Bulletin of Entomological Research (2019) **109**, 713–722
doi:10.1017/S0007485319000026

Genetic structure and range expansion of *Zeugodacus Cucurbitae* (Diptera: Tephritidae) in Africa

H. Delatte[1]*, M. De Meyer[2] and M. Virgilio[2]

[1]CIRAD, UMR PVBMT, F-97410 Saint-Pierre, Réunion, France: [2]Royal Museum for Central Africa, Leuvensesteenweg 13, B-3080 Tervuren, Belgium

Abstract

Hypotheses about the worldwide colonization routes of the melon fly, *Zeugodacus cucurbitae* (Diptera: Tephritidae), are mainly based on sparse historical records. Here we aim at reconstructing the colonization history of the African continent based on an improved description of the population structure of *Z. cucurbitae* and approximate Bayesian analyses. Individuals of *Z. cucurbitae* were sampled in 17 localities from East, West and Central Africa and genotyped at 19 microsatellite markers. Bayesian analyses showed intracontinental population structuring with populations from Uganda diverging from those of Tanzania and populations from Burundi and Kenya showing traces of admixture with West African samples. Approximate Bayesian Computation provided support to the hypothesis of a single introduction *Z. cucurbitae* into East Africa and subsequent expansion to West Africa, each colonization event was followed by a bottleneck that promoted population divergence within Africa. Parameter estimates suggested that these events are roughly compatible with the historical records of *Z. cucurbitae* presence in sub-Saharan Africa (*viz*. 1936 in East Africa and 1999 in West Africa) and allow excluding alternative hypotheses on older or multiple introductions of *Z. cucurbitae*.

Keywords: agricultural pests, Tephritidae, *Zeugodacus*, population structure, colonization history, ABC

(Accepted 8 January 2019; First published online 7 February 2019)

Introduction

Zeugodacus cucurbitae (Coquillett, 1899) (Diptera: Tephritidae), the melon fly (formerly named *Bactrocera cucurbitae*; see (Virgilio *et al*., 2015; San Jose *et al*., 2017)), is a polyphagous agricultural pest predominantly attacking a wide range of Cucurbitaceae but also other plant families. This species, native to Central Asia (Drew & Hancock, 2000) allegedly spread throughout the 20th century to other regions successfully establishing into Central and East Asia, Africa, Oceania, Hawaii and the Islands of the Indian Ocean (White & Elson-Harris, 1992). As most tephritid pests, the melon fly *Z. cucurbitae* causes important economic losses for farmers and traders and reduces the availability of essential dietary components to local populations (Mwatawala *et al*., 2009; Sood *et al*., 2016). *Z. cucurbitae* attacks more than 80 plant species including commercial crops such as pumpkin, cantaloupe, watermelon, squash, gourd and cucumber but also non-cucurbit plants such as tomato, eggplant and soft fruits such as mango, orange, papaya and peach (White & Elson-Harris, 1992). Significant differences in the dietary preferences of *Z. cucurbitae* have been described among populations from different geographic regions (Vayssières *et al*., 2007; De Meyer *et al*., 2015).

The colonization history of this pest has been reconstructed mainly based on historical records that are often sparse and sometimes ambiguous. In Africa, the species was first recorded in the East in Tanzania and Kenya in 1936 and 1937, respectively (http://www.gbif.org// (De Meyer *et al*., 2015)). Whether these dates are representative of the genuine arrival of *Z. cucurbitae* on the African continent or not is however still a matter of debate. The important historical links between the eastern coast of Africa (dominated by the so-called Swahili culture) and the Near East and Indian subcontinent from as early as 100 AD (Gilbert, 2004), and the genetic diversity found in a

*Author for correspondence
Phone: 00262262492735
Fax: 00262262492793
E-mail: delatte@cirad.fr

previous population worldwide analysis of this pest have led some authors to suspect that the pest could have been present in the continent for a much longer time (Virgilio *et al.*, 2010; De Meyer *et al.*, 2015). *Z. cucurbitae* has not been recorded in West Africa before 1999 (http://www.africamuseum.be/fruitfly/AfroAsia.htm), when it was found in the Gambia and Ivory Coast (De Meyer *et al.*, 2015). Over the last decade it has been recognized as a common agricultural pest in a number of other West African countries including Benin, Burkina Faso, Ghana, Guinea, Mali, Senegal and Togo (Vayssières *et al.*, 2007; Vayssières *et al.*, 2008). In 2006, *Z. cucurbitae* was recorded in the Democratic Republic of Congo (Congo DR) as well as several other countries of eastern Africa, including Sudan (2006), Uganda (2009), Ethiopia (2010), Malawi (2010) and Mozambique (2013) (De Meyer *et al.*, 2015).

Several studies using molecular markers have looked at different scales the genetic diversity and structuring patterns of this species in different countries. In a previous study (Virgilio *et al.*, 2010), a first attempt to unravel the worldwide population genetic structure of *Z. cucurbitae* has shown that populations sampled were split into five main population groups distributed over the African continent, Central Asia, East Asia, Hawaii and La Réunion, respectively. Another study, deciphering the genetic differentiation between African and non-African populations demonstrated a common ancestry of the African *Z. cucurbitae* and suggested that the invasive populations of the Indian Ocean islands (La Réunion, Mauritius, Seychelles) were a result of recent an introduction from Africa (Jacquard *et al.*, 2013). Then, two recent studies investigated the fine genetic structure with an integrative approach combining neutral nuclear markers, mitochondrial cytochrome oxidase I gene sequencing and morphometric measurements, of several populations in Southeast Asia considered as native vs. invaded areas in the West Pacific (Boontop *et al.*, 2017*a*, *b*). The authors confirmed the higher genetic diversity in the native range with genetic sub-structured populations and suggested that the Hawaii invasion was due to multiple introductions from mainland Asia. Then, a population genomic study using genome-wide single nucleotide polymorphisms (SNP) investigated the overall structure of the species, including several populations with an important sampling on the Asian/Pacific area and two localities in Africa (Dupuis *et al.*, 2017). This study confirmed more or less previously identified genetic clusters in former studies, with however discrepancies in the mainland and oceanic Southeast Asia confounded due to variability among samplings.

Yet, the genetic relationships between mainland African invasive populations and their possible invasion routes across East and West Africa had not been resolved, leaving open a number of questions about intracontinental population structure and colonization history of *Z. cucurbitae*. The objective of this paper is to provide a more detailed description of the African population structure of *Z. cucurbitae* and to investigate the spatio-temporal dynamics of its African colonization routes. In this respect, a higher number of microsatellite markers, a large sampling coverage in East, West and Central Africa and novel analyses such as the use of Approximate Bayesian Computation (ABC) were used.

Materials and methods

Sampling and DNA extraction

Specimens of *Z. cucurbitae* ($n = 332$) were sampled from 17 African locations (table 1) distributed throughout the whole distribution range of the species on the African continent (De Meyer *et al.*, 2015). Field sampling was made by baiting adults using traps with either male lures (cue-lure or methyl eugenol) or a protein lure (torula yeast).

DNA was extracted from ethanol-preserved adults, preserved in the collections of the Royal Museum for Central Africa (Belgium), via the DNeasy Blood and Tissue Kit (Qiagen) as per the manufacturer's instructions. Individual flies were genotyped using 19 microsatellite loci (B5.2, C3.3, E3.4, E4.3, F1.4, F1.6, F3.2, F3.4, G3.4 and BcCIRC3, BcCIRD3, BcCIRD11, BcCIRE8, BcCIRF3, BcCIRF4, BcCIRG1, BcCIRH7, BcCIRH9, BcCIRH10) developed for *Z. cucurbitae* by Wu *et al.* (2009) and Delatte *et al.* (2010). Primer sequences and protocols for DNA amplification, electrophoresis and allele scoring were performed as described in (2010). Electrophoretic analyses were conducted on an automated ABI Prism 3100 Genetic Analyzer (Applied Biosystem) and an individual was declared not amplified at a locus only after two amplification failures.

Genetic diversity

Microsatellite diversity within populations was estimated using observed (H_{obs}) and Nei's (1987) unbiased expected heterozygosity ($H_{n.b}$) in GENETIX 4.03 (Belkhir *et al.*, 1996) and within-population diversity was also estimated (H_s) (Nei, 1973) using FSTAT (Goudet, 2001). All pairs of loci were tested for linkage disequilibrium using the permutational probability test (10^4 iterations) of GENEPOP 3.4 (Raymond & Rousset, 1995). Single and multilocus Fis were estimated through the fixation index of Weir & Cockerham (1984). Deviations from Hardy–Weinberg equilibrium (HWE) were tested using a two-tailed Fisher's exact test based on Markov-chain randomization (10^3 dememorizations, 10^2 batches, and 10^3 iterations per batch) in GENEPOP. Probability values of repeated comparisons were corrected for Type I errors using the False Discovery Rate (FDR) procedure (Benjamini & Hochberg, 1995). FreeNA (Chapuis & Estoup, 2007) was used to estimate null allele frequencies (for each locus in each population) according to the expectation maximization (EM) algorithm of Dempster *et al.* (1977). Population differentiation was quantified by calculating pairwise F_{st} values (Weir & Cockerham 1984) and verifying their significance through the permutational test implemented in GENETIX.

Population structure

Population structure was revealed using the Bayesian clustering procedures implemented by STRUCTURE 2.2 (Pritchard *et al.*, 2000), and INSTRUCT (Gao *et al.*, 2007). The most informative number of genetic clusters was inferred according to the method of Evanno *et al.* (2005). The ad hoc statistic ΔK was calculated by running STRUCTURE for 10^6 generations (admixture model, burn-in of 5.10^5 generations) with five iterations for each value of K ranging from 1 to 17. In order to allow asymmetric patterns of admixture among populations the Dirichlet parameter for degree of admixture (α) was separately inferred for each population (Pritchard *et al.*, 2000). At $K = 4$, analyses in STRUCTURE were repeated using different priors, namely either by considering missing data as recessive homozygotes for the null alleles (Recessivealleles = 1) or including location information (Locprior = 1). CLUMPP v1.1.2 (Jakobsson & Rosenberg, 2007) was used to summarize the posterior estimates of cluster memberships of the ten best runs of STRUCTURE ($K = 4$), *viz.*

https://doi.org/10.1017/S0007485319000026 Published online by Cambridge University Press

Table 1. Sampling locations, sampling year and geographic coordinates (decimal degrees) of populations of *Z. cucurbitae* from Africa.

Location			Number of individuals	Code	latitude	longitude	Sampling year
1	Senegal	Niayes	24	Sene_Niay	14.8	−15.77	2008
2	Guinea	Foulaya	24	Guin_Foul	10.28	−14.45	2008
3	Ivory Coast	Korhogo	8	IvCo_Korh	9.45	−5.63	2008
4	Burkina Faso	Bobo Dioulasso	23	Burk_Bobo	11.2	−4.3	2008
5	Togo	Agome	24	Togo_Agom	6.95	0.64	2009
6	Togo	Tove Abessia	24	Togo_Abes	6.87	0.67	2009
7	Benin	Koro	24	Beni_Koro	8.9	2.57	2008
8	Congo D.R.	Kinshasa	16	Cong_Kins	−4.3	15.3	2008
9	Sudan	Singa	29	Suda_Sing	13.18	33.96	2009
10	Sudan	Abunaama	12	Suda_Abun	12.72	34.11	2009
11	Burundi	Kigwena	12	Buru_Kigw	−4.14	29.53	2009
12	Uganda	Jinja	25	Ugan_Jinj	0.44	33.2	2009
13	Uganda	Wakiso	17	Ugan_Waki	0.4	32.48	2010
14	Kenya	Nguruman	24	Keny_Ngur	−1.75	36.03	2008
15	Kenya	Nairobi	22	Keny_Nair	−1.28	36.82	2009
16	Tanzania	Morogoro	24	Tanz_Moro	−6.82	37.67	2008
17	Tanzania	Tanga	24	Tanz_Tang	−5.07	39.1	2006

All samples were from traps based in the different localities.

those runs with the highest log probability of the data (LnP (D)). We used the *Large K Greedy* algorithm of CLUMPP with random input order and 10^3 permutations to align runs and the G′ pairwise matrix similarity statistics. Admixture proportions of samples and individuals were then visualized using DISTRUCT v1.1 (Rosenberg, 2004). The admixture proportions resulting from STRUCTURE ($K = 4$) were then interpolated with the geographic distribution of individuals. Posterior predictive maps of admixture proportions among individuals were obtained using kriging as described in Virgilio *et al*. (2010). In parallel, INSTRUCT software, robust to HW disequilibrium was run with five chains for $K = 1$–20 (10^4 burn-in steps and 2×10^5 iterations).

Demographic models

An ABC analysis was conducted to infer the sub-Saharan colonization history of *Z. cucurbitae* in Africa. DIYABC v2.0 (Cornuet *et al*., 2008; Cornuet *et al*., 2010; Cornuet *et al*., 2014) was used to test a number of scenarios compatible and not compatible with the available historical records. All scenarios were based on introductions from a source population, possibly of Central Asian origin (see Virgilio *et al*., 2010 and Boontop *et al*., 2017*b*). The coalescent model implemented in DIYABC assumes the divergence of panmictic populations without recurrent migration. To avoid violating this assumption, we restricted our analysis to four representative populations (out of 17), each belonging to one of the main genetics clusters identified by STRUCTURE, and showing as few traces of admixture as possible. Based on co-ancestry coefficients computed in STRUCTURE, Tanz_Tang was used as representative of the East African group 1, Ugan_Waki was used for East African group 2. Congo_Kins was used for Central Africa, IvCoast_Khor represented West Africa and the source group Asia was represented by an Indian sample ($n = 32$, Virgilio *et al*., 2010). The aim of the analysis was to retrace the order of colonization events and estimate the extent of possible bottlenecks. Even with four populations, the number of possible scenarios (including bottlenecks and multifurcations) is relatively large. Scenarios without and with bottlenecks, were tested, among which the source population was from Asia as a unique event to Africa (any one of the four countries, and all the other invasions were derived from this source population), then including as well a scenario as Asia being the source population to each of them, testing the possible 'multiple introduction events', more complex scenario including multifurcations were also tested between all populations. Then, among those ones we discarded the ones with the lowest probabilities and kept the 25 best ones (SM1). In this analysis, all scenarios assuming an initial introduction into Africa different from Tanz_Tang had very low posterior probabilities (below 0.001). Similarly, scenarios omitting genetic bottlenecks for Tanz_Tang, Ugan Waki, or Congo_Kins received extremely low support.

For the clarity of presentation, all previous scenarios were discarded and the whole ABC scenario choice analysis was repeated on a short list of six scenarios with non-negligible posterior probabilities. All of these included East Africa as the entry point of *Z. cucurbitae* in the continent and differed with respect to the subsequent colonization history, i.e., to the order of colonization of Central-East, Central and West Africa (fig. 1). For each scenario and each population, the following demographic parameters were estimated: dates of founding (as the number of generations) (t_i), current effective population size (as the number of diploid individuals) (N_{ei}), number of founders at introduction (N_{bi}) and duration of the initial bottleneck (d_{bi}). For all these demographic parameters, prior distribution ranges, (shown in table 2) were estimated according to known records of *Z. cucurbitae* in sub-Saharan Africa (De Meyer *et al*., 2015). Time, expressed in generations before sampling, was translated in years assuming eight generations per year. This indicative value was inferred from the previously published studies on African *Z. cucurbitae* (Dhillon *et al*., 2005) as well as from the mean generation time (T) calculated on Hawaiian populations (8.4–9.2 generations per year with a temperature range from 24 to 35 °C, see Vargas *et al*. (2000)). One million simulations were conducted under each scenario. Posterior probabilities of each scenario were computed by performing a polychotomous weighted logistic regression on the 1% simulated datasets closest to the observed dataset (Cornuet *et al*., 2008; Cornuet *et al*., 2010) after linear discriminant analysis on summary statistics (Cornuet *et al*., 2014). Confidence in

https://doi.org/10.1017/S0007485319000026 Published online by Cambridge University Press

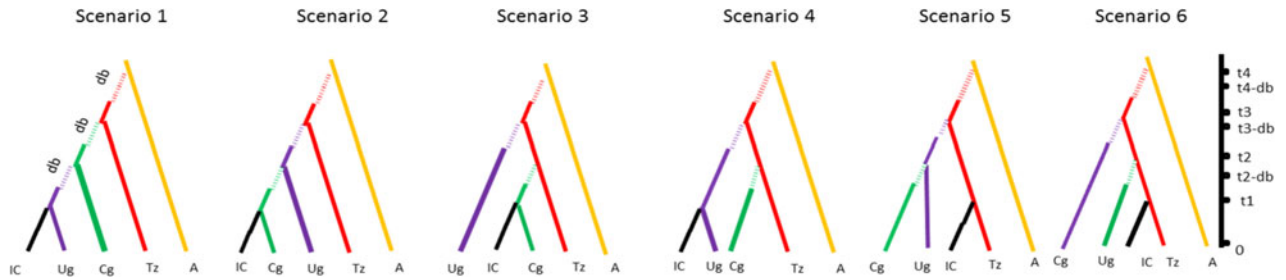

Fig. 1. Schematic drawing of the six scenarios used for the ABC method with time scale (t_1 to t_4) and bottleneck (db = duration of bottleneck). Country codes are the following: IC, Ivory Coast; UG, Uganda; CG, Congo; TZ, Tanzania; A, Asia.

Table 2. Summary of the DIYABC analysis based on 6,000,000 simulated datasets averaged over 60,000 selected datasets.

Parameter names	Prior range	Posterior parameter estimates	95% CI	Relative bias	Relative square root error
Population size effective					
$N_{1Ivory-Coast}$	(100–5000)	4390	(2420–4900)	−0.030	0.290
$N_{2Congo\ D.R.}$	(500–8000)	2210	(1270–7270)	−0.005	0.251
$N_{3Uganda}$	(500–8000)	3300	(1680–7330)	0.014	0.291
$N_{4Tanzania}$	(500–8000)	1450	(985–5110)	0.074	0.262
N_{5Asia}	(500–20,000)	17,400	(12,100–19,500)	0.104	0.276
Time in generation					
t_1	(100–1000)	144	(112–445)	−0.064	0.290
t_2	(200–1000)	318	(228–667)	0.039	0.205
t_3	(300–1000)	477	(357–921)	0.193	0.280
t_4	(300–1500)	998	(602–1440)	0.254	0.281
d_{b2}	(1–200)	29.9	(9.54–188)	−0.563	0.643
d_{b3}	(1–200)	24.6	(8.96–186)	−0.691	0.710
d_{b4}	(1–200)	155	(13.9–191)	−0.712	0.719
Genetic parameters (rate)					
μ_{mic}	$(1.00 \times 10^{-5}$–$1.00 \times 10^{-3})$	5.57×10^{-5}	$(3.64 \times 10^{-5}$–$1.36 \times 10^{-4})$	0.13	0.35
p_{mic}	$(1.00 \times 10^{-1}$–$3.00 \times 10^{-1})$	3.00×10^{-1}	$(1.69 \times 10^{-1}$–$3.00 \times 10^{-1})$	−0.05	0.35
sni_{mic}	$(1.00 \times 10^{-8}$–$1.00 \times 10^{-5})$	1.00×10^{-8}	$(1.29 \times 10^{-8}$–$2.54 \times 10^{-6})$	−0.70	0.99

Prior minimum and maximum values, posterior parameter estimates (as inferred from the mode of distributions), 95% confidence interval (CI), relative bias and relative square root error calculated for the best scenario (scenario 6).

scenario choice was further tested using additional simulations. Specifically, 500 pseudo-observed datasets (PODs) were simulated under each scenario and treated as real data, with their posterior probabilities computed as described above. The type II error rate was then computed as the proportion of PODs with the highest posterior probability (PP) for the retained scenario (Cornuet *et al.*, 2008). The choice of scenario was followed by parameter inference, with parameter posterior distributions and their 95% confidence intervals (CI) that were inferred via local linear regression on the 1% closest simulated data sets (logit transformation). The precision of parameter inference was then assessed by computing the Square Root of the Mean Square Error of each parameter and the Median of the Absolute Error (MAE), based on the 500 PODs.

In parallel to DIYABC analysis, we used GENECLASS 2 software (Piry *et al.*, 2004) to detect the probability of assignation or exclusion of individuals from a given population. For doing so, we used the Bayesian criterion implemented in this software (Rannala Mountain, 1997) with 1000 simulated individuals ($\alpha = 0.01$) using the Markov Chain Monte Carlo resampling method (Paetkau, 2004).

Results

The analysis of 19 microsatellite loci in 17 populations of *Z. cucurbitae* ($n = 18.6$, SD = 5.8) from 12 African countries showed relatively low levels of genetic variability, with mean numbers of alleles (N_a) ranging from 2.7 to 3.7 and observed (H_{obs}) and expected heterozygosity ($H_{n.b.}$) ranging from 0.30 to 0.49 and from 0.40 to 0.54, respectively (table 3). After FDR correction, the exact tests showed significant deviations from HWE in 61 out of 302 population/locus combinations. Multilocus estimates of F_{is} ranged from 0.07 to 0.28 and showed significant heterozygote deficiencies in 15 out of 17 populations (table 3). The average gene diversity (H_s) of each population was high for each population, with comparable values with He (table 3). The allelic richnesses (A_r) were low and similar between African populations. No linkage disequilibrium among microsatellite markers was detected through the 171 pairwise comparisons (SM2) therefore all loci were considered as independent. The average proportion of null alleles per locus was 0.08 (SD = 0.05) with null alleles per population ranging from 0.04 (SD = 0.08) in Buru_Kigw to 0.10 (SD = 0.10) in Suda_Sing (table 3).

https://doi.org/10.1017/S0007485319000026 Published online by Cambridge University Press

Table 3. Summary of genetic variability 17 African populations of *Z. cucurbitae* (numbered according to table 1) analyzed at 19 microsatellite loci.

	N	N_a	A_r	H_s	H_{obs}	$H_{n.b}$	F_{is}	An
Sene_Niay	27.31 (3.23)	3.68 (1.29)	3.02 (0.23)	0.52 (0.04)	0.39 (0.20)	0.51 (0.19)	0.24*	0.08 (0.10)
Guin_Foul	23.10 (1.94)	3.63 (1.42)	2.96 (0.22)	0.51 (0.04)	0.40 (0.23)	0.50 (0.19)	0.21*	0.08 (0.10)
IvCo_Korh	22.89 (2.23)	3.63 (1.34)	3.08 (0.23)	0.52 (0.04)	0.40 (0.24)	0.51 (0.19)	0.23*	0.08 (0.11)
Burk_Bobo	22.15 (2.08)	3.63 (1.25)	3.11 (0.22)	0.54 (0.04)	0.46 (0.29)	0.53 (0.18)	0.14*	0.07 (0.12)
Togo_Agom	11.47 (0.84)	3.47 (1.34)	3.11 (0.24)	0.54 (0.04)	0.41 (0.23)	0.53 (0.18)	0.23	0.07 (0.09)
Togo_Abes	11.73 (0.56)	3.31 (1.20)	3.01 (0.23)	0.50 (0.05)	0.45 (0.28)	0.50 (0.21)	0.10*	0.05 (0.09)
Beni_Koro	23.05 (1.77)	3.63 (1.25)	3.05 (0.22)	0.53 (0.05)	0.46 (0.23)	0.52 (0.19)	0.12*	0.05 (0.08)
Cong_Kins	21.36 (2.75)	2.73 (1.14)	2.35 (0.19)	0.40 (0.05)	0.30 (0.22)	0.39 (0.21)	0.24*	0.08 (0.08)
Suda_Sing	11.57 (0.90)	3.10 (1.10)	2.93 (0.22)	0.52 (0.04)	0.37 (0.26)	0.51 (0.18)	0.28*	0.10 (0.10)
Suda_Abun	11.78 (0.53)	3.26 (1.09)	2.92 (0.18)	0.53 (0.04)	0.49 (0.27)	0.52 (0.15)	0.07*	0.06 (0.09)
Buru_Kigw	7.78 (0.41)	2.78 (1.13)	2.73 (0.25)	0.48 (0.06)	0.41 (0.29)	0.47 (0.24)	0.12	0.04 (0.08)
Ugan_Jinj	21.31 (1.73)	3.36 (1.11)	2.87 (0.19)	0.53 (0.05)	0.39 (0.21)	0.52 (0.19)	0.24*	0.08 (0.07)
Ugan_Waki	23.00 (1.88)	3.31 (1.05)	2.93 (0.19)	0.53 (0.05)	0.41 (0.24)	0.53 (0.20)	0.23*	0.08 (0.11)
Keny_Ngur	22.89 (1.99)	3.68 (1.29)	3.03 (0.21)	0.52 (0.04)	0.37 (0.20)	0.51 (0.18)	0.27*	0.08 (0.10)
Keny_Nair	14.94 (2.34)	3.15 (1.01)	2.8 (0.18)	0.49 (0.04)	0.44 (0.26)	0.49 (0.18)	0.10*	0.06 (0.09)
Tanz_Moro	23.21 (1.31)	3.21 (1.13)	2.79 (0.19)	0.50 (0.04)	0.36 (0.22)	0.50 (0.18)	0.27*	0.09 (0.11)
Tanz_Tang	17.21 (1.13)	2.94 (1.02)	2.66 (0.20)	0.49 (0.04)	0.34 (0.23)	0.46 (0.19)	0.27*	0.09 (0.11)

N_a, mean number of alleles; A_r, allelic richness; H_s, gene diversity; H_{obs}, observed heterozygosity; $H_{n.b}$, expected unbiased heterozygosity; F_{is}, Weir & Cockerham's (1984) fixation index. In each population, asterisks indicate multilocus deviations from HWE (experiment-wise $P < 0.05$ after False Discovery Rate correction), An, mean null allele frequency based on Dempster *et al.* (1977). Standard deviations are in parentheses.

Table 4. Pairwise F_{st} values (above diagonal) and results of permutational tests (below diagonal) among 19 African populations of *Z. cucurbitae*.

	Sene_Nia	Guin_Fou	IvCo_Kor	Burk_Bob	Togo_Ago	Togo_Abe	Beni_Kor	Cong_Kin	Suda_Sin	Suda_Abu	Buru_Kig	Ugan_Jin	Ugan_Wak	Keny_Ngu	Keny_Nai	Tanz_Mor	Tanz_Tan
Sene_Nia	–	−0.002	−0.009	0.008	−0.007	0.008	−0.005	0.117	−0.011	0.014	0.046	0.041	0.067	0.017	0.075	0.058	0.108
Guin_Fou		–	0.010	0.002	−0.008	0.007	0.002	0.119	0.007	0.008	0.063	0.032	0.059	0.019	0.079	0.049	0.085
IvCo_Kor			–	0.002	0.007	0.014	−0.002	0.105	0.001	0.019	0.055	0.044	0.069	0.029	0.064	0.073	0.131
Burk_Bob				–	−0.001	0.019	0.000	0.099	0.024	0.024	0.071	0.032	0.046	0.017	0.077	0.049	0.098
Togo_Ago					–	0.022	−0.004	0.110	−0.012	−0.001	0.036	0.015	0.045	0.009	0.090	0.033	0.065
Togo_Abe				*		–	−0.007	0.119	0.012	0.042	0.067	0.029	0.060	0.041	0.075	0.072	0.133
Beni_Kor							–	0.124	0.002	0.028	0.050	0.025	0.056	0.029	0.069	0.070	0.115
Cong_Kin	*	*	*	*	*	*	*	–	0.130	0.113	0.237	0.138	0.161	0.145	0.193	0.160	0.222
Suda_Sin				*				*	–	0.006	0.037	0.019	0.053	0.018	0.061	0.067	0.115
Suda_Abu				*		*	*	*		–	0.086	0.043	0.048	0.018	0.090	0.054	0.106
Buru_Kig	*	*	*	*		*	*	*		*	–	0.046	0.082	0.038	0.093	0.051	0.101
Ugan_Jin	*	*	*	*		*	*	*		*	*	–	0.029	0.038	0.073	0.084	0.127
Ugan_Wak	*	*	*	*	*	*	*	*	*	*	*	*	–	0.030	0.099	0.070	0.125
Keny_Ngu	*	*	*	*		*	*	*			*	*	*	–	0.077	0.033	0.071
Keny_Nai	*	*	*	*	*	*	*	*	*	*	*	*	*	*	–	0.132	0.169
Tanz_Mor	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	–	0.042
Tanz_Tan	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	–

*significant at experimentwise $a = 0.05$. Population abbreviations as in table 1.

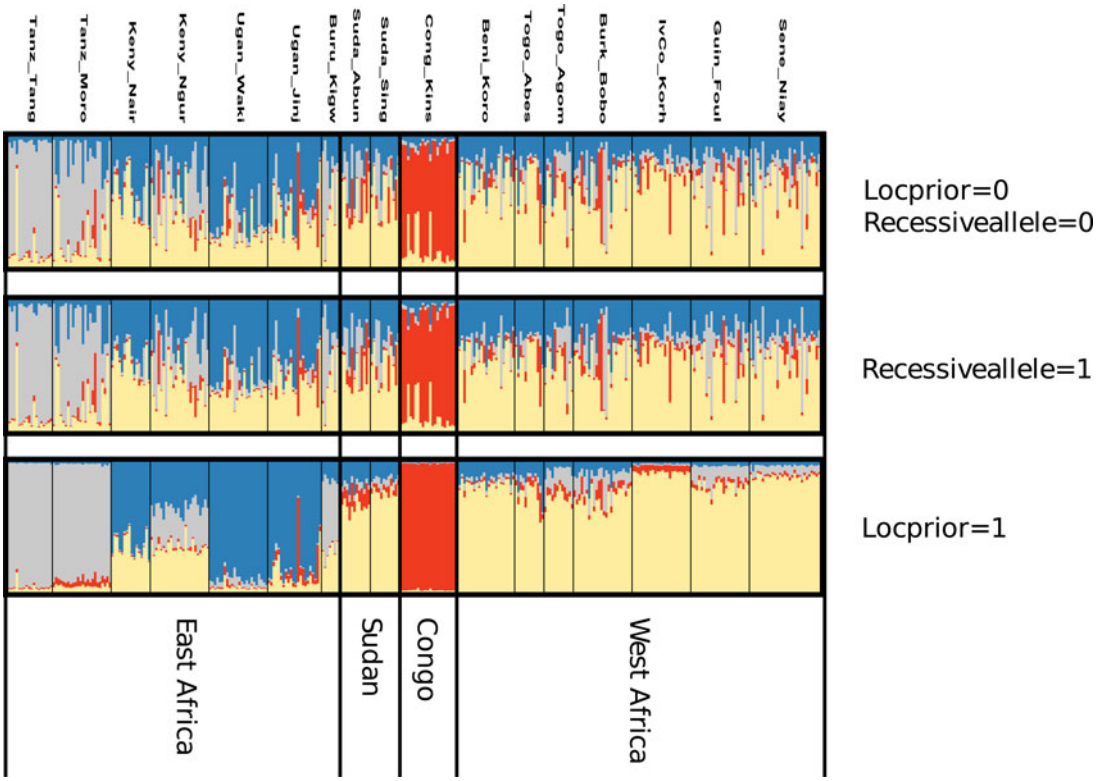

Fig. 2. Ancestry estimation based on the Bayesian clustering method STRUCTURE assuming four population clusters ($K = 4$). Each vertical line represents an individual, and each color represents a cluster. Individuals are grouped by sampling location (see table 1 for more details). Clustering STRUCTURE outcomes is presented at $K = 4$ (see results), with analyses in STRUCTURE repeated using different priors, namely either by considering missing data as recessive homozygotes for the null alleles (Recessivealleles = 1) or including location information (Locprior = 1).

After FDR correction, the majority (71.3%) of pairwise Fst values were significantly different from zero indicating that most *Z. cucurbitae* populations were genetically divergent (table 4). Analyses in STRUCTURE showed a break in the slope of likelihood values at $K = 4$, similar results were obtained with INSTRUCT, hence the reconstruction of the African population structure of *Z. cucurbitae* was based on four main population clusters (fig. 2, SM3). Most populations (13 out of 17) were assigned to one of the clusters with average admixture coefficients (Q) higher than 0.70 (table 5). All populations from West Africa (Sene_Niay, Guin_Foul, IvCo_Korh, Burk_Bobo, Togo_Agom, Togo_Abes, Beni_Koro) were primarily assigned to cluster 1 ($Q = 0.71$–0.92), as well as the two populations from Sudan (Suda_Sing, Suda_Abun, $Q = 0.77$ and 0.66, respectively). The population from Congo DR (Cong_Kins) was primarily assigned to cluster 2 ($Q = 0.97$), those from Uganda (Ugan_Jinj, Ugan_Waki) to cluster 3 ($Q = 0.79$, 0.90) and those from Tanzania (Tanz_Moro, Tanz_Tang) to cluster 4 ($Q = 0.89$, 0.95). Three samples showed relatively low average co-ancestry coefficients. These were the population Burundi (Buru_Kigw), which was primarily assigned to clusters 1 and 4 ($Q = 0.31$ and 0.54, respectively), and the two populations from Kenya (Keny_Ngur, Keny_Nair) primarily assigned to clusters 1 ($Q = 0.34$, 0.35) and 3 ($Q = 0.35$, 0.63).

Table 5. Average co-ancestry coefficients in 17 populations of *Z. cucurbitae* assigned to four clusters (population numbers and codes according to table 1).

		Genetic clusters			
Population number	Population code	1	2	3	4
1	Sene_Niay	0.89	0.02	0.03	0.06
2	Guin_Foul	0.81	0.04	0.03	0.13
3	IvCo_Korh	0.92	0.05	0.02	0.01
4	Burk_Bobo	0.71	0.06	0.09	0.13
5	Togo_Agom	0.72	0.06	0.06	0.16
6	Togo_Abes	0.73	0.07	0.15	0.04
7	Beni_Koro	0.82	0.02	0.12	0.03
8	Cong_Kins	0.01	0.97	0.01	0.01
9	Suda_Sing	0.77	0.05	0.13	0.05
10	Suda_Abun	0.66	0.12	0.17	0.05
11	Buru_Kigw	0.31	0.01	0.15	0.54
12	Ugan_Jinj	0.11	0.08	0.79	0.02
13	Ugan_Waki	0.03	0.02	0.90	0.05
14	Keny_Ngur	0.34	0.02	0.35	0.29
15	Keny_Nair	0.35	0.01	0.63	0.01
16	Tanz_Moro	0.04	0.05	0.02	0.89
17	Tanz_Tang	0.03	0.01	0.01	0.95

Coefficients were obtained from the STRUCTURE analysis illustrated in fig. 1 (see methods).

https://doi.org/10.1017/S0007485319000026 Published online by Cambridge University Press

Table 6. Confidence in scenario choice. The best-chosen scenario was the number 6 (embolden).

		Scenario with highest posterior probability						Probability to detect scenario 6 while data were generated using scenario *i*
		1	2	3	4	5	**6**	
Simulated under scenario	1	345	51	11	78	4	**11**	0.022
	2	46	364	67	13	4	**6**	0.012
	3	5	70	382	3	10	**30**	0.060
	4	59	3	4	414	3	**17**	0.034
	5	9	5	9	2	398	**77**	0.154
	6	10	5	32	22	55	**376**	0.752

Total type 2 error rate of scenario 6 = 0.282.

The GENECLASS 2 analysis aimed at detecting the probability of assignation or exclusion of individuals from a given population was run between our different populations (SM4). West African populations are showing quite high values between West African populations (and Sudan) reflecting recent exchanges of populations. These results are corroborating the STRUCTURE results showing high admixture levels between those populations.

The low posterior probabilities allowed confidently discarding 19 scenarios out of the 25 tested (see material and method) using the DIYABC software. The best scenario of the short list of 6, was scenario 6 (SM5) with a PP of 0.56 (95% CI [0.44–0.59]) followed by scenario 5 (PP = 0.10 [0.13–0.21]), and then by scenario 2 (PP = 0.11 [0.00–0.16]). The best scenario assumed that the representative populations from Congo DR, Uganda and Ivory Coast stem from independent introductions from Tanzania, with a source population from Asia (tables 2, 6; fig. 1). The estimated proportions of assignment rates obtained with GENECLASS2 are showing an asymmetric and unidirectional migration of the Asia population to all tested African populations (0.51 ± 0.03 to African populations and 0.03 ± 0.01 received from Africa), also corroborating this unidirectional origin of the African population.

The type II error rate (table 6) associated with scenario 6 (i.e. the probability of select scenario 6 though it is not correct) was relatively low (0.056). Parameter estimates, relative bias and relative precision indices for the best simulations are detailed in table 2. Representative populations from Congo DR, Tanzania and Uganda experienced bottlenecks lasting from 24 to 155 generations, while according to the model; the representative population from Ivory Coast did not experience bottlenecks (see Methods). Estimations suggested that 985–5100 individuals (N_e) were introduced to Tanzania from a larger ancestral population with an effective size of 12,100–19,500 individuals (fig. 3). DIYABC estimated that *Z. cucurbitae* was introduced in Tanzania 998 generations before sampling (95% CI = 602–1440), approximately corresponding to 1883 AD (95% CI = 1828–1933). From Tanzania, then *Z. cucurbitae* reached Congo DR 477 generations before sampling (approximately in 1948, 95% CI = AD 1893–1963) then Uganda 318 generations before sampling (approximately in 1969, 95% CI = AD 1926–1980) and Ivory Coast 144 generations before sampling (approximately in 1990, 95% CI = AD 1952–1994) (fig. 3).

Discussion

The colonization routes of invasive species are often inferred from historical records that are sparse or incomplete (Estoup & Guillemaud, 2010). The reliability of historical records to trace back the colonization routes of African fruit flies is particularly difficult to verify, due to the relatively recent development of comprehensive monitoring programs and sampling campaign in different parts of Africa, to the patchy spatial and temporal distribution of tephritid flies (see Mwatawala *et al*., 2009; Mwatawala *et al*., 2010) and to the scarcity or total lack of information from large areas of the African continent (http://data.gbif.org/occurrences/). Previous analyses on the genetic structure of *Z. cucurbitae* suggested that the range expansion of this pest originated in Asia and that, based on their results (Virgilio *et al*., 2010; Jacquard *et al*., 2013; Dupuis *et al*., 2017; Boontop *et al*., 2017*a*), African populations were clearly distinguishable from the Asian populations as well as from the other worldwide-distributed population groups. Increasing the number of African samples (from almost its whole range of distribution in Africa) provided a better resolution of the intra-continental population structure of *Z. cucurbitae*.

This study shows that the sub-Saharan population structure of the melon fly includes four main genetic clusters: corresponding to populations of central Africa, Tanzania, Uganda and West Africa, respectively. Two countries, i.e., Burundi and Kenya, showed admixed patterns between East and West African populations. A high number of our sampled populations presented significant heterozygote deficiencies, however, with a low percentage of null alleles detected (<10%). This might possibly be explained by the spatial scale chosen for sampling, which might have attracted individuals out of their true scale of population, inducing more homozygotes than expected under HW equilibrium creating a Wahlund effect (Selkoe & Toonen, 2006). The low genetic diversity and high significant F_{is} of these African populations are also strong signals of founder effects that often occurs in invasive populations.

ABC scenario choice analyses supported the hypothesis of a single introduction event occurring in East Africa between AD 1828 and 1933. This estimate is compatible with the first historical record, AD 1936, available for Tanzania (De Meyer *et al*., 2015). Colonization events were followed by bottlenecks, promoting genetic divergence between populations, and then, expansion to East Africa (between AD 1893 and 1963), Central Africa (AD 1893–1963) and West Africa (AD 1952–1994). The present results corroborate the hypothesis of a relatively recent introduction to East Africa and subsequent expansion to West Africa of *Z. cucurbitae* and suggest that these events happened following the timeline suggested by the historical records (*viz*. 1936 in East Africa and 1999 in West Africa).

The recent results based on genomic (SNP) data (Dupuis *et al*., 2017) using a worldwide sampling of *Z. cucurbitae*

https://doi.org/10.1017/S0007485319000026 Published online by Cambridge University Press

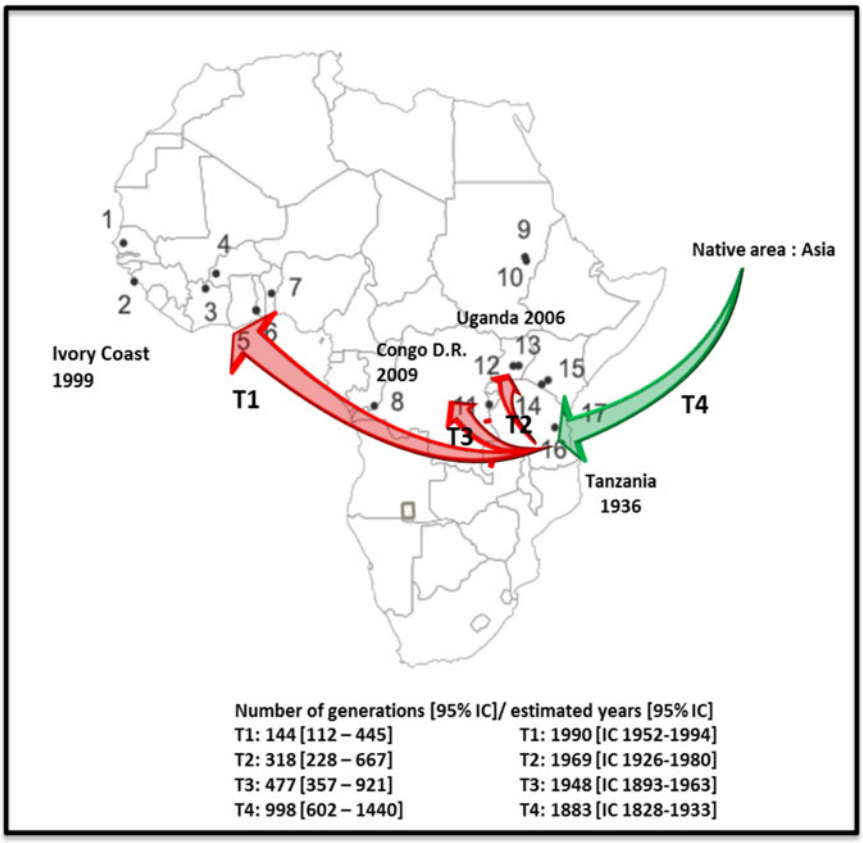

Fig. 3. Map of Africa with sampling locations (see table 1 for details), and dates of the first description of *Zeugodacus cucurbitae* from historical records (cf De Meyer *et al*., 2015), and most likely invasion routes of *Z. cucurbitae*, deduced from genetic analysis based on microsatellite markers variation and approximate Bayesian computation (DIYABC). The arrows indicate the most likely invasion pathways with the 95% confidence intervals indicated below the map. Estimated years evaluated from the generation time were also deduced, followed by 95% confidence intervals in brackets.

including two African populations showed that their sampled population from Tanzania was closer to the ones from Asia (Bangladesh/Nepal region) than the one from West Africa (Senegal; based on pairwise F_{st} values). Furthermore, their *Z. cucurbitae* population from Tanzania was closer to the one from Senegal than any other sampled population in their study. Those results are pointing toward a similar hypothesis of a single source of introduction in East Africa followed by an expansion to West Africa.

This allows excluding previous alternative hypotheses (Virgilio *et al*., 2010) considering older introductions of *Z. cucurbitae* possibly related to the first trade contacts between Africa and Asia (dating back as 100 AD, see De Meyer *et al*., 2015 and references therein). Range expansion promoted by a successful invasive population, which provides propagules outside the native distribution range has recently been described as the invasive bridgehead effect. Indeed this effect was first described on a coccinellid native to Asia (*Harmonia axyridis*), which is invasive worldwide, and where many of its invasions have stemmed not from its native range, but from a particularly successful invasive population (Lombaert *et al*., 2010). Similarly, we can hypothesize that the East African *Z. cucurbitae* might have played this role during the sub-Saharan range expansion of *Z. cucurbitae*. To a lesser extent, a similar pattern of invasion might have occurred for *B. dorsalis* in Africa, where it was suggested that the invasion of this pest across the whole African continent might have been realized by one or two source populations (Khamis *et al*., 2009).

These invasion patterns are driven by a more successful population, as suggested by Lombaert *et al*. (2010), suggest that we should increase vigilance against invasive bridgehead populations. Indeed, *Z. cucurbitae* might carry on its invasion wave in other countries where it has not been settled yet, and more careful control should be taken at borders to prevent its arrival in new territories.

Supplementary material

The supplementary material for this article can be found at https://doi.org/10.1017/S0007485319000026.

Acknowledgements

The authors wish to thank all colleagues and collectors who provided specimens for this study. A great thanks to V Ravigné for her advice for conducting the DIYABC methods, such as remarks on the manuscript. The authors also wish to

thank the two anonymous referees that improved our manuscript with their suggestions. This work was funded by the Conseil Régional de la Réunion, the European Regional Development Fund (ERDF), the Centre de Coopération International en Recherche Agronomique pour le Développement (CIRAD) and the Royal Museum for Central Africa. This work also had support from a FAO/IAEA Coordinated Research Project on 'Resolution of cryptic species complexes of tephritid pests to overcome constraints to SIT and international trade'.

References

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. (1996) GENETIX 4.05, logiciel sous WindowsTM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier (France).

Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B* (*Statistical Methodology*) **57**, 289–300.

Boontop, Y., Schutze, M.K., Clarke, A.R., Cameron, S.L. & Krosch, M.N. (2017*a*) Signatures of invasion: using an integrative approach to infer the spread of melon fly, *Zeugodacus cucurbitae* (Diptera: Tephritidae), across Southeast Asia and the West Pacific. *Biological Invasions* **19**, 1597–1619.

Boontop, Y., Kumaran, N., Schutze, M.K., Clarke, A.R., Cameron, S.L. & Krosch, M.N. (2017*b*) Population structure in *Zeugodacus cucurbitae* (Diptera: Tephritidae) across Thailand and the Thai–Malay peninsula: natural barriers to a great disperser. *Biological Journal of the Linnean Society* **121**, 540–555.

Chapuis, M-P. & Estoup, A. (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* **24**, 621–631.

Cornuet, J.M., Santos, F., Beaumont, M.A., Robert, C.P., Marin, J. M., Balding, D.J., Guillemaud, T. & Estoup, A. (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics (Oxford, England)* **24**, 2713–2719.

Cornuet, J-M., 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 Bioinformatics* **11**, 401.

Cornuet, J-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M. & Estoup, A. (2014) DIYABC v2. 0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30**, 1187–1189. https://doi.org/10.1093/bioinformatics/btt763

De Meyer, M., Delatte, H., Mwatawala, M., Quilici, S., Vayssières, J.-F. & Virgilio, M. (2015) A review of the current knowledge on *Zeugodacus cucurbitae* (Diptera: Tephritidae) in Africa. *ZooKeys* **540**, 539–557.

Delatte, H., Virgilio, M., Simiand, C., Risterucci, A.M., DeMeyer, M. & Quilici, S. (2010) Isolation and characterization of microsatellite markers from *Bactrocera cucurbitae* (Coquillett). *Molecular Ecology Ressources* **10**, 576–579.

Dempster, A., Laird, N. & DB, R. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B* **39**, 1–38.

Dhillon, M.K., Singh, R., Naresh, J.S. & Sharma, H.C. (2005) The melon fruit fly, *Bactrocera cucurbitae*: a review of its biology and management. *Journal of Insect Science* **5**, 1–16. https://doi.org/10.1093/jis/5.1.40.

Drew, R. & Hancock, D. (2000) Phylogeny of the tribe Dacini (Dacinae) based on morphological, distributional, and biological data. pp. 491–504 *in* Aluja, M. (Ed.) *Fruit Flies (Tephritidae): Phylogeny and Evolution of Behavior*. New York, CRC Press.

Dupuis, J.R., Sim, S.B., San Jose, M., Leblanc, L., Hoassain, M. A., Rubinoff, D. & Geib, S.M. (2017) Population genomics and comparisons of selective signatures in two invasions of melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae). *Biological Invasions* **20**, 1211–1228.

Estoup, A. & Guillemaud, T. (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology* **19**, 4113–4130.

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**, 2611–2620.

Gao, H., Williamson, S. & Bustamante, C.D. (2007) A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics* **176**, 1635–1651.

Goudet, J. (2001) Fstat, a Program to Estimate and Test Gene Diversities and Fixation Indices Version 2.9.3. Available online at http://www.unil.ch/izea/softwares/fstat.html.

Gilbert, E. (2004) *Dhows and the Colonial Economy of Zanzibar 1860–1970*, p. 167. Athens, Ohio, Ohio University Press.

Jacquard, C., Virgilio, M., David, P., Quilici, S., De Meyer, M. & Delatte, H. (2013) Population structure of the melon fly, *Bactrocera cucurbitae*, in Reunion Island. *Biological Invasions* **15**, 759–773.

Jakobsson, M. & Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics (Oxford, England)* **23**, 1801–1806.

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

Lombaert, E., Guillemaud, T., Cornuet, J.M., Malausa, T., Facon, B. & Estoup, A. (2010) Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *PLoS ONE* **5**, e9743.

Mwatawala, M.W., De Meyer, M., Makundi, R.H. & Maerere, A.P. (2009) Host range and distribution of fruit-infesting pestiferous fruit flies (Diptera, Tephritidae) in selected areas of Central Tanzania. *Bulletin of Entomological Research* **99**, 629–641.

Mwatawala, M., Maerere, A.P., Makundi, R. & De Meyer, M. (2010) Incidence and host range of the melon fruit fly *Bactrocera cucurbitae* (Coquillett) (Diptera: Tephritidae) in Central Tanzania. *International Journal of Pest Management* **56**, 265–273.

Nei, M. (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences* **70**, 3321–3323.

Nei, M. (1987) *Molecular Evolutionary Genetics*. New York, Columbia University Press.

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. *Molecular Ecology* **13**, 55–65.

Piry, S., Alapetite, A., Cornuet, J-M., Paetkau, D., Baudouin, L. & Estoup, A. (2004) Geneclass2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**, 536–539.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Rannala, B. & Mountain, J.L. (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences* **94**, 9197–9221.

Raymond, M. & Rousset, F. (1995) GENEPOP: a population genetics software for exact tests and ecumenicisms. *Journal of Heredity* **86**, 248–249.

Rosenberg, N.A. (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**, 137–138.

San Jose, M., Doorenweerd, C., Leblanc, L., Barr, N., Geib, S. & Rubinoff, D. (2017) Incongruence between molecules and morphology: a seven-gene phylogeny of Dacini fruit flies paves the way for reclassification (Diptera: Tephritidae). *Molecular Phylogenetics and Evolution* **121**, 139–149.

Selkoe, K.A. & Toonen, R.J. (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615–629.

Sood, P., Prabhakar, C.S., Yadav, D.S. & Thakur, S.K. (2016) Popularizing IPM of fruit flies in cucurbits and subtropical fruits through an area wide aproach in north western Himalaya. *Proceedings of the 9th ISFFEI*, 78–86.

Vargas, R.I., Walsh, W.A., Kanehisa, D., Stark, J.D. & Nishida, T. (2000) Comparative demography of three Hawaiian fruit flies (Diptera: Tephritidae) at alternating temperatures. *Annals of the Entomological Society of America* **93**, 75–81.

Vayssières, J.F., Rey, J.Y. & Traoré, L. (2007) Distribution and host plants of *Bactrocera cucurbitae* in West and Central Africa. *Fruits* **62**, 391–396.

Vayssières, J.F., Carel, Y., Coubes, M. & Duyck, P.F. (2008) Development of immature stages and comparative demography of two cucurbit-attacking fruit flies in Réunion Island: *Bactrocera cucurbitae* and *Dacus ciliatus*. *Environmental Entomology* **37**, 307–314.

Virgilio, M., Delatte, H., Backeljau, T. & De Meyer, M. (2010) Macrogeographic population structuring in the cosmopolitan agricultural pest *Bactrocera cucurbitae* (Diptera: Tephritidae). *Molecular Ecology* **19**, 2713–2724.

Virgilio, M., Jordaens, K., Verwimp, C., White, I.M. & De Meyer, M. (2015) Higher phylogeny of frugivorous flies (Diptera, Tephritidae, Dacini): localised partition conflicts and a novel generic classification. *Molecular Phylogenetics and Evolution* **85**, 171–179.

Weir, B.S. & Cockerham, C.C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.

White, I.M. & Elson-Harris, M.M. (1992) *Fruit Flies of Economic Significance: Their Identification and Bionomics*. Wallingford, CAB International.

Wu, Y., Li, Z. & Wu, J. (2009) Polymorphic microsatellite markers in the melon fruit fly, *Bactrocera cucurbitae* (Coquillett) (Diptera: Tephritidae). *Molecular Ecology Resources* **9**, 1404–1406.