a potential prezygotic hybridization barrier (male sex pheromone difference) due to the availability of the genome sequences.

Additionally, we used a combination of mapping and candidate gene approaches to identify genes responsible for the observed nuclear-cytoplasmic incompatibility resulting in partial postzygotic isolation between *N. vitripennis* and *N. giraulti*. (see also Poster by Gibson et al.).

These examples show that with the availability of the genome sequence *Nasonia* has become a powerful model organism to understand the genetic architecture and basis of species differences and speciation.

27

Signatures of positive selection in nuclear encoded genes of the oxidative phosphorylation pathway in holometabolic insects

*Gibson, Joshua D.; Niehuis, Oliver; Gadau, Jürgen

Arizona State University, School of Life Sciences, Tempe, AZ 85287

Almost all eukaryotes utilize the oxidative phosphorylation (OXPHOS) pathway in their mitochondria to produce adenosine 5'-triphosphate (ATP), the principal source of cellular energy. The pathway's function thereby critically depends on the tight interaction of genes encoded in both the nuclear and the mitochondrial genome. The mitochondrial genome accumulates substitutions at a higher rate than the nuclear genome. However, the ability of selection to act on mitochondrial genes is significantly reduced relative to nuclear genes because of the non-recombining nature of the mitochondrial genome. It has therefore been hypothesized that nuclear encoded genes of the OXPHOS pathway are under strong selective pressure to compensate for the accumulation of deleterious mutations in mitochondrial encoded OXPHOS genes; a process known as compensatory co-adaptation.

We tested this hypothesis by analyzing nuclear encoded OXPHOS genes for signatures of positive selection (dN/dS ratio) as well as evolutionary constraints (based on physiochemical properties) at amino acid sites. We considered OXPHOS gene sequences of holometabolic insects from the mitocomp database and the orthologs from three *Nasonia* parasitoid wasps that we annotated using their recently sequenced genomes. For at least four of the 59 studied nuclear encoded OXPHOS genes, we found evidence for positive selection. Three of the four genes further contain amino acid replacements at evolutionarily constrained sites between the *Nasonia* species. These genes could therefore account for the earlier reported disruption of the OXPHOS pathway in *Nasonia* hybrids in which they have to interact with mitochondrial encoded genes with which they did not co-evolve. Our results are consistent with the hypothesis of compensatory co-adaptation of the nuclear and mitochondrial genes of the OXPHOS pathway.

28

Aphid and Water flea have a High Rate of Gene Duplications Compared to Other Arthropods

Gilbert, Don*; Pfrender, Mike; in collaboration with Daphnia and Aphid genome consortia
(DG) Biology Dept., Indiana U., Bloomington, IN 47405, gilbertd@indiana.edu; (MP) Biology Dept., Utah State U., Logan, UT 8432, pfrender@biology.usu.edu

Many new genomes are being deciphered with the advent of rapid, low cost next generation sequencing. New genome discoveries are facilitated with ready access for biologists to informatics tools on shared cyberinfrastructure. Genome informatics tools have been integrated and are provided publicly for genome scientists on the US TeraGrid (http://gmod.org/Genome_grid). Arthropod genomes provide examples of rapid genome discovery using this, with interesting results from comparing aphid and water flea genomes.

Comparison of gene orthology and paralogy groups among thirteen arthropod genomes (<http://insects.eugenics.org/arthropods/>) finds the pea aphid *Acyrtosiphon pisum* and the water flea *Daphnia pulex*, both cyclical asexuals, have four times the gene duplications of other arthropods. These duplicates are not from whole genome duplication, and artifacts do not account for this.

This observation suggests a new phenomenon of interest to ecological, evolutionary and biomedical genomics. A catalog of overabundant, annotated genes indicates that both species have independently acquired extra genes for mitosis and chromosomal maintenance. Speculation on reasons for this expansion include mitosis-related genes have evolved with asexuality. High phenotypic plasticity and rapid environmental responses are exhibited by both species, and likely are facilitated by the many gene duplications. There are apparent effects of gene conversion, differentiation and gene dosage in the preservation of these duplicate genes.

29

Sequencing the genome of the cotton bollworm, *Helicoverpa armigera*: a model for the evolution of complex ecological traits.

Gordon, Karl H.J.; Oakeshott John G.; Batterham, Philip; Richards, Stephen; Gibb, Richard A.; Miller, Jason R.; East, Peter D.; Edwards, Owain R.*; Robin, G. Charles; Heckel, David G.; Vogel, Heiko; Feyereisen, René; D'Alencon, Emmanuelle; Sezutsu, Hideki; Legeai, Fabrice; Gagneur, Cyril C.; Gimenez, Sylvie; Volkoff, Anne-Nathalie; Fournier, Philippe; Brun-Barale, Alexandra; Suraporn, Siripuk S.; Shimomura, Michihiko; Mita, Kazuei; Scherer, Steven E.; Qu, Carson; Worley Kim C.
(KHJ, JGO, PDE) CSIRO Entomology, Canberra ACT 2601 Australia; (PB, GCR) Department of Genetics, University of Melbourne, Parkville VIC 3010 Australia; (SR, RAG, SES, CQ, KCW) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030; (JRM) J. Craig Venter Institute, Rockville, MD 20850; (ORE) CSIRO Entomology, Floreat, WA 6014 Australia; (DGH, HV) Max Planck Institute for Chemical Ecology, D-07745 Jena, Germany; (RF, AB-B, SSS) INRA-Sophia, 06903 Sophia-Antipolis, France; (FL) INRA-UMR-Bio3P, 35653 LeRheu, France; (ED'A, PF, CCG, SG, A-NV) INRA-AgroM, 34060 Montpellier, France; (HS, MS, KM) National Institute of Agrobiological Sciences, Tsukuba, Japan.

There is great interest in the genetic basis for the evolution of ecologically relevant complex traits. Genomic studies of closely related animal species that diverge significantly in their adaptations to the environment are likely to overcome our lack of understanding of such traits. Herbivorous insects provide ideal model systems due to their ecological diversity and limited genomes. In the Lepidoptera, the noctuid subfamily Heliiothinae includes host plant specialists and generalists, including *Helicoverpa armigera*, the world's most significant pest of agriculture. It has evolved resistance to every class of pesticide used against it, including the biopesticide Bt. The complete genome (400Mb) of this insect is now being sequenced using next generation technologies, and related species will follow. The sequenced genomes of *H. armigera* and its relatives will provide a platform addressing whether and how evolutionary adaptation has affected particular gene families, as well as to discover new genes that are targets of evolutionary adaptation. Comparison of genome organisation between noctuid insects and the lepidopteran model silkworm has shown that within a basic framework of microsynteny, pest genomes have extensive duplication of genes encoding detoxification enzymes such as the cytochrome P450s. *H. armigera* will be the first metazoan genome assembled entirely using short sequencing reads. We are investigating *de novo* assembly using several programs currently under active development. These include versions of the Newbler program (Roche) and the newest implementation of the Celera Assembler provided by the J.C. Venter Institute. The genome is also being assembled using the BCM HGSC's Atlas genome assembly suite. Our data is therefore proving invaluable as a test set for software engineers at Roche and the JCVI in achieving algorithms designed for short-read assembly to replace labour-intensive approaches such as that in the Atlas pipeline. Currently the project has generated ~ 16x coverage of the ~400Mb genome, comprising 7x coverage using 454 XLR (Titanium) technology and 6x coverage using XLR 20kb paired-end reads, with the remainder on the 454 FLX platform. Insights into assembly issues gained from this data set will be presented. Further sequencing, both fragment and paired-end, is now underway. For *Helicoverpa*, there are particular challenges due to the extensive dispersed repeats of varying types and lengths in the genome, as well as the high level of polymorphism that is a consequence of the insect's resistance to inbreeding.

30

Jasmonate-mediated Plant Defense Responses of *Nicotiana attenuata* Influence the Transcriptomes of a Specialist and Generalist Lepidopteran Herbivore Differently

Govind, Geetha*; Griebel, Thasso; Mittapalli, Omprakash; Allmann, Silke; Böcker, Sebastian and Baldwin, Ian T. (GG, OM, SA, ITB) Max Planck Institute for Chemical Ecology, Department of Molecular Ecology, Hans Knöll Straße 8, 07745 Jena, Germany; (TG, SB) Faculty of Mathematics and Computer Science, Friedrich-Schiller-University Jena, 07743 Jena, Germany

To understand how specialist and generalist lepidopteran herbivores transcriptionally respond to the defenses of their host plant, we created microarrays of normalized cDNA libraries of neonate *Manduca sexta* (*Ms*) and *Heliothis virescens* (*Hv*) larvae 24h after they had initiated feeding on their native host plant *Nicotiana attenuata*. These microarrays were then used to compare the transcriptional responses of larvae feeding on three different transformed *N. attenuata* genotypes, silenced in jasmonate (JA) signaling, nicotine production and protease inhibitor (PI) and nicotine (N) production, with those of larvae feeding on WT plants. The polyphagous *Hv* larvae regulated 16-times more genes than did the oligophagous *Ms* larvae when feeding on these defenseless plants, but both larval species regulated more genes in proportion to the number of different defenses that were silenced in the plants (i.e. JA>PI/N>N). While *Hv* larvae regulated more genes, the pattern of regulation did not differ dramatically amongst the different larval diets. In contrast, *Ms* larvae had down-regulated most of its genes, principally on JA-deficient plants with 1.9 and 4.5 fold changes in primary and secondary metabolism, 2.5 and 4 fold changes in peptidases and hydrolases in comparison to larvae reared on N or PI/N deficient diets. *Ms* larvae also down-regulated specific group of transporters, cellular processes and environmental information processing, but up-regulated genes involved in transcription, translation, maintenance of protein turnover and omega peptidases when the diet was deficient in JA mediated defense. Hence while the generalist regulated more genes in a diet-independent manner, the specialist regulated fewer genes but in a highly diet-specific manner.

31

Transcriptional evidence of microRNAs in the jewel wasp *Nasonia vitripennis*

*Hagen, Darren E.; Evans, Jay D.; Werren, John H.; Elsik, Christine G. (DEH, CGE) Department of Biology, Georgetown University, Washington DC 20057 (JDE) Bee Research Lab, USDA-ARS Bldg. 476, BARC-East Beltsville, MD 20705 (JHW) Department of Biology, University of Rochester, Rochester, New York 14627

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. We here describe the characterization of a set of putative miRNAs in the parasitic wasp *Nasonia vitripennis* using 454 sequencing technology in combination with data emerging as a result of the ongoing whole genome sequencing project. A total of 2,490 RNA molecules (19–24 nucleotides in length) were successfully mapped to 2,531 unique loci in the *N. vitripennis* genome, allowing us to identify 30 putative miRNAs conserved in other insects. Our analysis indicates that approximately 1,600 additional sequences fall within regions of the genome that are capable of forming stable hairpin structures if transcribed. Interestingly as many as 1,300 of these loci are found in areas of the genome known to contain repetitive DNA sequences.

32

Gene Families Encoding Proteins With Chitin-Binding Domains in *Tribolium castaneum*: Domain Organization, Phylogenetic Analysis and Expression During Development

*Jasrapuria, S.; Arakane, Y.; Beeman, R.W.; Kramer, K.J.; Muthukrishnan, S. (SJ, YA, SM, KJK) Kansas State University, Department of Biochemistry, Manhattan, KS, 66506; (RWB) Grain Marketing and Production Research Center, ARS-USDA, Manhattan, KS, 66502.

This study is focused on the characterization and expression of genes encoding proteins with cysteine-containing chitin-binding domains (CBD) belonging to the Peritrophin-A subfamily (CBM_14 = Pfam 01607 or SMART family 00494), in the red flour beetle, *Tribolium castaneum*. An exhaustive bioinformatics search of the genome of *Tribolium castaneum* queried with CBDs identified in previously characterized “peritrophins” and other insect cuticular proteins from several orders verified 27 genes encoding proteins with one to fourteen CBDs. Using phylogenetic analyses, these proteins were classified into two large families. One-group Peritrophic Matrix Proteins (PMPs) included eight proteins with multiple CBDs closely related to those found in peritrophins associated with the peritrophic matrix. The other larger group Cuticular Protein Analogous to Peritrophins (CPAPs) included several proteins similar in organization to the proteins encoded by the *gasp* gene in *Drosophila* (also known as *obtsructor*), which has three CBDs. This group includes a large assortment of diverse proteins each with only one CBD. In addition, several enzymes of chitin metabolism and several unidentified proteins with no other identifiable domain were present in this group. By comparing the gene sequences with available ESTs and/or cloned cDNAs, corresponding to these two groups of genes, their exon-intron organization as well as domain organization and sequence homologies of the encoded proteins were determined. Analysis of the expression of all members of these two groups of genes during *Tribolium* development and in different tissues provided experimental support for evolutionary divergence and specialization in the functions of these two groups of CBD-containing proteins. Members of the first group, which we denote as peritrophin-like, genes exhibit midgut-specific expression restricted to the actively feeding larval and adult stages. Members of the second group of genes are expressed predominantly in the carcass (whole body without the gut), throughout all of the developmental stages.

33

The Pyrokinin/PBAN-receptor Family of GPCRs and Peptide Ligands in Arthropoda: Comparison with Neuromedin U in Vertebrates.

*Jurenka, Russell; Rafaeli, Ada; Stern, Peter

(RJ) Iowa State University, Department of Entomology, Ames, IA 50011-3222; (AR) Institute for Technology and Storage of Agricultural Products, The Volcani Center, ARO, Bet Dagan 50250 Israel; (PS) Chemical Physics Department, Weizmann Institute of Science, 76100 Rehovot, Israel

The superfamily of G-protein coupled receptors (GPCRs) is one of the largest families of proteins found in all animals. Several subfamilies are known with the rhodopsin family being the largest. The recent crystal structures of rhodopsin and several other GPCRs have been reported, which greatly aids in the functional characterization of these membrane bound proteins. The GPCR family of receptors is characterized as having seven transmembrane domains with an extracellular and intracellular domain. The pyrokinin/PBAN family of receptors and ligands found in arthropods is characterized by sequence similarity with the neuromedin U receptors and ligands found in vertebrates. Neuromedin U is a peptide that is involved in a variety of physiological functions including feeding behaviors. It is found in the central nervous system as well as peripheral in a variety of tissues including the digestive tract. The pyrokinin/PBANs were first identified in insects as stimulating gut muscle contraction and later to stimulate pheromone biosynthesis in moths. Ligands and receptors have been found in all insects so far investigated based on annotated genomic sequences compared with known identified sequences. The peptide ligands are characterized by having an FXPRamide C-terminal ending that is critical for receptor activation (X = S, T, or G). Neuromedin U has an FXPRamide C-terminal ending. We demonstrated that these peptides are cross-reactive. We compared the peptide structure of these ligands across Insecta with those of vertebrates. Little similarity exists except in the C-terminal ending. The receptors on the other hand exhibit considerable sequence similarity indicating a conserved primary and secondary structure. Receptors and ligands have been found in 12 *Drosophila* species, 8 moth species, 3 mosquito species, *Tribolium castaneum*, *Apis mellifera*, *Pediculus humanus*, *Nasonia vitripennis*, *Acyrtosiphon pisum*, *Dermacentor variabilis*, and *Ixodes scapularis*. The structures will be illustrated with molecular models based on the rhodopsin model built at 2.8 angstrom resolution. The next closely related GPCR found in insects is the PRVamide-receptor. Interestingly the peptide ligand is produced by a gene that also produces an FXPRamide peptide. The PRVamide-receptor's closest homologue in vertebrates is also the neuromedin U receptor. This poster will illustrate the conserved nature of this family of GPCRs and peptide ligands across arthropods and vertebrates.

34

A Novel Aminopeptidase-like Gene Possibly Involved in Cry1Ab Toxicity and/or Resistance in European Corn Borer (*Ostrinia nubilalis*)

*Khajuria, Chitvan; Buschman, Lawrent L.; Chen, Ming-Shun; Siegfried, Blair D.; Coates, Brad S.; Sumerford, Douglas V.; Zhu, Kun Yan

(CK, MSC, LLB, KYZ) Department of Entomology, Kansas State University, Manhattan, KS 66506; (BDS) Department of Entomology, University of Nebraska-Lincoln, Lincoln, NE 68583; (BSC, DVS) Corn Insects and Crop Genetics Research, USDA/ARS, Ames, IA 50011

Transgenic crops expressing *Bacillus thuringiensis* (Bt) toxins have been widely used to control various insect pests. However, there is a growing concern regarding possible development of insect resistance against Bt crops. The identification of genes associated with Bt toxicity and/or resistance are considered to be crucial for preserving the Bt technology. In this study, we characterized 12 aminopeptidase-like genes revealed from 15,000 ESTs that were generated from the gut of European corn borer (ECB, *Ostrinia nubilalis*) larvae. The aminopeptidase gene (Amino 8) was expressed in all five larval instars as well as in pupae but was not expressed in eggs. Its expression in larval stages was observed mainly in the midgut with only very low expression in foregut and no detectable expression in hindgut, fat bodies, salivary glands, and carcass. Expression profiling of these genes in a Bt-susceptible and two resistant strains revealed that Amino 8 was virtually not expressed or expressed at very low level in the two resistant strains of ECB. Feeding-based RNA interference (RNAi) of Amino 8 using Amino 8 dsRNA resulted in a significant reduction in the expression of Amino 8 gene on days 4, 6, and 8, compared with the control larvae fed with the diet containing GFP dsRNA. Our bioassay using Cry1Ab furthermore showed that the RNAi of Amino 8 reduced the larval mortality by 22% compared with the larvae fed with GFP dsRNA and water. Our results suggest that the aminopeptidase-like Amino 8 may play an important role in Cry1Ab toxicity and resistance in ECB.

35

Genome Structures of Two Closely Related Bracoviruses (CpBV and CgBV) Sharing a Common Pathogenic Host, *Plutella xylostella*

Kim, Yonggyun*

Department of Bioresource Sciences, Andong National University, Andong, Korea

Polydnaviruses are an insect DNA virus group symbiotic to some hymenopteran subfamilies. This virus family is divided into two subfamilies of bracovirus (BV) and ichnovirus (IV). Even though their independent origins, both BV and IV exhibit an identical life maintenance pattern. They are present at proviral forms in their specific hymenopteran hosts. Their replication occurs only in the ovary at host pupal stage. The episomal viral particles are transferred to target lepidopteran insect hemocoel by parasitization of hymenopteran host. This horizontal transmission is the dead end for the viral life due to no viral replication. Thus the viral generation is accomplished by proviral form in the wasp host. *Cotesia plutellae* and *C. glomerata* can parasitize a common lepidopteran insect, *Plutella xylostella*. These wasps have their own BVs, CpBV and CgBV, respectively. Parasitized *P. xylostella* by either *C. plutellae* or *C. glomerata* exhibits altered physiological symptoms in development and immune reactions. Though several other parasitic factors such as ovarian proteins, venom, and teratocytes are identified, CpBV or CgBV has been more

focused on elucidating various host physiological alterations occurring due to the parasitism, which has driven the viral genome project. CpBV attains a typical bracovirus structure by its single unit membrane envelope, in which multiple nucleocapsids are enclosed. Its genome DNAs are segmented and located on the genome of *C. plutellae*. Its replication begins at adult tissue development during pupal stage. An apparent genome size is 471 kb estimated from 27 segments separated on 5% agarose gel. A current work on the genome has been completely sequenced 24 genomic segments and analyzed their genomic structure. The aggregated genome size is 351,299 bp long and exhibits an average GC content of approximately 34.6%. Average coding density is about 32.3% and 125 putative open reading frames are predicted. Though more than half (52.5%) of predicted genes are annotated as hypothetical, the annotated CpBV genes share amino acid sequence homologies with those of other bracoviral genomes. The annotated genes are classified into the known bracoviral families, in which a family of protein tyrosine phosphatase is the largest including 36 ORFs, suggesting a significant role during parasitization. In addition, 8 and 7 ORFs encode $\text{I}\kappa\beta$ -like and EP1-like, respectively. Some predicted genes are known only in *Cotesia*-associated bracoviral genomes. Finally, two homologous genes, CpBV15 α/β , are unique in CpBV genome, which are not matched to any other known polydnviral genes. Their homology with malarian circumsporozoite toxin and eukaryotic translation inhibition factors suggests their function in host translation inhibitory factor. CgBV genome consists of at least 12 segments in unequal amounts. Though CgBV has much small number of genome segments, most CpBV genes were detected in CgBV genome, but showed genus variation. For example, CpBV15 α/β -like, lectin, and histone H4 genes are detected in CgBV genome and show high homologies (over 90%) in amino acid sequences, suggesting a common gene composition of *Cotesia*-associated bracoviruses.

36

Identification and phylogenetics of two putative RNAi genes in the cricket, *Allonemobius socius*

*Kraus, Emily ;Marshall, Jeremy L.

Kansas State University, Department of Entomolgy, Manhattan, KS 66506

The RNA interference (RNAi) pathway is involved in post transcriptional gene silencing and can either knockout or knockdown a gene transcript. Molecular biologists have developed techniques to utilize this naturally occurring pathway to determine the function of a wide-array of genes. The RNAi pathway involves several pieces of machinery, including the genes dicer, argonaute, RNA-induced silencing complex (RISC), and small interfering RNA (siRNA). The objective of our study was to determine the genes involved in the RNAi pathway in the cricket *Allonemobius socius*. This species is an excellent laboratory organism, as it has a relatively quick generation time, simple colony maintenance, large numbers of EST sequences available, and many genetic tools have worked well including RNAi. In our study, we surveyed about 150,000 testis-specific ESTs to identify putative RNAi genes. From this search we identified three sequences believed to be part of the RNAi pathway due to genetic similarity with RNAi pathway genes in other species. Using BLAST analysis and multialignments to make phylogenies for each gene, we determined that our three sequences represent two putative genes of the RNAi machinery in crickets. The sequences appear to code for a Dicer and an Argonaute. More research is needed to determine if other RNAi machinery exists in *A. socius*.

37

RNAi revealing neuropeptide functions in the *Tribolium castaneum*

*Li, Bin; Akins, J. Michael; Cosme, V. Luciano; Park, Yoonseong,

(BL, MJA, LVC, YP) Kansas State University, Department of Entomology, Manhattan, KS 66506

Neuropeptides and their receptors are the most vital components in the signaling pathways for regulation and/or modulation of various physiological events and behavioral processes of organisms. The red flour beetle, *Tribolium castaneum*, provides an opportunity to examine hypomorph or amorph of these signaling pathways by using the RNA interference (RNAi) technique. Genome annotation of *T. castaneum* predicted more than 80 mature peptides in the preprohormones encoded by 41 genes. We performed RNAi targeting 43 transcripts produced from the 41 neuropeptide genes and analyzed the phenotypes in various developmental stages. Double stranded RNA (dsRNA) was injected at the penultimate larval stage and their abnormal phenotypes in pupal development, adult eclosion, reproduction, and survivorship in the offspring were analyzed. If early lethality occurred by the larval RNAi, dsRNA was injected into late pupae to investigate the effects of RNAi on the adult and their offspring. We found that RNAi of thirteen genes caused eclosion deficiencies or malformation of adult morphology, of which two peptides caused developmental delays in the pupal stage. RNAi of sixteen peptide genes resulted in deficiencies in reproduction and sixteen peptide genes were required for survival of the offspring. About two thirds of the neuropeptide genes were found to be required for normal development of the beetle. Our current RNAi protocol is being further evaluated and developed to investigate expanded life stages under environmental stresses and to examine a larger number of genes.

38

Defining the core apoptosis pathway in the disease vector *Aedes aegypti*

Liu, Qingzhen; *Clem, Rollie J.

(QL, RJC) Division of Biology, Kansas State University, Manhattan, KS; (QL) State Key Laboratory of Virology, College of Life Sciences, Wuhan University, People's Republic of China

To date, our knowledge concerning molecular pathways that regulate apoptosis in insects comes almost exclusively from studies done in the model organism *Drosophila melanogaster*. In *Drosophila*, the core apoptosis pathway includes the master regulator DIAP1, which is required to prevent spontaneous apoptosis in *Drosophila* cells due to its inhibition of the caspases Dronc and DrICE, while IAP antagonists such as Reaper, Hid and Grim negatively regulate DIAP1. Also required for apoptosis is the adaptor protein Ark, which binds and activates Dronc. Here we have studied the function of several *Aedes aegypti* genes that share

homology with known apoptosis regulatory genes in *Drosophila*. Silencing of the *Aelap1* gene in the *A. aegypti* cell line Aag2 caused spontaneous apoptosis, indicating that AeIAP1 plays a similar role to that of DIAP1. Silencing *AeArk* or *AeDronc* inhibited apoptosis triggered by several different apoptotic stimuli, while silencing effector caspases had varying effects. This is the first comprehensive study of apoptosis regulation in any insect other than *Drosophila melanogaster*. Our results suggest that the core pathway that regulates apoptosis in *A. aegypti* is similar to that of *Drosophila*, but there may be subtle differences in how these pathways are regulated in the yellow fever mosquito.

39

Progress Towards the Positional Cloning of the Maternal Selfish Gene *Medea*⁴ in *Tribolium castaneum*

Lorenzen, Marcé D.; Brown, Susan J.; Beeman, Richard W.

(MDL, RWB) USDA-ARS-Grain Marketing and Production Research Center, Manhattan, KS 66502

(SJB) Division of Biology, Kansas State University, Manhattan, KS 66506

Medea (**M**aternal-**E**ffect **D**ominant **E**mryonic **A**rrest) genes are maternally-acting, selfish genetic elements known only from two species of flour beetles in the genus *Tribolium*. In spite of their strictly limited taxonomic distribution they are of broad interest because of their potential to serve as gene drivers for genetic modification of target populations of arthropods. By far the most prevalent, naturally-occurring *Medea* element is *M^f*, which occurs on every continent, and is found in more than 40% of all *T. castaneum* populations sampled worldwide. *M^f* was initially subjected to low-resolution recombinational mapping using 24 marker loci that were evenly distributed over all nine autosomal linkage groups (chromosomes). This analysis localized *M^f* to one end of chromosome 3. A second round of higher-resolution mapping was done, which narrowed the *M^f* target region to a ~1Mb segment near the telomere. We screened our collection of *piggyBac* insertion lines and identified two EGFP-marked insertions located <1cM from *M^f*. These insertions are now being used for ultra-high resolution mapping of *M^f* in support of the positional cloning of this selfish genetic element. So far we have screened ~13,000 backcross progeny and have detected 28 recombination events (total = 0.2% recombination, or about 0.01 cM = 2.5 kb average spacing of recombination events) between the *AM460 piggyBac* insertion and the target *M^f* locus. These recombinants are being precisely mapped, and we will apply the same strategy to recombinants that occur between *M^f* and a second insertion (*AM351*). Once the site of the *M^f* lesion has been localized to a small segment (~several kb), this segment will be cloned from the *M^f* strain and its sequence will be compared with the corresponding region of the non-*M^f* GA2 strain. In order to obtain sequence variants from an *M^f* strain that could be used as a basis for this high-resolution mapping, we screened a genomic *M^f* bacterial artificial chromosome (BAC) library with 12 unique 40-mer "overlapping oligonucleotide" (overgo) probes spanning ~1.2 Mb in the *M^f* region of chromosome 3. BAC-end sequencing was performed on each of the 81 BAC clones identified. Useful dimorphisms were detected and are being used for recombinant mapping. The discovery of the DNA sequence of this unusual selfish element should provide insight into its molecular mechanism of self-selection.

40

Expression and Functional Analysis of Two Acetylcholinesterase Genes in Red Flour Beetle (*Tribolium castaneum*)

*Lu, Yanhui; Park, Yoonseong; Gao, Xiwu; Zhu, Kun Yan

(YL, YP, KYZ) Department of Entomology, Kansas State University, Manhattan, KS 66506;

(YL, XG) Department of Entomology, China Agricultural University, Beijing, China

Since the second acetylcholinesterase (AChE) gene was first reported in the greenbug (*Schizaphis graminum*) in 2002, two different AChE genes (*Ace1*, *Ace2*) have been found in many other insect species. In this study, we profiled development- and tissue-specific expression patterns of *Ace1* and *Ace2* and used a RNA interference (RNAi) technique to investigate biological functions of each gene in the red flour beetle (*Tribolium castaneum*), an important stored grain pest worldwide. Our results showed that both genes were virtually transcribed in all the developmental stages examined, including 1-d and 3-d eggs; 5-d and 20-d larvae; 1-d, 3-d and 6-d pupae; and 2-d and 14-d adults. The lowest expression levels of these genes were found in eggs, particularly for *Ace2*. As expected, these two genes were predominately expressed in brain although their expressions were also detected in the gut and carcass after the brain and ventral nerve cord were removed. Our preliminary studies showed that the gene-specific double-stranded RNA (dsRNA)-mediated transcript depletion can lead to the delayed pupation and adult emergence, abnormal molting and lethal phenotypes. Further studies are needed to elucidate possible novel functions of the two AChEs in insect growth and development.

41

Genetic Variation in Honey Bees from Central United States

*Magnus, Roxane M.; Szalanski, Allen L.

University of Arkansas, Department of Entomology, Fayetteville, AR 72701

We studied the genetic diversity of honey bees, *Apis mellifera* L., from central United States by DNA sequencing a portion of the mtDNA COI and COII intergenic regions, focusing primarily on feral and managed colonies by hobbyist and sideline beekeepers. Of the 394 samples sequenced, a total of 35 mitotypes were observed which represent all four *A. mellifera* lineages known to exist in the United States. The greatest amount of genetic variation was observed in the 'C' eastern European lineage which includes *A. m. ligustica*, *A. m. carnica* and *A. m. caucasica*. We also found evidence of the 'M' lineage (*A. m. mellifera*) still existing in this region, even though this subspecies has not been preferred by beekeepers for over 100 years. Interestingly, we found representatives of the 'O' North Africa / Middle East lineage, which have not been imported into the US since the 1880's. The commercial breeding industry uses relatively few subspecies as breeder queens which selection that provide resistance to pests, pathogens, and disease may be limited. Finding representatives of all four lineages and documenting genetic diversity within all 4

lineages demonstrates that there is a large amount of genetic variation within honey bees in the central United States. Future breeding programs utilizing these distinct lineages/mitotypes could be used for increasing the fitness of honey bee colonies maintained by beekeepers.

42

Distinguishing the roles of duplicated *Cyp12d1* genes in DDT resistance in *Drosophila melanogaster*

*McDonnell, Cynthia M.; Li, Hong-mei; Schuler, Mary A.; Berenbaum, May R.; Pittendrigh, Barry R.
(CMM, HML, MRB, BRP) University of Illinois Urbana Champaign, Department of Entomology, Urbana, IL 61801;
(MAS) University of Illinois Urbana Champaign, Department of Developmental and Cellular Biology, Urbana, IL 61801

Cyp12d1 is a cytochrome P450 that was first identified by its induced response to DDT in *Drosophila melanogaster*. Its response to DDT is predictable, to the extent that it can be used to indicate exposure to DDT in the laboratory. *Cyp12d1* is actually a pair of genes, *Cyp12d1-p* and *Cyp12d1-d*, which are nearly identical sequences, except for 3 non-synonymous nucleotide differences, and are arranged in tandem on chromosome 2R. A variety of xenobiotics appear to induce *Cyp12d1-d* or *Cyp12d1-p*, but it is not clear if only one or both genes respond to xenobiotic treatment. In order to gain a better understanding of the role of *Cyp12d1-d* or *Cyp12d1-p* in xenobiotic response we must first elucidate which of these two genes are induced by compounds such as DDT.

Despite 99.4% shared identity in their amino acid sequence, differentiation between the two genes is possible. The *Cyp12d1-d* gene is predicted to produce transcripts of two lengths, one identical to the *Cyp12d1-p* transcript and one with an additional 129 nucleotides beyond the stop codon. Variations in the *Cyp12d1* transcript length have been detected by northern analysis of adult flies exposed to DDT. A DDT-susceptible strain, 91-C, constitutively expressed the longer transcript, but in response to DDT, the shorter transcript is induced. Interestingly, in 91-R, a DDT-resistant strain derived from 91-C, only the shorter transcript of *Cyp12d1* is expressed, both constitutively and in response to DDT. Thus, the induced response of the shorter transcript in 91-C seems to have been selected for increased DDT resistance in 91-R. In order to determine if the observed differences between these strains corresponded to different *Cyp12d1* genes, we constructed primers for the unique 3' region of *Cyp12d1-d* to measure gene expression between 91-C and 91-R in response to DDT. Differential expression of *Cyp12d1-d* or *Cyp12d1-p* between the two strains would suggest that the two genes differ in their response to DDT. We discuss how our findings fit within the model of gene duplication and divergence used to explain the diversity of cytochrome P450 genes found in insects.

43

Mode of Action of the Insect Growth Regulator Diflubenzuron: a Comprehensive Study in *Tribolium castaneum* Utilizing Proteomic and Genomic Approaches

*Merzendorfer, Hans; Kim, Hee S.; Specht, Charlie A.; Beeman, Richard W.; Kramer, Karl J.; Muthukrishnan, Subbaratnam (HM) University of Osnabrück, Department of Biology/Chemistry, 49069 Osnabrück, Germany; (HM, HSK, KJK, SM) Department of Biochemistry, Kansas State University, Manhattan, KS 66506; (CAS) Department of Medicine, University of Massachusetts, Worcester, MA 01605, USA; (RWB, KJK) Grain Marketing and Production Research Center, ARS-USDA, Manhattan, KS 66502

Several benzoylurea-derived insecticides such as diflubenzuron (DFB, Dimilin®) are in wide use to control various insect pests. Although this class of compounds is known to disrupt molting and to affect chitin content, their precise mode of action is still not understood. To gain more insight into the mechanism underlying the insect growth regulating effects of benzoylurea compounds, we conducted a comprehensive study utilizing genomics, proteomics and the model beetle species and stored product pest *Tribolium castaneum* (red flour beetle). DFB was added to a wheat flour-based diet and fed to larvae and adults at various concentrations. Whereas larvae were significantly impaired by DFB (4 d LC₅₀ ~ 50 ppm), adults did not exhibit any increase in mortality when compared with control animals. Younger larvae, which were generally more susceptible to the insecticide, died during the larval-to-larval molts. Older larvae died primarily during the larval-to-pupal molt, but a few molted to the pupal stage after which they died during the pupal-to-adult molt. Oviposition and subsequent embryonic development appeared to be normal after DFB-treatment of adult females. However, egg-hatch was completely inhibited by DFB treatment and the unhatched embryos became twisted and enlarged. This embryonic phenotype was indistinguishable from that observed after chitin synthase A (*Chs-A*) knockdown in adult females by RNAi.

The chitin content of whole larvae or isolated larval midguts was reduced by ~ 50% after DFB treatment, an effect which could be also visualized by Calcofluor White staining of cuticles or staining of the peritrophic matrix with a chitin-binding protein domain conjugated to fluorescein isothiocyanate. Surprisingly, the chitosan content was elevated significantly after DFB treatment. To explain the observed changes in chitin/chitosan content, we performed RT-PCR to examine the expression levels of several genes involved in chitin metabolism, as well as a gene encoding a putative receptor for DFB. Gene expression for *Chs-A* and *Chs-B* was only marginally up-regulated. However, the transcript level of chitin deacetylase 2 (*Cda-2*) was considerably increased upon DFB treatment, corroborating our observation of increased chitosan content. Expression of the gene encoding a gut chitinase (*Cht-4*) was unchanged, while that of an epidermal chitinase (*Cht-5*) was reduced significantly, perhaps in compensation for the reduction in chitin synthesis. Expression of the gene for a putative sulfonylurea receptor (*Sur-1*), a potential target of DFB, was unaffected.

To conduct an unbiased survey of all of the genes that are potentially affected by DFB treatment, a genome-wide tiling array for hybridization with differentially labeled DNA probes derived from mRNA of control and DFB-treated larvae was utilized. Fluorescence intensity levels from two biological replicates were compared over an average of at least four tiles per gene. Among the genes that showed changes in their expression levels were genes known to be involved in chitin metabolism, cuticle formation and insecticide detoxification. *In silico* comparison of the proteins resolved by two-dimensional gel electrophoresis of midgut protein extracts from DFB-treated and untreated larvae revealed changes in protein levels due to the DFB treatment. Among the 388 proteins that were analyzed in duplicate gels, only 26 proteins (~7%) were significantly changed as determined by laser

densitometry, including 21 up-regulated and 5 down-regulated proteins. Those seven proteins whose levels changed more than 2.5-fold (~2%, all up-regulated upon DFB-treatment), were identified by MALDI-MS analysis of tryptic peptides. Among these proteins was a UDP-*N*-acetylglucosamine pyrophosphorylase (UAP-2), which catalyzes a step in the synthesis of the substrate used by chitin synthase.

In summary the red flour beetle is a good model organism for investigating the mode of action of bioactive materials such as insect growth regulators and other insecticides. Our study revealed a surprisingly small number of genes and proteins that are significantly changed in response to DFB treatment of *Tribolium*. Some of those are involved in chitin metabolism; others are involved in cuticle formation and insecticide metabolism. Therefore, DFB affects the expression of several different genes that are critical for normal insect growth and development.

44

The Tick Genome Landscape – Challenging Terrain for the Genome Explorer.

Meyer, Jason M.; Van Zee, Janice P.; Silva, Catherine E.; Schlueter, Shannon S.; Schlueter, Jessica; Kurtti, Tim J.; Dixon, Phil; *Hill, Catherine A.

(JMM, JPV, CES, CAH) Purdue University, Department of Entomology, West Lafayette, IN, 47907; (SS) Purdue University, Department of Computer and Information Technology, West Lafayette, IN, 47907; (JS) Purdue University, Department of Agronomy, West Lafayette, IN, 47907; (TJK) University of Minnesota, Department of Entomology, St. Paul, MN, 55108; (PD) Iowa State University, Ames, IA, 50011

Understanding the arrangement of repetitive and non-repetitive DNA on chromosomes is essential for studies of genome biology. Ixodid ticks (subphylum Chelicerata; family Ixodidae) are obligate ectoparasites that transmit a wide variety of disease-causing agents to humans and animals worldwide. We are exploring genome organization in ixodid ticks to complement genome sequencing efforts in this important group of arthropod vectors. The Lyme disease tick (*Ixodes scapularis*) genome project represents the first initiative to sequence a chelicerate species. The *Ixodes* genome assembly, completed in late 2008, is a useful tool to begin to analyze tick genome organization. The tick genome landscape is challenging terrain for the genome biologist. Tick genomes are large (> 1 Gbp) and repeat rich (> 70% repetitive DNA). Large blocks of heterochromatic DNA are localized to the telomeres and centromeres of tick chromosomes; these regions contain multiple families of highly conserved tandem repeats that account for more than 10% of the genome. The euchromatin comprises a large number of potentially duplicated genes. Estimates of divergence rates suggest that several recent and large-scale duplication events may have shaped the genomes of members in the ixodid pro- and metastriate lineages. Many duplicated genes that appear to be under positive selection are predicted to function in tick-host-pathogen interactions and xenobiotic detoxification. These findings have implications for the selection of acaricide and anti-tick vaccine targets, and the management of pesticide resistance. The euchromatin is also littered with tandem repeats and copies of type I and II transposable elements, most of which are non-functional remnants of ancient transposition events. In *Ixodes* (and other ticks being considered for genome projects), extensive repeat proliferation and high levels of genome polymorphism present challenges for genome assembly and annotation. We are developing physical and genetic mapping approaches to help genome biologists overcome these challenges and advance tick genome research.

45

AngstD - A domain-centric tool for the analysis of genes, sequences, trees and domains

Moore, Andrew D.*; Held, Andreas; Schueler, Andreas; Weiner, January 3rd; Bornberg-Bauer, Erich; Evolutionary Bioinformatics Group, Institute for Evolution and Biodiversity, University of Muenster Hüfferstr. 1, 48149 Münster, Germany

Proteins can be considered as a sequence of evolutionary blocks defined at sequence level by sensitive profiles. The study of particular protein families of interest can be simplified by studying domain wise re-arrangement events, as domain re-arrangements can contain considerable evolutionary signals. Here, we present the first version of a domain-centric software suite, AngstD, which aims at the visualization and analysis of protein families. With the increasing amount of genomic data available today, AngstD has the potential to expedite the analysis of families of interest and summarize the results in high-quality graphics. Using AngstD, the user can annotated proteins using InterPro, project arrangements onto an established phylogeny, uncover events such as domain gain and loss along lineages and calculate domain co-occurrence graphs. All steps are interactive and allow the user to manipulate the graphical output to produce publication-ready graphics. We illustrate the abilities of AngstD using a case of LGT from the intracellular parasite *Wolbachia* to its host, the parasitic hymenoptera *Nasonia vitripennis*.

46

The Hymenoptera Genome Database

*Munoz-Torres, Monica C.; Reese, Justin T.; Childers, Christopher P.; Hagen, Darren E.; Elsik, Christine G. Georgetown University, Department of Biology, Washington, DC 22209.

The Hymenoptera Genome Database (HGD, <http://HymenopteraGenome.org>) is an informatics resource supporting genomics of hymenopteran insect species. It currently provides access to genome sequences and annotation for honey bee, *Apis mellifera* (<http://BeeBase.org>) and the parasitic wasp, *Nasonia vitripennis* (<http://NasoniaBase.org>). Honey bee is the most important pollinator of food crops and is a useful model social behavior and immunity. *Nasonia* belongs to a large and extremely important group of Hymenoptera with more beneficial insects to humans than any other group. It is also being developed as a model organism in genetics and development. The combination of these two genomes into a single resource will allow biologists to leverage the genome information, and will enhance the value of genomic data for each species by facilitating cross-species comparisons. HGD provides a central location for hymenopteran genomics data. The genome viewer includes tracks for predicted

genes, manually annotated genes, linkage markers, transposable elements, SNPs, and alignments of protein homologs and ESTs. The comparative map viewer displays comparisons between linkage maps and genome assembly, highlighting markers that are common among maps. We seek to generate new computational data and add value to existing data, including creating consensus gene sets for each species. Most notably, we coordinated community gene annotation of the bee and wasp genomes. Remote connection to the Chado database by Apollo Genome Annotation client software allows biologists to annotate gene models. User-submitted annotations are then reviewed and incorporated into an Official Gene Set for each species. Our plans are to extend HGD to serve as data repository for the complete genome annotation for each species, which will be accomplished implementing a computational pipeline based on ortholog comparisons to assign Gene Ontology categories to each generated sequence.

47

RUI: Deciphering Plant-Herbivore Interactions

*¹Musser, Richard O., ¹Hum-Musser, Sue; and ²Vogel, Heiko

*Western Illinois University, Biological Sciences Department, Macomb, IL 61455; Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, Jena D-07745

Insect herbivory dramatically affects how plants defend themselves against insect feeding, pathogen infection, as well as plant growth habits. Plants respond differently when fed on by insects compared to mere physical wounding. This response is likely due to saliva produced and secreted by the insect. The complexity of plant defense pathways still need to be determined due to interactions between the two major plant defense mechanisms that can be stimulated by herbivores with different methods of feeding. Chewing insects produce large wounds and sucking insects produce smaller punctures in the plant tissue. In addition, the success of plant feeding insects is dependent on their ability to overcome plant produced toxic compounds. This new NSF funded project will investigate the role of caterpillar saliva and aphid secretions simultaneously stimulating the corn and tomato plant defense responses using microarray hybridization technology. This technology allows for the study of the behavior of practically all the plant genes in response to insect feeding. There have been no studies of this kind investigating the simultaneous feeding of two different types of plant herbivory that are known to stimulate the opposite and different plant defense pathways, compared to feeding from a single type of insect. Using the same technology, this research will concurrently address how the insect herbivore responds to consumption of toxic plant defense compounds. These experiments may reveal any connection between plant defense compounds in response to insect herbivory and the insect defense responses to plant defensive chemicals. Because insect pests cause billions of dollars of damage to agriculture annually, information from this research could provide the insight required to develop novel methods for protecting economically important plants.

48

Discovering SNPs in model organisms using the Illumina Genome Analyzer

*Noll, Aaron; Blumenstiel, Justin; Perera, Anoja; Walton, Kendra; Griffiths, Jennifer; Hawley, Scott; Staehling-Hampton, Karen Stowers Institute for Medical Research, Kansas City, Missouri 64110

Determining causative SNPs for phenotypic differences between strains in model organisms is typically achieved using standard genetic mapping and candidate gene sequencing. In situations where multiple causative SNPs create a phenotype, or where there are few polymorphic markers or candidate genes, this process can be arduous and depending on the organism, could consume months to years. Here we describe the rapid determination of homozygotic SNPs between strains in the model organisms *Saccharomyces Cerevisiae* and *Drosophila Melanogaster* using high throughput sequencing on the Illumina Genome Analyzer.

Differences in depth, alignment and sequencing errors can introduce artificial variations between strains. An elementary but effective approach of using thresholds for depth, mapping confidence and base quality can be used to call SNPs. Using this approach, average depth as low as 8X was found to be sufficient for calling homozygotic SNPs. This method allows for rapid and accurate identification of homozygotic SNPs specific to a given strain and could benefit traditional genetic labs substantially in terms of time and cost. Coupled with mutagenizing agents, which can result in phenotypic differences through random SNP generation, there is great potential for advancing model organism genetics.

49

Transcriptome profiling of *Bacillus thuringiensis* Cry3Aa intoxication of *Tenebrio molitor* larvae

*Oppert, Brenda; Toutges, Michelle; Lorenzen, Marcé; Bouffard, Pascal; Li, Lewyn; Dowd, Scot; Marshall, Jeremy; Huestis, Diana; Fabrick, Jeff; Oppert, Cris; and Jurat-Fuentes, Juan Luis.

(BO, MT, ML) USDA ARS Grain Marketing and Production Research Center; Manhattan, KS, USA; (PB, LL) 454 Life Sciences, Branford, CT, USA; (SD) Medical Biofilm Research Institute, Lubbock, TX, USA; (JM, DH) Department of Entomology, Kansas State University, Manhattan, KS, USA; (JF) USDA-ARS U.S. Arid Land Agricultural Research Center, Maricopa, AZ, USA; (CO, JLJF) Entomology and Plant Pathology, University of Tennessee, Knoxville, TN, USA.

Cry proteins from *Bacillus thuringiensis* (Bt) are highly specific insecticidal toxins. While the lepidopteran-specific Cry1A toxins are the best characterized, little is known about the intoxication process of coleopteran-specific Cry3 toxins. Therefore, transcriptome profiling was used to gain insight into Cry3Aa intoxication of the coleopteran *Tenebrio molitor* (the yellow mealworm). High-throughput sequencing of Cry3Aa-treated and/or untreated larval gut transcripts (Genome Sequencer FLX pyrosequencing system, 454 Life Sciences) resulted in approximately 280,000 *T. molitor* expressed sequence tags (ESTs) with an average read length of 238 bp. ESTs were assembled into 65,536 uniESTs using custom scripts and Roche GS de novo Assembler. Those with at least 2-fold coverage were utilized for downstream processing (28,164 tentative consensus sequences). Sequence comparisons to *Drosophila melanogaster* revealed 10,185 matches with significant e-values ($<e^{-25}$), and statistical comparison of relative

expression levels indicated that there were 2,126 differentially expressed genes in Cry3Aa-intoxicated larvae compared to control larvae, ranging from an 8-fold increase to a 42-fold decrease in gene expression. Many of the down-regulated genes were associated with energy production, including hydrolases and glycosidases, suggesting that larvae respond to intoxication through conservation of energy and down-regulation of food processing genes. Up-regulated genes included those associated with membrane functions, such as those associated with membrane reorganization. These data summarize the first high-throughput approach to Bt mode of action studies. Understanding the global responses to Cry3Aa intoxication in a beetle pest may ultimately lead to improved toxin-based pest management strategies.

50

Condition-dependent Transcription and Neofunctionalization: Insights from Genome-wide Expression Profiling in *Daphnia*

Pfrender, Michael*; Gilbert, Don; Colbourne, John; Rechtsteiner, Andreas; Choi, Jong-Hyeon; Lopez, Jacqueline; Skalitzky, Courtney; Manak, John; Cherbas, Peter

(MP) Biology Dept., Utah State Univ., Logan, Utah 84322, pfrender@biology.usu.edu; (DG) Biology Dept. and The Center for Genomics and Bioinformatics, Indiana Univ., Bloomington, IN 47405, gilbertd@indiana.edu; (JC) The Center for Genomics and Bioinformatics, Indiana Univ., Bloomington, IN 47405, jcolbour@cgb.indiana.edu; (AR) Dept. of MCD Biology, Univ. of California, Santa Cruz Santa Cruz, CA 95064, andreas@ucsc.edu; (JHC) The Center for Genomics and Bioinformatics, Indiana Univ., Bloomington, IN 47405, jeochoi@cgb.indiana.edu; (JL) The Center for Genomics and Bioinformatics, Indiana Univ., Bloomington, IN 47405, jaalopez@cgb.indiana.edu; (CS) Roche NimbleGen, Inc., 504 South Rosa Road, Madison, WI 53719, cskalitzky@nimblegen.com; (JM) Dept. of Biology and the Roy J. Carver Center for Genomics, Univ. of Iowa, Iowa City, IA 52242, john-manak@uiowa.edu; (PC) The Center for Genomics and Bioinformatics and Biology Dept., Indiana Univ., Bloomington, IN 47405, cherbas@cgb.indiana.edu

Advances in whole-genome tiling path microarrays are opening a window into the patterns of transcriptional activity in organisms occupying diverse ecological contexts. One such organism, the waterflea *Daphnia*, has long been a model for investigating phenotypic plasticity and sensitivity to chemical stressors in the environment. Ecologically relevant chemical exposures reveal condition-dependent patterns of transcription. Two observations from this study are: 1) the class of genes most highly differentially regulated in response to environmental cues are genes in this lineage lacking orthology to other arthropod genomes, and 2) a large fraction of differential regulation occurs in unannotated regions of the genome.

Taking advantage of the large number of duplicate genes in the *Daphnia* genome we compare sequence divergence among paralogous genes to condition-specific expression divergence. This comparison demonstrates the process of neofunctionalization in gene regulation. Divergence in transcription across environments varies among paralogous gene families and functional groups of genes. In concert with large numbers of duplicate genes in *Daphnia* is a high incidence of concerted evolution among duplicate genes. The effect of concerted evolution is to homogenize sequences retarding sequence divergence. Neofunctionalization of regulation (condition-specific expression) is decoupled from sequence divergence in paralogous gene families showing concerted evolution.

Our understanding the genomic responses of organism to ecologically relevant conditions is advancing rapidly with the growing availability of genomic tools across phylogenetically diverse taxa. Studies of this kind, examining condition-dependent patterns of genome-wide transcription, are a central component of ecological genomic investigations linking genome function and phenotype to environmental context.

51

Proteomic identification of serpin-1 isoforms in the caterpillar *Manduca sexta*

*Ragan, Emily J.; Kanost, Michael R.

Kansas State University, Department of Biochemistry, Manhattan, KS 66506

The Lepidopteran insect *Manduca sexta* is large and well-suited for biochemical analysis of proteins in the hemolymph (blood). We are interested in proteins involved in the immune response. Serine proteases function in proteolytic cascades that trigger effector responses, such as antimicrobial peptide production and prophenoloxidase activation. Serpins regulate specific active proteases by covalently binding and forming covalent serpin-protease complexes. Serpin-1, an abundant plasma protein, has 12 isoforms that are generated by mutually exclusive alternative splicing of the ninth exon. The twelve isoforms differ in reactive center loop sequence and, therefore, are predicted to inhibit different proteases. Previous work has identified serpin-1 isoform inhibitory specificity for commercially available proteases. We are interested in determining endogenous proteases inhibited by the serpin-1 isoforms as well as investigating and comparing the expression levels of the different isoforms. Using 2D-PAGE and MALDI-TOF/TOF I identified nine individual serpin-1 isoforms as well as several serpin-1 protease complexes. Serpin-1A, serpin-1E and serpin-1J were found to be complexed with hemolymph proteinase-8, and an unidentified isoform of serpin-1 formed a complex with hemolymph proteinase-1. Discovering the serpin-1 isoforms that inhibit specific proteases enhances our understanding of the regulation of proteolytic cascades in *M. sexta*.

52

Genomic analysis of the American cockroach bacterial endosymbiont

Sabree, Zakee L.*; Kambhampati, Srinivas; Moran, Nancy A.

(ZLS) Center for Insect Science and Department for Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721

(SK) Department of Entomology, Kansas State University, Manhattan, KS 66506 USA

(NAM) Department for Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721

Nitrogen acquisition and assimilation is a primary concern of insects feeding on diets largely composed of plant material. Reclaiming nitrogen from waste products provides a rich reserve for this limited resource, provided that mechanisms for recycling are in place. Cockroaches, unlike most terrestrial insects, excrete waste nitrogen within their fat bodies as uric acids, presumably as a supplement when dietary nitrogen is limited. The fat bodies of most cockroaches are inhabited by *Blattabacterium sp. (cuenoti)*, which are vertically-transmitted, Gram-negative bacteria that have been hypothesized to participate in uric acid degradation, nitrogen assimilation and nutrient provisioning. We have completely sequenced the *Blattabacterium* genome from the American cockroach *Periplaneta americana*. Genomic analysis confirms that *Blattabacterium* is a member of the Flavobacteriales and its closest relative is the cicada endosymbiont *Sulcia muelleri*. Metabolic reconstruction indicates that it lacks recognizable uricolytic enzymes, but it can recycle nitrogen from uric acid degradation products urea and ammonia using its glutamate dehydrogenase into glutamate. Subsequently, *Blattabacterium* can produce all of the essential amino acids, various vitamins and other nutrients from a limited palette of metabolic substrates. This ancient relationship between *Blattabacterium* and cockroaches has allowed the insect to successfully subsist on nitrogen-poor diets and exploit nitrogenous wastes, thereby facilitating its global distribution.

53

Site-specific recombination to improve transgenic medfly strains

*Schetelig, Marc F.; Scolari, Francesca; Handler, Alfred M.; Kittelmann, Sebastian; Gasperi, Giuliano; Wimmer, Ernst A.

(MFS, AMH) USDA/ARS, Center for Medical, Agricultural and Veterinary Entomology, 1700 SW 23rd Drive, Gainesville, FL 32608, USA; (FS, GG) Dipartimento di Biologia Animale, Università di Pavia, Piazza Botta 9, 27100 Pavia, Italy;

(SK, EAW) Department of Developmental Biology, Georg-August-University Göttingen, GZMB, Ernst-Caspari-Haus, Johann-

Friedrich Blumenbach Institute of Zoology and Anthropology, Justus-von-Liebig-Weg 11, 37077 Göttingen, Germany

The Sterile Insect Technique (SIT) is an environmentally friendly bio-control method used in area-wide pest management of the Mediterranean fruit fly (medfly) *Ceratitis capitata* (Wiedemann). Recently, we generated new transgenic strains to improve reproductive sterility and sex-specific marking of medfly. All the DNA constructs, which were used to generate these strains, carried an attachment P site (*attP*) – a short DNA sequence for site-specific recombination. The efficiency of the reproductive sterility and marking systems and the fitness of the transgenic flies were highly influenced by position effects of the transgenes. Strains that were successfully tested and evaluated as “beneficial” for the functionality of the transgenic system were then studied for further improvements by using site-specific recombination via their *attP* site. A two-step modification was tested in these transgenic medfly strains: 1) the combination of two transgenic systems at a positively evaluated genomic position and 2) the increase of transgene stability by the deletion of transposon ends with newly developed medfly jumpstarter strains. Such new possibilities in modifying medfly transgenes will enhance the development of new and safe applications for SIT programs as well as functional studies in Tephritid species.

54

Domain arrangements and modular evolution within the Arthropodic clade

Schuler, Andreas*; Moore, Andrew D.; Bornberg-Bauer, Erich;

Evolutionary Bioinformatics Group, Institute for Evolution and Biodiversity, University of Muenster Hüfferstr. 1, 48149 Münster, Germany

New sequencing technologies are rapidly producing vast amounts of genomic data. In the past, sequenced genomes from model organisms have been separated by large evolutionary distances, complicating the exact quantification of events in modular protein evolution. With the completed sequencing of 12 drosophila genomes, and the further addition of other arthropod genomes such as *Apis mellifera*, *Tribolium castaneum* or *Daphnia pulex*, a well resolved dataset has arisen which allows a detailed analysis of domain wise evolution. With this dataset, consisting of 17 arthropodic genomes, we are studying the events in modular evolution. Using a rigorous annotation strategy, we are focusing on events that facilitate the formation of (novel) domain arrangements such as gene fusion / fission and domain emergence. We plan to quantify such events within the well resolved time-scale this dataset provides, as well as examine the impact of such events on functional groups. Here, we show first results of the annotation and present evolutionary rates of fusion / fission and domain emergence.

55

Antibacterial and antiviral responses of larval *Heliothis virescens*

*Shelby, K.S.; Popham, H.J.R.

In order to better understand the immune response of *Heliothis virescens* against microbial entomopathogens several orthologs of antibacterial response proteins were identified and extracted from Heliobase (an EST database for *H. virescens*) for detailed study. Induction of these antibacterial protein transcripts by fungal, bacterial and baculoviral infection was quantitated using qPCR. Scolexin-B, Growth-blocking peptide binding protein, and C-type lectin were rapidly induced by bacterial elicitation; however *per os* HzSNPV infection elicited no significant increase in transcript level. Expression of additional antibacterial response proteins also will be reported.

56

Pea Aphid Glutathione Peroxidase, a Likely Component of Aphid Saliva

*Sheng, Changzhong; Cui, Feng; Edwards, Owain; Reese, John; *Reeck, Gerald
(CS, FC, GR) Kansas State University, Department of Biochemistry, Manhattan, KS 66506; (OE) CSIRO Entomology, Perth
(Floreat); (CS, FC, OE) CRC for Plant Biosecurity; (JR) Kansas State University, Department of Entomology, Manhattan, KS 66506

Based on EST frequencies, glutathione peroxidase appears to be one of the more abundant enzymes of saliva of the pea aphid, *Acyrthosiphon pisum*. We have cloned the cDNA encoding this enzyme. The ORF of 699bp encodes 233 amino acid residues, with a predicted signal peptide for extracellular secretion at its N-terminus. The predicted amino acid sequence has strong similarity to other glutathione peroxidases, especially from bacteria. The corresponding pea aphid gene has 5 introns. We have expressed the enzyme in *E. coli* and in Sf9 cells. The recombinant protein as expressed in *E. coli* was purified by chromatography on a Ni-NTA column followed gel filtration, and the enzyme expressed in Sf9 cells by successive use of several column chromatographies. The molecular mass of the mature recombinant protein is 29 kDa on SDS-PAGE. The enzyme expressed in Sf9 cells did not bind to ConA-Sepharose, suggesting that the protein is not glycosylated. The enzyme expressed in *E. coli* cells showed enzymatic activity using cumene hydroperoxide as the oxidizing substrate (and glutathione as the reducing substrate). We hope to study the activity using phosphatidylcholine hydroperoxide, which has been shown to be a preferred substrate for some glutathione peroxidases.

57

Single Inverted Terminal Repeats Of The *Junonia coenia* Densovirus Promotes Somatic Chromosomal Integration Of Vector Plasmids In Insect Cells And Supports High Efficiency Expression.

*Shirk, Paul D.; Furlong, Richard B.
USDA ARS CMAVE 1700 SW 23rd Drive Gainesville FL 32608

Plasmids that contain a disrupted genome of the *Junonia coenia* densovirus (*JcDNV*) integrate into the chromosomes of the somatic cells of insects. When subcloned individually, both the P9 inverted terminal repeat (P9-ITR) and the P93-ITR promote the chromosomal integration of vector plasmids in insect embryonic cells. Microinjection of plasmids that included the 517 bp of either the P9-ITR or the P93-ITR into embryos of *Drosophila melanogaster* resulted in somatically transformed larvae and adults with efficiencies comparable to the *JcDNV* vectors. A somatic transformation vector, pDP9, that includes the P9-ITR was constructed. The pDP9 vectors provide a simplified means of producing somatically transformed larvae or permanently transformed cultured insect cells. Transformation of *Bombyx mori* Bm5 cells or *Spodoptera frugiperda* Sf9 cells with pDP9 vectors result in the integration of the pDP9 vectors into genomic DNA of the Bm5 and Sf9 cells. In addition, the pDP9 vectors support high levels of expression of foreign genes. Insertion of foreign genes 3' of the densoviral P9 promoter results in a high level production of the gene by pDP9 transformed cells. P9 driven transcription within the pDP9 transformed Sf9 cells results in foreign gene transcript levels 5,000 fold higher than constitutively expressed actin 3 genes. The pDP9 vector transformation results in the efficient selection of high-producing clones for biosynthesis of target proteins.

58

Cis-Regulatory Element (CRE) Discovery in Mosquitoes using Comparative Genomics.

Sieglauff, Douglas H.; *Dunn, W. Augustine; Xie, Xiaohui S.; Megy, Karyn; Marinotti, Osvaldo; and James, Anthony A.
(DHS,WAD,OM,AAJ) University of California, Irvine, Department of Molecular Biology and Biochemistry, Irvine, CA 92697; (AAJ) University of California, Irvine, Department of Microbiology and Molecular Genetics, Irvine, CA 92697; (XSX) University of California, Irvine, Department of Computer Science, Irvine, CA 92697; (DHS,XSX) University of California, Irvine, Institute for Genomics and Bioinformatics, Irvine, CA 92697; (KM) European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SD United Kingdom

We report the results of 5'- and 3'-*cis*-regulatory element (CRE) discovery through a non-alignment based comparative genomics approach using one-to-one orthologs between *Anopheles gambiae*, *Aedes aegypti*, and *Culex quinquefasciatus*.

Mosquitoes contribute to the morbidity and mortality of human populations through their roles as vectors of diseases such as malaria and dengue fever. The past two decades have seen an increase in the application of modern research technologies and techniques to these vectors in an effort to better understand their basic biology as well as the cycle of disease transmission. The results have informed many mosquito based disease control strategies, including our focus: population replacement with genetically-engineered mosquitoes possessing pathogen refractory phenotypes.

An important component of engineering pathogen refractory phenotypes is controlling the expression of the transgene. We currently make use of multiple endogenous promoters to control transgenes; however, we aim to rationally adjust gene expression patterns to better suite our needs and perhaps eventually create fully synthetic promoters designed for specific applications. Toward these goals we must better understand the collection of CREs present in mosquitoes.

Comparative genomics provides a powerful tool for *in silico* prediction of CREs based on the observation that functional DNA sequences tend to be conserved. The well known strategy of globally aligning orthologous promoter sequences proved problematic for us due to sequence divergence and an elevated prevalence of mobile elements in the *Ae. aegypti* genome. We therefore used a word enumeration based method named MDOS (motif discovery using orthologous sequences). MDOS produces a "conservation z-score" for every *k*-sized word it encounters in the sequence data.

We applied MDOS to each pairwise combination of ortholog sets between the three mosquito genomes for both the 5'- and 3'-flanking DNA sequences (lengths of up to 2000 bp or 500 bp, respectively). Using the 5'-flanking regions, we identified 137 degenerate motifs, which we sorted into 18 groups based on sequence similarity. We provide support for these motifs with examples from the literature. Additionally, motifs are enriched in promoters of genes with specific temporal and spatial expression profiles following a bloodmeal in *An. gambiae*. Similar MDOS analyses were carried out using the 3'-flanking regions of orthologous genes. The highest scoring motifs were highly enriched with 7-mers that are complementary to the seed regions of *An. gambiae* miRNAs.

59

Colocalization of two novel neuropeptides in innervation of salivary gland of the black legged tick *Ixodes scapularis*

*Šimo, Ladislav; Park, Yoonseong

Kansas State University, Entomology Department, Manhattan, KS 66506

Salivary secretions of ticks are essential during the blood-feeding for manipulation and suppression of host defense responses, and might include key components in the transmission of pathogens. We have been interested in neural and hormonal mechanisms of the salivary secretion in the tick. We identified two novel neuropeptides in the nervous projections innervating salivary glands of the blacklegged tick *Ixodes scapularis* (Say, 1821). In our immunohistochemistry, which was later supplemented by *in situ* hybridization, these two neuropeptides myoinhibitory peptide (MIP) and SIFamide were colocalized in the protocerebral cells and their projections terminating on specific cells of salivary gland acini (types II and III). Immunoreactive substances were identified by matrix-assisted laser desorption/ionization-time of flight (MALDI TOF/TOF) analysis; a 1321.6 Da peptide with the sequence typical for MIP (ASDWNRLSGMWamide) and a 1395.7 Da SIFamide (AYRKPPFNGSIFamide), which is highly conserved among arthropods. Genes encoding these peptides were identified in the available *Ixodes* genome and expressed sequence tag (EST) database. In addition, the cDNA encoding the MIP pre-propeptide was isolated by rapid amplification of cDNA ends (RACE). In this report we describe the anatomical structure of specific central neurons innervating salivary gland acini. Our data provide evidences for the innervation of salivary gland by MIPs and SIFamide produced from the synganglion, thus leading a basis for functional studies of these two distinct classes of neuropeptides.

60

Towards a Seminal Fluid Proteome of the Mosquito, *Aedes aegypti*

Laura K. Sirot*, Michelle E. Helinski, Mari Kimura, José M.C. Ribeiro, Rebecca Poulson, Morgan C. McKenna, Mariana F. Wolfner, Laura C. Harrington

(LKS, MFW) Cornell University, Department of Molecular Biology and Genetics, Ithaca, NY 14853; (MEH, MK, RP, MCM, LCH) Cornell University, Department of Entomology, Ithaca, NY, 14853; (JMCR) National Institute of Allergy and Infectious Disease, National Institutes of Health, Laboratory of Malaria and Vector Research, Bethesda, MD, 20892

Aedes aegypti is the main vector of dengue fever and yellow fever viruses worldwide. An estimated 50 million people are infected with dengue each year and two thirds of the world's population is now at risk for infection. Therefore, development of safe and effective methods to control *Ae. aegypti* is of critical public health importance. Previous studies suggested that seminal fluid proteins (SFPs) are potent regulators of the fecundity, blood digestion, and mating and host-seeking behaviors of female *Ae. aegypti*. Consequently, exploration of SFPs may provide new tools and targets for control of the spread of this species.

Here, we present the results of our studies to identify the seminal fluid proteome of *Ae. aegypti*. Initially, we used bioinformatic comparisons to *Drosophila melanogaster* SFPs, and proteomic analysis of repro gland extracts to identify over 60 *Ae. aegypti* proteins that are predicted SFPs (Sirot et al. 2008 Insect Biochem. Molec. Biol.). Recently, we extended this analysis by identifying proteins that are transferred to females during mating, by using an isotope-labeling approach. Together our studies have identified 71 *Ae. aegypti* SFPs that are made in the male reproductive accessory glands and are transferred to females during mating. Some of the proteins we identified correspond to the products of previously-unidentified genes. Many of the classes of the SFPs we identified in *Ae. aegypti* overlap with those found in seminal fluid of *Drosophila* and other taxa (e.g., chaperones, proteases, protease inhibitors, and oxidoreductases); others show no similarities to previously-identified SFP classes. We are investigating the function of a subset of *Ae. aegypti* SFPs by RNA interference experiments.

61

The Inotocin System in Arthropods

*Stafflinger, Elisabeth; *Hansen, Karina K.; Hauser, Frank; Schneider, Martina; Cazzamali, Giuseppe; Williamson, Michael R.; Grimmelhuijzen, Cornelis J.P.

University of Copenhagen, Institute of Biology, Center for Functional and Comparative Insect Genomics, DK-2100 Copenhagen, Denmark

More than 20 years ago, an oxytocin/vasopressin like peptide, CLITNCPRGamide, was isolated from the locust *Locusta migratoria* by Proux and co-workers. However, no similar peptide could be identified in other insects, nor could its prohormone be cloned, or its physiological functions be established. We found that the recently sequenced genomes from both, the red flour beetle *Tribolium castaneum* and the parasitic wasp *Nasonia vitripennis*, contain a gene coding for an oxytocin/vasopressin-like peptide, identical to the locust peptide, which we named inotocin (for insect oxytocin/vasopressin-like peptide) and a gene coding for an inotocin G protein-coupled receptor (GPCR). However, this neuropeptide system could not be found in any other holometabolous insect with a completely sequenced genome (12 *Drosophila* species, the malaria mosquito *Anopheles gambiae*, the yellow fever mosquito *Aedes aegypti*, the silk worm *Bombyx mori*, and the honey bee *Apis mellifera*).

We cloned the *Tribolium* inotocin preprohormone and the inotocin GPCR and expressed the GPCR in CHO cells. This GPCR is strongly activated by low concentrations of inotocin (EC₅₀, 5 X 10⁻⁹ M), demonstrating that it is the inotocin receptor. Quantitative RT-PCR (qPCR) showed that in adult *Tribolium*, the receptor is mainly expressed in the head and much less in the hindgut and Malpighian tubules, suggesting that the inotocin system does not play a role in water homeostasis as does for example mammalian vasopressin. Surprisingly, qPCR also showed that the receptor is 30 X more expressed in the first larval stages than in adult animals, suggesting a potential role in development.

Additionally, we identified an oxytocin/vasopressin-like peptide and receptor in the recently sequenced genome from the water flea *Daphnia pulex* (Crustacea). *Daphnia pulex* belongs to the Branchiopoda and they are regarded to be the ancestor group of insects, from which insects evolved approximately 420 million years ago. The oxytocin/vasopressin hormonal system is an evolutionary old hormonal system, emerging before the split of Proto- and Deuterostomia about 700 million years ago. It is present in both invertebrates (as one system, e.g. inotocin in some insects) and vertebrates (as two systems, e.g. in mammals as vasopressin and oxytocin). This system has persisted until arthropods, where it is present in crustaceans, and in basal insects. In the holometabolous insects, however, it has been preserved only in the evolutionary lines leading to coleopterans and basal hymenopterans, whereas the other major orders have abandoned it. This loss of the oxytocin/vasopressin system must have occurred at least two times during evolution. Why some insects can live without this neuropeptide system and why others like *Tribolium castaneum* and *Nasonia vitripennis* need this system, still has to be clarified.

62

Avirulence Gene Discovery in the Hessian Fly

*Stuart, Jeff; Aggarwal, Rajat; Benatti, Thiago; Zhao, Chaoyang; Chen, Ming-Shun; Schemerhorn, Brandon; and Shukle, Richard (JS, RA, TB, CZ) Purdue University, Department of Entomology, West Lafayette, IN 47907; (M-SC) USDA-ARS, Kansas State University, Department of Entomology, Manhattan, KS 66506; (BS, RS) USDA-ARS, Purdue University, Department of Entomology, West Lafayette, IN 47907

Plant resistance (*R*) genes provide efficient and economical protection against many plant diseases and insect pests. Studies of plant pathogenic bacteria, fungi, and nematodes have determined that this protection depends on interactions between the *R* gene products and the products of "Avirulence" (*Avr*) genes in the pathogen. Loss-of-function mutations in specific *Avr* genes render their corresponding *R* genes completely ineffective. Moreover, these mutations have a selective advantage when the corresponding *R* genes are deployed in crop cultivars, and they are, therefore, a major limitation to the durability of the protection that the *R* genes afford. Plant *R* genes that provide defence against insects have been identified. In addition, elicitors of plant defense to insect attack have been discovered in the saliva of certain insects. However, the discovery of insect *Avr* genes has been elusive.

We used the small genome of the Hessian fly (*Mayetiola destructor*), its genetic tractability, and an FPC-based physical map of the Hessian fly genome to test the hypothesis that the Hessian fly has specific *Avr* genes that interact with specific *R* genes in wheat (*Triticum spp*). Hessian fly avirulence (the inability to survive and stunt wheat) and virulence (the ability to survive and stunt wheat) to three *R* genes in wheat (*H6*, *H9*, and *H13*) were examined. Based on our hypothesis that *Avr* genes exist in the Hessian fly, we expected to be able to map virulence to each *R* gene to a position in the genome corresponding to the length of a single bacterial artificial chromosome (BAC), and then, within these positions, identify the *Avr* genes using a combination of association mapping and functional analyses.

Virulence to *R* gene *H6* (*vH6*) was mapped to a 300-kb position near the centromere of Hessian fly chromosome X2. Fine-scale mapping was problematic due to the proximity of *vH6* to the centromere where rates of recombination are low. Virulence to *R* gene *H9* (*vH9*) was mapped to a 180-kb position near the telomere of the short arm of Hessian fly chromosome X1. Using association mapping, it was further localized to a 12-kb region. This position included only two small genes, each encoding a signal peptide. These genes were nearly identical and each was composed of only two exons. RT-PCR experiments found that both genes are expressed in the salivary glands of first instar Hessian fly larvae. Sequence comparisons discovered an insertion in the promoter region of one of the genes that appeared to confer the *vH9* phenotype. Virulence to *R* gene *H13* (*vH13*) was mapped to a 130-kb position near the telomere of the short arm of Hessian fly chromosome X2. It was further localized to a region containing a small signal peptide-encoding gene composed of two exons. RT-PCR experiments found that this gene is expressed in the salivary glands of avirulent first instar larvae, but is not expressed in virulent larvae. Sequence comparisons discovered three *H13*-virulence-associated and three *H13*-avirulence-associated alleles. The avirulence-associated alleles differed only with respect to the number of copies of a 14-residue motif. The virulence-associated alleles each had a different insertion that appeared to disrupt the transcription of the gene. RNA interference (RNAi) was used to test the putative avirulence function of the gene. A proportion (~38%) of genotypically avirulent larvae that were fed double stranded RNA copies of the gene were converted to an *H13*-virulent phenotype. In comparison, all control larvae had an *H13*-avirulent phenotype. Thus, we concluded that an *Avr* gene corresponding to the *H13* *R* gene in wheat had been identified. To our knowledge, this is the first *Avr* gene identified in an insect. Its discovery confirms that *Avr* genes exist in the genomes of certain insect pests of crop plants.

63

Study of Gene Expression in *Lycopersicon esculentum* Induced by *Helicoverpa zea* and *Manduca sexta* Herbivory

Thomas, Vickie.*; Musser, Richard O.

Western Illinois University, Biological Sciences Department, Macomb, IL 61455

Plants guard themselves against various pathogenic and herbivore attacks by implementing both general and specific induced defenses. Many plants have the ability to recognize different caterpillar-related elicitors in order to modify their defense response to target the specific herbivore. Studies have shown that some plants have a more consistent overall transcriptional response to generalist herbivores compared to specialist herbivores. In recent years, microarrays have proven to be useful tools in examining gene expression. In this experiment, microarrays were used to examine gene expression due to herbivory on tomato plants (*Lycopersicon esculentum*). *Helicoverpa zea* (generalist) and *Manduca sexta* (specialist) were allowed to feed on tomato leaves for a course of 48 hours. RNA was extracted and aRNA was labeled using Cy-3 and Cy-5 dyes. Approximately 482 genes were significantly altered ($P < 0.01$ ANOVA) by herbivory on tomato plants. Of these 482 genes, 166 are unknown. Of the 316 known genes, 251 were altered due to *H. zea* and 197 were altered due to *M. sexta* ($P < 0.05$ 2 tailed t-test). Of the 482 genes, *H. zea* herbivory stimulated 163 genes and suppressed 88 genes in comparison to non-wounded. *M. sexta* herbivory stimulated 119 genes and suppressed 78 genes in comparison to non-wounded. In comparison between *H. zea* and *M. sexta*, 22 % of the genes were significantly different. The results showed many similar transcriptional genes in categories such as defense, photosynthesis, biosynthesis, transcription, and stress; leading to the idea that many transcriptional genes are conserved between the two caterpillars.

64

***Anopheles* allocates 2% of its protein-coding genes for structural cuticular proteins. How come?**

Tsujimoto, Hitoshi; Cornman, R. Scott; Togawa, Toru; *Willis, Judith H.

(HT) University of Georgia, Department of Entomology, Athens, GA 30602; (RSC) USDA-ARS BARC, Beltsville, MD 20705; (TT) Nihon University, College of Humanities and Sciences, Tokyo 156-8550 Japan; (JHW) University of Georgia, Department of Cellular Biology, Athens, GA 30602.

Over 230 of the protein coding genes from *Anopheles gambiae* code for structural cuticular proteins. Data from qRT-PCR illustrating the temporal expression patterns for most of these genes are published or in press. We have examined the spatial distribution of a small number of cuticular protein genes belonging to several families with *in situ* hybridization using both conventional antisense RNA and locked NA (LNA) probes. Some genes in the CPR family belong to sequence clusters with linked almost identical genes, a feature at present found only in mosquitoes. Genes in sequence clusters are expressed in pharate stages when the animals are most vulnerable, suggesting that gene amplification facilitates rapid synthesis of a new cuticle. This postulate was supported when LNA probes allowed us to see that the genes in one sequence cluster had identical spatial (throughout the epidermis) as well as temporal (pharate pupa/young pupa) patterns of expression. Some single copy genes had very restricted expression and the results raised challenging issues about orthology. For example, *CPR132* shares 50% identity in its chitin binding domain but only 18% identity in the rest of the protein with a *Drosophila* gene (*cry*) that is expressed in the developing eye. The expression of *CPR132* is restricted to the eye. So, what properties of this protein make it suitable for forming an optical structure when its sequence is so different from *cry* and is it a genuine ortholog of *cry* with such limited sequence identity? *CPR60* is also expressed in the same region of the eye but its mRNA is found in other regions of the epidermis. *Anopheles* males obtain information about the wing beat frequency of conspecific females with Johnston's organ located in the second antennal segment. That organ is far larger in males than in females. Only one of the CPR genes, *CPR152*, had a clear difference in expression levels in males and females. *CPR152* has 73% identity in its chitin binding domain with *Drosophila* resilin. The remaining 80% percent of the mature protein is nothing like resilin; it has 20% histidine residues whereas *Drosophila* resilin has only one histidine. *CPR 152* is expressed exclusively in Johnston's organ, an ideal location for a flexible cuticle capable of sensing vibration. Some other genes for cuticular proteins are expressed over much of the epidermal surface, others have limited expression. Hence, the answer to why so many genes coding for structural cuticular proteins must be – in order to allow precise temporal and spatial patterns of expression of proteins that are required to be synthesized and secreted rapidly or in very precise locations to allow *Anopheles* to build a highly efficient flying and biting machine. (This work was supported by a grant from the National Institutes of Health AI55624.)

65

Comparison of Chitin Binding for Soft and Hard Cuticular Proteins

*Viau, David M.; Rebers, John E.

Northern Michigan University, Department of Biology, Marquette, MI 49855

Arthropod cuticles are composed of fibers of chitin embedded in a matrix of cuticular proteins. Many cuticular proteins include a conserved domain, the Rebers and Riddiford (R&R) consensus sequence, which can act as a chitin-binding domain. Although this domain is not present in all cuticular proteins, this motif is generally viewed as diagnostic for cuticular proteins, as noted in Pfam, Prosite, and other databases. Annotations of arthropod genomes often use this motif as a tool to help identify cuticular proteins. Detailed comparisons of proteins including the R&R consensus revealed three variations, generally designated as RR-1, RR-2, and RR-3. Proteins with the RR-1 variation are usually found in soft and flexible cuticle, whereas the RR-2 variation is usually found in proteins from hard and rigid cuticles. There is no obvious functional similarity among the small number of proteins that include the RR-3 variation. Proteins with each variation cluster together in sequence alignments. However, only a few studies have directly characterized cuticular protein-chitin interactions. This study compares the chitin binding of a soft (RR-1) cuticular protein from the tobacco hornworm *Manduca sexta* and a hard (RR-2) cuticular protein from the African malaria mosquito *Anopheles gambiae*. The soft cuticular protein MSCP14.6 binds chitin more weakly than the hard cuticular protein AGCP2b. Additional research is needed to determine if this difference in binding affinity is specific to these two proteins or if it reflects a general difference in the affinity of RR-1 and RR-2 proteins for chitin. Biochemical studies such as those reported here will help establish structure-function relationships for cuticular proteins.

66

Effects of silencing apoptosis regulatory genes on Sindbis virus replication and dissemination in *Aedes aegypti*

Wang, Hua; Penabaz, Taryn; Boyle, Dan; *Clem, Rollie J.

Division of Biology, Kansas State University, Manhattan, KS 66503

Sindbis virus (SINV) is a member of the *Togaviridae* family, and is the type species of the genus *Alphavirus*. SINV is associated with occasional outbreaks of disease in Africa, Europe and Asia. SINV is an arthropod-borne virus (arbovirus) which is normally vectored by mosquito species in the genus *Culex*, but is also capable of being vectored by the yellow fever mosquito, *Aedes aegypti*. Since *Ae. aegypti* is the major vector of yellow fever and dengue fever, and since SINV has been well characterized at the molecular level (including the development of gene expression systems based on the SINV genome), SINV infection of *Ae. aegypti* is often used as a model to study arbovirus-mosquito interactions. We are interested in determining whether apoptosis plays a role in the ability of arboviruses to replicate in mosquitoes. Genes which either positively or negatively regulate apoptosis in *Ae. aegypti* were silenced by RNA interference in adult female mosquitoes, and the mosquitoes were then fed a blood meal containing the SINV infectious clone p5'dsMRE16ic expressing green fluorescent protein (MRE/GFP). Effects were observed on the both the occurrence and intensity of expression of GFP in various mosquito tissues, depending on the gene silenced. Increased caspase activity and mosquito mortality were also observed following silencing of a negative regulator of caspase activity, even in the absence of virus infection. These results suggest that apoptosis may influence SINV replication in *Ae. aegypti*.

67

Assembly of DNA Components for Development of a Female Conditional-lethal Genetic System in New World Screwworm, *Cochliomyia hominivorax*

Wang, Ju; *Guerrero Felix; Rachinsky Anna

USDA-ARS, Knippling-Bushland US Livestock Insects Research Laboratory, Kerrville, Texas, 78028

The New World Screwworm (*Cochliomyia hominivorax*) is an important economic parasite that can cause great damage to domestic livestock and other warm-blooded animals. The female adult fly oviposits in open wounds of warm-blooded host animals and the developing larvae feed on raw flesh. New World Screwworm has been eradicated from the United States and Central America and a sterile release barrier is established in Panama to prevent re-infestation from South America. Our project is using molecular technologies to adapt a female conditional-lethal genetic system to New World Screwworm to increase the efficiency of the USDA Screwworm Eradication Program's sterile release program. One of the key elements of this project makes use of sex-specific differential splicing of an intron from the *transformer* gene (*tra*). Here we successfully cloned the first intron of *transformer* from Caribbean fruit fly (*Anastrepha suspense*). Two different sequences of the first intron of *tra* gene have been found and will be evaluated for potential roles in the female conditional lethal genetic system.

68

Macrophage migration Inhibitory Factor (MIF) in ticks; phylogeny and homology modeling from a parasitological perspective

*Wasala, N.B; Jaworski, D.C.

Oklahoma State University, Department of Entomology and Plant Pathology, 127 Noble Research Center, Stillwater, OK 74078

As blood feeding ectoparasites, ticks transmit various diseases to humans and domestic animals. Their prolonged blood feeding strategies on the host may enhance the pathogen transmission. Macrophage Migration Inhibitory Factor (MIF) is a pro-inflammatory cytokine and to date, the only "cytokine" identified from ticks. Homologs of MIF have also been identified in four Ixodid tick species (*Amblyomma americanum*, *Dermacentor variabilis*, *Haemaphysalis longicornis*, *Ixodes scapularis*). A MIF homolog has not been found in *Drosophila* genome or any other Dipteran genome. Our hypothesis is that MIF functions in tick blood meal acquisition and pathogen transmission. This is a bioinformatics-based comparison of tick MIF with MIFs from other arthropods (*Tribolium castaneum*, *Bombyx mori*, *Acyrtosiphon pisum*, *Maconellicoccus hirsutus*) and selected parasitic nematodes (*Onchocerca volvulus*, *Brugia malayi*, *Ascaris suum*, *Trichuris trichiura*, *Trichinella spiralis*), combined with a homology modeling study to identify and understand a potential model for structure conservancy in MIF. A PSI-BLAST search was performed using *A. americanum* MIF as the query to retrieve MIF sequences from NCBI database for arthropods and parasitic nematodes. From the resulting sequences a multiple sequence alignment (MSA) and phylogenetic analysis of MIF was performed using ClustalW2. A phylogenetic tree was constructed using the Tree Puzzle Version 5.2. Next, homology modeling was performed using SwissModel first approach modeling request (<http://swissmodel.expasy.org/>) and *A. americanum* MIF as the query sequence. The resulting model was used to perform a structure conservation analysis (Bayesian method) at Consurf Server (<http://consurf.tau.ac.il/>). MSA showed a high level of sequence conservation in MIF among different species and two common patterns were identified among the studied sequences. The CXXC pattern was observed in some but not all nematode sequences, but varied in ticks and in other arthropods. CXXC is the active-site motif of thiol-disulfide oxidoreductases and could grant protein disulfide oxidoreductase activity to MIF. In addition, all the ticks and some insects (*Bombyx mori*, *Acyrtosiphon pisum*, *Maconellicoccus hirsutus*) shared the function unidentified CXXA pattern. Additionally, another highly conserved amino acid sequence, possibly tick specific, was observed in the alignment. The percentage identity of this region ranged from 73- 91% among ticks and it was from 27- 50% in other arthropods and nematodes. The phylogenetic tree analysis predicted tick MIF sequences were more related to the nematode MIF sequences than to the other arthropod MIFs. The protein identified with Swiss Model was *Xenopus laevis* (PDB ID: 1uiz) MIF with a 42.609% identity to the query sequence. Our analysis inferred the highest conservation of amino acids as Pro 2, Thr 8/113, Asn 9/73, Pro 34/56, Ser 64 and Ile 97. All of these residues were found to be conserved among the tick, insect and nematode MIF sequences tested for this study. Both parasitic nematodes and ticks are dependent on their mammalian host to acquire nutrients and to survive from one life stage to the next. A high level of sequence conservation and phylogenetic proximity of MIF in these two groups is likely to be a result of selective pressure on parasitism during their evolution. While parasitic nematodes may use MIF for delaying immune recognition by lymphocytes during infection, ticks may require MIF to modify the cellular infiltrate in the feeding lesion in order to achieve an adequate blood meal. It appears that parasite-specific MIFs may be essential to evading host MIF responses to parasitism.

69

Prevalence of the Microsporidian Parasite *Nosema ceranae* in Turkish Honey Bees

Whitaker, Justin T.; Szalanski, Allen L.

University of Arkansas, Department of Entomology, Fayetteville, AR 72701

Nosema ceranae, which appears to have crossed over from the Asiatic honey bee *Apis ceranae* to the European honey bee *Apis mellifera*, is considered to a contributing factor to Colony Collapse Disorder, or CCD. *N. ceranae* rates of infection, virulence with regard to specific bee lineages, and geographic origin are among the unknowns regarding *N. ceranae*'s worldwide prevalence. Using a 16S PCR primer specific for *N. ceranae*, we conducted a broad survey for *N. ceranae* in Turkey using 'C' and 'O' lineage honey bee colonies from Artvin, Bursa, Hatay, Giresun, Gökçeada, Kirklareli, and Sakarya provinces. DNA sequencing of the 16S amplicon revealed that it is similar to *N. ceranae* from Europe and North America.

70

A recently sequenced model organism, *Nasonia vitripennis*, reveals new insights into developmental diapause

Wolschin, Florian ; Allan, Benjamin; Gadau, Jürgen

(FW, BA, JG) Arizona State University, School of Life Sciences, Tempe, AZ 85287;

(FW) Norwegian University of Life Sciences, Dept. of Biotechnology, Ås N-1432, Norway

Insect diapause is an alternative life-history strategy used to increase longevity and survival in harsh environmental conditions. Even though some aspects of diapause are well investigated, broader scale studies that elucidate the global metabolic adjustments required for this remarkable trait, are rare.

In order to better understand the metabolic changes during early insect diapause, we used a shotgun proteomics approach on early diapausing and non-diapausing larvae of the recently sequenced hymenopteran model organism *Nasonia vitripennis*. Our results deliver insights into the molecular underpinnings of diapause in *Nasonia* and corroborate previously reported diapause-associated features for invertebrates, such as a diapause-dependent abundance change for heat shock proteins, storage proteins, and ferritins. Furthermore, we observed a diapause-dependent switch in enzymes involved in glycerol synthesis and a vastly changed capacity for protein synthesis and degradation. The abundance of structural proteins and proteins involved in protein synthesis decreased with increasing diapause duration, while the abundance of proteins likely involved in diapause maintenance increased. Only few potentially diapause-specific proteins were identified suggesting that diapause in *Nasonia* relies to a large extent on a modulation of pre-existing pathways.

Studying a diapause syndrome on a proteomic level rather than isolated pathways or physiological networks, has proven to be an efficient and successful avenue to understand molecular mechanisms involved in diapause.

71

Characterization of Trypsin- and Chymotrypsin-like cDNAs from the Gut of European Corn Borer (*Ostrinia nubilalis*)

Jianxiu Yao, Chitvan Khajuria, Lawrent L. Buschman, Ming-Shun Chen, and Kun Yan Zhu Department of Entomology, Kansas State University, 123 West Waters Hall, Manhattan, KS 66506

Trypsin- and chymotrypsin-like serine proteases are the main digestive enzymes in lepidopteran midguts, and are considered the major proteases involved in *Bacillus thuringiensis* (Bt) protoxin activation or toxin detoxification. A previous report indicated that the reduced level of transcript T23 was associated with European corn borer (*Ostrinia nubilalis*) (ECB) resistance to Dipel® Bt formulations. Here we describe the sequence and the partial characterization of 17 trypsin- and 14 chymotrypsin-like proteases, which were identified from ECB gut-specific cDNA libraries. These putative trypsins and chymotrypsins have structural characteristics of serine proteases, including the catalytic triad of histidine, aspartic acid and serine, six conserved cysteine residues; and other amino acids involved in substrate specificity. They also shared a high percent similarity with those of other lepidopteran species- *Heliothis virescens*, *Manduca Sexta*, and *Helicoverpa armigera* based on bootstrap amino acid consensus analysis. RT-PCR results indicated that all of these proteases were highly expressed in the midgut of ECB; however, there existed different expression profiles in other tissues, such as foregut, hindgut, haemolymph, fatbodies, silk glands, Malpighian tubules, and carcass. In addition to their roles in digestion, some of these proteases might also have other important physiological functions in other tissues.

72

Molecular Properties and Expression Patterns of Two Chitin Synthases in African Malaria Mosquito (*Anopheles gambiae*)

Zhang, Xin; Zhang, Jianzhen; Zhu, Kun Yan

Department of Entomology, Kansas State University, Manhattan, KS 66506

Chitin synthases (CHSs) are important enzymes in arthropods and other chitin-containing organisms. However, current knowledge on chitin synthases, especially their structures, functions, and regulations in insects is very limited. We report molecular properties and expression patterns of two chitin synthases (AgCHS1 and AgCHS2) in African malaria mosquito (*Anopheles gambiae*). *AgCHS1* and *AgCHS2* genes were predicted to encode proteins of 1,578 and 1,610 amino acid residues, respectively. Their deduced amino acid sequences showed high similarities to those of other insect species. Both genes were expressed in egg, larval, pupal and adult stages but *AgCHS2* was expressed at relatively low levels, particularly during the larval stages as examined by RT-PCR and real-time quantitative PCR. Relatively high expression levels were detected in the carcass followed by the foregut and hindgut for *AgCHS1*, and the foregut followed by the midgut for *AgCHS2*. Preliminary analysis using immunohistochemistry methods in the pupae localized *AgCHS1* in the eyes and epidermal cells of the adult integument newly developed underneath of the pupal cuticle. In contrast, *AgCHS2* was localized in the eyes and abdominal inter-segmental regions of pupal cuticle. Further research is needed to elucidate biological implications of such localized expression patterns of these enzymes in the mosquito.

73

A comparative analysis of serpin genes in the silkworm genomeZhao, Picheng¹, Zen Zou¹, Hua Weng², Kazuei Mita³, and Haobo Jiang¹¹ Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078² Department of Biochemistry and Molecular Biology, Oklahoma State University, Stillwater, OK 74078³ National Institute of Agrobiological Science, Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan

Serine protease inhibitors (serpins) are a superfamily of proteins, most of which control protease-mediated processes by forming covalent complexes with their cognate enzymes. Sequencing of the silkworm genome provides us an opportunity to investigate serpin structure, function, and evolution at the genome level. There are thirty-four serpin genes in *Bombyx mori*, similar in number to *Drosophila melanogaster* (29) or *Tribolium castaneum* (31) but many more than *Anopheles gambiae* (17) or *Apis mellifera* (7). Most of the silkworm serpins contain a secretion signal peptide, and six are highly similar in sequence to their *Manduca sexta* orthologs that regulate innate immunity. Thirteen serpin genes are predicted to encode functional inhibitors, most of which regulate serine proteases with trypsin-like specificity. Three alternative exons in serpin1 gene and four in serpin28 encode a variable region including the reactive site loop (RSL). Splicing of serpin2 pre-mRNA generates minor variations in the RSL of serpin2A, 2A', and 2B. We have identified six phylogenetic groups in this family of proteins. Many Group A genes contain alternative exons to generate RSL variation. Some members of Groups B, C, and D inhibit proteases involved in the proteolytic activation of prophenoloxidase and pro-spätzle, leading to melanization and Toll-induced synthesis of antimicrobial proteins. All eleven members of Group F have evolved from a lineage-specific family expansion. We have reconstructed the evolutionary pathway of seven Group C genes. Reverse transcription-PCR revealed an increase in the mRNA levels of serpin1, 3, 5, 6, 9, 12, 13, 25, 27, 32 and 34 in both fat body and hemocytes of larvae injected with bacteria. The serpin23 transcript level increased in fat body after an immune challenge whereas the serpin4, 28 and 31 mRNA became more abundant in hemocytes. These results suggested that the silkworm serpins regulate innate immune responses.

74

De novo genome assemblies from the next generation sequencing data.

*Zimin, Aleksey V; Yorke, James A

University of Maryland, College Park, MD 20742, USA

Next generation sequencing technologies such as 454 sequencing by Roche, ABI SOLiD or Illumina sequencing provide large amounts of genomic data quickly and at significantly lower cost, compared to traditional Sanger sequencing. As the read lengths increase, one can apply the traditional genome assembly software, such as Celera Assembler to create *de novo* assemblies of the genomes sequenced utilizing the next generation technologies. We applied the Celera Assembler to the sequence data for domestic turkey, produced by Roche, VBI and USDA to create a *de novo* assembly. Preliminary results obtained from the early data set that includes 5x coverage by the 454 reads and 10x coverage by Illumina reads indicate, that the Celera Assembler can be used to create an assembly with scaffold N50 size of over 200Kb and contig N50 size of over 6Kb.

75

Comparative analysis of melanization cascade genes in the mosquitoes

*Zou, Zhen; Shin, Sang Woon; and Raikhel, Alexander S.

Department of Entomology and Institute for Integrative Genome Biology, University of California, Riverside, California 92521

Culex quinquefasciatus is an important disease vector of filariasis and West Nile virus. Recently, the genome of *Cu. quinquefasciatus* has been sequenced. As a part of genome project in the mosquito, we have annotated immune gene families, which then were compared to those from other mosquito and insect species. Based on sequence similarity to orthologues known to be involved in immune reactions, around 350 immunity-related proteins have been identified. Here, we present characterization of immune gene families linked to melanization activation pathway in *Cu. quinquefasciatus*. Melanization is an important defense mechanism in arthropods, in which prophenoloxidase (PPO) is a key enzyme. The activation of PPO is regulated by a serine protease cascade, which is composed from clip domain serine proteases (cSP), clip domain serine protease homologs (cSPH), and serpins. Extensive repertoire of melanization genes of *Cu. quinquefasciatus* consists of 10 PPOs, 41 CLIPB subfamily of cSPs, and 14 CLIPA subfamily of cSPHs, which is similar to that of *Aedes aegypti* (10 PPOs, 41 CLIPB-cSPs, and 12 CLIPA-cSPHs). *Anopheles gambiae* also has a large number of melanization genes (9 PPOs, 20 CLIPB-cSPs, and 15 CLIPA-cSPHs). This represents a mosquito specific expansion of melanization component genes, when compared to *Drosophila* (3 PPOs, 14 CLIPB-cSPs, and 11 CLIPA-cSPHs) and other insects such as *Tribolium castaneum* (2 PPOs, 15 CLIPB-cSPs, and 15 CLIPA-cSPHs). Functional analyses have identified a subgroup of the CLIPB subfamily of cSP, named Melanization Proteases (MPs). Further phylogenetic analysis revealed that mosquito MPs belongs to three clades. Dipteran PPO genes are divided into four distinct clades. Functional divergence evaluation for PPO gene and MP gene families has provided favorable evidence for site-specific altered selective constraints after gene duplication. Further maximum-likelihood analysis of codon substitution shows that although purifying selection is the predominant force during the evolution, point mutation with positive selection still has contributions to the individual clade of the PPO family in mosquitoes.