MAGGI F., MARZACHI, C. and D. BOSCO. 2013. A stage-structured model of *Scaphoideus titanus* in vineyards. Environmental Entomology **42**(2): 181-193.

MOORE, J., LIANG, S., AKULLIAN, A. and J. REMAIS. 2012. Cautioning the use of degree-day models for climate change projections in the presence of parametric uncertainty. Ecological Applications **22**(8): 2237–2247.

MOORE J., and J. REMAIS. 2014. Developmental models for estimating ecological responses to environmental variability: structural, parametric, and experimental issues. Acta Biotheor **62**: 69-90.

NACHMAN G., and T. GOTOH. 2015. Modeling the effects of constant and variable temperatures on the vital rates of an age-, stage-, and sex-structured population by means of the SANDY approach. Environmental Entomology **44** (3): 821-834.

NISBET, R. M., and W. S. C. GURNEY. 1982. The systematic formulation of population models for insects with dynamically varying instar duration. Theoretical Population Biology **23**: 114-135.

REGNIERE, J., POWELL, J., BENTZ, B. and V. NEALIS. 2012. Effects of temperature on development, survival and reproduction of insects: experimental design, data analysis and modeling. Journal of Insect Physiology **58**: 634-647.

SCHAALJE G. B., and H. R. VAART 1989. Relationships among recent models for insect population dynamics with variable rates of development. Journal of Mathematical Biology **27**: 399-428.

SCHNEIDER, S. M., and H. FERRIS. 1986. Estimation of stage-specific developmental times and survivorship from stage frequency data. Researches on Population Ecology **28**: 267-280.

SCHOOLFIELD, R. M., SHARPE, P. J. H. and C. E.MAGNUSON. 1981. Nonlinear regression of biological temperature-dependent rate models based on absolute reaction-rate theory. Journal of Theoretical Biology **88** (4): 719–731.

SHARPE, P. J. H., and D. W. DEMICHELE. 1977. Reaction kinetics of poikilotherm development. Journal of Theoretical Biology **64**(4): 649–670.

SKOVGARD, H., and G. NACHMAN. 2017. Modeling the temperature- and age-dependent survival, development, and oviposition rates of stable flies (*Stomoxys calcitrans*) (Diptera: Muscidae). Environmental Entomology **46** (5): 1130-1142.

STINNER, R. E., GUTIERREZ, A. P. and G. D. BUTLER. 1974. An algorithm for temperature-dependent growth rate simulation. The Canadian Entomologist **105**: 145- 156.

TURCHIN, P. 1990. Rarity of density dependence or population regulation with lags? Nature **344**: 660-663.

TURCHIN, P. and A TYLOR. 1992. Complex dynamics in ecological time series. Ecology **73**(1): 289-305.

VANSICKLE, J. 1977. Analysis of a distributed parameter population model based on physiological age. Journal of Theoretical Biology **64**: 571-586.

WAGNER, T. L., WU, H., SHARPE, P. J. H., SCHOOLFIELD, R. M. and R. N. COULSON. 1984a. Modeling insect development rates: a literature review and application of a biophysical model. Annals of the Entomological Society of America **77**: 208-225.

WAGNER, T. L., WU, H. I., SHARPE, P. J. H. and R. N. COULSON. 1984b. Modeling distributions of insect development time: a literature review and application of the Weibull function. Annals of the Entomological Society of America **77**: 475–487.

WANG, E., and T. ENGEL. 1998. Simulation of phenological development of wheat crops. Agricultural Systems **58**: 1–24.

WORNER, S. P. 1992. Performance of phenological models under variable temperature regimes: consequences of the Kaufmann or rate summation effect. Environmental Entomology **2l** (4): 689-699.

High Quality Genomic Resources for Stored Product Insects

Erin D. Scully[1]*, Scott M. Geib[2], Sheina B. Sim[2]

[1]Stored Product Insect and Engineering Research Unit, USDA-ARS-Center for Grain and Animal Health Research, Manhattan, KS 66502, USA.
[2]Tropical Crop and Commodity Protection Research Unit, USDA-ARS-Daniel K. Inouye U.S. Pacific Basin Agricultural Research Center, Hilo, HI 96720, USA.
*Corresponding authour: Erin.Scully@ars.usda.gov
DOI 10.5073/jka.2018.463.027

Abstract

The expansion of genomic resources for stored product insects has largely been hampered by cost, time required for inbreeding, and technical issues that can arise during genome assembly from pooling multiple individuals together for DNA isolation and library preparation. However, newer library methods, such as 10X Chromium libraries, largely overcome these issues in that sufficient DNA can be recovered from a single individual for library prep and allelic variants are assembled as separate phase blocks, eliminating the need for inbreeding. Using 10X Chromium libraries coupled with 150 x 150 bp HiSeqX sequencing to a depth of at least 60X coverage, we are developing high quality draft genome assemblies for eight different stored product insect species, including Dermestidae (*Trogoderma variabile*, *Trogoderma granarium*, and *Dermestes maculatus*), Tenebrionidae (*Tribolium confusum*), Anobiidae (*Lasioderma serricorne* and *Stegobium paniceum*), Bostrichidae (*Prostephanus truncatus*), and Pyralidae (*Plodia interpunctella*). Overall, BUSCO (Benchmarking Using Single Copy Orthologs) scores exceeded 95% in all assemblies with few fragmented or duplicated genes, suggesting a high quality assembly of the gene space. Further, scaffold N50s exceeded 1 Mb in many cases and further

improvements to these scaffolding metrics will be made using linkage maps and Hi-C libraries. Overall, this approach will yield high quality assemblies for eight different insects and could be used to quickly and efficiently generate draft assemblies of invasive or emerging stored product pests.

Keywords: khapra beetle, Bostrichidae, Dermestidae, Anobiidae, Pyralidae.

Introduction

Genome sequences have provided tremendous insight into the physiological and metabolic capabilities of various insect species, led to the identification of causative mutations associated with pesticide and fumigant resistance (Schlipalius et al., 2012), facilitated the identification of taxonomically informative loci for DNA barcoding (Chesters et al., 2015), and identified copy number expansions that may allow insects to exploit new ecological niches (McKenna et al., 2016). Despite these utilities, only one stored product insect genome is publicly available (*Tribolium castaneum*) (Tribolium sequencing consortium, 2008) while a small, but growing number of transcriptome assemblies, are available for other stored product species. Since its initial assembly, the *T. castaneum* reference assembly has been used to identify mutations in a gene coding for dihydrolipoamide dehydrogenase (DLD) associated with phosphine resistance (Schlipalius et al., 2012), biorational gene targets for pest control via RNAi (Dönitz et al., 2014), and causative mutations associated with sensory system defects (Angelini et al., 2009). This assembly has even facilitated the discovery of a mutation associated with phosphine resistance in lesser grain borer (*Rhyzopertha dominica*), which happens to occur in a DLD ortholog (Schlipalius et al., 2012). However, the biologies of stored product insects vary tremendously across various taxonomic groups. Further, different taxonomic groups may evolve different strageies for overcoming biotic and abiotic stresses. In this case, having genome references available would greatly facilitate genome wide association analyses to identifiy causative mutations associated with tolerance to stress. In other cases, some species are more inherently tolerant to certain biotic and abiotic stresses and genome sequences could lead to the identification of genetic factors associated with tolerance.

Historically, genome sequencing for insects has been cost prohibitive; however, new library approaches coupled with the reduced cost of sequencing is making genome assembly more affordable and accessible than ever. One major challenge faced by those working with insects is obtaining sufficient quantities of DNA for library preparation and assembly. For Illumina mate-pair and PacBio long-read libraries, over 10 ug of DNA must be provided. This requires pooling multiple individuals for sequencing. This practice leads to multiple allelic variants derived from the same locus in the DNA pool, which can introduce bubbles and breaks into the assembly graph, reducing overall contiguity. Although sequence variations among individuals can be reduced through inbreeding for several generations, this often involves a significant time investment and the number of backcrosses required to obtain homogeneity varies by species and their recombination rates. One new library approach (10X Chromium) largely overcomes these limitations in that sufficient quantities of DNA can be recovered from a single insect for library construction and haplotypes are assembled as separate phase blocks, reducing the number of bubbles in the assembly graph and improving contiguity. In addition, during library construction, HMW DNA is separated into microfluidic chambers designed to hold exactly one molecule of DNA per chamber. Within each chamber, DNA is fragmented and DNA derived from the same molecule is tagged with the same barcode. In this manner, sequencing reads with the same barcode can be linked together during the assembly stage to form long scaffolds and contigs. This approach greatly improves assembly contiguity compared to other short read assembly methods.

In order to improve genomic resources for stored product insects, we sequenced 10X Chromium libraries derived from eight different species of stored product insects from the families Dermestidae (*Trogoderma variabile*, *Trogoderma granarium*, and *Dermestes maculatus*), Tenebrionidae (*Tribolium confusum*), Anobiidae (*Lasioderma serricorne* and *Stegobium paniceum*), Bostrichidae (*Prostephanus truncatus*), and Pyralidae (*Plodia interpunctella*). Genomes were assembled using the program Supernova and assemblies will be superscaffolded to chromosome

scale using linkage maps and/or chromatin contact maps. Overall, these assemblies exceeded the quality of many publicly beetle genome assemblies and thus, represent a viable strategy for generating genome sequences for underreprented groups of insects. Not only will these assemblies be useful for mapping traits, conducting population genetics studies, and understanding genetic similarities and differences between various stored product species, but they will also allow for broader evolutionary analyses regarding gene order, gene duplications, and the evolution of different gene families across different taxonomic groups. Such broad scale analyses can lead to the identification of convergent strategies for overcoming stress, such as mutations in orthologous genes associated with stress response shared across species, and can also shed light on family-, genus-, or species-specific adaptations.

Materials and Methods

High molecular weight (HMW) DNA was isolated from single individuals using several different approaches. For *S. paniceum*, *T. confusum*, *T. granarium*, *D. maculatus*, and *P. interpunctella*, DNA was isolated using the Qiagen MagAttract HMW DNA Kit (Gaithersburg, MD) following the manufacturer's directions. Unfortunately, insufficient quantites of HMW DNA were recovered from *L. serricorne*, *P. truncatus*, and *T. variabile* for 10X Chromium library preparation using this approach, so other approaches were attempted. For *L. serricorne* and *T. variabile*, an agarose isolation method was used. In brief, single insects were macerated with a pestle in a nuclei isolation buffer containing 2% Triton X-100 (w/v), 10 mM EDTA, 100 mM KCl, 4mM spermidine, 1 mM spermine, and 17.1% sucrose. 50 µL of the supernatant was transferred to 75 µL molten InCert agarose (Lonza, Basel, Switzerland). The solution was placed in a gel mold and allowed to set for 10 mins at 4°C. The sample was lysed overnight at 50°C in a solution containing proteinase K and 1% sarkosyl. The gel plugs were rinsed in TE buffer and proteinase K was deactivated for 1 hour using phenylmethane sulfonyl fluoride (PMSF). Four washes in TE buffer were used to remove PMSF and DNA was recovered from the gel plug using an agarase treatment (Wieslander, 1979). For *P. truncatus*, a salting out approach was employed. In brief, insects were macerated in a lysis buffer containing 10 mM Tris-HCl, 400 mM NaCl, 2 mM EDTA, and 0.5% SDS, and incubated overnight at 37°C with mixing. 1.2 mL of 5 mM NaCl solution was added to 'salt out' the DNA and the sample was centrifuged for 15 mins at 4°C under low speed (1,000 x g). The supernatant containing the HMW DNA was washed with ethanol and centrifuged at medium speed (6250 x g) for 5 mins. Ethanol was removed and the pellet was resuspended in TE buffer. The full protocol can be found at https://support.10xgenomics.com/genome-exome/sample-prep/doc/demonstrated-protocol-salting-out-method-for-dna-extraction-from-cells.

In all cases, the final concentration of the DNA was validated using the dsDNA High Sensitvity Assay on the Qubit Fluorimeter (Thermo Fisher Scientific, Waltham, MA) and, when sufficient DNA quantities were available, the quality of the DNA was validated using Pulsed-Field Gel Electrophoresis (PFGE). 1-5 ng of HMW DNA were used to make 10X Chromium libraries at Hudson Alpha Biotechnology Institute (Huntsville, AL) and libraries were sequencing using 150 x 150 bp reads on the Illumina HiSeq X-ten instrument to a depth of at least 60X. All genomes were assembled using the Supernova assembler with barcode subsampling in order to normalize coverage across the barcodes and improve the contiguity of the assemblies. Subsampling was performed from 30 to 70% to determine how much subsampling was needed to produce the most contiguous assembly, which was gauged using the programs QUAST (Gurevich et al., 2013) and BBTools (https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/). Criteria that were used to select the most optimal assembly included scaffold N50, maximum scaffold length, recovery of conserved single copy orthologs (BUSCOs) (Simao et al., 2015), and scaffold length accumulation curves.

Once the Supernova assemblies were finalized, either linkage maps or HiC chromatin contact maps were prepared in order to obtain chromosome scale assemblies. Linkage maps were pursued for non-quarantine and non-invasive species with short generation times that were relatively easy to

rear, including *L. serricorne*, *T. variable*, *D. maculatus*, and *T. confusum*. In brief, crosses were set up between males and females collected from different and isolated populations of each insect species. F_1 individuals were collected and sibling crosses were performed from the F_1 through the F_3 generations to facilitate recombination. At the F_4 generation, four sets of F_3 parents and 40-50 offspring derived from each set of parents were collected for genotyping using ddRAD-Seq. Samples were barcoded and sequenced on a single lane on the Illumina HiSeq 2500 platform using 100 x 100 bp paired end libraries. Prior to mapping, the genome was masked for repeats using RepeatModeler for *de novo* repeat analysis and RepeatMasker. The *T. castaneum* repeat library was used for masking in addition to the predicted repeats from RepeatModeler. The Stacks pipeline is currently being used for genotyping and variant calling and LepMap-3 will be used to identify and order linkage groups (Rastas et al., 2015). For HiC contact maps of *T. granarium*, *P. interpunctella*, *P. truncatus*, and *S. paniceum*, pools of insects were macerated and treated with formaldehyde to cross-link DNA and chromatin complexes (Belton et al., 2012). Endonucleases and restriction enzymes were used to digest the uncrosslinked regions, cross linking was reversed, and DNA was sequenced to a depth of at least 50 million 100 x 100 paired end reads on a HiSeq 2500 platform. Reads will be eventually mapped to the 10X assemblies using BWA (Li and Durbin, 2009) and LACHESIS (Burton et al., 2013) will be used to identify regions of DNA that shared chromatin contacts.

Results

HMW DNA was successfully acquired from single insects in sufficient quantities to generate 10X Chromium libraries for all eight stored product species. No major differences in DNA quality were noted across species. It was more difficult to obtain HMW DNA from *P. truncatus* using either the gel plug extraction method or the Qiagen MagAttract kit; however, HMW DNA was obtained using a salting out approach that had been previously used for prepare HMW DNA for 10X Chromium libraries for other insect species. Additionally, HiSeqX yields across all eight species were relatively consistent and ranged from 700-875 million reads. All reads were initially used for Supernova assemblies; however, barcode subsampling was employed to normalize read coverage across barcodes, which can vary significantly for genomes smaller than 1 Gb. The amount of barcode subsampling required to produce the most contiguous assembly varied by predicted genome size, with larger genomes requiring less subsampling compared to smaller genomes. Genome size estimates ranged from 150 Mb (*L. serricorne*) to 500 Mb (*D. maculatus*). Examples of assembly improvements with subsampling are shown for *T. variabile* and *L. serricorne* in Tables 1 and 2. For *T. variabile*, subsampling 160 million reads generated the best assembly metrics, including longest contig and scaffold N50s, longest maximum contig and scaffold lengths and the highest percentage of the assembly in scaffolds > 50 Kb (Table 1). Additionally, a significantly higher percentage of the genome was present in long scaffolds when 160 million reads were subsampled (Figure 1a). Similarly, subsampling 350 million reads led to the best assembly metrics for *L. serricorne* (Table 2) while subsampling either 300 or 350 million reads led to the highest percentages of the assembly in long scaffolds (Figure 1b).

Tab. 1 Assembly improvement with barcode subsampling for *Trogoderma variabile*. Assemblies were performed using Supernova with various levels of subsampling to generate the most contiguous assembly. Subsampling 160 million reads of a total of 812 million reads led to the assembly with the highest contig and scaffold N50s, the highest maximum contig and scaffolds lengths, and the highest percentage of the genome in scaffolds > 50 kb. Thus, this assembly was selected as the most optimal.

	140M	**150M**	**160M**	**170M**	**All**
Number contigs	12,688	12,643	12,903	13,622	32,292
Contig N50	250 Kb	277 Kb	294 Kb	274 Kb	33 Kb
Max Contig Length	748 Kb	605 Kb	650 Kb	619 Kb	185 Kb
Number of Scaffolds	7,003	6,894	7,122	7,687	21,917
Scaffold N50	3.8 Mb	4.9 Mb	7.0 Mb	6.3 Mb	1.0 Mb
Max Scaffold Length	17 Mb	14 Mb	22 Mb	17 Mb	8.7 Mb

Number of Scaffolds > 10 Kb	640	679	586	722	1,727
% of Genome in Scaffolds > 50 Kb	89.3%	89.2%	89.7%	88.1%	71.4%
Total Contig Assembly Length	264 Mb	265 Mb	265 Mb	269 Mb	272 Mb
Total Scaffold Assembly Length	271 Mb	274 Mb	273 Mb	278 Mb	310 Mb
% Gap	2.6%	2.8%	2.8%	3.1%	12.4%

Tab. 2 Assembly improvement with barcode subsampling for *Lasioderma serricorne*. Assemblies were performed using Supernova with various levels of subsampling to generate the most contiguous assembly. Subsampling 350 million reads of a total of 870 million reads led to the assembly with the highest contig and scaffold N50s, highest max contig and scaffold lengths, and the largest percentage of the assembly present in scaffolds > 10 Kb. Thus, this assembly was selected as the most optimal.

	160M	**270M**	**300M**	**350M**	**400M**	**All**
Number contigs	9,705	10,270	10,288	10,057	10,405	17,892
Contig N50	52 Kb	61 Kb	69 Kb	79 Kb	72 Kb	27 Kb
Max Contig Length	776 Kb	3.8 Mb	2.1 Mb	3.8 Mb	2.0 Mb	1.6 Mb
Number of Scaffolds	9,266	9.653	9,594	9.315	9,540	16,588
Scaffold N50	55 kb	75 Kb	87 Kb	119 Kb	118 Kb	38 Kb
Max Scaffold Length	851 Kb	3.8 Mb	3.8 Mb	3.8 Mb	2.9 Mb	2.8 Mb
Number of Scaffolds > 10 Kb	2,562	2,313	2,105	1,963	2,125	3,061
% of Genome in Scaffolds > 50 Kb	56.7%	61.1%	64.4%	66.4%	64.2%	42.0%
Total Contig Assembly Length	143 Mb	154 Mb	154 Mb	156 Mb	161 Mb	160 Mb
Total Scaffold Assembly Length	144 Mb	155 Mb	155 Mb	154 Mb	159 Mb	164 Mb
% Gap	0.1%	0.2%	0.2%	0.2%	0.3%	0.4%

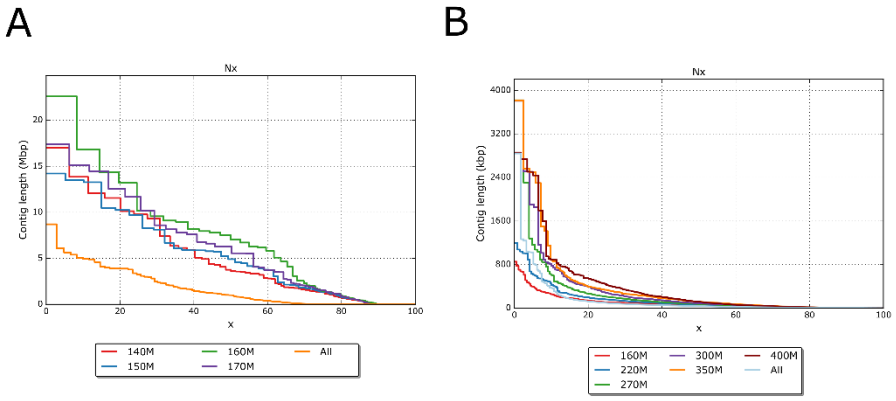

Fig. 1 Scaffold size distribution for a) *Trogoderma variabile* and b) *Lasioderma serricorne*. Y-axis represents scaffold length and X-axis represents number of scaffolds at each length. For *T. variabile*, subsampling 160 million reads led to the highest representation of long scaffolds in the assembly and for *L. serricorne,* subsampling either 300 or 350 million reads led to the highest representation of long scaffolds.

Assemblies for *T. granarium T. confusum*, and *D. maculatus* were similarly optimized. Although the assembly contiguity varied among these three insects, with contig N50s ranging from X to Y, scaffold N50s ranging from Z to Q, maximum contig lengths ranging from Y to Z, and maximum scaffold lengths ranging from blah to blah, all three assemblies had over 80% of their total assembly lengths in scaffolds > 50 kb and over 94% of conserved single copy orthologs were detected (Tables 3 and 4). While we are still awaiting sequencing data for *P. interpunctella, P. truncatus, and S. paniceum*, libraries have been prepared. We are also in the process of improving our scaffolding metrics using linkage maps and HiC contact maps.

Tab. 3 Assembly metrics for Trogoderma granarium, Tribolium confusum, and Dermestes maculatus optimized with barcode subsampling.

	T. granarium	***T. confusum***	***D. maculatus***

Number contigs	25,392	16,079	32,623
Contig N50	50 Kb	129 Kb	102 Kb
Max Contig Length	862 Kb	734 Kb	930 Kb
Number of Scaffolds	18,835	10,607	23,540
Scaffold N50	570 Kb	1.2 Mb	876 Kb
Max Scaffold Length	3.6 Mb	11 Mb	6.4 Mb
Number of Scaffolds > 10 Kb	1,596	702	1,747
% of Genome in Scaffolds > 50 Kb	80.1%	90.1%	83.1%
Total Contig Assembly Length	297 Mb	294 Mb	456 Mb
Total Scaffold Assembly Length	329 Mb	306 Mb	460 Mb
% Gap	9.5%	3.9%	3.8%

Tab. 4 BUSCO (Benchmarking Using Single Copy Orthologs) metrics for stored product assemblies completed to date.

	L. serricorne	***T. variabile***	***T.*** *granarium*	***T. confusum***	***D. maculatus***
Complete/Single Copy	94.3%	98.3%	94.9%	96.9%	96.6%
Duplicate/Single Copy	0.8%	0.7%	0.9%	0.8%	0.5%
Missing	2.8%	0.5%	0.5%	0.5%	2.0%
Fragmented	2.9%	0.5%	1.9%	1.8%	2.7%

Discussion

Overall, the 10X Chromium libraries alone produced high quality assemblies that recovered significant percentages of the predicted gene space as the recovery of BUSCOs ranged from 92 to 98%. Assembly qualities were also relatively consistent regardless of genome size or taxonomic group and in all cases, over 80% of the assembly was present in less than 1,000 scaffolds, suggesting that these libraries can be a good first approach for assembling high quality draft genomes from insect species from many underrepresented taxonomic groups. In addition to the assemblies presented here, 10X Chromium libraries have been also used to produce high quality assemblies of aphid, butterfly, and dipteran genomes, providing further support for the use of this technique (Talla et al., 2017). Using linkage maps or HiC analyses, higher order assemblies will be obtained and will be almost to chromosome scale. Assemblies of this caliber can lead to the identification of syntenic orthologs across species. The identification of syntenic orthologs is important because orthologous genes present in the same chromosomal locations across species often have conserved functions, which can greatly facilitate functional annotation for non-model species (Zheng et al., 2004). In addition, the identification of syntenic orthologs may also expedite the identification of genetic targets for pest control (Futahashi, et al., 2011). For example, if knocking down a syntenic ortholog in one species reduces fitness or causes lethality, knocking down the same gene in other species that share synteny will likely also cause similar phenotypes. Although long-read sequencing approaches, such as PacBio, can generate assemblies of similar contiguity, the cost of the 10X libraries and the accompanying seuqencing is substantially less, potentially facilitating larger-scale comparative genomics studies that can be used to address broader evolutionary questions. In addition, because inbreeding is not necessary, 10X libraries can be generated much more rapidly relative to long-read sequencing approaches, which may greatly expedite genome assemblies for emerging or invasive pests.

Acknowledgements

We thank Dr. Jim Campbell from USDA-ARS-SPIERU and Dr. Scott Myers from USDA-APHIS Buzzards Bay, MA for assistance obtaining specemins for genome sequencing, Kris Hartzer, Kathy Leonard, and Rich Hammel from USDA-ARS-SPIERU for assistance with the linkage maps, and Valerie Nguyen and Rachel Wilkins for assistance with DNA and RNA preparations.

References

Angelini, D.R., Kikuchi, M. and Jockusch, E.L., 2009: Genetic patterning in the adult capitate antenna of the beetle Tribolium castaneum. Developmental Biology 327, 240-251.

Belton, J.M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y. and Dekker, J., 2012: Hi-C: A comprehensive technique to capture the conformation of genomes. Methods 58, 268-276.

Burton, J.N., Adey, A., Patwardhan, R.P., Qiu, R., Kitzman, J.O. and Shendure, J., 2013: Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. Nature Biotechnology 31, 1119.

Chesters, D., Zheng, W.M. and Zhu, C.D., 2015: A DNA barcoding system integrating multigene sequence data. Method in Ecology and Evolution 6, 930-937.

Dönitz, J., Schmitt-Engel, C., Grossmann, D., Gerischer, L., Tech, M., Schoppmeier, M., Klingler, M. and Bucher, G., 2014: iBeetle-base: A database for RNAi phenotypes in the red flour beetle Tribolium castaneum. Nucleic Acids Research 43, D720-725.

Futahashi, R., Tanaka, K., Matsuura, Y., Tanahashi, M., Kikuchi, Y. and Fukatsu, T., 2011: Laccase2 is required for cuticle pigmentation in stinkbugs. Insect Biochemistry and Molecular Biology 41, 191-196.

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G., 2013: QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29, 1072-1075.

Li, H. and Durbin, R., 2009: Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics 25, 1754-1760.

McKenna, D.D., Scully, E.D., Pauchet, Y., Hoover, K., Kirsch, R., Geib, S.M., et al., 2016: Genome of the Asian longhorned beetle (Anoplophora glabripennis) reveals key functional and evolutionary innovations at the beetle-plant interface. Genome Biology 17, 227.

Rastas, P., Calboli, F.C., Guo, B., Shikano, T., and Merila, J., 2015: Construction of ultradense linkage maps with Lep-MAP2: stickleback F 2 recombinant crosses as an example. Genome Biology and Evolution 8, 78-93.

Schlipalius, D.I., Valmas, N., Tuck, A.G., Jagadeesan, R., Ma, L., Kaur, R., Goldinger, A., Anderson, C., Kuang, J., Zuryn, S. and Mau, Y.S., 2012: A core metabolic enzyme mediates resistance to phosphine gas. Science 338, 807-810. .

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V. and Zdobnov, E.M., 2015: BUSCO: On the proper publication of stored product protection research – conference proceedings are better than you thought. BUSCO: Assessing genome assembly and annotation completedness with single-copy orthologs. Bioinformatics 31, 3210-3212.

Talla, V., Suh, A., Kalsoom, F., Dincă, V., Vila, R., Friberg, M., Wiklund, C. and Backström, N., 2017: Rapid increase in genome size as a consequence of transposable element hyperactivity in wood-white (Leptidea) butterflies. Genome Biology and Evolution 9, 2491-2505.

Wieslander, L., 1979: A simple method to recover intact high molecular weight RNA and DNA after electrophoretic seprate in low gelling temperature agarose gels. Analytical Biochemistry The control, 305-309.

Zheng, X.H., Lu, F., Wang, Z.Y., Zhong, F., Hoover, J. and Mural, R., 2004: Using shared genomic synteny and shared protein functions to enhance the identification of orthologous gene pairs. Bioinformatics 21, 703-710.

DNA barcode of stored-product Pests based on Mitochondrial Cytochrome Oxidase I Gene

Yi Wu [1]*, Zhihong Li [2], Fujun Li[1], Václav Stejskal[3], Dan Zheng[1], Xin Chen[1], Yang Cao[1]

[1]Academy of State Administration of Grain, No. 11 Baiwanzhuang Street, Beijing 100037, China.
[2]Department of Entomology, College of Agronomy and Biotechnology, China Agricultural University, Yuanmingyuan West Rd. 2,Haidian District, Beijing 100193, China.
[3]Department of Pest Control of Stored Products and Food Safety, Crop Research Institute, Drnovská 507, Prague, Czech Republic.
*Corresponding author: wuyi@chinagrain.org
DOI 10.5073/jka.2018.463.028

Abstract

The stored-product pests are economically important that can be spread through grain trade. Most stored-product pests, including eggs, nymphs, and adults, are very small and difficult to identify morphologically. Also the classification and identification of them have always been hindered by the overwhelming number of species, widely distribution. Here, we collected 43 stored-product pests from 46 geographical locations in China and other countries. The mtDNA COI gene sequences were sequenced. Software MEGA 5 was used to analyze the sequence comploition and genetic distances. Three molecular phylogenetic trees of Platypodidae were recomstructed using PAUP4.0 according to distance/ the neighbour-joining (NJ) and maximum parsimony (MP). The molecular results were compared with the morphological taxonomy. The interspecific genetic distance of the stored-product pests was significantly higher than the intraspecific genetic distance according to the barcoding gap analysis. This work provides a practical approach for the precise and rapid diagnosis of stored-product pests.

Keywords: DNA Barconding, stored-product pests, mtDNA COI gene, phylogeny