Impact of Host Plants on Genetic Variation in the *Bactrocera tau* (Diptera: Tephritidae) Based on Molecular Markers[1]

Wei Shi, Teng-yu Yang[2], Hui Ye[4], and Jun Cao[3]

School of Agriculture Research, Yunnan University, 650091 Kunming, China

J. Entomol. Sci. 52(4): 411–426 (October 2017)

Abstract The adaptation to novel host plants is an important factor that facilitates spread of tephritid fruit flies and formation of new biotypes. *Bactrocera tau* (Walker) (Diptera: Tephritidae) is a significant pest and member of the Tephritidae family that has expanded its normal host range from Cucurbitaceae to many other plant families. The objective of this study was to monitor the impact of novel host plants on genetic variation in *B. tau*. In this study, under laboratory conditions, we examined genetic variation based on microsatellite molecular markers of *B. tau* when development occurred on one novel host (banana) and two traditional hosts (pumpkin and cucumber) for 15 continuous generations. Analysis of molecular variance showed that the genetic difference (10.39%) among populations feeding on the various hosts was significant, but not among populations from different generations. It appears that host food, more than population generations, played a critical role in genetic variation of *B. tau*. The effect of hosts on genetic variation of *B. tau* was greater on the novel host compared with the traditional Cucurbitaceae hosts. The significant genetic variation of the banana-feeding populations was reflected in difficulties of balancing in frequency of gene and genotype, reduced genetic diversity, and more divergent genetic difference.

Key Words *Bactrocera tau,* genetic variation, microsatellite markers, novel host species

The adaptation to novel host plants is an important factor that facilitates the geographic expansion of tephritid flies (Diptera: Tephritidae) (Ju et al. 2012, Wan 2009). Many studies have shown that geographic distribution of tephritids is closely related to host plant diversity (Malacrida et al. 2007, Nardi et al. 2005). As fruit flies expand their geographic range, they will encounter novel hosts with selection pressures possibly affecting their growth, development, reproduction, and, ultimately, their genetic composition (Faucci et al. 2007, Schwarz et al. 2005).

There are several examples where tephritid flies have adapted to new hosts with changes in their genetic composition. In the case of the apple maggot, *Rhagoletis pomonella* (Walsh), populations adapted to its traditional host apple (*Malus*) and those adapted to the novel host hawthorn (*Crateagus*) formed two sympatric host races (Han and McPheron 1994). Similarly, U.S. populations of *R. pomonella* that adapted to honeysuckle (Caprifoliaceae: *Lonicera*), a different plant family than the

[1]Received 09 February 2017; accepted for publication 09 May 2017.
[2]Life Science College, Yunnan University, 650091 Kunming, China.
[3]Lab of Ecology and Evolutionary Biology, Yunnan University, 650091 Kunming, China.
[4]Corresponding author (email: yehuiyd@126.com and yehui@ynu.edu.cn).

more typical hosts (e.g., Rosaceae), are considered a new hybrid species and often termed the "Lonicera fly" (Martel et al. 2003).

Bactrocera tau (Walker) (Diptera: Tephritidae), is a major economic pest that is listed as a quarantine pest in many countries (Ooi and Wee 2016). This fly was first reported in 1894 in Fujian, a southeast coastal province of China (Walker 1849). The current geographic range of *B. tau* includes several countries in tropical and subtropical Asia and the South Pacific (Singh et al. 2010), such as China, Thailand, India, Cambodia, Indonesia, and Pakistan (Hasyim et al. 2008, Huang et al. 2005, Huque 2006, Ohno et al. 2008). The geographic range of *B. tau* is still expanding in many regions. For example, in China, *B. tau* has recently spread to two northern provinces (i.e., Shanxi and Shaanxi) (Wan et al. 2010). Similarly, in Japan, *B. tau* has recently been reported for the first time in Okinawa and Ishigaki Islands (Ohno et al. 2008).

Bactrocera tau is polyphagous and mainly infests plants in the Cucurbitaceae family (Christenson and Foote 1960). However, the range of host plants utilized by *B. tau* has expanded dramatically over the last century (Allwood et al. 1999) and now includes more than 80 plant species (Huang et al. 2005), including species in the plant families Leguminosae (e.g., *Phaseolus vulgaris* L.), Moraceae (e.g., *Ficus racemosa* L.) (Sumrandee et al. 2011), Myrtaceae (e.g., *Psidium guajava* L.) (Hasyim et al. 2008), and Rutaceae (e.g., *Citrus*) (Zhang and Chen 2012).

Development and reproduction rates, as well as genetic composition of *B. tau* populations, vary considerably among geographic regions (Sumrandee et al. 2011, Dujardin and Kitthawee 2013, Jamnongluk et al. 2003a, 2003b, Saelee et al. 2006, Thanaphum and Thaenkham 2003). Such variation has been attributed to environmental factors including air temperature, and host species (Dujardin and Kitthawee 2013). Among these factors, host species is considered to be one of the critical factors that determine the geographical distribution of *B. tau* (Sumrandee et al. 2011). In order to enhance understanding of tephritid distribution and expansion mechanisms, it is necessary to determine the direct impact of host plants on these flies.

The aim of the present study was to examine the effect of host plants on genetic variation in *B. tau*. Our experiment was conducted under laboratory conditions where factors such as natural enemies, climatic conditions, and human activities could be excluded or dramatically reduced. The flies were fed with both traditional and novel host foods for 15 consecutive generations, and microsatellite markers were used to evaluate the genetic variation among different fly populations. The main objective was to elucidate any genetic variation in *B. tau* during development on different host plants.

Materials and Methods

Summer squash (*Cucurbita pepo* var. *fastigata* L.) infested with *B. tau* in the field were transported to our entomology laboratory at Yunnan University, China. These infested summer squash were placed in insect-rearing cages (60 $\times$ 40 $\times$ 45 cm) until *B. tau* adults emerged, and all emerging flies were subsequently maintained on summer squash. After 2 weeks of rearing, a large number of emerging adult flies were used for this study.

The test flies were divided into four groups, and each group was fed one of the following: cucumber (*Cucumis sativus* L.), pumpkin (*Cucurbita pepo var. pepo* L.), orange (*Citrus*), or banana (*Musa paradisiaca* Colla). Sixty adult flies including 30 males and 30 females were used in each group. The flies were reared continuously on each of the four host plants. However, the flies fed on orange died within three generations; therefore, tests with that host plant were discontinued. Flies on the other three host plants were reared for 15 continuous generations. After 10 d of adult emergence in each generation, we kept 30 male adults and 30 female adults to initiate the next generation. Rearing was conducted in the insect-rearing cages described above, which were placed in an environmental chamber maintained at 27 $\pm$ 1°C, 70 $\pm$ 5**%** relative humidity, and a photoperiod of L12:D12 h (lights 0700–1900). Each treatment was replicated three times. Survival rates were calculated for each host group based on the 5th, 10th and 15th generation as described by Vargas et al. (1984). Mean survival rates of the three host plant treatment groups were compared using one-way analysis of variance (Snedecor and Cochran 1989). Means that were significantly different ($P < 0.05$) were separated using *á posteriori* Tukey's honestly significant difference test (Zar 1999).

Twenty *B. tau* adults were selected from each of the 5th, 10th, and 15th generations for each host plant group and used for molecular analysis. These adults were preserved in 95**%** ethanol and stored at 4°C. The nine populations were named as follows: C5 (the 5th generation feeding on cucumber), B5 (the 5th generation feeding on banana), P5 (the 5th generation feeding on pumpkin), and similarly C10, B10, and P10, and C15, B15, and P15.

In total, 180 adults were used in the molecular analyses. DNA samples of 20 adults per generation per population were extracted using a DNeasy Blood and Tissue Kit (Qiagen, Boston, MA). Seven pairs of highly polymorphic microsatellite primers were used for PCR amplification (Table 1). The amplification schedule was 40 cycles of 15 s each at 94°C, 15 s at 55°C, and 30 s at 72°C, with an initial denaturation step of 10 s at 95°C and a final extension step of 30 min at 72°C, followed by storage at 4°C. PCR products were electrophoresed using an automated ABI PRISM 310 Genetic Analyzer, and allele calling was performed using GeneMapper. An individual allele was declared null for a given locus only after at least two amplification failures.

FreeNA software was used to evaluate the frequency of null alleles for each microsatellite loci (Chapuis and Estoup 2007). GENEPOP4.5 (Rousset 2008) was used to estimate the linkage disequilibrium between pairs of seven microsatellite loci in each *B. tau* population and deviation from Hardy–Weinberg equilibrium based on Fisher's method. The four genetic diversity indices of each *B. tau* population, including number of alleles (N_A), number of private alleles (N_P), frequency of private alleles (A_P), observed heterozygosity (H_O), expected heterozygosity (H_E) and gene diversity (H_S), were calculated using FSTAT2.9.3.2 (Goudet 2001).

In order to infer the genetic structure of *B. tau* populations feeding on different hosts, the Bayesian clustering method implemented in STRUCTURE2.3.4 (Pritchard et al. 2000) was used to first determine whether the nine *B. tau* populations could be subdivided into different groups. STRUCTURE software runs a model in which K is assigned to various clusters (K is known), and each K is marked with a collection of allele frequencies at each locus. Individuals from

Table 1. The seven pairs of microsatellite primers for the nine populations of *B. tau* feeding on the three plant hosts.

Loci	Primer Sequence	Length (bp)	Primer Source
BcuB4.3	F-CTCGCCGTAATAGCCTGT	149–168	Wu et al. (2009)
	R-GGGTCGTAAATTCCGTTG		
BcuB5.2	F-CCAAAACCAATCACGACG	125–135	Wu et al. (2009)
	R-AAACATACGCACGCAACT		
BcuF3.2	F-GGGAGAGTCCAGTGAAGGTG	107–115	Wu et al. (2009)
	R- CTTCCACGCAACAGCAAAC		
BcuF3.4	F-AACTCTGTCGAGTCGGCAAT	68–126	Wu et al. (2009)
	R-CAATCAACGCAAAAGTCCAA		
Bi1	F-CTCTTGACACTGGCCTCGTT	129–159	Khamis (2009)
	R-GTATGGCCGGAGACATCAGT		
Bi2	F-ACGAGAGCACTCACTCAACCT	107–118	Khamis (2009)
	R-GCAGCAATCATAACAAGTAGCA		
Bi7	F-CTCGCTCTTCATTCAATCCA	129–159	Khamis (2009)
	R-CGACACGTTAAGTGGCAAAA		

* bp, base pair.

different samples are assigned to one assumed cluster, or together to two or more clusters if they shared admixed genotypes (Aketarawong et al. 2007). To select the most likely number of *K* based on our samples, we arranged our data on various *K*s ranging from 1 to 10 and performed ten dependent runs for each *K* (Shi et al. 2012). Then, the estimated log probability of the data for the different *K*s was compared. The proportion of individuals assigned into different *K* clusters was displayed as Q-matrices. When running STRUCTURE2.3.4, we set a model of admixture with 100,000 burn-in steps followed by 100,000 MCMC simulation steps. Structure groups were visualized using the Distruct software (Rosenberg 2004).

After obtaining the optimal STRUCTURE groups by using the analysis described above, analysis of molecular variance (AMOVA) was conducted to test molecular variation between these groups by using Arlequin 3.5 (Excoffier and Lischer 2010). The same software was also used to perform an AMOVA to detect the sources of genetic variation based on groups categorized by generations and host fruits.

A neighbor-joining (NJ) tree was constructed with PHYLIP3.69 (Felsenstein 2005) based on the distances of pairwise proportion of shared alleles between populations as the second method to estimate the genetic structure of *B. tau* populations. The NJ tree was supported by 1,000 bootstrap re-samplings of the original data over loci.

Differentiation among populations, as measured by pairwise F_{ST} value, was the third approach used to determine the genetic structure of nine *B. tau* populations

Table 2. Mean percentage survival rate of the 5th, 10th, and 15th *B. tau* generations from each of three host plant groups.

Populations*	Survival Rate, M ± SD**
P5	93.45% ± 1.52 a
P10	88.33% ± 1.77 a
P15	91.63% ± 0.86 a
C5	90.80% ± 0.65 a
C10	92.78% ± 1.42 a
C15	89.81% ± 2.91 a
B5	69.89% ± 0.87 b
B10	72.36% ± 0.32 b
B15	74.67% ± 1.54 b

* C5 (the 5th generation feeding on cucumber), B5 (the 5th generation feeding on banana), P5 (the 5th generation feeding on pumpkin), and similarly C10, B10, and P10, and C15, B15, and P15
**Means followed by the different lower case letter are significantly ($P < 0.05$) different.

using Arleqquin3.5. The same software also was used to estimate the F_{ST} value among groups defined by STRUCTURE2.3.4.

Gene flow among populations was detected by using GENECLASS 2.0 (Piry et al. 2004). This software estimated the gene flow by calculating the probability of assignment of each individual to other reference populations or assignment in the population itself based on multilocus genotypes (Piry et al. 2004). Values of probability were computed by the standard criterion described by Rannala and Mountain (1997) with 10,000 simulated individuals and a $P = 0.01$.

Results

Survival. Mean survival among the *B. tau* populations reared on the three plant hosts for 5, 10, and 15 generations differed significantly ($F = 2.78$; df = 1,540; $P = 0.032 < 0.05$), with pumpkin = cucumber > banana (Table 2). Survival rates of flies fed on banana increased with progressive generations.

Microsatellite genotype characteristics. The number of alleles per locus, based on seven pairs of microsatellite loci, ranged from 13 to 25 for the nine *B. tau* populations that developed on three plant hosts (Table 1). The mean frequency of null alleles for each of seven loci was 0.0138 for locus BcuB4.3, 0.0435 for BcuB5.2, 0.003 for BcuF3.2, 0.023 for BcuF3.4, 0.017 for both Bi1 and Bi2, and 0.014 for Bi7. Overall, the mean frequency of null alleles for each locus was always below 0.10. No linkage disequilibrium was observed for any pair of loci. The Hardy–Weinberg equilibrium test showed that the P5 population deviated from the equilibrium at loci BcuB5.2 ($P = 0.034 < 0.05$) and BcuF3.2 ($P = 0.041 < 0.05$), the C5 population deviated from equilibrium at loci BcuB4.3 ($P = 0.009 < 0.01$) and Bi7 ($P = 0.026 < 0.05$), the B5 population deviated from equilibrium at loci BcuB5.2 ($P =$

Table 3. Genetic diversity index for the nine *B. tau* host populations.

Populations	N_A	N_P	H_O	H_E	H_S
P5	5.857	2	0.567	0.721	0.679
P10	7.571	3	0.495	0.782	0.659
P15	6.429	3	0.467	0.569	0.592
Mean value	6.619	2.7	0.510	0.691	0.643
C5	5.571	2	0.552	0.720	0.780
C10	5.857	2	0.495	0.665	0.674
C15	4.286	3	0.429	0.602	0.509
Mean value	5.238	2.3	0.492	0.657	0.654
B5	4.571	1	0.586	0.679	0.601
B10	5.714	1	0.514	0.646	0.647
B15	4.241	1	0.471	0.557	0.550
Mean value	4.842	1	0.524	0.627	0.599

C5 (the 5th generation feeding on cucumber), B5 (the 5th generation feeding on banana), P5 (the 5th generation feeding on pumpkin), and similarly C10, B10, and P10, and C15, B15, and P15
N_A, number of alleles; N_P, number of private alleles; H_O, mean observed heterozygosity; H_E, mean expected heterozygosity; H_S, gene diversity.

$0.018 < 0.05$) and Bi7 ($P = 0.022 < 0.05$), and the B10 population deviated from equilibrium at loci BcuB5.2 ($P = 0.004 < 0.01$) and Bi2 ($P = 0.031 < 0.05$). Other populations fit the Hardy–Weinberg equilibrium at all loci.

Genetic diversity. We used five indices, namely N_A, N_P, H_O, H_E, and *H*s, to measure genetic diversity within 10 *B. tau* populations (Table 3). Among the nine *B. tau* populations, the mean number of N_A, N_P, and *H*s of three banana populations were slightly lower than the values of three populations feeding on cucumber and pumpkin, respectively. The value of H_E, H_O, and *H*s for the three host populations of the 15th generation (i.e., P15, C15, and B15) decreased slightly compared to the values in populations of the 5th and 10th generations that developed on the same hosts.

Population group assessment. The STRUCTURE analysis showed that *B. tau* populations that developed on the three host plants could be subdivided into four genetic clusters (*K*), as shown by the likelihood curve of STRUCTURE (Fig. 1, 10 runs for each *K*). The curve reached a plateau when $K = 4$; therefore, optimal *K* was set as 4. Each of the 180 flies was subsequently assigned to one of the four clusters with a certain probability Q (Table 4). Flies from populations B5 and B10 were mostly assigned to cluster 1 ($Q > 0.45$). Individuals from populations P5, P10, and P15 were mostly assigned to cluster 2 ($Q > 0.70$). Flies from populations C5, C10, and C15 were primarily assigned to cluster 3 ($Q > 0.6$). Almost all flies of B15 were assigned to cluster 4 ($Q = 0.95$). Therefore, we defined the four clusters as four groups: G1, G2, G3, and G4 (Fig. 2). The G1 group was composed of populations of

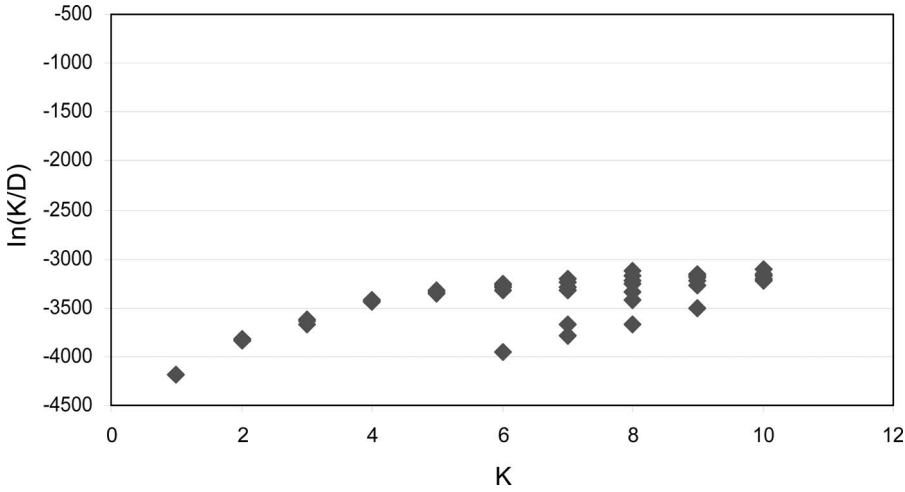

Fig. 1. Log-likelihood probability LnP(D) of the number of inferred clusters (*K*) as a function of *K* by using STRUCTURE for *K* = 1–10, with 10 independent runs averaged for each *K*.

Table 4. Average co-ancestry coefficients for the nine populations of *B. tau* assigned to four clusters.

	Cluster (*K*)			
Population	**1**	**2**	**3**	**4**
1. P5	0.088	**0.745**	0.132	0.034
2. P10	0.073	**0.864**	0.053	0.010
3. P15	0.021	**0.927**	0.025	0.027
4. C5	0.137	0.021	**0.796**	0.047
5. C10	0.148	0.024	**0.775**	0.053
6. C15	0.205	0.021	**0.639**	0.136
7. B5	**0.579**	0.168	0.022	0.231
8. B10	**0.453**	0.290	0.028	0.229
9. B15	0.014	0.021	0.017	**0.948**

C5 (the 5th generation feeding on cucumber), B5 (the 5th generation feeding on banana), P5 (the 5th generation feeding on pumpkin), and similarly C10, B10, and P10, and C15, B15, and P15.
The highest value of co-ancestry of each population in a cluster is in bold. Values higher than 0.10 are bold.

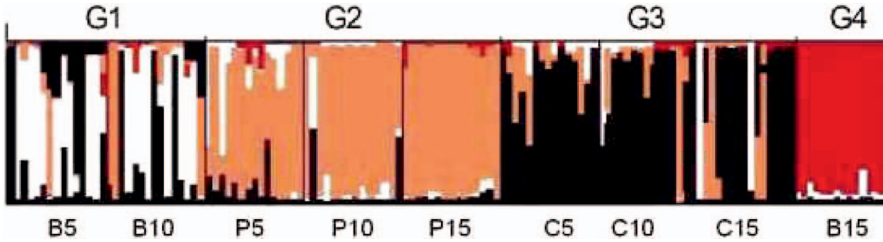

Fig. 2. Structure bar plot showing the four groups identified by the analysis (*K*=4). The groups G1, G2, G3, and G4 corresponded to the groups in Table 4.

the 5th and 10th generations that developed on banana (B5 and B10); the G2 group included populations of the 5th, 10th, and 15th generations that developed on pumpkin (P5, P10, and P15); the G3 group was composed of populations of the 5th, 10th, and 15th generations that developed on cucumber (C5, C10, and C15); and the G4 group was only one population, B15.

Molecular variation analysis. The fixation indices of the AMOVA run using the group of populations assigned by the STRUCTURE results previously described are shown in Table 5. Most of the molecular variation was found within populations (85.05%, $P < 0.01$), only 8.12% of the variation was found among groups ($P < 0.01$), and only 6.83% of the variation was found among populations within groups. This partitioning was significant ($P < 0.01$).

The nine *B. tau* populations were also grouped by generation and host (Table 6). When populations were grouped by host, most variation was detected within populations (80.85%), 8.76% was among population within groups, and only 10.39% was among groups, with all three values being significant ($P < 0.01$). When the nine populations were grouped by generation, a substantial proportion of genetic differentiation was within populations (83.64%), 13.68% of the variation was among populations within groups, and only a small proportion (2.68%) of the variation was among groups.

NJ tree construction. An unrooted NJ tree constructed with the microsatellite data (based on the pairwise proportion of shared alleles distances) also showed four monophyletic clades, which were similar to the STRUCTURE groups identified for $K = 4$ (see Fig. 3).

Pairwise F_{ST} values. Table 7 lists all microsatellite pairwise F_{ST} values of the nine *B. tau* populations, which ranged from 0.043 (C5–C10, P10–P15) to 0.256 (C15—B15). Among 36 F_{ST} values of the nine populations, we observed that six values were not significant. Slightly high values of differentiation were observed between the B15 population and other populations where nine F_{ST} values were significant and higher than 0.2. When populations were grouped according to the STRUCTURE results, pairwise F_{ST} values between the four groups varied from 0.010 to 0.202 (Table 5). All F_{ST} values among the four groups were significantly different, except for F_{ST} values among groups G2 and G3, which consisted of the two traditional Cucurbitaceae host feeding populations.

Table 5. The F_{ST} values for the four groups obtained from STRUCTURE 2.3.4 and analysis of molecular variances based on the four groups.

Populations	Group Name	Genetic Difference Among Groups				Source of Variation	Percentage Variation	Fix Index
		G1	G2	G3	G4			
B5 + B10	G1	0.000				Among groups	8.12**	0.0812
P5 + P10 + P15	G2	0.102	0.000			Among populations within groups	6.83**	0.0743
C5 + C10 + C15	G3	0.128	0.010NS	0.000		Within populations	85.05**	0.3423
B15	G4	0.202	0.195	0.174	0.000			

C5 (the 5th generation feeding on cucumber), B5 (the 5th generation feeding on banana), P5 (the 5th generation feeding on pumpkin), and similarly C10, B10, and P10, and C15, B15, and P15.
NS = no statistical significance; **, $P < 0.01$.

Table 6. The analysis of molecular variance results of the 10 *B. tau* populations that developed on one of three host plant fruits.

Source of Variation	Percentage of Variation	Fixation Index
Groups based on different host fruits		
Among groups	10.39	0.010**
Among populations within groups	8.76	0.129**
Among populations	80.85	0.227**
Groups based on different generations		
Among groups	2.68	0.044
Among populations within groups	13.68	0.106**
Among populations	83.64	0.217**

C5 (the 5th generation feeding on cucumber), B5 (the 5th generation feeding on banana), P5 (the 5th generation feeding on pumpkin), and similarly C10, B10, and P10, and C15, B15, and P15.
**, $P < 0.01$; *, $P < 0.05$.

Gene flow estimates. GENECLASS 2.0 was used to estimate the gene flow among nine *B. tau* populations. Gene flow was calculated by probability of individual assignment to different populations (Table 8). Most of the values of gene flow were low and not higher than 0.047. Only five values of gene flow were higher than 0.05, and four values of gene flow were higher than 0.1. The diagonal values of the assignment matrix indicated the average probability with which individuals were assigned to population itself. These probability values ranged from 0.544 (C15) to 0.625 (P5).

Discussion

Survival rate can be used as a direct measure to evaluate the effect of the various hosts on the tephritid flies (Liu et al. 2014). The present study showed that survival rates were higher for the *B. tau* populations fed on pumpkin and cucumber than those fed on banana, with no differences between the two Cucurbitaceae hosts (Christenson and Foote 1960) (Table 2). Survival rates of flies fed on the two non-Cucurbitaceae hosts (banana and orange) were low and survived only through three generations. Our study further showed that *B. tau* had higher fitness when feeding on a traditional host compared to the non-traditional hosts, thus, corroborating Liu (2014) who also reported that *Bactrocera correcta* (Bezzi) (Diptera: Tephritidae) and *Bactrocera dorsalis* (Hendel) (Diptera: Tephritidae) had high survival rates in their traditional hosts.

The impact of the host plant was also reflected in the genetic variation among the *B. tau* populations. Molecular variance analysis based on groups clustering by host-feeding populations (i.e., host groups) as well as the various generation populations (i.e., generation groups) demonstrated that significant differences were found

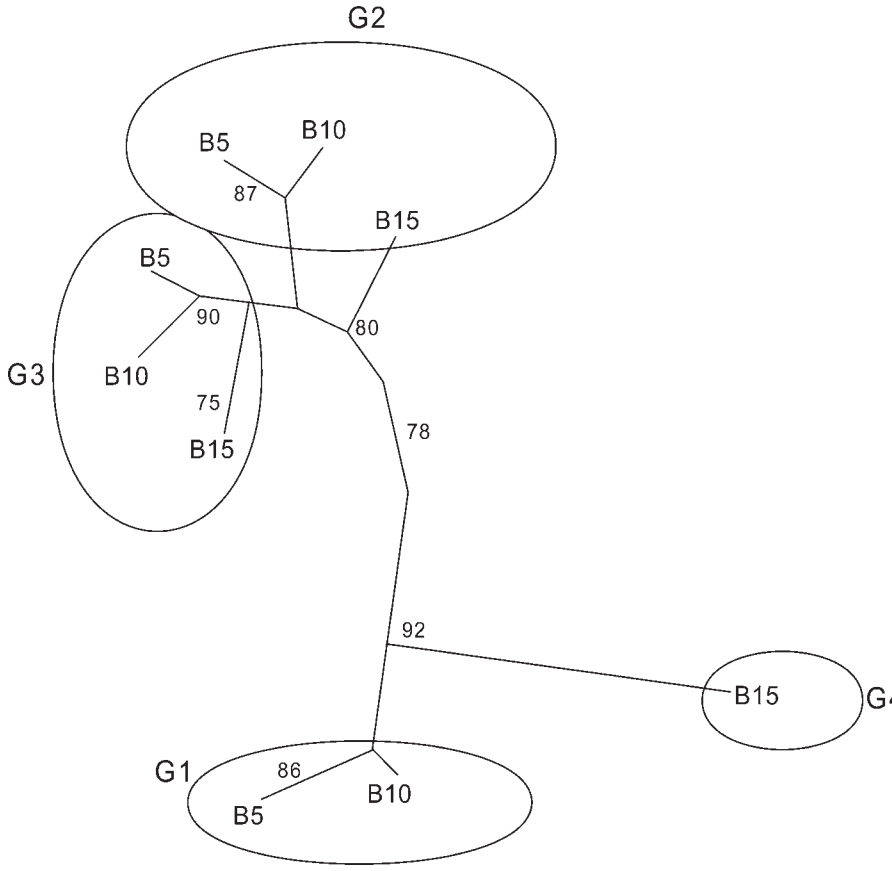

Fig. 3. Unrooted tree based on the proportion of shared alleles for microsatellite data. Bootstrap values after 1000 replicates are indicated by the number at each node. Only values above 70% are shown. The groups defined by five ovals correspond to the STRUCTURE groups defined by *K*=4.

among host groups with 10.39% variation; whereas, there were no significant differences among generation groups (Table 6). It appears that host food more than population generations played a critical role in the observed genetic variation of *B.tau* in our study.

The influence of the host plant on genetic variation in *B. tau* was primarily related to feeding on the non-traditional hosts. In our study, banana, belonging to Musaceae family, was considered a non-traditional host for *B. tau* (Huang et al. 2005). Moreover, banana could be considered a novel host for *B. tau*, given that there have been no reports of *B. tau* infesting banana under natural conditions (Huang et al. 2005). Compared with the two traditional hosts, the effect of banana on the genetic variation in *B. tau* was most obvious when considering the frequency balance of gene and genotype, genetic diversity variation, and genetic differences.

Table 7. Pairwise F_{ST} values of the nine *B. tau* populations.

Populations	P5	P10	P15	C5	C10	C15	B5	B10	B15
P5									
P10	0.046NS								
P15	0.051NS	0.043 NS							
C5	0.116	0.115	0.150						
C10	0.110	0.151	0.178	0.043 NS					
C15	0.104	0.173	0.194	0.072 NS	0.064 NS				
B5	0.133	0.157	0.170	0.084	0.094	0.097			
B10	0.111	0.147	0.159	0.171	0.126	0.077	0.084		
B15	0.236	0.222	0.204	0.218	0.255	0.256	0.228	0.212	

C5 (the 5th generation feeding on cucumber), B5 (the 5th generation feeding on banana), P5 (the 5th generation feeding on pumpkin), and similarly C10, B10, and P10, and C15, B15, and P15.
NS = no statistical significance.

Table 8. Mean assignment rate of individuals into (rows) and from (columns) each population based on microsatellite data as estimated by GeneClass 2.

Populations	P5	P10	P15	C5	C10	C15	B5	B10	B15
P5	**0.625**	0.030	0.027	0.016	0.015	0.000	0.047	0.046	0.000
P10	***0.059***	**0.569**	0.034	0.015	0.009	0.030	0.000	0.002	0.000
P15	0.011	*0.321*	**0.569**	0.003	0.002	0.011	0.000	0.002	0.000
C5	***0.056***	0.014	0.002	**0.595**	0.043	0.010	0.001	0.033	0.000
C10	0.041	0.011	0.008	*0.314*	**0.579**	0.042	0.033	0.015	0.002
C15	***0.055***	0.003	0.001	*0.241*	*0.235*	**0.544**	0.036	0.014	0.010
B5	0.044	0.005	0.001	0.037	***0.052***	0.046	**0.605**	0.041	0.023
B10	0.037	0.004	0.001	0.014	0.016	0.035	***0.072***	**0.599**	0.007
B15	0.004	0.000	0.000	0.003	0.006	0.000	0.042	0.035	**0.581**

C5 (the 5th generation feeding on cucumber), B5 (the 5th generation feeding on banana), P5 (the 5th generation feeding on pumpkin), and similarly C10, B10, and P10, and C15, B15, and P15.
Values in bold indicate the proportions of individuals assigned to the same population. Values of *m* above 0.05 are bold italic and above 0.1 are italicized. *m* is defined as mean assignment rate of individuals

In this study, banana-feeding fly populations in the 5th generation as well as the 10th generation had several microsatellite loci that deviated from Hardy–Weinberg equilibrium. Also, in the 15th generation, there remained an individual locus that deviated from equilibrium. By contrast, flies that developed on pumpkin and cucumber deviated from equilibrium only in the 5th generation, with the remaining generations in equilibrium at all loci. Hardy–Weinberg equilibrium could be used to estimate the state of population equilibrium by measuring frequency balance of genes and genotypes (Chen et al. 2005). For example, the 15th generation of the banana-feeding populations had gene and genotype frequencies that did not reach balance, suggesting that it was difficult for *B. tau* to reach equilibrium when developing on a novel host. The molecular results of this study were consistent with the survival rate results that indicated lower fitness for *B. tau* on banana, the non-traditional host.

Regarding genetic diversity, three genetic diversity indices (N_A, N_P, and H_S) were somewhat lower in banana-feeding populations compared with populations feeding on the two traditional Cucurbitaceae hosts; however, there were no differences in the three indices between the two Cucurbitaceae host-feeding populations. It should be noticed that a decreasing genetic diversity was also found in *B. dorsalis* when it infested apple (Rosaceae) and pear (Rosaceae), two non-traditional hosts (Shi et al. 2012).

Furthermore, the nine *B. tau* populations were divided into four groups in the genetic structure based on NJ tree and STRUCTURE (Figs. 2, 3). Among the four genetic groups, both the G1 and G4 groups were the fly populations feeding on banana, corresponding to the 5th and 10th generations, and the 15th generation,

whereas G2 and G3 were the flies feeding on pumpkin and cucumber, respectively. These results revealed that genetic differences were most obvious between banana-feeding populations and populations feeding on the two traditional hosts, with no significant differences between the flies feeding on the two Cucurbitaceae hosts (Table 5). This genetic divergence between banana-feeding populations and the two Cucurbitaceae-feeding groups increased with the increasing number of generations (Tables 5, 7). The most obvious genetic differences observed between the two banana-feeding groups (Table 5) were from different generations of banana-feeding flies. In contrast, the genetic difference between two Cucurbitaceae host-feeding groups did not diverge more with subsequent generations. Perhaps the greater genetic divergence observed in the banana-feeding populations arose from a bottleneck effect while developing on this novel host.

In summary, the results of our study indicate that the novel host banana had a greater impact on genetic variation in *B. tau* than the two traditional Cucurbitaceae hosts. Genetic variation is considered an important mechanism by which *B. tau* can adapt to the selective pressures of a novel host. A similar phenomenon also has been found in other insect species. For example, genetic variation was observed between two populations of *Ostrinia nubilalis* (Hübner) (Lepidoptera: Pyralidae) that were fed on corn and a novel host *Artemisia argyi* (H. Lév. & Vaniot) (Martel et al. 2003). Significant genetic difference also was observed between butterfly populations of the tribe Nymphalini when feeding on its traditional host *Urtica dioica* L. and the novel host *Ribes uva-crispa* L. (gooseberry) (Celorio-Mancera et al. 2013).

Under laboratory conditions, we estimated the impact of various host plants on the genetic variation of *B. tau* while attempting to exclude the influences of other environmental factors such as natural enemies. We found that *B.tau* can use a novel food, such as banana, if it enters a new area where traditional hosts are not present. The use of novel host plants appears to be an ecological genetic strategy to improve fitness when dispersing into new regions where traditional hosts may be lacking.

Acknowledgments

We are grateful to Robert A. Haack, who helped us improve a previous version of the manuscript. This research was supported by the National Key R&D Program of China (2017YFC0505200), the National Basic Research Program of China (grant no.2014, 31460163) and the Applied Basic Research Foundation Yunnan Province ([2009] CD010).

References Cited

Allwood, A.J., A. Chinajariyawong, S. Kritsaneepaiboon, R.A.I. Drew, E. L.Hamacek, D. L. Hancock, C. Hengsawad, J. C. Jipanin, M. Jirasurat, C. Kong Krong, C. T. S. Leong and S. Vijaysegaran. 1999. Host plant records for fruit flies (Diptera: Tephritidae) in southeast Asia. Raffles Bull. Zool. Suppl.7: 1–92.

Aketarawong, N., M. Bonizzoni, S. Thanaphum, M. Gomulski, G. Gasperi, A. R. Malacrida and C. R. Gugliemino. 2007. Inferences on the population structure and colonization process of the invasive oriental fruit fly, *Bactrocera dorsalis* (Hendel). Mol. Ecol. 16: 3522–3532.

Celorio-Mancera, M.D.L.P., C.W. Wheat, H. Vogel, L. Soderlind and S. Niklasjanz Nylin. 2013. Mechanisms of macroevolution: Polyphagous plasticity in butterfly larvae revealed by RNA-Seq. Mol. Ecol. 22: 4884–4895.

Chapuis, M.P. and A. Estoup. 2007. Microsatellite null alleles and estimation of population differentiation. Mol. Biol. Evol. 24: 621–631.

Chen, J.J., T. Duan and R. Single. 2005. Hardy–Weinberg testing of single homozygous genotype. Genetics 70: 1439–1442.

Christenson, L.D. and R.H. Foote. 1960. Biology of fruit flies. Annu. Rev. Entomol. 5: 171–192.

Dujardin, J.P. and K. Sangvorn. 2013. Phenetic structure of two *Bactrocera tau* cryptic species (Diptera: Tephritidae) infesting *Momordica cochinchinensis* (Cucurbitaceae) in Thailand and Laos. Zool. 116: 129–138.

Excoffier, L. and H.E.L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol.17: 431–449.

Faucci, A., R.J. Toonen and H.M.G. Adfield. 2007. Host shift and speciation in a coral-feeding nudibranch. Proc. Biol. Sci. 274: 111–119.

Felsenstein, J. 2005. PHYLIP (Phylogenyliference Package), Version3.6. Dept. Genome Sci., Univ. Washington, Seattle.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). 08 June 2016 (http://www2.unil.ch/popgen/softwares /fstat.htm)

Han, H.Y. and B.A. McPheron. 1994. Phylogenetic study of selected tephritid flies (Insecta: Diptera: Tephritidae) using partial sequences of the nuclear 18S ribosomal DNA. Biochem. Syst. Ecol. 22: 447–457.

Hasyim, A., M. Muryati and W.J. de Kogel. 2008. Population fluctuation of adult males of the fruit fly, *Bactrocera tau* Walker (Diptera: Tephritidae) in passion fruit orchards in relation to abiotic factors and sanitation. Indonesian J. Agr. Sci. 9: 29–33.

Huang, K.H., Q.X. Guo, Y. Yu and Z.H. Huang. 2005. Risk analysis of *Bactrocera* (Zeugodacus) *tau* (Walker). Wuyi Sci. J. 21: 77–80.

Huque, R. 2006. Comparative studies on the susceptibility of various vegetables to *Bactrocera tau* (Diptera: Tephritidae). Pak. J. Biol. 9: 93–95.

Jamnongluk, W., V. Baimai and P. Kittayapong. 2003a. Molecular phylogeny of tephritid fruit flies in the *Bactrocera tau* complex using the mitochondrial COI sequences. Genome 46: 112–118.

Jamnongluk, W., V. Baimai and P. Kittayapong. 2003b. Molecular evolution of tephitid fruit flies in the genus *Bactrocera* based on the cytochrome oxydase I gene. Genetica 119: 19–25.

Ju, R., H. Li, Ch. Shih and B. Li. 2012. Progress of biological invasions research in China over the last decade. Biodiv. Sci. 20: 581–611.

Khamis, F.M., N. Karam, S. Ekesi, M.D. Meyer, A. Bonomi, L.M. Gomulski, F. Scolari, P. Gabrieli, P. Siciliano, D. Masiga, E.U. Kenya, G. Gasperi, A.R. Malacrida and C. R. Guglielmino. 2009. Uncovering the tracks of a recent and rapid invasion: the case of the fruit fly pest *Bactrocera invadens* (Diptera: Tephritidae) in Africa. Mol. Ecol. 18: 4798–4810.

Liu, H., B.H. Hou, C. Zhang, R.R. He, F. Liang, M.F. Gu, M.T. Wu, J.P. Zhao and J. Ma. 2014. Oviposition preference and offspring performance of the oriental fruit fly *Bactrocera dorsalis* and guava fruit fly *B. correcta* (Diptera: Tephritidae) on six host fruits. Acta Ecologica Sinica 9: 2274–2281.

Malacrida, A., L. Gomulski, M. Bonizzoni, S. Bertlin, G. Gasperi and C.R. Guglielmino. 2007. Globalization and fruit fly invasion and expansion: the medfly paradigm. Genetica 131: 1–9.

Martel, C., A. Réjasse, F. Rousset, M.T. Bethenod and D. Bourguet. 2003. Host-plant-associated genetic differentiation in Northern French populations of the European corn borer. Heredity 90: 141–149.

Nardi, F., A. Carapelli, R. Dallai, G.K. Roderick and F. Frati. 2005. Population structure and colonization history of the olive fly *Bactrocera olea* (Diptera, Tephritidae). Mol. Ecol. 14: 2729–2738.

Ohno, S., Y. Tamura, D. Haraguchi and T. Kohama. 2008. First detection of the pest fruit fly, *Bactrocera tau* (Diptera: Tephritidae), in the field in Japan: evidence of multiple invasions of Ishigaki island and failure of colonization. Appl. Entomol. Zool. 43: 541–545.

Ooi, Y.T. and S.L. Wee. 2016. Sexual maturation, mating propensity and remating incidence of Zeugodacus tau (Walker) (Diptera:Tephritidae). J. Asia-Pac. Entomol. 19: 451–457.

Piry, S., A. Alapetite, J.M. Cornuet, D. Paetkau, L. Baudouin and A. Estoup. 2004. GeneClass2: A software for genetic assignment and first-generation migrant detection. J. Heredity 95: 536–539.

Pritchard, J. K., M. Stephens and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Rannala, B. and J.L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. 94: 9197–9201.

Rosenberg, N.A. 2004. Distruct: A program for the graphical display of population structure. Mol. Ecol. Notes 4: 137–138.

Rousset, F. 2008. Genepop007: A complete re-implementation of the GENEPOP software for Windows and Linux. Mol. Ecol. Res. 8: 103–106.

Saelee, A., S. Tigvattananont and V. Baimai. 2006. Allozyme electrophoretic evidence for a complex of species within the *Bactrocera tau* group (Diptera Tephritidae) in Thailand. Songklanakarin J. Sci. Technol. 28: 249–257.

Schwarz, D., B.M. Matta, N.L. Shakir-Botteri and B.A.H. Mcpheron. 2005. Host shift to an invasive plant triggers rapid animal hybrid speciation. Nature 428: 28–32.

Shi, W, C. Kerdelhue and H.Ye. 2012. Genetic structure and inferences on potential source areas for *Bactrocera dorsalis* (Hendel) based on mitochondrial and microsatellite markers. PLoS One *7*, Article ID e37083.

Singh, S.K., D. Kumar and V.V. Ramamurthy. 2010. Biology of *Bactrocera* (Zeugodacus) *tau* (Walker). Entomol. Res. 40: 259–263.

Snedecor, G.W., and W.G. Cochran. 1989. Statistical Methods. 8th ed. Ames, IA: Iowa State University Press.

Sumrandee, C., J.R. Milne and V. Baimai. 2011. Ovipositor morphology and host relations of the *Bactrocera tau* complex (Diptera: Tephritidae) in Thailand. Songklanakarin J. Sci. Technol. 33: 247–254.

Thanaphum, S. and U. Thaenkham. 2003. Relationships of forms within the *Bactrocera tau* (Walker) (Diptera: Tephritidae) taxon based on heat shock protein 70 cognate sequences. Ann. Entomol. Soc. Am. 96: 44–53.

Vargas, R.I., D. Miyashi and T. Nishida. 1984. Life history and demographic parameters of three laboratory-reared tephritids (Diptera: Tephritidae). Ann. Entomol. Soc. Am. 77: 651–656.

Walker, F. 1849. List of the specimens of dipterous insects in collection of the British Museum, part 4. Br. Mus. Lond. 77: 689–1172.

Wan, F.H. 2009. Research on Biological Invasions in China. Science Press, Beijing, China. 112 pp.

Wan, F.H. 2010. Biological Invasions: Risk Analysis and Early Prevention. Science Press, Beijing, China. 99 pp.

Wu, Y., Z.H. Li and J.J. Wu. 2009. Polymorphic microsatellite markers in the melon fruit fly, *Bactrocera cucurbitae* (Coquillett) (Diptera: Tephritidae). Mol. Ecol. Res. 9: 1404–1406.

Zhang, X.Y. and G.Q. Chen. 2012. Observation of *Bactrocera tau* infesting orange. Zhejiang Agric. Sci. 9: 1274–1275.

Zar, J. H. 1999. Biostatistical Analysis. Prentice Hall, Upper Saddle River, NJ.