

# The oriental fruitfly *Bactrocera dorsalis* s.s. in East Asia: disentangling the different forces promoting the invasion and shaping the genetic make-up of populations

N. Aketarawong · C. R. Guglielmino · N. Karam ·  
M. Falchetto · M. Manni · F. Scolari · L. M. Gomulski ·  
G. Gasperi · A. R. Malacrida

Received: 2 October 2013 / Accepted: 2 May 2014 / Published online: 11 May 2014  
© Springer International Publishing Switzerland 2014

**Abstract** The Oriental fruit fly, *Bactrocera dorsalis sensu stricto*, is one of the most economically destructive pests of fruits and vegetables especially in East Asia. Based on its phytophagous life style, this species dispersed with the diffusion and implementation of agriculture, while globalization allowed it to establish adventive populations in different tropical and subtropical areas of the world. We used nine SSR loci over twelve samples collected across East Asia, i.e. an area that, in relatively few years, has become a theatre of intensive agriculture and a lively fruit trade. Our aim is to disentangle the different forces that have affected the invasion pattern and shaped the genetic make-up of populations of this fruit fly. Our data suggest that the considered samples probably represent well established populations in terms of genetic variability and population structuring. The human influence on the genetic shape of populations and diffusion is evident, but factors such as breeding/habitat size and life history traits of the species may have determined the post introduction phases and expansion. In East Asia the origin of diffusion can most probably be allocated in the oriental coastal provinces of China, from where this fruit fly spread into Southeast

Asia. The spread of this species deserves attention for the development and implementation of risk assessment and control measures.

**Keywords** *Bactrocera dorsalis* s.s. · Oriental fruit fly in East Asia · Invasion patterns · Genetic diversity · Microsatellites · Pest management

## Introduction

Fruit flies within the Tephritidae family include more than 5,000 species, one-third of which develop in fresh fruits and represent important invasive pests (White and Elson-Harris 1992). Given their opportunistic phytophagous lifestyles (Diaz-Fleischer et al. 1999), host plants have a significant influence on their survival and play an important role in their dispersion. In other words, fruit flies dispersed with the diffusion and implementation of agriculture and, as secured passengers of trading goods, took further advantage from increasing human mobility and trading activities (Malacrida et al. 2007). Therefore, the global invasion of tephritid flies is emerging as the combined result of both natural and human processes. Examples are provided by pest species such as the Medfly *Ceratitidis capitata* (Malacrida et al. 2007) and *Bactrocera* species such as *Bactrocera cucurbitae* (Virgilio et al. 2010) and taxa pertaining to the *Bactrocera dorsalis* complex. The latter species have natural distributions confined to oriental regions, but are becoming worldwide invasive pests (White and Elson-Harris 1992; Aketarawong et al. 2007; Khamis et al. 2009; Krosch et al. 2012). Within this species complex, the Oriental fruit fly, *B. dorsalis sensu stricto*, is one of the most economically destructive species for fruits and vegetables across East Asia and Southeast Asia (Drew and

---

N. Aketarawong, C. R. Guglielmino have equally contributed.

---

N. Aketarawong · C. R. Guglielmino · N. Karam ·  
M. Falchetto · M. Manni · F. Scolari · L. M. Gomulski ·  
G. Gasperi · A. R. Malacrida (✉)  
Department of Biology and Biotechnology, University of Pavia,  
Via Ferrata 9, 27100 Pavia, Italy  
e-mail: malacrid@unipv.it

### Present Address:

N. Aketarawong  
Department of Biotechnology, Faculty of Science, Mahidol  
University, Rama VI Road, Bangkok 10400, Thailand

Hancock 1994; Clarke et al. 2005; Drew et al. 2008). The very broad host range of *B. dorsalis s.s.*, its dispersal capacity and its relatively wide climatic tolerance (Fletcher 1989; Duyck et al. 2004; Liu et al. 2011), have allowed it to establish adventive populations in different tropical and subtropical areas (Stephens et al. 2007). It was first recorded in 1912 as lectotype in Taiwan (Hardy 1973; Drew and Hancock 1994) and in the period 1934–1937 on Hainan Island and on the south eastern coast of China (Xie 1937). In the following years its presence was reported in most countries of the Asia–Pacific region, from South China to the northern areas of the Indian subcontinent (Wang 1996). Outside Asia, the oriental fruit fly has been reported in Hawaii (1945), Guam (1947) and occasionally in California and Florida (1960–1990) (White and Elson-Harris 1992). The spread of this species is receiving considerable attention for the development and implementation of risk assessment and control measures (Biosecurity Australia 2009). As a consequence, this species is mobilizing substantial ecological/demographical and molecular genetic research to assess its potential home place and to interpret its dispersion patterns. At the macro-geographic diffusion level, the analyses of microsatellite markers in 14 populations from Far East Asia, South Asia, Southeast Asia and the Pacific area, indicated a western-oriented migration route from China to the West, with a colonization process associated with a relatively stable population demographic structure of adventive populations (Aketarawong et al. 2007). Within Asia the necessity of implementing control strategies against this pest stimulated studies attempting to validate the *B. dorsalis s. s.* taxonomic status in relation to its sister pest taxa, *B. papayae* and *B. philippinensis*, with which it shared allopatric distribution in the Thai/Malay region (Schutze et al. 2012; Krosch et al. 2012). This area marks the southern diffusion limit of *B. dorsalis s.s.*. In parallel, strong efforts were devoted in the attempt to trace the invasion routes of this pest across and out the China/Southeast Asia regions (Wan et al. 2011, 2012; Shi et al. 2012). Each of these studies provided pictures of its complex colonization pattern. Two main invasion routes were proposed by Wan et al. (2012): one from Southeast China to Central China, another from Southeast China to Southeast Asia. A more extensive geographic sampling (Shi et al. 2012), which included most regions used by Aketarawong et al. (2007) and Wan et al. (2011), provided indirect evidence that the tropical regions of Southeast Asia and the Southern coast of China may be considered as the native range of the species. Moreover, the weak population genetic structure and the absence of isolation by distance effect detected across the considered areas were explained by the presence of human mediated long-distance dispersal events and miscellaneous host selection by this species. But specific approaches are necessary to disentangle the

different forces that affected the invasion pattern and shaped the genetic make up of its populations, in an area such as East Asia, which in relatively few years became a theatre of intensive agriculture and a lively fruit trade.

On this background, here we use microsatellite data to evaluate both large- and small- scale population genetic differentiation and connectivity within and among South East Asia and East Asia regions. Using both classical and Bayesian methods, we attempt to infer the origin and the routes of invasion of *B. dorsalis s.s.* across these areas, in which the propagule pressure is exacerbated and spread over great distances by various intense human activities.

## Materials and methods

### Samples considered

Attention has been centred on 12 *B. dorsalis s.s.* populations from East Asia sampled in 2006–2007 (Table 1). We took advantage of the available SSR genotype dataset previously obtained from five Southeast Asia populations (Thailand, Myanmar, Laos, Cambodia) and from Taiwan (Aketarawong et al. 2007). This dataset has been integrated with six new population samples collected in China and genotyped at the same SSR loci. The six samples from China were collected in the following provinces: Hunan (Hu), Fujian (Fj), Yunnan (Yn), Guangxi (Gx), Guangdong (Gd), and Hainan (Ha). The integration of this new data from China with the available Southeast Asia dataset was aimed at obtaining a genetic portrait of this species along the tropical/subtropical East Asia region. This area offers climatic conditions and continuous host resources, also due to intensive crop cultivation (e.g. mango, guava, and orange), suitable for stable populations of *B. dorsalis s.s.* (Biosecurity Australia 2009). The Chinese samples were obtained as ethanol preserved adults and genomic DNA was extracted individually from each fly using the method of Baruffi et al. (1995).

### Microsatellite analysis

Each individual from the Chinese samples was genotyped at the same nine microsatellite loci (Bd1, 7, 9, 15, 19, 39, 42, 54, and 85B) used for the South East Asia samples by Aketarawong et al. (2006, 2007). We employed the TANDEM program (Matschiner and Salzburger 2009) for the automated binning of microsatellite allele lengths (with additional manual checking) to overcome the problems of genotyping errors. Samples for which amplification was not successful, or scoring was uncertain, were rerun, and re-extraction of DNA was performed if necessary. The genotyping data obtained from the Chinese samples were

**Table 1** Collection data of *Bactrocera dorsalis* s.s. samples used in this study

Country	Province	Population abbreviation	Year	Latitude	Longitude	Sample size
China	Hunan (Aanhua)	China-Hu	2007	27°35'N	111°38'E	20
	Fujian (Xiamen)	China-Fj	2007	24°04'N	117°55'E	20
	Yunnan (Kunming)	China-Yn	2007	25°02'N	102°42'E	20
	Guangxi (Nanning)	China-Gx	2007	23°48'N	108°46'E	20
	Guangdong (Guangzhou)	China-Gd	2007	23°07'N	113°15'E	20
	Hainan (Wenchang)	China-Ha	2007	20°10'N	110°20'E	20
Taiwan	Kaohsiung	Taiwan	2006	23°02'N	120°35'E	18
Thailand	Udon Thani	Thailand-U	2006	17°25'N	102°51'E	18
	Ratchaburi	Thailand-R	2004	13°53'N	99°48'E	18
Myanmar	Taungoo	Myanmar	2006	18°56'N	96°43'E	18
Laos	Vientiane	Laos	2006	17°58'N	102°36'E	18
Cambodia	Phanom Phenh	Cambodia	2006	12°56'N	103°46'E	14

integrated into the dataset from the Southeast Asian samples.

#### Sample variability estimates

Genetic variability in each of the 12 populations was estimated in terms of number of alleles ( $n_a$ ), effective number of alleles ( $n_e$ ), variance of allele size ( $V_m$ ), and observed and expected heterozygosities ( $H_O$  and  $H_E$ , respectively) using MICRO-SATELLITE ANALYSER (MSA) V.4.05 (Dieringer and Schlötterer 2003). Allele richness ( $R_s$ ), gene diversity ( $H_S$ ) and the inbreeding index ( $F_{IS}$ ) were obtained using FSTAT V.2.9.3.2 (Goudet 2001). The frequency of null alleles ( $A_n$ ) was estimated following Brookfield (1996). Fisher's exact test, subprogram in GENEPOP V.4 (Rousset 2008), was used to test the departures from Hardy–Weinberg equilibrium at each locus/population combination. The same program was also used to examine the linkage disequilibrium between pairs of loci in each population (100 batches, 1,000 iterations per batch).

#### Population structure and demography analysis

The degree of genetic differentiation among populations was quantified by calculating pairwise- $F_{ST}$  values (Weir and Cockerham 1984) in the MSA V.4.05 software (Dieringer and Schlötterer 2003). The statistical significance of each value was assessed by the comparison of the observed value with the values obtained in 10,000 matrix permutations. The relationship between genetic ( $F_{ST}/1-F_{ST}$ ) and geographic distances [Ln distance (Km)] among populations was studied using subprogram ISOLDE in GENEPOP V.4. Principal Coordinate Analysis (PCoA) in the program GENALEX V.6.4 (Peakall and Smouse 2005) was applied to identify the relationships among populations on the basis of their allele frequencies. A plot of the first three principal coordinates

was conducted using subprogram MOD3D in NTSYSp V.2.1 (Rohlf 2005).

A model-based clustering method implemented in the program STRUCTURE V 2.3.2 (Pritchard et al. 2000; Falush et al. 2003, 2007) was used to infer the number of potential genetic clusters ( $K$ ) within the *B. dorsalis* samples. The simulation was performed using an admixture model assuming correlated allele frequencies (the  $F$  model). The  $F$  model allows frequencies in the different populations to be similar, probably because of ongoing migration or shared ancestry. The burn-in was set to 100,000 steps and was followed by 1,000,000 Markov Chain Monte Carlo (MCMC) replications. All runs were repeated 10 times for each number of possible clusters ( $K$ ), set between 1 and 12 (i.e. the number of samples), regardless of a priori information (e.g., sampling location, distance, etc.). To indicate the most likely number of genetic clusters, the natural logarithm of the likelihood of the data,  $\text{Ln } P(X/K)$ , was calculated. It is expected to be high with a low variance for the real cluster (Pritchard et al. 2000). We also implemented the Delta  $K$  method (Evanno et al. 2005) which uses the second-order rate of change of  $\text{Ln } P(X/K)$  given the data. This method is based on the rate of likelihood change between successive  $K$  values showing the best estimate of  $K$ .

The genetic distance based on the proportion of shared alleles ( $D_S$ ) (Bowcock et al. 1994) and Nei's genetic distance ( $D_A$ ) (Nei et al. 1983) were calculated using the MSA V.4.05 (Dieringer and Schlötterer 2003). Neighbour-joining trees were reconstructed based on  $D_S$  and  $D_A$ , after 1,000 bootstrap re-samplings of the original data, using the PHYLIP package (Felsenstein 2005).

Hierarchical analysis of molecular variance (AMOVA) was tested to study the genetic discontinuities among geographical areas. Populations were partitioned into five scenarios as shown in the results section. The partition of

genetic variation among and within regions and within populations (quantified through the sum of square size) were tested through the nonparametric permutation procedures of 10,000 iterations using ARLEQUIN V.3.0 (Excoffier et al. 2005).

Inference of population demography was analysed through the program GENECLASS V.2.0 (Piry et al. 2004). The former was used to assign or exclude reference populations as possible origins of individuals, on the basis of multilocus genotypes. The program calculates, for each individual of a population, the probability of belonging to each other reference population or to be a resident in the population where it was sampled. The standard criterion described by Rannala and Mountain (1997), which applies Bayesian statistics to compute probabilities, was used. The additional parameters chosen included (1) the simulation algorithm for population assignment described in Paetkau et al. (2004), (2) simulation of 10,000 genotypes for each population and (3) an arbitrary threshold probability value of 0.01 to determine the origin. In order to obtain detailed information of the routes of the invasion and colonization histories we designed four sets of evolutionary scenarios and we analysed them with ABC methods using the DIYABC v.2.0.3. program (Cornuet et al. 2014). We used prior distribution of demographic parameters as described in Table 2. The definition of scenarios was chosen on the basis of the scattered historical information and on the results of cluster analysis. The estimation of the time of events (in number of generations back in time), was based on the consideration that *B. dorsalis s.s.* can support from five to ten generations/year in most tropical areas, but less than four generations/year in most subtropical areas (Fletcher 1989; Shi et al. 2005; Ye and Liu, 2005; Liu et al. 2011). The current effective population size was assumed the same for all the sampled populations using a uniform distribution bounded between 10 and 100,000 diploid individuals. During colonization, the number of founder individuals for each colonization event was described as  $N_b$  and its size was drawn from a uniform distribution bounded between 2 and 500 individuals. Moreover, the duration of bottleneck ( $db$ ) was assumed to occur after the introduction events. The introduced populations might take several generations to establish their population size. Therefore, this parameter was set from one to ten generations. A total of 11 scenarios were considered (in four tests) to test the hypothesis that the invasion routes of *B. dorsalis s.s.* originated in South East China (i.e., China-Fj) and propagated to South East Asia (i.e., Thailand-U, Thailand-R, Cambodia, and Laos). For all the 11 scenarios we simulated  $10^6$  data/scenario. To compute posterior probabilities of the competing scenarios, we used the 1 % of the simulated datasets closest to the observed data to estimate the relative posterior probability (with 95 % CI) of each

**Table 2** Definition and prior distributions of parameters used in the definition of the 11 tested scenarios in ABC method

Parameter	Interval
<i>Effective population size</i>	
$N_i$ ( $i = 1, 2, \dots, 6$ )	Uniform [10–100,000]
<i>Number of founder for colonization event</i>	
$N_b$ ( $i = 1, 2, \dots, 6$ )	Uniform [2–500]
<i>Time of events (in generations back in time)</i>	
$t_1$	Uniform [1–100]
$t_2, t_3, t_4$	Uniform [100–1,000]
<i>Duration of bottleneck (db)</i>	
Mean mutation rate	Uniform [ $10^{-4}$ – $10^{-3}$ ]
Locus mutation rate	Gamma [ $10^{-5}$ – $10^{-2}$ , 2]
Mean coefficient P	Uniform [0.1–0.3]
Locus coefficient P	Gamma [0.01–0.90, 2]
Mean SNI rate	Log-Uniform [ $10^{-8}$ – $10^{-5}$ ]
Locus SNI rate	Gamma [ $10^{-9}$ – $10^{-4}$ , 2]

scenario with a logistic regression (Cornuet et al. 2014). For each set of scenarios, the most likely scenario was the one with the highest posterior probability value and non-overlapping 95 % confidence intervals. Confidence in scenario choice was assessed by evaluating Type I and Type II error rates, following the method described in Cornuet et al. (2010).

## Results

### Intrapopulation genetic diversity

Overall the considered nine microsatellite loci, detected in 224 flies across the 12 localities, appear to be very informative with a mean Polymorphic Information Content (PIC) estimate of 0.745. Indeed, these loci display high levels of polymorphism, both in terms of number of alleles (from 15 to 26) and allele size range: Bd1: 18 alleles (87–107 bp); Bd7: 15 alleles (76–126 bp); Bd9: 26 alleles (118–184); Bd15: 17 alleles (163–189 bp); Bd19: 19 alleles (142–175 bp); Bd39: 16 alleles (70–89 bp); Bd42: 21 alleles (147–188 bp); Bd54: 20 alleles (91–122 bp); Bd85B: 20 alleles (61–86 bp). Results of Fisher's exact test with the sequential Bonferroni correction (Rice 1989) revealed that the majority of populations conformed to Hardy–Weinberg equilibrium (HWE) at most loci. However, the locus/populations combinations that were not in HWE were not concentrated at any locus or in any population. Significant linkage disequilibrium was not detected between genotypes at the nine loci.

The overall genetic variability for each population, is summarized in Table 3. China (especially China-Fj) and

**Table 3** Genetic variability estimates in field-collected samples of Oriental fruit fly *Bactrocera dorsalis* s.s. from East Asia

Population	$n_a$	$n_r$	$n_p$	$A_p$	$R_S \pm SD$	$H_S \pm SD$	$V_m$	$n_e$	$H_O$	$H_E$	$F_{IS}$	$A_n$
China-Hu	6.44	15	1	0.13	5.40 ± 1.92	0.68 ± 0.22	1.84	4.00	0.44	0.67	0.36	0.14
China-Fj	7.44	19	9	0.06 ± 0.04	6.31 ± 1.51	0.75 ± 0.11	3.63	4.59	0.52	0.74	0.31	0.13
China-Yn	6.33	16	5	0.05 ± 0.04	5.29 ± 1.14	0.70 ± 0.09	1.27	3.46	0.52	0.69	0.25	0.10
China-Gx	7.33	20	1	0.03	5.95 ± 2.19	0.75 ± 0.15	2.97	5.05	0.55	0.74	0.27	0.11
China-Gd	6.11	11	2	0.21 ± 0.27	5.16 ± 1.73	0.66 ± 0.19	1.59	3.49	0.50	0.65	0.25	0.10
China-Ha	7.22	15	3	0.03 ± 0.00	5.94 ± 1.98	0.74 ± 0.12	2.34	4.50	0.56	0.74	0.25	0.10
Taiwan	6.33	19	4	0.04 ± 0.02	5.13 ± 2.16	0.61 ± 0.21	1.83	3.36	0.50	0.61	0.18	0.07
Thailand-U	7.89	17	8	0.04 ± 0.01	6.53 ± 2.03	0.74 ± 0.17	3.94	4.97	0.46	0.73	0.37	0.15
Thailand-R	8.33	24	13	0.03 ± 0.02	6.74 ± 2.02	0.76 ± 0.13	3.25	5.40	0.46	0.75	0.39	0.17
Myanmar	5.56	16	3	0.03 ± 0.00	4.56 ± 1.23	0.59 ± 0.17	1.81	2.83	0.42	0.59	0.29	0.11
Laos	6.44	12	3	0.03 ± 0.00	5.58 ± 1.46	0.75 ± 0.11	2.00	4.45	0.40	0.74	0.47	0.20
Cambodia	6.44	20	3	0.07 ± 0.04	5.74 ± 2.00	0.72 ± 0.13	1.75	4.12	0.52	0.71	0.27	0.11

$n_a$  average number of alleles,  $n_r$  number of rare alleles (frequency < 0.05),  $n_p$  number of private alleles,  $A_p$  average frequency of private alleles,  $R_S$  allele richness and standard deviation,  $H_S$  average gene diversity and standard deviation,  $V_m$  median variance in repeat numbers,  $n_e$  effective number of alleles [ $1/(1 - H_E)$ ],  $H_O$  average observed heterozygosity,  $H_E$  average expected heterozygosity,  $F_{IS}$  average inbreeding index,  $A_n$  frequency of null alleles [ $(H_E - H_O)/(H_E + 1)$ ] (Brookfield 1996)

China-Gx) and Thailand (especially Thailand-R) present relatively high levels of genetic variability on the basis of the number of alleles ( $n_a$ ), allelic richness ( $R_S$ ), median variance of allele size ( $V_m$ ), and expected heterozygosity ( $H_E$ ). Rare alleles (frequency less than 0.05) and private alleles were detected in all populations ranging from 11 to 24 and from 1 to 13, respectively. Within China, the China-Fj population showed the highest number of private alleles (9) with a relatively low average frequency (0.06 ± 0.04). The lowest number of private alleles (1) was observed in China-Hu and China-Gx populations, although the China-Hu private allele displays a quite high frequency (0.13). China-Gd, with only two private alleles, is the population with the highest frequency of private alleles (0.21 ± 0.27). Outside China, the highest number of private alleles (13) was observed in Thailand-R and the lowest number was detected in Myanmar, Laos and Cambodia (3), with relatively low frequencies (0.03–0.07). A deficiency in the average level of observed heterozygosity was detected in several populations. This observation may be related to the high number of rare alleles present in almost all populations, although it could be possibly associated with a Wahlund effect and/or to the presence of null alleles. The frequency of null alleles was estimated in all populations, and ranges from 0.07 to 0.20 (Table 3).

#### Population structure

Genetic differentiation among populations was measured by the fixation index  $F_{ST}$  (Table 4). The pairwise  $F_{ST}$  values among the sampled populations range from 0.000 (Thailand-U and Laos) to 0.272 (China-Gd and Myanmar). For several pairs of populations,  $F_{ST}$  values are not significantly different

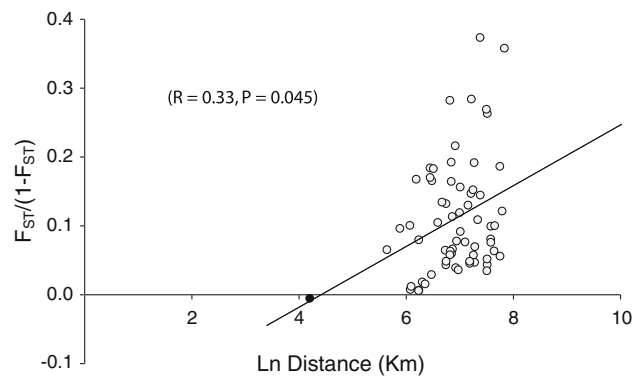
from zero. It is noteworthy that the Chinese island sample (China-Ha) displays non-significant estimates not only with two southern China samples, China-Fj and China-Gx, but also with Cambodia. This last population, in turn, shows non-significant differentiation with the other south-eastern populations. An additional case of non-significant differentiation is between the eastern China region (China-Fj) and Laos ( $F_{ST} = 0.034$ ). The most differentiated China samples, are China-Gd, China-Yn and the Taiwan with all  $F_{ST}$  significantly different from zero. Myanmar significantly differs from all the other populations, with the highest  $F_{ST}$  value with China-Gd (0.272) and the lowest with China-Yn ( $F_{ST} = 0.060$ ). The analysis of isolation by distance (IBD) results in a significant correlation ( $R = 0.33$ ,  $P = 0.045$ ) between geographical and genetic distances (Fig. 1). However, it was found that the relationship of the Thailand-U and Laos populations might have an impact on the correlation, as they present both the closest geographic distance and the smallest genetic distance ( $F_{ST} = 0.000$ ). If either Laos or Thailand-U, or both of them are excluded from the analysis, the correlation between geographical distance and genetic distance becomes non-significant ( $R_{No\ Laos} = 0.30$ ,  $P = 0.084$ ;  $R_{No\ Thailand-U} = 0.30$ ,  $P = 0.084$ ;  $R_{No\ Laos\ and\ Thailand-U} = 0.31$ ,  $P = 0.110$ ).

Principal Coordinate Analysis (PCoA) was performed as shown in Fig. 2. The first three axes of the PCoA explained a relatively high amount of the genetic variation (79 %). The first axis (44 %) separates China-Yn and Myanmar from the other populations. The second axis (20 %) mainly differentiates the remaining populations into two clusters: all Chinese samples versus Taiwan and Southeast Asia. In this representation the isolated position of China-Gd with respect to the third axis (15 % variation) is interesting.

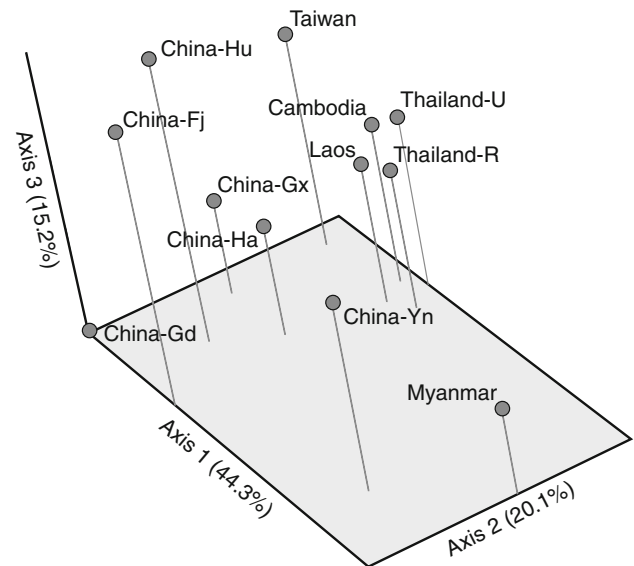
**Table 4** Pairwise- $F_{ST}$  values among 12 population samples of *B. dorsalis* s.s. as derived from MICROSATELLITE ANALYSER (Dieringer and Schlötterer 2003)

$F_{ST}$ -values:	China-Hu	China-Fj	China-Yn	China-Gx	China-Gd	China-Ha	Taiwan	Thailand-U	Thailand-R	Myanmar	Laos	Cambodia
China-Hu												
China-Fj	0.029 <sup>ns</sup>											
China-Yn	0.141	0.098										
China-Gx	0.074	0.062	0.156									
China-Gd	0.144	0.095	0.220	0.062								
China-Ha	0.061	0.038 <sup>ns</sup>	0.102	0.008 <sup>ns</sup>	0.088							
Taiwan	0.072	0.092	0.208	0.071	0.162	0.084						
Thailand-U	0.045	0.042	0.117	0.057	0.135	0.042	0.075					
Thailand-R	0.071	0.053	0.115	0.065	0.127	0.044	0.108	0.007 <sup>ns</sup>				
Myanmar	0.212	0.157	0.060	0.221	0.272	0.161	0.264	0.155	0.142			
Laos	0.055	0.034 <sup>ns</sup>	0.119	0.055	0.107	0.046	0.091	0.000 <sup>ns</sup>	0.018 <sup>ns</sup>	0.146		
Cambodia	0.049	0.060	0.128	0.046	0.132	0.035 <sup>ns</sup>	0.091	0.006 <sup>ns</sup>	0.012 <sup>ns</sup>	0.178	0.016 <sup>ns</sup>	

Some of the values are not significantly different from zero at  $P > 0.050$  (ns)

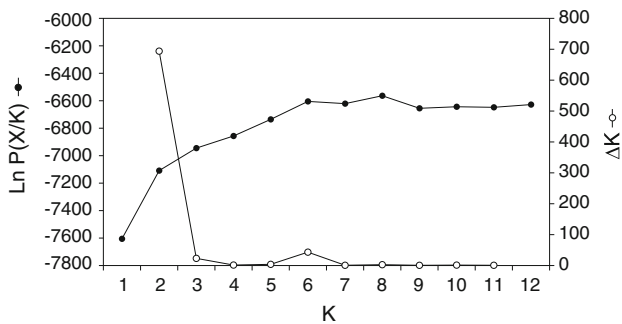


**Fig. 1** Regression of  $F_{ST}$  values ( $F_{ST}/(1 - F_{ST})$ ) on geographic distances (Ln distance in Km) among the 12 East Asian samples of *Bactrocera dorsalis* s.s. The highlighted point represents the relationship between Laos and Thailand-U. See text for discussion



**Fig. 2** Three-dimensional plot of principal coordinate analysis (PCoA) based on similarity matrix derived from *B. dorsalis* s.s. microsatellites data

The population structure of the 12 populations can be described in more detail when individual genotypes are considered using the Bayesian cluster analysis. In the STRUCTURE analysis the ancestry of each individual is referred to one of the likely hypothetical original populations defined as cluster ( $K$ ). The Evanno et al. (2005) method indicated 2 as an optimal  $K$  value (Fig. 3), as Delta  $K$  dropped sharply with  $K = 3$  and climbed slightly with  $K = 6$ . At  $K = 2$ , cluster 1 comprises China-Yn and Myanmar and cluster 2 all the other ten populations (Table 5); the same population subdivision is seen along the first principal axis of PCoA. In Fig. 3 the number of most probable  $K$  higher than the optimum of 2 is 6, which also corresponds to the number of clusters indicated by the



**Fig. 3** The two curves indicate the most likely number of genetic clusters ( $K$ ) of *B. dorsalis* s.s. according to the natural logarithm of the likelihood of the data  $\text{Ln } P(X/K)$  (Pritchard et al. 2000) and according to the Delta  $K$  method (Evanno et al. 2005)

plateau of  $\text{Ln } P(X/K)$  (Pritchard et al. 2000; Falush et al. 2003). Thus the genotype of each fly was assigned, entirely or in part, to each of six clusters with a certain probability value (Table 5). The Chinese ancestry is fragmented in all six clusters. Considering the highest proportion of membership, two clusters are centred in China, i.e., the first with  $Q_{\text{China-Gd}} = 0.865$  and the second with  $Q_{\text{China-Hu}} = 0.508$  and  $Q_{\text{China-Fj}} = 0.453$ . The third and fourth clusters are centred in Southeast Asia and share coancestry with China. The fifth cluster is mainly represented by Taiwan ( $Q_{\text{Taiwan}} = 0.705$ ) which shares coancestry with China (especially with China-Ha and China-Gx). In cluster 6 the coancestry is shared only between China-Yn and Myanmar.

The Neighbour-joining phylogenetic trees reconstructed from the proportion of shared alleles ( $D_S$ ) and from Nei’s distance (data not shown), provided similar topologies. The branching of the trees was consistent with the population

coancestry relationships of STRUCTURE, as shown in Fig. 4. The pie charts are related to the partitioning into six groups according to STRUCTURE results.

Tests of homogeneity among the populations were performed using AMOVA. Populations were partitioned following five criteria (Table 6). Significant differences among groups were observed only when groups were those corresponding to the STRUCTURE and tree results (cases 3, 4, and 5).

Demographic inference

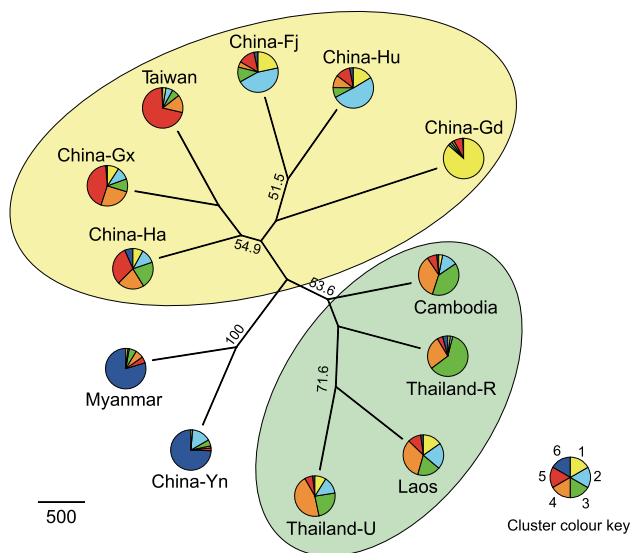
Assignment test

The proportion of assignment rates, estimated using GENECLASS 2.0, are presented in Table 7. In diagonal, values indicate the average probability with which individuals were assigned to the corresponding reference population. The probability values of self-assignment ranged from 0.914 for China-Fj to 0.519 for Laos. The highest average assignment value from one population to the others (column values) is for the eastern region of China, China-Fj with an overall contribution of  $0.368 \pm 0.137$ . On the other hand, migration rate from other populations into a population (row values) has the highest value in Taiwan island ( $0.162 \pm 0.157$ ). China-Fj shares its genomes with all populations, especially with China-Hu (0.619). Likewise, Thailand-R mainly shares its genetic composition with Southeast Asian populations. Asymmetric assignments with a predominant unidirectional migration from one to other populations, were detected in two Chinese populations, China-Fj and China-Gx, and Thailand-R.

**Table 5** Average coefficient of ancestry obtained from a STRUCTURE run with  $K = 2$  and  $K = 6$  for the 224 individuals of *B. dorsalis* s.s. from the 12 regions

Populations	$K = 2$		$K = 6$					
	1	2	1	2	3	4	5	6
China-Hu	<b><u>0.962</u></b>	0.038	<b>0.166</b>	<b><u>0.508</u></b>	0.079	<b>0.101</b>	<b>0.117</b>	0.029
China-Fj	<b><u>0.703</u></b>	<b>0.297</b>	<b>0.219</b>	<b><u>0.453</u></b>	<b>0.119</b>	0.045	<b>0.133</b>	0.032
China-Yn	0.091	<b><u>0.909</u></b>	0.016	<b>0.151</b>	0.052	0.017	0.022	<b><u>0.742</u></b>
China-Gx	<b><u>0.982</u></b>	0.018	0.094	<b>0.101</b>	<b>0.103</b>	<b>0.254</b>	<b><u>0.436</u></b>	0.013
China-Gd	<b><u>0.988</u></b>	0.012	<b><u>0.865</u></b>	0.019	0.019	0.019	0.071	0.008
China-Ha	<b><u>0.908</u></b>	0.092	0.083	<b>0.116</b>	<b>0.214</b>	<b>0.213</b>	<b><u>0.309</u></b>	0.065
Taiwan	<b><u>0.989</u></b>	0.011	0.025	0.057	0.065	<b>0.140</b>	<b><u>0.705</u></b>	0.008
Thailand-U	<b><u>0.906</u></b>	0.094	0.090	<b>0.137</b>	<b>0.241</b>	<b><u>0.452</u></b>	0.062	0.019
Thailand-R	<b><u>0.866</u></b>	<b>0.134</b>	0.021	0.019	<b><u>0.608</u></b>	<b>0.268</b>	0.044	0.040
Myanmar	<b>0.171</b>	<b><u>0.829</u></b>	0.017	0.011	0.060	0.066	0.049	<b><u>0.797</u></b>
Laos	<b><u>0.892</u></b>	<b>0.108</b>	<b>0.153</b>	<b>0.209</b>	<b>0.183</b>	<b><u>0.330</u></b>	<b>0.104</b>	0.022
Cambodia	<b><u>0.981</u></b>	0.019	0.033	<b>0.124</b>	<b><u>0.395</u></b>	<b>0.355</b>	0.078	0.015

Coancestry higher than 10 % of each population in a cluster is in bold. The highest value of coancestry of each population is underlined



**Fig. 4** Neighbour-joining unrooted tree based on genetic distances derived from the proportion of shared alleles (Bowcock et al. 1994). Numbers at each node indicate the bootstrap values after 1,000 replicates. Only values above 50 % are shown. The six coloured segments in the pies represent the coancestry distribution of the 224 flies in six hypothetical clusters

Unravelling the colonization scenario of *B. dorsalis s.s.* using ABC method

In order to unravel the routes of the invasion and colonization, four sets of evolutionary scenarios are presented. As shown in Table 8, in the four tests, a single scenario clearly dominated the others in terms of posterior probability. In each test, the most likely scenario is the one starting from the China-Fj, from which founder events lead to the formation of the South East Asia populations. Cambodia population (tests 1 and 2) is formed by introduction from

China-Fj through China-Ha and/or through Laos. A higher value of posterior probability is shown in test 3, where China-Fj is connected to Thailand-U through Cambodia. In the test 4, seven populations are taken into account to obtain an overall picture. The best scenario is that shown in Fig. 5, in which starting from China-Fj, China-Ha is a bridge to South East Asian populations.

We found that divergence times between populations, estimated by the median values of the posterior distribution, are reasonably congruent with the historical records. If *B. dorsalis s.s.* in sampled areas can complete ten generations per year, China-Ha firstly diverted from China-Fj during 1918–1920. In the following approximately 20 years, the China-Ha population had become established. Regarding the estimated time period, it is noteworthy that the presence of this fly had been recorded on this island in 1934–1937 (Xie 1937). In a relatively short time (between 671 and 679 generations ago) the colonization of South East Asia began (i.e., Thailand-R Thailand-U and Cambodia). Laos appears to have been recently colonized from Thailand-U (1999–2001). Type-II error rates of all the most probable scenarios are found to be low (ranging from 4.6 to 7.6 %) of wrong assignment. Likewise, type-I error rates are reasonable with low values (ranging from 3.6 to 7.6 %) (Table 8).

Discussion

An important key to explain the spatial heterogeneity in terms of genetic variability and population relationships is that *B. dorsalis s.s.* attacks fruits of high economic value, in addition to non agricultural plants (Clarke et al. 2005). The genetic variability of populations can be thus influenced by the presence of intensive crop cultivation, markets and the

**Table 6** Analysis of molecular variance (AMOVA)

Case	Among groups			Among populations within groups			Within populations		
	Va	Percentage	P	Vb	Percentage	P	Vc	Percentage	P
1	0.061	2.65	0.186	0.165	7.14	<0.0001	2.089	90.20	<0.0001
2	0.025	1.09	0.183	0.197	8.52	<0.0001	2.089	90.39	<0.0001
3	0.305	12.15	0.019	0.116	4.62	<0.0001	2.089	83.22	<0.0001
4	0.148	6.41	0.002	0.077	3.34	<0.0001	2.089	90.25	<0.0001
5	0.169	7.26	<0.0001	0.073	3.13	<0.0001	2.089	89.60	<0.0001

Case 1: Country subdivision, six groups: China, Taiwan, Thailand, Myanmar, Laos and Cambodia

Case 2: Geographical subdivision, two groups: Far East Asia versus Southeast Asia

Case 3: Genetic coancestry (STRUCTURE, K = 2): (China-Yn + Myanmar) versus the others

Case 4: Genetic coancestry (STRUCTURE, K = 6): (China-Hu + China-Fj), (China-Gd), (China-Ha + China-Gx + Taiwan), (Thailand-U + Laos), (Thailand-R + Cambodia) and (Yunnan + Myanmar)

Case 5: Genetic relationship (NJ tree), four groups: (China-Hu + China-Fj + China-Gd), (China-Gx + China-Ha + Taiwan), (Thailand-U + Thailand-R + Laos + Cambodia) and (Yunnan + Myanmar)



**Table 7** Average assignment rate (*m*) of individuals into (rows) and from (column) each population as calculated by GENECLASS2 (Piry et al. 2004)

Population	China-Hu	China-Fj	China-Yn	China-Gx	China-Gd	China-Ha	Taiwan	Thailand-U	Thailand-R	Myanmar	Laos	Cambodia	Ave.	SD
China-Hu	<b>0.815</b>	<b>0.619</b>	0.027	<b>0.264</b>	0.027	<b>0.187</b>	0.031	<b>0.139</b>	<b>0.140</b>	0.011	<b>0.118</b>	<b>0.145</b>	<b>0.155</b>	0.173
China-Fj	<b>0.193</b>	<b>0.914</b>	0.028	<b>0.136</b>	0.070	0.075	0.030	0.084	<b>0.107</b>	0.014	0.067	0.075	0.080	0.051
China-Yn	0.017	<b>0.142</b>	<b>0.709</b>	0.001	0.000	0.033	0.002	0.027	0.033	0.064	0.006	0.018	0.031	0.041
China-Gx	<b>0.138</b>	<b>0.396</b>	0.005	<b>0.698</b>	0.011	<b>0.196</b>	0.025	<b>0.127</b>	0.089	0.003	0.019	0.074	0.098	0.118
China-Gd	<b>0.103</b>	<b>0.445</b>	0.003	0.089	<b>0.594</b>	0.075	0.014	0.076	0.054	0.006	0.056	0.036	0.087	0.123
China-Ha	<b>0.157</b>	<b>0.433</b>	0.053	<b>0.261</b>	0.004	<b>0.694</b>	0.030	<b>0.131</b>	<b>0.167</b>	0.016	0.035	<b>0.107</b>	<b>0.127</b>	0.129
Taiwan	<b>0.197</b>	<b>0.498</b>	0.002	<b>0.353</b>	0.001	<b>0.127</b>	<b>0.632</b>	<b>0.263</b>	<b>0.120</b>	0.003	0.087	<b>0.132</b>	<b>0.162</b>	0.157
Thailand-U	<b>0.105</b>	<b>0.365</b>	0.014	<b>0.126</b>	0.004	<b>0.120</b>	0.020	<b>0.594</b>	<b>0.216</b>	0.015	<b>0.188</b>	<b>0.131</b>	<b>0.119</b>	0.109
Thailand-R	0.054	<b>0.258</b>	0.004	0.070	0.001	0.077	0.010	<b>0.163</b>	<b>0.592</b>	0.016	0.059	<b>0.136</b>	0.077	0.080
Myanmar	0.030	<b>0.187</b>	<b>0.412</b>	0.040	0.000	0.076	0.007	0.091	<b>0.164</b>	<b>0.582</b>	0.022	0.067	<b>0.100</b>	0.120
Laos	<b>0.175</b>	<b>0.362</b>	0.023	0.095	0.034	<b>0.118</b>	0.020	<b>0.333</b>	<b>0.227</b>	0.021	<b>0.519</b>	<b>0.155</b>	<b>0.142</b>	0.123
Cambodia	<b>0.201</b>	<b>0.347</b>	0.018	<b>0.137</b>	0.000	<b>0.106</b>	0.023	<b>0.167</b>	<b>0.217</b>	0.007	0.062	<b>0.654</b>	<b>0.117</b>	0.110
Ave.	<b>0.124</b>	<b>0.368</b>	0.054	<b>0.143</b>	0.014	<b>0.108</b>	0.019	<b>0.146</b>	<b>0.139</b>	0.016	0.065	0.098		
SD	0.068	0.137	0.120	0.102	0.022	0.049	0.010	0.087	0.066	0.017	0.052	0.046		

The proportion of individuals assigned to the source population is underlined. Values of *m* above 0.100 are in bold

use of pesticides, while the fly’s diffusion, and consequently the population structure, may be highly influenced/shaped by the intensity of human trade and mobility.

Genetic diversity and breeding habitat

A consistent level of genetic variability is observed both in China and in Southeast Asia (in Thailand) where the majority of the sampled localities are characterized by the presence of extensive cultivations of host fruits, rich markets, and an overall favourable climate for population growth and expansion. Examples are the Fujian sample from China and Ratchaburi in Thailand. The Fujian area has a highly developed fruit crop industry and, as an area of fresh fruit export, it ranks first in China with excellent enterprises for national and international markets with convenient transportation to Southeast Asia and other Chinese areas ([http://www.unapcaem.org/Activities%20Files/A22/p50\\_FruitChina.pdf](http://www.unapcaem.org/Activities%20Files/A22/p50_FruitChina.pdf)). Fujian ranks among the advanced regions for its contribution to the development of an environmentally friendly fruit market with strong limitations on insecticide use. Ratchaburi province in Thailand is also one of the biggest tropical fruit growing and trading areas in Southeast Asia (Aketarawong et al. 2007). But also the Guangdong sample, which was collected in one of the most cultivated provinces of China (Li 2009) that belongs to the main Chinese trade routes for fruit import/export (Biosecurity Australia 2009), would be expected to be highly variable. On the contrary, the sample from this province, China-Gd, has a relatively low level of genetic variability and appears to be genetically isolated from the geographically close eastern group of Chinese populations (see *F<sub>ST</sub>*, PCoA, STRUCTURE). But the only two private alleles detected in this population have very high frequencies. It is noteworthy that a high level of pesticide resistance (Jin et al. 2011) is registered in this region resulting in selection consequent to pesticide exposure. Therefore, the genetic peculiarity of the China-Gd sample may be the result of the intensive use of pesticides in Guangdong.

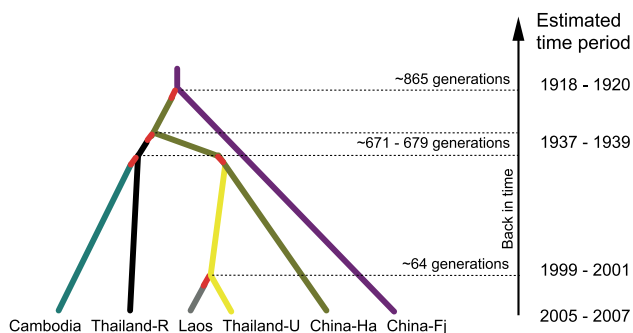
Dynamics of *B. dorsalis* s.s. in East Asia

The neighbour joining tree, PCoA, STRUCTURE and ABC analyses clearly show that genomes are shared between China and Southeast Asia, and there is also evidence that there is no isolation by distance between the populations from these two areas. To highlight the genetic connection between China and Southeast Asia and possibly to infer the direction of migration, we combined the results of two specific genetic analyses, i.e. *F<sub>ST</sub>* and GENECLASS. In Fig. 6, which represents the population sampling sites, the lines connecting the populations are based on non-

**Table 8** Type-I and Type-II error rates for simulated data

Test	Scenario	True scenario	Type-I error	Type-II error	Posterior probability (95 % CI)
1	<b>China-Fj → China-Ha → Cambodia</b>	<b>1</b>	<b>0.048</b>	<b>0.048</b>	<b>0.6707 (0.6391, 0.7024)</b>
	Cambodia → China-Ha → China-Fj	2	0.048	0.048	0.3293 (0.2976, 0.3609)
2	<b>China-Fj → Laos → Cambodia</b>	<b>1</b>	<b>0.036</b>	<b>0.064</b>	<b>0.6196 (0.5306, 0.7085)</b>
	Laos → Cambodia → China-Fj	2	0.072	0.062	0.1554 (0.0768, 0.2339)
	Cambodia → Laos → China-Fj	3	0.072	0.054	0.2251 (0.1585, 0.2917)
3	<b>China-Fj → Cambodia → Thailand-U</b>	<b>1</b>	<b>0.046</b>	<b>0.046</b>	<b>0.8133 (0.7360, 0.8907)</b>
	Thailand-U → Cambodia → China-Fj	2	0.050	0.050	0.1867 (0.1093, 0.2640)
4	China-Fj → China-Ha; China-Ha → Cambodia; China-Ha → Thailand-R; China-Ha → Laos; China-Ha → Thailand-U	1	0.082	0.126	0.0006 (0.0000, 0.0064)
	<b>China-Fj → China-Ha; China-Ha → Thailand-R; Thailand-R → Cambodia; China-Ha → Thailand-U; Thailand-U → Laos</b>	<b>2</b>	<b>0.076</b>	<b>0.076</b>	<b>0.8740 (0.8333, 0.9148)</b>
	Thailand-R → China-Fj; China-Fj → China-Ha; China-Ha → Cambodia; China-Ha → Laos; China-Ha → Thailand-U	3	0.152	0.138	0.0001 (0.0000, 0.0060)
	Thailand-R → China-Fj; China-Fj → China-Ha; China-Ha → Cambodia; China-Ha → Thailand-U; Thailand-U → Laos	4	0.150	0.120	0.1252 (0.0845, 0.1659)

The error rates are computed using the “confidence in model scenario” function, in the DIYABC program. The scenarios in bold font are those with the highest posterior probability values

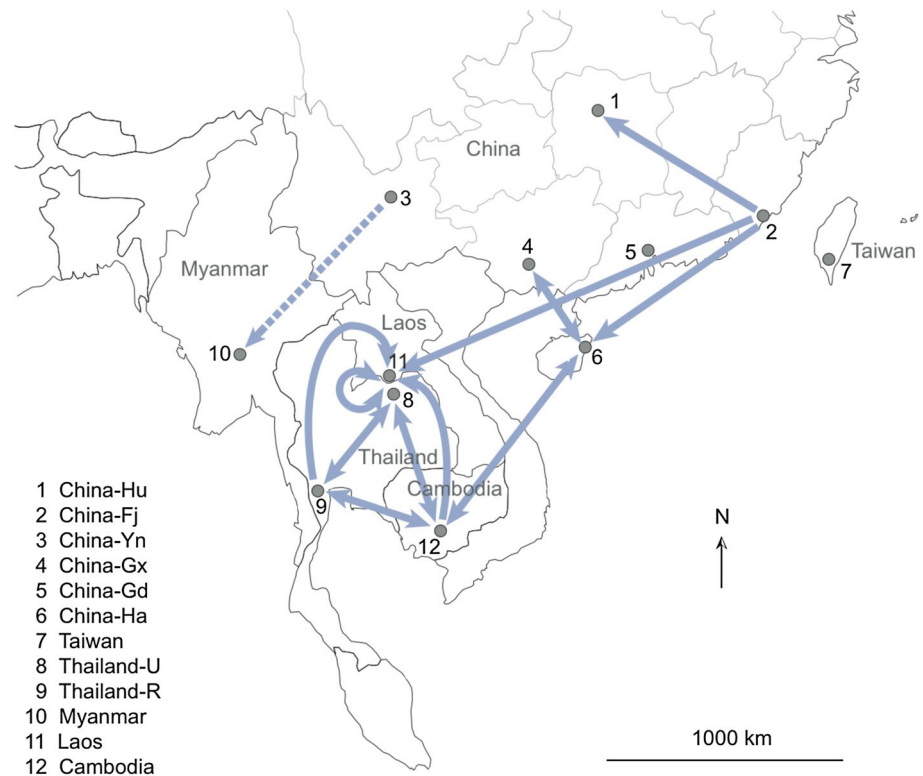


**Fig. 5** Most-likely scenario of test 4 using Approximate Bayesian Computational (ABC) method. The Y-axis indicates the estimated time of event (not to scale). The details of this scenario are described in Table 8

significant  $F_{ST}$  values, while the direction of each arrow is deduced from the migration rates estimated by GENECLASS. This picture may represent the outcome of propagule pressure, i.e. the number of introduced individuals and/or the number of independent introductions, which could have been responsible for the present population structure. Under this perspective the rich fruit growing area of China-Fj appears to have played a crucial role in distributing genomes not only within China, but also in Southeast Asia. A role in this spreading process is played by China-Ha island which appears to be a node towards both

inland China and Southeast Asia. This connection is confirmed also by ABC analyses (Fig. 5) and the coincidence between the estimated date of the founder event (1937–1939) and the date of the first observation in Hainan island (China-Ha) (1934–1937, Xie 1937). Hainan’s moist tropical climate and the intensive plantations of hosts preferred by the fly (Li et al. 2012) allowed, over time, the establishment of large populations of this pest which achieved medium/high levels of genetic variability (Table 3). On the other hand, the intensive trade of fruits to major mainland regions could have provided a portal for the introduction of the fly and for the maintenance of gene flow toward Southeast Asia. Indeed, as shown in the tests 1 and 4 of the ABC analysis, and Fig. 6 the two Chinese samples from Fujian and Hainan are tightly related to Laos and Cambodia: this result allows us to believe that Laos and Cambodia constitute a corridor, through China-Ha island, for the spread of *B. dorsalis s.s.* towards Southeast Asia. In turn, Southeast Asia appears to be an area with a high rate of internal gene flow in which the populations form an homogeneous group characterized by high levels of variability and a distinct molecular variance (Aketarawong et al. 2007). Indeed, in this area there is no geographical isolation of populations and there is intensive fruit crop planting without commercial restriction of tropical fruits among and within Thailand, Laos and Cambodia. Instead, when political events have limited human mobility and the

**Fig. 6** Geographic representation of the dynamics of *B. dorsalis s.s.* in East Asia. The lines connecting the sampling sites are based on non-significant  $F_{ST}$  values, while the direction of each arrow is deduced from the migration rates (greater than 10 %) estimated by GENECLASS. The dashed line connecting Myanmar with China-Yn is related to the lowest (but significant)  $F_{ST}$  value among those related to the two populations. The arrow indicates the high contribution ( $m = 0.412$ ) of China-Yn to Myanmar



exchange of merchandise, the genetic variability of this species is modelled in the form of genetic isolation. In this context, a major separation is represented by the Southeast Asian population of Myanmar which shares alleles only with the closest Chinese population represented by China-Yn, as the China borders are the only commercial partners of Myanmar. Yunnan is one of the provinces highly infested by the fly, but due to the eco-geographic conditions of this area, the habitat of the fly is highly fragmented (Shi et al. 2005). The Yunnan sample was collected in the Kunming area which is an important trade centre serving as transportation hub, linking by road Southwest China to Myanmar. Kunming is a significant horticultural centre located on a high plateau on which, due to the particular geographical and climate conditions, *B. dorsalis s.s.* has only two generations per year, but with considerable genetic overlap due to the continuous arrival of immigrant flies during the summer months (Ye and Liu 2005). These conditions explain both the genetic variability of this fly in this area and the genetic separation of China-Yn sample from the other Chinese samples. Similar conclusions are obtained by mitochondrial DNA analyses (Li et al. 2012).

Taiwan, which in Fig. 6 appears to be isolated, according to the GENECLASS data is characterized by fly immigration mainly from Chinese districts and especially from Fujian. Accordingly, since ancient times, Fujian has cooperated with Taiwan to develop an intensive fruit agriculture.

About the home range of *B. dorsalis s.s.*

When diffusion is mainly due to step by step natural migration, the resulting gradient of ancestral diversity loss along the expansion axes, due to successive founder events, permits the identification of the original home of the species (Malacrida et al. 2007). This is not the case for the oriental fruit fly expansion in East Asia, as the overall picture seems to be masked by a strong propagule pressure driven by the human activities. Indeed none of the considered East Asian populations are markedly genetically depauperate, as the large number of founding members may have been supplemented by continued migration events via commercial traffic.

Nevertheless, specific analyses suggest that the probable place of origin of *B. dorsalis s.s.* in East Asia is in the region of Southeast China. Indeed the STRUCTURE analysis shows that four (clusters 1, 2, 5 and 6) out of six clusters are centred in China. But more evidence derives from ABC and GENECLASS analyses, where the Southeast population China-Fj shares the greatest proportion of its genomes with all the other Chinese and Southeast Asian populations, as shown in Table 7 and depicted in Figs. 5 and 6. Therefore, China-Fj appears to represent in East Asia an ancestral population pertaining to the home range of the species and probably the most ancient among the analysed populations.

## Conclusion

In conclusion, this paper describes the different forces that affect the invasion process in East Asia. It complements and integrates the previous studies, mainly devoted to the inference of the colonization route and the native region of the species (Wan et al. 2011, 2012; Shi et al. 2012). Our data validate our previous results (Aketarawong et al. 2007) and are in agreement with Wan et al. (2011, 2012) and Shi et al. (2012) on indicating Southern East China as a possible source of invasion of this pest. We provide scientific background which clearly indicates the prevalent direction of the expansion pattern in East Asia. In this area, it appears that the genetic variability, the population structuring and dispersal are influenced by humans, but factors such as breeding/habitat size and life history traits of the species have influenced the post-introduction phases and expansion. In this context it is not secondary that *B. dorsalis* s.s. is a species with *K*-strategy traits which may have aided in its establishment and spread (Fletcher 1989) also in saturated habitats. Indeed in all recorded cases, *B. dorsalis* s.s. has invaded over other species in terms of food resources, exploitation and interspecific competition and displacement (Duyck et al. 2004).

From the perspective of management strategies for this pest in East Asia it appears evident that the high genetic interconnectivity and the high migration rates among populations suggest the planning of Area Wide integrated pest management (AW-IPM) strategies (Hendrichs et al. 2007) which take into account the spatial/temporal distribution of the species together with ecological and genetic information. The use of intensive insecticide spraying, in addition to causing environmental damage, may impact the structure of populations by inducing differentiative processes as we observed in the Guandong population. As insecticide resistance could be linked to the genetic differentiation, this may impact the application of chemical control methods.

**Acknowledgments** This research was performed in the framework of FAO/IAEA (Food and Agriculture Organization/International Atomic Energy Agency) Research Agreement No. 16046/RO in favour of A. R. Malacrida. N. Aketarawong was hosted at the Department of Biology and Biotechnology, University of Pavia, thanks to a fellowship within the framework of the Erasmus Mundus Action 2 Programme, MAHEVA Project.

## References

- Aketarawong N, Bonizzoni M, Malacrida AR, Gasperi G, Thanaphum S (2006) Seventeen novel microsatellite markers from an enriched library of the pest species *Bactrocera dorsalis sensu stricto*. Mol Ecol Notes 6:1138–1140
- Aketarawong N, Bonizzoni M, Thanaphum S, Gomulski LM, Gasperi G, Malacrida AR, Guglielmino CR (2007) Inference on the population structure and colonization process of the invasive oriental fruit fly, *Bactrocera dorsalis* (Hendel). Mol Ecol 16:3522–3532
- Australia Biosecurity (2009) Report of the assessment of northern China's fruitfly pest free areas: Hebei, Shandong and Xinjiang. Biosecurity Australia, Canberra
- Baruffi L, Damiani G, Guglielmino CR, Bandi C, Malacrida AR, Gasperi G (1995) Polymorphism within and between populations of *Ceratitis capitata*: comparison between RAPD and multilocus enzyme electrophoresis data. Heredity 74:425–437
- Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455–457
- Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol 5:453–455
- Clarke AR, Armstrong KF, Carmichael AE, Milne JR, Raghu S, Roderick GK, Yeates DK (2005) Invasive phytophagous pests arising through a recent tropical evolutionary radiation: the *Bactrocera dorsalis* complex of fruit flies. Annu Rev Entomol 50:293–319
- Cornuet JM, Ravigné V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). BMC Bioinform 11:401
- Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R et al. (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about populations history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics 1–3. doi:10.1093/bioinformatics/btt1763
- Diaz-Fleischer F, Papaj DR, Prokopy RJ, Norrbom AL, Aluja M (1999) Evolution of fruit fly oviposition behavior. In: Aluja M, Norrbom AL (eds) Fruit flies (Tephritidae): phylogeny and evolution of behavior. CRC Press, Boca Ranton, pp 811–842
- Dieringer D, Schlötterer C (2003) Microsatellite analyzer (MSA): a platform independent analysis tool for large microsatellite data sets. Mol Ecol Notes 3:167–169
- Drew RAI, Hancock DL (1994) The *Bactrocera dorsalis* complex of fruit flies (Diptera: Tephritidae: Dacinae) in Asia. Bull Entomol Res Suppl 2:1–68
- Drew RAI, Raghu S, Halcoop P (2008) Bridging the morphological and biological species concepts: studies on the *Bactrocera dorsalis* (Hendel) complex (Diptera: Tephritidae: Dacinae) in South-East Asia. Biol J Linn Soc 93:217–226
- Duyck PF, David P, Quilici S (2004) A review of relationships between interspecific competition and invasions in fruit flies (Diptera: Tephritidae). Ecol Entomol 29:511–520
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software Structure: a simulation study. Mol Ecol 14:74–75
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1578
- Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7:574–578
- Felsenstein J (2005) PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle
- Fletcher BS (1989) Movements of tephritid fruitflies. In: Robinson AS, Hooper G (eds) Fruit flies: their biology, natural enemies and control, vol 3B., Elsevier ScienceAmsterdam, The Netherlands, pp 209–219

- Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from <http://www2.unil.ch/popen/softwares/fstat.htm>.
- Hardy DE (1973) The fruit flies (Tephritidae-Diptera) of Thailand and bordering countries. *Pac Insects Monogr* 31:1–353
- Hendrichs J, Kenmore P, Robinson AS, Vreysen MJB (2007) Area-wide integrated pest management (AW-IPM): principle, practice and prospects. In: Vreysen MJB, Hendrichs J, Robinson AS (eds) Area-wide control of insect pests: from research to field implementation. Springer, Dordrecht, pp 3–33
- Jin T, Zeng L, Lin Y, Lu Y, Liang G (2011) Insecticide resistance of the oriental fruit fly, *Bactrocera dorsalis* (Hendel) (Diptera: Tephritidae), in mainland China. *Pest Manag Sci* 67:370–376
- Khamis FM, Karam N, Ekesi S, De Meyer M, Bonomi A, Gomulski LM, Scolari F, Gabrieli P, Siciliano P, Masiga D, Kenya EU, Gasperi G, Malacrida AR, Guglielmino CR (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
- Krosch MN, Schutze MK, Armstrong KF, Boontop Y, Boykin LM, Chapman TA, Englezou A, Cameron SL, Clarke AR (2012) Piecing together an integrative taxonomic puzzle: microsatellite, wing shape and aedeagus length analyses of *Bactrocera dorsalis* s.l. (Diptera: Tephritidae) find no evidence of multiple lineages in a proposed contact zone along the Thai/Malay Peninsula. *Syst Entomol* 38:2–13
- Li KW (2009) China's total factor productivity estimates by region, investment sources and ownership. *Econ Syst* 33:213–230
- Li Y, Wu Y, Chen H, Wu J, Li Z (2012) Population structure and colonization of *Bactrocera dorsalis* (Diptera: Tephritidae) in China, inferred from mtDNA COI sequences. *J Appl Entomol* 136:241–251
- Liu JH, Xiong X, Pan Y, Xiong Z, Deng Z, Yang L (2011) Predicting potential distribution of oriental fruit fly, *Bactrocera dorsalis* in Jiangxi Province, South China based on maximum entropy model. *Sci Res Assays* 6:2888–2894
- Malacrida AR, Gomulski LM, Bonizzoni M, Bertin S, Gasperi G, Guglielmino CR (2007) Globalization and fruitfly invasion and expansion: the medfly paradigm. *Genetica* 131:1–9
- Matschiner M, Salzburger W (2009) TANDEM: integrating automated allele binning into genetics and genomics workflows. *Bioinformatics* 25:1982–1983
- Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic tree from molecular data. *J Mol Evol* 19:153–170
- Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol Ecol* 13:55–65
- Peakall R, Smouse PE (2005) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
- Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2: a software for genetic assignment and first-generation migrant detection. *J Hered* 95:536–539
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotypic data. *Genetics* 155:945–959
- Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proc Natl Acad Sci USA* 94:9197–9221
- Rice WR (1989) Analysis tables of statistical tests. *Evolution* 43:223–225
- Rohlf FJ (2005) NTSYSpc numerical taxonomy and multivariate analysis system, Version 2.1. Exeter Software, Applied Biostatistics Inc., New York, USA
- Rousset F (2008) Genepop'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Notes* 4:137–138
- Schutze MK, Krosch MN, Armstrong KF, Chapman TA, Englezou A, Chomic A, Cameron SL, Hailstones D, Clarke AR (2012) Population structure of *Bactrocera dorsalis* s. s., *B. papayae* and *B. philippinensis* (Diptera: Tephritidae) in southeast Asia: evidence for a single species hypothesis using mitochondrial DNA and wing-shape data. *BMC Evol Biol* 12:130
- Shi W, Kerdelhué C, Ye H (2005) Population genetics of the Oriental fruitfly, *Bactrocera dorsalis* (Diptera Tephritidae), in Yunnan (China) based on mitochondrial DNA sequences. *Environ Entomol* 34:977–983
- Shi W, Kerdelhué C, Ye H (2012) Genetic structure and inferences on potential source areas for *Bactrocera dorsalis* (Hendel) based on mitochondrial and microsatellite markers. *PLoS ONE* 7:e37083
- Stephens AE, Kriticos DJ, Leriche A (2007) The current and future potential geographic distribution of the oriental fruit fly *Bactrocera dorsalis* (Diptera Tephritidae). *Bull Entomol Res* 97:369–378
- Virgilio M, Delatte H, Backeljau T, De Meyer M (2010) Macrogeographic population structuring in the cosmopolitan agriculture pest *Bactrocera cucurbitae* (Diptera: Tephritidae). *Mol Ecol* 19:2713–2724
- Wan XW, Nardi F, Zhang B, Liu YH (2011) The oriental fruit fly, *Bactrocera dorsalis*, in China: origin and gradual inland range expansion associated with population growth. *PLoS ONE* 6:e25238
- Wan X, Liu Y, Zhang B (2012) Invasion history of the oriental fruit fly, *Bactrocera dorsalis*, in the Pacific-Asia region: two main invasion routes. *PLoS ONE* 7:e36176
- Wang X (1996) The fruit flies (Diptera: Tephritidae) of the East Asian Region. *Acta Zootaxonomica Sinica* 21:49–54 [In Chinese]
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370
- White IM, Elson-Harris MM (1992) Fruit flies of economic significance: their identification and bionomics. Redwood Press Ltd., Melksham
- Xie YZ (1937) Study on the Trypetidae or fruit-flies of China. *Sinensia* 2:103–226
- Ye H, Liu JH (2005) Population dynamics of oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae), in the Kunming area, southwestern China. *Insect Sci* 12:231–240