Bumblebee family lineage survival is enhanced in high quality landscapes

Claire Carvell[1]*, Andrew F. G. Bourke[2], Stephanie Dreier[3,4], Stephen N. Freeman[1], Sarah Hulmes[1], William C. Jordan[3†], John W. Redhead[1], Seirian Sumner[3,4,5], Jinliang Wang[3] and Matthew S. Heard[1]

[1]NERC Centre for Ecology & Hydrology, Benson Lane, Crowmarsh Gifford, Wallingford, OX10 8BB, UK

[2]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK

[3]Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK

[4]School of Biological Sciences, Bristol Life Sciences Building, University of Bristol, 24 Tyndall Avenue, Bristol BS8 1TQ, UK

[5]Centre for Biodiversity and Environment Research, University College London, Gower Street, London WC1E 6BT, UK

*****Corresponding author Email: ccar@ceh.ac.uk

†deceased

Summary paragraph

Insect pollinators such as bumblebees (*Bombus* spp.) are in global decline[1,2], a major cause of which is habitat loss due to agricultural intensification[3]. A range of global and national initiatives aimed at restoring pollinator habitats and populations have been developed[4-6]. However, the success of these initiatives depends critically upon understanding how landscape change affects key population-level parameters, such as survival between lifecycle stages[7], in target species. Such understanding is lacking for bumblebees because of the difficulty of systematically finding and monitoring colonies in the wild. We used a novel combination of habitat manipulation, land-use and habitat surveys, molecular genetics[8] and demographic and spatial modelling to examine between-year survival of family lineages in field populations of three bumblebee species. Here we show that the survival of family lineages from the summer worker to the spring queen stage in the following year increases significantly with the proportion of high-value foraging habitat, including spring floral resources, within 250-1000 m of the natal colony. This is the first evidence of a positive impact of habitat quality on survival and persistence between successive colony cycle stages in bumblebee populations. The findings provide strong support for conservation interventions that increase floral resources at a landscape scale and throughout the season having positive effects on wild pollinators in agricultural landscapes.

Main text

The loss of semi-natural habitats and floral resources within intensively managed agricultural landscapes has been identified as a major driver of declines in insect pollinators[3,9,10], with negative consequences for crop pollination[11]. Habitat restoration (e.g. the planting of flowering hedgerows, meadows or flower strips along field margins under agri-environment schemes[12]) can mitigate these effects, increasing local pollinator abundance and species richness[13-15] and enhancing rates of persistence and colonization at the community level[16]. However, we lack understanding of the effects of restoration on key aspects of pollinator biology that may explain the mechanisms behind these responses. In particular, improving habitat quality might be expected to enhance the prospects of successful reproduction and between-year survival in targeted areas, but whether this occurs is unknown.

Bumblebees (*Bombus* spp.) are key pollinators of wild flowers and commercial crops[17,18]. Following a eusocial, annual colony cycle, new queens enter hibernation in the autumn and emerge in spring to search for a nest site and found a colony. Each colony may produce up to several hundred 'daughter' workers, which forage from spring to summer at flowers for nectar and pollen to rear new daughter queens and males[19]. The survival and dispersal patterns of bumblebee queens during hibernation and nest-searching periods are critical to overall population persistence, but remain undescribed in wild populations[8,20,21]. In addition, although the availability of floral resources within foraging distance of the nest has been shown to increase numbers of workers and males produced per colony, effects on queen production have been less clear[22] and there is no evidence regarding how queen production, survival and dispersal may be linked with underlying habitat quality and land-use[23].

Here, we investigated the effects of habitat quality and land-use on bumblebee survival and dispersal between colony cycle stages across two years. We first tested whether colonies located within or near high-value foraging habitats had a greater probability of producing daughter queens that survive the winter hibernation and spring emergence stages, henceforth termed 'family lineage survival'. Second, we tested whether the distances travelled by queens between hibernation and nest-searching periods (as a measure of minimum relative queen dispersal distances within our study landscape) were affected by the proportion of high quality habitat surrounding their natal colony. We sampled DNA non-lethally from 537 spring queens (in 2011 and 2012) and 2,101 workers (in 2011) of three widespread species (*Bombus terrestris*, *B. lapidarius* and *B. pascuorum*) at a fine spatial scale across a 20-km^2 agricultural landscape in southern England, UK. The landscape was dominated by arable fields and permanent intensive grassland but also included a range of habitat restoration measures for pollinators[24], which resulted in a heterogeneous matrix with areas of high and low proportions of flower-rich habitat[15]. Sampled queens and workers were genotyped at 13-14 microsatellite loci per species and family relationships between them were estimated using maximum likelihood sibship reconstruction[8]. This revealed a total of 1,665 family lineages within our sample. Relationships were detected within and between generations in the colony cycle: between spring (2011) queens and their daughter (2011) workers (i.e. within a single colony cycle), and between spring (2011) queens and their daughter (2012) queens sampled the following year or summer (2011) workers and their sister (2012) queens sampled the following year (i.e. between two successive colony cycles, Fig. 1; Extended Data Table 1). These relationships were used to estimate family lineage survival in terms of rates of 'apparent survival'[25], i.e. the probability that a family lineage survives and remains available for capture within the study landscape (see Methods and Extended Data Tables 2, 3). Family lineage survival between the summer (2011) worker and spring (2012) queen stages was then related to measures of habitat quality and land-use at four spatial scales (relative to colony locations estimated from the distributions of sampled sister workers[24]). We focused on the family lineage relationship between the summer worker and spring queen stage, since only data from summer workers permitted the estimation of colony locations (see Methods).

We found that, across all three bumblebee species, habitat quality and land-use variables were significantly positively correlated with the between-year survival of family lineages (Extended Data Tables 4-6). Mixed semi-natural vegetation cover, queen-visited spring flower cover (as provided, for example, by flowering trees and hedgerow plants) and summed spring and summer flower cover for bumblebee-visited species had highly significant positive effects on family lineage survival within radii of 250 m, 500 m and 1,000 m from estimated colony locations (Fig. 2; Extended Data Table 4). While the individual colony locations estimated using our approach may be subject to some error, we found no evidence to suggest any systematic bias[24] which would affect relationships with habitat across the large sample sizes that were analysed. The strongest relationship was for mixed semi-natural vegetation cover (including sown field margins) within a radius of 1,000 m (slope = 26.17; LRT statistic = 11.34; $P < 0.001$; Fig. 2a). We also found significant positive effects of high-value foraging habitats (queen-visited spring flower cover and summed spring and summer flower cover) within radii equal to the estimated colony-specific foraging distance (Fig. 2; Extended Data Table 4). Given that colony-specific foraging distances are reduced as floral resources increase[24], this result suggests that when workers were able to forage closer to their colonies, greater numbers of queens from those colonies survived the winter and spring emergence periods. Our findings also suggest that family lineage survival is particularly sensitive to small changes in landscape composition (Fig. 2).

Family lineage survival was not significantly influenced by all spring or all summer flower cover (flower cover of all surveyed plant groups in either season), worker-preferred summer flower cover or nesting habitat cover within the ranges of variation tested across the study landscape at any modelled radius (Extended Data Tables 4 and 5). This suggests that summer-flowering resources at the worker stage alone are not sufficient to sustain colonies throughout their cycle. Moreover, it suggests that flower cover of particular plant groups utilised as forage resources (Extended Data Table 6) is more important than overall flower cover within the habitats surrounding a colony. We did, however, find a weak but significant positive effect on family lineage survival of arable field cover within 1,000 m of estimated colony locations (Extended Data Table 4); this is most likely due to the presence of spring-flowering oilseed rape and summer-flowering field bean crops across the landscape (Extended Data Table 5).

Across all three bumblebee species, the mean (± s.e.) distance of sampled 2012 spring queens from their estimated natal colony location was 1227 ± 125 m . This distance, taken as a measure of minimum relative queen dispersal distance, is between two to three times greater than the typical foraging distances estimated for workers within the same landscape[24]. Mean minimum relative queen dispersal distances were greatest for *B. terrestris* (1553 ± 223 m, from n = 15 colonies), and lower for *B. pascuorum* (1149 ± 273 m, n = 12) and *B. lapidarius* (980 ± 148 m, n = 16). Although these differences were not statistically significant (ANOVA, $F_{2,42}$ = 2.07; P = 0.14), the three bumblebee species differed significantly in the proportion of 2012 spring queens from family lineages not represented in the previous year's dataset (82% in *B. terrestris* and 56% in both *B. pascuorum* and *B. lapidarius*, $\chi^2_2 = 31.06;\ P < 0.01$) (Extended Data Table 1). Taken together, these findings suggest that *B. terrestris* is the most widely dispersing of the three species.

Despite this variation in minimum relative queen dispersal distance within species and, potentially, between species, there were no significant correlations at any spatial scale between this distance and any seasonal flower cover variable, mixed semi-natural vegetation cover or arable field cover (Extended Data Table 5). However, there were significant positive correlations between queen dispersal distance and nesting habitat cover at radii of 250 m, 500 m and 1,000 m (Kendall's tau correlation coefficients = 0.309, 0.308 and 0.331, respectively; df = 41; P < 0.05). This suggests that while high quality foraging habitats may not extend the dispersal distances of queens, non-crop habitats suitable for nesting may facilitate queen movement into the wider landscape.

Our study is the first to demonstrate a positive impact of habitat quality on survival and persistence between successive colony cycle stages in wild pollinators. Given that two of the habitat variables having the strongest influence, mixed semi-natural vegetation and summed spring and summer flower cover, included sown field margins (Fig. 2; Extended Data Table 4), this impact includes effects of habitat restoration via the implementation of agri-environment schemes. The study provides evidence that floral resources in spring-flowering trees, hedgerow plants and crops are particularly vital for bumblebee queens during their emergence and colony founding, and, in combination with summer floral resources, enhance the representation of colonies in the population the following year. It therefore adds to previous studies in highlighting the importance of temporally sustained floral resources within 1 km of nests for within-season survival and performance of bumblebee colonies[22,26,27], especially spring-flowering resources, which are often overlooked in conservation intervention options[28]. Furthermore, our findings suggest that, as well as sustaining colonies, appropriately managed non-crop areas can act as a source of queens to the wider landscape[29]. While there is an urgent need for systematic monitoring of pollinator populations to provide more robust data on patterns and causes of decline[30], our study provides strong support for conservation interventions targeted at a landscape scale having a positive impact on wild pollinators in agricultural landscapes.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements. We thank the CEH field team (Lucy Hulmes, Jodey Peyton, Jo Savage, Sam Amy, Roselle Hyman, Gemma Baron and Rachel MacDonald) for sampling bumblebees and conducting habitat surveys; Robin Faccenda and Richard Franklin of Faccenda Farms, and other landowners, for access to the Hillesden Estate and surroundings; Richard Pywell and Marek Nowakowski for access to the Hillesden Experimental Platform; Hannah Dean for data management; Colin Harrower for help with graphics; H. Michael Lattorff for use of his primers for the molecular discrimination of *B. terrestris* and *B. lucorum* workers; Ian Warren for assistance with laboratory work; and James Bullock and Karsten Schonrogge for comments on an earlier version of the manuscript. This research was supported by the Insect Pollinators Initiative (grant BB/I000925/1). The Insect Pollinators Initiative was funded jointly by the Biotechnology and Biological Sciences Research Council, the Department for Environment, Food and Rural Affairs, the Natural Environment Research Council, The Scottish Government and The Wellcome Trust, under the Living with Environmental Change Partnership. Acquisition of remote sensing data was funded by Syngenta Plc.

Author Contributions. C.C., M.H., A.B., S.S. and W.J. (the late Bill Jordan) conceived the project and designed the study. C.C., M.H. and S.H. coordinated the fieldwork and modelling elements and C.C. prepared the manuscript. S.D. carried out the molecular genotyping and sibship assignments with guidance from J.W. J.R. developed and applied the spatial analyses and S.F. designed and undertook the statistical analyses. All authors contributed to writing and critiquing the manuscript.

References

1 Williams, P. H. & Osborne, J. L. Bumblebee vulnerability and conservation world-wide. *Apidologie* **40**, 367-387 (2009).

2 Cameron, S. A. *et al.* Patterns of widespread decline in North American bumble bees. *Proceedings of the National Academy of Sciences* **108**, 662-667, doi:10.1073/pnas.1014743108 (2011).

3 Vanbergen, A. J. & TheInsectPollinatorsInitiative. Threats to an ecosystem service: pressures on pollinators. *Frontiers in Ecology and the Environment* **11**, 251-259, doi:10.1890/120126 (2013).

4 Defra. The National Pollinator Strategy: for bees and other pollinators in England https://www.gov.uk/government/publications/national-pollinator-strategy-for-bees-and-other-pollinators-in-england. (2014).

5 http://www.nrcs.usda.gov/wps/portal/nrcs/main/national/plantsanimals/pollinate/.

6 Gill, R. J. *et al.* Protecting an Ecosystem Service: Approaches to Understanding and Mitigating Threats to Wild Insect Pollinators. *Advances in Ecological Research* **54**, 135-206, doi:http://dx.doi.org/10.1016/bs.aecr.2015.10.007 (2016).

7 Carvell, C. *et al.* Bumble bee species' responses to a targeted conservation measure depend on landscape context and habitat quality. *Ecological Applications* **21**, 1760 -1771 (2011).

8 Dreier, S. *et al.* Fine-scale spatial genetic structure of common and declining bumble bees across an agricultural landscape. *Molecular Ecology* **23**, 3384-3395, doi:10.1111/mec.12823 (2014).

9 Baude, M. *et al.* Historical nectar assessment reveals the fall and rise of floral resources in Britain. *Nature* **530**, 85-88, doi:10.1038/nature16532 (2016).

10 Winfree, R., Aguilar, R., Vazquez, D. P., LeBuhn, G. & Aizen, M. A. A meta-analysis of bees' responses to anthropogenic disturbance. *Ecology* **90**, 2068-2076, doi:doi:10.1890/08-1245.1 (2009).

11 Garibaldi, L. A. *et al.* Stability of pollination services decreases with isolation from natural areas despite honey bee visits. *Ecology Letters* **14**, 1062–1072 (2011).

12 Natural England. *Countryside Stewardship Manual*. (Natural England Catalogue Code NE608 (2015).

13 Carvell, C., Meek, W. R., Pywell, R. F., Goulson, D. & Nowakowski, M. Comparing the efficacy of agri-environment schemes to enhance bumble bee abundance and diversity on arable field margins. *Journal of Applied Ecology* **44**, 29-40 (2007).

14 Wood, T. J., Holland, J. M., Hughes, W. O. H. & Goulson, D. Targeted agri-environment schemes significantly improve the population size of common farmland bumblebee species. *Molecular Ecology* **24**, 1668-1680, doi:10.1111/mec.13144 (2015).

15 Pywell, R. F. *et al.* Wildlife-friendly farming increases crop yield: evidence for ecological intensification. *Proceedings of the Royal Society of London B: Biological Sciences* **282**, doi:10.1098/rspb.2015.1740 (2015).

16 M'Gonigle, L. K., Ponisio, L. C., Cutler, K. & Kremen, C. Habitat restoration promotes pollinator persistence and colonization in intensively managed agriculture. *Ecological Applications* **25**, 1557-1565, doi:10.1890/14-1863.1 (2015).

17 Klein, A.-M. *et al.* Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences* **274**, 303-313, doi:10.1098/rspb.2006.3721 (2007).

18 Garratt, M. P. D. *et al.* The identity of crop pollinators helps target conservation for improved ecosystem services. *Biological Conservation* **169**, 128-135, doi:http://dx.doi.org/10.1016/j.biocon.2013.11.001 (2014).

19 Benton, T. *Bumblebees: the natural history & identification of the species found in Britain*. (Collins, 2006).

20 Beekman, M., van Stratum, P. & Lingeman, R. Diapause survival and post-diapause performance in bumblebee queens (Bombus terrestris). *Entomologia Experimentalis Et Applicata* **89**, 207-214 (1998).

21 Lepais, O. *et al.* Estimation of bumblebee queen dispersal distances using sibship reconstruction method. *Molecular Ecology* **19**, 819-831 (2010).

22 Williams, N. M., Regetz, J. & Kremen, C. Landscape-scale resources promote colony growth but not reproductive performance of bumble bees. *Ecology* **93**, 1049 – 1058 (2012).

23 Jha, S. & Kremen, C. Urban land use limits regional bumble bee gene flow. *Molecular Ecology* **22**, 2483-2495, doi:10.1111/mec.12275 (2013).

24 Redhead, J. W. *et al.* Effects of habitat composition and landscape structure on worker foraging distances of five bumble bee species. *Ecological Applications* **26**, 726-739, doi:10.1890/15-0546 (2016).

25 Lebreton, J. D., Burnham, K. P., Clobert, J. & Anderson, D. R. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs* **62**, 67-118 (1992).

26 Goulson, D. *et al.* Effects of land use at a landscape scale on bumblebee nest density and survival. *Journal of Applied Ecology* **47**, 1207-1215 (2010).

27 Jha, S. & Kremen, C. Resource diversity and landscape-level homogeneity drive native bee foraging. *Proceedings of the National Academy of Sciences* **110**, 555-558, doi:10.1073/pnas.1208682110 (2013).

28 Dicks, L. V. *et al.* How much flower-rich habitat is enough for wild pollinators? Answering a key policy question with incomplete knowledge. *Ecological Entomology* **40**, 22-35, doi:10.1111/een.12226 (2015).

29 Carvell, C., Bourke, A. F. G., Osborne, J. L. & Heard, M. S. Effects of an agri-environment scheme on bumblebee reproduction at local and landscape scales. *Basic and Applied Ecology* **16**, 519-530, doi:http://dx.doi.org/10.1016/j.baae.2015.05.006 (2015).

30 Nieto, A. *et al.* European Red List of bees. (Luxembourg: Publication Office of the European Union., 2014).

Figure 1. **Bumblebee colony cycle stages (grey boxes) and family lineages sampled in the study with estimated survival parameters**. **a)** Relationships (dotted lines, italics) assigned using sibship reconstruction from non-lethal DNA sampling. **b)** Family lineage survival (arrows: between the mother queen and daughter worker stage in spring and summer 2011 (φ_1), and between the summer worker and sister queen stage in spring 2012 (φ_2), with estimated asymptotic 95% confidence limits) and recapture probability (p_1, p_2) parameters (see Extended Data Table 3). Parameter values shown are (as examples) mean values across 1,665 family lineages of all species (*Bombus terrestris*, *B. lapidarius* and *B. pascuorum*). Only females (shown in red) were sampled; hollow female symbol denotes queens, and female symbols containing a star denote workers; males (shown in blue) were not sampled.

Figure 2. **Effects of habitat quality and land-use variables on bumblebee family lineage survival from the summer worker to spring queen stage (parameter φ_2 in Fig. 1b).** Solid line shows model-fitted logistic regression; dashed lines show 95% confidence intervals. Predictor variables are proportions of: a) mixed semi-natural vegetation cover (including sown field margins) within 1,000 m of colony locations ($\chi^2_1 = 11.34$; $P < 0.001$); b) queen-visited spring flower cover within colony-specific foraging distance ($\chi^2_1 = 9.52$; $P < 0.01$); c) summed spring and summer flower cover within colony-specific foraging distance ($\chi^2_1 = 7.2$; $P < 0.01$); and d) arable field cover within 1,000 m of colony locations ($\chi^2_1 = 4.3$; $P < 0.05$). See Extended Data Table 5 for full descriptions of predictor variables. Relationships at radii showing best model fit for each variable are presented; see Extended Data Table 4 for model results for all variables at all radii. Data generated from 456 wild colonies of *Bombus terrestris* (n = 69), *B. lapidarius* (n = 267) and *B. pascuorum* (n = 120). Model comparisons showed no significant differences between species in apparent survival at this stage, hence data from all species were combined (see Methods).

METHODS

Mapping the study landscape. The study was conducted across an agricultural landscape covering 20 km^2, centred on the Hillesden Estate, Buckinghamshire, UK (1˚00'01"W; 51˚57'16"N)[8]. The Estate consists of a c. 1,000 ha intensive arable farm on which a number of experimental targeted habitat restoration options (including the sowing of wildflower mixtures for pollinators along field margins) have been established and managed since 2005 under the English agri-environment scheme[12]. These 'sown field margins' comprised 2% of the total area of the study landscape, although their density was manipulated spatially so that, per 50-60 ha of land, the area taken out of production varied systematically from 0-8%[15]. The landscape surrounding the Estate was predominantly arable, with some areas of permanent intensive grassland, woodland and small villages.

Detailed habitat maps of the study landscape were generated using a land use/land cover (LULC) map derived from two airborne remote sensed sources - Light Detection and Ranging (LiDAR) and hyperspectral imaging[31]. These data were combined and classified to form a high-resolution (0.5 x 0.5 m pixels) LULC map with each pixel assigned to one of 9 land cover classes (arable; short grass; non-woody semi-natural mixed vegetation; agri-environmental field margin; garden and urban vegetation; woody vegetation; road and building; water; and bare soil). The study landscape was surveyed systematically in terms of its value for bumblebees in both spring and summer (during and immediately following our sampling of queens and workers, respectively). The LULC map was converted to vector format and every resultant LULC polygon which could be distinguished as a discrete habitat parcel in the field was surveyed in July and August 2011, to estimate a) the percentage cover of each plant species, family or group (hereafter plant groups), and b) the proportion of that plant group in flower at the time of the survey. These variables (a and b) were multiplied for each plant group to give a measure of the proportion of each habitat parcel covered with flowers. To construct a spring habitat map, we surveyed a stratified random sub-sample of parcels across all habitat types in April 2011 and 2012. These data were then used to estimate plant and floral cover values in unsampled parcels by adjusting species cover estimates from the full summer dataset.

An assessment of the suitability of each land parcel as nesting habitat for bumblebees was made by estimating average vegetation height (m) across the whole parcel, whether tussocky vegetation was present, the extent of plant litter or moss within the sward and whether there were signs of small mammal activity such as the presence of burrows, runs or faeces. Species-specific nesting requirements (classified a priori using expert knowledge and published sources[19,32]) were then used to categorise each parcel as being of high, medium or low nesting habitat suitability for each of the three study species.

The vector LULC map was updated using the digitized, completed survey maps, and linked to the floral and nesting data. In total, 18.7 km^2 of the study landscape were surveyed in this way. Where parcels were not surveyed (due to access restrictions), plant cover values were estimated by taking the average value of covers from parcels of the same LULC class within a 500 m radius. The final habitat dataset thus consisted of all discrete parcels with information on cover and floral cover of all surveyed plant groups[31]. Floral cover of surveyed plant groups was further summed in terms of the plants' relative value as forage resources for bumblebees, according to whether they had been observed to be visited or not visited by foraging queens or workers during bee sampling, and, if visited, whether they were classified as 'preferred' in the worker dataset (Extended Data Table 6). Handling of the LULC map and survey data was performed in ArcMAP v10.0 (© ESRI, Redlands, CA, USA).

Study species. We studied three social bumblebee (*Bombus*) species that are common and widespread across much of the UK but vary in their forage plant choice and nesting behaviour. *Bombus terrestris* L. and *B. lapidarius* L. typically nest underground in large colonies (reaching up to 300 workers at maturity) and have shorter-tongued workers that visit a wide range of flowers. *B. pascuorum* (Scopoli) nests on the ground surface, usually within mossy grass tussocks tending to form smaller colonies (reaching between 100 and 200 workers) with longer-tongued workers that specialize in foraging at flowers with long corolla tubes[19]. There is limited evidence regarding numbers of queens produced in wild colonies in the UK, and these are likely to vary widely within and between species but, in *B. terrestris*, studies suggest averages of 4, 14 and 35 queens produced per colony[33-35].

Bumblebee sampling and genotyping. The study area was divided into 250 × 250 m grid cells and within every cell sampling intensity (i.e. search effort) for both queens and workers was proportional to the relative cover of suitable habitats present. Hence searches were more focussed on field boundaries and other non-crop habitat parcels (defined areas of continuous land use) but did include field interiors. All female individuals of the three study species encountered were caught for DNA sampling. At the same time, their locations were recorded using a GPS device and their behaviours noted (whether nest-searching, in flight or foraging and, if foraging, which plant species was being visited). Sampling was performed at this fine spatial scale to maximise the likelihood of detecting sister workers at multiple sites [36] and to ensure a high proportion of colonies were sampled across the landscape.

We obtained DNA samples non-lethally by clipping the tarsal tip of a mid-leg of each bee and preserving it in 100% ethanol[37]. If a bee was encountered that had already been sampled, which occurred in three queens and <10 workers, we identified it to individual level by taking a second DNA sample from the basitarsus of the same mid-leg that had already been clipped and seeking a match for its multilocus genotype in the dataset. Sampling was carried out between 09:00h and 17:00h during dry weather when ambient temperature was above 11°C with at least 60% clear sky, or above 15°C under any sky conditions.

Queens were sampled across the study landscape from 21 March to 18 April 2011 and from 19 March to 2 May 2012, reflecting the main periods of spring emergence and nest founding activity observed for the three study species. Workers of the three species were sampled continuously (during 4-5 days per week) between 20 June and 5 August 2011 across all the habitats in the study landscape. We used PCR-based molecular identification of mtDNA markers to differentiate *B. terrestris* workers from any *B. lucorum* workers present in our sample (H.M.G. Lattorff, personal communication), since workers of the two species are difficult to separate reliably in the field. DNA was isolated from each tarsal sample using the HotSHOT protocol, and all individuals were genotyped at 14 (*B. terrestris* and *B. pascuorum*) or 13 (*B. lapidarius*) microsatellite loci (see Dreier et al., 2014[8,38]). Missing data were rare across both queen and worker genotypes, with all except a single individual worker of *B. terrestris* included in analyses being successfully typed at 12 or more loci. The mean missing genotype frequency per individual across all species was 0.002.

Assigning family relationships. COLONY version 2.0[39] was used to reconstruct family relationships between the bees in our sample for each species (Fig. 1a). COLONY implements a full-likelihood approach to sibship analysis, and yields the best (maximum likelihood) and possible alternative (high likelihood) estimates of family relationships with corresponding estimated posterior probabilities. The inferred relationships with an estimated posterior probability of 0.8 or higher were accepted and used in downstream analyses [8]. Given previous work[40] we assumed a monogamous mating system for males and females, therefore allowing the assignment of full-siblings, mothers and daughters. We carried out a medium-length run with medium-likelihood precision, using genotyping error rates of 0-5% based on results of regenotyping and rescoring 10% of randomly selected individuals[8]. Two replicate COLONY runs were conducted across the full sample of queens and workers, each with a different random number seed but with all other parameters kept equal (membership of sibship families was identical in both COLONY runs and the variance between the estimated probabilities of inference was small (<0.002)). Reconstructed relationships included those between sister workers sampled in summer 2011 and between: a) founding queens sampled in spring 2011 and their daughter workers sampled in summer 2011; b) queens sampled in spring 2011 and their daughter queens sampled in spring 2012; and c) workers sampled in summer 2011 and their sister queens sampled in spring 2012 (Fig. 1a).

Estimating colony locations and surrounding landscape quality. The location of each sampled queen and worker was mapped from GPS locations in ArcGIS. We estimated the location of each colony from the distributions of full-sib sister workers, using a mean centre approach[24]. Model simulation showed that this approach was not biased by either outlying worker locations or clusters of workers at particular foraging patches, and it required no additional parameters or prior assumptions (e.g. regarding likely foraging distances)[24,36]. Colony locations were estimated only for colonies represented by inferred sibships of two or more workers, since it is not possible to assign a biologically meaningful nest location for colonies represented by single workers.

Mean centre locations were ‘snapped’ (ie. moved to coincide exactly with the coordinates of another feature) to the nearest LULC class that might have formed suitable nesting habitat for bumblebees, thus avoiding cropped arable fields, roads, buildings and water[19]. Overall, 214 estimated colony locations required snapping (47% of the sample); 208 of these were from cropped arable fields (mean snapping distance, i.e. distance between original location and ‘snapped’ location, = 47.2 m) and 6 were from the ‘roads and buildings’ land-use category (mean snapping distance = 4.9 m). The straight-line distance of each worker from its capture location to estimated colony location was calculated and the mean of these distances for all workers in a sibship was used as a measure of ‘colony-specific foraging distance’[24]. The snapping process made a mean difference to colony-specific foraging distances of only 4.3 m.

Furthermore, random resampling of pairs of sister workers from sibships with more than two workers showed that sibship size had only a minor influence on estimated colony locations and foraging distances; hence there was no evidence to suggest a systematic bias that might have affected relationships with habitat[24].

Measures of habitat quality and land-use at different spatial scales surrounding each colony were made by creating a series of buffers with the following radii around the colony location: colony-specific foraging distance; 250 m; 500 m; and 1,000 m. Within each buffer, the proportion of each LULC class, floral cover of different subsets of plant groups in spring and summer (including queen-visited and worker-preferred plant groups), and cover of suitable nesting habitats were then determined. Mixed, non-woody, semi-natural vegetation, agri-environmental sown field margins and other linear habitats were combined to a single 'mixed semi-natural vegetation' class (Extended Data Table 5). The range of variation in these variables across our study landscape, and correlations between them, are shown in Extended Data Figure 1 for the proportion of each variable within 1,000 m of colony locations.

Estimating queen dispersal distance. The distances travelled by queens of *B. terrestris*, *B. lapidarius* and *B. pascuorum* between departure from their natal nest in late summer and post-hibernation colony foundation in the spring were estimated by measuring geographic distances between the 2011 colony locations and their inferred sister queens sampled in spring 2012. For families in which two or more queens were sampled, the average distance between queens and the natal colony location was calculated. We did not estimate the nest-site locations of founding queens at the time of spring sampling, since it was not possible to accurately separate queens that had already founded nest-sites and were foraging from them from queens that were actively searching for nest sites i.e. still in their dispersal phase.

Statistical analyses. We developed a novel extension of the standard Cormack-Jolly-Seber (CJS) mark-recapture model[25] to estimate survival rates and 'recapture' probabilities (i.e. the probability of a daughter worker or queen being sampled from a given family lineage) using data on the numbers of individuals sampled at every stage for every family lineage of each species (Figure 1 and Extended Data Table 1). This allowed us to estimate family lineage survival, corrected for under-detection of individuals arising, for example, from long-distance dispersal or non-capture of queens. Where the standard CJS model has recaptures relating to individual animals, our modified model related recaptures to different individuals from the same family lineage and incorporated counts of the number of individuals (full-sib workers or sister queens) sampled at each stage. A full step-by-step account of model construction is given in Supplementary Information (SI) under section ‘A novel mark-recapture model for colonial species’, along with the code used for model fitting.

The standard CJS model of Equation (1) was first fitted for the three *Bombus* species separately and for all species combined (n = 1,665 family lineages) using the software package MARK[41]. For *B. lapidarius* no first-generation queens were captured and the model (1) collapsed to a simple binomial. We therefore estimated the single parameter for this species in isolation via a GLM routine. Performance of this simple, standard CJS model proved to be poor and only φ_1 and p_1 were uniquely estimable (Extended Data Table 2). We therefore improved the model using the modified form (SI Table 1; Extended Data Figure 2), which incorporated extra information contained in the colony counts of workers and queens the following spring and introduced the Poisson distribution to the model. All parameters were estimated with greater precision with data for all species combined, and with parameters common across all species (row 1, Extended Data Table 3), and the model captures the general variation in counts observed for the two castes (Extended Data Figure 3). Sampling correlations between the model parameters further demonstrate that survival and detection probability can be estimated separately (SI Table 2). Using the maximised log-likelihood values for these models (-log(L), Extended Data Table 3) suggested that there were between-species differences in the parameters (likelihood-ratio test statistic of 2 × (2045.6 – 467.6 – 538.6 – 992.6) = 93.6 , which is significant on the basis of a χ^2 distribution with 7 degrees of freedom ($P < 0.01$)). However, examination of the confidence limits suggested that the mean counts of workers (λ_1) and queens (λ_2) caught were estimated more precisely than the survival probabilities (φ_1 and φ_2). Thus data from all three species were combined for further analyses, with different mean counts but shared survival rates, to remove overparameterisation and increase precision.

Finally, relationships between habitat/land-use variables and survival (φ_2) were explored for families with $W_{1i}>1$ for which colony locations could be estimated (n = 456). To further test for any species-specific differences in survival among this reduced sample, we established a baseline model (Model A) in which φ_2 was constant but, motivated by the above likelihood-ratio test statistic and confidence limits, λ_1 and λ_2 were permitted to vary between species. A formal comparison of Model A with a model including species-specific differences in survival (Model B) confirmed that these were not significant (χ_2^2 = 3.95, P = 0.14). Therefore families of all three species (n = 456) were pooled for extensions of Model A with φ_2 regressed against each of the habitat or land-use variables at different radii from the colony locations in turn. Those models leading to a significant improvement in fit when compared to Model A were considered to demonstrate significant effects of habitat or land-use variables on the probability of family lineage survival (Fig. 2; Extended Data Table 4). These regressions were repeated using a complementary log-log link function which did not affect the model outcomes, with the sign of the slopes remaining the same and the maximum log-likelihood values remaining similar to those presented in Extended Data Table 4.

Data Availability: Datasets are available from the NERC Environmental Information Data Centre (EIDC) as follows: Family lineage and landscape quality data for wild bumblebee colonies across an agricultural landscape in Buckinghamshire, U.K. (http://doi.org/10.5285/6be00174-6544-4156-b1df-8678f6df2034); Map of land-use/land-cover and floral cover across an arable landscape in Buckinghamshire, UK (http://doi.org/10.5285/0667cf06-f2c3-45c1-a80a-e48539b52427); Microsatellite genotype data for five species of bumblebee across an agricultural landscape in Buckinghamshire, UK. (http://doi.org/10.5285/6a408415-0575-49c6-af69-b568e343266d); Location data of worker bumblebees across an agricultural landscape in Buckinghamshire, UK. (http://doi.org/10.5285/a60f52b8-0f9f-44f6-aca4-861cb461a0eb). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to ccar@ceh.ac.uk.

Code availability. The full R code for simulating a dataset and fitting the modified CJS model is available in Supplementary Information: Carvell_Nature_Supplementary Information_Msimulation.R.

References

31 Redhead, J. W. *et al.* Map of land-use/land-cover and floral cover across an arable landscape in Buckinghamshire, UK. *NERC Environmental Information Data Centre* doi:http://doi.org/10.5285/0667cf06-f2c3-45c1-a80a-e48539b52427 (2014).

32 Lye, G., Park, K., Osborne, J., Holland, J. & Goulson, D. Assessing the value of Rural Stewardship schemes for providing foraging resources and nesting habitat for bumblebee queens (Hymenoptera: Apidae). *Biological Conservation* **142**, 2023-2032 (2009).

33 Lopez-Vaamonde, C. *et al.* Lifetime reproductive success and longevity of queens in an annual social insect. *Journal of Evolutionary Biology* **22**, 983-996 (2009).

34 Goulson, D., Hughes, W. O. H., Derwent, L. C. & Stout, J. C. Colony growth of the bumblebee, Bombus terrestris, in improved and conventional agricultural and suburban habitats. *Oecologia* **130**, 267-273 (2002).

35 Whitehorn, P. R., O'Connor, S., Wackers, F. L. & Goulson, D. Neonicotinoid Pesticide Reduces Bumble Bee Colony Growth and Queen Production. *Science* **336**, 351-352, doi:10.1126/science.1215025 (2012).

36 Carvell, C. *et al.* Molecular and spatial analyses reveal links between colony-specific foraging distance and landscape-level resource availability in two bumblebee species. *Oikos* **121**, 734-742, doi:10.1111/j.1600-0706.2011.19832.x (2012).

37 Holehouse, K. A., Hammond, R. L. & Bourke, A. F. G. Non-lethal sampling of DNA from bumble bees for conservation genetics. **50**, 277-285 (2003).

38 Dreier, S. *et al.* Microsatellite genotype data for five species of bumblebee across an agricultural landscape in Buckinghamshire, UK. *NERC Environmental Information Data Centre* doi:http://doi.org/10.5285/6a408415-0575-49c6-af69-b568e343266d (2014).

39 Wang, J. Sibship reconstruction from genetic data with typing errors. *Genetics* **166**, 1963-1979, doi:10.1534/genetics.166.4.1963 (2004).

40 Bourke, A. F. G. Sex ratios in bumble bees. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **352**, 1921-1932 (1997).

41 White, G. C. & Burnham, K. P. Program MARK: survival estimation from populations of marked animals. *Bird Study* **46**, 120-139 (1999).

Extended Data Table 1. Numbers (and percentages) of bumblebee (*Bombus* spp.) colonies and lineages detected within each family relationship category. * includes from sampled queens and workers. *B. lapidarius* queens were not sampled in 2011.

Extended Data Table 2. **Initial model results showing estimated survival and detection parameters for the three study bumblebee (*Bombus*) species.** Column headings are: estimated colony survival rate φ_1 (mother queen to daughter worker stage); recapture probability p_1 (worker stage); and combined survival and recapture probability θ (summer worker to spring queen stage). Estimated via a Cormack-Jolly-Seber type mark-recapture model fitted in MARK[41] except *where estimated as a simple binomial GLM from the numbers of colonies identified at the worker stage that produced records of queens the following year. Estimated asymptotic 95% confidence limits (based on the Hessian matrix and back-transformed) given in parentheses.

Extended Data Table 3. **Estimated probabilities of survival and detection of bumblebee (*Bombus* spp.) family lineages using the modified mark-recapture model** (see Methods and Supplementary Information). Column headings are: apparent survival rates φ_1, $\varphi_{2;}$ expected numbers of workers (λ_1) and spring queens (λ_2) caught per surviving family lineage; probabilities of at least one individual worker (p_1) or queen (p_2) per surviving family lineage being captured in the study landscape (recapture probabilities are calculated as (1 - the probability of no individuals being caught), based upon the estimates of λ); combined probabilities of survival and recapture, calculated as $\varphi_1 p_1$ and $\varphi_2 p_2$, respectively; and -log(L) = maximised log-likelihood values for each model. Estimated asymptotic 95% confidence limits (based on the Hessian matrix and back-transformed) given in parentheses. Parameter values were estimated from all sampled families (*n* = 1,665), including those represented by single workers.

Extended Data Table 4. Model results for logistic regression of apparent survival (probability of bumblebee family lineage survival from the summer colony to spring queen stage (φ_2)) against habitat quality and land-use variables (Extended Data Table 5) at four spatial scales. *n* = 456 family lineages, excluding those represented by only queens or by single workers for which colony locations could not be estimated. Significant variables and *P* values are shown in bold. Foraging distance is the mean of the straight-line distances of each worker from its capture location to its estimated colony location.

Extended Data Table 5. **Habitat quality and land-use variables for which effects on bumblebee family lineage survival and queen dispersal distance were tested.** All variables were calculated as proportions of cover represented by the given category out of the total land area within a given radius (see Methods).

Extended Data Table 6. Plant groups used for field survey of habitats across the study landscape.

*Y/N denotes plant groups visited/not visited by foraging queens or workers during bee sampling.

#Y/N denotes preferred/not preferred forage plant groups of workers during sampling, preferred plants groups being identified as the five plant groups with the highest mean number of worker visits (across all three *Bombus* species) per plant species within that group.

Extended Data Figure 1. Variation and correlations between habitat and land-use variables across the study landscape. Plots show i) histograms to demonstrate variation within each habitat/land-use variable along the diagonal; ii) scatter plots showing correlations between variables (top right) with a linear model trend line fitted to the correlation data (in red, only extended to the limits of the data) and a 1:1 line (in pale grey) and iii) correlation coefficients with their significance (bottom left) where $P < 0.05$ = *, $P < 0.01$ = ** and $P < 0.001$ = ***. Axis values are standardised and represent proportional cover of the different habitat variables within 1,000 m of estimated colony locations, with variable names following the same order and shortened format as presented in Extended Data Table 5. Each point on the scatter plots represents one family lineage ($n = 456$).

Extended Data Figure 2. Simulation-based assessment of robustness of the modified CJS model. This shows that the estimated parameter values aggregate around the true values. Frequency distributions of parameter estimates are shown, from 1,000 simulated data sets, each of 2,000 families. Parameters plotted are (a) φ_1: true value = 0.6 (b) φ_2: true value = 0.5 (c) λ_1: true value = 3 (d) λ_2: true value = 2. To align with the real data in which some families were not detected at the founding queen (Q1) stage, if at all, data were simulated assuming a detection probability of 0.4 at the Q1 stage.

Extended Data Figure 3. **Goodness of fit for the model of Extended Data Table 3.** Frequency distributions across all species of a) observed counts of workers (W_{1i}); b) expected counts of workers (W_{1i}); c) observed counts of second-generation queens (Q_{2i}) and d) expected counts of second-generation queens (Q_{2i}).