Diversity and Distributions, (Diversity Distrib.) (2015) 21, 813–825

BIODIVERSITY RESEARCH

Deconstructing intercontinental invasion pathway hypotheses of the Mediterranean fruit fly (*Ceratitis capitata*) using a Bayesian inference approach: are port interceptions and quarantine protocols successfully preventing new invasions?

Minette Karsten[1], Bettine Jansen van Vuuren[2], Pia Addison[1] and John S. Terblanche[3]*

[1]*Department of Conservation Ecology and Entomology, Stellenbosch University, Stellenbosch, South Africa,* [2]*Centre for Invasion Biology, Department of Zoology, University of Johannesburg, Johannesburg, South Africa,* [3]*Centre for Invasion Biology, Department of Conservation Ecology and Entomology, Stellenbosch University, Stellenbosch, South Africa*

*Correspondence: John S. Terblanche, Department of Conservation Ecology and Entomology, Stellenbosch University, Private Bag X1, Stellenbosch 7602, South Africa.
E-mail: jst@sun.ac.za

ABSTRACT

Aim Knowledge of how effective interceptions and quarantine measures are in preventing new biological invasions is of paramount importance for maintaining ecosystem function in a rapidly changing world. Here, we determine current macrogeographic population structure and routes of invasion of the Mediterranean fruit fly (*Ceratitis capitata*) using genetic approaches and reconstruct and test invasion pathway hypotheses in a Bayesian framework.

Location Africa, Australia, Greece, Guatemala and Madeira.

Methods We sampled 323 *C. capitata* individuals from 14 locations worldwide and genotyped all individuals for 11 polymorphic microsatellite markers. We calculated measures of genetic diversity and determined population structure. Moreover, we reconstructed and tested eighteen invasion pathway scenarios in a Bayesian framework using ABC modelling.

Results We show a decrease in genetic diversity outside the native range (Africa) into the introduced range (Australia, Greece, Guatemala and Madeira). The most likely invasion pathway scenario closely matched the historical records, with an initial colonization of Europe from Africa and a secondary colonization of Australia from Europe. Moreover, we show an introduction from Greece to the Americas and, finally, a back introduction into South Africa from Europe.

Main conclusions Given the lack of new introductions into colonized (non-African) locations despite increasing trade, and apart from the initial invasion and establishment of the species outside of Africa, we conclude that quarantine and interception measures have been largely successful to date.

Keywords
Agricultural pest, biological invasions, fruit flies, medfly, microsatellites.

INTRODUCTION

With the increase in human population placing rising pressure on food security, together with the predicted effects of climate change on the productivity of agricultural areas (e.g. Godfray *et al.*, 2010; Bebber *et al.*, 2013), research on the biotic and abiotic factors influencing the likelihood, frequency and potential impacts of invasions is urgently required (Blackburn *et al.*, 2011; Kirk *et al.*, 2013). Invertebrate pests move around the world via different human-mediated pathways (reviewed in Hulme, 2009) and after establishment expand their range through natural dispersal

DOI: 10.1111/ddi.12333
http://wileyonlinelibrary.com/journal/ddi

A Journal of Conservation Biogeography
Diversity and Distributions

in the newly colonized region. These human-mediated pathways are closely linked to trade of commodities; transport by air and sea as well as subsequent movement on land by, for example, rail, road and canals being primary movement routes (Hulme, 2009; see also Gaston *et al.*, 2003).

The importance of understanding invasion pathways is made clear in the multiple practical implications thereof, especially surrounding the control of new invasions (i.e. risk management) and their further prevention (Wilson *et al.*, 2009; Estoup & Guillemaud, 2010; Blackburn *et al.*, 2011). For example, reconstruction of invasion pathways can assist in uncovering the native range which can, in turn, provide possibilities for biological control as natural enemies are typically sourced from the native range (Estoup & Guillemaud, 2010). Knowledge of invasion pathways can also be used to assess the effectiveness of current or previously implemented management strategies (e.g. quarantine and/or port interceptions) by assessing the extent of ongoing introductions of invasive species and how this affects genetic diversity of the introduced, newly established population.

After the initial introduction to the new location, the species avoiding or overcoming the geographic barrier faces many additional challenges before successful establishment and further spread can be achieved (Blackburn *et al.*, 2011). Broadly, propagule pressure [number and frequency of individuals introduced at a specific location over time (Lockwood *et al.*, 2009)] is considered the most important driver of successful invasions and not necessarily species-level characteristics (Lockwood *et al.*, 2005; Simberloff, 2009; Blackburn *et al.*, 2013). However, Szűcs *et al.* (2014) showed that although propagule pressure is important in the initial establishment stage, genetic processes are important for the subsequent spread and growth of the population. Predicting which species will become invasive has been an urgent and long-standing priority in invasion biology (e.g. Kimberling, 2004; Van Kleunen *et al.*, 2010; Higgins & Richardson, 2014) together with improved predictive models (Kirk *et al.*, 2013). The number of possible biological input parameters that can be considered in predictive modelling is extensive, but of primary importance are traits characterizing environmental niches, such as physiological parameters (e.g. thermal development requirements or activity limits) and dispersal capacity and/or distance travelled by natural or assisted means (e.g. Berthouly-Salazar *et al.*, 2013). Furthermore, traits of population demographics, which may be linked to environmental niches (e.g. Dixon *et al.*, 2009), are also significant, as rapid population growth rates can assist in establishment and evolutionary adaptation post-introduction (Gilchrist *et al.*, 2008; Rey *et al.*, 2012).

Arthropods (including the Insecta) are generally considered to be understudied in invasion ecology, but especially so on the African continent (Pyšek *et al.*, 2008). Some of the most successful invaders and agricultural pest insects worldwide are members of the Tephritidae (for example *Bactrocera dorsalis*, Aketarawong *et al.*, 2007; *Bactrocera invadens* (*Bactrocera dorsalis*), De Meyer *et al.*, 2010), commonly referred to as the 'true fruit flies', and comprising more than 5000 species (White & Elson-Harris, 1994). These fruit flies cause high levels of economic losses due to both direct damage of feeding larvae in fruit, as well as indirectly by placing restrictions on the export of fruit to certain trade-partner countries. Regardless of ongoing surveillance and rigorous quarantine measures to ensure fruit and vegetable consignments are pest free, infested consignments are still intercepted at port inspections and a likely source of new invasions (Papadopoulos, 2014). These new introductions can remain undetected for some time (Carey, 2010) and once established can spread rapidly, as in, for example, the recent spread of *B. invadens* (*B. dorsalis*) across the African continent (De Meyer *et al.*, 2010). Therefore, dispersal and movement of fruit flies by either natural or human-assisted means are receiving increased attention to determine both the native range and movement patterns of species and to identify sources of new introductions. This information feeds specifically into the development of risk assessments and management of phytosanitary programmes (Leung *et al.*, 2002).

Possibly the most notorious of the Tephritidae flies is *Ceratitis capitata* (Wiedemann), the Mediterranean fruit fly. *Ceratitis capitata* is highly invasive world-wide, likely because of its broad larval host range, tolerance of a wide range of climatic conditions and its relatively high dispersal capacity (Lance & McInnis, 2005; Nyamukondiwa *et al.*, 2010). *Ceratitis capitata* is of quarantine importance world-wide (Reyes & Ochando, 2004; De Meyer *et al.*, 2008) and has spread to many countries outside of its native range, which is believed to be Afrotropical (De Meyer *et al.*, 2002). Although Kenya (East Africa) is the likely native range (Malacrida *et al.*, 2007), there still remains some uncertainty as to the extent of the native range within Africa, as the levels of genetic diversity for South Africa (Karsten *et al.*, 2013) are similar to those estimated for Kenya and do not follow the decline in genetic variability seen in other derived populations world-wide (e.g. Bonizzoni *et al.*, 2004; Malacrida *et al.*, 2007). Although population genetics methods have been used previously to investigate the colonization process, routes of invasion and the underlying evolutionary forces that have shaped the population structure of tephritids (e.g. Bonizzoni *et al.*, 2001, 2004; Gilchrist & Meats, 2009; Aketarawong *et al.*, 2014), several pressing invasion biology questions remain.

Global studies using genetic markers have shown that *C. capitata* populations can be subdivided into three distinct groups: first, the ancestral population from sub-Saharan Africa; second, a Mediterranean Basin group; and finally, a group from Latin America and the Pacific (Gasperi *et al.*, 2002). The colonization history of *C. capitata* world-wide is relatively well documented (Gasperi *et al.*, 2002; Malacrida *et al.*, 2007). The proposed invasion pathway and historical chronological order of colonization events include, firstly an introduction of *C. capitata* to the Mediterranean coast (possibly on the coast of Spain) (De Breme, 1842; Fimiani, 1989)

likely human-mediated (Maddison & Bartlett, 1989) from Africa, after which the rest of the Mediterranean area was colonized (Fimiani, 1989). This initial colonization and spread throughout the Mediterranean area was confirmed using Slatkin's private allele approach (Gasperi *et al.*, 2002). This is thought to have been followed by a secondary introduction to Australia from Europe in the 1890s (Hooper & Drew, 1989), which was confirmed using microsatellites (Bonizzoni *et al.*, 2004). The next leg of introductions in *C. capitata's* global colonization was into the Americas (Costa Rica) (Harris, 1989). The source for the colonization of Guatemala remains unclear, although some evidence exist to show that it might have been a southwards spread from Costa Rica after which Mexico was colonized with the flies moving along the coffee belt (Harris, 1989). Given the nature of these records and observations, these colonization dates all likely reflect the earliest record of *C. capitata* in the location, but not necessarily the date of introduction. This incongruity can be due to fruit flies establishing in an area long before they are detected in routine monitoring. Previously, support for proposed routes of invasion has come from genetic diversity estimates, where the area with the highest level of genetic diversity would be assumed to be part of the native range (Malacrida *et al.*, 2007), or from gene flow estimates based on Slatkin's private allele method (Malacrida *et al.*, 1998). However, these previous estimates are potentially weakened by the reduction of genetic data to a single summary statistic of genetic diversity, which does not necessarily account for stochastic population events (Estoup & Guillemaud, 2010). To date, however, hypotheses of global colonization of *C. capitata* have not been examined and tested using a strong inference framework, with predefined hypotheses (e.g. Pascual *et al.*, 2007).

Here, we make use of microsatellites, genetic diversity estimates and individual fly sampling at broad geographic scales to investigate the macrogeographic (intercontinental) population structure with a strong focus on sub-Saharan Africa and whether this can inform us of the origins, pathways of invasion and then subsequent success of prevention or intervention strategies for further establishment of *C. capitata* in those countries. We hypothesize that, based on previous studies of global *C. capitata* structure (Malacrida *et al.*, 1998, 2007; Gasperi *et al.*, 2002), there will be a clear decrease in global genetic diversity moving away from East Africa, as well as strong genetic differentiation between native populations in Africa and derived populations elsewhere in the world, possibly due to founder effects and genetic drift.

Moreover, we investigate eighteen distinct but plausible scenarios for the invasion pathway of *C. capitata* using approximate Bayesian computations (ABC). All scenarios considered are based on an 'out-of-Africa' approach including information from *C. capitata* macrogeographic studies and historical information of the colonization. The different scenarios were centred on an introduction from Africa (all locations excluding Burgers Hall) to the Mediterranean region [Unsampled location (possibly Spain)] after which the range expanded to Madeira and Greece, with a secondary introduction from the Mediterranean region to Australia. In each scenario, we also test from which location Guatemala was colonized as this information is unavailable in the literature together with testing from where Burgers Hall (South Africa) was recolonized.

METHODS

Sampling and microsatellite genotyping

We sampled 323 *C. capitata* individuals from 14 populations across four biogeographic regions (Afrotropical, Neotropical, Palaearctic and Australasian), representing all major regions from where *C. capitata* are known to occur except the Indian Ocean Islands, the Middle East, South America as well as North America (only with sporadic outbreaks), with more extensive sampling focused on sub-Saharan Africa (Table 1). Flies were collected via trapping in orchards using baits as well as reared from infested fruit. We refer to individuals from the same sampling location (several traps across a few hundred metres) as a population in the rest of this study. After morphological confirmation of the identification of *C. capitata*, DNA was extracted using a DNeasy® Tissue Kit (Qiagen Inc., Hilden, Germany). All individuals collected were genotyped for 11 microsatellite markers following Karsten *et al.* (2013). In each plate, we included a negative and a positive control to check that plates were read consistently between different runs. Samples were genotyped using an ABI 3130 Automated Sequencer (Applied Biosystems, Foster City, CA, USA). Samples that remained unamplified after two independent reactions were assumed non-amplifiable and therefore not included. Four of the South African sites were part of the Karsten *et al.* (2013) study. Only a small subset was genotyped again, and all individuals were scored again with the additional samples included.

Microsatellite analysis

Estimates of sample variability

Alleles were scored in GENEMAPPER v3.7 (Applied Biosystems). All microsatellite markers were tested for linkage disequilibrium and deviations from Hardy–Weinberg equilibrium (HWE) using 10,000 permutations in GENEPOP v4.01 (Raymond & Rousset, 1995; Rousset, 2008). In all cases where multiple testings were performed, we adjusted significance levels using false discovery rates (FDR; QVALUE, Storey, 2002), and these are the values reported. Genetic diversity levels were investigated calculating basic statistics for all populations including expected heterozygosity (H_E), observed heterozygosity (H_O), number of alleles (N_A), number of private alleles (N_P), as well as the inbreeding coefficient (F_{IS}) (GENETIX v4.05.2, Belkhir *et al.*, 1996–2004;

Table 1 The locations of collection sites of *Ceratitis capitata* used in this study as well as sample size (N), number of alleles (N_A), number of private alleles (N_P), expected (H_E) and observed (H_O) heterozygosity (± standard error), allelic richness (A_R), the inbreeding coefficient (F_{IS}) and the mean null allele frequency (A_n, Dempster *et al.*, 1977; SD in parentheses)

				GPS coordinates									
	Country		ID	Latitude	Longitude	N	N_A	N_P	H_E	H_O	A_R	F_{IS}	A_n
Africa													
	Kenya		KEN	−1.29	36.83	27	10.727	0.636	0.796 ± 0.149	0.533 ± 0.143	5.657	0.348	0.139 (0.087)
	Mozambique		MOZ	−19.13	33.43	22	9.818	0.364	0.809 ± 0.113	0.606 ± 0.201	5.698	0.276	0.115 (0.110)
	Senegal		SEN	14.18	−16.56	9	3.909	0	0.588 ± 0.136	0.511 ± 0.247	5.569	0.194	0.061 (0.107)
	South Africa												
		Burgers Hall	BUR	−25.06	31.05	10	3.273	0.273	0.431 ± 0.224	0.404 ± 0.173	2.877	0.134	0.023 (0.053)
		Levubu	LV	−23.05	30.17	29	11.273	0.636	0.793 ± 0.126	0.575 ± 0.188	5.571	0.292	0.115 (0.103)
		Port Elizabeth	PE	−33.96	25.57	6	4.909	0.182	0.641 ± 0.205	0.530 ± 0.277	3.927	0.274	0.068 (0.090)
		Stellenbosch	CL	−33.88	18.74	30	11.273	1.182	0.814 ± 0.080	0.637 ± 0.188	5.787	0.235	0.099 (0.100)
		Upington	UP	−28.45	21.24	29	9.818	0.455	0.779 ± 0.103	0.611 ± 0.202	5.65	0.233	0.098 (0.083)
	Tanzania		TAN	−6.95	37.53	28	11.273	0.727	0.819 ± 0.091	0.662 ± 0.214	5.138	0.211	0.095 (0.106)
	Zimbabwe		ZIM	−17.86	31.04	23	10.273	0.273	0.800 ± 0.113	0.603 ± 0.212	5.228	0.269	0.116 (0.117)
Australia			AUS	−31.98	115.88	30	3.091	0	0.378 ± 0.232	0.303 ± 0.254	2.26	0.214	0.064 (0.095)
Europe													
	Greece		GRE	39.4	21.91	28	2.727	0	0.369 ± 0.228	0.287 ± 0.245	2.147	0.244	0.070 (0.112)
	Madeira		MAD	32.67	−16.84	22	3.364	0.091	0.487 ± 0.176	0.319 ± 0.197	2.718	0.373	0.125 (0.101)
Americas													
	Guatemala		GUA	14.89	−90.51	30	2.909	0	0.391 ± 0.193	0.359 ± 0.218	2.217	0.102	0.047 (0.070)

GENALEX v6.5, Peakall & Smouse, 2006, 2012). The frequency of null alleles (A_n) was estimated in FREENA v1.0 (Chapuis & Estoup, 2007).

Population structure analysis

Individuals were assigned to populations (genetic clusters) using STRUCTURE v2.3.4 (Pritchard *et al.*, 2000; Falush *et al.*, 2003) based on multilocus genotypes without including *a priori* information. The number of optimal clusters (K) was estimated using an admixture model and correlated allele frequencies which allow allele frequencies to be similar in different populations due to shared ancestry or continued migration. For each possible cluster (K), a number varying between 1 and 14 (total number of sampling localities), we performed 10 independent runs with 1,000,000 MCMC permutations and burn-in set at 100,000, which allowed the different parameters to reach stability and convergence. As this only detects the uppermost level of differentiation in the data, we subsequently investigated the substructure in the data following Coulon *et al.* (2008) by repeating the analyses until the number of clusters inferred was one. The optimal K (most likely number of clusters) was assessed using two methods. First, we inspected the log-probabilities of the different possible clusters; a high value with limited variance is an indication of the true K (Pritchard *et al.*, 2000). Second, we inspected the second-order rate change of Ln P(X/Y) calculated according to Evanno *et al.* (2005) implemented in the online resource STRUCTURE HARVESTER (Earl & Von Holdt, 2012). Thereafter, the number of runs (from STRUCTURE) of the chosen K value was averaged in CLUMPP v1.2.2 (Jakobsson & Rosenberg, 2007) and visualized in DISTRUCT v1.1 (Rosenberg, 2004).

We quantified the degree of population differentiation between the genetic clusters identified in STRUCTURE by calculating pairwise F_{ST} values between the clusters in MICROSATELLITE ANALYSER v4.05 (MSA; Dieringer & Schlötterer, 2003) as well as overall F_{ST} values in FREENA v1.0 running 10 000 replications including (INA) and excluding (ENA) null alleles (Chapuis & Estoup, 2007). The hypothesis of isolation by distance (IBD) within the two clusters identified in STRUCTURE was investigated using two separate Mantel tests in ARLEQUIN v3.5.1.2 (Excoffier & Lischer, 2010). In each case, we regressed the sum of squared size differences (R_{ST}) against the geographic distance (km). To further assess the relationships between different populations based on their allele frequency, we used a Principal Coordinate Analysis (PCoA) implemented in the program GENALEX v6.5 (Peakall & Smouse, 2006, 2012). The first three principal axes were plotted in STATISTICA v12 (Statsoft Inc., Tulsa, Oklahoma).

We constructed unrooted neighbour-joining trees and assessed statistical support for each branch using non-parametric bootstrapping (10,000 replicates) using the online POPTREEW (Takezaki *et al.*, 2014) based on Nei's genetic distance (D_A; Nei *et al.*, 1983), as well as the genetic distance calculated based on the proportion of shared alleles (D_S; Bowcock *et al.*, 1994).

Population demography analysis

To obtain in-depth information regarding the invasion pathway of *C. capitata,* we tested eighteen different hypothetical scenarios and analysed them using the ABC method implemented in the program DIYABC v2.03 (Cornuet *et al.*, 2014). The use of ABC has many advantages including using several summary statistics simultaneously as well as providing probability values with confidence intervals (Estoup & Guillemaud, 2010). Many different scenarios can be tested with the different populations in this study which can become computationally intensive. We therefore narrowed down the number of scenarios to test by utilizing historical information as well as results from our cluster analyses. On the basis of Bayesian cluster analyses, we based the different scenarios on six populations [Africa (without Burgers Hall), Burgers Hall, Madeira, Greece, Guatemala and Australia] as well as including an unsampled population. This unsampled population corresponds to an initial site of colonization of the Mediterranean coast at a location other than locations sampled in this study. We tested 18 different scenarios related to testing (1) from where Guatemala was colonized as this information is unknown as well as (2) from where Burgers Hall (South Africa) was recolonized (Appendix S1 in Supporting Information). In total, we considered eighteen scenarios to test as hypotheses for the route of invasion, specifically surrounding the overall theme that *C. capitata* originated in eastern and southern Africa and subsequently spread to the rest of the world. The prior distributions of parameters were uniform and outlined as follows: $1000 < N1 < 1000\,000$; $10 < N2,N3,N4,N5,N6,N7 < 100000$; $1 < db < 10$; $2 < Nib < 100$; $10 < t1,t2 < 500$; $100 < t3,t4 < 1500$; $1000 < t5,t6 < 10000$; $0.001 < ra < 0.999$; N is the effective population size, Nib is the number of founders in each colonization event, db is the duration of the bottleneck event (number of generations), ra is the admixture rate and t is the timing of an event (generations back in time) (Appendix S2). Prior distributions of parameters for the microsatellite mutation model were set to default values and therefore represented the generalized stepwise mutation model. We included all summary statistics in both the one-sample and two-sample summary statistics. Our timing of events was based on the approximate number of generations *C. capitata* can support in a year and colonization dates available. We based the number of generations per year on the number of days development takes from egg to egg in the laboratory at 18 °C (Grout & Stoltz, 2007). We assumed that the effective population size for the native populations was larger than those in the introduced range and the bottleneck event (db) occurred after the initial introduction. The newly introduced individuals might take several generations to establish a population, and this parameter was therefore bounded between

one and ten generations. In each test, we simulated 7,000,000 computations and for each scenario computed a posterior probability, including 95% confidence intervals (CI), with a logistic regression (Cornuet *et al.*, 2014). The scenario with the highest posterior probability and non-overlapping 95% CI was chosen as the most likely scenario.

RESULTS

Population genetic diversity

Overall, the *C. capitata* populations examined showed deviations from HWE. These departures from HWE can possibly be coupled with moderate levels of inbreeding (F_{IS}) ranging between 0.102 (Guatemala) and 0.373 (Madeira) (Table 1), the Wahlund effect (reduction in heterozygosity due to hierarchical structuring) and/or the presence of null alleles. The frequency of null alleles in all populations was low to intermediate and ranged between 0.023 and 0.139 (Table 1, Appendix S3). Null alleles are commonly found in taxa with high effective population sizes (Chapuis & Estoup, 2007). The F_{ST} estimation before and after null allele correction (see *Population structure* section) showed no difference, although this does not suggest that all other analyses are unbiased. Furthermore, no linkage disequilibrium was detected for any of the 11 microsatellite markers. All African populations possessed relatively high levels of genetic diversity based on the expected heterozygosity (H_E), number of alleles (N_A) and allelic richness (A_R) compared to those locations in the introduced range (Australia, Greece, Guatemala, Madeira) (Table 1). All African populations except Senegal had a number of private alleles (0.182–1.182; Table 1). The highest number of private alleles was found in Stellenbosch (South Africa) with many locations (Australia, Greece, Guatemala, Senegal) having no private alleles.

Population structure

Genetic differentiation was measured by overall and pairwise F_{ST} calculations between the genetic clusters inferred by the Bayesian clustering method (Appendix S4). All pairwise F_{ST} comparisons between the different clusters were significant. The overall F_{ST} value before null allele correction (F_{ST} = 0.143) as well as after ENA correction (F_{ST} = 0.142) indicated significant population differentiation. There was no significant pattern of IBD based on results from the Mantel test for the cluster that includes all populations in the introduced range as well as Burgers Hall (r = −0.149, P = 0.640). Moreover, significant IBD was detected over the African continent (excluding Burgers Hall) (r = 0.736, P = 0.031).

In the Principal Coordinate Analysis (Fig. 1a), the first 3 axes explained most of the genetic variation (63.22%). The first axis (45.33%) separates samples from Africa from the rest of the world including one sampling location from South Africa (Burgers Hall). The second axis separates the African group (Cluster 1) from a second group that includes Australia, Greece, Guatemala and Madeira. Moreover, detailed population structure can be investigated using the Bayesian clustering method implemented in STRUCTURE. Graphical representation of the results from Evanno *et al.* (2005) and the log-probabilities (Appendix S5) indicated that K = 2 was the optimal number of clusters in STRUCTURE (Fig. 2). The first cluster grouped all African populations together, except for Burgers Hall (South Africa), which groups more closely with populations from the introduced range (Cluster 2). The peak at K = 2 (considering all locations) also corresponded to the results from the first axis of the PCoA (Fig. 1a) separating African sampling localities (excluding Burgers Hall) with those from the rest of the world. This clustering corresponded to the unrooted neighbour-joining trees reconstructed based on D_S (proportion of shared alleles) (Fig. 1b) and Nei's distance (data not shown) forming two clear groups. These two clusters were further

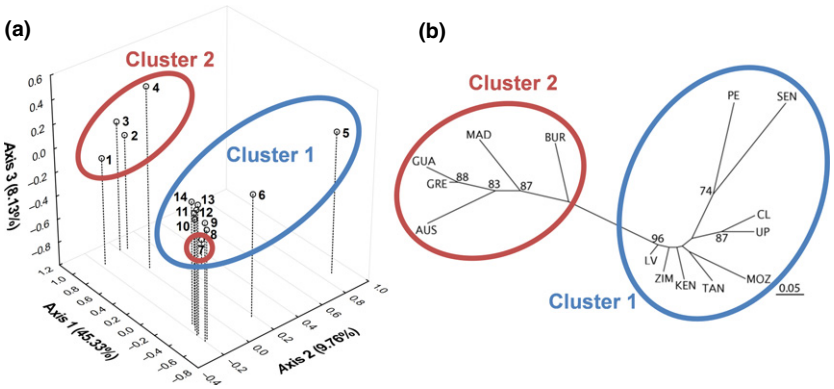

Figure 1 (a) Principal Coordinate Analysis (PCoA) plot for 14 *Ceratitis capitata* populations (1- Australia, 2- Greece, 3- Guatemala, 4- Madeira, 5- Senegal, 6- Port Elizabeth, 7- Burgers Hall, 8- Upington, 9- Stellenbosch, 10- Zimbabwe, 11- Levubu, 12- Mozambique, 13- Tanzania, 14- Kenya). (b) Unrooted neighbour-joining trees for genetic distance based on shared alleles (D_S). The number at each node indicates the bootstrap values after 10 000 bootstrap replicates. Only bootstrap values above 70% are shown. Coloured circles indicate the two clusters that correspond to clusters obtained in STRUCTURE.

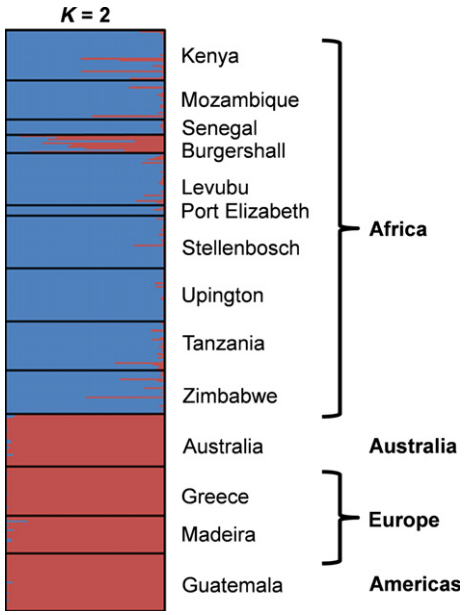

Figure 2 Assignment results from STRUCTURE for $K = 2$ for 14 *Ceratitis capitata* populations.

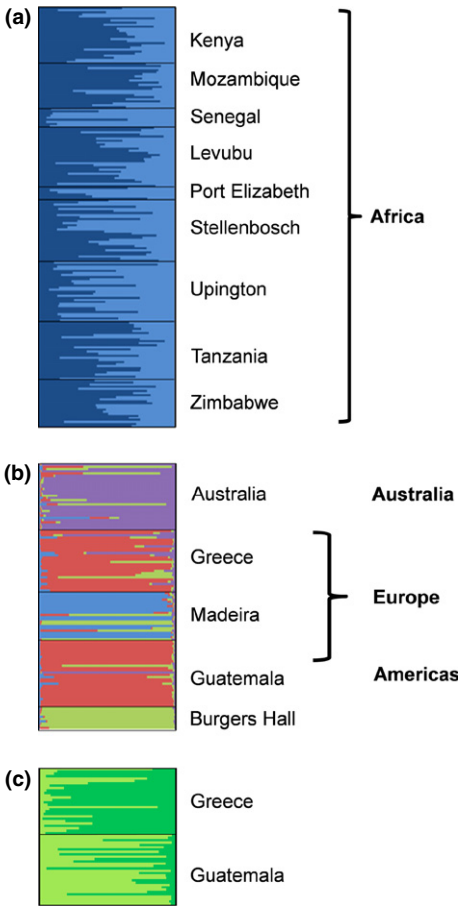

Figure 3 Assignment results from STRUCTURE for the African cluster at $K = 2$ (a), the world cluster at $K = 4$ (b) and the substructure in the Guatemala and Greece cluster (c).

investigated by running each cluster separately (Fig. 3). Substructuring in the African cluster (Cluster 1) (Fig. 3a) showed two further clusters based on the optimal K calculated in Evanno *et al.* (2005). However, no obvious geographic pattern was observed that correlated with the two clusters identified within the African group. Within the world cluster (Cluster 2), four additional clusters were identified (Fig. 3b). The first cluster included Australia, the second individuals from both Greece and Guatemala, the third individuals from Madeira, and the final cluster included all individuals from Burgers Hall. The cluster representing Greece and Guatemala was run again, and these two populations were subsequently clustered separately (Fig. 3c).

Invasion pathway scenarios from ABC

To investigate the routes of colonization, eighteen different scenarios were tested using approximate Bayesian computations. One scenario (Scenario 3), based on posterior probability values, was clearly superior to all other scenarios tested (Table 2; Appendix S6). This scenario closely matched the proposed historical routes of invasion (Fig. 4), with an initial colonization of Europe from Africa and a secondary colonization of Australia from Europe. Moreover, we show that Guatemala was likely colonized from Greece and the back introduction into South Africa at Burgers Hall came from the Unsampled location (Europe) included in the analysis. For the most likely scenario, the number of founders in each introduction event ranged between 37.9 and 66.70 and the duration of the bottleneck event was approximately nine generations (Appendix S7). Larger-than-expected effective population sizes (Appendix S7) were observed in some of the introduced populations (Australia, Guatemala, Unsampled location) as well as in Burgers Hall compared to populations from the native range. The timing of events (Appendix S7) was estimated as being much further in the past than indicated by historical records.

DISCUSSION

Global biological invasions of pest insects, such as that of *Ceratitis capitata*, are typically characterized by reduced genetic diversity due to a small number of founders colonizing the introduced range. This creates a unique set of genetic and demographic attributes which were investigated here making use of 11 microsatellites for 14 macrogeographic sampling locations of *C. capitata*. The colonization of *C. capitata* world-wide is well documented, and our results indicate high levels of genetic diversity in the native range and a decrease in this diversity in the introduced range. Moreover, clear genetic differentiation was found between Africa and the rest of the world. The reconstructed invasion pathway for *C. capitata* closely matches the proposed hypothesis of an initial colonization of Europe from Africa

Table 2 Results of eighteen different scenarios tested in DIYABC v2.03 (Cornuet *et al.*, 2014) with posterior probabilities and 95% confidence intervals (CI)

	Scenario	Posterior probability (95% CI)
1	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Greece-Guatemala; Madeira-Burgers Hall	0.001 (0.000, 0.546)
2	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Greece-Guatemala; Greece-Burgers Hall	0.001 (0.000, 0.549)
3	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Greece-Guatemala; Unsampled-Burgers Hall	0.998 (0.997, 0.999)
4	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Greece-Guatemala; Australia-Burgers Hall	0.000 (0.000, 0.545)
5	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Greece-Guatemala; Guatemala-Burgers Hall	0.000 (0.000, 0.545)
6	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Madeira-Guatemala; Greece-Burgers Hall	0.000 (0.000, 0.545)
7	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Madeira-Guatemala; Madeira-Burgers Hall	0.000 (0.000, 0.545)
8	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Madeira-Guatemala; Unsampled-Burgers Hall	0.001 (0.000, 0.545)
9	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Madeira-Guatemala; Australia-Burgers Hall	0.000 (0.000, 0.545)
10	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Madeira-Guatemala; Guatemala-Burgers Hall	0.000 (0.000, 0.545)
11	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Unsampled-Guatemala; Madeira-Burgers Hall	0.000 (0.000, 0.545)
12	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Unsampled-Guatemala; Greece-Burgers Hall	0.000 (0.000, 0.545)
13	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Unsampled-Guatemala; Unsampled-Burgers Hall	0.000 (0.000, 0.545)
14	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Unsampled-Guatemala; Australia-Burgers Hall	0.000 (0.000, 0.545)
15	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Africa-Guatemala; Madeira-Burgers Hall	0.000 (0.000, 0.545)
16	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Africa-Guatemala; Greece-Burgers Hall	0.000 (0.000, 0.545)
17	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Africa-Guatemala; Unsampled-Burgers Hall	0.000 (0.000, 0.545)
18	Africa-Unsampled-Madeira-Greece;Unsampled+Madeira+Greece-Australia; Africa-Guatemala; Australia-Burgers Hall	0.000 (0.000, 0.545)

and a secondary colonization of Australia from Europe. We also show that Guatemala was likely colonized from Greece and the reintroduction into Burgers Hall was from the Unsampled location on the Mediterranean coast.

The genetic diversity estimates from this research (expected heterozygosity, allelic richness, number of alleles and number of private alleles) are similar to those found in other studies of *C. capitata* (Bonizzoni *et al.*, 2001, 2004; Gasperi *et al.*, 2002). We found high levels of genetic diversity (expected heterozygosity, number of private alleles, allelic richness) in all African populations, except Senegal and Burgers Hall, with a further decline in the genetic diversity estimates of populations in the introduced range (Australia, Greece, Guatemala, Madeira). The high level of genetic diversity found in this study for *C. capitata* in Africa is possibly due to their large effective population sizes. By contrast, both Senegal and Burgers Hall populations have diversity estimates similar to those of populations in the introduced range. As eastern and southern Africa has been identified as the native range of *C. capitata* (Malacrida *et al.*, 1998, 2007; Gasperi *et al.*, 2002), it is perhaps unsurprising that Senegal (West Africa) has similar levels of genetic diversity as other populations in the introduced range. The diversity estimates for Senegal are in fact higher than those of other locations in the introduced range, which might indicate that flies moved to Senegal from the native range via natural movement as well as by human assistance rather than only by jump dispersal. Interestingly, Burgers Hall, a location within South Africa, also has a degree of genetic diversity similar to that of populations in the introduced range, whilst other locations in South Africa seem to form part of the native range. These results are further supported by STRUCTURE and the

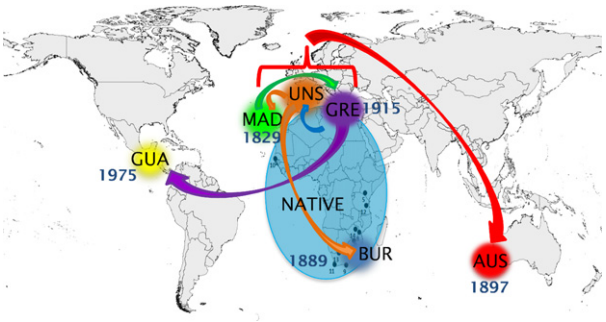

Figure 4 The world-wide route of invasion of *Ceratitis capitata*. Each coloured circle represents a sampling area in our data (GUA = Guatemala, MAD = Madeira, GRE = Greece, AUS = Australia, NATIVE = All African populations, UNS = Unsampled location, BUR = Burgers Hall), and the bracket indicates admixture between populations. Arrows indicate the most likely scenario hypothesized (Table 2) and is supported by posterior probabilities and 95% confidence intervals [0.998 (0.997, 0.999)]. Dates indicated are those available as the earliest record.

neighbour-joining trees which show that the Burgers Hall population groups closely with those from the introduced range. All populations showed deviations from HWE, and this can possibly be attributed to inbreeding, the Wahlund effect, as well as null alleles. All populations sampled showed some null alleles which can arise because of their polymorphic genome (e.g. Malacrida *et al.*, 2007). These levels of inbreeding indicate that non-random mating is occurring despite the high effective population sizes, which may well explain the observed deviations from HWE.

Furthermore, we found a clear pattern of genetic differentiation between the native African populations and those from the introduced range broadly forming two groups. The PCoA, STRUCTURE and NJ trees largely support the 'out-of-Africa' hypothesis, grouping all the African populations (Cluster 1) together as well as all the introduced locations elsewhere in the world including Burgers Hall (Cluster 2). Moreover, Cluster 2 was further structured into four groups: Australia, Madeira, Burgers Hall and a group including Greece and Guatemala. This group was subsequently further sub-divided into two separate groups, each group pertaining to a sampling location. These six additional clusters identified in Cluster 2 were supported by the pairwise F_{ST} values. The lack of differentiation throughout Africa indicates high levels of movement of *C. capitata* either by natural range expansion or via humans through corridors for movement supplied by the continuous trade of goods (formal and informal) or human travel. This high level of movement of flies is further aided by limited quarantine restriction within areas on the African continent that are not pest free. *Ceratitis capitata* is assumed to have travelled human-mediated to the Mediterranean coast (Maddison & Bartlett, 1989), and the reduction in genetic diversity here seems to support the idea of a small number of founders introduced to this region, and is further supported by results of the ABC analysis. The European populations (Greece) in our study also group with Guatemala, and although it has been proposed that Australia was a secondary colonization event from Europe (Hooper & Drew, 1989; Bonizzoni *et al.*, 2004), our data (STRUCTURE, NJ trees, ABC) indicate that the same is also true for Guatemala. The most likely scenario chosen based on ABC calculations closely matches the proposed invasion pathway for *C. capitata* with some additions (Fig. 4). First, the Mediterranean coast was colonized at an Unsampled location from the native range (Africa) after which Madeira was colonized and then *C. capitata* spread to Greece. These European locations formed an admixed population that secondarily colonized Australia. Guatemala was colonized from Greece, and Burgers Hall was recolonized from Europe (Unsampled location). To our knowledge, this is the first time that the invasion pathway (albeit for the majority but not all continents where the flies are known to occur) of *C. capitata* has been tested and confirmed using ABC. Moreover, we show that the recorded historical dates of first introduction are likely much more recent than those indicated by our ABC analysis (Appendix S7). This result therefore highlights the challenges of early detection to prevent new invasions.

Consequently, it seems reasonable to conclude that there are low levels of connectivity between the African continent and the introduced range except for the reintroduction into Burgers Hall from Europe, although high levels of connectivity exist on a regional scale. This information, in turn, is important for the management of *C. capitata* world-wide. This pattern of differentiation on a macrogeographic scale and a lack thereof on a more regional scale has also been evident in other tephritid studies including *Bactrocera cucurbitae* (Virgilio *et al.*, 2010) and *Bactrocera oleae* (Nardi *et al.*, 2005). The lack of connectivity indicates that quarantine measures for export consignments from Africa to elsewhere in the world are broadly successful in limiting the movement of fruit flies intercontinentally. Moreover, all locations sampled in the introduced range belong to separate genetic clusters, indicating a lack of movement between these locations. However, there is evidence for high levels of movement of fruit flies on the African continent. This does not bode well for the fruit industry on the African continent, and especially in the prevention of new invasions into a previously pest-free region. A case in point is that of *Bactrocera invadens* (*B. dorsalis*) which, after its introduction to the African continent in 2003 (Lux *et al.*, 2003), has spread over large parts of the continent despite extensive quarantine and eradication efforts (De Meyer *et al.*, 2010). Information from the reconstruction of the routes of invasion for *C. capitata* is important to understand the different evolutionary and environmental factors that influence successful invasions, and that can then be incorporated into strategies for the control and prevention of new invasions (Estoup & Guillemaud, 2010). Future studies should therefore focus on including more samples from native and introduced locations for *C. capitata* as well as higher numbers of molecular markers to infer a more

complete picture of the global invasion pathways and whether this is dynamic over time.

ACKNOWLEDGEMENTS

Microsatellite data were genotyped at the Central Analytical Facility, Stellenbosch University. We are grateful to farmers in South Africa, Luís Dantas, Walther Enkerlin, Ernie Steiner, Sunday Ekeshi, Luis Bota, Louisa Makumbe, Maulid Mwatawala, Nikos Papadopoulos, Ibrahim Solo and Pierre-Alexandre Rastorgueff. This research was funded by THRIP from the National Research Foundation (NRF) of South Africa [(UID) 71909], SAAPPA, SASPA, as well as SATGI. The Grantholder acknowledges that opinions, findings and conclusions or recommendations expressed in any publication generated by NRF supported research are that of the author (s), and that the NRF accepts no liability whatsoever in this regard. We are grateful for constructive comments and discussion of this work by Matthew Hill, Anthony Clarke, Jaco le Roux and the anonymous referees.

REFERENCES

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

Aketarawong, N., Guglielmino, C.R., Karam, N., Falchetto, M., Manni, M., Scolari, F., Gomulski, L.M., Gasperi, G. & Malacrida, A.R. (2014) 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. *Genetica*, **142**, 201–213.

Bebber, D.P., Ramotowski, M.A. & Gurr, S.J. (2013) Crop pests and pathogens move polewards in a warming world. *Nature Climate Change*, **3**, 985–988.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. (1996–2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Montpellier: Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II.

Berthouly-Salazar, C., Hui, C., Blackburn, T.M., Gaboriaud, C., Van Rensburg, B.J., Jansen van Vuuren, B. & Le Roux, J.J. (2013) Long-distance dispersal maximizes evolutionary potential during rapid geographic range expansion. *Molecular Ecology*, **22**, 5793–5804.

Blackburn, T.M., Pyšek, P., Bacher, S., Carlton, J.T., Duncan, R.P., Jarošík, V., Wilson, J.R.U. & Richardson, D.M. (2011) A proposed unified framework for biological invasions. *Trends in Ecology and Evolution*, **26**, 333–339.

Blackburn, T.M., Prowse, T.A., Lockwood, J.L. & Cassey, P. (2013) Propagule pressure as a driver of establishment success in deliberately introduced exotic species: fact or artefact? *Biological Invasions*, **15**, 1459–1469.

Bonizzoni, M., Zheng, L., Guglielmino, C.R., Haymer, D.S., Gasperi, G., Gomulski, L.M. & Malacrida, A.R. (2001) Microsatellite analysis of medfly bioinfestations in California. *Molecular Ecology*, **10**, 2515–2524.

Bonizzoni, M., Guglielmino, C.R., Smallridge, C.J., Gomulski, M., Malacrida, A.R. & Gasperi, G. (2004) On the origins of medfly invasion and expansion in Australia. *Molecular Ecology*, **13**, 3845–3855.

Bowcock, A.M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J.R. & Cavalli-Sforza, L.L. (1994) High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, **368**, 455–457.

Carey, J.R. (2010) The Mediterranean fruit fly (*Ceratitis capitata*). *American Entomologist*, **56**, 158–163.

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

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.

Coulon, A., Fitzpatrick, J.W., Bowman, R., Stith, B.M., Makarewich, C.A., Stenzler, L.M. & Lovette, I.J. (2008) Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (Aphelocoma cœrulescens). *Molecular Ecology*, **17**, 1685–1701.

De Breme, F. (1842) Note sur le genre *Ceratitis* de M. Mac Heay (Diptera). *Annales de la Société Entomologique de France*, **11**, 183–190.

De Meyer, M., Copeland, R.S., Wharton, R.A. & McPheron, B.A. (2002) On the geographic origin of the Medfly *Ceratitis capitata* (Weidemann) (Diptera: Tephritidae). Proceedings of the 6th International Fruit Fly Symposium, Stellenbosch, South Africa, pp. 45–53.

De Meyer, M., Robertson, M.P., Peterson, A.T. & Mansell, M.W. (2008) Ecological niches and potential geographical distributions of Mediterranean fruit fly (*Ceratitis capitata*) and Natal fruit fly (*Ceratitis rosa*). *Journal of Biogeography*, **35**, 270–281.

De Meyer, M., Robertson, M.P., Mansell, M.W., Ekesi, S., Tsuruta, K., Mwaiko, W., Vayssières, J.-F. & Peterson, A.T. (2010) Ecological niche and potential geographic distribution of the invasive fruit fly *Bactrocera invadens* (Diptera, Tephritidae). *Bulletin of Entomological Research*, **100**, 35–48.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, **39**, 1–38.

Dieringer, D. & Schlötterer, C. (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167–169.

Dixon, A.F., Honěk, A., Keil, P., Kotela, M.A.A., Šizling, A.L. & Jarošík, V. (2009) Relationship between the minimum

and maximum temperature thresholds for development in insects. *Functional Ecology*, **23**, 257–264.

Earl, D.A. & Von Holdt, B.M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

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.

Excoffier, L. & Lischer, H.E.L. (2010) Arlequin suite version 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Falush, D., Stephens, M. & Pritchard, J.K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Fimiani, P. (1989) Mediterranean region. Fruit flies: their biology, natural enemies and control. Elsevier Science Publishers.

Gasperi, G., Bonizzoni, M., Gomulski, L.M., Murelli, V., Torti, C., Malacrida, A.R. & Guglielmino, C.R. (2002) Genetic differentiation, gene flow and the origin of infestations of the medfly, *Ceratitis capitata*. *Genetica*, **116**, 125–135.

Gaston, K.J., Jones, A.G., Hänel, C. & Chown, S.L. (2003) Rates of species introduction to a remote oceanic island. *Proceedings of the Royal Society B*, **270**, 1091–1098.

Gilchrist, A.S. & Meats, A.W. (2009) The genetic structure of populations of an invading pest fruit fly, *Bactrocera tryoni*, at the species climatic range limit. *Heredity*, **105**, 165–172.

Gilchrist, G.W., Jeffers, L.M., West, B., Folk, D.G., Suess, J. & Huey, R.B. (2008) Clinal patterns of desiccation and starvation resistance in ancestral and invading populations of *Drosophila subobscura*. *Evolutionary Applications*, **1**, 513–523.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M. & Toulmin, C. (2010) Food security: the challenge of feeding 9 billion people. *Science*, **327**, 812–818.

Grout, T.G. & Stoltz, K.C. (2007) Developmental rates at constant temperatures of three economically important *Ceratitis* spp. (Diptera: Tephritidae) from Southern Africa. *Environmental Entomology*, **36**, 1310–1317.

Harris, E.J. (1989) Hawaiian islands and north America. Fruit Flies: Their Biology, Natural Enemies and Control. Elsevier Science Publishers.

Higgins, S.I. & Richardson, D.M. (2014) Invasive plants have broader physiological niches. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 10610–10614.

Hooper, G.H.S. & Drew, R.A.I. (1989) Australia and South Pacific islands. Fruit flies: their biology, natural enemies and control. Elsevier Science Publishers.

Hulme, P.E. (2009) Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology*, **46**, 10–18.

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*, **23**, 1801–1806.

Karsten, M., Van Vuuren, B.J., Barnaud, A. & Terblanche, J.S. (2013) Population genetics of *Ceratitis capitata* in South Africa: implications for dispersal and pest management. *PLoS ONE*, **8**, e54281.

Kimberling, D.N. (2004) Lessons from history: predicting successes and risks of intentional introductions for arthropod biological control. *Biological Invasions*, **6**, 301–318.

Kirk, H., Dorn, S. & Mazzi, D. (2013) Molecular genetics and genomics generate new insights into invertebrate pest invasions. *Evolutionary Applications*, **6**, 842–856.

Lance, D. & McInnis, D. (2005) Biological basis of the sterile insect technique. *Sterile insect technique. Principles and practice in area-wide integrated pest management* (ed. by V.A. Dyck, J. Hendrichs and A.S. Robinson), pp.69–94. Springer, Netherlands.

Leung, B., Lodge, D.M., Finnoff, D., Shogren, J.F., Lewis, M.A. & Lamberti, G. (2002) An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings of the Royal Society B*, **269**, 2407–2413.

Lockwood, J.L., Cassey, P. & Blackburn, T. (2005) The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution*, **20**, 223–228.

Lockwood, J.L., Cassey, P. & Blackburn, T.M. (2009) The more you introduce the more you get: the role of colonization pressure and propagule pressure in invasion ecology. *Diversity and Distributions*, **15**, 904–910.

Lux, S.A., Copeland, R.S., White, I.M., Manrakhan, A. & Billah, M.K. (2003) A new invasive fruit fly species from the *Bactrocera dorsalis* (Hendel) group detected in East Africa. *International Journal of Tropical Insect Science*, **23**, 355–361.

Maddison, P.A. & Bartlett, B.J. (1989) A contribution towards the zoogeography of the Tephritidae. Fruit Flies: their biology, natural enemies and control. Elsevier Science Publishers.

Malacrida, A.R., Marinoni, F., Torti, C., Gomulski, L.M., Sebastiani, F., Bonvicini, C., Gasperi, G. & Guglielmino, C.R. (1998) Genetic aspects of the worldwide colonization process of *Ceratitis capitata*. *Heredity*, **89**, 501–507.

Malacrida, A.R., Gomulski, L.M., Bonizzoni, M., Bertin, S., Gasperi, G. & Guglielmino, C.R. (2007) Globalization and fruitfly invasion and expansion: the medfly paradigm. *Genetica*, **131**, 1–9.

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

Nei, M., Tajima, F. & Tateno, Y. (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*, **19**, 153–170.

Nyamukondiwa, C., Kleynhans, E. & Terblanche, J.S. (2010) Phenotypic plasticity of thermal tolerance contributes to the invasion potential of Mediterranean fruit flies (*Ceratitis capitata*). *Ecological Entomology*, **35**, 565–575.

Papadopoulos, N.T. (2014) Fruit Fly Invasion: Historical, Biological, Economic Aspects and Management. In Trapping and the Detection, Control, and Regulation of Tephritid Fruit Flies. Springer.

Pascual, M., Chapuis, M.P., Mestres, F., Balanya, J., Huey, R.B., Gilchrist, G.W., Serra, L. & Estoup, A. (2007) Introduction history of *Drosophila* subobscura in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology*, **16**, 3069–3083.

Peakall, R. & Smouse, P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Peakall, R. & Smouse, P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics*, **28**, 2537–2539.

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

Pyšek, P., Richardson, D.M., Pergl, J., Jarošík, V., Sixtová, Z. & Weber, E. (2008) Geographical and taxonomic biases in invasion ecology. *Trends in Ecology and Evolution*, **23**, 237–244.

Raymond, M. & Rousset, F. (1995) An exact test for population differentiation. *Evolution*, **49**, 1280–1283.

Rey, O., Estoup, A., Vonshak, M., Loiseau, A., Blanchet, S., Calcaterra, L., Chifflet, L., Rossi, J.-P., Kergoat, G.J., Foucaud, J., Orivel, J., Maurice, L., Schultz, T. & Facon, B. (2012) Where do adaptive shifts occur during invasion? A multidisciplinary approach to unravelling cold adaptation in a tropical ant species invading the Mediterranean area. *Ecology Letters*, **15**, 1266–1275.

Reyes, A. & Ochando, M.D. (2004) Mitochondrial DNA variation in Spanish populations of *Ceratitis capitata* (Wiedemann) (Tephritidae) and the colonization process. *Journal of Applied Entomology*, **12**, 358–364.

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

Rousset, F. (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Simberloff, D. (2009) The role of propagule pressure in biological invasions. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 81–102.

Storey, J.D. (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B*, **64**, 479–498.

Szűcs, M., Melbourne, B.A., Tuff, T. & Hufbauer, R.A. (2014) The roles of demography and genetics in the early stages of colonization. *Proceedings of the Royal Society B*, **281**, 20141073.

Takezaki, N., Nei, M. & Tamura, K. (2014) POPTREEW: web version of POPTREE for constructing population trees from allele frequency data and computing some other quantities. *Molecular Biology and Evolution*, **31**, 1625–1637. doi:10.1093/molbev/msu093.

Van Kleunen, M., Weber, E. & Fischer, M. (2010) A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters*, **13**, 235–245.

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.

White, I.M. & Elson-Harris, M.M. (1994) *Fruit flies of economics significance: their identification and bionomics*. ACIAR, CAB International, Wallingford Oxon.

Wilson, J.R., Dormontt, E.E., Prentis, P.J., Lowe, A.J. & Richardson, D.M. (2009) Something in the way you move: dispersal pathways affect invasion success. *Trends in Ecology and Evolution*, **24**, 136–144.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Schematic representations of the eighteen different scenarios tested using DIYABC v2.03 (Cornuet *et al.*, 2014).

Appendix S2 Definitions and prior distributions of parameters used in the testing of eighteen invasion scenarios of *Ceratitis capitata* using the ABC method implemented in DIYABC v2.03 (Cornuet *et al.*, 2014).

Appendix S3 Frequency of null alleles calculated according to Dempster *et al.* (1977) in FREENA 1.0 (Chapuis & Estoup, 2007) for 323 *Ceratitis capitata* individuals.

Appendix S4 Pairwise F_{ST} values calculated in MICROSATELLITE ANALYSER v4.05 (Dieringer & Schlötterer, 2003) between the six genetic clusters [Australia, Greece, Guatemala, Madeira, Burgers Hall, Africa (all other African locations)] inferred from Bayesian Clustering for *Ceratitis capitata*.

Appendix S5 The lines indicate the most likely number of clusters (K) in STRUCTURE of *Ceratitis capitata* populations according to the natural logarithm of the likelihood (Ln P (X/K) (Pritchard *et al.*, 2000) and the delta K method (Evanno *et al.*, 2005).

Appendix S6 The most likely scenario (3) making use of approximate Bayesian computations (ABC) in DIYABC v2.03 (Cornuet *et al.*, 2014).

Appendix S7 Means of the posterior probabilities for demographic parameters estimated under Scenario 3.

BIOSKETCH

Minette Karsten is a PhD student at the University of Stellenbosch, South Africa. Her interests include population genetics, invasion biology and the management of fruit flies. She is part of two larger research groups, an Integrated Pest Management group and the Applied Physiological Ecology Lab (http://insectphysiologicalecology.blogspot.com/).

Author contributions: J.S.T. and M.K. conceived and designed the experiments. M.K. performed the experiments. J.S.T., B.J.V.V. and P.A. contributed reagents/materials/analysis tools. M.K. wrote the first draft of manuscript, and all authors contributed to interpretation of results and helped writing the manuscript.

Editor: Brian Leung