

Comprehensive transcriptome of the maize stalk borer, Busseola fusca, from multiple tissue types, developmental stages, and parasitoid wasp exposures

Kayla Hardwick, Gladys Bosibori Bichang'a, Andnet Bayleyegn Abtew, Ryan Musumba Awori, Leah Cepko, Lorna Jemosop Chebon-Bore, Alistair Darby, Jon Devries, Jonathan Filée, Muna Fuad, et al.

▶ **To cite this version:**

Kayla Hardwick, Gladys Bosibori Bichang'a, Andnet Bayleyegn Abtew, Ryan Musumba Awori, Leah Cepko, et al.. Comprehensive transcriptome of the maize stalk borer, Busseola fusca, from multiple tissue types, developmental stages, and parasitoid wasp exposures. Genome Biology and Evolution, 2020, 10.1093/gbe/evaa195 . hal-02998098

HAL Id: hal-02998098
https://hal.science/hal-02998098

Submitted on 16 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ARTICLE TYPE: Genome Report for Genome Biology and Evolution

TITLE: Comprehensive transcriptome of the maize stalk borer, *Busseola fusca*, from multiple tissue types, developmental stages, and parasitoid wasp exposures

AUTHORS: Kayla M. Hardwick[1,2,3], Gladys Bosibori Bichang'a[3], Andnet Bayleyegn Abtew[3], Ryan Musumba Awori[3], Leah C. S. Cepko[3], Lorna Jemosop Chebon-Bore[3], Alistair Darby[3], Jon D. deVries[1,3], Jonathan Filée[3], Muna Fuad[3], Grace Gachara[3], Dedan K. Githae[3], Patrick Gunga[3], Madison Held[1,3], Hellen Wambui Kariuki[3], Evans Sioma Kataka[3], Susan Diana Kerfua[3], Kelvin Muteru Kimenyi[3], Bruno Le Ru[3], Ernest Mukweyi Lutomia[3], Elizabeth Ajema Luvai[3], Maureen K. Luvanda[3], Beatus Modest Lyimo[3], Eunice Magoma Machuka[3], Daniel G. Maeda[3], Solomon Maina[3], Peterson Gitonga Mathenge[3], Damaris Matoke-Muhia[3], Caitlin H. Miller[3], Silviane Aoko Miruka[3], Alfred Mitema[3], Antoinette Aluoch Miyunga[3], Sarah Akwabi Mukolwe[3], Edward Kirwa Muge[3], Simon Ngao Mule[3], Mary Kanyiri Murithi[3], Jean Pierre Musabyimana[3], Saphan Muzoora[3], Esther Wangui Mwangi[3], Harrison Ndung'u Mwangi[3], Nduta Mwangi[3], Ann Njeri Mwaura[3], Nelly Ndungu[3], Bidii Stephen Ngalah[3], Peter Njenga Ng'ang'a[3], Harun N. Njoroge[3], Brenda Nyaboke Nyandika[3], Johnson Ounya Nyasani[3], Faith Akinyi Obange[3], Shem Joabh Ochieng[3], Willingtone Otieno Odhiambo[3], Hellen Adhiambo Ogot[3], Maureiq Awino Edith Ojwang'[3], Fred Odhiambo Osowo[3], Billy Omboki Ratemo[3], Tolbert Sonda[3], Alexander Ssamula[3], Sharon Towett-Kirui[3], Augustin Penda Twizerimana[3], Kelvin Mwangi Wachiuri[3], Alex Paul Wacoo[3], Kevin Wamae[3], Mark Kilongosi Webale[3], Reuben Mangi Yaa[3], Dawn E. Gunderson-Rindal[3], Francesca Stomeo,[3] Appolinaire Djikeng[3], Paul-Andre Calatayud[3], Sarah Schaack*[1,3]

Author affiliations:

[1]Department of Biology, Reed College, Portland, OR 97202
[2]current address: Phylos Bioscience, Portland, OR 97214
[3]*Busseola fusca* Genomics Consortium (for a complete list of all author addresses and affiliations, please see Appendix A

***Corresponding Author, email:** schaack@reed.edu

Running title: *Busseola fusca* transcriptome

Keywords: agricultural pest, Lepidoptera, insect genomics

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

Abstract

Busseola fusca (Fuller) (Lepidoptera: Noctuidae), the maize stalk borer, is a widespread crop pest in sub-Saharan Africa that has been the focus of biological research and intensive management strategies. Here, we present a comprehensive annotated transcriptome of *B. fusca* (originally collected in the Western Province of Kenya) based on ten pooled libraries including a wide array of developmental stages, tissue types, and exposures to parasitoid wasps. Parasitoid wasps have been used as a form of biocontrol to try and reduce crop losses with variable success, in part due to differential infectivities and immune responses among wasps and hosts. We identified a number of loci of interest for pest management, including genes potentially involved in chemoreception, immunity, and response to insecticides. The comprehensive sampling design used expands our current understanding of the transcriptome of this species and deepens the list of potential target genes for future crop loss mitigation, in addition to highlighting candidate loci for differential expression and functional genetic analyses in this important pest species.

Significance Statement

Despite growing interest in the role of genomics as a method of ensuring food security, there are few examples where new technologies have been applied to insect species that devastate staple food crops in sub-Saharan Africa. Here, we present the transcriptome of *Busseola fusca*, the maize stalk borer, a lepidopteran known to be responsible for massive crop losses. The assembly and annotation provided represents a key step towards identifying candidate genes for the control and mitigation of infestations by *B. fusca* throughout regions where it is found.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

Introduction

The maize stalk borer, *Busseola fusca* (Fuller) (Lepidoptera: Noctuidae; Figure 1), is an important pest of cereal crops in sub-Saharan Africa. Due to its abundance and distribution, it represents a major constraint to food production where infestation rates are high (Kfir et al. 2002). Crop losses resulting from *B. fusca* feeding activity vary by region, but can result in a total loss in some areas (Van den Berg et al. 1991, Calatayud et al. 2014). Its impact on the food security and economic well-being of people has led to a number of studies into the physiology and ecology of the species (reviewed in Calatayud et al. 2014). With the release of the genome (from the same lineage used in this study; Hardwick et al. 2019) and the publication of the first transcriptome (derived from neonate tissue from a genotype collected in South Africa; Peterson et al. 2019), we now have basic genomic information for this species. With pest species, in particular, the identification of candidate genes related to important traits such as developmental timing, reproduction, insecticide resistance, adaptation to plant defense mechanisms, immunity, and chemoreception requires sampling multiple time points, exposures, and populations. This is particularly relevant in *B. fusca*, where certain strains are known to be differentially vulnerable to infection by parasitoid wasps (Kfir 1995). This information, in turn, can inform or enhance management strategies (e.g., those developed for *Plutella xylostella* [You et al. 2013]).

Female *B. fusca* typically deposit eggs (Figure 1A) between the stem and leaf sheet of the host plant. Larvae hatch (Figure 1B), feed on young leaves, and penetrate the plant stem during the third and fourth instar (Figure 1C) where they remain until pupation. Feeding during the larval stage, which is also when the animals are vulnerable to parasitoid wasps (Figure 1D), damages the host plant and reduces yield or kills the plant. After pupation, adult moths (Figure 1E) use chemosensory cues and receptors to attract and find mates, food, and suitable places to lay eggs. Pest management strategies can include, for example, introducing substances that can interrupt sending or receiving chemosensory cues to/from mature adults or can include the introduction of biocontrol species to which the focal species is unable to mount a defense response. Because it involves fewer chemicals, biological control is a pest management strategy that has been employed with *B. fusca* over the last several decades. Beginning in the 1990s, a biocontrol program was launched in Kenya using the indigenous larval parasitoid of *B. fusca*, *Cotesia sesamiae* (Cameron; Hymenoptera: Braconidae; Figure 1D). Prevalence of parasitism by *C. sesamiae* can vary (ranging from <5% to 75% [Kfir 1995; Sallam et al. 1999; Jiang et al. 2006; Songa et al. 2007]), in part due to differences among wasp strains (Mochiah et al. 2002; Gitau et al. 2010; Branca et al. 2011). *Busseola fusca* are resistant to infection by *C. sesamiae* from Mombasa (coastal Kenya), referred to hereon as Cs-Coast, but vulnerable to *C. sesamiae* from Kitale (a site in inland Kenya; referred to hereon as Cs-Inland; Ngi-Song et al. 1995).

Here, we present a comprehensive transcriptome of *B. fusca* based on data from 10 libraries representing different tissues, developmental stages, and parasitoid-exposure treatments. Phylogeographic studies (Sezonlin et al. 2006; Dupas et al. 2014) and reports of dominant inheritance of field-evolved resistance to *Bt* maize (Campagne et al.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

2013) represent the current frontier, in terms of studies of individual genes. We use the transcriptome to identify important candidate genes and gene families, as well as investigating the transcriptomic profiles of non-genic regions, for future research and ongoing management efforts aimed at curbing the effects of this devastating crop pest.

Methods

See Supplemental Methods for complete details. Briefly, sample collection, preparation, and sequencing were performed at the International Centre of Insect Physiology and Ecology (*icipe*) and Biosciences eastern and central Africa genomics facility at the International Livestock Research Institute in Nairobi, Kenya (BecA-ILRI Hub). Specimens were obtained from a colony initiated from larvae collected in maize fields from the Western Province of Kenya in 2008, and maintained on an artificial diet under laboratory conditions at *icipe*. RNA extractions were performed on 10 different *B. fusca* tissues types, (thoraces, antennae [both sexes], and female ovipositor), developmental stages (eggs, neonates, fourth instar larvae, and adults [both sexes]), and parasitoid-exposure conditions (unexposed and exposed to two strains of wasp; Table S1). We used a TruSeq RNA Library Prep Kit (Illumina) to prepare libraries and performed next-generation sequencing using Illumina MiSeq technology, resulting in 10 libraries of 300 bp paired-end reads (Table S2). Sequence data are available on NCBI SRA (BioProject: PRJNA553865). We used the RNA-seq *de novo* assembly program Trinity (version 2.4; Grabherr et al. 2011) to assemble the transcriptome (Table 1), and generated a set of non-redundant unigenes from our transcripts using the program CD-HIT (version 4.7; Li and Godzick 2006; Fu et al. 2012). The transcriptome assembly is available on NCBI (BioProject: PRJNA553865). We also assembled each library individually, and generated library-specific sets of non-redundant unigenes with CD-HIT. We ran BUSCO (version 3; Waterhouse et al. 2017) to assess the completeness of our assembly and Trinotate (v. 3.0.2; Bryant et al. 2017) to annotate the transcriptome. We quantified the distribution of genes within different gene ontology (GO) categories using WEGO (version 2.0; Ye et al. 2018; Figure S1). Finally, we used OrthoFinder (version 2.3.3; Emms and Kelly 2019) to identify orthologous gene clusters among Lepidopteran transcriptomes and identified the 20 largest gene families in *B. fusca.* We identified the top 25 most highly expressed transcripts for each RNA library by mapping reads from each library to the pooled transcriptome assembly unigenes using the program STAR (version 2.7; Dobin et al. 2013) and generated unique lists of the 25 transcripts with the highest RPKM values for each.

Results

Our *B. fusca* transcriptome (Table 1 and Table S2; BioProject PRJNA553865), which contained >99% of the conserved single-copy orthologs used by BUSCO to assess completeness, contains approximately 1/3 more transcripts than previously known for this species, greatly expanding the number of genes available for further investigation (Peterson et al. 2019). The median number of isoforms per gene was one for all individual libraries, with library 5_S5_L001 (infection treatment: exposed Cs-Inland) having the highest maximum number of isoforms for a single transcript (transcript

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

TRINITY_DN8760_c0_g1, with 45 isoforms reported by Trinity). Furthermore, the antenna libraries (4_S4_L001 and 3_S3_L001) had the highest percentage of genes with greater than one isoform per gene (21.1% and 19.4%, respectively), while the egg library (2_S2_L001) had the lowest (6.6%). We annotated the assembled *B. fusca* transcriptome using Trinotate and identified 22,707 protein-coding genes and generated a set of 39,445 proteins overall. We identified more protein-coding genes than were predicted in the previously published *B. fusca* whole genome assembly (19,417; Hardwick et al. 2019). Of the 22,707 predicted protein-coding genes found in this study, 19,401 aligned to the genome assembly. The additional 4,306 genes identified in this study may represent alternative transcripts or fusion transcripts that the transcriptome annotation software does not recognize as variants produced by a given locus or loci. Alternatively, it possible these sequences were unrecognized by the genome annotation software because the depth of sequencing coverage for the whole genome sequencing project was insufficient to fully assemble the genome given its size (~500 Mb) and repeat content (~50%).

Figure S1 gives an overview of the GO terms associated with all annotated transcripts in the *B. fusca* transcriptome assembly after 20,118 of genes were assigned to gene ontology categories using Trinotate. Using protein sets from *Bombyx mori*, *Manduca sexta*, and *Plutella xylostella,* we searched for orthologs in *B. fusca* and characterized gene families*.* The top 20 largest gene families identified by OrthoFinder are listed in Table S3. We report the most abundant transcripts across all libraries (Table S4) and in each individual tissue type, developmental stage, and infection treatment (Tables S5-7) and between male and female antennae (Table S8) and thoraces (Table S9). We also characterize the non-genic transcriptomic profiles by reporting the diversity of transposable elements (TEs) transcribed in each library (Table S10) and the evidence for transcription of putatively horizontally-transferred regions of the genome based on high identity between sequences in *B. fusca* and *Cotesia sesamiae* (both strains; Table S11).

Discussion

Because food security is an essential component of public health worldwide, managing pests that target major crop species is crucial (World Bank 2008). Here, we sequenced, assembled, and annotated the transcriptome of *B. fusca* to try and identify candidate genes for pest management strategies and to understand more about the biology of this species. This study expands on previously published transcriptomic resources in several ways. In particular, we analyzed data from 10 different RNA-Seq libraries spanning multiple life stages, tissue types, and sexes. The inclusion of these groups allowed us to identify additional transcripts relative to previous studies. Our full transcriptome assembly includes 240,022 sequences. In contrast, the only other *B. fusca* transcriptome published to date using the larval stage only identified 170,756 transcript sequences (Peterson et al. 2019). The current study also expands on the geographical range of *B. fusca* populations surveyed; while the previously published transcriptome is based on tissues from individuals that originated in South Africa, our work focuses on individuals originating from Western Province, Kenya. The inclusion of

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

these additional populations, developmental stages, and tissues will aid in the development of targeted pest control based on life cycle and behavior in this destructive species.

Based on our transcriptome (see Supplemental Methods and Results for complete details), we identified genes of interest putatively involved in chemoreception. A number of the transcripts with potential chemosensory function were abundant in the antennae relative to other tissue types. Specifically, 10 out of 25 transcripts uniquely highly expressed in the antennae had significant homology to General odorant-binding proteins (Table S7). General odorant-binding proteins are distributed in the sensilla, and function to bind odorants and transport them to odorant receptors. General odorant-binding proteins have been shown to be involved in the detection of plant volatiles (Vogt et al. 2002; Nardi et al. 2003; Maida et al. 2005; Liu et al. 2015). We also observed differences between male and female libraries in genes expressed in the antennae, which may be of interest for determining sex-specific chemosensory genes. For example, while both males and females showed high expression of General odorant-binding proteins, we identified suites of uniquely abundant General odorant-binding proteins that were distinct for each sex (Table S8). In addition, the female antenna library exhibited elevated abundance of two transcripts with high similarity to Pheromone-binding protein. Pheromone binding proteins are typically associated with the detection of female sex pheromones by males, but some studies have identified high levels of expression of these genes in female antennae of moths in the family Noctuidae (e.g., Callahan et al. 2000). The abundance of these transcripts in females may suggest a role of Pheromone-binding protein genes in detecting odorants other than sex pheromone in *B. fusca*.

We also sought to better understand immune response in *B. fusca*, by identifying transcripts expressed in larvae infected by the parasitoid *C. sesamiae*. *Busseola fusca* exhibits differential susceptibility to infection by *C. sesamiae* from coastal and inland subpopulations. To shed light on the biological mechanisms underlying this difference, we assessed gene expression in larva exposed to Cs-Inland and Cs-Coast, with the goal of determining immune genes highly expressed in each treatment. Fourteen of the 25 most highly expressed transcripts in the Cs-Coast treatment had significant sequence similarity to the gene Arylphorin (Table S5). Arylphorin was also the largest *B. fusca* gene family identified via our analysis of orthologous gene clusters (Table S4). Arylphorin is a storage protein produced throughout larval instars in Lepidoptera (Riddiford 1995), and has also been proposed to play a role in humoral immune defense in response to bacterial infection (Freitak et al. 2007). In contrast to the Cs-Coast treatment, larvae in the Cs-Inland treatment showed high expression of the genes Gloverin and Lebocin, both of which have antimicrobial function in Lepidoptera (reviewed in Yi et al. 2014). These results indicate that Arylphorin, Gloverin, and Lebocin may be good targets for future research to understand why Cs-Inland are able to successfully parasitize *B. fusca* larva, while Cs-Coast are not.

We identified several large gene families potentially of interest for understanding *B. fusca*'s response to insecticides. One of the largest *B. fusca* gene clusters we identified

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

was the Xanthine dehydrogenase family, encompassing five genes and nine isoforms. Xanthine dehydrogenase is known to be involved in the production of nitrogenous excretory compounds (Bergmann and Dikstein 1955; Cochran 1985), and has also been implicated in phytochemical detoxification (Slansky 1993). Previous research has found that duplication of the Xanthine dehydrogenase gene is related to resistance to organophosphorus insecticide resistance in mosquitoes (Coleman and Hemingway 1997). We also identified a large *B. fusca* gene family corresponding to the Transient receptor potential channel pyrexia gene. This family encompassed five genes and six isoforms. Transient receptor potential (TRP) genes are related to environmental perception in insects (Venkatachalam and Montell 2007; Fowler and Montell 2013), and are common targets of insecticides (Nesterov et al. 2015; Kandasamy 2017; Kwon et al. 2010). Additional research to identify the specific functions of the large Xanthine dehydrogenase and TRP pyrexia gene families in *B. fusca* may shed light on the potential for insecticide resistance in this species, and inform future control efforts. Future studies may also utilize our transcriptomic resources to better understand Cry toxin resistance by looking at tissue-specific expression of genes such as alkaline phosphatase, aminopeptidase N, cadherin Cry toxin receptor, ATP-binding cassette transporters, and mitogen-activated protein kinases (Peterson et al. 2017).

We identified a number of highly abundant transcripts in the antennae which were annotated as Replicase polyproteins (Table S7). Upon further investigation, these sequences had significant BLAST hits to iflavirus sequences. Iflaviruses are single-stranded RNA viruses that to date have been shown to exclusively infect arthropods (Van Oers 2010). RNA derived from iflaviruses occur at high concentrations in the antennae of infected individuals (Dos Santos et al. 2019). To our knowledge, iflavirus infection has not previously been observed in *B. fusca* and the significance of infection should be investigated further. To characterize the non-genic components of the transcriptome, we used the previously curated repeat library (from Hardwick et al. 2019) to search against the transcriptomes (Table S10). The majority of transposable elements in *B. fusca* are expressed in each of the individual transcriptomes, with the greatest diversity transcribed in male antennae. We also used low coverage draft sequences generated from the two parasitoid wasp strains (*C. sesamiae* Kitale and Mombasa) to identify candidate sequences that may have been horizontally-transferred to or from *B. fusca* (Supplemental Data File S6) and are transcribed (14 of 17 identified candidates; Table S11). Interestingly, three of the candidate regions appeared to be of bracoviral origin, suggesting that, as with many other insects (e.g., Gasmi et al. 2015), viruses may play an instrumental role in both the genomic and transcriptomic profile of *B. fusca*.

Conclusion

By sequencing a comprehensive set of ten libraries representing different developmental stages, tissue types, and parasitoid exposures, our transcriptome contains the largest number of expressed transcripts available to date and has allowed us to identify a number of loci of interest for potential pest management. Namely, loci potentially involved in chemoreception, response to infection, and response to

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

insecticides could serve as promising targets for future research. We hope insights from the *B. fusca* transcriptome will aid ongoing efforts to develop control measures that can be deployed as part of an integrated pest management strategy in order to reduce *B. fusca's* impact on food security across the continent.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

Acknowledgements

We would like to thank the leadership and staff at the International Livestock Research Institute in Nairobi, Kenya who supported the *Busseola fusca* Genomics Consortium genome annotation jamboree and workshop which took place in April 2014. We would also like to gratefully acknowledge the scientists and staff at the BecA-ILRI Hub, specifically, for their support for this project, including Joyce Njuguna, Jason Rogena, and Alan Orth. We would also like to acknowledge our funding support from the Institute of Research for Development (*Icipe*-NSBB-B4405B; to PAC) and from the National Science Foundation (MCB-1150213), the Reed College sabbatical fellowship program, and the Fulbright Foundation (to SS).

Data Availability

Data files are accessible as Supplementary Data file and sequences are in GenBank (BioProject PRJNA553865). Sequence data and assemblies for individual libraries can be found under the following accessions: 2_S2_L001: SRR12107984, GITF00000000; 1_S1_L001: SRR12107983, GITJ00000000; S2_S1_L001: SRR12107982, GITA00000000; 5_S5_L001: SRR12107981, GITI00000000; S3_S2_L001: SRR12107980, GITB00000000; S10_S5_L001: SRR12107985, GISZ00000000; 4_S4_L001: SRR12107986, GITH00000000; 3_S3_L001: SRR12107987, GITG00000000; S7_S4_L001: SRR12107989, GITE00000000; S6_S3_L001: SRR12107988, GITD00000000. The comprehensive assembly is under accession GISL00000000.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

References

Bergmann, F., & Dikstein, S. (1955). The relationship between spectral shifts and structural changes in uric acids and related compounds. Journal of the American Chemical Society, 77(3), 691-696.

Branca, A., Le Ru, B. P., Vavre, F., Silvain, J. F., & Dupas, S. (2011). Intraspecific specialization of the generalist parasitoid *Cotesia sesamiae* revealed by polyDNAvirus polymorphism and associated with different Wolbachia infection. Molecular Ecology, 20(5), 959-971.

Bryant, D. M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M. B., Payzin-Dogru, D., Lee, T. J., Leigh, N. D., Kuo, T. H., Davis, F. G., Bateman, J., Bryant, S., Guzikowski, A. R., Tsai, S. L., Coyne, S., Ye, W. W., Freeman, R. M. Jr, Peshkin, L., Tabin, C. J., Regev, A., Haas, B. J., & Whited, J. L. (2017). A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. Cell Reports, 18(3), 762-776.

Calatayud, P. A., Le Ru, B. P., Van den Berg, J., & Schulthess, F. (2014). Ecology of the African maize stalk borer, *Busseola fusca* (Lepidoptera: Noctuidae) with special reference to insect-plant interactions. Insects, 5(3), 539-563.

Callahan, F. E., Vogt, R. G., Tucker, M. L., Dickens, J. C., & Mattoo, A. K. (2000). High level expression of “male specific” pheromone binding proteins (PBPs) in the antennae of female noctuiid moths. Insect Biochemistry and Molecular Biology, 30(6), 507-514.

Campagne, P., Kruger, M., Pasquet, R., Le Ru, B., & Van den Berg, J. (2013). Dominant inheritance of field-evolved resistance to Bt corn in *Busseola fusca*. PLoS One, 8(7), e69675.

Cochran, D. G. (1985). Nitrogen excretion in cockroaches. Annual Review of Entomology, 30(1), 29-49.

Coleman, M., & Hemingway, J. (1997). Amplification of a xanthine dehydrogenase gene is associated with insecticide resistance in the common house mosquito *Culex quinquefasciatus*. Biological Society Transactions, 526S.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics, 29(1), 15-21.

Dos Santos, E. R., Trentin, L. B., Ecker, A., Silva, L. A., Borges, M., Mowery, J. D., Bergmann, M. R., Harrison, R. L., & Ardisson-Araujo, D. M. P. (2019). An iflavirus found in stink bugs (Hemiptera: Pentatomidae) of four different species. Virology, 534, 72-79.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

Dupas, S., Le Ru, B., Branca, A., Faure, N., Gigot, G., Campagne, P., Sezonlin, M., Ndemah, R., Ong'amo, G., Calatayud, P. A. & Silvain J. F. (2014). Phylogeography in continuous space: coupling species distribution models and circuit theory to assess the effect of contiguous migration at different climatic periods on genetic differentiation in *Busseola fusca* (Lepidoptera: Noctuidae). Molecular Ecology, 23(9), 2313-2325.

Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biology, 20(1), 1-14.

Fowler, M. A., & Montell, C. (2013). Drosophila TRP channels and animal behavior. Life Sciences, 92(8-9), 394-403.

Freitak, D., Wheat, C. W., Heckel, D. G., & Vogel, H. (2007). Immune system responses and fitness costs associated with consumption of bacteria in larvae of *Trichoplusia* ni. BMC Biology, 5(1), 56.

Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics, 28(23), 3150-3152.

Gasmi L., Boulain H., Gauthier J., Hua-Van A., Musset K., Jakubowska A.K., Aury J.M., Volkoff A.N., Huguet E., Herrero S., Drezen J.M. (2015) Recurrent domestication by Lepidoptera of genes from their parasites mediated by bracoviruses. PLoS Genetics, 11(9):e1005470.

Gitau, C.W, Schulthess, F. & Dupas, S. (2010). An association between host acceptance and virulence status of different populations of *Cotesia sesamiae*, a braconid larval parasitoid of lepidopteran cereal stemborers in Kenya. Biol. Control 54, 100-106.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., & Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology, 29(7), 644-652.

Hardwick, K. M., Ojwang', A. M. E., Stomeo, F., Maina, S., Bichang'a, G., Calatayud, P. A., Filée, J., Djikeng, A., Miller, C., Cepko, L., Darby, A. C., Le Ru, B., & Schaack, S. (2019). Draft genome of *Busseola fusca*, the maize stalk borer, a major crop pest in sub-Saharan Africa. Genome Biology and Evolution, 11(8), 2203-2207.

Jiang, N., Zhou, G., Overholt, W. A., & Schulthess, F. (2006). The synchrony of the stemborer and parasitoid populations of coastal Kenya. Ann. Soc. Entomol. Fr, 42, 381-388.

Kandasamy, R., London, D., Stam, L., von Deyn, W., Zhao, X., Salgado, V. L., & Nesterov, A. (2017). Afidopyropen: new and potent modulator of insect transient receptor potential channels. Insect Biochemistry and Molecular Biology, 84, 32-39.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

Kfir, R. (1995). Parasitoids of the African stem borer, *Busseola fusca* (Lepidoptera: Noctuidae), in South Africa. Bulletin of Entomological Research, 85(3), 369-377.

Kfir, R., Overholt, W. A., Khan, Z. R., & Polaszek, A. (2002). Biology and management of economically important lepidopteran cereal stem borers in Africa. Annual Review of Entomology, 47(1), 701-731.

Koizumi, N., Imamura, M., Kadotani, T., Yaoi, K., Iwahana, H., & Sato, R. (1999). The lipopolysaccharide-binding protein participating in hemocyte nodule formation in the silkworm *Bombyx mori* is a novel member of the C-type lectin superfamily with two different tandem carbohydrate-recognition domains. FEBS letters, 443(2), 139-143.

Koizumi, N., Morozumi, A., Imamura, M., Tanaka, E., Iwahana, H., & Sato, R. (1997). Lipopolysaccharide-binding proteins and their involvement in the bacterial clearance from the hemolymph of the silkworm *Bombyx mori*. European Journal of Biochemistry, 248(1), 217-224.

Kwon, Y., Shen, W. L., Shim, H. S., & Montell, C. (2010). Fine thermotactic discrimination between the optimal and slightly cooler temperatures via a TRPV channel in chordotonal neurons. Journal of Neuroscience, 30(31), 10465-10471.

Li, W., & Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics, 22(13), 1658-1659.

Liu, N. Y., Yang, K., Liu, Y., Xu, W., Anderson, A., & Dong, S. L. (2015). Two general-odorant binding proteins in *Spodoptera litura* are differentially tuned to sex pheromones and plant odorants. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 180, 23-31.

Maida, R., Mameli, M., Müller, B., Krieger, J., & Steinbrecht, R. A. (2005). The expression pattern of four odorant-binding proteins in male and female silk moths, *Bombyx mori*. Journal of Neurocytology, 34(1-2), 149.

Mochiah, M. B., Ngi-Song, A. J., Overholt, W. A., & Stouthamer, R. (2002). Variation in encapsulation sensitivity of *Cotesia sesamiae* biotypes to *Busseola fusca*. Entomologia experimentalis et applicata, 105(2), 111-118.

Nardi, J. B., Miller, L. A., Walden, K. K., Rovelstad, S., Wang, L., Frye, J. C., Ramsdell, K., Deem, L. S., & Robertson, H. M. (2003). Expression patterns of odorant-binding proteins in antennae of the moth *Manduca sexta*. Cell and Tissue Research, 313(3), 321-333.

Nesterov, A., Spalthoff, C., Kandasamy, R., Katana, R., Rankl, N. B., Andrés, M., Jahde, P., Dorsch, J. A., Stam, L. F., Braun, F., Warren, B., Salgado, V. L., & Gopfert, M. C. (2015). TRP channels in insect stretch receptors as insecticide targets. Neuron, 86(3), 665-671.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

Ngi-Song, A. J., Overholt, W. A., & Ayertey, J. N. (1995). Suitability of African gramineous stemborers for development of *Cotesia flavipes* and *C. sesamiae* (Hymenoptera: Braconidae). Environmental Entomology, 24(4), 978-984.

Peterson, B., Bezuidenhout, C. C., & Van den Berg, J. (2017). An overview of mechanisms of Cry toxin resistance in lepidopteran insects. Journal of Economic Entomology, 110(2), 362-377.

Peterson, B., Sanko, T. J., Bezuidenhout, C. C., & Van den Berg, J. (2019). Transcriptome and differentially expressed genes of *Busseola fusca* (Lepidoptera: Noctuidae) larvae challenged with Cry1Ab toxin. Gene, 710, 387-398.

Riddiford, L. M. (1995). Hormonal regulation of gene expression during lepidopteran development, in Molecular Model Systems in the Lepidoptera, editors M. R. Goldsmith and A. S Wilkins. Cambridge University Press, pp. 293-322.

Sallam, M. N., Overholt, W. A., & Kairu, E. (1999). Comparative evaluation of *Cotesia flavipes* and *C. sesamiae* (Hymenoptera: Braconidae) for the management of *Chilo partellus* (Lepidoptera: Pyralidae) in Kenya. Bulletin of Entomological Research, 89(2), 185-191.

Sezonlin, M., Dupas, S., Le Rü, B., Le Gall, P., Moyal, P., Calatayud, P. A., Giffard, I., Faure, N. & Silvain, J. F. (2006). Phylogeography and population genetics of the maize stalk borer *Busseola fusca* (Lepidoptera, Noctuidae) in sub-Saharan Africa. Molecular Ecology, 15(2), 407-420.

Slansky, F. (1993). Xanthine toxicity to caterpillars synergized by allopurinol, a xanthine dehydrogenase/oxidase inhibitor. Journal of Chemical Ecology, 19(11), 2635-2650.

Songa, J. M., Jiang, N., Schulthess, F., & Omwega, C. (2007). The role of intercropping different cereal species in controlling lepidopteran stemborers on maize in Kenya. Journal of Applied Entomology, 131(1), 40-49.

Van den Berg, J., Van Rensburg, J. B. J., & Pringle, K. L. (1991). Comparative injuriousness of *Busseola fusca* (Lepidoptera: Noctuidae) and *Chilo partellus* (Lepidoptera: Pyralidae) on grain sorghum. Bulletin of Entomological Research, 81(2), 137-142.

van Oers, M. M. (2010). Genomics and biology of Iflaviruses. Insect Virology, 231-250.

Venkatachalam, K., & Montell, C. (2007). TRP channels. Annual Review of Biochemistry, 76, 387-417.

Vogt, R. G., Rogers, M. E., Franco, M. D., & Sun, M. (2002). A comparative study of odorant binding protein genes: differential expression of the PBP1-GOBP2 gene cluster

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

in Manduca sexta (Lepidoptera) and the organization of OBP genes in *Drosophila melanogaster* (Diptera). Journal of Experimental Biology, 205(6), 719-744.

Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E. V., & Zdobnov, E. M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. Molecular Biology and Evolution, 35(3), 543-548.

World Bank (2008) World Development Report 2008: Agriculture for Development. The World Bank, Washington, District of Columbia.

Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., Xu, H., Huang, X., Li, S., Zhou, A., & Zhang, X. (2018). WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. Nucleic Acids Research, 46(W1), W71-W75.

Yi, H. Y., Chowdhury, M., Huang, Y. D., & Yu, X. Q. (2014). Insect antimicrobial peptides and their applications. Applied Microbiology and Biotechnology, 98(13), 5807-5822.

You, M., Yue, Z., He, W., Yang, X., Yang, G., Xie, M., Zhan, D., Baxter, S. W., Vasseur, L., Gurr, G. M., & Douglas, C. J. (2013). A heterozygous moth genome provides insights into herbivory and detoxification. Nature Genetics, 45(2), 220-225.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

List of Abbreviations

Cs-Coast: *Cotesia sesamiae* strain from Mombasa (coastal Kenya)
Cs-Inland: *Cotesia sesamiae* strain from Kitale (inland Kenya)
Bt maize: transgenic (*Bacillus thuringiensis*) corn that produces Cry proteins (crystal-like proteins)
NCBI SRA: National Center Biotechnology Information Short Read Archive
RPKM: Reads Per Kilobase of transcript, per Million mapped reads

List of Figures and Tables

Figure 1. *Busseola fusca* (A) eggs, (B) neonates after hatching, (C) 4th instar larva, (D) parasitoid wasp, *Cotesia sesamiae* (female), used for exposure treatment, and (E) adult (female).

Table 1. Assembly statistics for *B. fusca* transcriptome (Bioproject PRJNA553865*)*.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

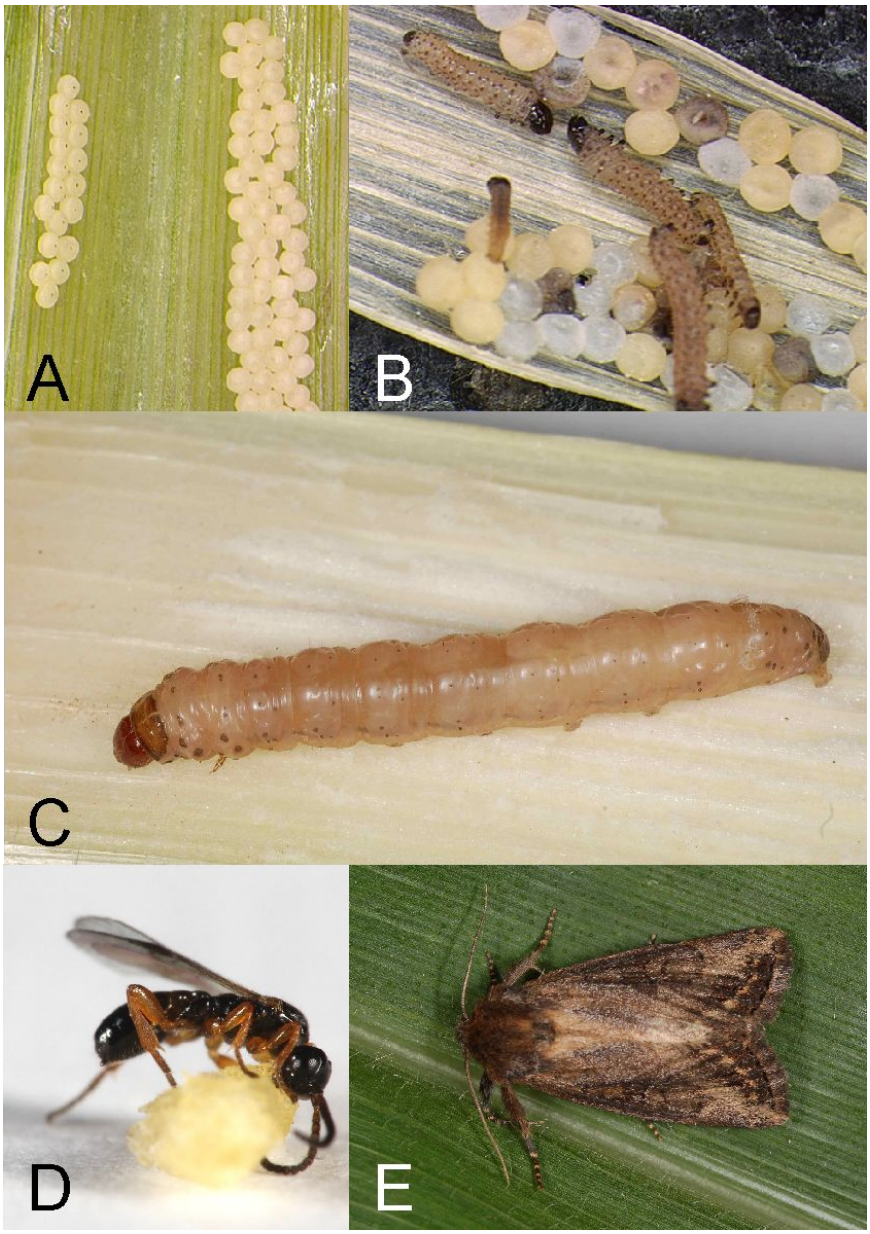


Figure 1. *Busseola fusca* (A) eggs, (B) neonates after hatching, (C) 4th instar larva, (D) parasitoid wasp, *Cotesia sesamiae* (female), used for exposure treatment, and (E) adult (female).

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020

Table 1. Assembly statistics for *B. fusca* transcriptome from 10 pooled libraries derived from tissue sampled from specimens originally collected in Western Province, Kenya.

	Transcriptome (Full Assembly)	**Transcriptome (Unigenes)**
Total size (bp)	171,925,619	116,455,446
Number of transcripts	240,022	185,159
Number of genes	129,814	123,344
Number of protein-coding genes	23,374	22,707
Number of proteins	59,341	39,445
Maximum transcript length (bp)	28,079	28,079
Median transcript length (bp)	398	377
% GC	40.1	40.2
% BUSCO genes- Eukaryota	99.6	99.6
% BUSCO genes- Arthropoda	99.1	99.2
% BUSCO genes- Insecta	98.7	98.6

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa195/5908520 by guest on 16 November 2020