Genomics Data 9 (2016) 97–99

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Data in Brief

Transcriptomic and functional resources for the small hive beetle *Aethina tumida*, a worldwide parasite of honey bees

Matthew R. Tarver [a,1], Qiang Huang [b,c], Lilia de Guzman [a], Tom Rinderer [a,2], Beth Holloway [a,2], Justin Reese [d], Daniel Weaver [d], Jay D. Evans [b,*]

[a] USDA-ARS Honey Bee Breeding, Genetics and Physiology Laboratory, Baton Rouge, LA 70820, United States
[b] USDA-ARS Bee Research Laboratory, Beltsville, MD 20705, United States
[c] Institute of Bee Health, Vetsuisse Faculty, University of Bern, CH-3003, Switzerland
[d] Genformatic, Austin, TX 78731, United States

ARTICLE INFO

Article history:
Received 8 June 2016
Accepted 18 June 2016
Available online 02 July 2016

Keywords:
Pollination
Host-parasite
Transcriptome
Honeybee

ABSTRACT

The small hive beetle (SHB), *Aethina tumida*, is a major pest of managed honey bee (*Apis mellifera*) colonies in the United States and Australia, and an emergent threat in Europe. While strong honey bee colonies generally keep SHB populations in check, weak or stressed colonies can succumb to infestations. This parasite has spread from a sub-Saharan Africa to three continents, leading to immense management and regulatory costs. We performed a transcriptomic analysis involving deep sequencing of multiple life stages and both sexes of this species. The assembled transcriptome appears to be nearly complete, as judged by conserved insect orthologs and the ability to find plausible homologs for 11,952 proteins described from the genome of the red flour beetle. Expressed genes include each of the major metabolic, developmental and sensory groups, along with genes for proteins involved with immune defenses and insecticide resistance. We also present a total of 23,085 high-quality SNP's for the assembled contigs. We highlight potential differences between this beetle and its honey bee hosts, and suggest mechanisms of future research into the biology and control of this species. SNP resources will allow functional genetic analyses and analyses of dispersal for this invasive pest. All resources are posted as Supplemental Tables at https://data.nal.usda.gov/dataset/data-transcriptomic-and-functional-resources-small-hive-beetle-aethina-tumida-worldwide, and at NCBI under Bioproject PRJNA256171.

Specifications	
Organism/cell line/tissue	*Aethina tumida*/larval and adult tissues
Sex	Both
Sequencer or array type	ILLUMINA Hi-Seq
Data format	Assembled transcriptome
Experimental factors	Pooled RNA sample form 17 tissues
Experimental features	Transcriptome from equal-molar extractions from various samples
Consent	Not applicable
Sample source location	Baton Rouge, Louisiana

1. Direct link to deposited data

https://data.nal.usda.gov/dataset/data-transcriptomic-and-functional-resources-small-hive-beetle-aethina-tumida-worldwide

2. Introduction

The small hive beetle (SHB), *Aethina tumida* was first recorded in the United States in Florida in 1998 [1] and arrived perhaps as early as 1996 in South Carolina [2]. This accidental introduction into the United States presumably originated from sub-Saharan Africa [3]. While the SHB is not a major honey bee pest in its native environment [4], US honey bees have been strongly impacted, leading to spoiling of honey and often colony losses. SHB larvae feed on all hive products including honey, pollen and brood [5]. SHB are now found throughout North America, although the southeastern US has been particularly hard hit by SHB infestations. In Australia, a point introduction in New South Wales ten years ago has led to extensive infestation and colony impacts throughout eastern Australia. Finally, a recent introduction into Italy has led to the first stable European population of this pest [6,7].

* Corresponding author.
E-mail address: jay.evans@ars.usda.gov (J.D. Evans).
[1] Current affiliation: BayerCrop Science, West Sacramento, CA 95605, United States.
[2] Retired.

http://dx.doi.org/10.1016/j.gdata.2016.06.003

SHB are opportunistic pests of both honey bees and bumble bee colonies [8]. They are also capable of reproducing and developing on different fruits under laboratory conditions [9,5,10,11] and have been observed feeding in compost containing decomposing watermelon rinds [10]. While SHB have not been observed to complete a life cycle in the absence of bees in natural settings, their ability to feed on plant food sources has likely enabled their rapid increase in numbers and geographic range. In-hive chemical treatments that control SHB but spare honey bees have been elusive, in part because of the high chemical sensitivity of honey bees to such chemicals. Mechanical controls have been developed to lure, bait or trap beetles whereas ground drenching chemicals, as well as diatomaceous earth, purportedly inhibit SHB pupation in soil surrounding colonies [12]. Since adult beetles can fly about 10 km, re-invasion of SHB from nearby untreated apiaries occurs quickly. Currently, the best method of control comes directly from the honey bees themselves. Strong healthy honey bee colonies are able to maintain low SHB populations by exhibiting aggression toward adult beetles. Aggressive behavior may limit food consumption and reproduction inside the colonies. Nevertheless, aggression toward SHB varies among honey bee genotypes [13,14,15].

Here we describe an extensive transcriptomic analysis of SHB based on RNA gathered from various life stages and both sexes. We show the breadth and utility of this transcriptome through assignments of orthology-based on Benchmarking Universal Single Copy Orthologs (BUSCO; [16]) developed for holometabolous insects. We also provide plausible orthologs based on alignments to the protein-coding genes found in the genome of the red flour beetle *Tribolium castaneum* [17]. Finally, we provide an estimate of sequence polymorphisms within the tested pool, and describe a tenfold increase in the number of known microsatellite loci for this species. Along with these tools for future research, we suggest possible mechanisms that can be targeted for in-hive SHB control.

3. Materials and methods

3.1. Samples

Small hive beetle adults were collected from multiple apiaries in Baton Rouge, Louisiana, USA. Adult beetles were maintained in the laboratory until sacrificed (approximately within two weeks). Adult males and females were then differentiated, and 30–50 adults per sex were frozen for analyses. The other adult beetles (30–50 per sex) were dissected to sample different body parts, including midguts, heads, antennae, thorax (excluding the forelegs), abdomens (along with the mid- and hind-legs), and ovaries of females. In order to obtain eggs and different larval stages for analyses, adult males and females were placed in a rearing container while being fed *ad libitum* honey, pollen and brood as described by de Guzman and Frake [18]. The rearing container was then placed inside an incubator at 34 °C. After 24 h, all adults were removed and eggs (100–150) were sampled at that time. When the remaining eggs hatched, 50–100 first-instar (L1) larvae were collected. Instars L2 and L3 (both feeding and wandering stages, n = 30–50 per stage) were also collected as larvae developed. Sampling was conducted on 5 cohort sets. Total RNA was isolated from each sample type and replicate using the Maxwell 16 nucleotide purification system according to the manufactures protocol (Promega Corporation, Madison, WI). RNA was purified from the samples by loading 400 μl of homogenate into the Maxwell 16 system. Purified RNA was eluted in 50 μl of nuclease-free water then assessed for quantity (ng/μl) and quality (NIN value, BioAnalyzer, Agilent Corp.), respectively. The quality and quantity of each sample were then compared and the top three replicates for each sample type were used. For each replicate, equal amount of each RNA sample type (2000 ng) was then combined to create an all-encompassing RNA pool.

3.2. De novo assembly and analyses

The RNA pool was subjected to ILLUMINA paired-end sequencing, generating a total of 176,246,845 paired 101-basepair reads (100 bp gap between reads, total of 35,601,862,690 nucleotides) after trimming and cleaning. Adapter sequences were removed using Fastq-mcf and adapters were removed using DynamicTrim. Fastq-mcf was run using default parameters for adapter trimming and quality trimming was disabled. DynamicTrim was used for quality trimming with default parameters. To initiate a *de novo* transcriptome assembly, Khmer was run separately on the forward and reverse files to normalize the reads *in silico*. Pseudo-reads were created to ensure that all of the remaining reads had a mate. The reads were assembled using a multi-kmer (k = 23–50) approach using Oases. Oases assemblies were made for each kmer, and then merged into an optimized non-redundant set. The resulting transcripts were then filtered to remove all transcripts less than 200 bp (two times the initial read length). This assembly was further collapsed using the contig program IDBA [19] under default conditions. Microsatellites were identified from this consensus assembly using MSDB (Microsatellite Search and Building Database) package [20]. SNP positions were identified with Tophat2-Picard-GATK pipeline [21,22] Identification of biologically important genes was aided by comparisons against the gene set generated for the red flour beetle, *Tribolium castaneum* (Beetlebase.org), and by matching this gene set to the BUSCO set of proteins found in each of the holometabolous insects, as well as to the OrthoDB catalogue of arthropod proteins [23].

4. Results

4.1. De novo assembly and polymorphism

A total of 176,246,845 paired reads (35,601,862,690 nucleotides) were submitted to the NIH-NCBI Short Reads Archive under Bioproject PRJNA256171 and BioSample SAMN02940944. Following OASES sequence assembly, a total of 259,543 transcript contigs were identified representing 36,649 loci and ranging in length from 200 bp to 50,917 bp (mean = 1681.65 bp). These contigs were collapse to 42,761 consensus contigs (NCBI Transcript-supported assembly GCKB00000000.1). GC content was determined based on analysis of the longest transcript for each locus and found to be 38.6% of the total potential transcriptome size of 32.84 Mb. In total, 969 microsatellite markers were identified, a frequency of one locus per 26 kbp (Supplementary Table 1). Of these, 156 were flanked by over 50 nucleotides on both sides of the repetitive region, enabling straightforward primer design for screening and scoring polymorphisms. Additionally, 23,085 SNPs were identified, an average distance of one SNP per 1 kbp (Supplementary Table 2).

4.2. Validation and annotation

Benchmarking sets of Universal Single-Copy Orthologs (BUSCO), a core set of single-copy genes present across the holometabolus insects [16] were used to test the assembled contigs for completeness. Of the BUSCO set consisting of 2477 arthropod genes, only 24 were not identified in the SHB transcriptome. Those present in the SHB transcriptome were found as both single copy (n = 132) and multiple copy (n = 2321) matches. The unexpected high number of multi-copy genes reflects, in large part, expression of multiple isoforms from the test population, listed as multiple transcripts in the OASES assembly. The consensus loci generated by a secondary assembly (IDBA, [18]) were queried using the protein set of the red flour beetle, *Tribolium castaneum* by tBLASTN, generating 11,952 matches to 8586 loci (Supplemental Table 3). These matches fell into all of the major functional protein groups (134 KEGG pathway terms, 2647 INTERPRO domains, 1390 distinct Gene Ontology Biological Process terms). While a complete genomic analysis is required to assess paralog counts of key gene families,

there were no obvious discrepancies relative to *Tribolium* beetles for key groups. As one example, the transcriptome presented 109 distinct cytochrome P450 detoxification enzyme candidates, a number close to the 134 members of this group identified in whole-genome sequencing of *Tribolium* [17]. Consensus loci and their associated matches in *T. castaneum*, OrthodDB, Flybase, Gene Ontology Biological process and Cellular location, INTERPRO, and KEGG pathway are given in Supplementary Table 3.

5. Discussion

This transcriptome resource offers insights into biology of a nitidulid beetle that is a significant pest of managed honey bees. All resources are available at Ag Data Commons (https://data.nal.usda.gov/dataset/transcriptomic-and-functional-resources-small-hive-beetle-aethina-tumida-worldwide-parasite). The transcriptome that we describe reflects members of a multitude of physiological pathways, including reproduction, digestion, respiration, behavior, and morphology. A specific assault on any of these pathways may lead to a less fit parasite, increasing the likelihood that natural aggression exhibited by honey bees can keep beetle populations below destructive levels. An important aspect of our transcriptome study is that both sexes, all life stages, and all adult body structures were analyzed. There are nevertheless likely to be genes that are expressed so briefly that our study failed to capture them. However, a successful targeted pest management strategy would employ a prolonged assault of gene targets to ensure a prolonged and sufficient incapacitation of the pest. Simple and cost-effective quantitative PCR can now be used to analyze potential gene targets in each sample type to determine the best treatment strategy based on gene regulation and function.

As expected, the SHB transcriptome is highly similar to another beetle, *Tribolium*, a well-described genome and model organism. The similarity between the species indicates that pest management controls developed for *Tribolium* and other beetles will function equally well for SHB. Indeed, RNAi technology is proving successful against pest insects such as *Tribolium* [24]. Assuming that this mechanism also exists in SHB, a multi-factored strategy can be used to combat the pest at several life stages and through the disruption of several important physiological pathways.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gdata.2016.06.003.

Acknowledgements

This work was funded by the USDA Agricultural Research Service. The funding source played no role in experimental design, implementation, and interpretation.

References

[1] P.J. Elzen, J.R. Baxter, D. Westervelt, C. Randall, K.S. Delaplane, L. Cutts, W.T. Wilson, Field control and biology studies of a new pest species, *Aethina tumida* Murray (Coleoptera, Nitidulidae), attacking European honey bees in the Western Hemisphere. Apidologie 30 (1999) 361–366.

[2] M. Hood, Overview of the small hive beetle *Aethina tumida* in North America. Bee World 81 (2000) 129–137.

[3] J.D. Evans, J.S. Pettis, H. Shimanuki, Mitochondrial DNA relationships in an emergent pest of honey bees: *Aethina tumida* (Coleoptera:Nitidulidae) from the United States and Africa. Ann. Entomol. Soc. Am. 93 (2000) 415–420.

[4] P. Neumann, P.J. Elzen, The biology of the small hive beetle (*Aethina tumida*, Coleoptera: Nitidulidae): Gaps in our knowledge of an invasive species. Apidologie 35 (2004) 229–247.

[5] J.D. Ellis, P. Neumann, H.R. Hepburn, P.J. Elzen, Longevity and reproductive success of *Aethina tumida* (Coleoptera: Nitidulidae) fed different natural diets. J. Econ. Entomol. 95 (2002) 902–907.

[6] F. Mutinelli, F. Montarsi, G. Federic, A. Granato, A.M. Ponti, Detection of *Aethina tumida* Murray (Coleoptera: Nitidulidae) in Italy: outbreaks and early reaction measures. J. Apic. Res. 53 (5) (2014) 569–575.

[7] V. Palmeri, G. Scirto, A. Malacrino, F. Laudani, O. Campolo, A scientific note on a new pest for European honeybees: first report of small hive beetle *Aethina tumida*, (Coleoptera: Nitidulidae) in Italy. Apidologie 46 (2015) 527–529.

[8] S. Spiewok, P. Neumann, Infestation of commercial bumblebee (*Bombus impatiens*) field colonies by small hive beetles (*Aethina tumida*). Ecol. Entomol. 31 (2006) 623–628.

[9] J.J. Keller, Testing Effects of Alternative Diets on Reproductive Rates of the Small Hive Beetle, *Aethina tumida*. MS thesis North Carolina State University, Raleigh, NC, 2002.

[10] R.T. Arbogast, B. Torto, S. Willms, P.E.A. Teal, Trophic habits of *Aethina tumida* (Coleoptera: Nitidulidae): their adaptive significance and relevance to dispersal. Environ. Entomol. 38 (2009) 561–568.

[11] R.T. Arbogast, B. Torto, P.E.A. Teal, Potential for population growth of the small hive beetle *Aethina tumida* (Coleoptera: Nitidulidae) on diets of pollen dough and oranges. Fla. Entomol. 93 (2010) 224–230.

[12] S. Buchholz, K. Merkel, S. Spiewok, J.S. Pettis, M. Duncan, R. Spooner-Hart, C. Ulrichs, W. Ritter, P. Neumann, Alternative control of *Aethina tumida* Murray (Coleoptera: Nitidulidae) with lime and diatomaceous earth. Apidologie 40 (2009) 535–548.

[13] P.J. Elzen, J.R. Baxter, P. Neumann, A. Solbrig, C. Pirk, H.R. Hepburn, D. Westervelt, C. Randall, Behavior of African and European subspecies of *Apis mellifera* toward the small hive beetle, *Aethina tumida*. J. Apic. Res. 40 (2001) 40–41.

[14] L.I. de Guzman, T.E. Rinderer, A.M. Frake, H. Tubbs, P.J. Elzen, D. Westervelt, Some observations on the small hive beetle, *Aethina tumida* Murray in Russian honey bee colonies. Am. Bee J. 146 (2006) 618–620.

[15] A.M. Frake, L.I. de Guzman, T.E. Rinderer, Comparative resistance of Russian and Italian honey bees (Hymenoptera: Apidae) to small hive beetles (Coleoptera: Nitidulidae). J. Econ. Entomol. 102 (2009) 13–19.

[16] F.A. Simao, R.M. Waterhouse, P. Loannidis, E.V. Kriventseva, E.V. Zdobnov, BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-copy Orthologs(Bioinformatics) 2015, http://dx.doi.org/10.1093/bioinformatics/btv351.

[17] T.G.S. Consortium, The first genome sequence of a beetle, Tribolium castaneum, a model for insect development and pest biology. Nature 452 (2008) 949–955.

[18] L.I. de Guzman, A.M. Frake, Temperature affects *Aethina tumida* (Coleoptera: Nitidulidae) development. J. Apic. Res. Bee World 46 (2007) 88–93.

[19] Y. Peng, H.C. Leung, S.M. Yiu, F.Y. Chin, IDBA–A Practical Iterative de Bruijn Graph de Novo Assembler. In *Research in Computational Molecular Biology* Springer, Berlin Heidelberg, 2010 426–440.

[20] L. Du, X. Zhang, B. Yue, MSDB: a user- friendly program for reporting distribution and building databases of microsatellites from genome sequences. J. Hered. (2013), http://dx.doi.org/10.1093/jhered/ess082.

[21] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20 (2010) 1297–1303.

[22] D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S.L. Salzberg, Tophat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. (2013), http://dx.doi.org/10.1186/gb-2013-14-4-r36.

[23] E.V. Kriventseva, F. Tegenfeldt, T.J. Petty, R.M. Waterhouse, F.A. Simão, I.A. Pozdnyakov, P. Ioannidis, E.M. Zdobnov, OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software. Nucleic Acids Res. (2015) 250–256, http://dx.doi.org/10.1093/nar/gku1220.

[24] J. Ulrich, V.A. Dao, U. Majumdar, C. Schmitt-Engel, J. Schwirz, D. Schultheis, Large scale RNAi screen in *Tribolium* reveals novel target genes for pest control and the proteasome as prime target. BMC Genomics (2015), http://dx.doi.org/10.1186/s12864-015-1880.