

FIVE INSECT GENOMES

Introduction to Genomics 0497-485
Fall 2019

Presented by:

Mariam M. Hasan

Baty M. Abdilfatah

Munera A. Alkhudher

Habiba S. Alshammari

Nouf M. Boqurais



Presentation Outline

1. The genome papers
2. Facts about the species
3. Sequencing strategy
4. Sequencing method
5. Assembly
6. Annotation
7. Interesting facts
8. Questions

The Genome Papers

McKenna et al. *Genome Biology* (2016) 17:227
DOI 10.1186/s13059-016-1088-8

Genome Biology

RESEARCH

Open Access



Genome of the Asian longhorned beetle (*Anoplophora glabripennis*), a globally significant invasive species, reveals key functional and evolutionary innovations at the beetle–plant interface

Duane D. McKenna^{1,2*}, Erin D. Scully³, Yannick Pauchet⁴, Kelli Hoover⁵, Roy Kirsch⁴, Scott M. Geib⁶, Robert F. Mitchell^{7,8}, Robert M. Waterhouse^{9,10}, Seung-Joon Ahn⁴, Deanna Arsalan¹¹, Joshua B. Benoit¹², Heath Blackmon¹³, Tiffany Bledsoe², Julia H. Bowers¹⁴, André Busch⁴, Bernarda Calla⁶, Hsu Chao¹⁵, Anna K. Childers¹⁶, Christopher Childers¹⁷, Dave J. Clarke¹, Lorna Cohen¹¹, Jeffery P. Demuth¹³, Huyen Dinh¹⁵, HarshaVardhan Doddapaneni¹⁵, Amanda Dolan¹⁸, Jian J. Duan¹⁹, Shannon Dugan¹⁵, Markus Friedrich²⁰, Karl M. Glastad²¹, Michael A. D. Goodisman², Stephanie Haddad¹, Yi Han¹⁵, Daniel S. T. Hughes¹⁵, Panagiotis Ioannidis⁹, J. Spencer Johnston²², Jeffery W. Jones²⁰, Leslie A. Kuhn²³, David R. Lance²⁴, Chien-Yueh Lee^{17,25}, Sandra L. Lee¹⁵, Han Lin^{17,25}, Jeremy A. Lynch¹¹, Armin P. Moczek²⁶, Shwetha C. Murali¹⁵, Donna M. Muzny¹⁵, David R. Nelson²⁷, Subba R. Palli²⁸, Kristen A. Panfilio²⁹, Dan Pers¹¹, Monica F. Poelchau¹⁷, Honghu Quan¹¹, Jiaxin Qu¹⁵, Ann M. Ray³⁰, Joseph P. Rinehart¹⁶, Hugh M. Robertson³¹, Richard Roehrdanz¹⁶, Andrew J. Rosendale¹², Seunggwang Shin¹, Christian Silva¹⁸, Alex S. Torson¹⁴, Iris M. Vargas Jentzsch²⁹, John H. Werren¹⁸, Kim C. Worley¹⁵, George Yocum¹⁶, Evgeny M. Zdobnov⁹, Richard A. Gibbs¹⁵ and Stephen Richards^{15*}

Vol 443|26 October 2006|doi:10.1038/nature05260

nature

ARTICLES

Insights into social insects from the genome of the honeybee *Apis mellifera*

The Honeybee Genome Sequencing Consortium*

Cong et al. *BMC Genomics* (2015) 16:639
DOI 10.1186/s12864-015-1846-0



RESEARCH ARTICLE

Open Access



Skipper genome sheds light on unique phenotypic traits and phylogeny

Qian Cong², Dominika Borek², Zbyszek Otwinowski² and Nick V. Grishin^{1,2*}

nature
COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-018-03281-1

OPEN

The genomic and functional landscapes of developmental plasticity in the American cockroach

Sheng Li¹, Shiming Zhu¹, Qiangqiang Jia¹, Dongwei Yuan^{2,3}, Chonghua Ren¹, Kang Li¹, Suning Liu¹, Yingying Cui¹, Haigang Zhao^{2,3}, Yanghui Cao², Gangqi Fang^{2,3}, Daqi Li⁴, Xiaoming Zhao⁴, Jianzhen Zhang⁴, Qiaoyun Yue⁵, Yongliang Fan⁶, Xiaoqiang Yu¹, Qili Feng¹ & Shuai Zhan²

Draft genome of the globally widespread and invasive Argentine ant (*Linepithema humile*)

Christopher D. Smith^{a,1}, Aleksey Zimin^b, Carson Holt^c, Ehab Abouheif^d, Richard Benton^e, Elizabeth Cash^f, Vincent Croset^g, Cameron R. Currie^{h,i}, Eran Elhaik^j, Christine G. Elsik^k, Marie-Julie Fave^l, Vilaiwan Fernandes^m, Jürgen Gadauⁿ, Joshua D. Gibson^o, Dan Graur^p, Kirk J. Grubbs^q, Darren E. Hagen^r, Martin Helmkamp^s, Jo-Anne Holley^t, Hao Hu^u, Ana Sofia Ibarra-Viniegra^v, Brian R. Johnson^w, Reed M. Johnson^x, Abderrahman Khila^y, Jay W. Kim^z, Joseph Laird^{aa}, Kaitlyn A. Mathis^{ab}, Joseph A. Moeller^{ac}, Monica C. Muñoz-Torres^{ad}, Marguerite C. Murphy^{ae}, Rin Nakamura^{af}, Surabhi Nigam^{ag}, Rick P. Overson^{ah}, Jennifer E. Placek^{ai}, Rajendhran Rajakumar^{aj}, Justin T. Reese^{ak}, Hugh M. Robertson^{al}, Chris R. Smith^{am}, Andrew V. Suarez^{ao}, Garret Suen^{ap}, Elissa L. Suhr^{aq}, Shu Tao^{ar}, Candice W. Torres^{as}, Ellen van Wilgenburg^{at}, Lumi Viljakainen^{au}, Kimberly K. O. Walden^{av}, Alexander L. Wild^{aw}, Mark Yandell^{ax}, James A. Yorke^{ay}, and Neil D. Tsutsui^{az}

Departments of ^aBiology and ^cComputer Science, San Francisco State University, San Francisco, CA 94132; ^bInstitute for Physical Science and Technology and ^fDepartment of Mathematics, University of Maryland, College Park, MD 20742; ^dDepartment of Human Genetics, Eccles Institute of Human Genetics, University of Utah School of Medicine, Salt Lake City, UT 84142; ^eDepartment of Biology, McGill University, Montreal, QC, Canada H3A 1B1; ^gCenter for Integrative Genomics, University of Lausanne, CH-1015 Lausanne, Switzerland; ^hSchool of Life Sciences, Arizona State University, Tempe, AZ 85287; ⁱDepartment of Bacteriology and ^jDepartment of Energy Great Lakes Bioenergy Research Center, University of Wisconsin, Madison, WI 53706; ^kThe Johns Hopkins University School of Medicine, Baltimore, MD 21205; ^lDepartment of Biology, Georgetown University, Washington, DC 20057; ^mDepartment of Biology and Biochemistry, University of Houston, Houston, TX 77204; ⁿDepartment of Entomology and ^oInstitute of Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801; ^pDepartment of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720-3114; ^qDepartment of Biology, Earlham College, Richmond, IN 47374; and ^rDepartment of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853

Edited* by Gene E. Robinson, University of Illinois, Urbana, IL, and approved January 11, 2011 (received for review June 17, 2010)

1



THE AMERICAN COCKROACH (*Periplaneta americana*)

American Cockroach: Facts

- The American cockroach was introduced to the U.S. from Africa in the early 16th century.
- It prefers indoor environments with access to food. It can be found outdoors in moist, shady, and warm places.
- It eats different types of substances but prefers fermenting foods.
- It can carry disease causing organisms and may trigger allergic reactions with some individuals.
 - The largest body size is 5.3 cm.
 - Molts 6-14 times before metamorphosis.
 - Life cycle is approximately 700 days.
 - It reproduces periodically up to 600 days.



American Cockroach Genome

- **Specimen**

- Genome source = single female adult.

- **Sequencing strategy**

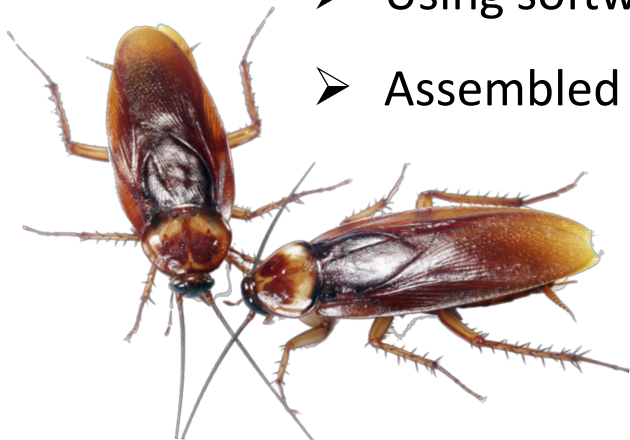
- Whole genome shotgun (WGS).

- **Sequencing method**

- Illumina.

- **Genome assembly**

- Using softwares (Discover Denovo, SSPACE, SOAP denovo).
- Assembled genome size = 3,380 Mb.



American Cockroach Genome

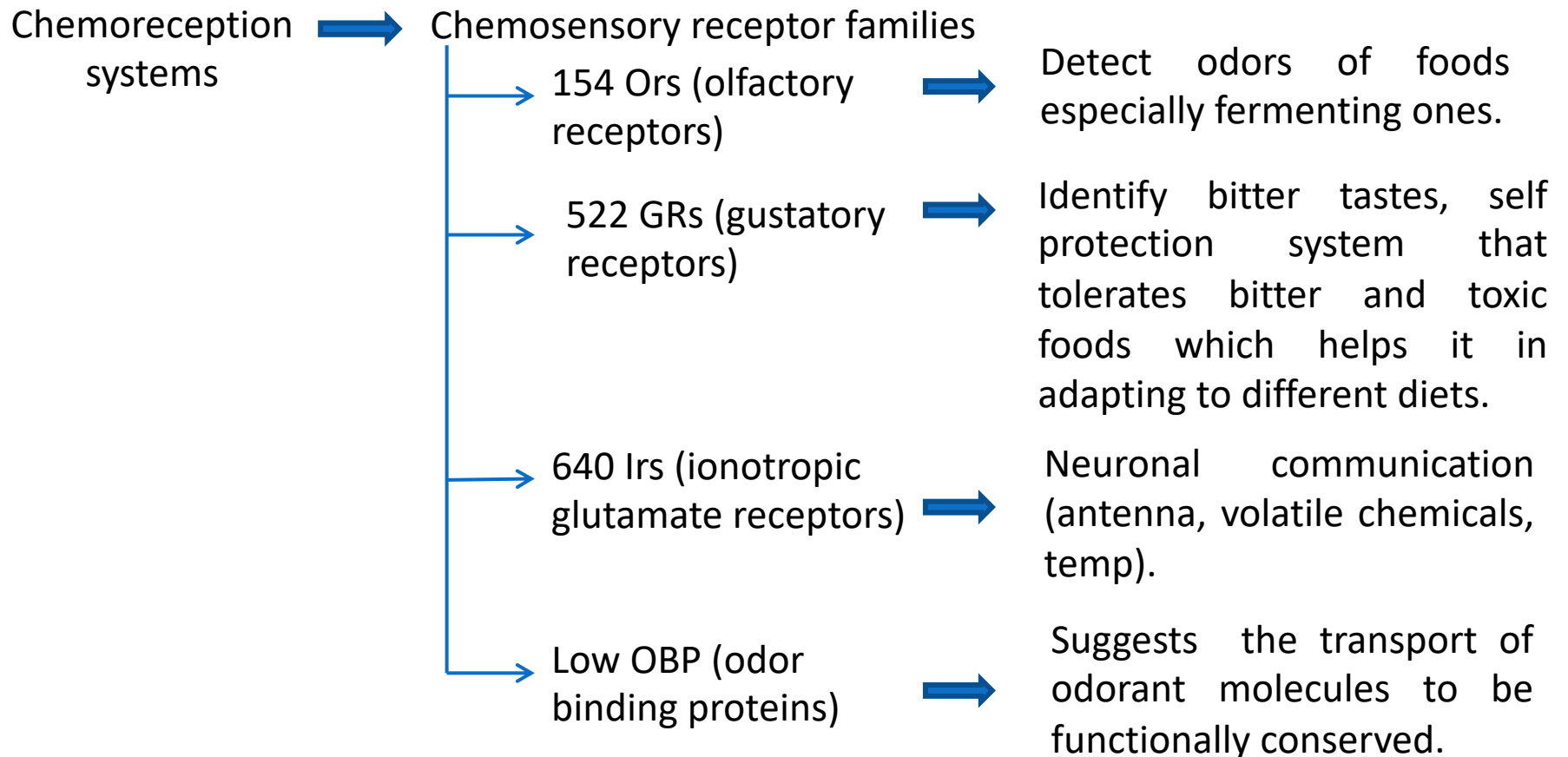
- **Annotation**

- Repeats including TE were identified using Repeat Masker v 4.0.5.
- Contig N50 of 21kb and scaffold N50 of 333Kb.
- Non-coding % = $100 - 0.88 = 99.12\%$.
- GC % = 35.7%.



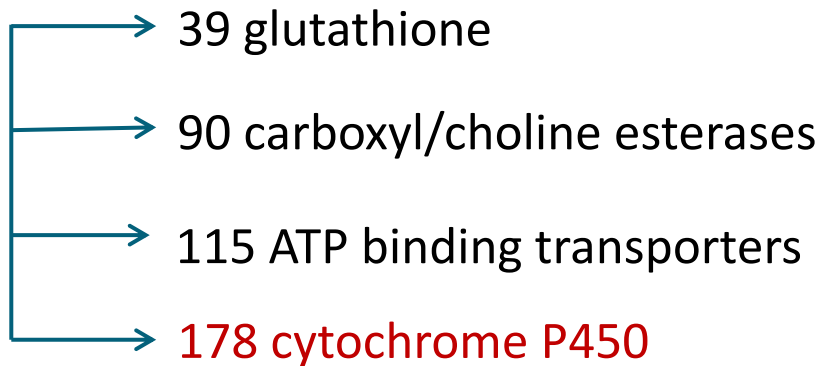
American Cockroach Genome

The American cockroach is an omnivorous scavenger. Its environmental adaptation is mainly mediated by **chemical communication** and **tolerance of chemical and biological factors (toxins & pathogens)**.

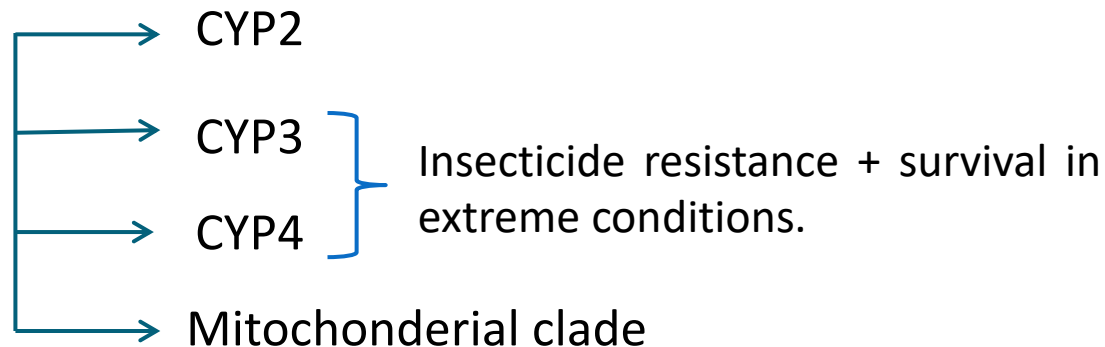


American Cockroach: Detoxification System

Detoxification system includes many enzymes and receptors.

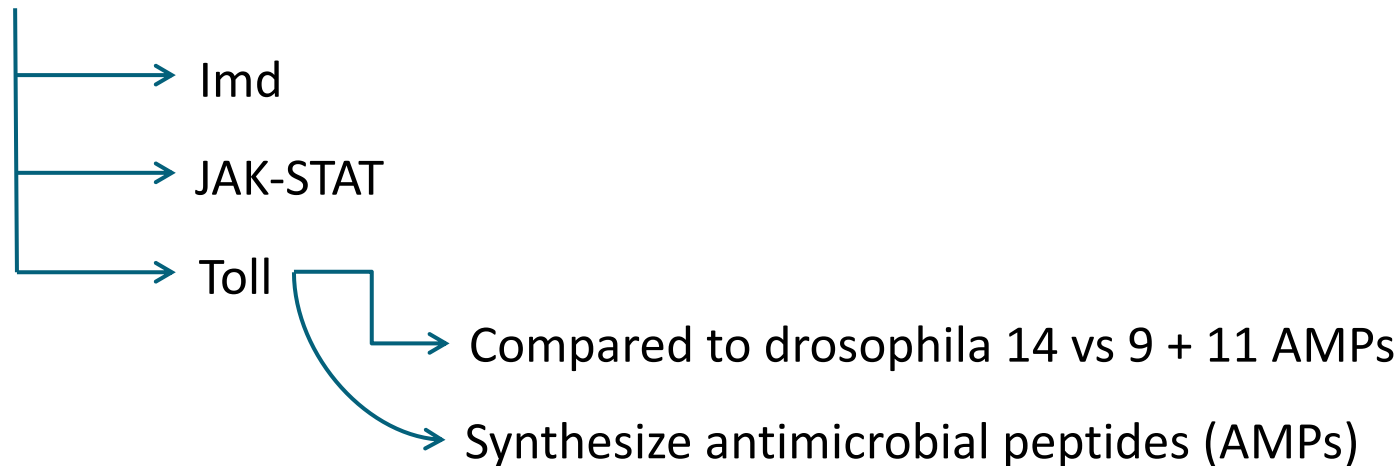


Major clans



American Cockroach: Innate Immunity

Innate immunity is mediated by 3 pathways:




- By performing RNA interference knockdown of 4 important genes in Toll, they found that Toll & AMPs play a major role in fighting invading pathogens.




American Cockroach: Interesting Facts

- The American cockroach's ethanol extract has been developed as a drug for wound healing and tissue repair in the traditional Chinese medicine.
- It has the ability to regenerate lost appendages.
- The specimen provided for the study was inbred for 30 years.
- A cockroach can live almost a month without food.
- A cockroach can live for about 2 weeks without water.
 - Some female cockroaches only mate once and stay pregnant for life.
 - A cockroach can live up to 1 week without its head, it dies from starvation.



American Cockroach: Development

 Insulin/insulin like growth factors + 20E (20-hydroxyecdysone)
Major mechanisms that define the final body size by regulating larval and nymph growth.

- 20E  regulation of molting
- JH (juvenile hormone)  metamorphosis
- Insulin  growth

American Cockroach: Question

- **What are the factors that mediate the environmental adaptation?**
 1. chemical communication
 2. tolerance of chemical and biological factors (toxins & pathogens).



2



THE ASIAN LONGHORNED BEETLE
(*Anoplophora glabripennis*)

Asian Longhorned Beetle: Facts

- Beetles (order: Coleoptera) accounts for more than 20% of metazoans (animals with multiple differentiated cells).
- Longhorned beetles (beetle family: Cerambycidae Latreille) are among the most diverse xylophagous (wood-feeding) animals.
- Asian longhorned beetles (ALB) have recently been established as a model for studies related to xylophagy in beetles.
- ALB is a globally significant invasive species.



Asian Longhorned Beetle Genome

- **Specimens**

- DNA was extracted from a single female (homogametic) larva, due to larvae producing larger DNA fragments than adults.
- RNA ,for annotation, was extracted from four larvae.

- **Sequencing strategy**

- Whole genome shotgun (WGS).

- **Sequencing method**

- Illumina.



Asian Longhorned Beetle Genome

- **Genome assembly**

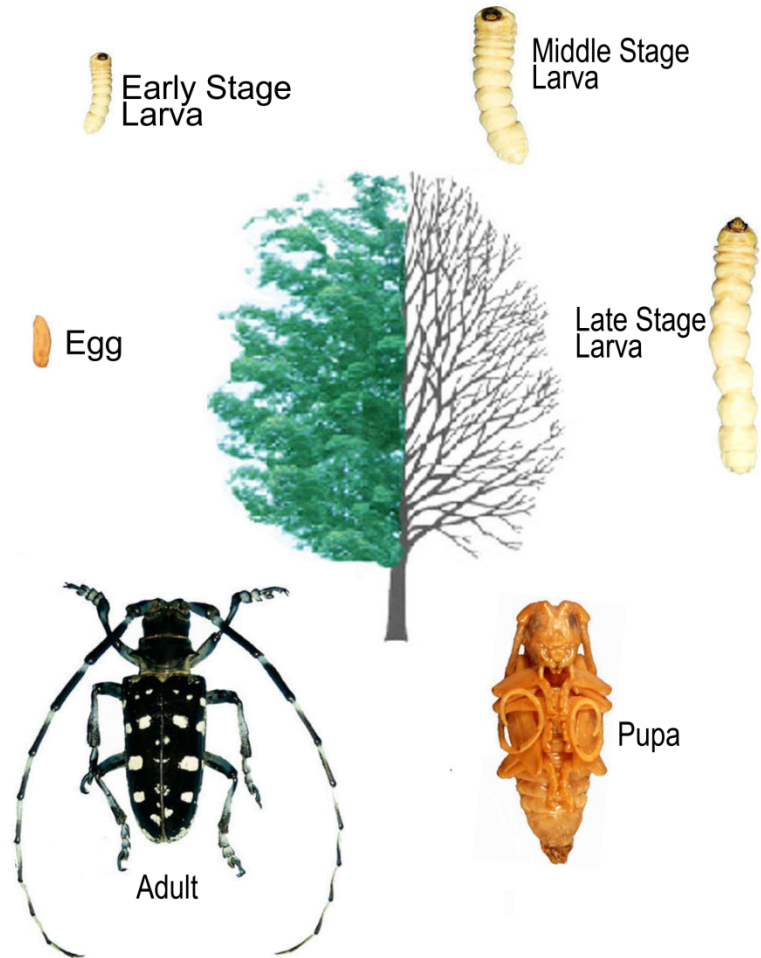
- Via ALLPATHS-LG.
- Further scaffolded and gap-filled using Atlas link and Atlas gap-fill.
- Size of assembly 710 Mb.
- Contig N50 of 16.5 kb and scaffold N50 of 659 Mb.

- **Annotation**

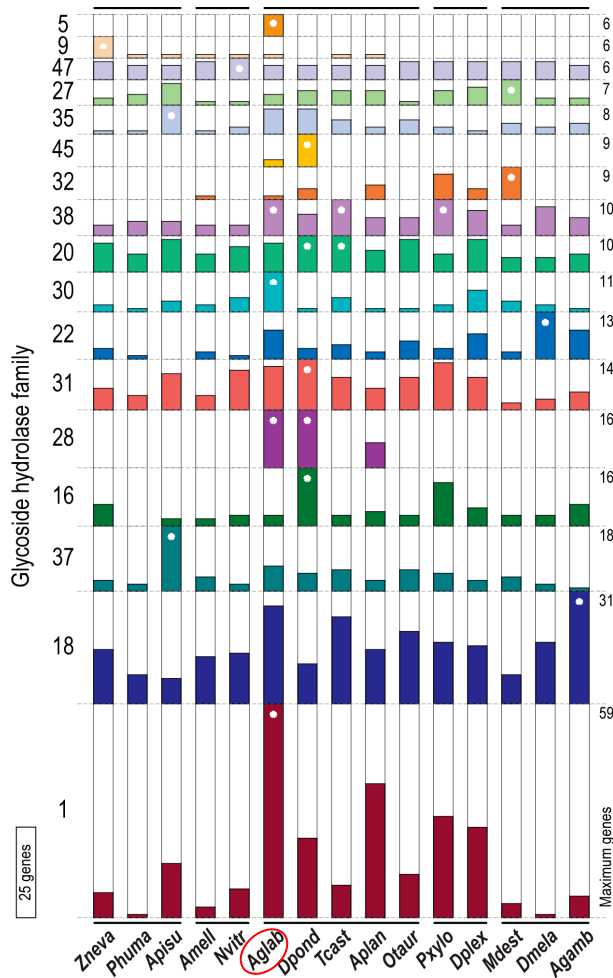
- Automatic gene annotation using MAKER 2.0.
- Both protein and RNA-seq from extant arthropod gene sets were used to guide gene models.



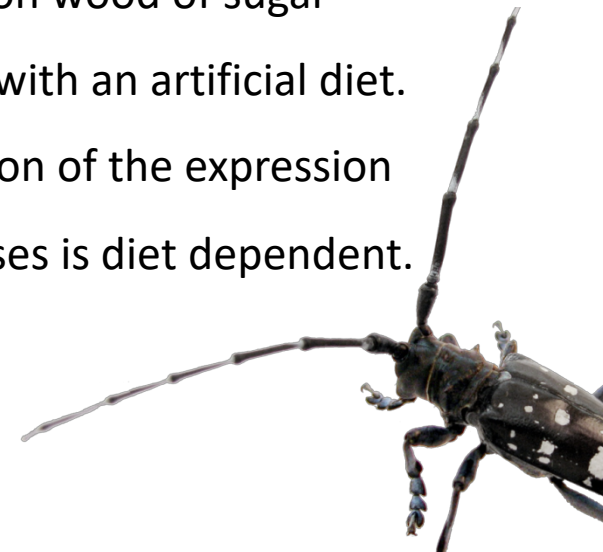
Asian Longhorned Beetle: Life Cycle



Asian Longhorned Beetle: Plant Cell Wall Degradation



- 86 glycoside hydrolase family genes were manually annotated in the *A. glabripennis* genome. This is more than is known for any other insect.
- Some of these hydrolases were expressed twice as much in larvae feeding on wood of sugar maple trees than in those with an artificial diet.
- This indicates that regulation of the expression of these glycoside hydrolases is diet dependent.



Asian Longhorned Beetle: Detoxification

- Allelochemicals are biochemicals produced by plants that have a physiological affect on other organisms.
- In cases in which parasites attack a plant, these allelochemicals can have detrimental effects.
- The invasiveness of *A. glabripennis* is aided by genes coding many enzymes that detoxify these allelochemicals.
- Example: Cytochrome P450s encode the most prevalent detoxification enzymes in insects. A total of 106 genes predicted to encode CYP450s were manually annotated in the *A. glabripennis* genome



Asian Longhorned Beetle: Interesting Fact

- ALB is capable of inflicting severe damage on many species of trees. If uncontrolled, its economic damage in the U.S. alone could reach up to \$889 billion. This is approximately equal to 270 billion Kuwaiti Dinar.



Asian Longhorned Beetle: Question

- **What are some of the traits that enable ALB to be an efficient xylophagus invasive species?**
 1. Enzymes that are capable of degrading plant cell walls.
 2. The ability to detoxify plant allelochemicals.

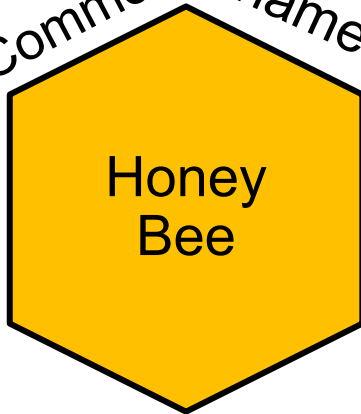


3



THE HONEYBEE
(*Apis mellifera*)

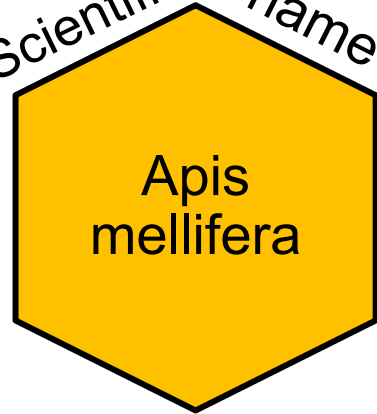
Common name



Honey
Bee



Scientific name



Apis
mellifera



Honeybee Genome



Worker



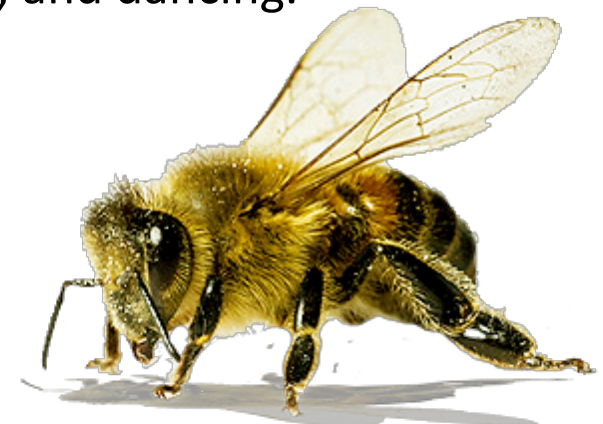
Queen



Drone

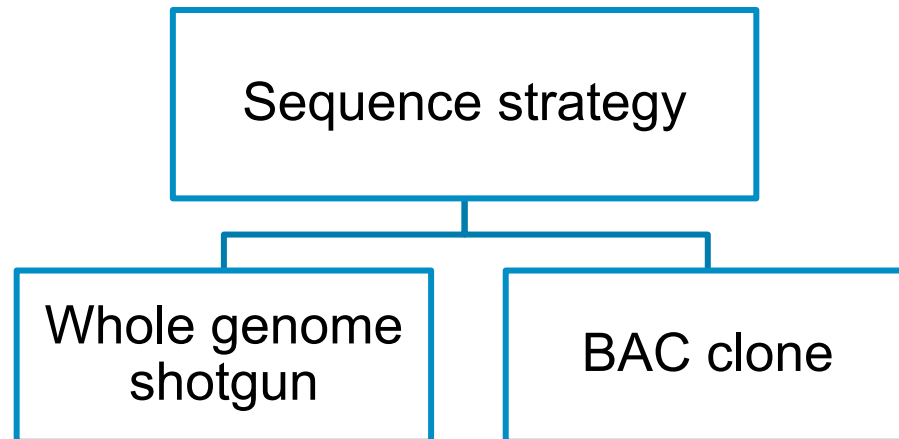
Honeybee: Facts

- Native to Middle East, Asia, and Africa.
- Feed on pollen, nectar, honey, and secretions.
- A queen lives for 1-2 years, workers live for weeks or months, drones live for 2 months
- The queen can lay up to 2000 eggs per day.
- Worker bees do most of the work in the colony.
- Drones only mate.
- Communication is based on chemical signals, vibrations, and dancing.
- Visit about 2 million flowers to make 1 pound of honey.



Honeybee Genome

- **Specimens**
 - Multiple drones derived from a single slightly inbred queen (DH4 strain).
- **Sequencing method**
 - Illumina.



Honeybee Genome

- **Genome assembly**

- Via Atlas assembler.
- Expected genome size 236 Mb.
- Contig N50 of 41 kb and scaffold N50 of 362 Mb.
- GC contents: 31.55%

- **Annotation**

- Number of genes: 10157.
- Number of protein coding: 893.
- % of the genome that contains genes: 71% with domains.





Honeybee: Interesting Fact

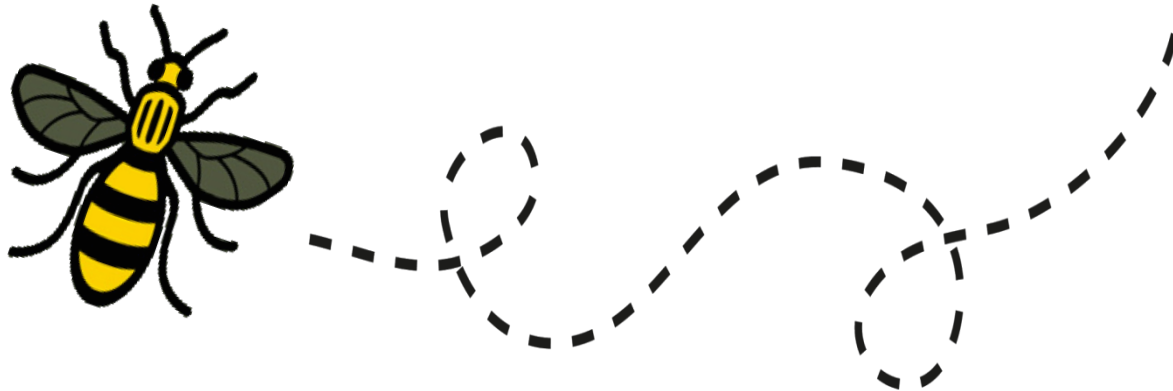
- Exocrine glands in the head of adult worker bees secrete royal jelly which is responsible for the development of a queen bee.



Honeybee: Question

- **Can female workers help the queen in laying eggs?**

No, because they are sterile due to the queen's secreted chemicals called pheromones.



4



THE CLOUDED SKIPPER

(*Lerema accius* -Lac-)

The Clouded Skipper: Facts

- The clouded skipper, is a butterfly of the family Hesperiidae.
- Has unique phenotypic traits including fast flight with rapid wing beats.
- It has Hox genes that are crucial for development.



The Clouded Skipper Genome

- **Specimens**
 - The source of DNA is the wings and abdomen of frozen male (Lac) specimen.
- **Sequencing strategy**
 - Whole genome shotgun (WGS).
- **Sequencing method**
 - Illumina platform.
 - Used synthetic adapters to build a library.



The Clouded Skipper Genome

■ Genome assembly

- Three methods to assemble the sequence read of (Lac):
 - De novo assembly .
 - Reference-based assembly.
 - Reference- guided assembly.
- Genome size is 310 Mb.
- Scaffold N50 of 513 Kb.

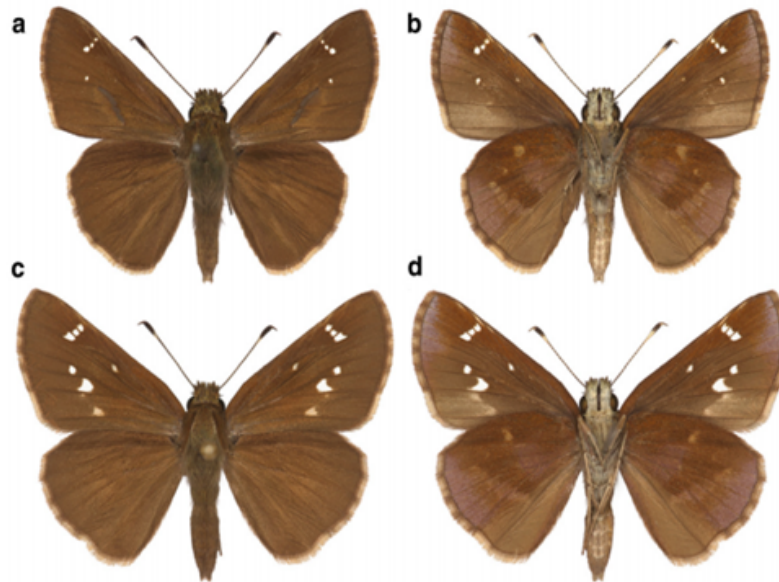
■ Annotation

- They predict the protein coding genes in the Lac to be 17,416.
- Number of species specific genes is 4586.
- GC content is 15.5%.



The Clouded Skipper: Interesting Fact

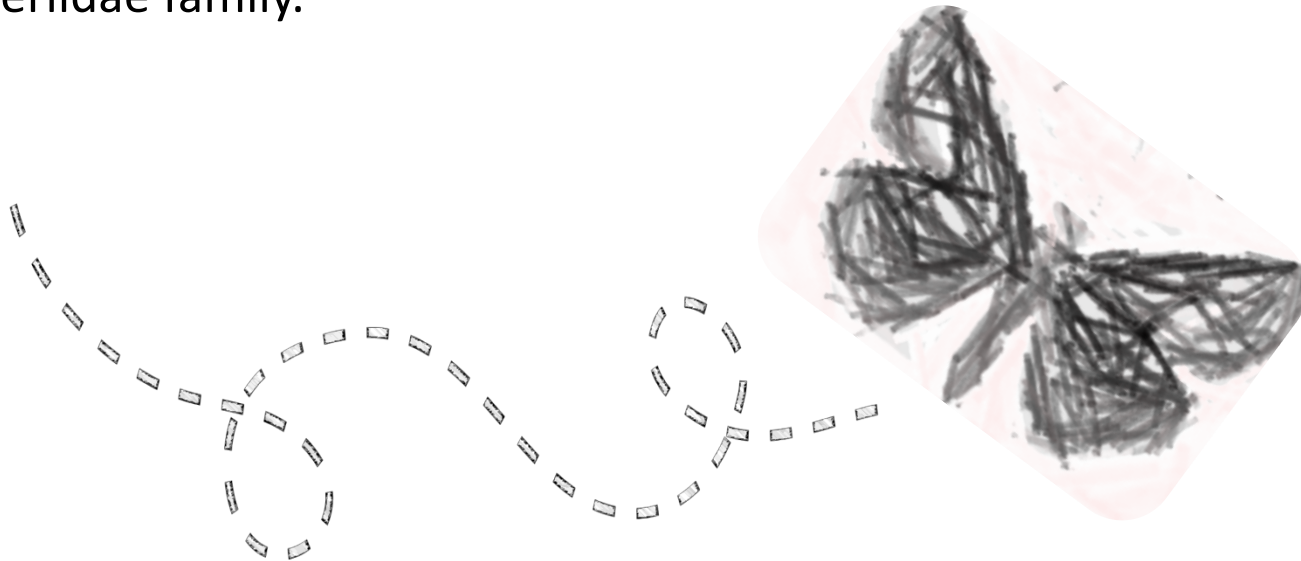
- The Lac genome encodes 56 ORs (Odorant Receptors), that are particularly important for the feeding and mating behaviors of insects.



Clouded Skipper: Question

- **Why does the genome of the clouded skipper offer rich data for comparative genomics and phylogenetic studies?**

Because the Clouded Skipper is the first sequenced genome from the HesperIIDae family.



5



THE ANT

(Linepithema humile)

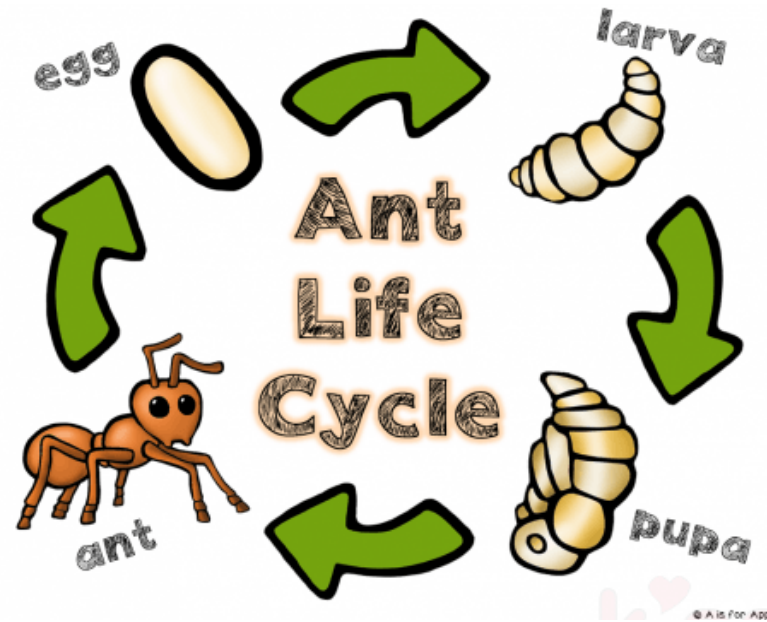
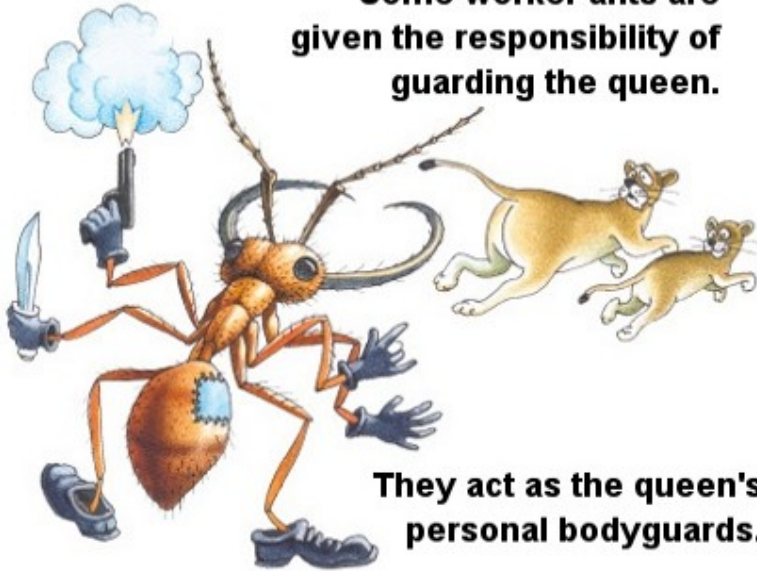
The Ant: Facts

- Ants use their two stomachs to share food with other members of the colony.
- Ants normally range in size from 2 to 7 mm in length.
- Ants include >14,000 described species, comprising about half of all insect biomass in the tropics.
- The sophisticated social structures of ant colonies are regulated by a complex interplay of chemical signaling, perception of those signals, and behavioral responses.
- To clarify the genetic and genomic contributions to these aspects of ant sociality, genes for the production of chemical signals (**desaturases**) were analyzed.



The Ant: Facts

Some worker ants are given the responsibility of guarding the queen.



The Ant Genome

- **Specimens**

- The DNA was extracted and purified from a single queen pupa (Saratoga).
- Collected from a residential orchard in Santa Clara Country.

- **Sequencing strategy**

- Whole genome shotgun (WGS).

- **Sequencing method**

- 454 and Illumina platform.



The Ant Genome

■ Genome assembly

- De novo from whole genome shotgun (WGS).
- Via Newbler and CABOG assembler (13).
- Genome size is 20,828,920 bp.
- Contig N50 of 35,858 bp and scaffold N50 of 1,386,360 bp.

■ Annotation

- 83 genes were revealed from cytoplasmic ribosomal protein genes, including the full set of 79 cytoplasmic ribosomal proteins (16, 17), and 4 duplicated genes (RpS16, RpS23, RpS28, and RpS30).
- Annotation of the 67 nuclear-encoded oxidative phosphorylation genes shows that the *L. humile* genome assembly is only missing *cox7a*.



The Ant: Questions

- **How do ants communicate with each other?**

Through chemical signaling.

- **What is the name of the molecules used for this type of communication?**

Desaturases.



Comparison of Genomic Features

	Genome Size	%GC	Sequencing Strategy	Sequencing Method
American Cockroach	3,380 Mb	35.7%	WGS	Illumina
ALB	710 Mb	32.7%	WGS	Illumina
Honeybee	236 Mb	31.55%	WGS Bac clone	Illumina
Clouded Skipper	310 Mb	15.5%	WGS	Illumina
The Ant	20.8 Mb	38%	WGS	Illumina