GENETIC STRUCTURE OF *TRIBOLIUM CASTANEUM* POPULATIONS IN AND AROUND A GRAIN STORAGE FACILITY

Shamjana U, Rukshana Mahamood, Tony Grace*

Department of Genomic Science, Central University of Kerala, Kasaragod, Kerala, India
Corresponding author: tonygrace@cukerala.ac.in

ABSTRACT

The presence and distribution of *Tribolium castaneum* outside of food storage facilities have the potential to be an important influence on the distribution of insects inside facilities. The distribution and dispersal distance of insects should also be considered when implementing integrated pest management (IPM) and monitoring programs. The present study aims to determine if the outdoor populations of *T. castaneum* are the source for infestation within the grain storage facility or if it is from grain that is transported into the facility for storage from elsewhere. Samples of *T. castaneum* were collected from a grain storage facility as well as various other locations in the vicinity. High-throughput genotyping of all samples were carried out by using species-specific microsatellites; the genetic differentiation, gene diversity per locus, extent of genetic inbreeding, and gene flow of the species were analyzed. F_{ST} values (Genetic differentiation) of the different populations indicate a moderate level of genetic differentiation between the populations chosen. This study reveals the level of genetic differentiation among *Tribolium castaneum* populations in the locale under consideration while providing evidence that populations do not have the same source of origin. The competence of external populations in influencing the internal populations implies the importance of this study for IPM and in understanding how to prevent re-infestation of grain storage facilities.

Keywords: *Tribolium castaneum*, Microsatellites, Genotyping

1. INTRODUCTION

Genetic differentiation is substantial, and each individual of a species owns a unique DNA sequence except monozygotic twins. The investigation about genetic variation among populations of stored grain insects may be used as discriminating criterion for estimating levels of isolation and gene flow of these populations at the regional, national or global level. Genetic differences among the populations are the result of evolutionary forces (mutation, selection, random genetic drift, and migration), ecological (e.g., mating system, social structure, dispersal) and environmental (e.g., population bottleneck, landscape fragmentation, and physical barriers). Grain processing and storage facilities such as mills, warehouses, retail buildings influence the

population genetic structure, population dynamics and dispersal behavior of stored product insects [1-2]. Stored product insects occur in food processing structures and inhabit resource patches that are separated from each other by landscape fragmentation and physical barriers within facilities and among food facilities [3].

The food processing facilities, which are spatially isolated, suggest that individual populations of stored-product insects within a facility may develop unique genetic fingerprints. The genetic isolation and genetic differentiation among the populations of stored product insects are the results of long-range dispersal by either their own behavioral mechanism or by human-aided dispersal. Moreover, the processes occurring within the food processing facilities such as frequent fumigations or heat treatment [4] and population rebound after treatment [5-7] lead to the genetic differentiation of populations. Although some studies have documented the genetic structure of stored product pests [8, 9], study to evaluate the genetic structure of stored product pests in and around the storage facility are not well understood.

Tribolium castaneum is a serious and widespread pest of stored products that cause significant damage to cereal products, flour, grain, and rice bran. The dispersion of *T. castaneum* is primarily by walking but the lack of food or lack of conspecifics triggers the *Tribolium* for flight. But it is not considered as a strong flier [10, 11]. Ecological and molecular studies have proved that beetles do leave

storage locations and disperse into surrounding areas by their own flight or walking behavior [12].Campbell and Arbogast [5] monitored the abundance of *T. castaneum* inside and outside of the mills. They observed that the profusion of beetles is less outside of the flour mill. The dispersal level of *T. castaneum* within storage facility contributes in defining the potential for invasion and colonization, the spatial extent of populations and interpreting patterns in pheromone trap capture, insect density, insect age, and food quality. Isolation by distance (IBD) emerges from the tendency of individuals to interact with the populations which are in closer geographic proximity than those farther apart [13]. Thus, isolation by distance would be one of the major forces for the genetic divergence in *T. castaneum* since the beetles have a tendency to immigrate into the storage facility nearby than the distant facility to find their mates. This causes the accumulation of genetic differences among these populations. The worldwide distribution of this insect and its association with stored food suggest that human-aided dispersal was an important determinant of its current distribution and population structure.

T. castaneum is the beetle with a completely sequenced genome and it is an important model for the study of molecular and developmental genetics [14, 15]. Unique and polymorphic molecular markers, including microsatellite markers, have been developed for *T. castaneum* [9, 16] and some assessments of the genetic structure of this population have been confirmed by these markers. Microsatellites provide a necessary source of genetic variation. The variation in microsatellite alleles in coding regions is thought to be the cause of adaptation in different environments and in a population, there may be many alleles of a single microsatellite locus [17].Our current study focuses on understanding the genetic structure of *T. castaneum* population collected from both inside and outside of the storage facility using microsatellites. The population genetic structure of *T. castaneum* revealed by microsatellites can be of great importance in monitoring the pest infestations in the flour mills and also introducing pest management strategies.

Ascertaining reasons for the continuous presence of *Tribolium castaneum* inside grain storage facilities is critical for assessing the effectiveness of management strategies and it is the primary step for improving management tactics. The state of food and agriculture [18] summarizes that cereal production needs to increase by 50% between 2000 and 2030 to meet the world's food demand. It has also been estimated that by 2050 developing countries' net imports of cereals will more than double from 135 million metric tons in 2008/09 to 300 million in 2050. This sharp increase in food demand will require maintenance of food quantity and quality during processing and storage. There is also a rising demand by the public for reduced use of pesticides. This condition highlights the importance of integrated pest management programs (IPM) in order to reduce insect infestation. Understanding the genetic structure and gene flow among *Tribolium castaneum* populations helps to discover the source of the infestation. This has important ramifications for pest control as treating the source of the infestation is important to control the pest problem and to prevent the re-infestation of the facility.

2. MATERIAL AND METHODS
2.1. Sample Collection
Samples of *Tribolium castaneum* were collected from the Food storage facility located in Nileshwar, Kerala. The samples were also collected from the 4 sites located on West, North, East and South directions of this central facility. The East population was collected from a retail store that sells grains located about 1 kilometer away from the food storage facility. North populations were collected from the flour mill located 1.5 km away from the central facility. The samples were also collected from each one of the households located in the West, and South directions of this facility. These households are located about 1 kilometer away from the food storage facility. These individuals of *T. castaneum* were captured by using the Dome traps which are baited with a *Tribolium* species pheromone lure (Storegards lures, Tre′ce′ Inc., Adair, OK, USA), and attractant oil (Storegards oil, Tre′ce′ Inc., Adair, OK, USA). The samples obtained from the storage facility were more in number as compared to the samples from other locations. Individual beetles so collected were placed in 1.5 mL centrifuge tubes with 75% ethanol and frozen at -80°C until DNA extraction.

2.2. DNA extraction and Amplification
Prior to DNA extraction, individual beetles were washed with distilled water for 2-3 times in order to remove its ethanol content. Genomic DNA was extracted from *T. castaneum* using DNeasy blood and tissue kit from QIAGEN Inc. The two microsatellite markers used in this study were MS3 and MS5 and were selected based on the polymorphism content and the efficiency of

amplification [16]. The primer sequences for two microsatellites are shown in Table 1.

Each set of primers optimized the PCR conditions by varying the annealing temperature between 45-55°C in a thermocycler (Eppendorf). After determining the optimum temperature of two primers (45°C for MS3 and 46°C for MS5 primer) amplification reactions were carried out in a 20 μl final volume containing 1X Taq buffer, 2.5mM MgCl2, 0.2mM dNTP, 0.2 μM forward primer, 0.4 μM reverse primer, and 0.4 μM F-primer M13, 1U Taq DNA polymerase, 2 μl of DNA and nuclease-free water. PCR amplification was carried out for *Tribolium* individuals from 5 populations under the following conditions by using MS3 and MS5 primers. The cycling program included for MS3 primer is 94°C, 3 minute for initial denaturation; 94°C 30 sec; 45°C 30 second; 72°C 30 second for 30 cycles; 94°C 30 sec; 53°C 30 second; 72°C 30 second; for 12 cycles; and a final extension at 72°C for 5 min. Similarly, for MS5 primer the following PCR program was applied. The cycling program included for MS5 primer is 94°C, 3 minutes for initial denaturation; 94°C 30 sec; 46°C 30 second; 72°C 30 second for 30 cycles; 94°C 30 sec; 53°C 30 second; 72°C 30 second; for 12 cycles; and a final extension at 72°C for 5 min.

Table 1: The primer sequences for two microsatellites used in this study

Sl.No	Code	Primer Sequences (5′→3′)	Dye	Motif
1	TC MS3-F	5′TATCCGAAATTTTATCTACTCAT3′	6 - FAM	AAT (7)
	TC MS3R	3′AGGACCCTTTTTACTTTTTCAG5′		
2	TC MS5-F	5′AAGTGCTGCTGTATTTTATT3′	6 - FAM	TAA(16)
	TC MS5R	3′TCAGACTCCGTATCCTTTATT5′		

2.3. Genotyping

The amplified products were then run on an agarose gel and the reactions showing the presence of the amplified products were selected for genotyping. The reaction mixture for genotyping consisted of 9 μl of master mix (formamide and Liz) and 1 μl of the PCR product. Denaturation step was carried out at 95°C for 4 min followed by snap chill for 5 min. This was then transferred into a 96 well plate. The PCR fragments that were fluorescently labeled (6- FAM) at 3 primer PCR reactions [19] were loaded for genotyping into ABI 3500 Genetic Analyzer. The results thus obtained were processed for data analysis.

2.4. Data Analysis

Samples were genotyped on an ABI 3500 Genetic Analyzer and analysis was performed using GeneMapper 4.0. Various analysis parameters were set in the software and output was generated in the form of peaks corresponding to the allele sizes of the PCR amplified fragments. After the data was collected from the machine, further analysis was performed in the GeneMapper software. Genotype peaks of samples from five locations were analyzed and allele sizes were scored as per standard protocol. The data obtained after genotyping was entered into an excel sheet which conforms to the format requirements (Two-column diploid format) for microsatellite toolkit which is the population genetics software used in this study. The genetic differentiation statistics from polymorphic microsatellite markers were evaluated by choosing the FSTAT options in the microsatellite tool kit [20].

3. RESULTS

3.1. Genetic Differentiation among Populations

Genetic differentiation was estimated from F_{ST} values among and within the populations. Pairwise F_{ST} comparisons were done and the overall F_{ST} ranging from 0.013 to 0.0904. The average genetic differentiation between the central facility and East direction was 0.01.

Table 2: Genetic differentiation among populations using F_{ST} value

Locations	FCI	East	West	South	North
FCI	0	0.0152	0.0761	0.0894	0.0904
East		0	0.0297	0.0406	0.0153
West			0	0.0193	0.013
South				0	0.0392
North					0

This shows that there was no significant genetic differentiation among the central location and East location. F_{ST} values (Genetic differentiation) between FCI and other three sampling locations (West, South, and

North) showed an average F_{ST} value of 0.07, 0.09, and 0.09 respectively (Table 2). These are values of F_{ST} that indicate a moderate level of genetic differentiation among these populations and the Central location.

3.2. Genetic Diversity per Locus and Population

For each population, the number of alleles per locus (A), allelic richness (R) and gene diversity per locus and population was calculated using FSTAT. 55 alleles were detected across two loci from five populations. Loci MS5 showed the highest number of alleles per locus (i.e., 28). Most alleles were not unique to one population and no private alleles were found in any of the population (Table 3). Allelic richness per locus and population were found between 7.746 and 12.05 in five populations (Table 4). Gene diversity per locus and population were also estimated (Table 5). The mean genetic diversity across MS3 loci in FCI and East population was equal (0.750). While genetic diversity across MS5 loci in these two populations was found different value (0.455 and 0.655).

Table 3: Number of alleles per locus

Primer	FCI	East	West	South	North	Total
MS3	5	4	7	6	5	27
MS5	5	4	5	6	8	28
Total	10	8	12	12	13	55

Table 4: Allelic richness per locus and population

Marker	FCI	East	West	South	North	Total
MS3	4.45	3.9	5.949	5.642	5	5.939
MS5	4	3.846	4.25	5.549	7.05	5.171
Total	8.45	7.746	10.199	11.191	12.05	11.11

Table 5: Gene diversity per locus and population

Locus Name	FCI	East	West	South	North
MS3	0.75	0.75	0.848	0.884	0.85
MS5	0.455	0.655	0.696	0.762	0.92

Table 6: F_{IS} per population

Marker	FCI	East	West	South	North
MS3	0.333	0.5	0.558	0.434	0.608
MS5	0.176	0.127	0.282	0.25	0.592
All	0.274	0.326	0.434	0.349	0.6

3.3. Inbreeding Coefficient F_{IS}

The extent of genetic inbreeding within subpopulation was measured by using F_{IS} (Table 6). F_{IS} can range from -1.0 (all individuals heterozygous) to +1.0 (no observed heterozygotes). F_{IS} per population across two loci was estimated at 0.274, 0.326, 0.434, 0.349, and 0.600 showing a high level of inbreeding.

3.4. Heterozygosity and Hardy Weinberg Equilibrium

Average Heterozygosity of all population across two loci was estimated as per Nei [21]. Observed Heterozygosity (Ho), Expected Heterozygosity (He) across the 5 populations varied significantly and statistically suggesting that none of these populations in the study were in Hardy Weinberg Equilibrium. The average expected heterozygosity was 0.59 in FCI populations whereas it was 0.68, 0.75, 0.80 and 0.84 in East, West, South and North populations. Similarly, the average observed heterozygosity was 0.44 in FCI populations and 0.47, 0.44, 0.54 and 0.35 in East, West, South and North populations (Table 7). Tests for linkage disequilibrium between the 2 loci did not show any linkage between the two loci MS-3 and MS-5 as the values for linkage disequilibrium were not statistically significant.

Table 7: Heterozygosity and Hardy Weinberg Equilibrium

Expected Heterozygosity					
Locus	FCI	East	West	South	North
MS3	0.733	0.725	0.817	0.858	0.803
MS5	0.45	0.648	0.683	0.747	0.883
Average	0.592	0.687	0.75	0.803	0.843
PIC values					
MS3	0.636	0.624	0.735	0.776	0.692
MS5	0.404	0.553	0.584	0.664	0.809
Observed Heterozygosity					
MS3	0.5	0.375	0.375	0.5	0.333
MS5	0.375	0.571	0.5	0.571	0.375
Average	0.438	0.473	0.438	0.536	0.354

3.5. Gene Flow among Populations

Gene flow was estimated among populations from the F_{ST} values and the overall gene flow among the 5 populations was calculated. It was found that the number of migrants into populations was found to be 6 migrants per generation. This is a very high amount of gene flow into the populations overall. However, at the level of individual population's comparisons, high gene flow was observed among grain storage facility and populations on the East, whereas lower gene flow was observed among the other populations and grain storage facility.

4. DISCUSSION

Although the activity of stored product insects outside the storage facilities has been widely examined [22-24], its genetic structure estimation is not well understood. The population's genetic structure can be measured in many ways. The method most frequently used to assess population structure is the calculation of F_{ST}, a summary statistic, first introduced by Sewall Wright [13]. For estimating the probabilities of shared alleles, we calculated the frequency of different alleles and genotypes of randomly collected individuals from a specific geographic area. Under an assumption of Hardy Weinberg equilibrium, we also measured the difference between observed and expected heterozygosity for this diploid sexual system.

F_{ST} can range from 0.0 (no genetic differentiation) to 1.0 (complete differentiation-subpopulations fixed for different alleles). Although F_{ST} has a theoretical range of 0 to 1.0, the observed maximum is usually much less than 1.0. Wright [25] suggested that the range 0.0 to 0.05 may be considered as indicating little genetic differentiation, the range 0.05 to 0.15 as moderate genetic differentiation, the range 0.15 to 0.25 as great genetic differentiation, and values of F_{ST} above 0.25 indicating as very great genetic differentiation. F_{ST} values also indirectly indicate the migration and gene flow among subpopulations

This study is focused on understanding the genetic structure and gene flow among *Tribolium castaneum* populations collected from a Grain storage facility and from 4 other locales around this facility- East, West, North, and South. The grain storage facility selected was the Food Corporation of India (FCI) Godown at Nileshwar, Kerala. Results of this study reveal the level of genetic differentiation among *Tribolium castaneum* populations. F_{ST} value of 0.01 did not reveal any evident genetic differentiation between the central facility and East populations. From this result, it can be clearly assumed that there is likely a high level of continued migration between these two subpopulations. Genetic differentiation between FCI and other three directions (West, South, and North) shows average F_{ST} values of 0.07, 0.09, and 0.09 respectively. This shows that there is a moderate level of genetic differentiation between these populations and a lower level of gene flow among subpopulations. The lack of strong differentiation among populations can lead to an inability of assigning all the individuals to their population of origin. Interpopulation genetic differentiation was measured in *T.castaneum* [9, 26] and other insect species [27, 28].

Studies of Drury *et al.*, [9] found that there was a moderate level of genetic differentiation in *T. castaneum* populations originating within the U.S. Reports of Drury *et al.*, [9] and Semeo *et al.*, [26] suggest that there was no relationship between geographic distance and genetic distance. Similarly, a low level of genetic differentiation was observed in *A. gambiae* populations that were as much as 6000 km apart [29].

Gene diversity per locus and population across two loci and five populations were estimated. Loci MS3 in FCI and East population showed high genetic diversity at 0.75. While genetic diversity across MS5 loci in these two populations was found to be 0.45 and 0.65. For MS3 loci, there was no diversity between these two populations, whereas MS5 loci show significant genetic diversity. Similarly, genetic diversity of remaining populations across two loci revealed significant genetic diversity among them. Allelic Richness per locus and population estimate shows that there are no overall significant differences in richness across the two loci. A total of 59 alleles were amplified across two loci in five populations. MS5 showed the highest number of alleles per locus of eight. Most alleles were not unique to one population and as such no private alleles were found in any of the populations in the study.

Accounting for non-random mating and divergence of subpopulation allele frequency necessitates several new versions of the fixation index. Inbreeding coefficient F_{IS} [30] compares the average observed heterozygosity of individuals in each subpopulation and the average Hardy−Weinberg expected heterozygosity for all subpopulation. It can range from -1.0 (all individuals heterozygous) to +1.0 (no observed heterozygotes). F_{IS} values in the study populations were 0.274, 0.326, 0.434, 0.349, and 0.600 in FCI, East, West, South and North populations respectively. This result illustrates that there are few heterozygotes in each subpopulation. Thus, there is more homozygosity or fixation within the subpopulations under random mating. The subpopulations on average have a deficit of heterozygotes as expected if there is consanguineous mating. This denotes to some level of inbreeding among the individuals of all subpopulations. This shows that there is continued migration of *T. castaneum* between the central facility and the surrounding environment.

The main objective of the study was to ascertain the source of the infestation in the Grain Storage Facility

(FCI Godown). Based on the genetic data, we aimed at finding if the outdoor populations on the East, West, North, and South were the source for infestation within the storage facility or if it was from Grain that is transported into the facility for storage from elsewhere (via truck or train). This has important ramifications for pest control as treating the source of the infestation is important to control the pest problem and to prevent the re-infestation of the facility. If treatment is done without accounting for the source, multiple treatments may be required to keep the problem under control. This will add to the cost of effective control and time needed for control. *T. castaneum* is resistant to almost all the insecticides used to control it. This also has ramifications for insecticide resistance management and in preventing the spread of resistance by identifying the origin of insects.

Based on the genetic differentiation estimates, migration/gene flow values, it is most likely that populations in the West, South, and North are not likely the source of infestation of the Central Facility. This is because these populations show a moderate level of genetic differentiation with populations in the Central Facility even though they are separated only by 1 kilometer. However, the population on the East shows a very low level of genetic differentiation with the populations in the Central Facility. This is also probably not the source of infestation of the Central Facility because East population is also separated by 1 kilometer like West, South and North populations. One possibility is infestation originating with insects migrating from the Central Facility to East. The East population was collected from a retail Store that sells grains, whereas the other 2 outdoor populations were from homes around the facility and one from flour mills. So another possibility is that the same grains that are brought into the FCI are also supplied to the Store on the East or grains were supplied from FCI to the Store. From this, it is clear that the source of infestation of the FCI Godown and the Grain Store on the East are from grain supplied via train to these facilities and not from local populations existing outside the facility. This has huge implications for insect pest management.

5. REFERENCES

1. Turner MG. *Annual Review of Ecological Systems,* 1989; **20(1)**:171-197.
2. Wiens JA. Metapopulation dynamics and landscape ecology. In *Metapopulation biology*: Academic Press. 1997; 43-62.
3. Campbell JF. *Insect Management for Food Storage and Processing*, 2005; 39-51.
4. Fields PM. &White NDG. *Annual Review of Entomology,* 2002; **47(1)**:331-359.
5. Campbell JF, Arbogast RT. *Entomologia Experimentaliset Applicata*, 2004; **112(3)**:217-225.
6. Campbell JF, Toews MD, Arthur FH, Arbogast, RT. *Journal of Economic Entomology,* 2010a; **103(3)**:991-1001.
7. Campbell JF, Toews M, Arthur FH, Arbogast RT. *Journal of Economic Entomology*, 2010b; **103(3)**:1002-1011.
8. Dowdy AK, McGaughey WH. *Environmental Entomology,* 1996; **25(2)**:396-400.
9. Drury DW, Siniard AL, Wade, MJ. *Journal of Heredity*, 2009; **100(6)**:732-741.
10. Sinclair ER, Haddrell RL. *Journal of Australian Entomological Society,* 1985; **24(1)**:9-15.
11. Daglish GJ, Ridley AW, Walter GH. *Julius-Kühn-Archiv.*, 2010; **425**:104-109.
12. Ridley AW, Hereward JP, Daglish GJ, Raghu S, *et al. Molecular Ecology*, 2011; **20(8):**635-646.
13. Wright S. *Genetics,* 1943; **28(2)**:114-138.
14. Lorenzen MD. *Genetics*, 2005; **170(2)**:741-747.
15. *Tribolium* Genome Sequencing Consortium. *Nature.* 2008; **452(7190)**:949-955.
16. Demuth JP, Drury DW, Peters ML, Dyken D, *et al. Molecular Ecology Notes,* 2007; **7(6)**:1189-1195.
17. Shamjana U, Thashi Bharadwaj, Tony Grace. *International Journal of Advanced Biological Research,* 2015; **5(2)**:84-96.
18. FAO, How to feed the world in 2050. Report from the High-level expert forum. http://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World in_2050.pdf. Rome, October, 12-13. Accessed on 12/20/2010.
19. Schuelke M. *Nature Biotechnology,* 2000*;* **18(2)**:233-234.
20. Goudet J, Hierfstat. *Molecular Ecology Notes*, 2005; **5(1)**:184-186.
21. Nei M. *The American Naturalist*, 1972; **106(949)**: 283-292.
22. Throne JE, Cline LD. *Journal of Agricultural Entomology*, 1989; **6**:183-192.
23. Dowdy AK, McGaughey WH. *Journal of Economic Entomology*, 1994; **87(5)**:1351-1358.
24. Doud CW, Phillips TW. *Journal of Economic Entomology*, 2000; **93(6)**:1842-1847.

25. Wright, S. *Evolution and the Genetics of Populations,Variability Within and Among Natural Populations*, Chicago:University of Chicago Press. 1978.

26. Semeao AA, Campbell JF, Beeman RW, Lorenzen MD *et al*. *Environmental entomology*, 2012; **41(1):**188-199.

27. Paupy C, Orsoni A, Mousson L. *Journal of Medical Entomology,* 2004; **41(4)**:664-671.

28. Roos CL, Markow TA. *Journal of Evolutionary Biology,* 2006*;* **19(5)**:1691-1700.

29. Lehmann T, Hawley WA, Kamau L, Fontenille D *et al*. *Heredity,* 1996*;* **77(2)**:192-208.

30. Slatkin M. *Genetics Research*, 1991; **58(2)**:167-175.