University of Plymouth

PEARL **https://pearl.plymouth.ac.uk**

Faculty of Science and Engineering School of Biological and Marine Sciences

2023-01-28

The tragedy of the common? A comparative population genomic study of two bumblebee species

Huml, JV

http://hdl.handle.net/10026.1/20223

10.1111/icad.12626
Insect Conservation and Diversity
Wiley

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Received: 21 September 2022 | Accepted: 4 January 2023

DOI: 10.1111/icad.12626

ORIGINAL ARTICLE

Insect Conservation and Diversity

The tragedy of the common? A comparative population genomic study of two bumblebee species

Jana V. Huml[1] | Jonathan S. Ellis[1] | Sarah Rustage[1] | Mark J. F. Brown[2] | Richard Billington[1] | Mairi E. Knight[1]

[1]School of Biological and Marine Sciences, University of Plymouth, Plymouth, UK

[2]School of Biological Sciences, Royal Holloway, University of London, Surrey, UK

Correspondence
Jana V. Huml, School of Biological and Marine Sciences, University of Plymouth, Plymouth PL4 8AA, UK.
Email: vanessa.huml@plymouth.ac.uk

Funding information
Natural Environment Research Council; Seale Hayne Educational Trust; Natural England; Natural Resources Wales; Scottish Natural Heritage

Editor: Nusha Keyghobadi and Associate Editor: Sandra Rehan

Abstract

1. Within the theoretical framework of the small population paradigm, we investigated the population genomics and parasite load of two bumblebee species across the UK and Ireland. *Bombus pratorum* is widespread and common throughout its range while *Bombus monticola* is restricted to higher altitudes and shows a more fragmented distribution.
2. *Bombus monticola* showed stronger population structuring, isolation-by-distance, and a deficit of heterozygotes in the most isolated population in the south of its range (Dartmoor). Heterozygosity and inbreeding coefficients (F_{IS}) were comparable between both species, but the proportion of polymorphic sites was much greater in *B. pratorum*. Notably, both species have suffered significant declines in Ne over the last 100 generations and estimates and declines for both species were of similar orders of magnitude. No pattern of increased parasite prevalence in populations of lower heterozygosity was observed. Instead, ecological and demographic factors (age, latitude, date, habitat suitability) were the main drivers of parasite prevalence.
3. Distinct patterns of selection were observed in both species in regions involved in regulation of transcription and neurotransmission and in particular pathways targeted by neonicotinoid insecticides.
4. Our results highlight the pressing need for monitoring to include common as well as rare species. This should not focus solely on census population counts, but include estimates of Ne. We also highlight the need for further work to establish adaptive shifts in globally important pollinator communities.

KEYWORDS
biodiversity loss, *Bombus*, effective population size, parasite load, pollinators, RAD-seq, small population paradigm

INTRODUCTION

Conservation efforts and priorities are generally focused on species that exist in small and/or declining populations since they are the most vulnerable to extinction. This is underpinned by the 'small population paradigm' (SPP, Gilpin & Soulé, 1986; Caughley, 1994), which predicts that small populations are vulnerable to the extinction vortex. The vortex is created when populations decline to a small effective

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *Insect Conservation and Diversity* published by John Wiley & Sons Ltd on behalf of Royal Entomological Society.

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

size (Ne), leading to high drift load (Willi et al., 2013) and decreasing adaptive potential. The consequent impacts on individual fitness become increasingly acute through each generation and without intervention, the interaction of demographic and genetic stochasticity can lead to extinction.

Conservation genetics has since been dominated by efforts to measure standing genetic diversity, drift, inbreeding, gene flow and effective population size (Ne) (Ouborg et al., 2010). These parameters are quantified by screening neutral genetic markers since they are free from the complex potential influences of selection and are expected to behave in an evolutionarily predictable way in populations through time. Typically, such studies are restricted to single species and often in very limited geographic locations (i.e., those populations/species of concern that are rare and localised).

While this framework has led to a wealth of insightful research, criticisms have followed. First, the single rare species approach often lacks taxonomically relevant and directly comparable datasets for context, hampering the ability to correctly interpret and identify key conservation threats in wild populations. Taxonomic groups vary markedly in levels of standing genetic diversity (Ellegren & Galtier, 2016) so population genetic metrics are best interpreted in the context of comparative data from related species with similar life histories (Leffler et al., 2012). Such comparative studies are rare, however, and while such taxonomic context may be available through other independent studies, the use of different marker sets can often make direct comparisons difficult. A single focus on rarity can also obfuscate trends in species that have historically been (and may appear to remain) relatively common. From some conservation perspectives, such species are (far more significant than those that are rare or localised since the impact of their loss or decline may be more ecologically severe (Gaston & Fuller, 2008)). Despite the fate of the passenger pigeon, among many (see e.g., Roycroft et al., 2021), and regular calls for more widespread monitoring to avoid catastrophic and rapid extirpations of more common organisms (e.g., Gaston & Fuller, 2008), the conservation lens remains strongly focused on rare species or populations.

Secondly, there are a number of scenarios where populations may not behave in accordance with the theoretical predictions of the SPP. Habel and Schmitt (2012), for example, argue that highly localised 'specialist' species occurring in relatively low abundances may carry much lower genetic load due to purging or other strong selection. There is certainly good evidence that purging through selection can remove genetic load (e.g., Gloag et al., 2016; Robinson et al., 2018; Khan et al., 2021; Xie et al., 2022, although see also Wang et al., 2021) but how effective purging will be will depend strongly on how rapid and severe declines in population sizes are (i.e., historical demography): selection driven purging processes can be over-powered by the strength of drift if that reduction is dramatic and fast (see e.g., López-Cortegano et al., 2016; Pérez-Pereira et al., 2022).

Thirdly, whether genetic factors present a real extinction threat to small in situ populations has been the subject of debate since the inception of the theoretical framework. Several authors have suggested that demographic and environmental factors are more likely to drive wild populations to extinction before these genetic effects become so extreme that they directly cause extirpation (see e.g., Caughley, 1994; Lande, 1988). There is now compelling evidence from a wide variety of studies that the increased mutational (genetic) load created by drift and inbreeding in small populations can be linked directly to extinction risk in wild populations (Nonaka et al., 2019; O'Grady et al., 2008; Saccheri et al., 1998; Willi et al., 2013). Of course, these studies do not establish a general rule, but they do provide examples of where genetic factors are likely to have been important.

While there has been an emphasis on studies assessing the effects of genetic load on fitness (inbreeding depression), the parallel issue of maintaining genetic diversity to enable adaptive potential has been less controversial. Although adaptive variation is more difficult to study given our limited toolbox for identifying, functional genetic variation, clear links have been established between low genetic diversity and extinction risk (e.g., Evans & Sheldon, 2008; Spielman, Brook, Briscoe, & Frankham, 2004; Spielman, Brook, & Frankham, 2004; see also Lanfear et al., 2010). Thus, while the ability to assess adaptive genetic diversity, particularly in non-model systems, is still in its infancy, the pressing calls (e.g., Ralls et al., 2018; Razgour et al., 2019) for maintaining genetic diversity are uncontentious, particularly in the face of current climate change predictions (Bush et al., 2016).

A final aspect of the SPP that has not been without controversy is which component of genetic diversity is most appropriate to assess. Recently, Teixeira and Huber (2021) asserted that variation at neutral genetic markers is at best only loosely linked to significant population declines and instead advocate a focus on functional genomic regions. DeWoody et al. (2021) and Kardos et al. (2021), however, present compelling evidence to the contrary and, as noted above, from a practical perspective we currently lack the sophisticated level of understanding of functional genetic diversity to enable an exclusive functional approach. There is also good evidence however, that, in some circumstances, differences in functional markers can arise between populations without notable patterns in neutral diversity being observed (see e.g., Allendorf et al., 2010; Krohn et al., 2019). Fortunately, much more comprehensive genome-wide screening has recently become accessible and affordable for population-level studies. This allows both neutral and functional markers to be screened simultaneously, and at much greater scale, significantly enhancing the resolving power of neutral markers while also giving important initial insights into variation at functional markers in non-model organisms (see e.g., Hoelzel et al., 2019).

Within this framework, we took a comparative, conservation genomic approach to assess paired samples from populations of two congeneric species of bumblebee (*Bombus* spp.) across their British and Irish ranges, relating aspects of both neutral and functional genomic diversity to measures of fitness (parasite load) and habitat suitability. While data remain surprisingly limited for many species, there is strong evidence for global declines in bumblebee species (e.g., Cameron & Sadd, 2020; Goulson et al., 2005; Vray et al., 2019) along with indications that specialised species may be most vulnerable (e.g., Casey et al., 2015), and that declines may be phylogenetically structured (Arbetman et al., 2017; Cameron et al., 2007, 2011). Furthermore, as eusocial, haplodiploid organisms, bumblebees are

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

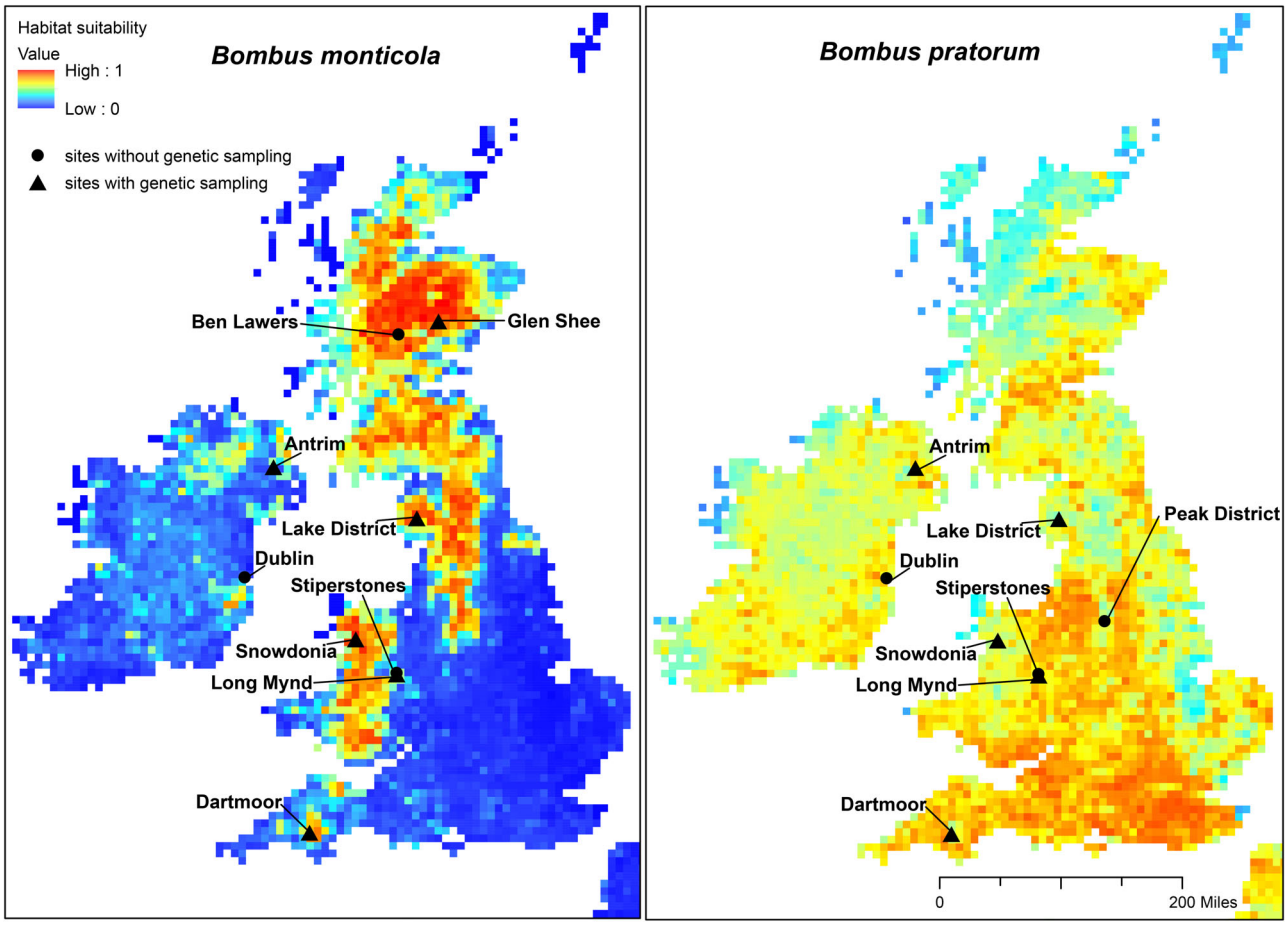

FIGURE 1 Sampling locations for *Bombus pratorum* (right) and *Bombus monticola* (left) with and without genetic information; species specific habitat suitability estimates from Polce et al. (2018) are also shown

especially prone to extinction (Zayed & Packer, 2005) and census-based abundance estimates can be particularly misrepresentative of genetic effective population sizes (Ne). As key ecosystem service providers pollinating both crops and wildflowers, this is a group of particular conservation concern. Bumblebees also provide ideal models for testing predictions of the small population paradigm since species are variably threatened and isolated, ranging from the very rare and highly localised to those that are widespread and common.

The two congeners studied belong to the subgenus *Pyrobombus*, a group showing resilience to overall patterns of decline (Arbetman et al., 2017). *Bombus monticola* is highly localised in its distribution and included on English Nature's Programme for Recovery due to recent significant declines (BWARS, 2022; BCT, 2022). In contrast, *Bombus pratorum* is one of the most ubiquitous species found across the UK (BWARS, 2022; BCT, 2022). Several species in this subgenus have recently expanded their range (Biella et al., 2021; Huml et al., 2021; Rasmont et al., 2015; Richardson et al., 2019) including *B. monticola* and *B. pratorum*, which were first recorded in Ireland in 1974 and 1947, respectively (Fitzpatrick et al., 2007; Speight & MCD, 1974). Despite these range expansions, *B. monticola* has historically been restricted to areas of higher altitude across Europe, while *B. pratorum* represents one of the most widespread and abundant species found across a range of habitats in Europe (Rasmont et al., 2015). Neither has been the subject of landscape-scale population genetic studies to date.

Specifically, our aims were:

1. To assess and interpret genetic diversity (heterozygosity and % polymorphic sites), drift, inbreeding, gene flow and effective population size (Ne);
2. To comparatively assess any associations between population genetic parameters and environmental (habitat) and biological (parasite) factors;
3. To identify whether particular and/or consistent functional regions of the genome are under strong selection;
4. To interpret the results in the context of the factors potentially driving declines and key conservation priorities.

MATERIALS AND METHODS

Sample collection

Bombus monticola and *B. pratorum* workers (females) were sampled from nine and eight locations, respectively (Figure 1) between June

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2010 and September 2011 (*B. monticola*) and between June 2010 and October 2011 (*B. pratorum*) (at some sites the species did not co-occur). Precautions were taken to minimise the probability of collecting sisters within sampling sites (Goulson et al., 2011), although putative sisters were also filtered from the subsequent analysis (see below). Individuals were frozen until processing using a portable freezer.

Measurements of wing wear and parasite load

Wing wear

As immunocompetence can decline with age in *Bombus* (Doums & Schmid-Hempel, 2000), the relative age of individual bees was estimated by assessing the level of wing wear, following a modified version of Mueller and Wolf-Mueller's (1993) method where 0 was the lowest and 4+ the highest level of wing wear (as in Goulson et al., 2012). Wing wear is frequently used as a correlate of age in bumblebees in the absence of other reliable individual age estimators (e.g., Colla et al., 2006; Allen et al., 2007; Whitehorn et al., 2011) however, it is important to note that other factors, such as individual foraging behaviour or predation pressure also affect wing wear (Foster & Cartar, 2011).

Measuring parasite load

The tracheal tubes and tissues of the abdomens of all samples were inspected for the presence or absence of the tracheal mite *Locustacharus buchneri*.

All samples were also screened for the presence or absence of the microparasites *Apicystis bombi*, *Crithidia bombi* and *Vairimorpha bombi* using species-specific molecular markers (SSUrRNA, Klee et al., 2006; Neo, Meeus et al., 2010; ITS1, Schmid-Hempel & Tognazzo, 2010). The contents of the abdomen were added to 500 µl of 10% Chelex® (10% volume Chelex® resin, 50 mM TRIS) and homogenised using a sterile micropestle. These were heated at 56 °C for 30 min, then, boiled for 8 min at 98 °C, before being centrifuged at 13,000*g* for 5 min. The supernatant was removed and 3 µl was used for subsequent PCR reactions. The PCR conditions were: 3 µl template, 1.25 µM each primer (0.625 µM for *C. bombi* primers), 2X Biomix (Bioline), made up to 10 µl using molecular grade water. The PCR cycle was: 98 °C for 2 min; 50 cycles of 98 °C for 30 s, 45 s annealing at 48 °C for SSUrRNA, 57 °C for Neo and 55 °C for ITS1, 72 °C for 1 min; 72 °C for 5 min. After PCR, 5 µl of the reaction was run for 1 h at 110 V on a 1% agarose gel stained with 5 µl of SYBRSafe (Life Technologies).

RAD library preparation

DNA was extracted from individuals from a sub-sample of six and five locations for *B. monticola* and *B. pratorum*, respectively (Figure 1) from thorax and leg tissue using a CTAB ammonium acetate protocol (Nicholls et al., 2000). Not all originally sampled locations could be genetically tested due to issues with sample quality. Extractions were quantified on a Qubit 3.0 Fluorometer (Thermo Fisher Scientific) after RNAse treatment. DNA extracted from individual bees was digested using either 100–200 ng of DNA and 5 Units of MluCI (New England Biolabs), 10 Units of MspI (New England Biolabs) in a 10 µl volume, or 200–400 ng of DNA, 10 Units of MluCI (New England Biolabs), 20 Units of MspI (New England Biolabs) in a 20 µl volume at 37 °C for 3 h. Library preparation followed Peterson et al. (2012) in randomised batches of 12 samples including 15 technical replicates for each species ($N = 324$, individually tagged samples in the library). Paired-end sequencing (150 bp) was carried out on a NovaSeq6000 by BGI-Hong Kong Co. Ltd.

Data processing and variant calling

Raw reads were de-multiplexed, trimmed and quality filtered using the process_radtags.pl script in Stacks 2.48 (Rochette et al., 2019). Reads were filtered and removed using default values in Stacks 2.48 (Rochette et al., 2019; i.e., those without intact restriction sites or an average phred score below 10 in any 22 bp window). All reads were trimmed to a uniform length of 140 bp to remove low-quality base calls at the end of the reads as standard (O'Leary et al., 2018). Individuals with low sequencing coverage and a representation below 1 M paired reads were excluded from the analysis ($N = 54$).

Filtered reads were assembled de novo using the denovo.pl script in Stacks 2.48 (Rochette et al., 2019). Assembly parameters were optimised following the recommendations of Paris et al. (2017) using a subset of samples including all replicates and a minimum of two samples per population (*B. monticola* $N = 33$, *B. pratorum* $N = 29$). For each parameter set, SNPs were called using the populations.pl script in Stacks 2.48 (Rochette et al., 2019) if they were present in at least 80% of individuals and showed a minor allele frequency above 1%. Consistency of variant calls (SNP, allele and locus error rates) across replicates was assessed using the R-scripts provided by Mastretta-Yanes et al. (2015) to identify the best performing parameter settings. After establishing the optimal parameter values that minimised error rates across replicate samples, these were used to build the final catalogue using eight samples from each population that showed the highest coverage, following general recommendations (Rochette et al., 2019). As additional filtering was carried out on the original Stacks output (see below), the calculation of allele and locus error rates only considered variable contigs in contrast to the original implementation by Mastretta-Yanes et al. (2015). If anything, this is expected to result in an overestimation of error rather than an underestimation. Next, de novo contigs were mapped against the *Bombus terrestris* genome (Bter_1.0 assembly, Ensembl) using the integrate alignment tool in Stacks (Rochette et al., 2019) in order to take advantage of the linkage group and functional annotations this assembly offers. Although by aligning to the genome of a different species some information is inevitably lost, for example, where annotations are

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

incomplete or where alignment fails in areas of high divergence, this is a necessary approach when working with non-model species (Shafer et al., 2017). Alignments of individual sample files against the catalogue, which contains all contigs resulting from the de novo assembly, were used to build an mpileup file using samtools (Li et al., 2009). The mpileup file was exported to vcf format including both variant and non-variant sites using bcftools (Li et al., 2009) and filtered in the same way as SNPs described below.

The generated vcf file was further filtered using vcftools (Danecek et al., 2011) (Table S1) following the recommendations of O'Leary et al. (2018; SNPs with a minimum allele count below 3 and a minimum mean depth of coverage across all individuals below 15 were filtered). A minimum genotype depth of five reads was required to call a SNP in any individual to avoid a bias towards homozygote calls in areas of low sequencing coverage (O'Leary et al., 2018). The percentage of missing data was minimised through iteratively removing SNPs and individuals to achieve a maximum of 5% of missing data for any SNP and a maximum of 25% of missing data in any individual (O'Leary et al., 2018). SNPs that showed a mean coverage across all individuals higher than double the mode average were also filtered, as these are likely to be derived from paralogous loci (Willis et al., 2017). Finally, loci significantly deviating from Hardy–Weinberg equilibrium after correction for multiple testing were also filtered, as these are likely to be affected by genotyping errors (Attia et al., 2010). Corrections for multiple testing were carried out in the R package qvalue v2.16 (Storey et al., 2015) applying a false discovery rate of 0.05 (Storey & Tibshirani, 2003). Coefficients of relatedness between all individuals were calculated in vcftools (Danecek et al., 2011) and for each pair showing a coefficient of relatedness >0.25, the individual with fewer reads was removed from all population genetic analyses (since the absence of kin is an assumption and would bias parameter estimates).

Assessment of genetic diversity, population structure and effective population size

Basic measures of population genetic diversity and pairwise F_{ST} (Weir & Cockerham, 1984) were calculated using the R package hierfstat (Goudet, 2005). To generate confidence intervals for pairwise F_{ST} values and inbreeding coefficients bootstrapping over loci was performed for 1000 cycles.

Population structure and admixture was assessed by principal coordinate analysis (PCoA), which is preferable to a principal component analysis when there are fewer individuals than characters (Rohlf, 1972). PCoA tests were performed in the R package ade4 v.1.7-13 (Dray & Dufour, 2007) based on Euclidean distance (Gower, 1966). A Bayesian approach implemented in the programme fastStructure (Raj et al., 2014) was also implemented to assess population structure and admixture. FastStructure assumes the independence of loci and therefore only one SNP was used per RAD-tag. To find the number of components (K) that best explain the structure of the data, fastStructure was run for values of K from 1 to 10. The Python script 'chooseK.py' provided within fastStructure was then used to identify the value of K that best fitted the data. The genetic clusters identified were used in further analysis for the definition of population demes. A Mantel test (Mantel, 1967) was further conducted in Genepop to test for isolation by distance based on pairwise F_{ST} (Raymond & Rousset, 1995).

To estimate long-term effective population size, Watterson's theta (θ_W) (Watterson, 1975) was calculated in SambaR (de Jong et al., 2021) as has been estimated for other *Bombus* species (Lattorff et al., 2016). This parameter is equivalent to $3N_e\mu$ in haplodiploids. For the mutation rate μ, Liu et al.'s (2017) estimate of 3.6×10^{-9} (CI: 2.38×10^{-9} 5.37×10^{-9}) for a single nucleotide per generation in *Bombus terrestris* was used. To assess trends in the recent trajectories of the effective population size, linkage disequilibrium (LD) across different inter-locus distances can be exploited (Waples & Do, 2010). LD patterns over larger distances are informative for recent N_e while LD patterns over shorter distances are informative for the more distant past (see also Nadachowska-Brzyska et al., 2022, for an informative overview). This analysis was carried out using SNeP v1.1 (Barbato et al., 2015), which is designed to estimate trends in effective population size using LD. First, Beagle 5.1 (Browning & Browning, 2007) was used to phase genotypes and then vcftools (Danecek et al., 2011) was used to calculate the correlation coefficient (r^2) between any two SNPs as a measure of LD used in SNeP for each population. SNeP was run using default parameters with an adjusted recombination rate of $8.7\ e^{-008}$ in accordance with the recombination rate (8.7 cM/M) observed in bumblebees (Liu et al., 2017). As methods based on LD can greatly underestimate absolute values of N_e in haplodiploid species (Wang et al., 2016) we emphasise that the focus of this analysis is on the *relative* change in N_e over time. Results are presented for generations 10–100, given the uncertainties around estimates in the most recent and most distant generations (Barbato et al., 2015; Corbin et al., 2012).

Investigation of predictors of parasite richness

Relationships between observed parasite richness (number of parasite species detected), wing wear, longitude and latitude as well as species-specific habitat suitability of the sampling site (estimated using Polce et al. (2018)) and date (represented as numerical value corresponding to the day and month of sampling) were investigated using a general linear modelling approach in the stats R package framework (Chambers & Hastie, 1992). Habitat suitability estimates were derived using predictor variables such as land cover, distance from natural or semi-natural habitat and climate variables (Polce et al., 2018). Values representing the probability of presence were then extracted for each sampling site. An additional GLM was performed with observed heterozygosity as an additional potential predictor including all locations with genetic data available. In both cases, a Poisson distribution was implemented. Furthermore, relationships were investigated independently for each parasite species based on presence and absence data using a binomial distribution. Akaike's information criterion (AIC) was used to compare models with different

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Measurements of population genetic diversity per site and species: number and proportion of polymorphic sites, number of private alleles, observed (H_o) and expected heterozygosity (H_e), inbreeding coefficients (F_{IS}) and nucleotide diversity (pi); estimates of long-term effective population size are given per population-cluster

Sample site	Species	*N* Individuals	*N* Polymorphic sites	Proportion of polymorphic sites	Private alleles	H_o	H_e	F_{IS} (CI)	Pi	Ne (CI)
Antrim	*Bombus monticola*	16	9843	0.0015	370	0.176	0.174	−0.0084 (−0.0164, −0.0055)	0.175	23,319 (15,633– 35,272)
Dartmoor	*B. monticola*	12	7147	0.0011	67	0.129	0.137	**0.0470 (0.0462, 0.0609)**	0.137	13,444 (9013–20,336)
Glenshee	*B. monticola*	15	15,485	0.0024	2195	0.246	0.242	−0.0103 (−0.0425, −0.0334)	0.243	50,700 (33,989–76,689)
Lake District	*B. monticola*	23	12,577	0.0019	338	0.178	0.175	−0.0127 (−0.0246, −0.0159)	0.176	30,296 (20,310–45,826)
Long Mynd	*B. monticola*	12	9256	0.0014	47	0.168	0.167	−0.0061 (−0.0102, 0.0026)	0.168	23,107 (15,491–34,952)
Snowdonia	*B. monticola*	19	9547	0.0014	75	0.163	0.160	−0.0107 (−0.0200, −0.0094)	0.161	
Antrim	*Bombus pratorum*	26	22,731	0.0037	555	0.133	0.134	**0.0051 (0.0018, 0.0081)**	0.134	43,981 (29,485–66,527)
Dartmoor	*B. pratorum*	19	39,049	0.0063	265	0.169	0.169	0.0003 (−0.0026, 0.0032)	0.170	113,907 (76,363–172,297)
Lake District	*B. pratorum*	9	31,456	0.0051	147	0.171	0.171	−0.0032 (−0.0029, 0.0047)	0.171	
Long Mynd	*B. pratorum*	22	40,481	0.0065	279	0.170	0.170	−0.0006 (−0.0061, −0.0009)	0.170	
Snowdonia	*B. pratorum*	25	41,065	0.0066	393	0.169	0.170	**0.0062 (0.0054, 0.0104)**	0.170	

Note: Values of F_{IS} significantly larger than zero are highlighted in bold.

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Population pairwise F_{ST} on the upper and confidence limits on the lower triangular of the matrix for *Bombus monticola* (top) and *Bombus pratorum* (bottom)

B. monticola	Antrim	Dartmoor	Glenshee	Lake District	Long Mynd	Snowdonia
Antrim	NA	0.265	0.16	0.159	0.184	0.19
Dartmoor	0.259, 0.271	NA	0.229	0.197	0.117	0.137
Glenshee	0.156, 0.164	0.224, 0.234	NA	0.151	0.165	0.186
Lake District	0.155, 0.163	0.192, 0.203	0.147, 0.155	NA	0.103	0.103
Long Mynd	0.179, 0.189	0.112, 0.121	0.161, 0.169	0.099, 0.106	NA	0.024
Snowdonia	0.185, 0.195	0.132, 0.141	0.182, 0.190	0.100, 0.106	0.022, 0.026	NA

B. pratorum	Antrim	Dartmoor	Lake District	Long Mynd	Snowdonia
Antrim	NA	0.119	0.123	0.116	0.116
Dartmoor	0.117, 0.121	NA	0.007	0.002	0.004
Lake District	0.120, 0.125	0.006, 0.008	NA	0.006	0.006
Long Mynd	0.114, 0.118	0.001, 0.002	0.005, 0.007	NA	0.002
Snowdonia	0.114, 0.118	0.003, 0.004	0.005, 0.007	0.002, 0.003	NA

numbers of explanatory variables (Table 2) using the step AIC function in the R-package gamlss, which implements a stepwise reduction in variables based on AIC (Rigby & Stasinopoulos, 2005).

Detection of signatures of selection

F_{ST}-outlier approach

In accordance with recommendations to combine several methods for the identification of F_{ST}-outliers to minimise the false positive detection rate (De Mita et al., 2013), here three different approaches were adopted. BayeScan (Foll & Gaggiotti, 2008), which implements a basic regression model to differentiate between locus and population specific effects on the distribution of F_{ST}–values, was run using default parameters. Secondly, FstHet (Flanagan & Jones, 2017), which identifies outliers based on F_{ST}–heterozygosity relationships, was run using the betahat option for the calculation of F_{ST} (Cockerham & Weir, 1993). This method implements a correction for sample size. Finally, pcadapt (Luu et al., 2017), which identifies outliers as loci that deviate from patterns of population structure defined by principal component analysis, was used. Overlap between the different methods was visualised using Venn diagrams (Oliveros, 2007). Outliers that were identified by at least two approaches were investigated using gene ontology (GO) analysis.

Selective sweep detection

RAiSD (Alachiotis & Pavlidis, 2018) was used to detect signatures of selective sweeps. This programme has the advantage of combining the three most common approaches used in the detection of selective sweeps, which are based on changes in the level of polymorphism, the site frequency spectrum (SFS) and LD, into a single statistic (μ). Data were generated across sliding windows of 10 SNPs to balance information content and resolution for the detection of signals of selective sweeps (Vatsiou et al., 2016). Outliers were then identified as those above the 95th percentile of the empirical distribution. In the absence of detailed knowledge of the demography of the study populations needed to generate a simulated neutral distribution, this approach has been shown to lead to comparable results (Vatsiou et al., 2016).

Gene ontology analysis

All SNPs were annotated to specific genes if they were located within a distance of 5000 bp and therefore in high LD, which is documented to decay after 10 kb (Stolle et al., 2011) in bumblebees. GO terms for the *B. terrestris* genome were obtained from the Ensembl database (metazoa genes 55; Yates et al., 2022). Furthermore, orthologous sequences were identified in the *Apis mellifera* and *Drosophila melanogaster* genome using Biomart (Kinsella et al., 2011), where no GO annotation could be obtained from the *B. terrestris* genome. Enrichment tests were conducted using the R package topGO (Alexa & Rahnenfuhrer, 2016) using Fisher's exact test on gene count (the 'classic' algorithm, Alexa et al., 2006) and applying the 'weight' and 'elim' algorithms, which account for dependencies within the GO hierarchies (Alexa et al., 2006). A minimum node size of 10 was required to prune our hierarchy from nodes with the support of less than 10 annotated genes, which represents a frequently applied threshold (e.g., Ahrens et al., 2013; Rademacher et al., 2017). Correction for multiple testing (FDR < 5%, q-value < .05) was carried out for the 'classic' algorithm using the qvalue package in R (Storey et al., 2015), but not the 'weight' and 'elim' algorithms for which multiple testing theory does not directly apply as tests are not independent and raw p-values ≤.05 were considered as significant (Alexa & Rahnenfuhrer, 2016). The background gene set used in each analysis represented only areas actually sampled (e.g., only regions within

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

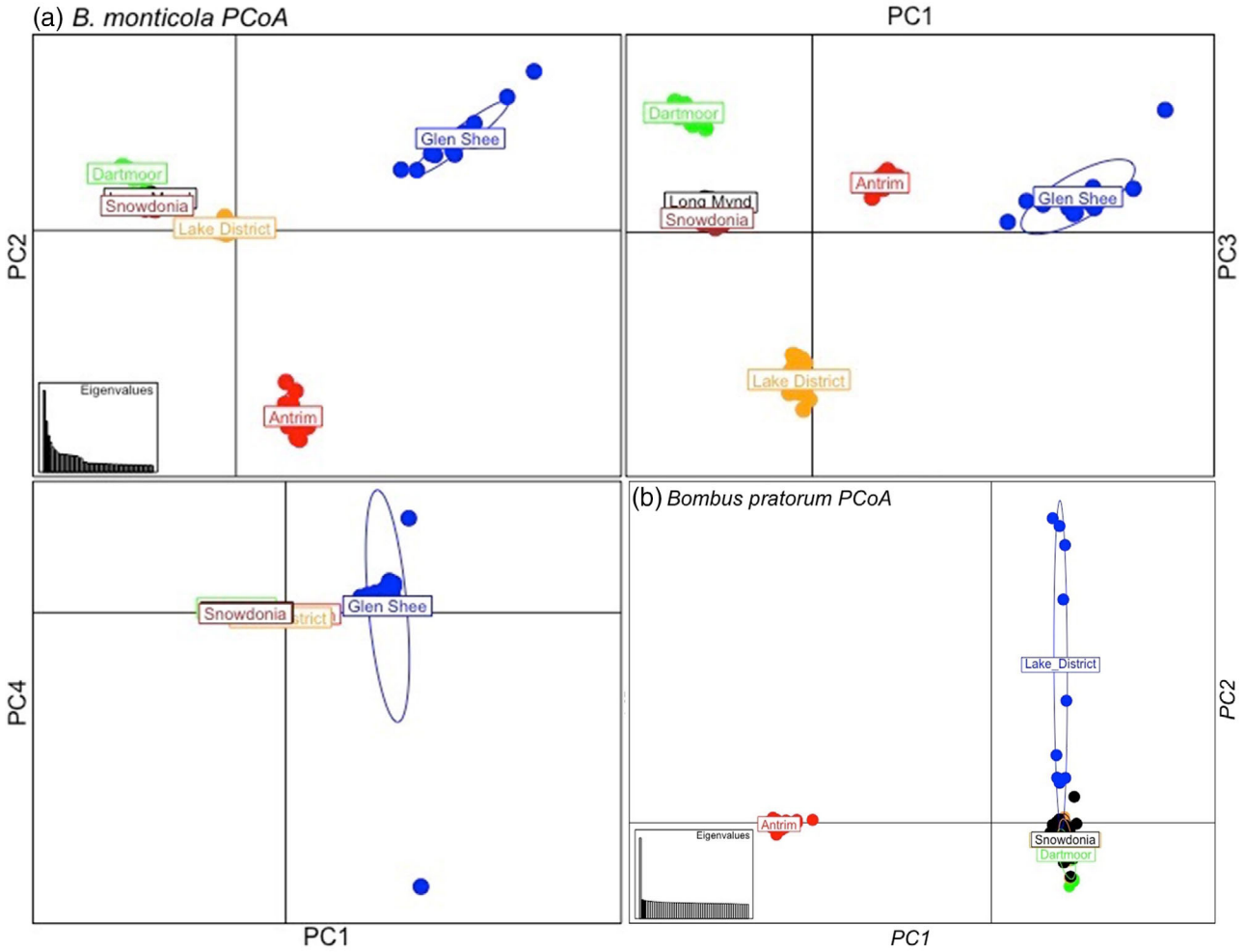

FIGURE 2 PCoA on Euclidian distance for *Bombus monticola* (a) and *Bombus pratorum* (b).

windows that qualified in the RAiSD analysis, present in the number of populations/species specified or only genes within 5000 bp of SNPs for the F_{ST}-analysis).

RESULTS

Population genetic parameters

After stringent filtering, 47,027 SNPs were identified across 21,423 contigs in *B. pratorum* (101 individuals retained) and 17,974 SNPs across 13,003 contigs in *B. monticola* (97 individuals retained). The SNP, locus and allele error rates (Mastretta-Yanes et al., 2015) were 0.006, 0.11 and 0.007 for *B. pratorum* and 0.014, 0.1 and 0.02 for *B. monticola*, respectively, these values are slightly lower than reported for the optimal filtering setting by Mastretta-Yanes et al. (2015). The total number of sites (including invariable sites) retained after filtering were 6,222,670 in *B. pratorum* and 6,585,511 in *B. monticola*, giving an average SNP density of 7.6 and 2.7 per kb, respectively. All sequenced sites were distributed across 17,236 and 17,395 linkage blocks of 10 kb (Stolle et al., 2011), giving a genome coverage of 63% (Lowry et al., 2017; McKinney et al., 2017) in both *B. pratorum* and *B. monticola*.

The number and proportion of polymorphic sites identified per population was considerably higher in *B. pratorum* than in *B. monticola*, while measurements of observed and expected heterozygosity were similar (Table 1). For *B. monticola* measurements of genetic diversity were highest in Glenshee and lowest in Dartmoor, while in *B. pratorum* Antrim showed the lowest genetic diversity compared to all other sample sites (Table 1). Inbreeding coefficients were significantly larger than zero for Dartmoor in *B. monticola* and for Antrim and Snowdonia in *B. pratorum* (Table 1).

Bombus monticola exhibited a significantly higher degree of population differentiation than *B. pratorum* (Mann–Whitney–Wilcoxon test: $p = .0008$) (Table 2). A Mantel test detected significant isolation by distance for *B. monticola* ($p = .002$), but not for *B. pratorum*. The PCoA on Euclidean distance (Figure 2a) indicates that for *B. monticola* most of the variation is explained by the first three axes with four distinct clusters separated on the first axis. This is consistent with the results of fastStructure, which identified four distinct clusters as the

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Insect Conservation and Diversity

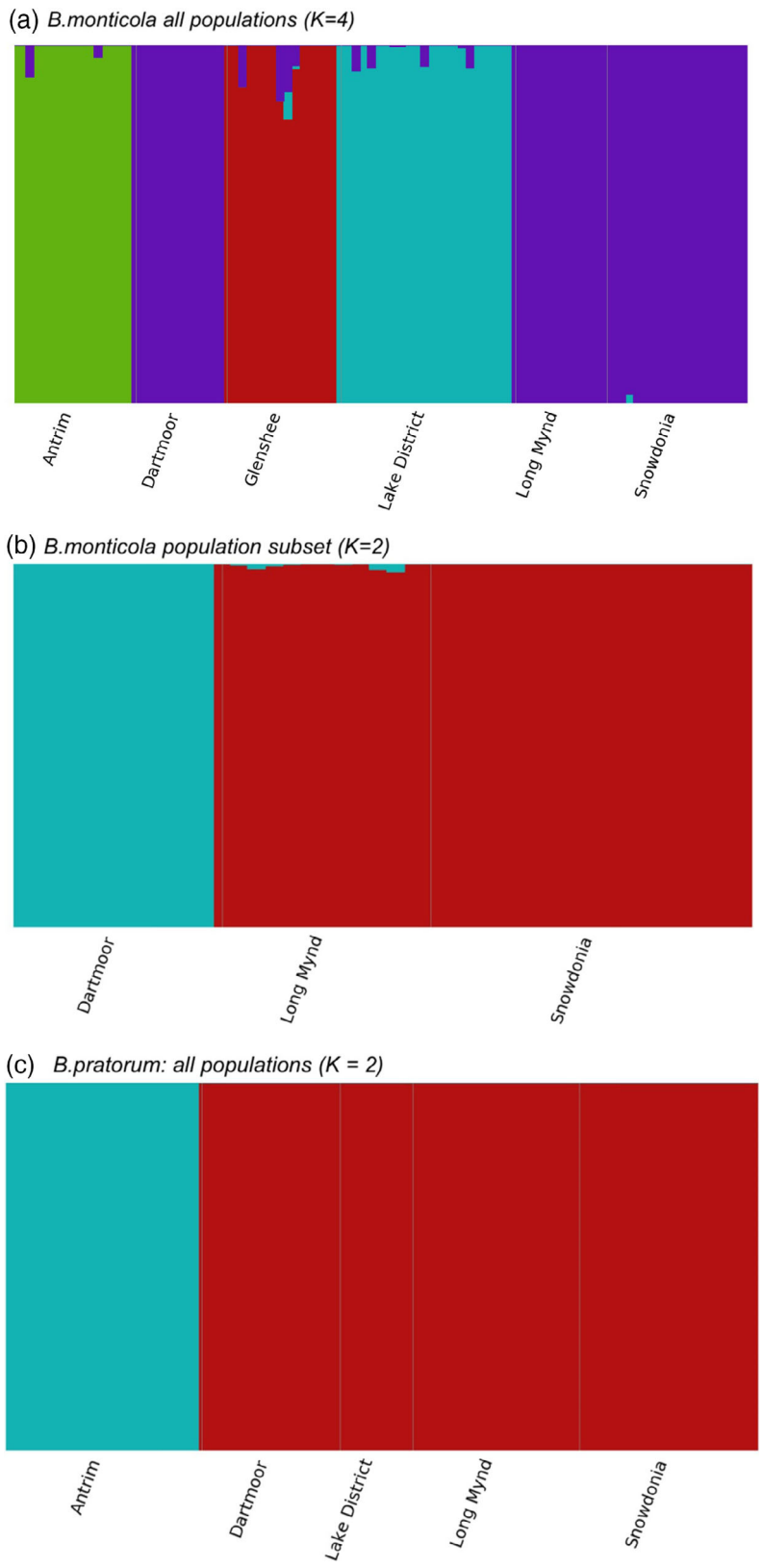

FIGURE 3 FastStructure analysis for all populations (a) and a population subset (b) for *Bombus monticola* and all populations for *Bombus pratorum* (c).

best fit to the data (Figure 3a). A further separation of Dartmoor from Long Mynd and Snowdonia becomes apparent on the second and third axes for *B. monticola* (Figure 2a). The presence of further substructure between these three populations is also confirmed by fastStructure analysis on this population subset (Figure 3b). For *B. pratorum*, the PCoA shows that variation is predominantly

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

explained along the first axis, which separates Antrim from all other populations sampled (Figure 2b), which is consistent with the fastStructure analysis identifying two clusters as the best fit to the data (Figure 3c). A separate analysis in fastStructure to assess finer population structure across all *B. pratorum* populations except Antrim found no support for any further population structure.

Long-term estimates of the genetic effective population sizes in *B. monticola* ranged from 13,444 in Dartmoor to 50,700 in Glenshee (Table 1). The long-term estimates of the genetic effective population size in *B. pratorum* were 43,981 for the Antrim population and 113,907 for the panmictic population across England and Wales (Table 1). Trajectories of genetic effective population size changes over the last 100 generations show a strong decrease in both species (Figure 4). This was 9-fold for the population in England and Wales in *B. pratorum* and 2–7 fold for *B. monticola* populations (Figure 4). In both species, the most recently colonised sampling site in Ireland (Antrim) exhibits the smallest population size over recent generations, and the least decrease in size (Figure 4).

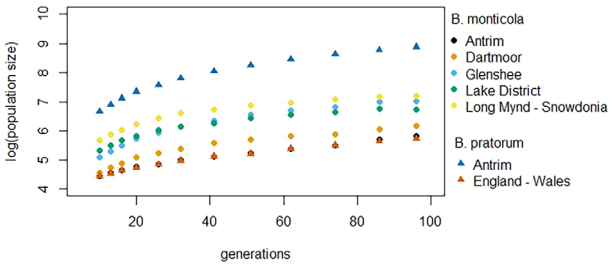

FIGURE 4 SNeP analysis for *Bombus monticola* and *Bombus pratorum* showing the trajectory of effective population size (Ne) change in the past. Note that Ne is plotted on a log scale and generations into the past increase along the x-axis. Results are presented for generations 10–100, given the uncertainties around estimates in the most recent and most distant generations (Barbato et al., 2015; Corbin et al., 2012).

Parasite prevalence

Across all sites, the prevalence of *L. buchneri* was significantly higher in *B. pratorum* than in *B. monticola* (Mann–Whitney–Wilcoxon test: $p < .001$) but no differences were found for the other parasites investigated (Figure 5). Patterns of parasite prevalence varied greatly between locations and were not correlated between *B. monticola* and *B. pratorum* (i.e., there was no pattern of parasite prevalence with geographic location, Figure S1). The number of parasites observed showed a significant relationship with habitat suitability in *B. monticola*, with lower suitability values being associated with higher parasite loads in both the model including all sampled locations (Table 3.1A) and the model restricted to only locations with genetic data (Table 3.1B). The 'all sample' model also indicated a positive association between age (estimated using wing wear) and parasite load. Date was also significant and showed a trend of decreasing parasite load throughout the year (Table 3.1A). In *B. pratorum* a significant positive effect of latitude was observed in both models (Table 3.1A, B). Significant effects of predictor variables were also identified for infections with the parasites *C. bombi*, *V. bombi* and *A. bombi* for both *B. monticola* and *B. pratorum* (Table 3.2–3.5).

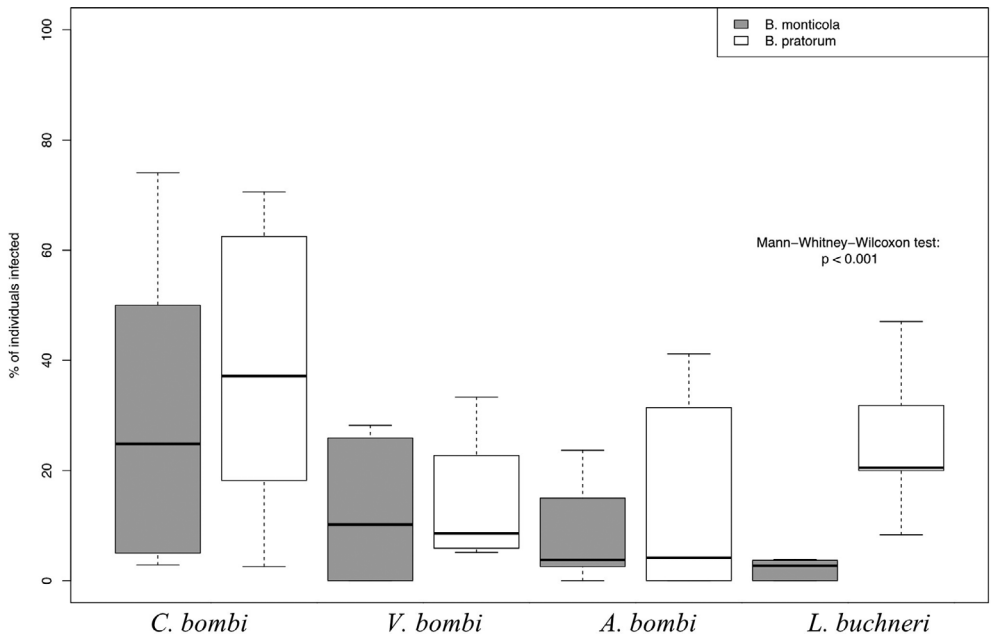

FIGURE 5 Percentage of individuals infected by *Crithidia bombi*, *Vairimorpha bombi*, *Apicystis bombi* and *Locustacharus buchneri* for *Bombus monticola* ($N = 308$) and *Bombus pratorum* ($N = 251$) across sampling locations.

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Results of the general linear model used to test relationships between: number of parasites with the predictor variables longitude, latitude, habitat suitability, wing wear and sampling date for all sampled locations (1A) and with the additional predictor of heterozygosity (1B) for all locations that included genetic information; the individual infection status with *C. bombi* (2), *V. bombi* (3), *A. bombi* (4) and *L. buchneri* (5)

	Bombus monticola					***Bombus pratorum***				
		Estimate	**SE**	**z value**	**Pr (>|t|)**		**Estimate**	**SE**	**z-Value**	**Pr (>|t|)**
	(1) Response variable: number of parasite species									
(A)	Suitability	−1.25	0.34	−3.69	0.0002	Latitude	0.21	0.09	2.32	0.0202
	Wing wear	0.12	0.06	2.11	0.0347	Suitability	−2.19	1.14	−1.92	0.0551
	Date	−0.01	0.00	−2.08	0.0373	Longitude	−0.07	0.05	−1.47	0.1430
	Longitude	0.13	0.07	1.75	0.0797					
(B)	Suitability	−1.61	0.46	−3.52	0.0004	Latitude	0.34	0.10	3.50	0.0005
	(2) Response variable: infection with *C. bombi*									
	Suitability	−2.99	0.61	−4.89	1.00 E-06	Latitude	0.73	0.22	3.29	0.0010
	Longitude	0.53	0.16	3.35	0.0008	Longitude	−0.29	0.11	−2.66	0.0079
	Wing wear	0.27	0.10	2.60	0.0093	Suitability	−5.20	3.00	−1.73	0.0833
	Latitude	0.21	0.11	1.82	0.0686					
	Date	−0.01	0.01	−1.79	0.0732					
	(3) Response variable: infection with *V. bombi*									
	Longitude	−1.09	0.17	−6.50	8.14 E-11	Latitude	3.00	0.76	3.94	8.08 E-05
	Latitude	−0.56	0.14	−4.16	3.12 E-05	Date	−0.26	0.07	−3.78	0.0002
						Longitude	0.99	0.32	3.08	0.0020
	(4) Response variable: infection with *A. bombi*									
	Longitude	0.80	0.25	3.14	0.0017	Suitability	−7.83	2.70	−2.90	0.0037
	Date	−0.01	0.01	−1.90	0.0580					
	(5) Response variable: infection with *L. buchneri*									
	Date	0.01	0.01	0.98	0.3275	Suitability	−2.32	2.58	−0.90	0.3690

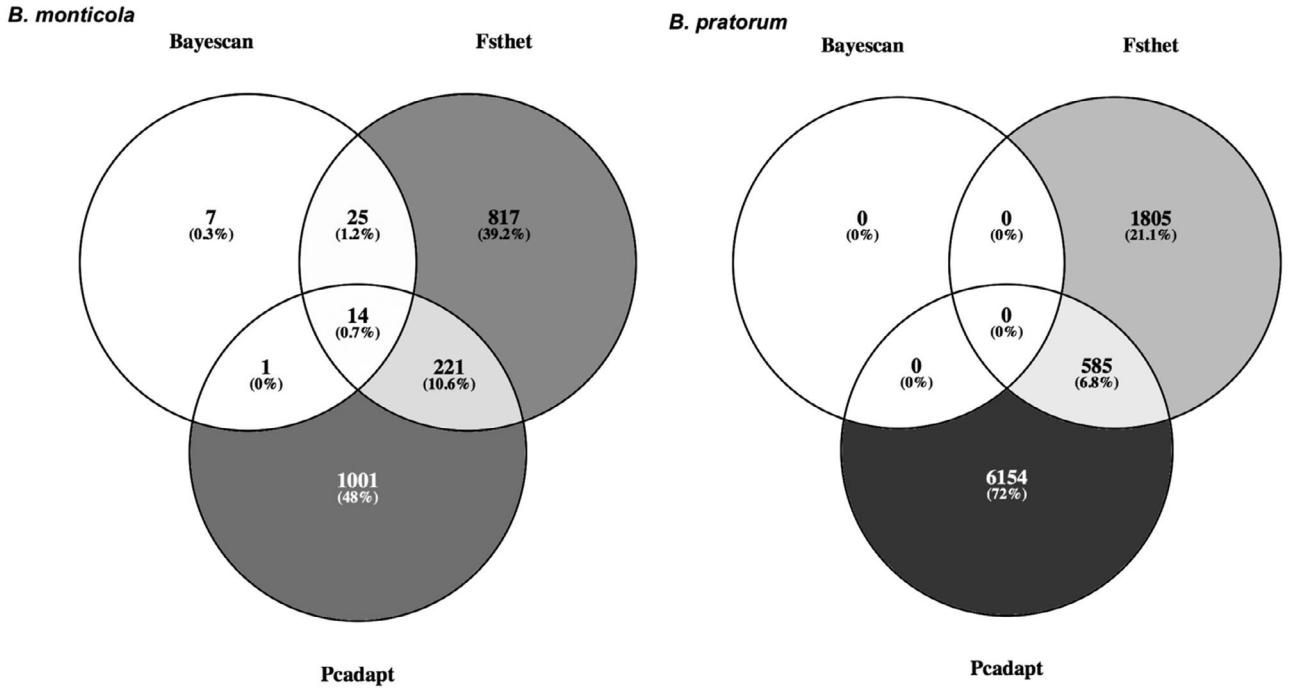

FIGURE 6 Venn diagram of outliers identified by each of the three methods applied for *Bombus monticola* (left) and *Bombus pratorum* (right).

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Detection of signatures of selection

F_{ST}-outlier approach

Of the three methods used, Bayescan identified the lowest number of outliers in both species (Figure 6) with no outliers identified for *B. pratorum*. This was followed by Fsthet and then Pcadapt, which identified the largest number of outliers. For *B. monticola* a small proportion of outliers were consistently identified by all three methods (Figure 6). Each of these fell within 5000 bp of one of eight genes, including a transcription factor (100647190), an RNA binding protein (100644912), an acetyl esterase (100649264), a receptor tyrosin kinase (100645523) and an inorganic phosphate cotransporter (100644143) (Table 4). Another subset of outliers identified were consistently called by at least two approaches (12.5% for *B. monticola* and 6.8% for *B. pratorum*) (Figure 6) and were considered for GO analysis. This revealed these were significantly overrepresented for the molecular function of DNA binding transcription factor activity in both species (Table 5).

TABLE 4 F_{ST}-outliers identified by all three methods in *B. monticola*

GeneID	Name	Chromosome
100646769	Uncharacterized LOC100646769	LG B01
100643686	Plexus	LG B04
100645523	Discoidin domain-containing receptor 2	LG B07
100649264	Neuroligin-4, X-linked	LG B09
100644143	Putative inorganic phosphate cotransporter	LG B11
100647190	Transcription factor Sox-19b	LG B11
100647937	Protein sprouty homologue 2	LG B12
100644912	RNA-binding protein Musashi homologue Rbp6	Unknown

Selective sweep detection

Signals of selective sweeps detected by the RaisD programme were similar across populations of the same species, but also showed some congruency across species. Of all genes within areas identified as outliers in any population, 70% and 19% were consistently found to be outliers across both populations in *B. pratorum* and across at least four out of five populations in *B. monticola*, respectively. Molecular functions and biological processes, over-represented in areas that showed signals of a selective sweep in both species, were predominantly related to neurotransmission and regulation of transcription (Table 6).

DISCUSSION

Population structure and genetic diversity

Our results are largely consistent with the predicted evolutionary and population genetic effects affecting species with different demographic histories. Gene flow is much higher in the widespread and more common *B. pratorum* than in *B. monticola*, with no population structure detected across mainland UK populations. In contrast

TABLE 5 Gene ontology analysis of F_{ST}-outliers identified by at least two methods for *Bombus monticola* and *Bombus pratorum* respectively; significant *p*-values are highlighted in bold; results are presented where at least two evaluation metrics (elim/weight/classic) were significant

	GO.ID	Term	Annotated	Significant	Expected	Elim	Weight	Classic
Molecular function								
B. monticola	GO:0005515	Protein binding	580	48	36.62	**0.013**	**0.015**	0.013
	GO:0003700	DNA-binding transcription factor activity	135	15	8.52	**0.019**	**0.019**	0.019
	GO:0043565	Sequence-specific DNA binding	46	7	2.90	**0.023**	**0.023**	0.023
	GO:0008236	Serine-type peptidase activity	30	5	1.89	**0.037**	**0.037**	0.037
B. pratorum	GO:0004175	Endopeptidase activity	20	11	3.24	**0.0054**	**0.0054**	0.0054
	GO:0000981	DNA-binding transcription factor activity	13	8	3.41	**0.0073**	**0.0073**	0.0073
	GO:0016829	Lyase activity	13	7	3.41	**0.0302**	**0.0302**	0.0302
	GO:0042626	ATPase-coupled transmembrane transporter activity	11	6	2.88	**0.0417**	**0.0417**	0.0417
Biological processes								
B. monticola	GO:0007275	Multicellular organism development	39	7	2.62	**0.013**	**0.0048**	0.0128
B. pratorum	GO:0006412	Translation	21	10	5.91	**0.043**	**0.043**	0.043

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 Gene ontology analysis of RaisD outliers identified in each species separately and combined (in a minimum of four population clusters for *B. monticola* and 2 population clusters for *B. pratorum*); significant *p*-values are highlighted in bold; results are presented where at least two evaluation metrics (elim/weight/classic) were significant

	GO.ID	Term	Annotated	Significant	Expected	Elim	Weight	Classic
Molecular functions								
Both species	GO:0022824	Transmitter-gated ion channel activity	20	7	1.68	**0.00082**	**0.00082**	0.00082
	GO:0099094	Ligand-gated cation channel activity	12	5	1.01	**0.00195**	**0.00195**	0.00195
	GO:0140101	Catalytic activity, acting on a trna	49	10	4.11	**0.00645**	**0.00645**	0.00645
	GO:0052689	Carboxylic ester hydrolase activity	21	5	1.76	**0.02676**	**0.02676**	0.02676
	GO:0017171	Serine hydrolase activity	79	12	6.63	**0.03036**	**0.03036**	0.03036
	GO:0008270	Zinc ion binding	184	23	15.44	**0.03264**	**0.03264**	0.03264
	GO:0061630	Ubiquitin protein ligase activity	31	6	2.60	**0.04087**	**0.04087**	0.04087
B. monticola	GO:0099094	Ligand-gated cation channel activity	12	8	2.69	**1.6 e-05**	**0.0013**	0.0013
	GO:0022824	Transmitter-gated ion channel activity	24	10	5.38	**0.00029**	**0.0271**	0.0271
B. pratorum	GO:0005198	Structural molecule activity	155	123	99.87	**0.00058**	**3,00 E-05**	**3,00 E-05**
	GO:0004519	Endonuclease activity	16	15	10.31	**0.00859**	**0.0086**	0.0086
	GO:0005102	Signalling receptor binding	36	31	23.20	**0.00342**	**0.0153**	0.0034
	GO:0004930	G protein-coupled receptor activity	63	49	40.59	**0.01569**	**0.0157**	0.0157
	GO:0016627	Oxidoreductase activity, acting on the CH—CH group of donors	25	21	16.90	**0.02786**	**0.0279**	0.0279
	GO:0004867	Serine-type endopeptidase inhibitor activity	12	11	7.73	**0.03886**	**0.0389**	0.0389
Biological process								
Both species	GO:0071805	Potassium ion transmembrane transport	22	9	1.90	**0.0104**	**0.0104**	**4.2 e-05**
	GO:0030431	Sleep	11	5	0.95	**0.0014**	**0.0014**	**0.00138**
	GO:0032222	Regulation of synaptic transmission, cholinergic	11	5	0.95	**0.0014**	**0.0014**	**0.00138**
	GO:1903818	Positive regulation of voltage-gated potassium channel activity	11	5	0.95	**0.0014**	**0.0014**	**0.00138**
	GO:0019318	Hexose metabolic process	12	4	1.03	**0.0154**	**0.0154**	0.01537
	GO:0006418	Trna aminoacylation for protein translation	26	6	2.24	**0.0205**	**0.0205**	0.02049
	GO:0016570	Histone modification	23	5	1.98	**0.0426**	**0.0426**	0.04256

(Continues)

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 (Continued)

	GO.ID	Term	Annotated	Significant	Expected	Elim	Weight	Classic
	GO:0045892	Negative regulation of transcription, DNA templated transcription	16	4	2.85	**0.0430**	**0.0430**	0.04297
B. monticola	GO:0032222	Regulation of synaptic transmission, cholinergic	11	6	2.49	**0.00094**	**0.021**	0.021
	GO:0030431	Sleep	11	6	2.49	**0.00094**	**0.021**	0.021
	GO:1903818	Positive regulation of voltage-gated potassium channel activity	11	6	2.49	**0.00094**	**0.021**	0.021
B. pratorum	GO:0006302	Double-strand break repair	11	11	7.1	**0.0080**	**0.008**	0.0080
	GO:0010948	Negative regulation of cell cycle processes	10	10	5.45	**0.0124**	**0.012**	0.0124
	GO:0007178	Transmembrane receptor protein serine/threonine kinase signalling pathway	10	10	5.45	**0.0124**	**0.012**	0.0124
	GO:0009967	Positive regulation of signal transduction	14	13	9.3	**0.0187**	**0.019**	0.0187
	GO:1903311	Regulation of mrna metabolic process	12	11	7.74	**0.0393**	**0.039**	0.0393
	GO:2001259	Positive regulation of cation channel activity	12	11	7.74	**0.0393**	**0.039**	0.0393
	GO:0050804	Modulation of chemical synaptic transmission	16	14	10.32	**0.0412**	**0.041**	0.0412
	GO:0044770	Cell cycle phase transition	15	12	7.68	0.1624	**0.046**	0.1624

B. monticola exhibited strong population structuring with low levels of gene flow. The only sampling locations where significant differentiation was not observed were Long Mynd and Snowdonia, which are less than 100 km apart and connected largely by habitat of high suitability according to the estimates of Polce et al. (2018).

Consistent with a larger effect of genetic drift in *B. monticola*, the number and proportion of polymorphic sites was on average more than three times higher in *B. pratorum* (Table 1). The relatively smaller proportion of polymorphic sites observed in *B. monticola* may be of concern given the lower adaptive genetic potential this species therefore has to respond to future environmental change. However, estimates of heterozygosity and nucleotide diversity across polymorphic sites were comparable between both species and only the Dartmoor population showed any strong evidence of inbreeding in *B. monticola*. Estimates of genetic diversity are also similar to those observed in other bumblebee species assessed using a RAD-seq approach (Huml et al., 2021; Jackson et al., 2018; Lozier, 2014).

Trends in effective population size

The effective population size is formally defined as the 'the size of the ideal (Wright-Fisher) population (*N*) that will result in the same amount of gentic drift as in the actual population being considered' (Allendorf et al., 2013). Its relationship with the census population size (Nc) varies greatly among taxa but is generally expected to be considerably lower than the census population size in most scenarios, given that not all individuals in a population successfully breed into the next generation and that there is likely to be variance in family size in natural populations. In eusocial insects, Ne is expected to be particularly lower than Nc, since the unit of reproduction is the colony and not

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the individual (Chapman & Bourke, 2001). The 'threshold' of concern for Ne is a subject of much debate and the use of 'magic numbers' has been much derided (Caughley, 1994; Flather et al., 2011; Lande, 1988, 1995). Regardless, the long-term values of Ne reported for both species are comfortably higher than the thresholds generally discussed in the literature (Lande, 1995; see also Frankham et al., 2014).What is surprising, however, is that the Ne values observed for *B. pratorum* are of similar orders of magnitude to *B. monticola* (Table 1), despite the former being generally considered as widespread and common in the UK (and Ireland). Lattorff et al. (2016) reported much higher Ne values than those observed here for other common bumblebee species using the same approach, although as those were continental European populations this would be expected (*B. terretris* 2.7×10^6; *B. lapidarius* 6.4×10^5, estimates that would be higher when applied to the *Bombus* mutation rate that has since become available (Liu et al., 2017)). We note the danger of over-interpreting raw values using single point datasets (Nadachowske-Bryska et al., 2022), and instead focus on the general pattern of decline observed in both species over recent generations (Figure 4). While these analyses show significant declines over the last 100 generations in both species, the decline in Ne of *B. pratorum* is more pronounced than that of *B. monticola*. Importantly, our results therefore highlight that significant declines are affecting common species (here *B. pratorum*) as well as historically range-restricted, more specialist species, which have tended to be the focus of conservation genetic studies to date. Although *B. pratorum* is perceived to be widespread and common, we again note that there is an important distinction between Nc and Ne. This pattern of high apparent abundance (Nc), but relatively low effective population size (Ne) may reflect a lower annual reproductive success or higher variance in nest success of *B. pratorum* colonies than has been assumed. Few studies have addressed the proportion of nests that successfully reproduce due to the challenges with finding and monitoring nests. However, both Carvell et al. (2017) and Boone et al. (2022) report highly variable annual lineage survival in other *Bombus* species. The former study also observed a strong association of colony reproductive output with habitat quality. Furthermore, Boone et al. (2022) note that in *B. affinis* a colony of 1341 individuals produced only 9 gynes and a colony of 251 individuals, only 2. Clearly, determining variance and mean nest success with regard to the production of reproductive individuals in the wild is challenging, but merits further study.

Patterns of parasite prevalence

The number of infections with different parasites was best explained by the distribution of habitat suitablility in *B. monticola* and latitude in *B. pratorum* (Table 3). Habitat suitability estimates likely relate to the distribution, quality and seasonal availability of ressources. Maintenance of immunity is energetically costly (Schmid-Hempel, 2005) and pollen deprivation has been shown to reduce both the constitutive immunity (Roger et al., 2017) and the specificity of the immune response in *B. terrestris* (Brunner et al., 2014) although food deprivation can also lead to reduced parasite loads where these rely on host nutrition (Conroy et al., 2016; Logan et al., 2005). Consistent with our results of higher infection rates for multiple parasites in *B. pratorum* at higher latitudes, McArt et al. (2017) also identified a trend for greater prevalence of *V. bombi* and increased local losses in four bumblebee species in North America at higher latitudes. Although fungicide usage showed a similar geographic pattern, and was the strongest predictor of *V. bombi* prevalence identified by McArt et al. (2017), a higher prevalence of pathogens at higher latitudes may be a factor contributing to the generally observed lack of bumblebees tracking changing climate conditions at northern range limits (Kerr et al., 2015).

Although inbreeding and loss of genetic diversity have been reported to represent a major factor for increased susceptibility to disease (Ekroth et al., 2019; Spielman, Brook, Briscoe, & Frankham, 2004), here we did not find any associations between parasite prevalence and genetic diversity. This contrasts with results from studies at the colony-level. For example, lower parasite loads have been observed for *C. bombi* and *V. bombi* in *B. terrestris* within more genetically diverse colonies (Baer & Schmid-Hempel, 1999; Huth-Schwarz et al., 2012; Liersch & Schmid-Hempel, 1998; Parsche & Lattorff, 2018). Similarly, Cameron et al. (2011) observed higher prevalence of *V. bombi* in declining populations with lower genetic diversity. Although in the absence of polyandrous mating a large proportion of colony diversity should be captured by sampling an individual worker, colony parasite loads may be significantly under- or over-estimated from just one worker, as reported in other studies (Whitehorn et al., 2014). Although we also investigated population level effects of heterozygosity on susceptibility, the number of sampled populations was too small ($n = 6$ and 5 for *B. monticola* and *B. pratorum*, respectively) to derive reliable conclusions.

Signatures of selection

There were some striking similarities between the two species in signatures of selective sweeps within genomic regions relating to neurotransmission (e.g., ion channel activities, synaptic transmission and sleep) and regulation of transcription (Table 6). A recent study on another species of *Pyrobombus*, *Bombus hypnorum*, also reported evidence for selection acting on regions related to ion channel activities and regulation of transcription (Huml et al., 2021). Signatures of selection on genes involved in neurotransmission have also been reported across UK populations of *Bombus terrestris* (Colgan et al., 2022). Sun et al. (2021) found ionotropic glutamate receptor activity, important in neurotransmission (Benton et al., 2009), to be among the top 20% of genes showing elevated dN/dS ratio across a wider range of *Bombus* species, which may indicate functional divergence in respect to neurotransmission in bumblebees. Such areas of high divergence among species are of particular interest, as these are likely to be involved in lineage specific adaptations (Ellegren, 2014). Ionotropic glutamate receptors were among the genes showing evidence for positive

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

selection in both species here (Table S2). Among the biological processes showing evidence for a selective sweep, we identified regulation of cholinergic synaptic transmission (Table 6), well known to represent the primary target for neonicotinoids (Matsuda et al., 2001). These results support a growing body of evidence for strong selection on neurological genes in bumblebees and other insects, which may be indicative of adaptive shifts in response to modern selection pressures, such as pesticides, in this group. We note, however, the need to be cautious to avoid over- or mis-interpreting genomic patterns in this context, and highlight the need for further research in this area.

A RAD-seq approach inevitably only samples a proportion of the genome. While this inevitably means that many loci are not screened this does not negate the results reported (Lowry et al., 2017; Tiffin & Ross-Ibarra, 2014). Unless a species exhibits exceeding levels of polymorphism, RAD-seq is an effective method to obtain unbiased estimates of genomic diversity across the genome (Cariou et al., 2016), and is a powerful tool for the study of population genomics and natural selection (McKinney et al., 2017).

Conservation

The significantly declining trajectories reported here for both species follow national and international warnings of widespread insect declines (e.g., Hallman et al., 2017; Powney et al., 2019; Sánchez-Bavo & Wyckhuys, 2019, but see also Simmons et al., 2019).

Here, habitat suitablility strongly affected the range-restricted species, *B. monticola*, both in terms of the highly restricted population connectivity and parasite prevalence. This suggests that this species is particularly vulnerable to habitat degradation, and at one site (Dartmoor) we observed some potential indicators of population stress (inbreeding). The conservation actions relevant for *B. monticola* would seem relatively straight-forward: conservation and restoration of upland heathland habitats, already identified as a conservation priority under the EU Habitats Directive 1994 due to widespread losses over the last two centuries, particularly in the UK (Fagúndez, 2013). The picture for *B. pratorum* is less tractable and arguably more concerning. Wide areas of habitat suitable for this species are available, but either that habitat is of declining quality, or other factors relating to declines are operating, or both. There are several possible and non-exclusive possibilities, for example, changes in weather patterns, long-term exposure to agrochemcials and pollutants, and decreased availability of suitable forage and nesting sites. Although the potential genomic regions under selection include targets for neonicotinoids, which may be indicative of the particular stresses that these populations are under, these are by no means exhaustive and require further study for causative links to be established. Decreases in the abundance of common species are of particular concern as these disproportionally affect ecosystem services (Winfree et al., 2015). We might take heed from the warning of *Bombus sylvarum*: now one of the UK's rarest species it was common across large parts of the the UK just a century ago (Sladen, 1912). Similar patterns of declines in common *Bombus* species have been reported in North America (Cameron et al., 2011).

Evidently, there is a presssing need for effective monitoring systems and research focus on the drivers of declines not just on rare and already threatened species, but for those species that are apparently of least conservation concern (Gaston & Fuller, 2008; Lindenmayer et al., 2011; Roycroft et al., 2021). Identifying differences in the potential of species to adaptively respond to current threats and environmental change is likely to enhance our understanding of species vulnerability to decline, and help in directing conservation efforts more effectively (Hoelzel et al., 2019; Lavergne et al., 2013; Salamin et al., 2010).

AUTHOR CONTRIBUTIONS

Jana V. Huml: Writing – original draft; writing – review and editing; visualization; methodology; formal analysis; investigation; funding acquisition. **Jonathan S. Ellis:** Conceptualization; investigation; funding acquisition; writing – review and editing; supervision; writing – original draft. **Sarah Rustage:** Resources; methodology; writing – review and editing. **Mark J. F. Brown:** Conceptualization; funding acquisition; writing – review and editing. **Richard Billington:** Writing – review and editing; conceptualization; funding acquisition. **Mairi E. Knight:** Conceptualization; investigation; funding acquisition; project administration; supervision; writing – original draft; writing – review and editing.

ACKNOWLEDGEMENTS

We would like to thank our funders (listed below) and The National Trust, Natural England, Natural Resources Wales (NRW) and Scottish Natural Heritage (SNH, now NatureScot) for field sampling permissions. We also thank both Anna Harte and David Hemprich-Bennett for their help with sampling.

FUNDING INFORMATION

This project was funded initially through a NERC PhD studentship (Sarah Rustage) followed by funding from the Seale Hayne Educational Trust (Mairi E. Knight, Jonathan S. Ellis and Jana V. Huml) for the genomic screening.

CONFLICT OF INTEREST

The authors declare no conflict of interests.

DATA AVAILABILITY STATEMENT

Demultiplexed reads generated in this study are available on GenBank (SRA) (PRJNA918549).

ORCID

Jana V. Huml https://orcid.org/0000-0003-4914-0007
Sarah Rustage https://orcid.org/0000-0002-4923-070X

REFERENCES

Ahrens, M., Ammerpohl, O., von Schönfels, W., Kolarova, J., Bens, S., Itzel, T. et al. (2013) DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. *Cell Metabolism*, 18, 296–302.

Alachiotis, N. & Pavlidis, P. (2018) RAiSD detects positive selection based on multiple signatures of a selective sweep and SNP vectors. *Communications Biology*, 1(1), 1–11.

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Alexa, A. & Rahnenfuhrer, J. (2016) topGO: Enrichment analysis for Gene Ontology. R package version 2.28. 0. Cranio.

Alexa, A., Rahnenführer, J. & Lengauer, T. (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22, 1600–1607.

Allen, G.R., Seeman, O.D., Schmid-Hempel, P. & Buttermore, R.E. (2007) Low parasite loads accompany the invading population of the bumblebee, Bombus terrestris in Tasmania. *Insectes Sociaux*, 54(1), 56–63.

Allendorf, F.W., Hohenlohe, P.A. & Luikart, G. (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics*, 11, 697–709.

Allendorf, F.W., Luikart, G. & Aitken, S.N. (2013) *Conservation and the Genetics of Populations*, 2nd edition. Oxford: Wiley-Blackwell, p. 118.

Arbetman, M.P., Gleiser, G., Morales, C.L., Williams, P. & Aizen, M.A. (2017) Global decline of bumblebees is phylogenetically structured and inversely related to species range size and pathogen incidence. *Proceedings of the Royal Society B*, 284, 20170204.

Attia, J., Thakkinstian, A., McElduff, P., Milne, E., Dawson, S., Scott, R.J. et al. (2010) Detecting genotyping error using measures of degree of Hardy-Weinberg disequilibrium. *Statistical Applications in Genetics and Molecular Biology*, 9(1), 1–17.

Baer, B. & Schmid-Hempel, P. (1999) Experimental variation in polyandry affects parasite loads and fitness in a bumble-bee. *Nature*, 397(6715), 151–154.

Barbato, M., Orozco-terWengel, P., Tapio, M. & Bruford, M.W. (2015) SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Frontiers in Genetics*, 6, 109.

BCT. (2022) Bumblebee Conservation Trust. https://www.bumblebeeconservation.org

Benton, R., Vannice, K.S., Gomez-Diaz, C. & Vosshall, L.B. (2009) Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell*, 136(1), 149–162.

Biella, P., Ćetković, A., Gogala, A., Neumayer, J., Sárospataki, M., Šima, P. et al. (2021) Northwestward range expansion of the bumblebee *Bombus haematurus* into Central Europe is associated with warmer winters and niche conservatism. *Insect Sci.*, 28(3), 861–872.

Boone, M.L., Evans, E., Wolf, A., Halley, M.H., Watson, J. & Smith, T.A. (2022) Notes from rusty patched bumble bee (*Bombus affinis* Cresson) nest observations. *Insect Conservation and Diversity*, 2022, 1–5.

Browning, B.L. & Browning, S.R. (2007) Efficient multilocus association testing for whole genome association studies using localized haplotype clustering. *Genetic Epidemiology*, 31(5), 365–375.

Brunner, F.S., Schmid-Hempel, P. & Barribeau, S.M. (2014) Protein-poor diet reduces host-specific immune gene expression in *Bombus terrestris*. *Proceedings of the Royal Society B*, 281(1786), 20140128.

Bush, A., Mokany, K., Catullo, R., Hoffmann, A., Kellerman, V., Sgrò, C. et al. (2016) Incorporating evolutionary adaptation in species distribution modelling reduces projected vulnerability to climate change. *Ecology Letters*, 19(12), 1468–1478.

BWARS. (2022) Bees, Wasps and Ants Recording society. https://www.bwars.com

Cameron, S.A., Hines, H.M. & Williams, P.H. (2007) A comprehensive phylogeny of the bumble bees (Bombus). *Biological Journal of the Linnean Society*, 91(1), 161–188.

Cameron, S.A., Lozier, J.D., Strange, J.P., Koch, J.B., Cordes, N., Solter, L.F. et al. (2011) Patterns of widespread decline in North American bumble bees. *Proceedings of the National Academy of Sciences*, 108(2), 662–667.

Cameron, S.A. & Sadd, B.M. (2020) Global trends in bumble bee health. *Annual Review of Entomology*, 65, 209–232.

Cariou, M., Duret, L. & Charlat, S. (2016) How and how much does RAD-seq bias genetic diversity estimates? *BMC Evolutionary Biology*, 16(1), 1–8.

Carvell, C., Bourke, A.F.G., Dreier, S., Freeman, S.N., Hulmes, S., Jordan, W.C. et al. (2017) Bumblebee family lineage survival is enhanced in high-quality landscapes. *Nature*, 543, 547–549.

Casey, L.M., Rebelo, H., Rotheray, E. & Goulson, D. (2015) Evidence for habitat and climatic specializations driving the long-term distribution trends of UK and Irish bumblebees. *Diversity and Distributions*, 21(8), 864–875.

Caughley, G. (1994) Directions in conservation biology. *Journal of Animal Ecology*, 63(2), 215–244.

Chambers, J.M. & Hastie, T.J. (1992) *Statistical Models in S*. London: Chapman & Hall.

Chapman, R.E. & Bourke, A.F. (2001) The influence of sociality on the conservation biology of social insects. *Ecology Letters*, 4(6), 650–662.

Cockerham, C.C. & Weir, B.S. (1993) Estimation of gene flow from F-statistics. *Evolution*, 47, 855–863.

Colgan, T.J., Arce, A.N., Gill, R.J., Rodrigues, A.R., Kanteh, A., Duncan, A.J. et al. (2022) Genomic signatures of recent adaptation in a wild bumblebee. *Molecular Biology and Evolution*, 39(2), msab366.

Colla, S.R., Otterstatter, M.C., Gegear, R.J. & Thomson, J.D. (2006) Plight of the bumble bee: pathogen spillover from commercial to wild populations. *Biological Conservation*, 129(4), 461–467.

Conroy, T.J., Palmer-Young, E.C., Irwin, R.E. & Adler, L.S. (2016) Food limitation affects parasite load and survival of *Bombus impatiens* (Hymenoptera: Apidae) infected with *Crithidia* (Trypanosomatida: Trypanosomatidae). *Environmental Entomology*, 45(5), 1212–1219.

Corbin, L.J., Liu, A.Y.H., Bishop, S.C. & Woolliams, J.A. (2012) Estimation of historical effective population size using linkage disequilibria with marker data. *Journal of Animal Breeding and Genetics*, 129, 257–270.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A. et al. (2011) The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158.

de Jong, M.J., de Jong, J.F., Hoelzel, A.R. & Janke, A. (2021) SambaR: An R package for fast, easy and reproducible population-genetic analyses of biallelic SNP data sets. *Molecular Ecology Resources*, 21(4), 1369–1379.

De Mita, S., Thuillet, A.-C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J. et al. (2013) Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, 22, 1383–1399.

DeWoody, J.A., Harder, A.M., Mathur, S. & Willoughby, J.R. (2021) The long-standing significance of genetic diversity in conservation. *Molecular Ecology*, 30(17), 4147–4154.

Doums, C. & Schmid-Hempel, P. (2000) Immunocompetence in workers of a social insect, *Bombus terrestris* L., in relation to foraging activity and parasitic infection. *Canadian Journal of Zoology*, 78(6), 1060–1066.

Dray, S. & Dufour, A.B. (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, 22(4), 1–20.

Ekroth, A.K., Rafaluk-Mohr, C. & King, K.C. (2019) Host genetic diversity limits parasite success beyond agricultural systems: a meta-analysis. *Proceedings of the Royal Society B*, 286(1911), 20191811.

Ellegren, H. (2014) Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, 29, 51–63.

Ellegren, H. & Galtier, N. (2016) Determinants of genetic diversity. *Nature Reviews Genetics*, 17(7), 422–433.

Evans, S.R. & Sheldon, B.C. (2008) Interspecific patterns of genetic diversity in birds: correlations with extinction risk. *Conservation Biology*, 22(4), 1016–1025.

Fagúndez, J. (2013) Heathlands confronting global change: drivers of biodiversity loss from past to future scenarios. *Annals of Botany*, 111(2), 151–172.

Fitzpatrick, Ú., Murray, T.E., Paxton, R.J., Breen, J., Cotton, D., Santorum, V. et al. (2007) Rarity and decline in bumblebees–a test of causes and correlates in the Irish fauna. *Biological Conservation*, 136(2), 185–194.

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Flanagan, S.P. & Jones, A.G. (2017) Constraints on the F ST–heterozygosity outlier approach. *Journal of Heredity*, 108(5), 561–573.
Flather, C.H., Hayward, G.D., Beissinger, S.R. & Stephens, P.A. (2011) Minimum viable populations: is there a 'magic number' for conservation practitioners? *Trends in Ecology & Evolution*, 26(6), 307–316.
Foll, M. & Gaggiotti, O. (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, 180, 977–993.
Foster, D.J. & Cartar, R.V. (2011) What causes wing wear in foraging bumble bees? *Journal of Experimental Biology*, 214(11), 1896–1901.
Frankham, R., Bradshaw, C.J.A. & Brook, B.W. (2014) Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation*, 170, 56–63.
Gaston, K. & Fuller, R. (2008) Commonness, population depletion and conservation biology. *Trends in Ecology & Evolution*, 23, 14–19.
Gilpin, M.E. & Soulé, M.E. (1986) Minimum viable populations: processes of species extinction. In: Soulé, M.E. (Ed.) *Conservation Biology: The Science of Scarcity and Diversity*. Sunderland, Mass: Sinauer, pp. 19–34.
Gloag, R., Ding, G., Christie, J.R., Buchman, G., Beekman, M. & Oldroyd, B. P. (2016) An invasive social insect overcomes genetic load at the sex locus. *Nature Ecology Evolution*, 1(1), 11.
Goudet, J. (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5(1), 184–186.
Goulson, D., Hanley, M.E., Darvill, B., Ellis, J.S. & Knight, M.E. (2005) Causes of rarity in bumblebees. *Biological Conservation*, 122, 1–8.
Goulson, D., Kaden, J.C., Lepais, O., Lye, G.C. & Darvill, B. (2011) Population structure, dispersal and colonization history of the garden bumblebee *Bombus hortorum* in the Western Isles of Scotland. *Conservation Genetics*, 12(4), 867–879.
Goulson, D., Whitehorn, P. & Fowley, M. (2012) Influence of urbanisation on the prevalence of protozoan parasites of bumblebees. *Ecological Entomology*, 37(1), 83–89.
Gower, J.C. (1966) Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53(3–4), 325–338.
Habel, J.C. & Schmitt, T. (2012) The burden of genetic diversity. *Biological Conservation*, 147(1), 270–274.
Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H. et al. (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS One*, 12(10), e0185809.
Hoelzel, A.R., Bruford, M.W. & Fleischer, R.C. (2019) Conservation of adaptive potential and functional diversity. *Conservation Genetics*, 20, 1–5.
Huml, J.V., Ellis, J.S., Lloyd, K., Benefer, C.M., Kiernan, M., Brown, M.J.F. et al. (2021) Bucking the trend of pollinator decline: the population genetics of a range expanding bumblebee. *Evolutionary Ecology*, 35(3), 413–442.
Huth-Schwarz, A., Settele, J., Moritz, R.F. & Kraus, F.B. (2012) Factors influencing *Nosema bombi* infections in natural populations of *Bombus terrestris* (Hymenoptera: Apidae). *Journal of Invertebrate Pathology*, 110(1), 48–53.
Jackson, J.M., Pimsler, M.L., Oyen, K.J., Koch-Uhuad, J.B., Herndon, J.D., Strange, J.P. et al. (2018) Distance, elevation and environment as drivers of diversity and divergence in bumble bees across latitude and altitude. *Molecular Ecology*, 27(14), 2926–2942.
Kardos, M., Armtrong, E., Fitzptric, S.W., Hauser, S., Hedrick, P.W., Miller, J.M. et al. (2021) The crucial role of genome-wide genetic variation in conservation. *PNAS*, 118(48), e2104642118.
Kerr, J.T., Pindar, A., Galpern, P., Packer, L., Potts, S.G., Roberts, S.M. et al. (2015) Climate change impacts on bumblebees converge across continents. *Science*, 349(6244), 177–180.
Khan, A., Patel, K., Shukla, H., Viswanathan, A., van der Valk, T., Borthakur, U. et al. (2021) Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers. *PNAS*, 118(49), e2023018118.
Kinsella, R.J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G. et al. (2011) Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database*, 2011, bar030.
Klee, J., Tay, W.T. & Paxton, R.J. (2006) Specific and sensitive detection of *Nosema bombi* (Microsporidia: Nosematidae) in bumble bees (*Bombus* spp.; Hymenoptera: Apidae) by PCR of partial rRNA gene sequences. *Journal of Invertebrate Pathology*, 91(2), 98–104.
Krohn, A.R., Diepeveen, E.T., Bi, K. & Rosenblum, E.B. (2019) Local adaptation does not lead to genome-wide differentiation in lava flow lizards. *Ecology and Evolution.*, 9(12), 6810–6820.
Lande, R. (1988) Genetics and demography in biological conservation. *Science*, 241(4872), 1455–1460.
Lande, R. (1995) Mutation and conservation. *Conservation Biology*, 9(4), 782–791.
Lanfear, R., Ho, S., Love, D. & Bromham, L. (2010) Mutation rate is linked to diversification in birds. *Proceedings of the National Academy of Sciences*, 107(47), 20423–20428.
Lattorff, H.M.G., Popp, M., Parsche, S., Helbing, S. & Erler, S. (2016) Effective population size as a driver for divergence of an antimicrobial peptide (Hymenoptaecin) in two common European bumblebee species. *Biological Journal of the Linnean Society*, 119, 299–310.
Lavergne, S., Evans, M.E., Burfield, I.J., Jiguet, F. & Thuiller, W. (2013) Are species' responses to global change predicted by past niche evolution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1610), 20120091.
Leffler, E.M., Bullaughey, K., Matute, D.R., Meyer, W.K., Segurel, L., Venkat, A. et al. (2012) Revisiting an old riddle: what determines genetic diversity levels within species? *PLoS Biology*, 10, e100138.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N. et al. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.
Liersch, S. & Schmid-Hempel, P. (1998) Genetic variation within social insect colonies reduces parasite load. *Proceedings of the Royal Society of London B*, 265(1392), 221–225.
Lindenmayer, D.B., Wood, J.T., McBurney, L., MacGregor, C., Youngentob, K. & Banks, S.C. (2011) How to make a common species rare: a case against conservation complacency. *Biological Conservation*, 144(5), 1663–1672.
Liu, H., Jia, Y., Sun, X., Tian, D., Hurst, L.D. & Yang, S. (2017) Direct determination of the mutation rate in the bumblebee reveals evidence for weak recombination-associated mutation and an approximate rate constancy in insects. *Molecular Biology and Evolution*, 34(1), 119–130.
Logan, A., Ruiz-González, M.X. & Brown, M.J.F. (2005) The impact of host starvation on parasite development and population dynamics in an intestinal trypanosome parasite of bumble bees. *Parasitology*, 130(6), 637–642.
López-Cortegano, E., Vilas, A., Caballero, A. & García-Dorado, A. (2016) Estimation of genetic purging under competitive conditions. *Evolution*, 70(8), 1856–1870.
Lowry, D.B., Hoban, S., Kelley, J.L., Lotterhos, K.E., Reed, L.K., Antolin, M. F. et al. (2017) Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, 17(2), 142–152.
Lozier, J.D. (2014) Revisiting comparisons of genetic diversity in stable and declining species: assessing genome-wide polymorphism in North American bumble bees using RAD sequencing. *Molecular Ecology*, 23(4), 788–801.
Luu, K., Bazin, E. & Blum, M.G. (2017) pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17(1), 67–77.
Mantel, N.A. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27, 209–220.

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T.H., Piñero, D. & Emerson, B.C. (2015) Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources*, 15(1), 28–41.

Matsuda, K., Buckingham, S.D., Kleier, D., Rauh, J.J., Grauso, M. & Sattelle, D.B. (2001) Neonicotinoids: insecticides acting on insect nicotinic acetylcholine receptors. *Trends in Pharmacological Sciences*, 22(11), 573–580.

McArt, S.H., Urbanowicz, C., McCoshum, S., Irwin, R.E. & Adler, L.S. (2017) Landscape predictors of pathogen prevalence and range contractions in US bumblebees. *Proceedings of the Royal Society B*, 284(1867), 20172181.

McKinney, G.J., Larson, W.A., Seeb, L.W. & Seeb, J.E. (2017) RAD seq provides unprecedented insights into molecular ecology and evolutionary genetics: comment on Breaking RAD by Lowry et al. (2016). *Molecular Ecology Resources*, 17(3), 356–361.

Meeus, I., De Graaf, D.C., Jans, K. & Smagghe, G. (2010) Multiplex PCR detection of slowly-evolving trypanosomatids and neogregarines in bumblebees using broad-range primers. *Journal of Applied Microbiology*, 109(1), 107–115.

Mueller, U.G. & Wolf-Mueller, B. (1993) A method for estimating the age of bees: Age-dependent wing wear and coloration in the Wool-Carder bee *Anthidium manicatum* (Hymenoptera: Megachilidae). *Journal of Insect Behavior*, 6(4), 529–537.

Nadachowska-Brzyska, K., Konczal, M. & Babik, W. (2022) Navigating the temporal continuum of effective population size. *Methods in Ecology & Evolution*, 13(1), 22–41.

Nonaka, E., Sirén, J., Somervuo, P., Ruokolainen, L., Ovaskainen, O. & Hanski, I. (2019) Scaling up the effects of inbreeding depression from individuals to metapopulations. *Journal of Animal Ecology*, 88(8), 1202–1214.

Nicholls, J.A., Double, M.C., Rowell, D.M. & Magrath, R.D. (2000) The evolution of cooperative and pair breeding in thornbills *Acanthiza*(Pardalotidae). *Journal of Avian Biology*, 31, 165–176.

O'Grady, J., Brook, B.W., Reed, D.H., Ballou, J.D., Tonkyn, D.W. & Frankham, F. (2008) Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation*, 133(1), 42–51.

O'Leary, S.J., Puritz, J.B., Willis, S.C., Hollenbeck, C.M. & Portnoy, D.S. (2018) These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, 27, 3193–3206.

Oliveros, J.C. (2007–2015) Venny. An Interactive Tool for Comparing Lists with Venn's Diagrams. https://bioinfogp.cnb.csic.es/tools/venny/index.html

Ouborg, N.J., Pertoldi, C., Loeschcke, V., Bijlsma, R.K. & Hedrick, P.W. (2010) Conservation genetics in transition to conservation genomics. *Trends in Genetics*, 26(4), 177–187.

Paris, J.R., Stevens, J.R. & Catchen, J.M. (2017) Lost in parameter space: a road map for stacks. *Methods in Ecology and Evolution*, 8(10), 1360–1373.

Parsche, S. & Lattorff, H.M.G. (2018) The relative contributions of host density and genetic diversity on prevalence of a multi-host parasite in bumblebees. *Biological Journal of the Linnean Society*, 125, 900–910.

Pérez-Pereira, N., Caballero, A. & García-Dorado, A. (2022) Reviewing the consequences of genetic purging on the success of rescue programs. *Conservation Genetics*, 23, 1–17.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. & Hoekstra, H.E. (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, 7(5), e37135.

Polce, C., Maes, J., Rotllan-Puig, X., Michez, D., Castro, L., Cederberg, B. et al. (2018) Distribution of bumblebees across Europe. *One Ecosystem*, 3, e28143.

Powney, G.D., Carvell, C., Edwards, M., Morris, R.K.A., Roy, H.E., Woodcock, B.A. et al. (2019) Widespread losses of pollinating insects in Britain. *Nature Communications*, 10, 1018.

Rademacher, N., Wrobel, T.J., Rossoni, A.W., Kurz, S., Bräutigam, A., Weber, A.P.M. et al. (2017) Transcriptional response of the extremophile red alga Cyanidioschyzon merolae to changes in CO2 concentrations. *Journal of Plant Physiology*, 217, 49–56.

Raj, A., Stephens, M. & Pritchard, J.K. (2014) fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, 197(2), 573–589.

Ralls, K., Ballou, J.D., Dudash, M., Eldridgew, M.D.B., Fenster, C.B., Lacy, R. C. et al. (2018) Call for a paradigm shift in the genetic management of fragmented populations. *Conservation Letters*, 11(2), 1–6.

Raymond, M. & Rousset, F. (1995) An exact test for population differentiation. *Evolution*, 49(6), 1280–1283.

Rasmont, P., Franzén, M., Lecocq, T., Harpke, A., Roberts, S.P., Biesmeijer, J.C. et al. (2015) Climatic risk and distribution atlas of European bumblebees. *BioRisk*, 70-71(90–91), 106–107.

Razgour, O., Taggart, B., Juste, J., Ibáñez, C., Puechmaille, S.J., Novella-Fernandez, R.N. et al. (2019) Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *PNAS*, 116(21), 10418–10423.

Richardson, L., McFarland, K., Zahendra, S. & Hardy, S. (2019) Bumble bee (*Bombus*) distribution and diversity in Vermont, USA: a century of change. *Journal of Insect Conservation*, 23(1), 45–62.

Rigby, R.A. & Stasinopoulos, D.M. (2005) Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3), 507–554.

Robinson, J.A., Bernard, C.B., Kim, Y., Lohmueller, K.E. & Wayne, R.K. (2018) Purging of strongly deleterious mutations explains long-term persistence and absence of inbreeding depression in Island foxes. *Current Biology*, 28, 3487–3494.

Rochette, N.C., Rivera-Colón, A.G. & Catchen, J.M. (2019) Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28(21), 4737–4754.

Roger, N., Michez, D., Wattiez, R., Sheridan, C. & Vanderplanck, M. (2017) Diet effects on bumblebee health. *Journal of Insect Physiology*, 96, 128–133.

Rohlf, F.J. (1972) An empirical comparison of three ordination techniques in numerical taxonomy. *Systematic Zoology*, 21(3), 271–280.

Roycroft, E., MacDonald, A.J., Moritz, C., Moussali, A., Miguez, R.P. & Rowe, K.C. (2021) Museum genomics reveals the rapid decline and extinction of Australian rodents since European settlement. *PNAS*, 118(27), e2021390118.

Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Forteliust, W. & Hanski, I. (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature*, 392, 491–494.

Salamin, N., Wüest, R.O., Lavergne, S., Thuiller, W. & Pearman, P.B. (2010) Assessing rapid evolution in a changing environment. *Trends in Ecology & Evolution*, 25(12), 692–698.

Sánchez-Bavo, F. & Wyckhuys, K.A. (2019) Worldwide decline of the entofauna: a review of its drivers. *Biological Conservation*, 232, 8–27.

Schmid-Hempel, P. (2005) Evolutionary ecology of insect immune defenses. *Annual Review of Entomology*, 50, 529–551.

Schmid-Hempel, P. & Tognazzo, M. (2010) Molecular divergence defines two distinct lineages of *Crithidia bombi* (Trypanosomatidae), parasites of bumblebees. *Journal of Eukaryotic Microbiology*, 57(4), 337–345.

Shafer, A.B., Peart, C.R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C.W. et al. (2017) Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, 8(8), 907–917.

Sladen, F.W.L. (1912) *The Humble-bee: its life-history and how to domesticate it*. London: Macmillan and Co. Ltd.

Simmons, B.I., Balmford, A., Bladon, A.J., Christie, A.P., De Palma, A., Dicks, L.V. et al. (2019) Worldwide insect declines: an important message, but interpret with caution. *Ecology & Evolution*, 9(7), 3678–3680.

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Speight, M.C. & MCD, S. (1974) *Bombus lapponicus*, *Parasyrphus lineola* and *Phaonia exoleta*: insects new to Ireland. *Irish Naturalists' Journal*, 18, 123–124.

Spielman, D., Brook, B.W., Briscoe, D.A. & Frankham, R. (2004) Does inbreeding and loss of genetic diversity decrease disease resistance? *Conservation Genetics*, 5(4), 439–448.

Spielman, D., Brook, B.W. & Frankham, R. (2004) Most species are not driven to extinction before genetic factors impact them. *PNAS*, 101(42), 15261–15264.

Stolle, E., Wilfert, L., Schmid-Hempel, R., Schmid-Hempel, P., Kube, M., Reinhardt, R. et al. (2011) A second generation genetic map of the bumblebee *Bombus terrestris* (Linnaeus, 1758) reveals slow genome and chromosome evolution in the *Apidae*. *BMC Genomics*, 12(1), 1–17.

Storey, J.D., Bass, A.J., Dabney, A. & Robinson, D. (2015) qvalue: Q-value estimation for false discovery rate control. *R package version*, 2(0), 10–18129.

Storey, J.D. & Tibshirani, R. (2003) Statistical significance for genome wide studies. *PNAS*, 100(16), 9440–9445.

Sun, C., Huang, J., Wang, Y., Zhao, X., Su, L., Thomas, G.W. et al. (2021) Genus-wide characterization of bumblebee genomes provides insights into their evolution and variation in ecological and behavioral traits. *Molecular Biology and Evolution*, 38(2), 486–501.

Teixeira, J.C. & Huber, C.D. (2021) The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences*, 118(10), e2015096118.

Tiffin, P. & Ross-Ibarra, J. (2014) Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution*, 29, 673–680.

Vatsiou, A.I., Bazin, E. & Gaggiotti, O.E. (2016) Detection of selective sweeps in structured populations: a comparison of recent methods. *Molecular Ecology*, 25(1), 89–103.

Vray, S., Rollin, O., Rasmont, P., Dufrêne, M., Michez, D. & Dendoncker, N. (2019) A century of local changes in bumblebee communities and landscape composition in Belgium. *Journal of Insect Conservation*, 23(3), 489–501.

Wang, P., Burley, J.T., Liu, Y., Chang, J., Chen, D., Lu, Q. et al. (2021) Genomic consequences of long-term population decline in brown eared pheasant. *Molecular Biology and Evolution*, 38(1), 263–273.

Wang, J., Santiago, E. & Caballero, A. (2016) Prediction and estimation of effective population size. *Heredity*, 117(4), 193–206.

Waples, R.S. & Do, C. (2010) Linkage disequilibrium estimates of contemporary *Ne* using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3(3), 242–262.

Watterson, G.A. (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7(2), 256–276.

Weir, B. & Cockerham, C. (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, 38(6), 1358–1370.

Whitehorn, P.R., Tinsley, M.C., Brown, M.J., Darvill, B. & Goulson, D. (2011) Genetic diversity, parasite prevalence and immunity in wild bumblebees. *Proceedings of the Royal Society B: Biological Sciences*, 278(1709), 1195–1202.

Whitehorn, P.R., Tinsley, M.C., Brown, M.J., Darvill, B. & Goulson, D. (2014) Genetic diversity and parasite prevalence in two species of bumblebee. *Journal of Insect Conservation*, 18(4), 667–673.

Willi, Y., Griffin, P. & Van Buskirk, J. (2013) Drift load in populations of small size and low density. *Heredity*, 110, 296–302.

Willis, S.C., Hollenbeck, C.M., Puritz, J.B., Gold, J.R. & Portnoy, D.S. (2017) Haplotyping RAD loci: an efficient method to filter paralogs and account for physical linkage. *Molecular Ecology Resources*, 17(5), 955–965.

Winfree, R., Fox, J., W., Williams, N.M., Reilly, J.R. & Cariveau, D.P. (2015) Abundance of common species, not species richness, drives delivery of a real-world ecosystem service. *Ecology Letters*, 18(7), 626–635.

Xie, H.X., Liang, X.X., Chen, Z.C., Li, W.M., Mi, C.R., Li, M. et al. (2022) Ancient demographics determine the effectiveness of genetic purging in endangered lizards. *Molecular Biology and Evolution*, 39(1), msab359.

Yates, A.D., Allen, J., Amode, R.M., Azov, A.G., Barba, M., Becerra, A. et al. (2022) Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. *Nucleic Acids Research*, 50(D1), D996–D1003.

Zayed, A. & Packer, L. (2005) Complementary sex determination substantially increases extinction proneness of haplodiploid populations. *Proceedings of the National Academy of Sciences*, 102(30), 10742–10746.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1: Percentage of samples infected with at least one of the parasite species: *Locustacharus buchneri*, *Apicystis bombi*, *Crithidia bombi* or *Vairimorpha bombi*, for all populations screened.

Table S1: details of the filtering steps performed and number of SNPs/individuals retained at each step

Table S2.1: List of genes within GO-classifications significantly enriched (see Table 6) for molecular functions

Table S2.2: List of genes within GO-classifications significantly enriched (see Table 6) for biological processes

How to cite this article: Huml, J.V., Ellis, J.S., Rustage, S., Brown, M.J.F., Billington, R. & Knight, M.E. (2023) The tragedy of the common? A comparative population genomic study of two bumblebee species. *Insect Conservation and Diversity*, 1–20. Available from: https://doi.org/10.1111/icad.12626

17524598, 0, Downloaded from https://resjournals.onlinelibrary.wiley.com/doi/10.1111/icad.12626 by Plymouth University, Wiley Online Library on [01/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License