Varroa Genome Project

Summary, Jan. 7, 2010

Varroa Genome Sequencing Consortium

Contacts:

Jay D. Evans, USDA-ARS Bee Research Lab Beltsville, MD USA

Jay.evans@ars.usda.gov, 301-504-5143

Christine Elsik, Georgetown University, Washington DC, USA

chris.elsik@gmail.com, 202-687-4485

Lelania Bourgeois, USDA-ARS Honey Bee Breeding, Genetics, and Physiology Lab, Baton Rouge, LA USA

Lanie.bourgeois@ars.usda.gov, 225-767-9288

Greg Hunt, Purdue University, West Lafayette, IN USA

ghunt@purdue.edu, 765-494-4605

Title: Genome characterization of the mite *Varroa destructor*, the primary pest of honey bees

Non-technical summary

Varroa mites remain the central pest and economic burden for beekeepers, having a strong impact on the production of bee products and the availability of honey bees for pollination. This genome project will provide direct tools for the control of Varroa mites through directed bee breeding, specific targeting of Varroa genes, and the development of baits and repellents that can be used to deter mites. The project will start delivering genomic information to public databases in year one and will be committed to public release of all data as it is generated. Milestones will be a sequenced and analyzed mite genome, the identification of mite genetic variants that affect virulence and acaricide resistance, and the determination of mite biology as relates to finding and exploiting bee hosts. By developing a large consortium this project will also bring fresh talents and resources into the fight to control Varroa mites.

Goals/Objectives: The parasitic mite *Varroa destructor* is the central pest of domesticated and free-living honey bees, causing direct impacts on bee health as well as indirect effects caused by vectoring viruses and other bee disease agents. Genomic information for this pest can be used to develop gene-based control strategies, determine weak points for conventional (chemical) controls, identify targets for biological control, and show how mites find honey bee hosts and reproduce in resistant and susceptible bee lineages. We will use emerging high-throughput sequencing techniques to sequence, assemble and annotate (describe key traits of) the genome of this mite. Sequencing will be completed in the first year of the project and released publically into genomic databases. Gene expression work aimed at discovering mite weaknesses will be completed in the second year and targets for mite control will be identified throughout the project. We will increase project impacts through a cost-effective partnership across existing sequencing and informatics centers and by choosing appropriate sequencing techniques for specific questions. We will leverage this project by engaging ca. 40 academic and governmental researchers in a volunteer consortium. The resulting insights will be used to improve honey bee health and crop pollination.

Approach: We will carry out genomic sequencing exceeding 60X coverage of the 565 million base-pair genome of Varroa destructor with an 'optimal' mix of straight and end-pair SOLEXA genomic sequencing reads, followed by genome assembly and pipeline annotation with resources provided by the US National Institutes of Health, NCBI. Specifically, we will carry out three SOLEXA flow cell runs with straight (72 bp) reads generated from Varroa genomic

DNA templates and one SOLEXA endpair run (72 bp) generated from the same source material. These will be complemented by longer-insert paired runs, either or 8kbp ILLUMINA or 6 and 20 kbp Roche. Transcriptome surveys will be carried out using ROCHE Titanium pyrosequencing and the SOLEXA platform, focused on 1) nymphal development, 2) host finding (tarsal library), 3) immune responses (gut with and without virus infection), 4) gut microbes, and 5) iniaition of reproduction. Single-nucleotide polymorphisms (SNPs) and protein polymorphism will be discovered using the ABI SOLiD platform on the mite transcriptome. These data will be aligned with homologous sequence data from the genome reads and from other transcriptome reads to give an abundance of SNPs. Development and testing of a canonical gene set will be followed by integration and posting of emerging data via Beebase, NCBI, and other public databases

Timetable for major landmarks:

Project Goal	2010		2011		2012		2013	
Genomic Sequencing	■	■						
RNA collection for transcript work		■						
Genome Assembly		■	■					
Transciptomic Pyrosequencing			■					
ABI SOLiD sequencing			■	■	■			
Consensus Gene Set				■				
Expression microarrays					■	■		
SNP surveys					■	■	■	
Publication and Beebase hosting				■	■	■	■	
Genbank Release of Sequence data	■	■	■	■	■	■	■	

Tentative Annotation and Translation Groups: 01/7/2010

<u>**Stakeholder Advisors**</u>

Randy Oliver (Scientific Beekeeping, Grass Valley, CA)

Dan Weaver (Bee Weaver, Navasota, TX)

David Mendes (American Beekeeping Federation, FL)

Chris Heintz (Almond Board CA)

Troy Fore (ABF, GA)

<u>**Sampling and stock selection (Leads: Greg Hunt, Purdue, Denis Anderson, CSIRO)**</u>

U.S. *Varroa* virulence and resistance (Hunt)

V. destructor and beyond (Denis Anderson, CSIRO)

European clone (Maria Navajas, INRA, France

Beltsville stock (Jeff Pettis, USDA-ARS)

Sequencing (Jay Evans, USDA-ARS)

ILLUMINA genomic/end pair runs (Tad Sonstegard/Curt van Tassell, USDA, Hugh Robertson, UI)

Roche transcript/ILLUMINA transcript (Evans)

Data aggregation/transfer to Elsik/Genbank (Scott Cornman, USDA-ARS)

Assembly, Genomic and transcriptomic (Christine Elsik, Georgetown)

Uber-consultant (Stephen Richards, Baylor College Med)

Chromosome structure (Spencer Johnston, TAMU)

Heterochromatin (Spencer Johnston)

Conserved and novel repeats

Transposable elements

Telomeres and Telomerase (Hugh Robertson, UI)

Microsatellites

Official Gene Set (Elsik, Evans, Ryszard Malezska, ANU)

Leverage with Ixodes (Catherine Hill, Purdue)

NCBI/GNOMON pipeline (Terrence Murphy, NCBI)

Transcript analysis and array expression (Lanie Bourgeois, USDA-ARS Baton Rouge)

Small RNAs (Elsik)

Automated prediction (Juan Anzola, Christine Elsik)

Small RNA SOLEXA lane (Evans)

Development (Maleszka)

Embryonic development

Yellow family (Maleszka)

Dnmt's (Maleszka/Robertson)

Cuticular Proteins (Scott Cornman, USDA-ARS)

Host recognition (Christina Grozinger, PSU, Yves Le Conte, INRA, Karl Gordon, ANU Peter Teal, USDA-ARS Gainesville)

Odorant Binding Proteins

Gustatory Receptors (Harland Patch, PSU, Hugh Robertson, UI?)

Pheromones/kairomones (Teal, Karl Gordon ANU)

Reproduction (Bourgeois, Zachary Huang, Michigan State Univ, and ??)

Ovariole activation (Ben Oldroyd, U Sydney, Kevin Donohue, Syngenta?)

Non-reproductive vs. Reproductive mites (Jeff Harris, USDA-BR)

Sex Determination

Mating and Gametogenesis

Male and female reproductive proteins and hormones (Grozinger/Patch?)

<u>Neurobiology (Christina Grozinger, PSU, Gene Robinson, UI?)</u>

      Receptors and Signaling molecules (ion channels, neurotransmitters, neuropeptides, biogenic amines)

      Mite Behavior (Yves Le Conte)

<u>Digestion and Metabolism</u>

      Proteases

      Sialome (Wayne Hunter, USDA-ARS, Diana Cox-Foster, PSU?)

<u>Microbes and Defenses (Jay Evans, USDA-ARS)</u>

      Viruses, transmission and replication (Judy Chen, USDA-ARS)

      Immune Pathways (Evans)

      Bacteria and DNA viruses (Cornman/Evans)

      Fungal associates (Boucias)

<u>Acaricide/xenobiotic resistance (</u>Reed Johnson, Univ. Nebraska ?<u>)</u>

      Ion Channels (Z. Huang, MSU?)
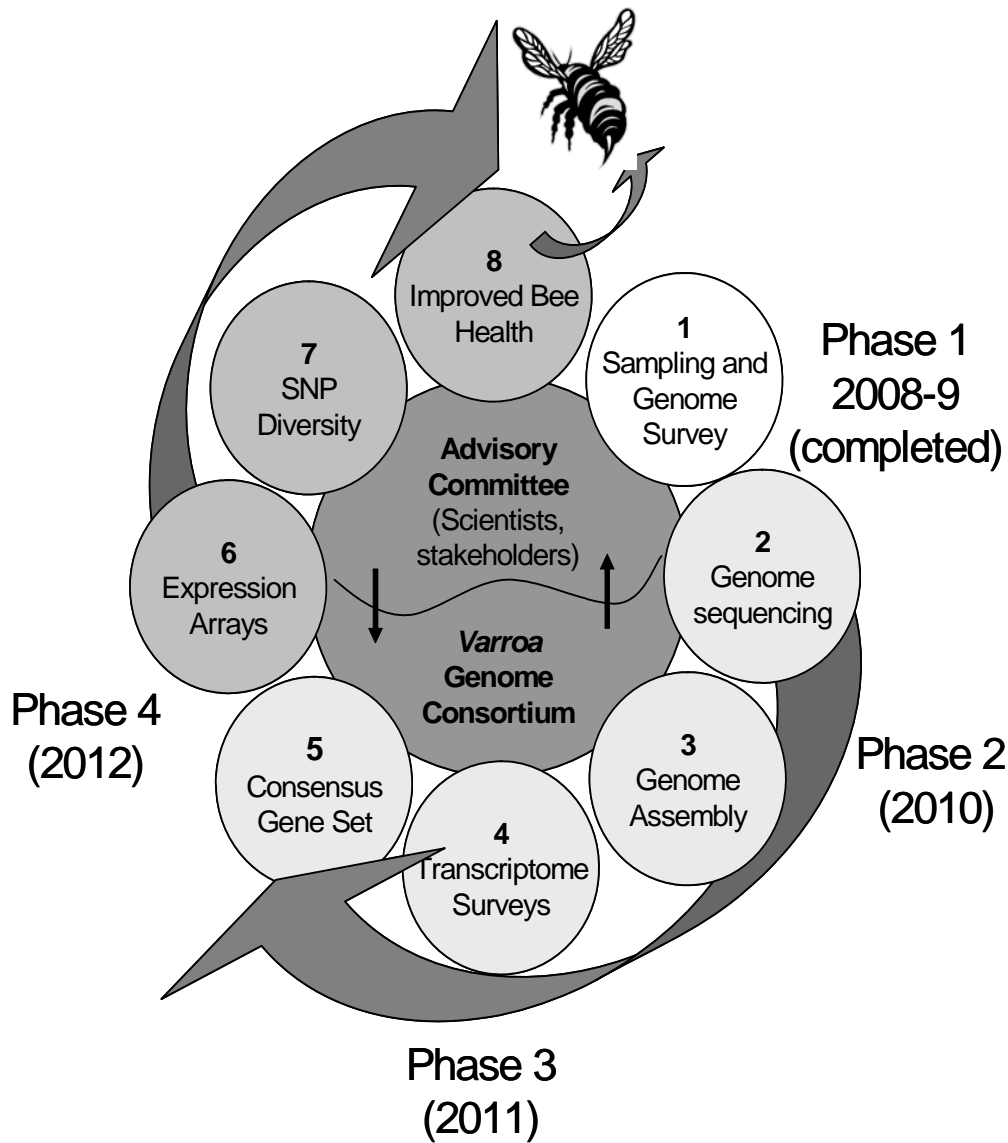
      Cp450's (Reed Johnson, Univ. Nebraska, Gordon?)

Bee side (Lead?)

      Hygienic behavior (Marla Spivak, U Minn)

      Integrated resistance (Tom Rinderer, USDA-ARS Baton Rouge)

      Quantitative genetics (Steve Sheppard, WSU)

Flow diagram for proposed *Varroa destructor* Genome Project.