

## Genetic structure of *Tribolium castaneum* populations in mills

Semeao, A.A.<sup>\*1</sup>, Campbell, J.F.#<sup>2</sup>, Beeman, R.W.<sup>2</sup>, Whitworth, R.J.<sup>1</sup>, Sloderbeck, P.E.<sup>3</sup>, Lorenzen, M.D.<sup>4</sup>

<sup>1</sup> Department of Entomology, Kansas State University, Manhattan KS 66506. Email: aasemeao@ksu.edu

<sup>2</sup> USDA ARS Center for Grain and Animal Health Research, 1515 College Ave, Manhattan KS 66502.

<sup>3</sup> KSU-Southwest Research and Extension Center, 4500 E. Mary St., Garden City KS 67846.

<sup>4</sup> Department of Entomology, North Carolina State University, Raleigh NC 27695

\* Corresponding author

# Presenting author

DOI: 10.5073/jka.2010.425.312

### Abstract

We investigated the genetic diversity and differentiation among nine populations of *Tribolium castaneum* using eight polymorphic loci, including microsatellites and other insertion-deletion polymorphisms (=“indels”). Samples were collected in food processing/storage facilities located in Kansas, Nebraska, California, Louisiana, Florida and Puerto Rico. Standard population genetic analysis was applied, and an assignment test was used to assign individuals to their genetic population. All loci were polymorphic across populations, with the number of alleles per locus-population combination varying from three to fourteen. Among 72 locus-by-population combinations, 31 deviated significantly from Hardy-Weinberg equilibrium, which was associated with a deficiency in heterozygosity. *Tribolium castaneum* populations show some level of genetic structuring. Genetic differentiation between populations, using  $F_{ST}$  estimates, was significant, with  $F_{ST}$  varying from 0.018 to 0.149. AMOVA indicated that 8.32% of the variation in allele frequency resulted from comparisons among populations. Genetic distance was not significantly correlated with geographic distance. Correct assignment to the genetic population was possible in only 56% of all individuals. Together, these results revealed that geographically distinct populations of *T. castaneum* had low to moderate levels of genetic differentiation that was not correlated with geographic distance, and the genotypic profile of the individuals did not provide enough information for fingerprinting them with their source population.

Keywords: *Tribolium castaneum*, Population genetics, Genetic structure,  $F_{ST}$ , Genetic fingerprinting

### 1. Introduction

The red flour beetle, *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae) is one of the major pests of wheat and rice mills and has a worldwide distribution. Its distribution and long association with stored food suggests that passive movement through human activities is an important determinant of population structure in both large and small spatial scales. *Tribolium castaneum* is capable of flight and can be captured far from storage locations, but appears to disperse by flight less readily than some other stored-product pests. Monitoring studies in flour mills suggest that *T. castaneum* populations are relatively self contained within individual flour mills and that these populations go through frequent genetic bottlenecks due to regular structural fumigations (Campbell and Arbogast, 2004). This suggests that *T. castaneum* populations within mills may be genetically differentiated populations. Genetic analysis of *T. castaneum* populations from mills would enable us to develop a better understanding of population structure of this important pest, and could potentially be useful in developing tools that would facilitate the identification of sources of infestation in food distribution channels.

The potential for more accurately differentiating populations by using genetic approaches and assigning individuals to their source population exists due to the recent advances in *T. castaneum* molecular studies. The recent genome sequence (*Tribolium* Genome Sequencing Consortium, 2008) makes possible the development of more appropriate molecular markers (Demuth et al., 2007) by providing an extensive collection of markers for screening, from which selection can be made based on characteristics such as polymorphism, codominance and abundance throughout the genome.

Here, we summarize part of a study investigating the genetic structure of nine *T. castaneum* populations collected from wheat and rice mills throughout the USA using microsatellites and other insertion-deletion polymorphisms (i.e., indels). Specifically, we wanted to determine if there was genetic structure

to these populations, if levels of differentiation were associated with geographic distance, and if it was possible to accurately assign individuals to their source population.

## 2. Materials and methods

### 2.1. Sample collection

*Tribolium castaneum* was collected from rice and wheat mills across the United States and the territory of Puerto Rico (Table 1). Beetles were collected using Storgard Dome traps baited with *Tribolium* spp. pheromone lure and food oil (Trécé, Adair, OK, USA). A sample from each location ( $\geq 30$  beetles) was obtained from a collection of individuals captured in multiple traps distributed throughout each facility during the same monitoring period ( $\sim 2$  wks) and should represent an unbiased proportion of the population present at the period of collection. After being collected, beetles were kept in 1.5 mL centrifuge tubes with 75% ethanol, and frozen at  $-80^{\circ}\text{C}$  until the DNA could be extracted.

**Table 1** Collection details and summary statistics for the nine populations of *Tribolium castaneum*.

Population sample ID	Location	Sampling period	Sample size	H <sub>exp</sub> *
SaCA1	Sacramento, CA	July, 2007	32	0.64 (0.15)
SaCA2	Sacramento, CA	July, 2007	31	0.67 (0.10)
FrNE	Fremont, NE	July, 2007	32	0.67 (0.12)
OmNE	Omaha, NE	July-Aug, 2007	31	0.66 (0.15)
MhKS	Manhattan, KS	Aug, 2007	30	0.60 (0.20)
HdKS	Hudson, KS	Feb-Mar, 2005	36	0.61 (0.19)
LcLA	Lake Charles, LA	Aug, 2007	32	0.68 (0.11)
TpFL	Tampa, FL	July, 2007	32	0.68 (0.14)
GnPR	Guaynabo, PR	Aug, 2007	32	0.73 (0.12)

\*H<sub>exp</sub>, expected heterozygosity. Values represent mean and standard deviation across all loci.

### 2.2. DNA extraction and fragment analysis

A modified protocol for DNA extraction and PCR amplification was developed to deal with issues associated with using beetles collected from pheromone traps. Genomic DNA from individual beetles was extracted from entire specimen homogenates using Wizard<sup>®</sup> Genomic DNA Purification Kit (Promega, Madison, WI, USA). Genotypes were determined based on eight unique polymorphic loci including six microsatellites (MS1 through MS6) and two other insertion-deletion polymorphisms (ID1 and ID2). These molecular markers, which were either a result of our own screening of the *T. castaneum* genome or based on the literature (Demuth et al., 2007), were unique, distributed across multiple linkage groups, and confirmed to be polymorphic. For high-throughput genotyping, fluorescent labelled PCR fragments were produced using a M13 oligonucleotide adaptor sequence attached to the 5' end of the forward primers that allowed the incorporation of the fluorescent dye to the fragments. Allele sizes at each locus were scored using an ABI 3730 DNA analyzer (Applied Biosystems, Foster City, CA, USA) and GeneMarker version 1.85 software (SoftGenetics, State College, PA, USA).

### 2.3. Statistical analysis

For each population-by-locus combination, the expected and observed heterozygosities were calculated using GDA v. 1.1 (Lewis and Zaykin, 2001). Linkage disequilibrium and deviation from Hardy-Weinberg equilibrium were tested using GENEPOP v. 4.0 (Raymond and Rousset, 1995). For multiple comparisons, sequential Bonferroni was applied to determine the significance level (Rice, 1989). The genetic structure of the populations of *T. castaneum* was determined on the basis of analysis of molecular variance (AMOVA) using Arlequin v. 3.11 (Schneider et al., 2000) with two levels of hierarchy, among and within populations. Since the presence of null alleles and the failure to account for their presence can underestimate the within-population genetic diversity and overestimate differentiation among populations (Avisé and Dakin, 2004), we used the software Freena (Chapuis and Estoup, 2007) for the estimation of  $F_{ST}$  values. Isolation by distance was investigated with Mantel test using semi-matrices of genetic distance ( $F_{ST}/1-F_{ST}$ ) and of geographic distance (ln (km)) as implemented in GENEPOP. Finally, we used an assignment approach to verify the likelihood of correctly assigning individuals to their population of origin. The assignment test was carried out using the software GENECLASS 2 (Piry et al., 2004).

### 3. Results

All loci included in this study segregate independently, since no significant linkage disequilibrium was observed in any of the pair-wise comparisons. Parameters used to describe the genetic diversity of *T. castaneum* populations showed a wide range of values for most loci and populations. There was a great level of individual genetic variability within populations for most loci. Observed heterozygosity ranged from 0.063 to 0.839 and tended to be, in many cases, lower than the expected. These values were within the range found for other populations of *T. castaneum* (Demuth et al., 2007; Drury et al., 2009), and other species of beetles (Brouat et al., 2003; Schrey et al., 2008). Expected heterozygosities varied among populations, with values ranging from 0.60 to 0.73 (Table 1). Many locus-by-population combinations (31 of 72) showed significant deviation from Hardy-Weinberg equilibrium after Bonferroni correction ( $P < 0.01$ ), caused by a deficit in heterozygotes.

All eight loci were polymorphic in all populations, with the number of alleles per locus within populations varying from three to fourteen. Allele frequencies were highly variable among populations regardless of allele size or locus with frequencies ranging from 0.00 to 0.867. Most alleles were found to be common to more than one population, but when considering all loci at least one private (i.e., only occurring in one population) allele was present in each population. Private allele frequencies within populations ranged from 0.0156 to 0.391. The population from Puerto Rico, GnPR, had the greatest number of private alleles (total of eight). Null alleles (i.e., undetectable alleles due to factors such as mutations at the primer sites) were estimated to be present in all loci tested (mean frequency in each population ranged from 0.08 to 0.16).

Analysis of molecular variance (AMOVA) showed a significant level of genetic differentiation ( $P < 0.01$ ) among the populations (among populations: d.f. = 8, sum of squares = 88.36, variance component = 0.14; within populations: d.f. = 571, sum of squares = 921.26, variance component = 1.61): among populations accounted for 8.32% of the total variance. After correction for the presence of null alleles, corrected  $F_{ST}$  values ranged from 0.018 to 0.149 (Table 2, upper diagonal), with the Global pair-wise  $F_{ST}$  for all loci and population pairs being 0.082. To evaluate how much of the variation among populations could be explained by the distance between mills, the correlation between genetic and geographic distance was evaluated. There was no significant correlation between geographic distance (Table 2, lower diagonal) and genetic distance ( $F_{ST}/1-F_{ST}$ ) using Mantel test ( $P = 0.61$ ). This resulted in a lack of isolation by distance which means that the increase in geographic distance did not result in an increase of genetic distance in most cases. For example, the  $F_{ST}$  value for the pair-wise comparisons between FrNE and MhKS (distance = 247 km), LcLa and TpFL (distance = 1072 km), and SaCA1 and FrNE (distance = 2148 km) were 0.069, 0.064, and 0.059, respectively.

**Table 2** Pair-wise  $F_{ST}$  values (upper diagonal) and geographic distances in km (lower diagonal) for populations of *T. castaneum* with a global  $F_{ST}$  of 0.082.

	SaCA1	SaCA2	FrNE	OmNE	MhKS	HdKS	LcLA	TpFL	GnPR
SaCA1		0.051	0.059	0.097	0.106	0.094	0.078	0.107	0.073
SaCA2	< 0.3		0.027	0.075	0.081	0.093	0.059	0.109	0.069
FrNE	2147.7	2147.7		0.071	0.069	0.060	0.018	0.078	0.025
OmNE	2194.7	2194.7	49.9		0.089	0.145	0.107	0.149	0.094
MhKS	2152.6	2152.6	247.0	236.4		0.149	0.091	0.118	0.096
HdKS	1990.1	1990.1	412.9	422.2	217.3		0.075	0.116	0.073
LcLA	2746.7	2746.7	1277.2	1249.6	1042.5	1007.2		0.064	0.038
TpFL	3787.1	3787.1	1967.4	1922.3	1806.5	1880.4	1071.7		0.040
GnPR	5762.5	5762.5	3852.4	3803.7	3731.9	3834.2	3030.0	1975.5	

To determine how well individual beetles could be assigned to their population of origin, we used a GENECLASS-based assignment test. Unfortunately, only 56% of the individuals were correctly assigned. When the Bayesian probability was averaged among all individuals within a population, the values ranged from  $0.429 \pm 0.06$  to  $0.596 \pm 0.06$ . These values indicate that the likelihood of correctly assigning individuals to their population of origin varies among the populations in this study.

#### 4. Discussion

Results of this study provide evidence of genetic differentiation among populations of *T. castaneum* in different mills. Even though significant among-population variance was detected, it was relatively low (8.32%) considering our expectations of low gene flow at least among some of the populations we used in this study. Conversely, the significant and high within-population variance detected by the analysis of molecular variance (AMOVA) is a characteristic of microsatellites because of their highly polymorphic nature (Carbannelle et al., 2007), which may increase the genetic diversity within a population and may reduce the power of differentiation among populations. Analysis of other systems has also detected low but significant among-population variance when using these types of markers (Paupy et al., 2004; Roos and Markow, 2006).

Drury et al. (2009) analyzing microsatellites from other populations of *T. castaneum*, including populations from other countries, found  $F_{ST}$  values that ranged from 0.0289 to 0.353 with a Global pair-wise  $F_{ST}$  of 0.18. When they selected one population from Africa, Central America, South America, and North America and compared them, the average pair-wise  $F_{ST}$  was 0.09. Although they used a larger number of different microsatellite loci, the Drury et al. (2009) study  $F_{ST}$  values were only slightly greater than in our study, when including only populations in the USA ( $F_{ST} = 0.127$ ). In both studies, most pairwise comparisons had low to moderate levels of differentiation. According to Balloux and Moulin (2001), a value lying between 0-0.05 indicates little genetic differentiation; 0.05-0.15, moderate differentiation; 0.15-0.25, great differentiation; and above 0.25 indicates very great genetic differentiation. A difference between these two studies is that all of our beetles were collected directly from natural populations, while in Drury et al. (2009) there were populations that came from laboratory colonies, some of which had been collected more than 20 years earlier.

The low genetic differentiation among populations was also indicated by the lack of relationship between geographic distance (ln (km)) and genetic distance ( $F_{ST}/1-F_{ST}$ ). Drury et al. (2009) also did not find a significant correlation between genetic distance and geographic distance, even though they had a large range of geographic distances. These findings could indicate that our markers are not appropriate for detecting population differentiation at these scales or that sufficient individuals are moving among populations even at a national and global scale that sufficient genetic differentiation cannot occur. Across these broad geographic distances this movement is likely to be human aided dispersal. Anthropogenic transport of commodities such as flour has been shown to play an important role in mixing populations of stored-product species (Ryne and Bensch, 2008). Our population locations were selected so that we would have different predicted levels of potential human movement among locations. For example, we predicted greater differentiation among populations from rice mills and wheat mills or between West Coast and Midwest mills, but this was not observed. Other factors could also play important roles in determining levels of differentiation, such as, the maternally acting selfish genes (MEDEA) (Beeman et al., 1992). Another possibility is that the likelihood of genetic mixture of these populations is increased by a combination of human aided dispersal of infested materials for long distances and active dispersal by the beetles themselves from surrounding areas into the mills.

The lack of strong differentiation among populations contributed to our inability of assigning all the individuals to their population of origin with accuracy (only 56% were correctly assigned). Although our inability to assign individuals may have been caused by the lack of unique genotypes shared by individuals of the same populations, it could also have resulted from the presence of null alleles. Null alleles can affect assignment tests by reducing the power of the tests and consequently the proportion of correctly assigned individuals, and this effect will be stronger if the total number of loci is low (Carlsson, 2008). It is still possible that alternative markers or loci can be found that could eventually be used for assigning beetles to source populations. A better understanding of population structure and gene flow can improve our pest management programs by determining the level of movement of individuals among populations and the spatial scale of movement. Moreover, it can help with the evaluation of IPM program effectiveness and resistance management by allowing the discrimination between individuals that originate from rebounding of local populations, or other founder populations.

## Acknowledgments

The authors would like to thank Richard Hammel and Teresa O'Leary for technical support and Zuzana Kucerova for reviewing an earlier version of the manuscript. This study was supported by in part by USDA CSREES RAMP grant (Agreement No. 2005-51101-02358). Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture or Kansas State University.

## References

- Avise, J.C., Dakin, E.E., 2004. Microsatellite null alleles in parentage analysis. *Heredity* 93, 504-509.
- Balloux, F., Lugon-Moulin, N., 2001. The estimation of population differentiation with microsatellite markers. *Molecular Ecology* 11, 155-165.
- Beeman, R.W., Friesen, K.S., Denell, R.E., 1992. Maternal-effect selfish genes in flour beetles. *Science* 256, 89-92.
- Brouat, C., Sennedot, F., Audiot, P., Leblois, R., Rasplus, J.-Y., 2003. Fine-scale genetic structure of two carabid species with contrasted levels of habitat specialization. *Molecular Ecology* 12, 1731-1745.
- Campbell, J.F., Arbogast, R.T., 2004. Stored-product insects in a flour mill: population dynamics and response to fumigation treatments. *Entomologia Experimentalis et Applicata* 112, 217-225.
- Carbonnelle, S., Hance, T., Migeon, A., Baret, P., Cros-Arteil, S., Navajas, M., 2007. Microsatellite markers reveal spatial genetic structure of *Tetranychus urticae* (Acari: Tetranychidae) populations along a latitudinal gradient in Europe. *Experimental & Applied Acarology* 41, 225-241.
- Carlsson, J., 2008. Effects of microsatellite null alleles on assignment testing. *Journal of Heredity* 99, 616-623.
- Chapuis, M.-P., Estoup, A., 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24, 621-631.
- Demuth, J.P., Drury, D.W., Peters, M.L., Dyken, D.V., Priest, N.K., Wade, M.J., 2007. Genome-wide survey of *Tribolium castaneum* microsatellites and description of 509 polymorphic markers. *Molecular Ecology Notes* 7, 1189-1195.
- Drury, D.W., Siniard, A.L., Wade, M.J., 2009. Genetic differentiation among wild populations of *Tribolium castaneum* estimated using microsatellite markers. *Journal of Heredity* 100, 732-741.
- Lewis, P.O., Zaykin, D., 2001. Genetic data analysis: computer program for the analysis of allelic data. (<http://lewis.eeb.uconn.edu/lewishome/software.html>).
- Paupy, C., Orsoni, A., Mousson, L., Huber, K., 2004. Comparisons of amplified fragment length polymorphism (AFLP), microsatellite, and isoenzyme markers: population genetics of *Aedes aegypti* (Diptera: Culicidae) from Phnom Penh (Cambodia). *Journal of Medical Entomology* 41, 664-671.
- Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., Baudouin, L., Estoup, A., 2004. GeneClass 2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* 95, 536-539.
- Raymond, M., Rousset, F., 1995. GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86, 248-249.
- Rice, W.R., 1989. Analyzing tables of statistical levels. *Evolution* 43, 223-225.
- Roos, C.L., Markow, T.A., 2006. Microsatellite variation among diverging populations of *Drosophila mojavensis*. *Journal of Evolutionary Biology* 19, 1691-1700.
- Ryne, C., Bensch, S., 2008. Do anthropogenic transports facilitate stored-product pest moth dispersal? A molecular approach. *Naturwissenschaften* 95, 155-159.
- Schneider, S., Roessli, D., Excoffier, L., 2000. Arlequin ver. 2000: software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.
- Schrey, N.M., Schrey, A.W., Heist, E.J., Reeve, J.D., 2008. Fine-Scale population structure of southern pine beetle (Coleoptera: Curculionidae) in Mississippi forests. *Molecular Ecology and Evolution* 37, 271-276.
- Tribolium* Genome Sequencing Consortium, 2008. The genome sequence of the model beetle and pest *Tribolium castaneum*. *Nature* 452, 949-955.