Reconstructing demographic events from population genetic data: the introduction of bumblebees to New Zealand

G. C. LYE, O. LEPAIS and D. GOULSON
School of Biological and Environmental Sciences, University of Stirling, Stirling, FK9 4LA, UK

Abstract

Four British bumblebee species (*Bombus terrestris*, *Bombus hortorum*, *Bombus ruderatus* and *Bombus subterraneus*) became established in New Zealand following their introduction at the turn of the last century. Of these, two remain common in the United Kingdom (*B. terrestris* and *B. hortorum*), whilst two (*B. ruderatus* and *B. subterraneus*) have undergone marked declines, the latter being declared extinct in 2000. The presence of these bumblebees in New Zealand provides an unique system in which four related species have been isolated from their source population for over 100 years, providing a rare opportunity to examine the impacts of an initial bottleneck and introduction to a novel environment on their population genetics. We used microsatellite markers to compare modern populations of *B. terrestris*, *B. hortorum* and *B. ruderatus* in the United Kingdom and New Zealand and to compare museum specimens of British *B. subterraneus* with the current New Zealand population. We used approximate Bayesian computation to estimate demographic parameters of the introduction history, notably to estimate the number of founders involved in the initial introduction. Species-specific patterns derived from genetic analysis were consistent with the predictions based on the presumed history of these populations; demographic events have left a marked genetic signature on all four species. Approximate Bayesian analyses suggest that the New Zealand population of *B. subterraneus* may have been founded by as few as two individuals, giving rise to low genetic diversity and marked genetic divergence from the (now extinct) UK population.

Keywords: approximate Bayesian computation, conservation, invasive species, museum specimens, population bottleneck, re-introduction

Received 8 October 2010; revision received 18 April 2011; accepted 22 April 2011

Introduction

Understanding the genetic processes associated with species relocation is fundamental to the study of conservation biology (Frankham *et al.* 2004). To assess the likely outcomes of species relocations, it is important to understand longer-term effects on the genetic diversity and structure of introduced populations. One way to do this is to investigate these in relation to known historic introductions. Whilst the exact circumstances of such introductions are often unknown, modern population genetic techniques allow not only investigation of dif-

ferences in genetic structure and diversity among source and introduced populations, but also allow inference of the demographic histories of these populations (Lopes & Boessenkool 2010). This enables the reconstruction of past introduction events. The use of museum specimens in conservation genetic studies provides an additional tool for the reconstruction of historic events, allowing assessment of the genetic composition of populations prior to important demographic events (Wandeler *et al.* 2007). Here, these techniques are employed to investigate changes in genetic diversity and structure and to estimate the demographic events associated with the introduction of British bumblebee species into New Zealand.

Correspondence: G. C. Lye, Fax: +44 1786 467843;
E-mail: gillian.lye@gmail.com

British bumblebees were introduced into South Island, New Zealand, at the turn of the last century for the pollination of the fodder crop, *Trifolium pratense* (Hopkins 1914). Four species became established (*Bombus terrestris*, *Bombus hortorum*, *Bombus ruderatus* and *Bombus subterraneus*), and these still persist in New Zealand today. Following their introduction, these species spread rapidly across the South Island and by 1965 all but *B. subterraneus* were also present in the North Island (MacFarlane & Gurr 1995).

Current New Zealand populations of British bumblebee species derive from two successful introduction attempts consisting of 93 bumblebee queens in 1885 and a further 143 bumblebee queens in 1906. However, the specific identity of these queens is unknown. It is believed that at least six species of bumblebee were included in the 236 bumblebee queens brought to New Zealand (Hopkins 1914), suggesting that the founder populations of each species were very small. In addition, there is likely to have been high initial mortality of transported queens after release, further reducing the number of individuals contributing to the current New Zealand populations.

Severe population bottlenecks such as those presumably experienced by New Zealand bumblebee populations can lead to a number of deleterious genetic effects. A bottleneck event inevitably results in the loss of genetic diversity, and this initial loss of genetic variation is likely to result in a reduced ability of the population to adapt to environmental change. Small populations are also more susceptible to genetic drift, which can cause chance fixation of deleterious, or loss of beneficial alleles from the population, and to inbreeding depression (Frankham *et al.* 2004).

The presence of British bumblebees in New Zealand provides an unique opportunity to investigate the genetic effects of an initial bottleneck event followed by over one hundred years of isolation in four replicate species. Whilst these species are closely related, current knowledge suggests that each species is likely to have a different demographic history. At the time of their introduction into New Zealand, *B. terrestris*, *B. hortorum* and *B. ruderatus* were all common in England and *B. subterraneus* was also described as abundant in many localities in the south (Sladen 1912). Today, *B. terrestris* and *B. hortorum* remain common and ubiquitous throughout the United Kingdom, but *B. ruderatus* and *B. subterraneus* have both suffered severe declines and *B. subterraneus* was declared extinct in the United Kingdom in 2000 (Edwards & Jenner 2005). The genetic consequences of the bumblebee introductions to New Zealand are of particular relevance in the light of a current collaborative project led by British conservation organisations seeking to re-introduce *B. subterraneus* from New Zealand into the United Kingdom. The ability of the New Zealand population to re-adapt to the conditions in the United Kingdom is crucial for the success of the re-introduction project because the population is likely to have become adapted to different environmental conditions and will have experienced relaxed selection for defences against natural enemies, many of which do not occur in New Zealand.

The magnitude of the population bottleneck experienced by New Zealand bumblebee populations and its impacts on the genetic structure and diversity of these populations are largely unknown. Recent data presented by Schmid-Hempel *et al.* (2007) suggest that New Zealand populations of *B. terrestris* exhibit similar levels of genetic diversity to, but significant differentiation from, the UK population. *B. terrestris* has always been extremely abundant in England so it is likely that this species may have represented a reasonably large proportion of the surviving queens introduced into New Zealand. The founding populations of the other three species are likely to have been smaller.

Here, we use molecular markers to compare the genetic diversity and structure of current British and New Zealand populations of *B. terrestris*, *B. hortorum* and *B. ruderatus* to study the genetic effects of a population bottleneck followed by approximately 110 generations of isolation. We compare the current New Zealand population of *B. subterraneus* with the original British population, represented by museum specimens, in an attempt to assess the divergence of the genetic structure of this population from the source population. *Bombus subterraneus* of Swedish origin were also genotyped to provide a comparison with a current European population. Finally, we use approximate Bayesian computation (ABC) to estimate the demographic parameters of the four bumblebee species and provide an estimate of the number of founder queens that gave rise to the current New Zealand bumblebee populations.

Methods

Sample collection

Nonlethal tarsal clips (Holehouse *et al.* 2003) were collected from live workers or queens of *Bombus terrestris*, *Bombus hortorum*, *Bombus ruderatus* and *Bombus subterraneus* in the MacKenzie District of New Zealand (within a radius of 40 km from S 44°12′49″, E 170°12′20″) and from *B. terrestris*, *B. hortorum* and *B. ruderatus* in the south of England in the summers of 2003 (within a 30 km radius of N 52°08′19″, W 01°06′20″ for *B. hortorum* and *B. ruderatus* and of N 51°48′56″ E 00°21′38″ for *B. terrestris*). Sample sizes of *B. hortorum* and *B. ruderatus* in England and New Zealand, and

of *B. subterraneus* in New Zealand were supplemented by additional collections in the summer of 2007. The maximum distance between sampling points for any given species was 60 km. Presumably, as a result of the high dispersal ability of male and queen bumblebees (Kraus *et al.* 2009; Lepais *et al.* 2010), bumblebee populations usually show very low genetic differentiation across continuous landmass areas (for example in Europe: Estoup *et al.* 1996; Widmer & Schmid-Hempel 1999; and in introduced *B. terrestris* in New Zealand and Tasmania: Schmid-Hempel *et al.* 2007). In comparison, our samples were collected from within a small enough area to exclude the possibility of sampling from multiple demes. Tarsal clips from individuals of the original British population of *B. subterraneus* were taken from dried workers or queens held at the Museum of Natural History in Oxford. All museum specimens sampled originated from the south of England but attributed to the low numbers available, dates of collection associated with individuals sampled ranged from 1940 to 1965. An additional sample consisting of workers and queens of *B. subterraneus* collected from the Uppland province of Sweden (within a 35 km radius of N 59°53′03″, E 17°55′45″) in the summers of 2007 and 2008 was also analysed.

All samples were preserved in 100% ethanol. Sample sizes are presented in Table 1.

Molecular techniques

DNA was extracted from fresh bees using the HotShot protocol (Truett *et al.* 2000). However, this protocol was inadequate for the extraction of DNA from museum specimens, so the QIAGEN QIAamp DNA Micro kit was employed for DNA extraction from these individuals.

All bees were genotyped at eight microsatellite loci (B100, B132, B11, B10, B96, B126, B124 and B121) using primers developed by Estoup *et al.* (1995, 1996). Amplification at these loci was achieved using the QIAGEN Multiplex PCR kit. PCRs were 10 µL in volume and consisted of 1 µL Q-solution, 5 µL PCR Master Mix, 1 µL primer solution (0.2 µM of each primer, forward primers labelled with NED, HEX or 6-FAM dyes, Applied Biosystems; multiplex set 1: B100-HEX, B132-NED, multiplex set 2: B11-6-FAM, B10-HEX, B96-NED, multiplex set 3: B126-6-FAM, B124-HEX, B121-NED), 1 µL template DNA (of variable concentration dependent on the extraction technique used) and 2 µL HPLC H$_2$O. Samples were denatured at 95 °C for 15 min, and this was followed by 34 cycles consisting of the following: a denaturing step at 94 °C for 30 s, an annealing step at 49 °C for 90 s and an extension step at 72 °C for a further 90 s. This was then followed by a final extension step at 72 °C for 10 min. An ABI PRISM 377 semi-automated slab gel sequencer was used to visualize PCR products, and fragment size was determined using an internal size standard (GeneScan ROX 350; Applied Biosystems). Fragments were scored using Genotyper (Applied Biosystems). Samples for which amplification was not successful, or scoring was uncertain, were rerun, and re-extraction of DNA was carried out if necessary. For all museum specimens, the amplification procedure was repeated twice and data were compared between amplifications to test for the consistency of scoring. If genotypes were not scored consistently, the

Table 1 Raw sample sizes, colonies represented in each sample (as detected by analysis of data using Colony (Wang, 2004)) and final sample sizes of bumblebees of English, New Zealand and Swedish origin for genetic analysis

Species	Location	Year	Sample size	Colonies represented	Final sample size
Bombus ruderatus	England	2003	33	24	28
	England	2007	4	4	
	New Zealand	2003	16	14	54
	New Zealand	2007	81	40	
Bombus hortorum	England	2003	19	18	46
	England	2007	31	28	
	New Zealand	2003	30	28	37
	New Zealand	2007	9	9	
Bombus terrestris	England	2003	209	141	141
	New Zealand	2003	66	56	56
Bombus subterraneus	England	1940–1965	58	41	41
	New Zealand	2003	44	24	38
	New Zealand	2007	25	14	
	Sweden	2007	17	13	46
	Sweden	2008	35	33	

individual was discarded. Individuals were also removed from the data set if amplification failed at more than three loci, because level of genetic degradation within these individuals was likely to be high (Lozier & Cameron 2009).

Data analysis

Genetic diversity and population genetic structure. Data sets were checked for unexpected mutation steps, large gaps in fragment lengths and unusually sized fragments using MSA version 4.05 (Dieringer & Schlotterer 2003). Colony version 2.0.0.1 (Jones & Wang 2009) was then used to identify sister pairings within each time period, species and population, accounting for a genotyping error of 0.5% at each locus. For each sisterhood identified, all but one individual was removed from the data set prior to further analysis. Final sample sizes used for analyses are given in Table 1. Genetic differentiation between samples collected in different years at the same locations was assessed for each species by the calculation of Weir and Cockerham's estimator of F_{st} (θ). Significance was determined following 10 000 allele permutations implemented in MSA. Deviations from Hardy–Weinberg equilibrium and linkage disequilibrium between loci were tested for using Genepop version 4.04 (Rousset 2008). To minimize type I errors, strict sequential Bonferroni corrections were applied.

Genetic diversity within populations was assessed by means of allelic richness and Nei's unbiased measure of gene diversity, calculated for each species and population at each locus using Fstat version 2.9.3 (Goudet 1995). Wilcoxon signed-rank tests were used to assess differences in allelic richness and gene diversity among populations for each species, with the exception of *B. subterraneus* for which a Friedman test was employed as more than two populations were considered. These analyses were implemented in R (R Development Core Team 2009), with Wilcoxon tests utilizing the 'coin' package (Hothorn et al. 2008). Wright's measure of population differentiation, F_{st}, was used to assess genetic differentiation between New Zealand and British populations for each species (Wright 1951). These were calculated in Fstat according to the Weir & Cockerham (1984) estimator (θ). Global θ values were calculated for all species, and means and standard deviations were calculated by jack-knifing over loci. Pairwise θ values were also calculated for all combinations of the three populations of *B. subterraneus* sampled. A permutation procedure (10 000 allele permutations) was employed to test for the departure of global and pairwise θ values from 0 using MSA. As F_{st} estimates are dependent on levels of genetic variation displayed at the markers used, these values cannot be used to make comparisons between species. Global values for the standardized measure G'_{ST} were therefore also calculated (following Hedrick 2005). To allow direct comparison among populations of *B. subterraneus*, pairwise values of G'_{ST} were calculated for the English and New Zealand populations and for the English and Swedish populations.

Demographic parameter estimation. Approximate Bayesian computation implemented in the software DIYABC version 1.04.33beta (Cornuet et al. 2008, 2010) was used to estimate the effective number of individuals of each of the four bumblebee species introduced into New Zealand and the current effective population sizes of each species in the United Kingdom and New Zealand. A common introduction scenario was used for all species: An effective number of queens (*Nb*) taken from an UK population of effective population size *N1* was introduced into New Zealand *T* generations ago. After a period of *db* generations, the effective population size in New Zealand increased to the current population size, *N2* (Fig. 1a).

For the genetic parameters, the Generalized Stepwise Mutation model (Estoup et al. 2002) was used to simulate mutations at the microsatellite loci. The microsatellites used in this study are interrupted repeats (several repeated regions separated by nonrepeated regions), with some of them being characterized by a small number of repeats (Estoup et al. 1995, 1996). It has been shown that such microsatellites typically show a lower mutation rate than perfect short tandem repeats (Viard et al. 1998). For instance, Yokoyama et al. (2004) estimated a slippage mutation rate of about $6–7 \times 10^{-6}$ mutations per generation for a 4 bp repeat for locus B11 in *Bombus diversus*. As mutation rate priors are likely to have a strong influence on estimates of effective population size, three models using different prior interval specifications for the mean microsatellite mutation rate across loci (μ_{mic}) were tested. Models M1, M2 and M3 include a low ($\mu_{mic} \sim U[10^{-6}–10^{-5}]$), a medium ($\mu_{mic} \sim U[10^{-5}–10^{-4}]$) and a high ($\mu_{mic} \sim U[10^{-4}–10^{-3}]$) mean mutation rate prior interval over loci. In all cases, the parameter P_{mic} of the geometric distribution modelling the number of repeat motifs added or removed from the microsatellite in each mutation step was specified with a uniform prior distribution bounded between 0.1 and 0.3 ($P \sim U[0.1–0.3]$). Individual loci mutation parameters were allowed to vary following a hierarchical scheme and using Gamma distributions ($\mu_{mic} \sim G[10^{-7}–10^{-4}, 2]$ for model M1, $\mu_{mic} \sim G[10^{-6}–10^{-3}, 2]$ for model M2, $\mu_{mic} \sim G[10^{-5}–10^{-2}, 2]$ for model M3 and $P_{mic} \sim G[0.01–0.9, 2]$ for all models) allowing heterogeneity in mutation rate and model between loci to be accounted for, as is essential in the case of interrupted microsatellites. Finally, single-nucleotide mutations in
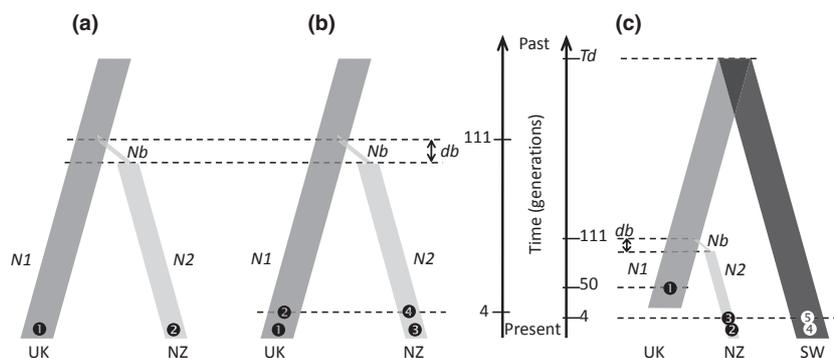
Fig. 1 Schematic representation of the introduction scenario modelled in DIYABC for (a) *Bombus terrestris*, (b) *Bombus ruderatus* and *Bombus hortorum* and (c) *Bombus subterraneus*. Parameters estimated are indicated in italics ($N1$: British effective population size, $N2$: New Zealand effective population size, $N3$: Swedish effective population size, Nb: effective number of founders, db: duration of bottleneck, Td: divergence time between Swedish and British populations). Circled numbers represent population samples (numbers refer to population number in the genetic data set submitted to DRYAD (doi:10.5061/dryad.sk22v). The interruption of the symbol representing the British *B. subterraneus* population indicates population extirpation. Time is expressed in number of generations. Times and population sizes are not to scale.

the microsatellite sequence, which are likely to have happened in interrupted microsatellites, were accounted for by the parameter μ_{sni} ($\mu_{sni} \sim G[10^{-9}–10^{-3}, 2]$).

Model demographic parameter priors were defined based on historical knowledge of the introductions. As bumblebees were introduced into New Zealand in two steps, the model was constructed to incorporate two distinct introduction events for each species. However, preliminary power analysis based on simulated data sets indicated that it would not be possible to produce accurate parameter estimates from a two-step introduction scenario (results not shown). We therefore kept our introduction model very simple, including only one introduction event at the mid-point between the two known introduction dates ($T\sim111$ generations ago, Hopkins 1914; MacFarlane & Griffin 1990). These bumblebee species have an annual life cycle (Goulson 2010) so it was assumed that a generation is equal to 1 year. In addition, as no more than 236 queens of at least six different species were introduced (Hopkins 1914), we defined an uniform prior bounded by 1 and 100 for the number of introduced queens for each species ($Nb \sim U$ [1–100]). This prior was selected because, whilst proportionate representation by each species is expected to be different, there is no reason to believe that collection in the United Kingdom was very strongly biased towards a single species and also because the likelihood of successful establishment following release is low making it very unlikely that more than 100 individuals from any one species managed to found successful colonies in New Zealand. Bumblebees in New Zealand were reported to demonstrate rapid population growth following their release and were observed as far as 100 miles away from the introduction points a year after introduction (Hopkins 1914). Similar rapid popula-

tion expansion has also been observed following the more recent introduction of bumblebees to Tasmania (Hingston *et al.* 2002). We thus assumed a very short duration of the bottleneck period in our model with an uniform prior bounded by 1 and 5 generations ($db \sim U$ [1–5]). As this parameter is not of primary interest, its estimation is not discussed. However, it is included in the model to account for uncertainty regarding the duration of the bottleneck. As the number of introduced queens (Nb) may partly depend on the duration of the bottleneck (db), the effect of prior specification of db on the estimate of Nb was further investigated by fixing db at one generation (model M4) and five generations (model M5). Furthermore, to get a comparable estimation of the strength of the bottleneck of the different species during the introduction, we computed a composite parameter, bottleneck intensity, expressed as $\log10(db/Nb)$ in the case of model M2 (medium mutation rate prior). This parameter accounts for the potential dependence of db on Nb. As scant information is available on the likely effective population sizes of each species in each location, broad priors were used. An uniform prior bounded by 10 and 100 000 was used for $N1$ and $N3$ (the effective population size of the Swedish population, included only in the model for *B. subterraneus*) and by 10 and 50 000 for $N2$.

The basic model was modified for each species to account for differences in sampling. *B. ruderatus* and *B. hortorum* were sampled in two different years, and this was incorporated into the model for these species (Fig. 1b). For *B. subterraneus*, UK museum samples originating from different collections could not be individually dated so the model was built around a sampling date of 50 generation ago, the average between known sampling dates (Fig. 1c). In addition, the

Swedish population (whose effective size is $N3$) was considered to have diverged from the UK populations Td generations ago (Fig. 1c) with a uniform prior distribution bounded by 10 and 30 000 generations ($Td \sim$ [10–30 000], Table 1).

For each species, 4 000 000 genetic data sets were simulated using the coalescent approach implemented in DI-YABC with parameters drawn from their prior distributions. The observed genetic data set and each simulated genetic data set were summarized using the following summary statistics: mean number of alleles, mean gene diversity (Nei 1987), mean allele size variance, mean M index (Excoffier et al. 2005) across loci computed for each population and for each pair of populations: F_{st} (Weir & Cockerham 1984), $(\sigma\mu)^2$ distance (Goldstein et al. 1995) and shared allele distance (Chakraborty & Jin 1993) between pairs of populations. The normalized Euclidian distance between the observed and simulated sets of summary statistics was then computed, and the 4000 closest simulated data sets (0.1%) were selected to estimate the posterior distribution of the parameter using a local linear regression technique (Beaumont et al. 2002). Estimations using the 1000, 20 000 or 40 000 closest simulations produced similar results (not shown). The *locfit* package (Loader 1996) implemented in R (R Development Core Team 2009) was used to plot the posterior density distribution of the demographic parameters of primary interest ($N1$, $N2$ and Nb).

When using the ABC framework, analysis of model and prior specifications and performance of estimation is an important step to ensure the reliability of conclusions (Bertorelle et al. 2010). The goodness of fit of the models was first checked by establishing that genetic data generated using parameters drawn from the posterior distributions were similar to the observed genetic data (Gelman & Meng 1996). The discrepancies between summary statistics of the observed data set and the same summary statistics obtained from 1000 simulated data sets were tested by computing the proportion of simulated data sets with corresponding summary statistics lower than those of the observed data set (Gelman & Meng 1996).

The performance of parameter estimations was then assessed by generating 1000 test data sets simulated based on known parameter values ('true values') drawn from the priors. Previously, simulated data sets were then used to estimate the parameters of these pseudo-observed data sets, and the estimations were compared with the true values to assess the bias and precision of the estimations (Cornuet et al. 2008). Using the mode as a point estimate, the relative bias, the relative root mean square error, the 50% and 95% coverage (the proportion of simulations in which the 'true' value lies within the 50% and 95% credible interval around the estimate)

and the factor 2 statistic (the proportion of estimated values falling within the interval between 50% and 200% of the 'true' value) were calculated.

Results

Deviation from Hardy–Weinberg equilibrium and linkage disequilibrium

Following removal of sister pairs from the microsatellite data, no significant deviation from Hardy–Weinberg equilibrium was found for any locus, nor was there any evidence for linkage disequilibrium among any locus pairs for *Bombus terrestris* and *Bombus ruderatus.* A significant deviation from Hardy–Weinberg was found at the B100 locus within *Bombus hortorum*, although this was only apparent within the New Zealand population. Further investigation demonstrated that this deviation was attributable to heterozygote deficit. For *Bombus subterraneus*, no deviation from Hardy–Weinberg was found at any locus within the New Zealand sample. Five of the eight microsatellite loci were out of Hardy–Weinberg equilibrium within the British sample of *B. subterraneus* because of heterozygote deficit. Swedish individuals also demonstrated significant deviation from Hardy–Weinberg equilibrium at B96 and B121, again as a result of heterozygote deficit at these loci. Significant linkage disequilibrium was detected between B100 and B11 in the British sample of *B. subterraneus*, and between B132 and B11 in the New Zealand sample. All further analyses on both *B. hortorum* and *B. subterraneus* were conducted with and without problematic loci but because differences between corresponding analyses were negligible, statistics presented are those calculated across all loci.

Genetic diversity

Allelic richness was lower in New Zealand than in the United Kingdom for all species except *B. terrestris* in which no difference in allelic richness was found (Fig. 2a). Similarly, gene diversity was lower in the New Zealand than the UK populations of *B. hortorum* and *B. subterraneus*. However, no significant difference in gene diversity was found for *B. terrestris* or *B. ruderatus* (Fig. 2b). The allelic richness and gene diversity found within the Swedish population of *B. subterraneus* appeared to be similar to those found within the original UK population of this species (Fig. 2).

Population genetic structure

No genetic differentiation was found among population samples collected at the same localities in different
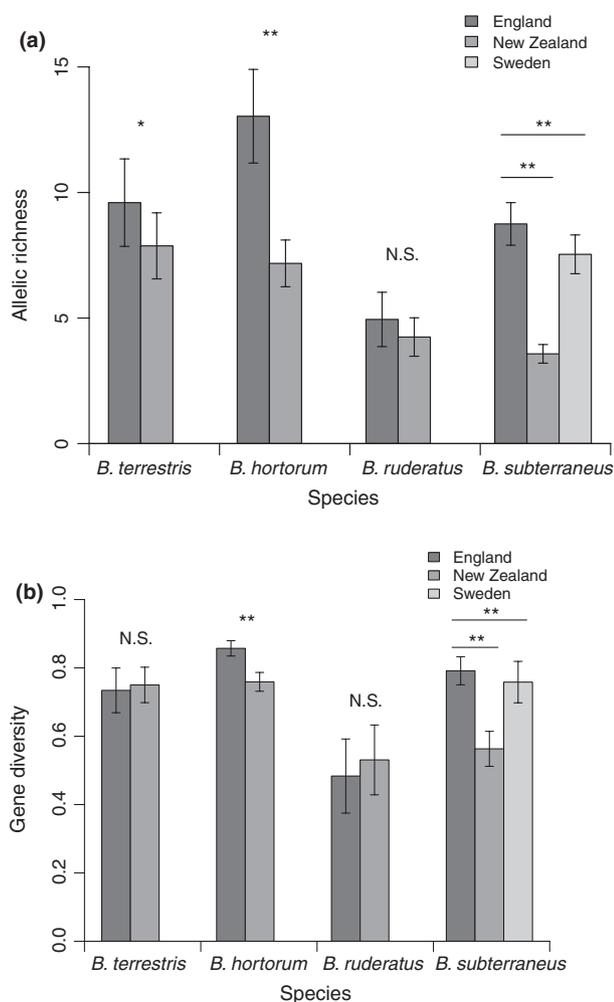
Fig. 2 Average allelic richness (a) and gene diversity (b) across eight microsatellite loci in New Zealand and UK populations of *Bombus terrestris*, *Bombus hortorum* and *Bombus ruderatus* and in New Zealand, United Kingdom and Swedish populations of *Bombus subterraneus* (±standard error). Calculated based on a minimum sample of 55, 34, 26 and 19 individuals, respectively. Results of statistical comparisons are shown (N.S., not significant, *$P < 0.05$, **$P < 0.001$). An additional pairwise comparison was used to demonstrate differences among the English and New Zealand population of *B. subterraneus*.

years for any species (results not shown). Population differentiation between United Kingdom and New Zealand populations of *B. terrestris* was relatively low although significant ($\theta = 0.019 \pm 0.004$, $P < 0.001$), probably attributed to the large number of alleles represented at each locus. Relatively, moderate differentiation was found among New Zealand and UK populations of *B. hortorum* and *B. ruderatus* ($\theta = 0.07 \pm 0.01$, $P < 0.001$ and $\theta = 0.083 \pm 0.025$, $P < 0.001$, respectively). For *B. subterraneus*, global θ among populations

was relatively high and significant ($\theta = 0.197 \pm 0.031$, $P < 0.001$) suggesting high levels of genetic differentiation between the three populations. Pairwise comparisons revealed that differentiation between Britain and New Zealand and also between Sweden and New Zealand is comparatively high ($\theta = 0.256$, $P < 0.001$ and $\theta = 0.225$, $P < 0.001$, respectively) but that differentiation between Sweden and the United Kingdom is relatively moderate ($\theta = 0.113$, $P < 0.001$).

Calculation of global G'_{ST} confirmed the patterns suggested above. *B. terrestris* demonstrates the lowest differentiation among populations of the four species included in the study ($G'_{ST} = 0.06$), and *B. ruderatus* also shows comparatively low population differentiation ($G'_{ST} = 0.14$). Differentiation between the populations of *B. hortorum* was higher than that of both *B. terrestris* and *B. ruderatus* ($G'_{ST} = 0.35$), but *B. subterraneus* shows by far the greatest population differentiation of the four species ($G'_{ST} = 0.75$), with this species demonstrating extremely high levels of differentiation between British and New Zealand populations. The Swedish population of *B. subterraneus* is also relatively highly differentiated from the British population ($G'_{ST} = 0.47$), but differentiation is much lower than between that of the British and New Zealand populations.

Demographic parameter estimation using ABC

Goodness-of-fit tests found no statistical difference between observed summary statistics and summary statistics simulated from posterior distributions of the parameters for the introduction models for *B. ruderatus*, *B. hortorum* and *B. terrestris* (result not shown). Discrepancies in 2 of 100 summary statistics after Bonferroni correction were found for *B. subterraneus*. The accuracy and precision of the parameter estimations using the model M2 (medium mutation rate priors) varied between parameters (Table S1, Supporting information). Notably, effective population size in New Zealand (*N2*) was usually poorly estimated as it showed high bias and relative mean square error leading to a low proportion of estimations within the interval 50–200% of the true value (Factor 2 ~30% only, Table S2, Supporting information). Otherwise, except μ_{sni} which is a nuisance parameter and those estimate is not of interest here, all other parameters are reasonably well estimated, especially for the demographic parameters (*N1*, *N2* and *N3*).

The numerical point estimates of the demographic and genetic parameters for all tested models are listed in Table S2 (Supporting information). The posterior distributions of the demographic parameters of primary interest are shown in Fig. 3. The prior specification of
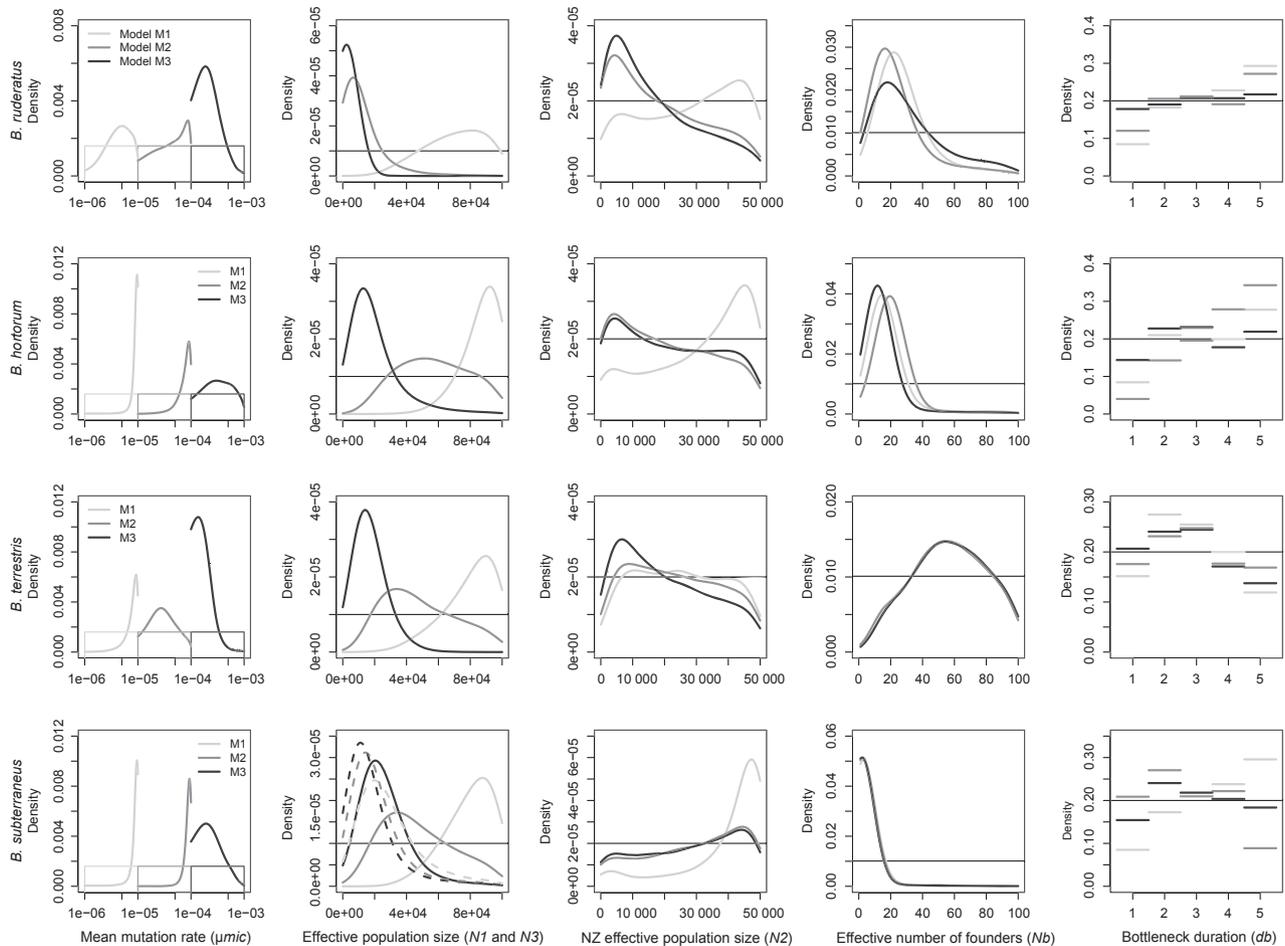
Fig. 3 Parameter posterior density estimates and effect of the microsatellite mutation rate prior specification. Light grey: model M1 with low microsatellite mutation rate, medium grey: model M2 with medium microsatellite mutation rate and dark grey: model M3 with high microsatellite mutation rate. Horizontal thin lines represent the prior distribution, the thick lines represented the posterior distribution and the dashed thick lines corresponded to *Bombus subterraneus* Swedish effective population size (*N3*). Mean mutation rate x axis is represented with a log scale.

the mutation rate had a clear impact on the estimation of UK effective population size (*N1*) and to a lesser extend Swedish and New Zealand effective population sizes (*N3* and *N2*, Fig. 3). A low mutation rate prior (model M1) led to high estimations of these parameters (Fig. 3). Differences in parameter estimates between the medium and the high mutation rate priors (model M2 and M3) were restricted to *N1*, which is lower with high mutation rate priors. Similar estimates for the other parameters in these two models may be because of the fact that mutation rate posterior tended to converge towards a similar value (in the region of 10^{-4} mutations by generation), with the exception of *B. terrestris* whose mutation rate seemed to be lower (Fig. 3). Finally, estimation of the number of founders (*Nb*) is relatively unaffected by the prior specification of the mutation rate (Fig. 3).

The prior specification of *db* had only a limited impact on parameter estimation (Fig. S1, Supporting information). Comparison of the estimates with the model M2 (*db* variable between 1 and 5 generations) and the models M4 and M5 (*db* fixed to 1 and 5 generations respectively) only showed noticeable differences for the estimation of *Nb* in *B. terrestris*, which varied between 19 [14–78] (mode [5–95% quantiles]) for model M4 to 80 [58–98] for model M5. Note that estimation prior interval of *db* in model M2 seemed to integrate this uncertainty as shown by the wider confidence interval of the estimate (57 [16–93], Table S1, Supporting information).

Given these results, we used the model M2 with a medium prior interval for the mutation rate to estimate the demographic parameters (Table S1, Supporting information).

Effective UK population size was estimated at 3990 [2050–36 100] for *B. ruderatus*, 33 700 [17 800–90 400] for *B. terrestris* and 57 100 [22 700–93 000] for *B. hortorum* (Table S1, Supporting information). The effective population size of *B. subterraneus* in the United Kingdom before extinction, 37 600 [16 600–89 200] appeared to be higher than the present effective population size in Sweden (11 400 [5330–56 110]) although this may be because of the pooling of UK museum individuals from different collection years. The divergence time between the British and Swedish populations was estimated at 7170 [2450–22 600] years (Table S1, Supporting information), which could correspond to the end of northward expansion of the species at the last glaciation period when the Swedish and British populations became effectively isolated from each other. Estimates of the effective number of individuals introduced into New Zealand varied between 2 [1–5] individuals for *B. subterraneus* and 58 [16–93] individuals for *B. terrestris* with intermediate values of 14 [6–67] and 19 [10–39] individuals, respectively, for *B. ruderatus* and *B. hortorum* (Fig. 3, Table S1, Supporting information). The current effective population sizes in New Zealand were estimated to be generally lower than those of the United Kingdom, but these estimates were found to be overestimated and inaccurate (Table S1, Supporting information). Finally, the bottleneck intensity during the introduction in New Zealand seems to have been particularly severe for *B. subterraneus*, strong for *B. ruderatus* and *B. hortorum* and relatively moderate for *B. terrestris* (Fig. 4).

Discussion

The introduction history of bumblebees in New Zealand is largely unknown because of poor documentation regarding the specific identity of the individuals that were released. However, advances in population genetic techniques allow retrospective assessment of demographic parameters associated with introduction events. An understanding of the dynamics of introduction events is important in the study of invasive species, with implications for the purposeful or accidental introduction of species outside of their native ranges. Additionally, population genetic analysis can be used to determine parameters that may shed light on the suitability of the New Zealand population of *Bombus subterraneus* for the current re-introduction programme for this species in the United Kingdom.

Limitations of the methodology used

An inherent difficulty with museum samples is the typically low sample size available and the mixed temporal
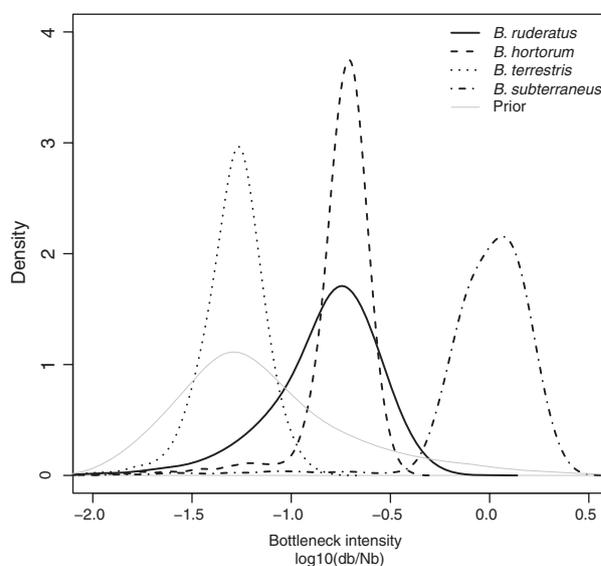
Fig. 4 Bottleneck intensity during the introduction of the four bumblebee species in New Zealand.

and geographical origin of the specimens. We chose to pool all museum specimens to obtain a useful number of representatives of the *B. subterraneus* UK population before it became extinct, with the drawback of a potential Wahlund effect which may account for the high number of alleles of this sample and the observed heterozygote deficit. In addition, pooling samples collected from temporally isolated demes can result in patterns usually associated with population expansion (Depaulis *et al.* 2009; Mona *et al.* 2010). As a deficit of heterozygotes is associated with expanding populations, this provides another potential explanation for these results. However, whilst uncertainty regarding the number of separate demes sampled from the British population of *B. subterraneus* may have led to some inaccuracies, the results are extremely close to those that might have been predicted based on current knowledge of the history and demography of the species.

Species-specific differences in genetic divergence and diversity and population history

Results presented in this study are consistent with the known and presumed population histories of the four bumblebee species present in New Zealand. *Bombus terrestris* is common and ubiquitous in both the United Kingdom and in New Zealand and accordingly demonstrates the highest estimated population sizes of all four species in both locations. As this species has always been one of the most common bumblebee species in Britain, it seems likely that it should have been well represented in a sample of British bumblebees taken for

introduction into New Zealand, and indeed, the effective number of founders was estimated to have been 58 [16–93], the highest estimate across all four species. Genetic diversity has remained high within the New Zealand population as a result of the buffering capacity of its relatively large size. However, consistent with the data presented by Schmid-Hempel *et al.* (2007), *B. terrestris* populations in New Zealand do exhibit slightly reduced genetic diversity in comparison with the United Kingdom and slight but significant differentiation from the UK population, demonstrating changes in the genetic composition of this species as a result of the introduction into New Zealand and/or the subsequent isolation of the British and New Zealand populations.

Bombus hortorum is also common and widespread throughout England although it is not as common as *B. terrestris*. Accordingly, this species seems to have been less well represented during the introduction of bumblebees into New Zealand. Whilst the current range of this species in New Zealand is unreported, it once demonstrated a reduced range compared with other species, having been largely confined to the south-east of the South Island (MacFarlane & Gurr 1995). *B. hortorum* demonstrates lower genetic diversity in New Zealand than in the United Kingdom and also exhibits a higher level of differentiation from the UK population than both *B. terrestris* and *Bombus ruderatus*. The estimated effective number of individuals of *B. hortorum* introduced into New Zealand is similar to *B. ruderatus* it is possible that these findings are the result of a combination of a lower initial founder population and lower initial success of this species in colonizing New Zealand leading to fluctuations in population sizes and resultant genetic drift. It is important to note that the estimated parameter *Nb* could reflect not only the original founder population size, but also temporal bottlenecking of the population subsequent to the initial introduction.

Bombus ruderatus was probably similarly common to *B. hortorum* at the time of introduction into New Zealand (Sladen 1912), and accordingly, these species appear to have been relatively equally represented during the introduction of bumblebees into New Zealand. Following the introductions, *B. ruderatus* spread rapidly throughout the South Island and also the North Island (MacFarlane & Gurr 1995). However, this species has shown rapid declines in the United Kingdom because the time of its introduction into New Zealand has become exceedingly scarce, being restricted to a handful of scattered sites in the south and midlands of England (Goulson 2010). This is reflected by the low-effective population size of this species in the United Kingdom as estimated by DIYABC. It is probable that the genetic diversity present in the New Zealand population of *B. ruderatus* represents a fraction of that of the original UK population (as in *B. hortorum*), but that declines experienced by *B. ruderatus* in the United Kingdom have resulted in the loss of genetic diversity here, such that the populations now exhibit similar diversity. This is supported by the fact that several alleles found in the New Zealand population were not present in the British samples of this species (data not shown).

Bombus subterraneus has always had a restricted range within the United Kingdom (Alford 1975) and has probably always been less abundant than *B. terrestris*, *B. hortorum* and *B. ruderatus*. It is likely that the large original UK effective population size estimated for *B. subterraneus* is an artefact of the sampling method, which involved combining several temporally isolated demes. However, the very low estimate of the effective number of *B. subterraneus* introduced into New Zealand (2 [1–5]) is supported by the number of alleles found at each microsatellite locus genotyped (Fig. 2). No single locus was found to have >7 alleles, and the majority had fewer than 5 (each queen could give rise to offspring with up to three different alleles at any given locus: two belonging to herself and one to the male with which she mated) in the New Zealand samples, whereas the British and Swedish populations of the same species demonstrated far greater allele numbers at each locus, and there was greater variation in allele number across loci. Given the strong bottleneck effect associated with the introduction of *B. subterraneus* into New Zealand (Fig. 4) and the subsequent existence of this species within relatively small populations (Goulson & Hanley 2004; Lye *et al.* 2010), it might have been predicted that genetic diversity would be low and that similarity to the original British population is likely to be limited, and this is indeed the case. The New Zealand population of *B. subterraneus* exhibits extremely low genetic diversity in comparison with both the Swedish and original UK population of the same species, and the New Zealand population of *B. subterraneus* is also significantly and highly genetically differentiated from both European populations.

That *B. subterraneus* has been able to persist at all in New Zealand following such a dramatic population bottleneck is something of a surprise and is reflective of other recent findings regarding the invasion success of insects. For example, it has been found that no more than three queens may have given rise to the current and thriving Tasmanian population of *B. terrestris* (Schmid-Hempel *et al.* 2007) and that the entire American population of *Drosophila subobscura* may derive from as few as seven founding individuals (Pascual *et al.* 2007). The ability of invasive species to persist within an introduced range despite low levels of genetic diversity is likely to be strongly influenced by the similarity of the

region to the species native range (Fernandez Iriarte *et al.* 2009). Indeed, the success of British bumblebees in New Zealand must have been greatly facilitated by the similar climate in many parts of New Zealand and the presence of an abundance of introduced bumblebee forage plant species (Hanley & Goulson 2003; Goulson & Hanley 2004; Lye *et al.* 2010). In addition, bumblebees may be better able to cope with very low genetic diversity as a result of haplodiploidy, because deleterious alleles will become subject to selection in haploid males and are therefore more likely to be removed from the population (Werren 1993). Finally, adaptive potential may also be higher than that assumed whether nonallelic diversity such as chromosome inversions contributes to the ability of the populations to respond to changes in environmental conditions (Balanyá *et al.* 2006).

The New Zealand population of *B. subterraneus* is currently being used as a source population for a re-introduction attempt to the United Kingdom. Consideration of genetic factors is a key in planning successful re-introduction attempts because the level of genetic diversity is likely to be important in determining the adaptive potential of a population and thus its ability to thrive despite novel environmental conditions associated with the introduction site. Whilst the success of *B. terrestris* in Tasmania and *B. subterraneus* in New Zealand suggests that low levels of genetic diversity do not necessarily prevent the establishment of healthy bumblebee populations, it should be borne in mind that colonizing a new environment may in some cases present fewer challenges than re-invading a home range. In New Zealand and Tasmania, bumblebees are free from many of the natural enemies and competitors present in their native range, so it remains to be seen whether New Zealand *B. subterraneus* will survive when re-exposed to these organisms. However, that the Swedish population of *B. subterraneus* demonstrates greater genetic variability than the New Zealand population and is no more genetically differentiated from the original United Kingdom than the New Zealand population suggests that supplementation of the introduced population with individuals of other origins could provide a viable additional source of genetic variation.

Conclusions

Species relocation is widespread either as a result of accidental introduction or through intentional intervention, for example, during re-introduction or biological control. Whilst the situations under which species relocation might occur are diverse, the value of the ability to accurately predict the population dynamics of newly invaded populations and in particular, the likelihood of successful establishment and spread, is common to all.

An understanding of the demographic history associated with past introductions and the effect of such introductions on genetic diversity and structure can provide a valuable tool to help achieve this. This study demonstrates the power of modern analytical techniques for reconstructing past events based on population genetic data, allowing us to gain an insight into the demographic processes associated with the establishment of invasive species.

References

Alford DV (1975) *Bumblebees*. Davis Poynter Ltd, London.

Balanyá J, Oller JM, Huey RB, Gilchrist GW, Serra L (2006) Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science*, **313**, 1773–1775.

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025–2035.

Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, **19**, 2609–2625.

Chakraborty R, Jin L (1993) A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. In: *DNA Fingerprinting: State of the Science* (eds Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ), pp. 153–175. Birkhäuser, Switzerland.

Cornuet JM, Santos F, Beaumont MA *et al.* (2008) Inferring population history with DIYABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**, 2713–2719.

Cornuet JM, Ravigné V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics*, **11**, 401.

Depaulis F, Orlando L, Hänni C (2009) Using classical population genetic tools with heterochroneous data: time matters! *PLoS ONE*, **4**, e5541. doi:10.1371/journal.pone.0005541

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

Edwards M, Jenner M (2005) *Field Guide to the Bumblebees of Great Britain and Ireland*. Ocelli, UK.

Estoup A, Scholl A, Pouvreau A, Solignac M (1995) Monoandry and polyandry in bumble bees (Hymenoptera; Bombinae) as evidenced by highly variable microsatellites. *Molecular Ecology*, **4**, 89–93.

Estoup A, Solignac M, Cornuet JM, Goudet J, Scholl A (1996) Genetic differentiation of continental and island populations of *Bombus terrestris* (Hymenoptera: Apidae) in Europe. *Molecular Ecology*, **5**, 19–31.

Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, **12**, 657–674.

Excoffier L, Estoup A, Cornuet JM (2005) Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics*, **169**, 1727–1738.

Fernandez Iriarte PJ, Balanyá J, Pascual M *et al.* (2009) Tracking the origin of an invasive species: *Drosophila subobscura* in Argentina. *Journal of Evolutionary Biology*, **22**, 650–658.

Frankham R, Ballou JD, Briscoe DA (2004) *A Primer of Conservation Genetics*. Cambridge University Press, Cambridge, UK.

Gelman A, Meng X-L (1996) Model checking and model improvement. In: *Markov Chain Monte Carlo in Practice* (eds Gilks W, Richardson S, Spiegelhalter D), pp. 189–201. Chapman and Hall, London, UK.

Goldstein DB, Linares AR, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics*, **139**, 463–471.

Goudet J (1995) FSTAT (vers. 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.

Goulson D (2010) *Bumblebees: Behaviour, Ecology and Conservation*. Oxford University Press, Oxford.

Goulson D, Hanley ME (2004) Distribution and forage use of exotic bumblebees in South Island, New Zealand. *New Zealand Journal of Ecology*, **28**, 225–232.

Hanley ME, Goulson D (2003) Introduced weeds pollinated by introduced bees: cause or effect? *Weed Biology and Management*, **3**, 204–212.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633–1638.

Hingston AB, Marsden-Smedley J, Driscoll DA *et al.* (2002) Extent of invasion of Tasmanian native vegetation by the exotic bumblebee *Bombus terrestris* (Apoidea: Apidae). *Australian Ecology*, **27**, 162–172.

Holehouse KA, Hammond RL, Bourke AFG (2003) Non-lethal sampling of DNA from bumble bees for conservation genetics. *Insectes Sociaux*, **50**, 277–285.

Hopkins I (1914) History of the bumblebee in New Zealand: its introduction and results. *New Zealand Department for Agriculture Industry and Commerce*, **46**, 1–29.

Hothorn T, Hornik K, van de Wiel MA, Zeileis A (2008) Implementing a class of permutation tests: the *coin* package. *Journal of Statistical Software*, **28**, 1–23.

Jones OR, Wang J (2009) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, **10**, 551–555.

Kraus FB, Wolf S, Moritz RFA (2009) Male flight distance and population substructure in the bumblebee *Bombus terrestris*. *Journal of Animal Ecology*, **78**, 247–252.

Lepais O, Darvil B, O'Connor S *et al.* (2010) Estimation of bumblebee queen dispersal distances using sibship reconstruction method. *Molecular Ecology*, **19**, 819–831.

Loader CR (1996) Local likelihood density estimation. *The Annals of Statistics*, **24**, 1602–1618.

Lopes JS, Boessenkool S (2010) The use of approximate Bayesian computation in conservation genetics and its application in a case study on yellow-eyed penguins. *Conservation Genetics*, **11**, 421–433.

Lozier JD, Cameron SA (2009) Comparative genetic analyses of historical and contemporary collections highlight contrasting demographic histories for the bumble bees *Bombus pennsylvanicus* and *B. impatiens* in Illinois. *Molecular Ecology*, **18**, 1875–1886.

Lye GC, Kaden JC, Park KJ, Goulson D (2010) Forage use and niche partitioning by non-native bumblebees in New Zealand: implications for the conservation of their populations of origin. *Journal of Insect Conservation*, **14**, 607–615.

MacFarlane RP, Griffin RP (1990) New Zealand distribution and seasonal incidence of the nematode *Sphaerularia bombi* Dufour, a parasite of bumblebees. *New Zealand Journal of Zoology*, **17**, 191–199.

MacFarlane RP, Gurr L (1995) Distribution of bumble bees in New Zealand. *The New Zealand Entomologist*, **18**, 29–36.

Mona S, Catalano G, Lari M *et al.* (2010) Population dynamic of the extinct European aurochs: genetic evidence of a north-south differentiation pattern and no evidence of post-glacial expansion. *BMC Evolutionary Biology*, **10**, 83. doi:10.1186/1471-2148-10-83.

Nei M (1987) *Molecular Evolution Genetics*, Columbia University Press, New York, pp. 512.

Pascual M, Chapuis MP, Mestres F *et al.* (2007) Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology*, **16**, 3069–3083.

R Development Core Team (2009) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

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

Schmid-Hempel P, Schmid-Hempel R, Brunner PC, Seeman OD, Allen GR (2007) Invasion success of the bumblebee, *Bombus terrestris*, despite a drastic genetic bottleneck. *Heredity*, **99**, 414–422.

Sladen FWL (1912) *The Humble-bee, Its Life History and How to Domesticate It*. Macmillan and Co. Ltd, London, UK.

Truett GE, Heeger P, Mynatt RL, Truett AA, Walker JA, Warman ML (2000) Preparation of PCR quality mouse genomic DNA with hot sodium hydroxide and tris (HotShot). *BioTechniques*, **29**, 52–54.

Viard F, Frank P, Dubois MP, Estoup A, Jarne P (1998) Variation of microsatellite size homoplasy across electromorphs, loci and populations in three invertebrate species. *Journal of Molecular Evolution*, **47**, 42–51.

Wandeler P, Hoeck PEA, Keller LF (2007) Back to the future: museum specimens in population genetics. *Trends in Ecology and Evolution*, **22**, 634–642.

Wang J (2004) Sibship reconstruction from genetic data with typing errors. *Genetics*, **166**, 1963–1979.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Werren JH (1993) The evolution of inbreeding in a haplodiploid organism. In: *The Natural History of Inbreeding and Outbreeding* (ed. Thornhill NW), pp. 42–94. University of Chicago Press, Chicago.

Widmer A, Schmid-Hempel P (1999) The population genetic structure of a large temperate pollinator species, *Bombus pascuorum* (Scopoli) (Hymenoptera: Apidea). *Molecular Ecology*, **8**, 387–398.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.

Yokoyama J, Fukuda T, Yokoyama A, Nakajima M (2004) Extensive size homoplasy at a microsatellite locus in the Japanese bumblebee, *Bombus diversus*. *Entomological Science*, **7**, 189–197.

G.C.L. is a biologist interested in the ecology and conservation of insects. This article forms part of her PhD thesis. O.L. is a molecular ecologist with a wide interest in application of molecular markers for studying the history, ecology, evolution and adaptation of species. D.G. is an insect ecologist specializing in studies of bumblebees, their behaviour, ecology and conservation.

Data accessibility

Data deposited at Dryad: doi:10.5061/dryad.sk22v.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Confidence in parameter estimated by the mode using Model 2.

Table S2 Demographic parameters estimated by DIYABC for the four studied species and the five models with different prior specification.

Fig. S1 Effect of the duration of the bottleneck prior specification on the parameter estimation.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.