MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 819–831 doi: 10.1111/j.1365-294X.2009.04500.x

Estimation of bumblebee queen dispersal distances using sibship reconstruction method

OLIVIER LEPAIS,* BEN DARVILL,* STEPHANIE O'CONNOR,* JULIET L. OSBORNE,† ROY A. SANDERSON,‡ JOHN CUSSANS,† LOUIS GOFFE‡ and DAVE GOULSON*

**School of Biological & Environmental Sciences, University of Stirling, Stirling FK9 4LA, UK, †Department of Plant & Invertebrate Ecology, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK, ‡Institute for Research on Environment and Sustainability, Devonshire Building, University of Newcastle, Newcastle-upon-Tyne NE1 7RU, UK*

Abstract

Dispersal ability is a key determinant of the propensity of an organism to cope with habitat fragmentation and climate change. Here we quantify queen dispersal in two common bumblebee species in an arable landscape. Dispersal was measured by taking DNA samples from workers in the spring and summer, and from queens in the following spring, at 14 sites across a landscape. The queens captured in the spring must be full sisters of workers that were foraging in the previous year. A range of sibship reconstruction methods were compared using simulated data sets including or no genotyping errors. The program Colony gave the most accurate reconstruction and was used for our analysis of queen dispersal. Comparison of queen dispersion with worker foraging distances was used to take into account an expected low level of false identification of sister pairs which might otherwise lead to overestimates of dispersal. Our data show that *Bombus pascuorum* and *B. lapidarius* queens can disperse by at least 3 and 5 km, respectively. These estimates are consistent with inferences drawn from studies of population structuring in common and rare bumblebee species, and suggest that regular gene flow over several kilometres due to queen dispersal are likely to be sufficient to maintain genetic cohesion of ubiquitous species over large spatial scales whereas rare bumblebee species appear unable to regularly disperse over distances greater than 10 km. Our results have clear implications for conservation strategies for this important pollinator group, particularly when attempting to conserve fragmented populations.

Keywords: *Bombus*, kinship, microsatellite, population structure, social insects

Received 25 August 2009; revision received 16 November 2009; accepted 23 November 2009

Introduction

Bumblebee species (*Bombus*) are important pollinators for a wide range of flowering plants in temperate and subarctic regions of the northern hemisphere. Many bumblebee species have exhibited range contractions and regional extinctions in recent decades (Goulson *et al*. 2008; Williams & Osborne 2009). Due to their haplodiploid life cycle (i.e. the sexual female is diploid while the reproductive male is haploid) and their single locus complementary sex determination system (Crozier 1971; Bull 1981), bumblebee populations may be particularly sensitive to low population size (Chapman & Bourke 2001). In small, inbred populations, males and queens are more likely to share one allele at the sex determination locus; if this occurs, half of the diploid offspring develop as sterile males that do not participate in foraging or other useful tasks and thus are costly to the colony. Such colonies are very likely to fail (Cook & Crozier 1995). Another consequence of the haplodiploid life cycle is that gene flow is more dependent on the movement of queens than of males; the dispersal of a diploid queen contributes twice as much

Correspondence: David Goulson, Fax: +44 1786 467843; E-mail: dave.goulson@stir.ac.uk

as the dispersal of a haploid male, all else being equal. The relative importance of queen dispersal is increased further if part of her dispersal happens after mating so that she is helping to disperse sperm (Berg *et al.* 1998). Furthermore, by being the founders of new nests, queens govern the effective dispersal of the species.

Gene flow and dispersal are important biological parameters that determine how species cope with habitat fragmentation and climatic change (Malcolm *et al.* 2002). However, they are often difficult to quantify. Many studies of bumblebees have attempted to estimate the foraging distance of nonreproductive workers (Chapman *et al.* 2003; Darvill *et al.* 2004; Knight *et al.* 2005). Such dispersal does not involve genetic exchange and therefore is not a component of gene flow. Indirect estimation of the amount of gene flow between populations can be obtained from analyses of population genetic structure. In the common bumblebee species studied to date, no significant spatial genetic structuring has been found over wide geographic areas, suggesting large population sizes and high connectivity (gene flow) (Estoup *et al.* 1996; Widmer & Schmid-Hempel 1999; Ellis *et al.* 2006). In contrast, studies focusing on rare bumblebee species have found significant isolation by distance and genetic differentiation between subpopulations indicating that gene flow is restricted to a few tens of kilometres (Darvill *et al.* 2006; Ellis *et al.* 2006). However, in addition to the level of gene flow, population structure is influenced by population size, genetic drift and bottlenecks, so that only broad inferences can be obtained about how far individuals move.

A more direct approach to estimate dispersal distances in bumblebees is the observation of the speed of spread of introduced bumblebees. One of the best documented examples is the invasion of Tasmania by *Bombus terrestris* via a single introduction event in 1992 (Hingston *et al.* 2002; Hingston 2006; Schmid-Hempel *et al.* 2007). The invasive species was able to colonize most of the island (i.e. a spread of 300 km) in about 10 generations, which indicates a high dispersal capacity of the bumblebee queens. However, the particular context of this expansion is perhaps not representative of the dispersal process in the native range of the species. Only one recent study (Kraus *et al.* 2009) estimates the dispersal distance of male bumblebees, and suggests that in *B. terrestris* they move between 2.6 km and 9.9 km. Queen dispersal in bumblebees has never been directly studied.

Queen dispersal could occur at different stages in her life: before mating; between mating and hibernation; or post-hibernation. In terms of gene flow and colonization distance, it is the sum of these movements which is important. Plant population structure may also be influenced by queen dispersal as, if queens do move long distances (and are likely to visit flowers along the way), they may transfer pollen far further than do workers. This might be particularly important for early spring-flowering plants, many of which are dependent on pollination by queen bumblebees (Macior 1968, 1994; Washitani *et al.* 1994).

In this study, we directly estimated effective queen dispersal distance by applying genetic fullsib reconstruction (FSR) methods on workers sampled in summer and queens sampled in the following spring. By comparing the locations of workers and queens that were sisters, we estimated the distance travelled by queens. FSR is an active area of research and numerous approaches have been recently developed. Among them, several can handle haplodiploid species (Konovalov *et al.* 2004; Wang 2004; Kokuvo *et al.* 2007a,b). For the researcher who wishes to apply such approaches, this choice results in the difficult task of selecting the most appropriate method. The accuracy of the FSR usually depends on numerous parameters including the number of loci available, the degree of polymorphism of the loci, the amount of genotyping error and missing data, and the reconstruction method being used. In addition, the particularly low family structure typically encountered in bumblebee species samples (only a few individuals belong to the same colony) may render FSR more difficult because the methods are expected to perform better in the case of moderate to high family structure (many sampled individuals originating for the same colony; Wang 2004, 2006). Although methods comparisons can usually be found in studies introducing new FSR methods (Wang 2004; Konovalov *et al.* 2005; Konovalov 2006; Kokuvo *et al.* 2007b, 2008; Wang & Santure 2009), data set simulations generally encompass a wide range of parameters (for example numbers of loci and alleles by loci, varying family structure) to demonstrate the applicability of the new method. In this study we adopted an end user point of view in addressing the difficult choice of the FSR method by comparing the power of most available methods and estimating the quality of our genetic data for FSR. For this purpose we used a 'real world' simulation mimicking as closely as possible our genetic marker characteristics, sampling scheme and species biology. In addition to providing an informed decision on the best software to use for FSR on a particular genetic data set, this approach allows us to estimate the level of accuracy that can be expected from a FSR. This is particularly important in our study because potential FSR errors could have a pronounced impact on dispersal distance estimation.

In this study we adopted a conservative approach to estimating queen dispersal distance. We first used a

simulated data set to compare and test the relative performance of available methods for sibship reconstruction. Our sampling was performed at a sufficiently wide spatial scale that we would expect sister workers to occur rarely across multiple sites. This allowed us to compare the estimated dispersal distance of workers (reflecting the consequence of sibship reconstruction errors) to those observed for queens. This resulted in a queen dispersal distance estimate that incorporates and accounts for errors in sibship reconstruction. The observed pattern of queen dispersal distance is compared to previous indirect estimation of gene flow and discussed with regard to the potential implications for the functioning and conservation of bumblebee populations.

Materials and methods

Sample collection

The study was carried out in a 10×20 km rectangle centred on Rothamsted Research (Harpenden, Hertfordshire, UK), an area that has previously been used for a number of related studies on bumblebee ecology (Knight *et al.* 2005, 2009; Osborne *et al.* 2008). Sampled sites consisted of field margins where *Bombus lapidarius* and *B. pascuorum* workers were collected along a 200×10 m strip of the field margin. We sampled the same sites three times: workers were sampled between 22 May 2007 and 22 June 2007 and a second time between 25 July and 9 August. The following year, queens were sampled during March and April 2008 at the same sites. All sites were at least 1 km apart from one another and 1 km from a substantial urban area. A total of 1083 *B. lapidarius* and 1660 *B. pascuorum* workers and 13 *B. lapidarius* and 220 *B. pascuorum* queens were caught (Table 1). A nonlethal tarsal sample (Holehouse *et al.* 2003) was taken from the mid-leg of each individual, and these samples were preserved in ethanol until subsequent DNA extraction.

Molecular methods

DNA was extracted using the HotShot protocol (Holehouse *et al.* 2003). Individuals were genotyped at ten or nine (*B. lapidarius* and *B. pascuorum*, respectively) microsatellite loci with a multiplex protocol allowing us to amplify up to five loci (Table 2) in one polymerase chain reaction (PCR). Amplification was carried out in a 10 μL final volume using QIAGEN Multiplex PCR kits. Each reaction contained 1 μL Q-solution, 5 μL PCR Master Mix, from 0.2 to 1 μM of each primer (Table 2) and approximately 10 ng template DNA. Samples were initially denatured at 95 °C for 15 min, followed by 35 cycles of denaturing at 94 °C for 30 s, annealing at either 52 or 47 °C (primer sets PS1 and PS2, respectively) for 90 s, extension at 72 °C for 90 s and a final extension step at 72 °C for 10 min.

PCR products were visualized on an ABI 3730 capillary DNA sequencer (The Sequencing Service, University of Dundee) with a 1:80 dilution before the run and using a GeneScan 500 LIZ internal size standard (Applied Biosystems). Fragments were sized using STRand software (Veterinary Genetics Laboratory, University of California at Davis, http://www.vgl.ucda-

Table 1 Sampling description

			Bombus lapidarius		*Bombus pascuorum*	
Sampled site	Latitude	Longitude	*N* workers	*N* queens*	*N* workers	*N* queens*
T1	673680.8	5746544.2	27	0	102	0
T2	675618.1	5743870.0	38	0	107	36
T3	674851.0	5741458.7	128	0	129	16
T4	677858.2	5740999.8	39	2	102	23
T5	678404.7	5744908.5	105	2	116	24
T7	683754.8	5741380.9	103	0	112	22
T8	684077.9	5746987.1	123	1	120	16
C1	671513.5	5744113.7	132	2	121	16
C2	671451.9	5741312.0	55	0	136	11
C3	672168.5	5740121.5	53	3	131	15
C4	673102.7	5744935.8	92	1	117	19
C5	676688.0	5746085.3	56	1	108	10
C7	689107.4	5744955.5	7	1	100	12
C8	687699.0	5745536.3	122	0	140	0
Total			1080	13	1641	229

*Queens were sampled the following year.

Table 2 Multiplex design and characteristics of the microsatellite loci

Bombus lapidarius								*Bombus pascuorum*							
Primer set	SSR	Ref.	Conc. (μM)	Dye	*A*	A_e	F_{is}	Primer set	SSR	Ref.	Conc. (μM)	Dye	*A*	A_e	F_{is}
PS1	BL11	a	0.20	FAM	19	5.2	0.08	PS1	BL03	a	0.40	FAM	21	5.0	**0.01**
PS1	BL06	a	0.20	NED	21	4.0	−0.01	PS1	BT10	a	0.30	PET	23	8.6	0.03
PS1	BT24	a	0.20	FAM	14	4.7	0.02	PS1	BT26	a	0.20	NED	14	3.0	**0.05**
PS1	BT09	a	0.60	PET	24	6.9	0.02	PS1	BT18	a	0.20	VIC	16	2.3	0.07
PS1	BT18	a	0.20	VIC	19	7.8	0.05	PS2	B124	b	0.15	VIC	17	3.8	−0.01
PS2	B126	a	0.20	FAM	17	4.1	−0.04	PS2	B126	b	0.15	FAM	14	2.0	**0.05**
PS2	B96	b	0.60	FAM	14	4.7	0.01	PS2	B96	b	0.60	FAM	22	3.5	0.04
PS2	B10	b	0.20	VIC	17	4.6	0.03	PS2	B132	b	1.00	PET	18	4.4	0.01
PS2	B11	b	0.20	PET	19	5.2	−0.01	PS2	B118	b	0.60	NED	18	6.1	**0.05**
PS2	B118	b	0.60	NED	9	2.2	0.03								
				mean	17.3	4.9						mean	18.1	4.3	

A, observed number of alleles; A_e, effective number of alleles [computed as $1/(1 - Hs)$]; F_{is}, deficit in heterozygosity with significant value in bold (after Bonferroni correction: $P < 0.0056$ for nine comparisons and *B. pascuorum*). Ref.: (a): Funk *et al.* (2006); (b): Estoup *et al.* (1995, 1996).

vis.edu/informatics/strand.php) and raw alleles sizes were binned into discrete classes using FlexiBin macro (Amos *et al.* 2007).

We regenotyped 22% of *B. lapidarius* and 46% of *B. pascuorum* individuals at least once because they failed to give a full genotype at all loci on the first attempt. We repeated the genotyping until the data set was nearly complete; bees were only retained for analysis if data were available from a minimum of six microsatellites (hence excluding three *B. lapidarius* and 19 *B. pascuorum* workers). All queens were fully genotyped at all loci. Missing genotypes at the end of this process amounted to 0.5% in *B. lapidarius* and 0.8% in *B. pascuorum*. Table 1 gives the number of individuals that were used in the analyses. Based on the repeated genotyping, we estimated a genotyping error rate that is an uppermost estimation because reanalysed individuals were more prone to contain genotyping errors due to nonoptimal microsatellite amplifications. Overall, we detected 0.5% rate of allele dropout and 1.5% of other types of error in both species based on comparison of repeated genotyping. This estimated genotyping error rate was furthermore reduced as all detected errors were corrected to compose the final data set. The genetic data are archived in the Dryad database (available at http://hdl.handle.net/10255/dryad.1113).

Statistical methods

Performance of sibship reconstruction methods. Power of microsatellite markers for sibship reconstruction methods. We first used the Descending Ratio (DR) algorithm implemented in KinGroup software (Konovalov *et al.* 2004) to identify workers that probably originated from the same nest, keeping only one worker from each nest in order to estimate allele frequencies at each locus for both species. KinInfor software version 1.0 (Wang 2006) was used to compute the power for relationship inference (PW_R estimated by the simulation procedure) based on estimated allele frequencies of our sets of microsatellite markers. We ran the software assuming fullsib ($\Delta_1 = 0.5$, $\Delta_2 = 0.5$) as the primary hypothesis and unrelated ($\Delta_1 = 0$ and $\Delta_2 = 0$) as null hypothesis, a prior Dirichlet distribution of (1,1,1) for Δ_0, Δ_1 and Δ_2, respectively, confidence level set at 0.05 and 1 000 000 simulated pairs of genotypes. We performed the analysis assuming either no genotyping error or a genotyping error rate of 0.02 at each locus.

Genetic data set simulation. We simulated a genetic data set to test the different sibship reconstruction methods. The simulation procedure included two steps: (i) computing the expected family structure given our sampling scheme and the species biology (nest density and dispersal distance) and (ii) simulating genetic data of the virtually sampled individuals taking into account the specific allele frequency and genotyping error rate of our genetic markers. First, we simulated a nest landscape which comprised a 15×30 km area within which 14 sample sites were placed with the same spatial distribution as our real sample sites (Table 1). Bumblebee nests were randomly located within this area at an overall density of 10 nests per square kilometre. We assumed that each nest contained the same unknown number of workers and that the probability of sampling

a worker from a particular nest decreased exponentially with the distance from the nest:

$$P_{\text{worker}} = \exp\left(\frac{-d^2}{2V^2}\right)$$

with d the distance between the nest and the sample site and V the parameter controlling the foraging distance of the workers (T.G. Chapman, unpublished). Sampling was simulated by computing the distance between each sampling site and each nest in the landscape, resulting in a probability vector of sampling an individual from each nest in the landscape. Indeed, for each site i, the relative probability of sampling a worker from the nest j is:

$$P_{ij} = \frac{\exp\left(-d_{ij}^2/2V^2\right)}{\sum_j \exp\left(-d_{ij}^2/2V^2\right)}$$

where d_{ij} is the distance between the sampling site i and the nest j. The numerator represents each nest's relative contribution of foragers to the sampling at site i while the denominator scales these contributions so that the sum of probabilities across all nests for each site equals 1. We set the V parameter to 400 m which means that 99% of the workers foraged within 800 m from their nest. Based on the probability vector of sampling nests at each sampling site, we randomly sampled 50 workers at each site with replacement (i.e. we allowed the sampling of more than one worker by nest). This simulation resulted in 14 sites sampled for a total of 700 workers belonging to 321 nests. Despite the somewhat arbitrary parameters for the nest density in the simulated nest landscape and the worker foraging distance function, the resulting family structure is typical of previous studies of bumblebees (e.g. Chapman *et al.* 2003; Darvill *et al.* 2004; Knight *et al.* 2005) with most nests represented by only one worker and a decreasing number of nests with an increasing number of workers (140 nests were represented by one worker, 81 nests had two workers, 48 nests had three workers, 27 nests had four workers, 12 nests had five workers, five nests had six workers and eight nests had seven workers; lambda parameter of the corresponding Poisson distribution: 1.8). The same family structure was used for the two species.

Secondly, we used allele frequencies at each locus, as estimated above, to add the genetic aspect to this simulated sampling. We assigned a queen and a male to each virtually sampled nest and randomly sampled from the known allele frequencies two alleles at each locus for the queens and one allele at each locus for the males. Worker genotypes were generated by randomly drawing one allele from the queen and the allele of the male at each locus. We thus knew the precise nest origin of each simulated worker and the genotype of the parents. This data set constituted the perfect data set (called PF hereafter). The allele assignment procedure was repeated independently for each species according to its particular allele frequencies: the *B. pascuorum* simulated data set consisted of nine microsatellites and *B. lapidarius* 10 microsatellites. In order to further mimic a typical microsatellite data set, we altered these PF data sets by randomly adding 0.5% and 0.8% of missing genotypes for *B. lapidarius* and *B. pascuorum*, respectively, and 2% of genotyping error (including 0.5% of allelic dropout and 1.5% of other kind of error) for the two species. These second 'realistic' data sets (called ER hereafter) allowed us to estimate the robustness of the sibship reconstruction methods to genotyping errors and missing data typically encountered when dealing with large microsatellite data set.

Testing the available sibship reconstruction methods. We tested the main full sibship reconstruction (FSR) methods available for haplodiploid species (Table 3). We used mSLCA package (Kokuvo *et al.* 2007b), KinGroup version 2 (Konovalov *et al.* 2004) and Colony version 1.2 (Wang 2003) as recommended in their respective manuals. For Colony, we assumed no genotyping error in the analysis when testing the PF-simulated data set while we assumed 0.5% of allele dropout and 1.5% of other error types in the analysis of the ER-simulated data set. In addition, this software can infer parental genotypes, a feature that was recently used in bumblebee studies (Herrmann *et al.* 2007; Kraus *et al.* 2009).

The accuracy of each FSR method was estimated with several parameters. We reported the estimated number of nest and the number of correctly reconstructed nests. We also computed with R software (R Development Core Team 2005) and the package *clue* (Hornik 2005) the minimum number of moves which corresponds to the minimum number of individuals which need to be moved from one nest to another (or to a new nest) to transform the reconstructed partition into the real one (Rubin 1967).

Full sibship reconstruction and estimation of queen dispersal distance. We used Colony version 1.2 (Wang 2004), the most accurate FSR method given our studied species (see Results section), on the full data set including workers and queens sampled the following year to cluster individuals into nests. We accounted for a genotyping error rate of 0.5% of allele dropout and 1.5% of other types of error in the analysis.

Given the sampled location of each individual and the nest they originated from, we assigned a dispersal

Table 3 Characteristics of the tested fullsib reconstruction (FSR) methods

FSR method	Description	Particularity	Reference
mSLCA	Correspondence analysis on a genetic similarity matrix between pairs of individuals to cluster individuals into fullsib groups	Does not accept missing data	Kokuvo *et al.* (2007a,b)
KinGroup-Simpson	Searches for the best sibship partition by randomly moving one individual into a different group at each iteration and tests the newly formed sibship group using Mendelian rules of inheritance to determine if all individuals of the sib group could have been generated by the same pair of parental genotypes. The best partition is the one that yields the highest Simpson index	Strictly based on Mendelian inheritance tests. Not robust to genotyping error	Konovalov *et al.* (2004)
KinGroup-DR	Based on a pairwise likelihood ratio test which assesses the probability of a relationship being either unrelated (the null hypothesis) or haplodiploid full-sibling (the primary hypothesis)	Expected to be robust to genotyping error	Konovalov *et al.* (2004)
KinGroup-SRD	Combines a modified version of the Simpson algorithm and the DR algorithm to improve the FSR	Expected to be robust to genotyping error	Konovalov (2006)
Colony	Explicitly takes genotyping error into account in its maximum likelihood search algorithm	Robust to genotyping error. Infers parental genotypes	Wang (2003)

distance to each individual following a four cases rule depending on the nest configuration. (i) First of all, nests represented by only one individual were not used in the analysis because the location of these nests remained unknown. In the remaining cases where a nest was represented by more than one individual, the following options were considered to assign a dispersal distance to individuals. (ii) In the easier to handle cases, a nest could be located at a particular site when a majority of the individuals from a nest originated from a single sampled site, and the remaining individuals were found in a single second site. In such instances, the majority were assigned a dispersal distance of 0 m while individuals originating from the same nest but sampled at the second site were considered as dispersers with a dispersal distance equal to the distance between the sample site and the nest site. (iii) Another configuration was a nest with two or more workers with an equal number of individuals originated from two different sites. Here it was impossible to assign a sampled site to the nest and all individuals were considered as dispersers with a dispersal distance equalling half the distance between the sites. (iv) Finally, in a few cases where a nest were represented by individuals originating from more that two sites, the corresponding individuals were not used in the analysis because of the uncertainty in the nest location.

Our dispersal distance assignment is conservative as we considered individuals independently of their nature, i.e. we did not assume that queens could disperse further than workers, and we did not consider ambiguous cases where nest location was difficult to estimate.

Despite the accuracy of Colony in FSR, some unrelated individuals will be assigned to the same nest by error (see Results section) and thus false dispersal distances will be introduced into our analysis, inflating dispersal distance estimation. However, the rate of this error will be the same for workers and queens, so that we can compare dispersal distance between these two kinds of individuals. In the case of limited dispersal, queen dispersal distance should not be greater than observed worker foraging distance, while if queens dispersed further than workers, we should observe a greater dispersal distance for queens than for workers. We first compared both dispersal distance distributions using a two-sample Wilcoxon test. We then divided dispersal distances into 1000 m class intervals and tested for differences between the frequency of queen dispersal and worker dispersal (our null hypothesis) in each class with a binomial exact test (Darvill *et al.* 2004). Sequential Bonferroni corrections were applied to minimize type I error (Rice 1989).

Results

Microsatellites power for relationship inference

The 10 microsatellite loci in *Bombus lapidarius* yielded an average of 17.3 alleles (min = 9, max = 24) and a mean effective number of alleles of 4.9 per locus (min = 2.2, max = 7.8; Table 2). There was not a significant deficit of heterozygotes (no F_{is} values were significantly different from 0; Table 2). Assuming no genotyping error, we get a power for relationship inference (PW_R)

of 1.0, reducing to 0.9999955 with a genotyping error rate of 0.02 at each locus for the 10 *B. lapidarius* microsatellites.

The nine microsatellite loci in *B. pascuorum* yielded an average of 18.1 alleles (min = 14, max = 23) and a mean effective number of alleles of 4.3 per locus (min = 2.0, max = 8.6; Table 2). The lower number of effective alleles despite a higher number of alleles by loci in this species was explained by a higher proportion of rare alleles compared to *B. lapidarius* microsatellites. Four loci showed a significant, although low, deficit in heterozygotes (significant F_{is} ranging from 0.01 to 0.05; Table 2) which could be due to the presence of family structure in the data set. For the nine microsatellites used for *B. pascuorum*, PW_R was 1.0 assuming no genotyping error and 0.9998976 assuming an error rate of 0.02. Thus, the *B. lapidarius* genetic data set appeared to be more powerful than the *B. pascuorum* genetic data set because of the difference in microsatellite number (10 against 9, respectively) and the more balanced allele frequencies in *B. lapidarius*.

Relative performance of sibship reconstruction methods

Number of correct classifications. Accuracy in FSR was higher for *B. lapidarius* than for *B. pascuorum* (Table 4). When applied to the PF, all but one FSR method underestimated the number of nests (Table 4) and only the Simpson method overestimated the number of nests. The percentage of correctly reconstructed nests varied greatly between FSR methods, from about 12% for the Simpson method to 88% and 98% for Colony in *B. pascuorum* and *B. lapidarius*, respectively. The DR method yielded a fairly accurate reconstruction (79% and 95% for *B. pascuorum* and *B. lapidarius*, respectively) while Simpson-assisted Descending Ratio (SDR) was found to be relatively accurate in *B. lapidarius* (77%) but gave poor reconstruction for *B. pascuorum* (37%). The reconstructions by the mSLCA method were inaccurate (30% and 60% of correctly reconstructed nest for *B. pascuorum* and *B. lapidarius*, respectively). The minimum number of moves confirmed the relative performances of the FSR methods with only three *B. lapidarius* and 20 *B. pascuorum* individuals needed to be moved from one nest to another to reach the correct classification for Colony while more than 100 individuals had to be moved with the reconstructions from Simpson and mSLCA to obtain the true classification (Table 4).

When applied to the simulated data set containing missing data and genotyping errors (ER), performances of the methods were lower (Table 4). Contrasting with the estimation obtained with the PF data set, all methods but Colony overestimated the number of nest while Colony underestimated the number of nests. Only Colony coped well with genotyping errors as the percentage of correctly reconstructed nest was 75% and 91% for *B. pascuorum* and *B. lapidarius*, respectively (translating to 17 and 46 minimum number of moves). It is

Table 4 Performance of different sibship reconstruction methods (321 simulated nests)

	PF data set				ER data set			
Method	*N* nests	*N* correct nests	% correct nests	Minimum number of moves†	*N* nests	*N* correct nests	% correct nests	Minimum number of moves†
Bombus lapidarius								
mSLCA*	282	194	60.4%	114	358	183	57.0%	136
KinGroup DR	317	306	95.3%	12	412	239	74.5%	102
KinGroup SDR	312	247	76.9%	48	401	191	59.5%	140
KinGroup Simpson	438	36	11.2%	394	433	32	10.0%	387
Colony	318	314	97.8%	3	306	292	91.0%	17
Bombus pascuorum								
mSLCA*	221	98	30.5%	248	358	80	24.9%	268
KinGroup DR	294	255	79.4%	44	363	212	66.0%	115
KinGroup SDR	268	119	37.1%	176	344	110	34.3%	216
KinGroup Simpson	443	41	12.8%	388	442	45	14.0%	387
Colony	303	283	88.2%	20	280	240	74.8%	46

N nests, estimated nest number; *N* and % correct nest, number and percentage of correctly reconstructed nests; PR data set, simulated microsatellite data set without error or missing data; ER data set, simulated microsatellite data set with an error rate of 0.02 and a 0.005 or 0.008 rate of missing data for *B. lapidarius* and *B. pascuorum*, respectively.

*Because the mSLCA method does not handle missing data, the ER data set only contained error in this case (no missing).

†The minimum number of individuals to move from a reconstructed nest to another nest to obtain the true classification.

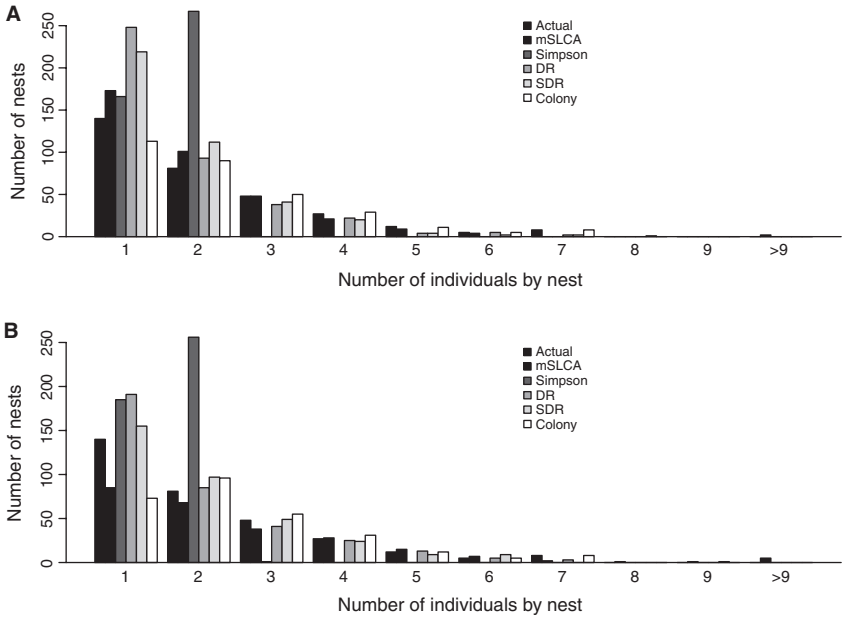

Fig. 1 Actual (simulated) and detected family structure using five FSR methods for (A) 10 microsatellites in *Bombus lapidarius* and (B) nine microsatellites in *B. pascuorum* based on the ER-simulated data sets (which included missing data and genotyping errors).

worth noting that the DR method still gave acceptable FSR performance with 66% and 75% of correctly reconstructed nests resulting in about 100 minimum number of moves. All other methods gave poor accuracy (Table 4).

Finally, despite the good performance of Colony in FSR, only 50.6% of monolocus and as few as 6% of multilocus queen genotypes were correctly reconstructed in the more favourable case of the simulated PF *B. lapidarius* data set. Queen genotypes were found to be correctly reconstructed for nests represented by at least five individuals.

Consequences of FSR errors in the detected family structure. It is important to understand the consequences of FSR error on sibship assignments and detected family structure. Figure 1 illustrates the estimated family structure by each FSR method used with the ER-simulated data sets. The poor performance of the Simpson method was explained by the fact that this method falsely identified nests with only one or two workers (Fig. 1A, B). In the presence of genotyping errors, DR and SRD methods tended to split individuals from the same nest into different nests as shown by the excess of nests with one or two individuals (Fig. 1A, B). mSLCA had the same behaviour with the simulated *B. lapidarius* data set (Fig. 1A) while the method grouped unrelated individuals in the same nest in the case of the *B. pascuorum*-simulated data set (Fig. 1B). Moreover, mSLCA was the only method that falsely grouped more than 10 individuals in the same nest irrespective of the data set. Finally, Colony tended to group unrelated individuals in the same nest in a constant manner as indicated by the deficit of nests represented by one individual and the excess of nest with two to four individuals. This effect was more pronounced for the *B. pascuorum* (Fig. 1B) than the *B. lapidarius* (Fig. 1A) simulated data set. Note however that nests with more than four individuals were correctly resolved.

Queen dispersal distance

Fullsib reconstructions and nest configuration among sampled sites. The 1093 *B. lapidarius* individuals were classified in 433 nests (Table 5). A total of 164 nests were represented by one individual (case 1; 161 nests represented by one worker and three nests one queen). Among the remaining 269 nests, 181 could be assigned to a particular site (case 2) while 85 contained the same number of individuals sampled in two different sites (case 3). Finally, three nests contained individuals sampled in more than two sites (case 4). Nests represented by only one individual comprised 15% of the workers and 23% of the queens (Table 5). About 38% of the queens belonged to a nest containing individuals from more than one sampled site compared to 16% for workers.

The 1861 *B. pascuorum* individuals were classified in 654 nests (Table 5) using Colony. A total of 88 nests were represented by one individual (case 1; 76 nests represented by one worker and 12 nests one queen). Among the remaining 566 nests, 319 were assigned to a particular site (case 2) because all or a majority of individuals originated from a single site, while 197 contained an equal number of individuals sampled at two

Table 5 Fullsib reconstruction in *Bombus lapidarius* and *B. pascuorum*

	B. lapidarius				*B. pascuorum*			
Cases	N_{nests}	N_{indiv}	$N_{workers}$	N_{queens}	N_{nests}	N_{indiv}	$N_{workers}$	N_{queens}
(1) Nest not used because only one individual	164	164	161 14.9%	3 23.1%	88	88	76 4.6%	12 5.5%
(2) Nests assigned to a site	181	747	742 68.7%	5 38.5%	319	1222	1127 68.7%	95 43.1%
(3) Nests at mid-point between two sites	85	172	168 15.6%	4 30.8%	197	398	316 19.3%	82 37.3%
(4) Nest not used because from more than two sites	3	10	9 0.8%	1 7.7%	50	153	122 7.4%	31 14.1%
Total	433	1093	1080	13	654	1861	1641	220

N_{nests}, number of nests; N_{indiv}, number of individuals; $N_{workers}$, number of workers; N_{queens}, number of queens. For details about the four different cases, see Materials and methods section.

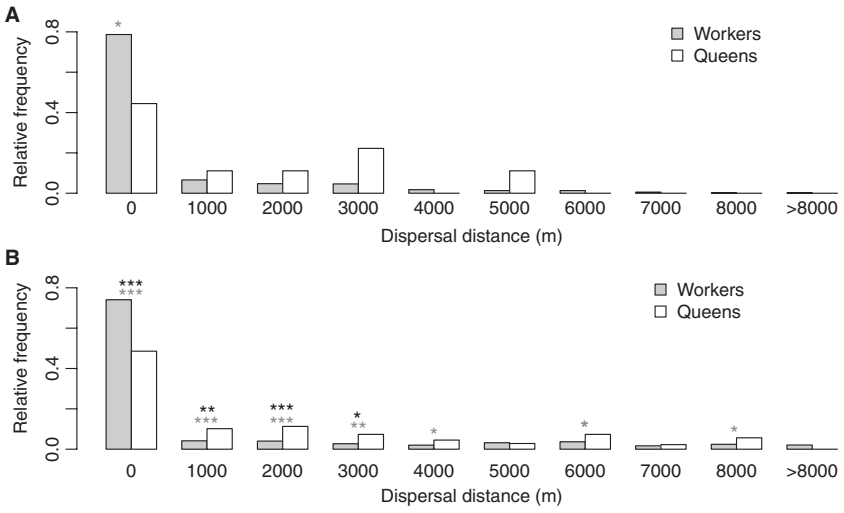

Fig. 2 Comparison of worker and queen dispersal distances in *Bombus lapidarius* (A) and *B. pascuorum* (B). Significance values are from a binomial exact test comparing the frequency of worker dispersal (null hypothesis) against queen dispersal in each 1000 m distance class. Stars indicate the *P*-value of the test before (grey) and after (black) sequential Bonferroni correction for multiple tests, *<0.05, **<0.01, ***<0.001.

different sites (case 3). Lastly, 50 nests contained individuals sampled in more than two sites (case 4). Although nests represented by only one individuals comprised 5% of workers and queen individuals, 51.4% of the queens (37.2% case 3 and 14.1% case 4) compared to 26.7% of the workers (19.3% case 3 and 7.4% case 4) were found to belong to nests that had individuals located in more than one sample site (Table 5).

Estimation of queen dispersal distance. A total of 919 *B. lapidarius* (910 workers but only nine queens) were used to estimate queen dispersal distance (Table 5, Fig. 2A). The median dispersal distances of workers and queens were 0 and 1820.4 m, respectively, with a significant difference between the two groups (two-sample Wilcoxon test, $W = 2690.5$, $P = 0.008$). In this case, the relative frequency of dispersal decreased with increasing dispersal distance (Fig. 2A). While a significant number of workers seemed to forage over a distance of 3 km, a few dispersal distances higher than 6 km were observed, consistent with the high power of the genetic data set for FSR in this species. We did not find a significant difference in movement distance between workers and queens, primarily because of the low number of sampled queens for this species. We detected however one dispersal event at 1 km, two at 3 km, one at 2 km and one at 5 km for *B. lapidarius* queens (Fig. 2A).

A total of 1620 *B. pascuorum* (1443 workers and 177 queens) were used to estimate queen dispersal distance (Table 5, Fig. 2B). The median dispersal distances of workers and queens were 0 and 1265 m, respectively, with significant difference between the two groups (two-sample Wilcoxon test, $W = 97860.5$, $P < 0.0001$). Although few *B. pascuorum* workers are expected to forage further than 1 km, we detected apparent worker movement for distances greater than 8 km (Fig. 2B),

consistently with false sister-pairs due to the relatively low power of the genetic data set for FSR in this species. The relative frequency of apparent dispersal of each distance class is the same, showing no trend to decrease with distance as would be expected in the case of real dispersal. We also detected apparent queen dispersal up to 8 km, but with a decreased in relative frequency of dispersal up to 5 km and fluctuations beyond this distance (Fig. 2B). Queen dispersal up to 3 km was significantly greater than worker foraging range, indicating that queens can disperse over at least 3 km in this species.

Discussion

Queen dispersal is likely to be a major contributor to gene flow in bumblebee species, an important pollinator group that will have to face increasing habitat fragmentation and range adjustment in response to climate warming. This study is the first attempt to directly estimate queen dispersal in bumblebee species. We evaluated FSR methods to find the most accurate method in our particular case of low family structure, and we applied the most accurate method to classify workers and queens into nests. We then used a conservative approach to estimate queen dispersal, taking into account the uncertainty in FSR. In *Bombus pascuorum* we demonstrated that queens disperse at least 3 km.

Application of FSR methods to haplodiploid species with low family structure

It seems that the genetic resolution needed for FSR in haplodiploid species was previously underestimated. Despite the apparent ease of testing for a relatedness of 0.75 against 0 when FSR is attempted in haplodiploid species, the analysis is particularly sensitive to missing data and genotyping error, both being types of uncertainty that are typically found when using large sample sizes and microsatellite markers. As we demonstrated, even with highly polymorphic markers, at least 10 microsatellites are needed to get accurate FSR results when the sample contains a low family structure. Even in a situation with a highly informative genetic data set, most methods appeared to be poor at reconstructing fullsib families. Only the KinGroup's DR method and Colony were found to approach accurate FSR for a simulated PF. We acknowledge that we did not include the likelihood ratio test implemented in Kinship (Goodnight & Queller 1999) and widely used for FSR in bumblebee species (Darvill *et al.* 2004; Knight *et al.* 2005, 2009; Ellis *et al.* 2006; Kokuvo *et al.* 2008). This was due to the pairwise nature of the method that renders its application very difficult with a large data set. Pairwise relationships have to be manually clustered into fullsib groups and noncircular relationships (individual 1 is sister of individual 2 and 3 but individual 2 is not sister to individual 3) have to be resolved (Knight *et al.* 2005) which is impracticable when analysing thousands of individuals. However, KinGroup DR is based on the same likelihood ratio test as Kinship (Goodnight & Queller 1999) and a test on a small simulated data set found that results using Kinship and KinGroup DR led to similar FSR (data not presented).

Besides the number and polymorphism of the genetic markers used, missing data and genotyping errors are critical factors that greatly reduced FSR accuracy for most of the tested methods. However, Colony (which takes into account genotyping error in its computations) yielded reasonable accurate FSR even in the situation of genotyping error rates as high as 2%. Previous simulations had demonstrated that an even higher genotyping error rate could be successfully handled by Colony using a high number of microsatellite markers (Wang 2004). However, as expected in the case of low family structure (i.e. most of the nests were represented by only one or a small number of individuals), parental genotype reconstructions were not accurate (Wang 2004). We found that parental genotype reconstruction was accurate only in the case of nests represented by more than five individuals which comprised a small proportion of the total nests (Fig. 1). Obviously, when a nest is represented by only one individual, which is the most frequent case, only one half of the genotype of the queen is observed, while it is impossible to discriminate between queen and male alleles. Thus such analysis should not be used to reconstruct parental genotypes in order to study population genetic structure as has been proposed and applied recently (Herrmann *et al.* 2007; Kraus *et al.* 2009). Inaccurately reconstructed genotypes could introduce biases in apparent population genetic structure and hence give misleading, spurious and inaccurate conclusions. This perhaps in part explains the unusual findings of Herrmann *et al.* (2007), who describe population structuring over small spatial scales in *B. pascuorum*. Using this approach would necessitate dealing with a sample containing at least five individuals from each nest to give accurate parental genotype reconstructions. This is unattainable in typical sampling of bumblebee species. A better alternative is to remove all but one individual from each sampled nest, and to perform population genetic analyses on this subsample.

Most previously published FSR analyses in bumblebees have aimed to describe the underlying family structure to estimate nest density or worker movements at the landscape level (Chapman *et al.* 2003; Darvill *et al.* 2004; Knight *et al.* 2005). Remarkably, the low efficiency of some FSR methods does not lead to a marked

family structure estimation bias. Most of the methods slightly overestimate the number of nests when the genetic data contain genotyping errors. On the contrary, Colony is inclined to underestimate the number of nests (Chapman *et al.* 2003; Wang 2004). However, because in our case we needed the most accurate reconstruction possible to estimate queen dispersal distance, we used Colony, bearing in mind that some detected fullsib relationships will be wrong.

Queens as an important gene flow route in Bombus *species*

Despite a quite similar genetic data set for the two studied species, the results from *B. lapidarius* were more consistent than those from *B. pascuorum*. This is somewhat frustrating because we were able to obtain many more queens in the latter species compared to the very small number of queens sampled for the former one. However, several results indicated that queen dispersal exceeds worker foraging distance in our FSR analyses. First, we found a higher proportion of queens, compared to workers, that were assigned to a nest outside the site where they were sampled (Table 5). We also found a higher proportion of queens that were classified alone in a nest (Table 5), probably indicating that more queens than workers migrated in from outside the study sites. Finally, we found evidence of a significant queen dispersal distance up to 3000 m for *B. pascuorum* and probably up to 5000 m for *B. lapidarius* (Fig. 2). The estimated queen dispersal distances are markedly higher than those of worker foraging distance estimated from less than 312 m (Darvill *et al.* 2004) to 449 m (Knight *et al.* 2005) for *B. pascuorum* and at 450 m for *B. lapidarius* (Knight *et al.* 2005). This is not surprising given the fact that our estimation of queen dispersal encompasses several dispersal steps (Goulson 2009) including the search for a mate, a hibernating place, a nest site and finally the foraging range. Thus, our estimation of dispersal includes the queen foraging once she has established a nest, which is not related to a component of the gene flow process. However, assuming that the foraging queens dispersed with the same probability in all directions centred around their nest, the observed dispersal distance should, on average, reflect dispersal *before* the nest founding by the queens.

It must be noted that our approach to estimating queen dispersal is unlikely to detect rare, long-distance movements, particularly as such movements could take queens out of the study area entirely. Hence our estimates are best regarded as a *minimum* estimate of the *maximum* dispersal distance. However, our estimation of queen dispersal over a few kilometres is in accordance with patterns of gene flow and population structure described in previous genetic studies of bumblebees. In the ubiquitous bumblebee species *B. terrestris* (Estoup *et al.* 1996) and *B. pascuorum* (Widmer & Schmid-Hempel 1999; Ellis *et al.* 2006), no evidence was found for isolation by distance or significant genetic differentiation between populations over large geographic areas, indicating ongoing gene flow and suggesting that these species exist as more-or-less continuous populations over much of Europe. The same is likely to be true for *B. pascuorum* and *B. lapidarius* in Britain, where one can envisage annual dispersal by queens over several kilometres being sufficient to maintain genetic cohesion across the country.

Interestingly, our estimation of queen dispersal distance falls within the range of those estimated for male *B. terrestris* from 2.2 to 9.9 km (Kraus *et al.* 2009). Similar dispersal distances for male of *B. lapidarius* and *B. pascuorum* would help explain the general pattern of high gene flow observed in the population structure of common bumblebee species.

Analysis of the population genetic structure in the rare and declining *B. muscorum* in Scottish islands demonstrated significant isolation by distance with significant genetic differentiation for populations 10 km apart (Darvill *et al.* 2006). This indicates that distances of this magnitude represent an efficient barrier to gene flow for this species, at least over water. Studies of the rare *B. sylvarum* showed a significant genetic differentiation between populations located more than 100 km apart (none of the sampled populations were closer together) (Ellis *et al.* 2006). These results are also in accord with our own, as if queen dispersal is typically in the range 1–5 km then we would expect populations 10 km or more apart to be isolated.

The annual rate of dispersal of invasive bumblebee species can also be compared to our estimation of queen dispersal distance. *Bombus terrestris* was introduced into Tasmania where it spread to occupy most of the island in ~7 years (Hingston *et al.* 2002; Hingston 2006; Schmid-Hempel *et al.* 2007). This represents a colonization rate from about 20 to 43 km per year (assuming either two or one generations per year, respectively). Although this fast expansion could in part be explained by release from the parasites found in the native range of *B. terrestris* (Allen *et al.* 2007), this observation indicates that bumblebee queens are able to migrate tens of kilometres from their natal nest to the nest foundation site. Such estimates are also consistent with the speed of invasion of this species in Hokkaido island in Japan (Kadoya *et al.* 2009), although in this case, the invasion history is not as clearly documented as the invasion of Tasmania because in Japan there are likely to have been multiple sites in which the species escaped from glasshouses.

Using fullsib reconstruction methods applied to a large sample of workers followed by subsequent sampling of queens the following year, we were able to estimate for the first time a minimum queen dispersal distance in two bumblebee species. We found that queens were able to disperse at least a few kilometres from the nest they originated from to establish their own nest. Such an estimated distance is in accordance with previous genetic analyses in bumblebee that indicate extensive gene flow in the common species and a geographic distance of tens of kilometres as an efficient barrier to gene flow for rare and declining species. It would clearly be valuable to carry out similar studies on rare species but our analysis of the accuracy of FSR suggests that this will not be possible with the inevitably smaller sample sizes and the reduced genetic variation which rare bumblebee species exhibit.

Our results have clear implications for management strategies for this important pollinator group, improving our understanding of the degree of habitat fragmentation over which common bumblebee species are likely to be able to maintain population cohesion.

Acknowledgements

We thank Niccolo Alfonso, Samantha Bailey, Giles Darvill, Penny Frith and Alex Wainwright for their help in collecting bumblebees. Help from Andrew Martin is much appreciated. Thanks to Rhys Green for providing details that helped to implement the nest landscape simulation. We thank two anonymous reviewers and the Subject Editor for their suggestions that improved the manuscript. This work was funded by BBSRC grant BB/E000932/1.

References

Allen GR, Seeman OD, Schmid-Hempel P, Buttermore RE (2007) Low parasite loads accompany the invading population of the bumblebee, *Bombus terrestris* in Tasmania. *Insectes Sociaux*, **54**, 56–63.

Amos W, Hoffman JI, Frodsham A *et al.* (2007) Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes*, **7**, 10–14.

Berg AM, Lascoux M, Pamilo P (1998) The infinite island model with sex-differentiated gene flow. *Heredity*, **81**, 63–68.

Bull JJ (1981) Coevolution of haplo-diploidy and sex determination in the Hymenoptera. *Evolution*, **35**, 568–580.

Chapman RE, Bourke AFG (2001) The influence of sociality on the conservation biology of social insects. *Ecology Letters*, **4**, 650–662.

Chapman RE, Wang J, Bourke AFG (2003) Genetic analysis of spatial foraging patterns and resource sharing in bumble bee pollinators. *Molecular Ecology*, **12**, 2801–2808.

Cook JM, Crozier RH (1995) Sex determination and population biology in the Hymenoptera. *Trends in Ecology & Evolution*, **10**, 281–286.

Crozier RH (1971) Heterozygosity and sex determination in haplo-diploidy. *The American Naturalist*, **105**, 399–412.

Darvill B, Knight ME, Goulson D (2004) Use of genetic markers to quantify bumblebee foraging range and nest density. *Oikos*, **107**, 471–478.

Darvill B, Ellis JS, Lye GC, Goulson D (2006) Population structure and inbreeding in a rare and declining bumblebee, *Bombus muscorum* (Hymenoptera:Apidae). *Molecular Ecology*, **15**, 601–611.

Ellis JS, Knight ME, Darvill B, Goulson D (2006) Extremely low effective population sizes, genetic structuring and reduced genetic diversity in a threatened bumblebee species, *Bombus sylvarum* (Hymenoptera: Apidae). *Molecular Ecology*, **15**, 4375–4386.

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.

Funk CR, Schmid-Hempel R, Schmid-Hempel P (2006) Microsatellite loci for *Bombus* spp. *Molecular Ecology Notes*, **6**, 83–86.

Goodnight KF, Queller DC (1999) Computer software for performing likelihood tests of pedigree relationship using genetic markers. *Molecular Ecology*, **8**, 1231–1234.

Goulson D (2009) *Bumblebees: Their Behaviour, Ecology and Conservation*, 2nd edn. Oxford University Press, Oxford, UK.

Goulson D, Lye GC, Darvill B (2008) Decline and conservation of bumble bees. *Annual Review of Entomology*, **53**, 191–208.

Herrmann F, Westphal C, Moritz RFA, Steffan-Dewenter I (2007) Genetic diversity and mass resources promote colony size and forager densities of a social bee (*Bombus pascuorum*) in agricultural landscapes. *Molecular Ecology*, **16**, 1167–1178.

Hingston AB (2006) Is the exotic bumblebee *Bombus terrestris* really invading Tasmanian native vegetation? *Journal of Insect Conservation*, **10**, 289–293.

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). *Austral 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.

Hornik K (2005) A CLUE for CLUster ensembles. *Journal of Statistical Software*, **14**, 1–25.

Kadoya T, Ishii HS, Kikuchi R, Suda S, Washitani I (2009) Using monitoring data gathered by volunteers to predict the potential distribution of the invasive alien bumblebee *Bombus terrestris*. *Biological Conservation*, **142**, 1011–1017.

Knight ME, Martin AP, Bishop S *et al.* (2005) An interspecific comparison of foraging range and nest density of four bumblebee (*Bombus*) species. *Molecular Ecology*, **14**, 1811–1820.

Knight ME, Osborne JL, Sanderson RA *et al.* (2009) Bumblebee nest density and the scale of available forage in arable landscapes. *Insect Conservation and Diversity*, **2**, 116–124.

Kokuvo N, Toquenaga Y, Goka K (2007a) A simple visualization method to reconstruct nest-mate patterns among bumble bees (Hymenoptera: Apidae) using genetic data. *Applied Entomology and Zoology*, **42**, 137–141.

Kokuvo N, Yoshimura S, Mano H, Toquenaga Y, Goka K (2007b) Modified SLCA algorithm for full-sib reconstruction of haplodiploid species. *Environmental Entomology*, **36**, 1523–1528.

Kokuvo N, Toquenaga Y, Goka K (2008) Estimating colony number of *Bombus terrestris* (Hymenoptera, Apidae) queens foraging in Biratori, Hokkaido, Japan. *Applied Entomology and Zoology*, **43**, 19–23.

Konovalov DA (2006) Accuracy of four heuristics for the full sibship reconstruction problem in the presence of genotype errors. *Proceedings of the 4th Asia-Pacific Bioinformatics Conference*, **3**, 7–16.

Konovalov DA, Manning C, Henshaw MT (2004) KINGROUP: a program for pedigree relationship reconstruction and kin group assignments using genetic markers. *Molecular Ecology Notes*, **4**, 779–782.

Konovalov DA, Bajema N, Litow B (2005) Modified SIMPSON $O(n^3)$ algorithm for the full sibship reconstruction problem. *Bioinformatics*, **21**, 3912–3917.

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.

Macior LW (1968) *Bombus* (Hymenoptera, Apidae) queen foraging in relation to vernal pollination in Wisconsin. *Ecology*, **49**, 20–25.

Macior LW (1994) Pollen-foraging dynamics of subalpine bumblebees (*Bombus* Latr.). *Plant Species Biology*, **9**, 99–106.

Malcolm JR, Markham A, Neilson RP, Garaci M (2002) Estimated migration rates under scenarios of global climate change. *Journal of Biogeography*, **29**, 835–849.

Osborne JL, Martin AP, Carreck NL *et al.* (2008) Bumblebee flight distances in relation to the forage landscape. *Journal of Animal Ecology*, **77**, 406–415.

R Development Core Team (2005) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Rubin J (1967) Optimal classification into groups – an approach for solving taxonomy problem. *Journal of Theoretical Biology*, **15**, 103–144.

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.

Wang JL (2003) Maximum-likelihood estimation of admixture proportions from genetic data. *Genetics*, **164**, 747–765.

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

Wang JL (2006) Informativeness of genetic markers for pairwise relationship and relatedness inference. *Theoretical Population Biology*, **70**, 300–321.

Wang J, Santure AW (2009) Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics*, **181**, 1579–1594.

Washitani I, Kato M, Nishihiro J, Suzuki K (1994) Importance of queen bumble bees as pollinators facilitating inter-morph crossing in *Primula sieboldii*. *Plant Species Biology*, **9**, 169–176.

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

Williams PH, Osborne JL (2009) Bumblebee vulnerability and conservation world-wide. *Apidologie*, **40**, 367–387.

O.L. is a molecular ecologist interested in the use of genetic markers to study hybridisation and dispersal. B.D.'s research is concerned with the population structure, ecology and conservation of bumblebees. S.O. is currently conducting research on the nesting ecology of bumblebees. J.L.O. has diverse research interests relating to the ecology of bees. R.A.S. is an ecological modeller. J.C. is a plant and pollination ecologist. L.G. is focussed on using remote-sensed data to examine habitat use by bumblebees. D.G. is an insect ecologist specializing in studies of bumblebees, their behaviour, ecology and conservation.
