

Genetic Diversity and Population Structure of *Zeugodacus Tau* Walker (Diptera: Tephritidae) in Southern Thailand

Jingyoh ZAELOR¹, Duangta JULSIRIKUL² and Sangvorn KITTHAWEE^{1,*}

¹Department of Biology, Faculty of Science, Mahidol University, Phaya Thai, Bangkok 10400, Thailand

²Department of Biology, Faculty of Science, Burapha University, Chonburi 20131, Thailand

(*Corresponding author's e-mail: sangvorn.kit@mahidol.ac.th)

Received: 6 August 2019, Revised: 13 February 2020, Accepted: 6 March 2020

Abstract

The population structure of a fruit fly, *Zeugodacus tau*, was studied by a Single-Strand Conformation Polymorphism (SSCP) technique. Twenty haplotypes of Cytochrome Oxidase unit I (COI) sequence were found in flies collected from Southern Thailand. The phylogenetic tree and haplotype network revealed gene flow across a large geographic range. With the aid of winds, their gene flow diminished population structure. The population size of *Z. tau* in Southern Thailand seemed to be large and stable, but the other populations in some locations had experienced a bottleneck effect, leading to local genetic differentiation. Fruit flies from the eastern areas had large effective population sizes, whereas the populations from other areas were smaller. This pattern matched the ecological niche centroid model, in which fruit flies disperse from high population areas to lower ones.

Keywords: COI, Dispersion, Effective population size, SSCP, Urbanization

Introduction

Genus *Zeugodacus* was considered to be a subgenus of genus *Bactrocera*, but recent genetic evidence suggests it to be another taxonomic rank [1,2]. Some species in this genus, such as *Zeugodacus cucurbitae* (Coquillett) (Syn: *Bactrocera cucurbitae* (Coquillett)), *Zeugodacus diversus* (Coquillett) (Syn: *Bactrocera diversus* (Coquillett)), and *Zeugodacus tau* (Walker) (Syn: *Bactrocera tau* (Walker)), cause serious agricultural damage, especially in cucurbit fruits [3-5]. The most abundant species in Thailand are *Z. cucurbitae* and *Z. tau* [6,7]. Although the *Z. cucurbitae* receives more attention due to its abundance [8,9], *Z. tau* is higher in diversity [4,10,11].

The estimation of fruit fly population size was crucial for effective pest management. However, the dynamics of fruit fly populations and their dispersal rates are not fully understood. Therefore, we studied genetic variation using molecular techniques to provide evidence concerning population changes and dispersal [12]. The sequence of Cytochrome Oxidase unit I (COI) was analyzed because it was influenced by a change in population size. Besides, COI is a widely used marker, and thus results from this study can be compared with those from other species [8,13-16]. The Single-Strand Conformation Polymorphism (SSCP) was used to visualize variation in DNA sequences, as it is efficient in analyzing large sample sizes [17,18].

The study of population genetics was carried out in Southern Thailand because genetic diversity in *Zeugodacus* spp. likely increases toward the equator [8,19,20]. Additionally, human disturbances, such as cucurbit cultivation and transportation, are relatively lower than in other areas. Hence, Southern Thailand was chosen because of its potentially high genetic diversity and low disturbances.

Materials and methods

Sample collections

Fruit flies were collected near cucurbit orchards at 18 locations within Southern Thailand (**Table 1, Figure 1**). At each location geographic coordinates were recorded by a Global Positioning System (GPS) and fruit fly traps were hung under tree shade for a week. Fruit fly traps, modified from Vargas *et al.* [21], were made of reused 300 - 500 ml bottles with 2 lateral holes. Fruit flies were lured into the bottle by cue-lure (4-(3-Oxobutyl) phenyl acetate), which was absorbed by a sponge glued inside each bottle. Propylene glycol was added to the bottom to preserve the flies. *Z. tau* was identified by using taxonomic keys [3,4], however, no *Z. tau* sibling species were found in the collected samples. The samples of *Z. tau* were preserved in labeled bottles in 70 % ethanol and stored at the Faculty of Science, Mahidol University, Bangkok.

Table 1 Details of sampled locations presented by groups, locations, GPS coordinates, provinces, dates of collection, and sample size.

Group ^a	Location	Coordinates		Province	Date of collection (dd/mm/yy)	Sample size
		Latitude	Longitude			
North	L01	9°10'9.9"N	99°48'9.2"E	Nakhon Si Thammarat	01/05/15	31
	L02	9°6'4.6"N	99°26'43.8"E	Surat Thani	01/05/15	30
	L03	9°5'27.9"N	99°46'23.6"E	Surat Thani	01/05/15	33
	L04	8°43'57.4"N	99°44'36.3"E	Nakhon Si Thammarat	22/05/15	36
	L05	8°31'52.1"N	99°50'57.4"E	Nakhon Si Thammarat	14/05/15	26
	L06	8°24'6.6"N	99°17'10.1"E	Nakhon Si Thammarat	24/05/15	60
	L07	8°19'40.4"N	99°47'0.5"E	Nakhon Si Thammarat	21/04/15	35
	L08	8°17'41.7"N	99°30'48.2"E	Nakhon Si Thammarat	25/04/15	48
	L09	8°15'59"N	99°53'58.5"E	Nakhon Si Thammarat	21/04/15	32
South	L10	8°15'58.2"N	99°6'23.1"E	Krabi	10/05/15	31
	L11	8°9'47.2"N	99°35'22.8"E	Nakhon Si Thammarat	26/04/15	24
	L12	8°9'23.9"N	99°32'18.8"E	Nakhon Si Thammarat	26/04/15	35
	L13	8°8'34.4"N	99°48'24.5"E	Nakhon Si Thammarat	22/04/15	29
	L14	8°6'5.8"N	99°48'28.6"E	Nakhon Si Thammarat	22/04/15	37
	L15	8°3'40"N	99°43'28.6"E	Nakhon Si Thammarat	26/04/15	36
	L16	8°1'8.7"N	99°32'49"E	Nakhon Si Thammarat	26/04/15	31
	L17	7°51'18.5"N	99°39'17.2"E	Trang	26/04/15	46
	L18	7°46'31.5"N	99°55'14.4"E	Phatthalung	22/04/15	30

^a locations were grouped based on result of cluster analysis of genetic distances among locations

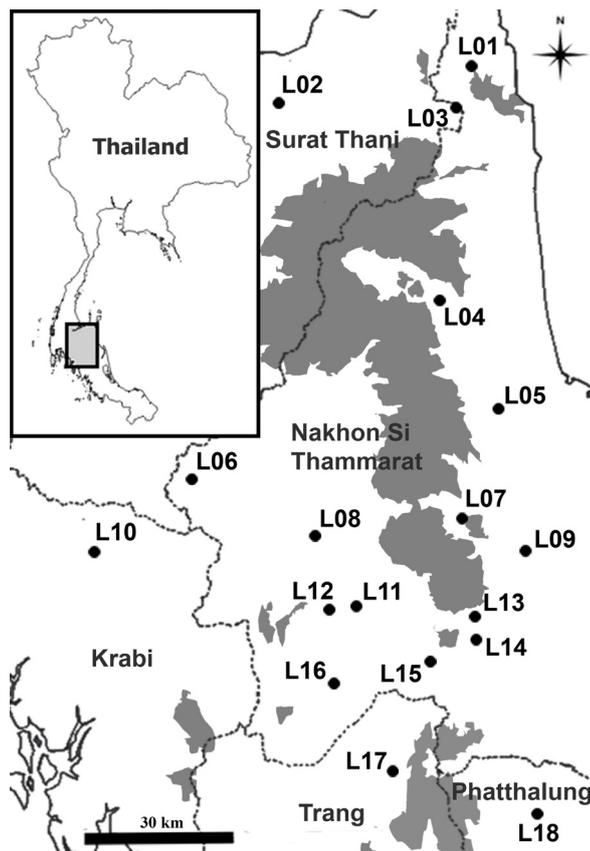


Figure 1 Map of studied region in Southern Thailand (upper left inset), the sampling locations (L01-L18) were black dots, whereas areas more than 200 m above sea level, province boundaries, and shorelines are shown by gray areas, dash lines, and solid lines, respectively.

Molecular processes

Sample flies from each location were dissected under a stereomicroscope. The head and thorax were separated from the abdomen, and then DNA material from the head and thorax was extracted by phenol-chloroform method [22]. DNA was partially amplified by Polymerase Chain Reaction (PCR) with primers: COI-F (5' TGG AGA CGA CCA AAT CTA 3') and COI-R (5' GAG GAA ATA CCA GCT AAA TG 3') [23]. The PCR cycle consists of 3 min at 95 °C of initiation, 35 cycles of 1 min each at 95, 50, and 72 °C, and extension for 10 min at 72 °C.

During the SSCP technique, PCR products were denatured and electrophoresed as follows. A mixture of 3 µl PCR products and 10 µl formamide was heated for 10 min at 95 °C and bathed on ice. The mixture as well as DNA ladders were run in 7 % of acrylamide gel (Acrylamide: bis-acrylamide 49:1) and 0.5X TBE buffer. The DNA was electrophoresed with 155 V for 165 min at 4 °C and then stained by silver nitrate solution. Similar band patterns were grouped as the same haplotype. The haplotype sequences were bi-directionally read by Sanger sequencing with modified primers that cover the sequence analyzed by SSCP: LCO1490(2) (5' CCA TTT AAT CGC GAC AAT GGC 3') and HC02198(2) (5' AAG AAT CAA AAT AAG TGT TGG TA 3') [24].

Data analysis

Haplotype sequences were aligned and edited for SSCP analyzed position (base 71st to 360th) in BIOEDIT ver.7.1.9 [25]. A haplotype network was generated by the minimum spanning method in POPART [26]. The network was presented as clustered groups, shown in **Table 1**. Similar sequences of *Z. tau* and *Z. cucurbitae* from the National Center for Biotechnology Information (NCBI) [27] database were included for phylogenetic tree construction using the maximum likelihood method (5000 bootstraps) in MEGA ver. 6 [28].

All population genetic values were computed in ARLEQUIN ver. 3.5 [29]. Genetic diversities were calculated in terms of numbers of haplotypes, haplotype diversity (Hd), and nucleotide diversity (Pi). Population neutrality was tested by Tajima's D, Fu's Fs, and Harpending's raggedness index (Hri). The first two neutrality tests were verified by 10,000 simulated samples, whereas the third test was examined by 10,000 bootstrap repeats. The effective population sizes (Theta H) were computed for both within locations and total samples. Genetic distances among locations were presented by Pi(xy). A cladogram of Pi(xy) values was constructed by Hierarchical cluster analysis in the R program [30]. The cladogram was then divided at different levels to generate different numbers of groups. Each grouping pattern was verified by Analysis of MOlecular VAriance (AMOVA) in terms of genetic distance among groups (Fct) and genetic distance among locations within the same group (Fsc) in ARLEQUIN.

Results and discussion

Haplotype analysis

Among 630 samples from 18 locations, 20 haplotypes were found. The haplotype sequences in this study were submitted to NCBI (MH500583 to MH500602). The polymorphic sites consisted of 15 singletons and five parsimony-informative sites (**Table 2**). The low levels of polymorphism in haplotype sequences could be due to the conservation of the protein sequences vital for fruit fly's survival. Based on analysis by NCBI, all polymorphic sites were silent mutations; hence, all fruit flies had equal chances of survival.

The minimum spanning method revealed a haplotype network, dominated mainly by H01 and H02 (**Figure 2**). Haplotypes H01 and H02 were likely the ancestral haplotypes since they were ubiquitous and surrounded by other haplotypes. The 20 haplotypes were equally distributed across all locations (**Figure 2** and **Table 3**). However, H01 was the most abundant haplotype in the South group (L01-L09), while H02 was the most abundant haplotype in the North group (L10-L18). H11, H20, and other rare haplotypes were found mainly in the N2 group. These variations would lead to significant Fct and difference in Theta H in further analyses.

Multiple loops within the network were evidence of reverse mutation which only accumulate in large populations over several generations. Loops were formed by parsimony-informative sites, such as positions 122 and 326 which caused a loop in H02, H04, and H07 (**Table 2** and **Figure 2**). Several other loops in the network were also results of parsimony-informative sites. These sites were products of reverse mutations that likely accumulated over many generations [31,32].

Combining the haplotypes from this study with sequences from the NCBI database gives a total of 56 sequences of *Z. tau* (**Figure 3**). A sequence from China (Hunan) was distinct from other *Z. tau*, but the difference was insignificant. Regardless of the high diversity as shown by phylogeny, the divergence among sequences was only several nucleotides. This supported the hypothesis that most of the mutations in COI sequences were likely fatal. Seven out of 20 haplotypes, H01, H02, H05, H06, H09, H11, and H15, were also reported from other geographical areas. Haplotypes H01 to H20 were scattered across the phylogenetic tree, and not clustered on the same clade. This phylogenetic pattern prevents the study of relations among sequences and does not allow *Z. tau*'s origin to be identified.

Some features of the haplotype network and phylogenetic tree suggest a high dispersal ability of *Z. tau*. First of all, some haplotypes in this study were also present in China, India, and Malaysia (**Figure 3**). Moreover, the topology of the phylogenetic tree is not concordant with the geographical origins of the sequences. This suggests that they dispersed across the region and were not limited by distances or barriers. According to NCBI records, haplotype H01 was the most commonly reported haplotypes across the geographic range. This suggests a long-range gene flow that will maintain genetic similarity among regions.

Table 2 The 20 polymorphic sites in SSCP analyzed sequences, periods indicated invariant sites from H01, white and gray columns were singleton and parsimony-informative sites, respectively.

Haplotype	Nucleotide position																			
	116	122	140	153	162	167	176	188	218	227	251	257	269	272	299	317	326	329	332	353
H01	G	G	A	T	C	A	A	G	A	T	G	T	C	A	T	C	T	A	C	T
H02	G
H03	G	.	.
H04	.	A	.	.	.	G	T	G
H05	A	.	.	.
H06	A	A	.	.	.
H07	.	A	G	.	.	A	.	.	.
H08	A	C
H09	.	.	G
H10	A	.
H11	A
H12	T	.	A	.	.	.	C
H13	.	A	C
H14	A	A	T
H15	.	A
H16	.	A	G
H17	A	G
H18	A	G
H19	A	C
H20	A	.	.	C	A

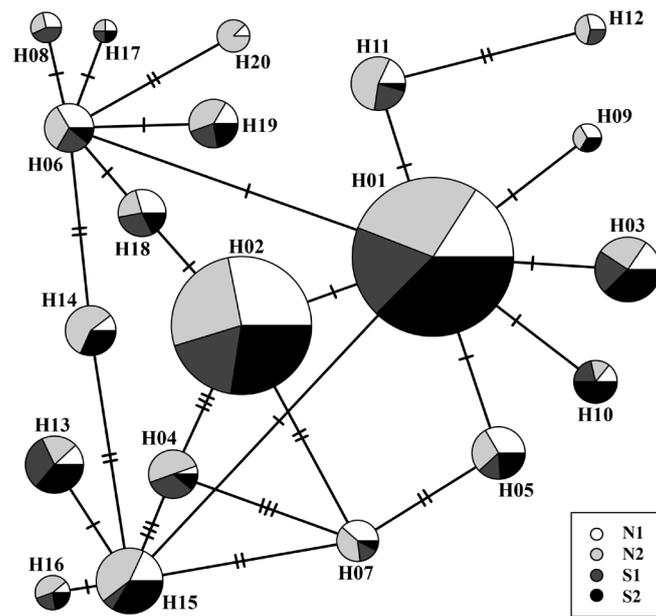


Figure 2 Minimal spanning networks of 20 haplotypes in SSCP analysis. Circle sizes correlate with haplotype abundance. Sizes of pies in each circle indicate the proportional contributions of groups, namely N1, N2, S1, and S2, and hatch marks indicate number of nucleotide substitutions between haplotype.

Table 3 Summary of haplotype (H01-H20) abundances in sampled locations (L01-L18).

Haplotype	Location																		Total
	L01	L02	L03	L04	L05	L06	L07	L08	L09	L10	L11	L12	L13	L14	L15	L16	L17	L18	
H01	7	7	10	7	7	23	2	14	8	9	7	10	10	8	13	10	26	15	193
H02	12	8	15	6	6	9	8	9	7	5	6	9	6	9	8	7	12	4	146
H03	1	1	1	2	3	1	2	2	0	4	1	1	1	3	4	2	2	1	32
H04	0	0	1	0	1	2	4	1	1	2	0	2	2	0	0	1	0	1	18
H05	3	3	0	1	0	3	0	3	0	0	0	2	1	1	0	2	0	2	21
H06	1	0	0	5	1	1	2	1	1	0	1	2	1	1	1	0	0	0	18
H07	2	1	0	2	0	1	2	1	1	0	0	1	1	0	1	0	0	0	13
H08	1	1	0	0	0	0	2	0	0	1	0	1	1	0	0	0	0	0	7
H09	0	1	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0	1	6
H10	1	0	0	1	0	0	0	2	0	1	0	1	1	2	1	2	1	1	14
H11	0	2	0	2	1	5	3	2	1	2	3	0	0	0	0	0	1	0	22
H12	0	2	0	0	0	1	1	0	1	0	2	0	0	0	0	0	0	0	7
H13	0	0	0	3	0	1	1	2	1	2	1	3	2	3	2	2	1	1	25
H14	0	0	2	0	2	3	0	2	4	0	0	0	0	1	3	1	0	1	19
H15	2	1	0	3	3	2	3	3	3	1	1	0	0	5	1	1	2	2	33
H16	0	0	0	1	0	2	0	0	2	1	0	0	1	1	0	1	0	0	9
H17	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	4
H18	0	1	2	2	1	2	0	0	1	2	0	2	1	1	0	2	0	0	17
H19	1	1	0	1	1	3	3	0	0	0	2	1	1	1	2	0	0	1	18
H20	0	0	1	0	0	0	0	6	1	0	0	0	0	0	0	0	0	0	8

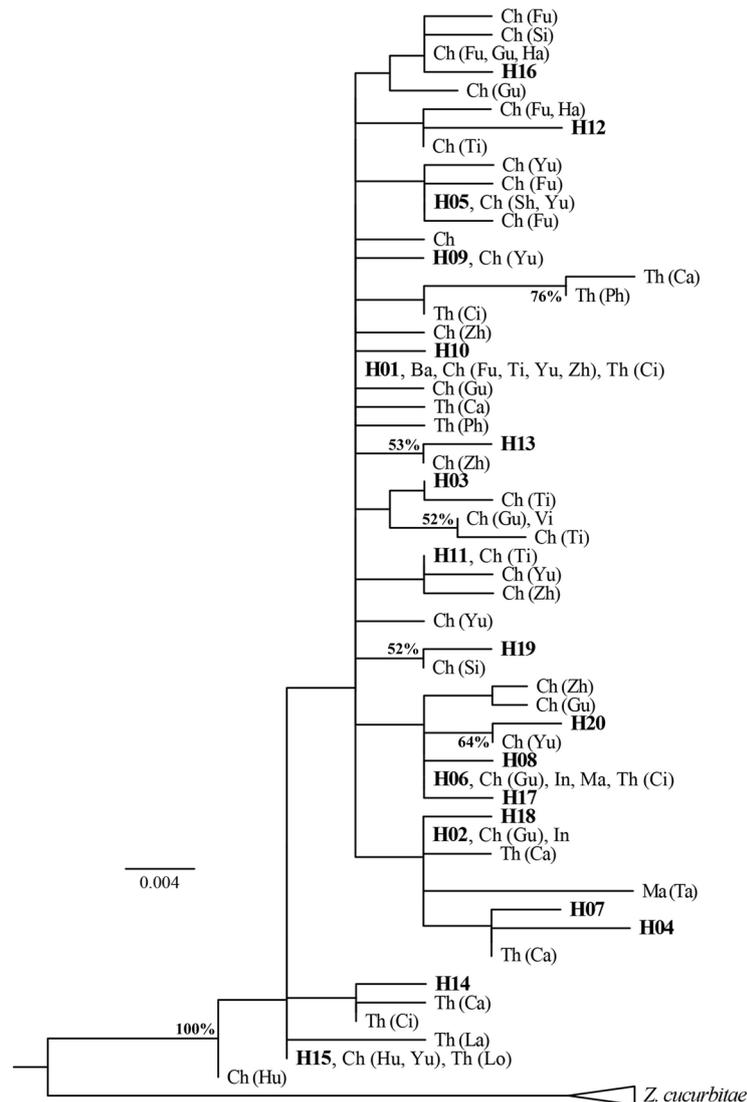


Figure 3 Maximum likelihood phylogenetic tree of COI sequences of *Z. tau* (General time-reversible model, 5000 bootstraps) with *Z. cucurbitae* as an outgroup. Values of bootstraps support higher than 50 % are shown on nodes. Scale bar presents 0.004 substitutions per site. Haplotype name (bold) and sequence origin are labeled at each branch. The abbreviations of countries and provinces in parentheses are: Ba Bangladesh; Ca Chaiphum; Ch China; Ci Chiang rai; Fu Fujian; Gu Guizhou; Ha Hainan; Hu Hunan; In India; La Lampang; Lo Loei; Ma Malaysia; Ph Phayao; Sh Shenzhen; Si Sichuan; Ta Tanah Rata; Th Thailand; Ti Tibet; Vi Vietnam; Yu Yunnan and Zh Zhejiang.

Genetic diversities

The numbers of haplotypes ranged from 8 to 16 among studied locations (**Table 4**). The genetic diversities (H_d and P_i) varied among locations, but most did not deviate from estimations of the total sample. According to **Table 4**, the tendencies of P_i and H_d values were similar, for example, L07 had the highest H_d and P_i values, whereas L17 had the lowest. However, the genetic diversities at each location

were not influenced by differences in sample sizes and were already larger than the recommended sample size [33]. In addition, further statistical analysis shows an insignificant correlation between sample sizes and genetic diversities (Hd: $r = -0.217$, $p = 0.37$ and Pi: $r = -0.121$, $p = 0.63$).

The genetic diversities of *Z. tau* (Hd = 0.9193 - 0.6203; **Table 4**) were similar to *Z. cucurbitae* (Hd = 1.0000 - 0.2857) [8]. However, some fruit fly species could have higher genetic diversity than found in this study, such as *Bactrocera dorsalis* (Hendel) (Hd = 1.000 - 0.8095) and *Bactrocera correcta* (Bezzi) (Hd = 0.9890 - 0.6670), while some other species, such as *Bactrocera oleae* (Rossi) had lower genetic diversity (Hd = 0.9778 - 0.4667) [14,15,34-36]. The difference in genetic diversity might reflect differences in host plant range [37]. Species with a wider host plant range, such as *B. dorsalis* and *B. correcta*, which infest 39 and 28 families of plants, respectively [38], had higher genetic diversities. Species with a narrower host plant range, such as *B. oleae*, which infests one family of plants [38], had lower genetic diversities. *Z. tau* and *Z. cucurbitae*, which infest 9 and 12 families of plants, respectively [38], had middle levels of genetic diversities.

Table 4 Summary of genetic analyses at all locations, including genetic diversities, neutrality tests, and Theta H.

Group and location	Genetic diversities			Neutrality tests			Theta H ^a	
	N	Hd	Pi	Tajima's D	Fu's Fs	Hri		
N1	L01	10	0.8022	0.0051	-0.813	-4.613 *	0.088 *	3.218 ± 1.203
	L02	13	0.8759	0.0066	-1.220	-7.284 *	0.058 *	5.898 ± 2.381
	L03	8	0.7121	0.0050	-1.304	-2.318	0.102	1.898 ± 0.579
	L04	13	0.9111	0.0064	-0.708	-6.555 *	0.100 *	8.892 ± 2.807
N2	L05	10	0.8677	0.0063	-0.757	-4.105 *	0.067 *	5.442 ± 2.105
	L06	16	0.8237	0.0062	-1.317	-8.569 *	0.037	3.755 ± 1.164
	L07	14	0.9193	0.0093	-0.861	-5.359 *	0.099 *	9.986 ± 3.912
	L08	13	0.8644	0.0072	-0.878	-4.634 *	0.052 *	5.273 ± 1.549
	L09	13	0.8810	0.0078	-1.164	-5.753 *	0.038	6.222 ± 2.270
S1	L10	12	0.8796	0.0068	-1.115	-5.458 *	0.048	6.126 ± 2.472
	L11	9	0.8515	0.0064	-0.873	-0.873	0.053	4.694 ± 1.880
	L12	12	0.8521	0.0061	-0.955	-0.955	0.044	4.720 ± 1.637
	L13	13	0.8473	0.0067	-1.210	-1.210	0.032	4.530 ± 2.005
S2	L14	13	0.8784	0.0059	-1.104	-1.104	0.093 *	6.051 ± 1.914
	L15	10	0.8143	0.0058	-0.673	-0.673	0.046	3.503 ± 1.154
	L16	11	0.8473	0.0059	-1.209	-1.209	0.061 *	4.530 ± 1.751
	L17	8	0.6203	0.0028	-1.522 *	-1.522 *	0.120	1.232 ± 0.339
	L18	11	0.7402	0.0047	-1.792 *	-1.792 *	0.067	2.204 ± 0.986
Total	20	0.8388	0.0062	-0.901	-5.474	0.043 *	4.898 ± 1.784	

Definitions: *n* number of haplotypes, *Hd* Haplotype diversity, *Pi* nucleotide diversity, *Hri* Harpending's raggedness index, *Theta H* effective population size

* significant Tajima's D and Hri ($p < 0.05$), and significant Fu's Fs ($p < 0.02$)

^a Theta H values are presented by average ± standard error of bootstrap repeats

Population structure

Regarding the relationships among locations, cluster analysis divided locations by their Pi(xy) into multiple groups (**Figure 4**). The first division resulted in North (L01-L09) and South (L10-L18) groups (**Table 5**). The Fct indicated insignificant isolation between the 2 groups, but differences within groups (Fsc) were significant. In the 3 group scenario, consisting of N1 (L01-L04), N2 (L05-L09), and South (L10-L18), there was a significant genetic difference among groups (Fct). Further analysis concluded that

genetic differences were significant in all pairs of groups. Moreover, an insignificant F_{sc} revealed the lack of genetic difference among locations within each group. Further division led to isolation within the South group between S1 (L10-L13) and S2 (L14-L18). The 4 group scenario also resulted in significant F_{ct} but insignificant F_{sc} . However, the pairwise analysis did not detect a significant genetic distance between S1 and S2 groups, indicating that they were not genetically differentiated. Therefore, the 3 group scenario was considered to be the most valid configuration.

The influence of human activities on the result of this study is believed to be negligible, not only because cucurbit cultivation is low, but also because the population structure of *Z. tau* does not coincide with the pattern of human activities. In the studied region, 2 main roads lie in a north-south direction on each side of the main mountain range, so north-south transportation provides the main avenue of dispersal of the fruit flies. However, the pattern of genetic structure was perpendicular to this path, which suggests that fly dispersal is independent of human transportation paths. The most significant genetic structure was the 3 group pattern (N1, N2, and South), concordant with their latitude (**Figures 1 and 4**). This is likely a result of the winds that seasonally blow in the east-west direction [39]. If adult fruit flies were carried by seasonal winds, their genes would be introduced across the mountains. Various insect species are also aided by winds during their dispersal [40,41]. Hence, the prediction of fruit fly population structure should take wind directions into account, because they influence their dispersal ability.

Despite the significance of F_{ct} , it occupied less than 1 % of the total genetic distance (**Table 5**). The small percentage of F_{ct} suggested that the population structure was illusive, and that *Z. tau* intensively dispersed among locations. Similar percentages of F_{ct} could be found in other fruit fly species, such as *Z. cucurbitae* (0.31 - 8.82 %), *B. correcta* (0.95 - 1.96 %), *B. dorsalis* (2.56 - 12.86 %) and *B. oleae* (8.83 - 53.75 %) [8,9,14,16,33,34]. The variation in F_{ct} percentages was likely a result of differences in species and geographic scale of studies. In *B. dorsalis*, 12.86 % of F_{ct} resulted from samples from China and 5 countries in Southeast Asia, on a large geographic scale [16]. An extreme case, 53.75 % of F_{ct} (*B. oleae*) was likely a result of host plant limitation as well as the large scale of the study [34]. The less than 1 % of F_{ct} in this study relates to the lower chance of dispersing in the north-south direction rather than the east-west direction, without significant isolation (**Figures 1 and 4**). Therefore, the low percentage of F_{ct} in *Z. tau* reflects their high dispersal ability, rather than dispersal limitation.

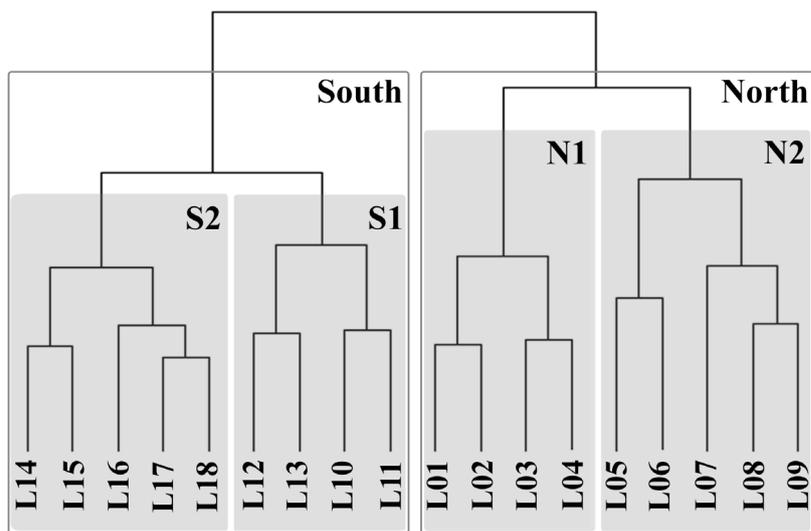


Figure 4 The cladogram of $Pi(xy)$ among locations, the locations were divided into North group, consisting of N1 and N2, and South group, consisting of S1 and S2. They were named based on the latitude of locations.

Table 5 AMOVA results among groups, consisted of North (N1 and N2) and South (S1 and S2) groups as described in **Figure 4**, whereas asterisks were an indication of significance at $p < 0.05$.

Grouping pattern	Groups of comparison	Genetic distance (percentage)	
		F _{ct}	F _{sc}
North and South	North and South	0.00345 (0.34)	0.00742* (0.74)
	N1, N2 and South	0.00752* (0.75)	0.00420 (0.42)
N1, N2 and South	N1 and N2	0.00867* (0.87)	0.00903 (0.89)
	N1 and South	0.00776* (0.78)	0.00464 (0.46)
	N2 and South	0.00680* (0.68)	0.00109 (0.11)
N1, N2, S1 and S2	N1, N2, S1 and S2	0.00711* (0.71)	0.00367 (0.36)
	S1 and S2	0.00453 (0.45)	-0.00315 (0.00)

Effective population size and neutrality tests

The overall Theta H value was 4.898, but the tendency of Theta H in each location was similar to its genetic diversities (**Table 4**), for example, L07 had the highest genetic diversities and also the largest Theta H value. L17 had the lowest genetic diversities and the smallest Theta H value. Most of the locations with high Theta H values were on the east side of the mountains (**Figure 1**). West of this area, Theta H values gradually decreased with distance. Based on our ecological hypothesis, locations with high Theta H should be considered to the centroid of ecological niches of the studied area [16,42,43]. The hypothesis describes the centroid as the area with the most favorable environment for population growth. Based on this hypothesis, *Z. tau* should be most abundant on the east side due to more favorable factors such as more diverse and abundant host plants [37]. Fruit flies from the centroid continually dispersed to nearby habitats. However, the lack of a suitable environment in nearby areas limited both population sizes and the diversity of haplotypes. The high diversity of haplotypes found on the east side of the mountains supports this interpretation. The pattern of Theta H in this study is similar to that of the population of *B. dorsalis* in China, in which genetic diversity was highest in the southern samples [15,16]. The authors of these studies suggested that the fruit flies dispersed northward into more hostile environments, an explanation similar to ours regarding the dispersal pattern of *Z. tau* in Southern Thailand.

Since the neutrality tests were computed on different models and hypotheses, they gave different results for different locations (**Table 4**). While Tajima’s D and Fu’s F_s are based on the stable population hypothesis, H_{ri} is based on a population expansion hypothesis. Tajima’s D was the least sensitive test and only detected significant population fluctuations in L17 and L18. However, different locations were also influenced by different causes. Locations L09 and L10 were significantly disturbed from neutrality, but the high genetic diversities in these locations suggested the possibility that fruit flies from those locations had dispersed there from nearby locations. In contrast, L17 and L18 had presumably experienced bottlenecks since the tests showed disturbance from neutrality but with low genetic diversities. The bottleneck effect in these locations may result in significant population structure by increasing genetic differences among groups (**Figure 4**). A similar phenomenon has also been described in *Z. cucurbitae* [8]. However, *Z. tau* from southern Thailand was under neutral selection pressure. This leads to a conclusion of a large and stable population in the studied area, despite disturbances in some locations.

To some extent, results of neutrality tests suggested that the genetic diversity of *Z. tau* was sensitive to local effects. Effects of human disturbance have also been reported in other species. *B. oleae* originated in Africa, but it expanded into the Mediterranean region as olive cultivation expanded [44,45]. Its genetic

diversity was also reduced as an insecticide was applied to the fields. *Z. cucurbitae* is another example of a fruit fly that has been being dispersed across continents by human transportation [46].

Conclusions

Overall, genetic evidence revealed great dispersal ability, aided by both winds and human activities. At the local scale, the significant AMOVA and neutrality tests revealed selection pressure in some locations, likely caused by human activities. On a larger scale, genetic similarity between regions supports an idea of wide-ranging gene flow. Mutation accumulation and Theta H suggested that the population size of *Z. tau* in southern Thailand was large and stable. All in all, the flies could be under selection pressure at the local scale, but the effects of selection are diminished by their dispersal ability, as revealed by analysis at larger scales.

Acknowledgements

We would like to thank Warren Y. Brockelman, Institute of Molecular Bioscience, Mahidol University for the critical reading of the manuscript and providing useful comments. We are grateful to the Department of Biology, Faculty of Science, and the Faculty of Graduate Studies, Mahidol University, for providing facilities. This research has been supported by the Center of Excellence on Biodiversity (BDC), Office of Higher Education Commission (BDC-PG2-160006-7).

References

- [1] M Virgilio, K Jordaens, C Verwimp, IM White and MD Meyer. Higher phylogeny of frugivorous flies (Diptera, Tephritidae, Dacini): localised partition conflicts and a novel generic classification. *Mol. Phylogenet Evol.* 2015; **85**, 171-9.
- [2] JR Dupuis, FT Bremer, A Kauwe, MS Jose, L Leblanc, D Rubinoff and SM Geib. HiMAP: Robust phylogenomics from highly multiplexed amplicon sequencing. *Mol. Ecol. Resour.* 2018; **18**, 1000-19.
- [3] IM White and MM Elson-Harris. *Fruit flies of economic significance: their identification and bionomics*. CAB International, Wallingford, 1992, p. 600.
- [4] RA Drew and MC Romig. *Tropical fruit flies of South-East Asia: (Tephritidae: Dacinae)*. CAB International, Wallingford, 2013, p. 664.
- [5] SA Ganie, ZH Khan, RA Ahangar, HA Bhat and B Hussain. Population dynamics, distribution, and species diversity of fruit flies on cucurbits in Kashmir valley, India. *J. Insect Sci.* 2013; **13**, 1-7.
- [6] A Chinajariyawong, S Kritsaneepaiboon and RAI Drew. Efficacy of protein bait sprays in controlling fruit flies (Diptera: Tephritidae) infesting angled luffa and bitter gourd in Thailand. *Raffles B. Zool.* 2003; **51**, 7-15.
- [7] MM Ramadan and RH Messing. A survey for potential biocontrol agents of *Bactrocera cucurbitae* (Diptera: Tephritidae) in Thailand. *Proc. Hawaii Entomol. Soc.* 2003; **36**, 115-22.
- [8] C Kunprom and P Pramual. Genetic structure and demographic history of the melon fly *Zeugodacus cucurbitae* (Coquillett) (Diptera: Tephritidae) in Thailand. *Agr. For. Entomol.* 2017; **20**, 180-90.
- [9] Y Boontop, N Kumaran, MK Schutze, AR Clarke, SL Cameron and MN Krosch. Population structure in *Zeugodacus cucurbitae* (Diptera: Tephritidae) across Thailand and the Thai-Malay peninsula: Natural barriers to a great disperser. *Biol. J. Linn. Soc.* 2017; **121**, 540-55.
- [10] V Baimai, J Phinchongsakuldit, C Sumrandee and S Tigvattananont. Cytological evidence for a complex of species within the taxon *Bactrocera tau* (Diptera: Tephritidae) in Thailand. *Biol. J. Linn. Soc.* 2000; **69**, 399-409.
- [11] J Zaelor and S Kitthawee. Geometric morphometric and molecular evidence suggest a new fruit fly species in *Bactrocera (Zeugodacus) tau* complex (Diptera: Tephritidae). *Zool. Systemat.* 2018; **43**, 27-36.
- [12] MB Hamilton. *Population genetics*. Welley-Blackwell, Malaysia, 2009, p. 424.

- [13] CS Prabhakar, P Sood, PK Mehta and PN Sharma. Population genetic structure of the pumpkin fruit fly, *Bactrocera tau* (Walker) (Diptera: Tephritidae) in Himachal Pradesh, India. *Biochem. Syst. Ecol.* 2013; **51**, 291-6.
- [14] J Liu, W Shi and H Ye. Population genetics analysis of the origin of the Oriental fruit fly, *Bactrocera dorsalis* Hendel (Diptera: Tephritidae), in northern Yunnan Province, China. *Entomol. Sci.* 2007; **10**, 11-9.
- [15] P Chen and H Ye. Relationship among five populations of *Bactrocera dorsalis* based on mitochondrial DNA sequences in western Yunnan, China. *J. Appl. Entomol.* 2008; **132**, 530-7.
- [16] W Shi, C Kerdelhué and H Ye. Genetic structure and inferences on potential source areas for *Bactrocera dorsalis* (Hendel) based on mitochondrial and microsatellite markers. *PloS One* 2012; **7**, 1-15.
- [17] F Nasernakhaei, MR Rahiminejad, H Saedi and M Tavassoli. PCR-SSCP: A method to fingerprint sequence variability in plant gene pools using the Iranian wild diploid *Triticum* gene pool as a case study. *J. Cell Mol. Res.* 2014; **6**, 38-43.
- [18] P Sunnucks, ACC Wilson, LB Beheregaray, K Zenger, J French and AC Taylor. SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Mol. Ecol.* 2000; **9**, 1699-710.
- [19] J Hu, JL Zhang, F Nardi and RJ Zhang. Population genetic structure of the melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae), from China and Southeast Asia. *Genetica* 2008; **134**, 319-24.
- [20] W Shi, C Kerdelhué and H Ye. Genetic structure and colonization history of the fruit fly *Bactrocera tau* (Diptera: Tephritidae) in China and Southeast Asia. *J. Econ. Entomol.* 2014; **107**, 1256-65.
- [21] RI Vargas, JD Stark, MH Kido, HM Ketter and LC Whitehand. Methyl eugenol and cue-lure traps for suppression of male oriental fruit flies and melon flies (Diptera: Tephritidae) in Hawaii: Effects of lure mixtures and weathering. *J. Econ. Entomol.* 2000; **93**, 81-7.
- [22] AF Cockburn and GA Fritz. *Isolation and purification of insect DNA*. In: JP Clapp (Ed.) Species Diagnostics Protocols. Humana Press, New York, 1996, p. 15-23.
- [23] P Somsill, D Julsirikul, P Rakthong and S Kitthawee. Morphology and genetic diversity of the *Bactrocera tau* complex in Southern Thailand. *Rajabhat J. Sci., Hum. Soc. Sci.* 2015; **16**, 334-46.
- [24] O Folmer, M Black, W Hoeh, R Lutz and R Vrijenhoek. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* 1994; **3**, 294-9.
- [25] TA Hall. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Res.* 1999; **41**, 95-8.
- [26] HJ Bandelt, P Forster and A Rohl. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 1999; **16**, 37-48.
- [27] NCBI Resource Coordinators. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 2017; **45**, D12-D17.
- [28] K Tamura, G Stecher, D Peterson and S Kumar. MEGA. Molecular evolutionary genetics analysis. Version 6.0. *Mol. Biol. Evol.* 2013; **30**, 2725-9.
- [29] L Excoffier and HE Lischer. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 2010; **10**, 564-7.
- [30] R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Available at: <https://softlibre.unizar.es/manuales/aplicaciones/r/fullrefman.pdf>, accessed December 2017.
- [31] WM Brown, M George and AC Wilson. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 1979; **76**, 1967-71.
- [32] BS Arbogast, SV Edwards, J Wakeley, P Beerli and JB Slowinski. Estimation divergence time from molecular data on phylogenetic and population genetic timescale. *Annu. Rev. Ecol. Evol. Syst.* 2002; **33**, 707-40.
- [33] WP Goodall-Copestake, GA Tarling and EJ Murphy. On the comparison of population-level estimates of haplotype and nucleotide diversity: A case study using the gene *cox1* in animals. *Heredity* 2012; **109**, 50-6.

- [34] CN Eti, E Dogac, BG Taskin, G Gokdere and V Taskin. Population structure and patterns of geographic differentiation of *Bactrocera oleae* (Diptera: Tephritidae) in eastern Mediterranean basin. *Mitochondrial DNA A*. 2017; **29**, 1051-62.
- [35] YJ Qin, N Buahom, MN Krosch, Y Du, Y Wu, AR Malacrida, YL Deng, JQ Liu, XL Jiang and ZH Li. Genetic diversity and population structure in *Bactrocera correcta* (Diptera: Tephritidae) inferred from mtDNA *cox1* and microsatellite markers. *Sci. Rep.* 2016; **8**, 1-10.
- [36] C Kunprom, PN Sopaladawan and P Pramual. Population genetics and demographic history of guava fruit fly *Bactrocera correcta* (Diptera: Tephritidae) in northeastern Thailand. *Eur. J. Entomol.* 2015; **112**, 227-34.
- [37] L Laukkanen, P Mutikainen, A Muola and R Leimu. Plant-species diversity correlates with genetic variation of an Oligophagous seed predator. *PLoS One* 2014; **9**, 1-9.
- [38] Plant Health Australia. *Fruit fly ID Australia*, Available at: <https://fruitflyidentification.org.au>, accessed January 2020.
- [39] International Research Institute for Climate and Society. *Monthly wind climatology*, Available at: https://iridl.ldeo.columbia.edu/maproom/Global/Climatologies/Vector_Winds, accessed March 2019.
- [40] A Otuka, K Nagayoshi, S Sanada-Morimura, M Matsumura, D Haraguchi and R Kakazu. Estimation of possible sources for wind-borne re-invasion of *Bactrocera dorsalis* complex (Diptera: Tephritidae) into islands of Okinawa Prefecture, southwestern Japan. *Appl. Entomol. Zool.* 2016; **51**, 21-35.
- [41] JA Byers. Analysis of vertical distributions and effective flight layers of insects: Three-dimensional simulation of flying insects and catch at trap heights. *Behavior* 2011; **40**, 1210-22.
- [42] E Martínez-Meyer, DF Díaz-Porras, AT Peterson and C Yañez-Arenas. Ecological niche structure and rangewide abundance patterns of species. *Biol. Lett.* 2013; **9**, 1-5.
- [43] PG Martínez-Gutiérrez, E Martínez-Meyer, F Palomares and N Fernández. Niche centrality and human influence predict rangewide variation in population abundance of a widespread mammal: The collared peccary (*Pecari tajacu*). *Divers. Distrib.* 2017; **24**, 103-15.
- [44] MD Ochado and A Reyes. Genetic population structure in olive fly *Bactrocera oleae* (Gmelin): gene flow and patterns of geographic differentiation. *J. Appl. Entomol.* 2000; **124**, 177-83.
- [45] F Nardi, A Carapelli, R Dallai, GK Roderick and F Frati. Population structure and colonization history of the olive fly, *Bactrocera oleae* (Diptera, Tephritidae). *Mol. Ecol.* 2005; **14**, 2729-38.
- [46] M Virgilio, H Delatte, T Backeljau and MD Meyer. Macrogeographic population structuring in the cosmopolitan agricultural pest *Bactrocera cucurbitae* (Diptera: Tephritidae). *Mol. Ecol.* 2010; **19**, 2713-24.