


Limited introgression from non-native commercial strains and signatures of adaptation in the key pollinator *Bombus terrestris*

Paolo Franchini¹  | Carmelo Fruciano^{2,3} | Thomas J. Wood^{4,5} | Vivaswat Shastry⁶ |
Dave Goulson⁴ | William O. H. Hughes⁴ | Julia C. Jones^{4,7}

¹Department of Ecological and Biological Sciences, University of Tuscia, Viale dell'Università s.n.c, Viterbo, Italy

²Institute for Marine Biological Resources and Biotechnology, National Research Council (IRBIM-CNR), Messina, Italy

³NBFC, National Biodiversity Future Center, Palermo, Italy

⁴School of Life Sciences, University of Sussex, Brighton, UK

⁵Laboratory of Zoology, Research Institute for Biosciences, University of Mons, Mons, Belgium

⁶Committee on Genetics, Genomics and Systems Biology, University of Chicago, Chicago, Illinois, USA

⁷School of Biology and Environmental Science, University College Dublin, Dublin, Ireland

Correspondence

Paolo Franchini, Department of Ecological and Biological Sciences, University of Tuscia, Viale dell'Università s.n.c, 01100 Viterbo, Italy.

Email: paolo.franchini@unitus.it

Julia C. Jones, School of Biology and Environmental Science, University College Dublin, Dublin 4, Ireland.

Email: julia.jones@ucd.ie

Present address

Carmelo Fruciano, Department of Biological, Geological and Environmental Sciences, University of Catania, Catania, Italy

Funding information

Marie Curie Fellowship

Handling Editor: Christian Schlötterer

Abstract

Insect pollination is fundamental for natural ecosystems and agricultural crops. The bumblebee species *Bombus terrestris* has become a popular choice for commercial crop pollination worldwide due to its effectiveness and ease of mass rearing. Bumblebee colonies are mass produced for the pollination of more than 20 crops and imported into over 50 countries including countries outside their native ranges, and the risk of invasion by commercial non-native bumblebees is considered an emerging issue for global conservation and biological diversity. Here, we use genome-wide data from seven wild populations close to and far from farms using commercial colonies, as well as commercial populations, to investigate the implications of utilizing commercial bumblebee subspecies in the UK. We find evidence for generally low levels of introgression between commercial and wild bees, with higher admixture proportions in the bees occurring close to farms. We identify genomic regions putatively involved in local and global adaptation, and genes in locally adaptive regions were found to be enriched for functions related to taste receptor activity, oxidoreductase activity, fatty acid and lipid biosynthetic processes. Despite more than 30 years of bumblebee colony importation into the UK, we observe low impact on the genetic integrity of local *B. terrestris* populations, but we highlight that even limited introgression might negatively affect locally adapted populations.

KEYWORDS

adaptation, *Bombus terrestris*, bumblebees, hybridization, introgression, pollinators

1 | INTRODUCTION

Insect pollinators play a vital role in sustaining natural ecosystems and in the productivity of agricultural crops (Biesmeijer et al., 2006; Gill et al., 2016; Klein et al., 2007; Potts et al., 2010). Evidence of widespread declines in pollinators in recent decades has been attributed to a combination of several human-mediated changes in the environment at local and regional scales, including habitat loss, intensification of agriculture and pesticide use and the introduction of parasites and pathogens (Goulson et al., 2015; Ollerton et al., 2014; Powney et al., 2019). In addition, the use of managed and commercial bee colonies, including bumblebee colonies, for crop pollination and honey production has promoted the transfer of diseases (Goulson, 2010), and facilitated the spread of non-native species and subspecies, posing a risk to endemic species (Jaffe et al., 2016; Lecocq et al., 2016; Seabra et al., 2019). Declines in wild bumblebee species, in particular, have been highlighted (Cameron & Sadd, 2020; Goulson et al., 2015; Grixti et al., 2009; Potts et al., 2010; Soroye et al., 2020; Williams & Osborne, 2009), and bumblebee colonies are mass produced for the pollination of more than 20 crops and are imported into over 50 countries, including those outside the native ranges of reared bumblebee species (Velthuis & van Doorn, 2006).

The successful deployment of commercially reared bumblebees around the world, particularly *Bombus terrestris*, is largely due to the development of effective mass-rearing techniques (Velthuis & van Doorn, 2006), and the ability of these bees to efficiently pollinate crops such as tomatoes, strawberries and melons through 'buzz pollination', where the bees release pollen from flowers by shaking the anthers at high frequency (Buchmann, 1983; Nayak et al., 2020; Wahengbam et al., 2019). Tomatoes, *Solanum lycopersicum*, which are cultivated on a large scale and pollinated by bumblebees worldwide, have an annual economic export value of 10.06 billion USD (Cooley & Vallejo-Marin, 2021). However, the risk of invasion by commercial non-native bumblebees is considered one of 15 emerging issues for global conservation and biological diversity (Sutherland et al., 2017). Commercial bees can potentially escape from greenhouses and successfully establish new populations and/or hybridize with endemic species or subspecies. Such invasions may drive changes in plant-pollinator interactions and local bumblebee population diversity, as well as result in the spread of new pathogens (reviewed in Mallinger et al., 2017). *Bombus terrestris* is considered to have high biological invasion potential because these bees are ecologically flexible, emerge early in the season and have highly polylectic foraging habits (Dafni & Shmida, 1996; Hingston et al., 2002; Hingston & McQuillan, 1998; Matsumura et al., 2004). The native range of the different subspecies of *B. terrestris* comprises Europe, the Middle East and parts of North Africa (Rasmont et al., 2008). International trade of commercially reared *B. terrestris* colonies is reported to be the main cause of invasions into several countries including Chile, China, Israel, Japan, Mexico, South Africa, South Korea and Taiwan (Acosta et al., 2016; Amots Dafni et al., 2010; Hingston et al., 2002; Howlett & Donovan, 2010; Inari et al., 2005; Matsumura et al., 2004; Schmid-Hempel et al., 2014; Velthuis, 2002; Velthuis & van Doorn, 2006),

and *B. terrestris* is also an invasive species in New Zealand and Tasmania (Australia) (Lye et al., 2011; Semmens et al., 1993). Invasions of *B. terrestris* are thus considered significant at a global scale, and this may be especially due to the rate of spread, which has been up to 200 km per year in Chile and Argentina (Acosta et al., 2016; Dafni et al., 2010; Schmid-Hempel et al., 2014).

The importation of commercial *B. terrestris* into regions where conspecific populations occur has also been highlighted as concerning but has been less investigated (Ings et al., 2006; Lecocq et al., 2016). *Bombus terrestris* includes nine subspecies that are classified by their body hair colour patterns, semiochemical profile, genetic composition and distributional ranges (Cejas et al., 2018; Rasmont et al., 2008), with two of the most widely used subspecies in artificial rearing being *Bombus terrestris dalmatinus* Dalla Torre and *Bombus terrestris terrestris* (Linnaeus) (Velthuis & van Doorn, 2006). A key example concerns the UK where the endemic subspecies is *B. t. audax* (Britain, Ireland and associated smaller islands), but commercially produced *B. t. terrestris* and *B. t. dalmatinus* have been imported for pollination (from the 1980s until 2015), particularly in central and southern England (Natural England). In such commercially bred species, artificial selection and hybridization may cause changes to life-history traits and ecological niche characteristics, and potentially increase their invasive potential (Facon et al., 2011; Lecocq et al., 2016). Furthermore, if invasion leads to the establishment of populations of commercially bred bumblebee subspecies, and/or hybridization between these bees and native subspecies occurs, a multitude of key biological traits in native populations may be affected, with impacts on their genetic diversity, ecological adaptation, as well as key plant-pollinator interactions. However, the effects of commercially bred subspecies on native pollinator genetic integrity and evolution have been the focus of few studies to date (Cejas et al., 2021; Kardum Hjort et al., 2022; Seabra et al., 2019).

Hybridization between native and commercially bred subspecies in the UK, and the establishment of populations of the commercial subspecies is plausible. For instance, hybridization between different *B. terrestris* subspecies under laboratory conditions and the production of viable colonies has been demonstrated (De Jonghe, 1986; Ings et al., 2005; Lecocq et al., 2015), and mate choice studies suggest that although mating is not random, *B. t. dalmatinus* can and do mate with *B. t. audax* (Ings et al., 2005). Further, recent studies have detected introgression between commercial and native subspecies in the Iberian peninsula (Bartomeus et al., 2020; Cejas et al., 2018, 2020; Seabra et al., 2019), and between *B. terrestris* lineages in Britain and continental Europe potentially resulting from the import of commercial colonies (Moreira et al., 2015). In addition, reports of winter activity have increased in the UK since the early 1990s (Hart et al., 2021; Robertson, 1991; Stelzer et al., 2010). Most northern European *B. terrestris* populations are active during the warmer parts of the year, when queens emerge from diapause in the spring and produce a colony over the summer, and then new queens enter diapause at the end of the summer (Løken, 1973; Rasmont et al., 2008). Expression of winter-active behaviour is usually restricted to populations occurring in the Mediterranean Basin, which includes the

native range of *B. t. dalmatinus*, but also *B. t. terrestris* in southern France and *B. t. lusitanicus* in the Iberian Peninsula. However, it has been suggested that warmer temperatures due to climate change may enable winter activity, and facilitate the successful establishment of commercial *B. t. dalmatinus* colonies in the south of the UK (Owen et al., 2016). On the other hand, introgression was not detected between commercial *B. t. dalmatinus* and wild *B. t. audax* in the UK (Hart et al., 2021), or between commercial and wild *B. impatiens* in New England, USA (Sun et al., 2017), and no evidence was found for recent introgression of commercial *B. terrestris* into local wild conspecific populations in southern Sweden (Kardum Hjort et al., 2022). It may simply be the case that the behaviour is plastic, and that *B. t. audax* can express winter-active behaviour under the right environmental conditions, such as can be seen with the facultative expression of social behaviour in *Halictus* sweat bees (Field et al., 2010).

Here, we use genome-wide SNPs to detect and quantify introgression of commercial *B. terrestris* subspecies into native UK bumblebees, and evaluate the consequences of importing commercially bred non-native bumblebee subspecies into the UK. We test the hypothesis that importing such bees into the UK has facilitated hybridization between commercial and native subspecies, and that *B. terrestris* throughout the island of the UK has become a single genetically homogenous population. We conduct population genomic analyses of natural populations of *B. terrestris* close to (<6 km) and far from (>20 km) farms using commercial colonies for pollination, as well as commercial populations, and test for hybridization. We estimate the genetic diversity of natural and commercial populations and identify private alleles in the wild populations that can be used to help distinguish the native and commercial subspecies of *B. terrestris*. We also identify genomic regions that may be associated with global and local adaptation in this key pollinator.

2 | MATERIALS AND METHODS

2.1 | Sampling

We studied 129 bumblebee specimens (*Bombus terrestris*), including drones (males) and workers (females), from seven natural populations throughout the United Kingdom (Figure 1a), and three populations representative of the commonly used commercial subspecies/strains (Table S1). Sampling was conducted from farms and from wild populations later in the season when drones were flying; therefore, the sampled populations include both drones (haploid males) and workers (diploid females), and differences in ploidy were accounted for in analysis (see below). All bees were collected in 2014 when both native and non-native subspecies were supplied to UK farms by companies that rear bumblebees for pollination. Natural populations were sampled from sites close to (<6 km) and far from (>20 km) farms using commercial colonies (data on farms and commercial colonies reported by farmers and collected by Natural England, Welsh Nature Policy and Science and Advice for Scottish Agriculture departments), and therefore deemed more and less likely, respectively, to be interacting with bees from commercial strains (Table S1). *Bombus terrestris* reportedly has a foraging range of up to 2 km (Walther-Hellwig & Frankl, 2000), with queen dispersal estimated as 3–5 km (estimated for other species only to date – *B. pascuorum* and *B. lapidarius* (Lepais, 2010)), and with males able to disperse up to 9.9 km (Kraus et al., 2009). With the permission of the farmers, commercial populations were sampled on site at two large tomato farms, using multiple commercial colonies designated as *B. t. terrestris*-like (non-native subspecies) and *B. t. audax*-like (native subspecies) respectively. Samples of *B. t. dalmatinus*, the third commonly commercially bred and supplied subspecies, were supplied as museum samples originally collected in Turkey, the native region of this subspecies.

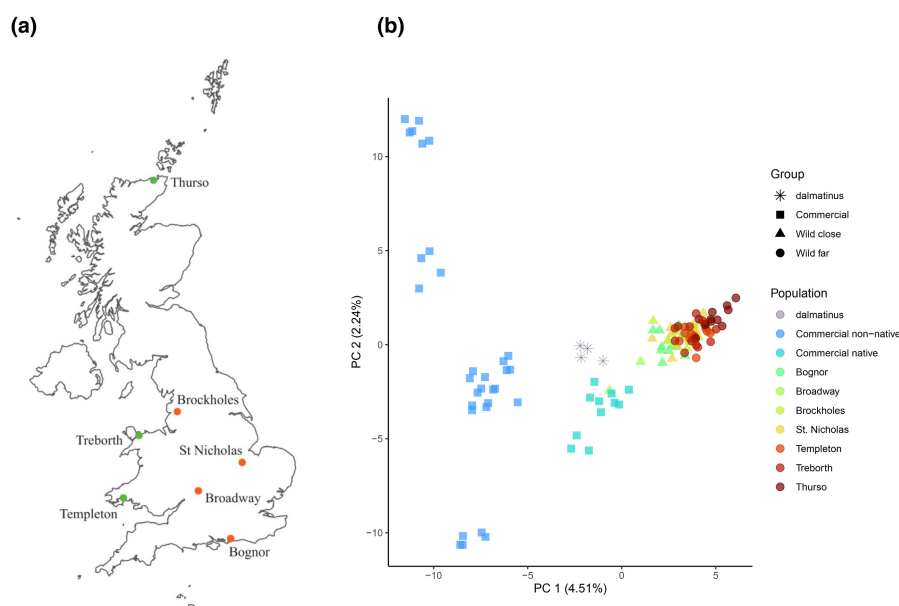


FIGURE 1 (a) Map of UK showing the sampling localities of the wild-caught bees used in this study. Orange and green dots indicate locations close to and far from farms using commercial colonies respectively. (b) The two main axes of genetic variation of a PCA.

2.2 | Sequencing, mapping and genotyping

Genomic DNA was extracted from the whole body of each bee stored in ethanol and refrigerated using a Qiagen DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's protocol. Gel electrophoresis and a Qubit v4.0 fluorometer (Life Technologies) were used to assess the integrity and quantify the DNA of each sample respectively. The species identity of all samples was confirmed by sequencing the mtDNA COI region and aligning sequences against entries in the NCBI database using BLAST tools.

Approximately 100 ng of DNA per sample was used to construct double-digest restriction site-associated DNA (ddRAD) using the quaddRAD protocol (Franchini et al., 2017) following the modifications introduced in Franchini et al. (2018). Double digestion was carried out with a rare cutter, PstI, and a frequent cutter, MspI, restriction enzyme. A single quaddRAD library including 129 barcoded individuals was size-selected from 450 to 550 bp in a Pippin Prep system (Sage Science) and paired-end sequenced (2 × 150 bp) in an Illumina HiSeq 2500 platform at the genome facility of Tufts University (TUCF Genomics).

A total of 117.67 million (M) raw reads were processed by the *clone_filter* and *process_radtags* scripts implemented in the Stacks v2.59 (Catchen et al., 2011) package in order to remove PCR duplicates (*clone_filter* option: "--inline_inline") and demultiplex individuals while discarding low-quality reads (*process_radtags* options: "-c -q"). A final data set of 107.89 M paired reads (averaging 0.84 M reads per individual) was aligned to the *Bombus terrestris* genome (Bter_1.0; NCBI GenBank assembly accession: GCA_000214255.1) using the program bwa-mem v0.7.17 (Li & Durbin, 2009) with default parameters. A total of 125 individual mapping files, compressed to BAM format and sorted with Samtools v1.9 (Danecek et al., 2021), were retained after excluding four samples due to exceptionally low number of aligned reads. Downstream variant and genotype calling was performed using GATK v4.3.0.0 (Van der Auwera & O'Connor, 2020) in a three-step procedure: (i) the *HaplotypeCaller* module was used to call variants per-sample in GVCF mode by genotyping each sample based on its ploidy, and thus allowing us to maximize the number of variants across ploidy levels; (ii) per-sample GVCF files produced by *HaplotypeCaller* were combined into a multi-sample GVCF with the module *CombineGVCFs*; (iii) joint genotyping on all individual samples, pre-called with *HaplotypeCaller*, was performed using *GenotypeGVCFs* with the 'all-sites' option in order to output both variant and invariant sites, required to calculate nucleotide diversity (π) and absolute divergence (dxy). To produce a VCF file including only variant sites, the GATK output in standard VCF format was handled and filtered with VCFtools v0.1.15 (Danecek et al., 2011) by removing short indels ('--remove-indels'), keeping only biallelic SNP loci ('--min-alleles 2 --max-alleles 2') with a mapping quality (MQ) > 30 ('--minQ 30') and minimum read depth of 5 ('--minDP 5'). Finally, sites with a minimum allele count of 3 ('--mac 3') and genotyped in at least 50% of the individuals in each population were retained. This procedure resulted in a highly complete matrix including 8953 variable sites

with only 3.47% missing genotypes across all samples (missing genotypes per population ranging from 2.03% to 5.54%). To reduce the effect of non-independence of markers due to physical linkage and linkage disequilibrium (LD), and thus satisfying one of the main assumptions of model-based algorithms used in this study, we generated a new SNP data set by pruning linked variants showing a r^2 (a commonly used measure for LD) greater than 0.15. To this end, we used the program Plink v1.9 (Purcell et al., 2007) with the option '--indep-pairwise 50 10 0.15' and ended up with a data set including 6774 unlinked SNPs. Invariant sites, selected by applying a minimum allele frequency filter ('--max-maf 0'), were processed by removing indels, applying a minimum read depth of 5 and a 50% missing data threshold. Independently filtered variant and invariant sites were then combined in a single VCF file using BCFtools v1.16 (Li, 2011) resulting in a total of 650,772 loci. As a final step, we "haploidized" the diploid individuals by randomly assigning heterozygote alleles as either homozygote reference or homozygote alternative using the "pseudoHaploidize" function included in the correctKin package (Nyerki et al., 2023). This procedure was implemented to create a bias-free SNP data set to be processed by algorithms that do not deal with mixed-ploidy populations. These new genotype data made of real haploids and haploidized diploid individuals were used for all downstream analyses except for the Entropy runs (see details below).

2.3 | Genetic diversity and differentiation

We used the program PIXY v1.2.5 (Korunes & Samuk, 2021) to compute sliding-window estimates of both nucleotide diversity within populations (π), relative divergence between populations (F_{ST}) and absolute divergence between populations (dxy). Sliding-window computations were performed across the genome using non-overlapping sliding windows of 200 kbp. The algorithm implemented in PIXY allows to calculate π and dxy from VCF files encoding both variant and invariant sites, and provides unbiased estimates of these genetic parameters (Korunes & Samuk, 2021) using, for each window, only the actual sequence data in the window for the computation. As PIXY computes per-window point estimates of π and dxy, we also computed bootstrapped estimates for each window in each population by obtaining bootstrapped VCF files (sampling with replacement the individuals in each population; 100 bootstraps) and subjecting the resulting VCF files to separate PIXY runs. This allowed us to obtain a 95% confidence interval for each window of the genome containing sites (for each population, we excluded windows with more than 10% missing data). Using the same PIXY output (which includes number of variant sites and number of comparisons), we also obtained genome-wide 95% confidence intervals for π for each population and dxy for each comparison between populations. We then deemed the windows whose confidence interval did not overlap with the genome-wide 95% confidence interval for a given population/comparison, as either high or low nucleotide diversity/absolute divergence regions.

Similarly, using the R package *hierfstat* v0.5-11 (Goudet, 2005), we used bootstrapping to obtain 95% confidence intervals for F_{ST} both genome-wide among populations and for each of the same windows used in the analysis of nucleotide diversity. As this analysis was performed bootstrapping *loci*, we restricted this analysis to combinations of comparison and windows containing at least 5 SNPs. We designated 'high F_{ST} windows' for a given comparison between two populations (e.g. Bognor vs. Thurso), those windows whose confidence interval entirely exceeded the genome-wide 99% confidence interval for that specific comparison. These windows can be interpreted as 'high F_{ST} ' in the sense that they exceed the level of divergence normally observed between two given populations. The location of 'high F_{ST} ' windows can be checked against other features (e.g. locations of high genetic diversity) to uncover putative local or global adaptation (Booker et al., 2021); therefore, for this analysis, we focused on wild UK bees and chromosomal regions only. For instance, patterns of π and dxy in high F_{ST} regions may reveal whether local or global adaptation is the most likely process giving rise to high F_{ST} (Booker et al., 2021). Specifically, following Booker et al. (2021), among high F_{ST} regions, those which had low π (for both populations) and dxy were deemed as putatively under global adaptation, and those which had high dxy were deemed as putatively under local adaptation. As local adaptation can produce both regions of high and low nucleotide diversity (Russ & Jasper, 2020), we did not consider π to identify regions under local adaptation.

Genomic regions found to contain adaptive *loci* were then screened for protein-coding genes that were subjected to gene function (based on Gene Ontology, GO) and pathways enrichment analyses (using the Kyoto Encyclopedia of Genes and Genomes, KEGG, database). To this end, we first intersected the coordinates of the genomic windows under adaptation with those of the *Bter_1.0* annotation using the *intersect* module of the BEDtools v2.31.0 package (Quinlan & Hall, 2010) and extracted the genes included in the intersected regions using Bash commands. Second, these genes (test sets) were compared with the whole *B. terrestris* genes (baseline set, source: Ensembl Genomes Metazoa release 54) to identify significantly overrepresented GO terms and KEGG pathways. To test for significance, a Fisher's exact test implemented in *g:Profiler* v1.07.0 (eg54_p17_bf42210 (Raudvere et al., 2019) was used (*g:SCS* multiple testing correction method, significance threshold of .05).

2.4 | Overall population structure

To assess the degree of overall population structure and genetic differentiation in our samples, we applied a combination of different approaches. First, we obtained point estimates of global F_{ST} for each population comparison using *hierfstat* and considered a given comparison between populations significant if the 95% confidence interval obtained through bootstrap (see above) did not include zero. Then, a principal component analysis (PCA) was performed in R based on the genotype data of the haploidized and pruned data

set after mean substitution of missing data and centring. Scatterplots of scores along the first few principal components can be highly misleading particularly when these account for a small proportion of total variation (Bookstein, 2017; Jombart et al., 2009; Lever et al., 2017). For this reason, we performed additional tests to check whether patterns of dispersion observed along the first two principal components were also observed when analysing all principal components. To achieve this, we used two approaches typically employed in morphometrics and ecology. Using the R package *GeometricMorphometricsMix* (Fruciano, 2018), we used bootstrap resampling to obtain population-specific 95% confidence intervals for multivariate variance (the variance across all the principal components). In this approach, which is commonly used in morphometrics (e.g. Fruciano et al., 2014), if a given population is more variable in PC space than the others, we expect its confidence interval to be disjunct from the others. The second approach involved the PERMDISP2 procedure (Anderson, 2006), as implemented in the R package *vegan* (Oksanen et al., 2016), using Euclidean distances and the median distance from the centroid of each population as a measure of dispersion. An analysis of variance (ANOVA) on these measures of dispersion was then employed to test whether there was at least one population with larger dispersion.

To investigate the possible presence of genetic clustering in the whole data set, we used the widely used model-based algorithms implemented in *Admixture* v1.23 (Alexander et al., 2009). *Admixture* was run with default parameters and allowing the number of predefined genetic clusters (K) to range from 1 to 10. Statistical support for the different number of inferred clusters was assessed using the cross-validation function included in the package. To further assess the genetic relationships among samples, we used *RADpainter* v 0.3.2_r109 (Malinsky et al., 2018) which, differently from widely used clustering approaches (e.g. *Admixture*), infers the ancestry of genomic segments in a population. The program then sums the coancestry values across all *loci* and generates a coancestry matrix for the full data set.

To ensure an accurate estimation of the genotype and ancestry parameters in our data set composed of mixed-ploidy (haploid and diploid) bumblebees, we used the hierarchical Bayesian model implemented in the program *Entropy* v2.0 (Shastry et al., 2021). This program is appositely devised to deal with data sets that include samples with different ploidy levels. *Entropy* was run on the genotype likelihoods encoded as "PL" field in the GATK VCF file that were first converted into a readable format (MPGL) using the *vcfR* v1.13.0 R package (Knaus & Grunwald, 2017), for 100,000 steps with 50,000 burn-in and stored every 10th step in memory. For each K model (putative number of demes) we ran three independent chains and assessed convergence using the *assessconvergence.R* script provided in the package. The number of K that best fits the data was evaluated using the widely applicable information criterion (WAIC) (Watanabe, 2010), and estimates were extracted from the raw *Entropy* HDF5 output files using the *estpost.entropy* script. The admixture proportions (Q scores) for all individuals and all K values in a stacked barplot fashion were drawn using the *plotadmix.R* scrip.

Finally, to visualize genotypic relationships among individuals, we used the NetViewR v2.1.0 program (Steinig et al., 2016) at multiple k-NN values (k-NN = 10–60). Optimization of k-NN values was performed by plotting each k-NN value against the number of communities detected using a “Fast-greedy”, an “Infomap” and a “Walk-trap” clustering algorithm. Networks were constructed at increasing values of k-NN (20, 40 and 60) allowing us to detect both fine- and large-scale genetic structure. As input data for NetviewR, we used a shared-allele distance matrix (1-IBS) that was generated by Plink from the full set of filtered variants in VCF format.

We used the VCF-contrast module of the VCFtools package to detect any private alleles in the wild bees. Target (wild) and background (commercial bees plus *B. t. dalmatinus*) samples were specified using the +<list> and -<list> options respectively. Major and minor allele frequency for each private allele was calculated using VCFtools with the ‘freq’ option. Raw numbers of private alleles per population were normalized to account for differences in population size.

2.5 | Overlap of genomic regions of interest

To test for the significance of the overlap of genomic regions of interest, including regions of high and low nucleotide diversity and putative local and global adaptation, we used the permutational procedure implemented in the package regioneR v1.30.0 (Gel et al., 2016) using the number of overlaps as test statistic and 100 permutations. This procedure performs permutation tests by randomizing genomic regions, therefore controlling for chromosome size and distribution of regions in the genome. Importantly, for comparisons of regions of interest, we restricted the analysis to chromosomal regions in the reference genome only (i.e. we excluded regions in unplaced scaffolds). We also masked those regions (i.e. windows) which did not contain sequence data in our data set – which is crucial because not masking those regions would result in elevated type I errors (i.e. finding erroneous ‘significant overlaps’).

The first tests of overlap we ran were pairwise among populations testing for the overlap of, respectively, low nucleotide diversity and high nucleotide diversity regions, reflecting the hypothesis that – across wild populations and between commercial and wild populations – the same regions of the genome tend to harbour genetic diversity or show reductions in genetic diversity. Further, we tested for the overlap of those regions identified in the wild samples as under global adaptation, and windows of low nucleotide diversity for the commercial samples. This test reflects the question of whether regions putatively under global adaptation in the UK (characterized, among other things, by low π) correspond to regions of low genetic diversity specific to the UK (in which case we expect a non-significant overlap with commercial samples). Vice versa, a significant overlap would indicate that regions putatively under global adaptation may be either false positives or reflect global adaptation at a scale broader than the UK. Finally, we tested the overlap for wild samples of regions showing signals of local adaptation, with the

location of SNPs containing private alleles for the wild samples. The rationale for this test is that in regions under local adaptation we expect to find private alleles more frequently (Sjöstrand et al., 2014).

3 | RESULTS

3.1 | Sequencing, mapping and genotyping

A total of 125 individuals were kept after excluding those with low number of reads aligned to the reference bumblebee genome (average number of paired-reads per individual that aligned concordantly: 700,139; SD: 169,189). After genotype calling and filtering according to the criteria detailed above (see Section 2), a matrix of 650,772 SNP loci (including 8953 variants, of which 6774 had reduced linkage between nearby sites) was retained for downstream genomic analyses.

3.2 | Genetic diversity and differentiation

An exploratory PCA (Figure 1b) identified distinguishable groups formed by wild and commercial individuals, with clearly distinct scores along the first axis of variation (PC1). The commercially reared non-native (*B. t. terrestris*-like) individuals were the most divergent, while the commercial native (*B. t. audax*-like) individuals formed a separate group, equally distant from commercial non-native and the wild bees, these latter clustered tightly together, except for two individuals from Brockholes (a more defined clustering that mirrors their sampling locality can be observed in the PCA conducted on the wild samples only; see Figure S1). The four *B. t. dalmatinus* bees were also located between the commercial non-native and the wild samples in the scatterplot of the scores along the first two principal components; however, the commercial native specimens had distinct scores along the second principal component (PC2). In addition to variation in location among groups, the scatterplot of the scores along the first two principal components showed substantial variation in dispersion among populations. In particular, the commercially reared non-native population appeared approximately as dispersed as the variation among all other samples. However, an investigation of dispersion using all principal components revealed a much lower degree of dispersion, with the commercial non-native population showing a tendency for higher dispersion, but never to the extent one would have expected from the scatterplot of the scores along the first two principal components. Indeed, confidence intervals for multivariate variance of each population are never fully disjunct from each other (Figure S2, Table S2). A similar result is obtained with the procedure based on Euclidean distances in which a permutational-based analysis rejected the null hypothesis that all populations have the same dispersion and the pattern was mainly driven by the commercial non-native samples having higher dispersion but not to the extent implied by the scatterplot along the first two principal components (Figure S3, Table S3).

The analysis of intra-population genetic diversity revealed similar patterns to the ones observed in dispersion in the full PCA space, with no single population having its confidence interval for genome-wide genetic diversity disjunct from the ones from other populations (Figure S4). However, the confidence interval for the commercial non-native sample only overlapped with the commercial native sample and *dalmatinus* (Figure S4), further confirming the idea that this population had a tendency for higher genetic variability (note how *dalmatinus* has always very broad confidence intervals due to the low sample size). The sliding window analysis of nucleotide diversity was based, considering all individuals in the analysis, on an average of 566.77 (SD 381.69) base pairs per window, which increased to 622.3 (SD 387.4) restricting the computation to windows mapping to chromosomes in the reference genome. This analysis revealed a large number of windows deemed to have either high or low genetic diversity, with low genetic diversity windows noticeably more frequent than high genetic diversity windows in all populations (Figure S5). Among all samples, the commercial non-native sample had the largest number of windows – both high and low diversity (Figure S5).

3.3 | Overall population structure

The general population structure pattern highlighted by the PCA was confirmed by the Admixture analysis (Figure 2). When the Bayesian algorithm was set to two genetic clusters, the commercially reared non-native and the wild individuals were assigned to different groups with high assignment scores (admixed individuals were exceptions, and none of these had assignment score (Q) < 0.5 to the correct group of origin). The commercial native and the *B. t. dalmatinus* bees were found to be admixed with similar proportion to the two defined genetic clusters. When further potential genetic structure was evaluated by increasing the number of K ($K=3$), the commercial native bees showed the highest assignment score (Q ranging

from 0.8049 to 1, average $Q=0.9523$) to a new genetic cluster, and several wild individuals were also assigned to the cluster with low Q , with the exception of one bee from Broadway ($Q=0.8359$). Evident substructure in the wild cluster (so identified with $K=2$) arose when K was set to 4, with several wild-caught individuals showing substantial admixture with multiple ancestries. Exploring further potential clustering across the whole sample set failed to clearly identify a match between inferred genetic clusters and sampling locality, and thereby confirmed the low genetic structure across the wild bees in the UK. The Admixture plots corresponding to $K=2-10$ are provided in the Figure S6 along with the results of the cross-validation techniques that suggested that the numbers of clusters with highest statistical support are $K=2-3$ (Figure S7).

The NetView results support the results of the previously applied population clustering methods and provide deeper insight into the interconnectedness within and between bumblebee populations. Setting K -NN to 20, 30 and 40, values that have been shown to be effective in detecting both large- and fine-scale genetic structures (Steinig et al., 2016), allowed us to distinguish groups including the commercial native, the non-native and the wild bees. Among the wild bees, individuals that tended to show more similarity to the commercially reared bees (i.e. higher number of connections) were highlighted (Figure 3).

Finally, F_{ST} statistics were congruent with the patterns of population structure revealed by the clustering approaches. Generally, we found low pairwise levels of relative genetic differentiation across the whole data set (F_{ST} ranging from 0.0002 to 0.13), with the highest F_{ST} values shown by the commercial populations. In the pairwise comparisons among the wild samples, Thurso showed the highest level of genetic divergence with the other populations. Similar patterns were detected using an absolute measure of genetic divergence (dx). All genome-wide pairwise F_{ST} values are reported in Table S4.

Across all pairwise comparisons among the wild samples, we found a total of 67 windows deemed as 'high F_{ST} ' windows (Table S5),

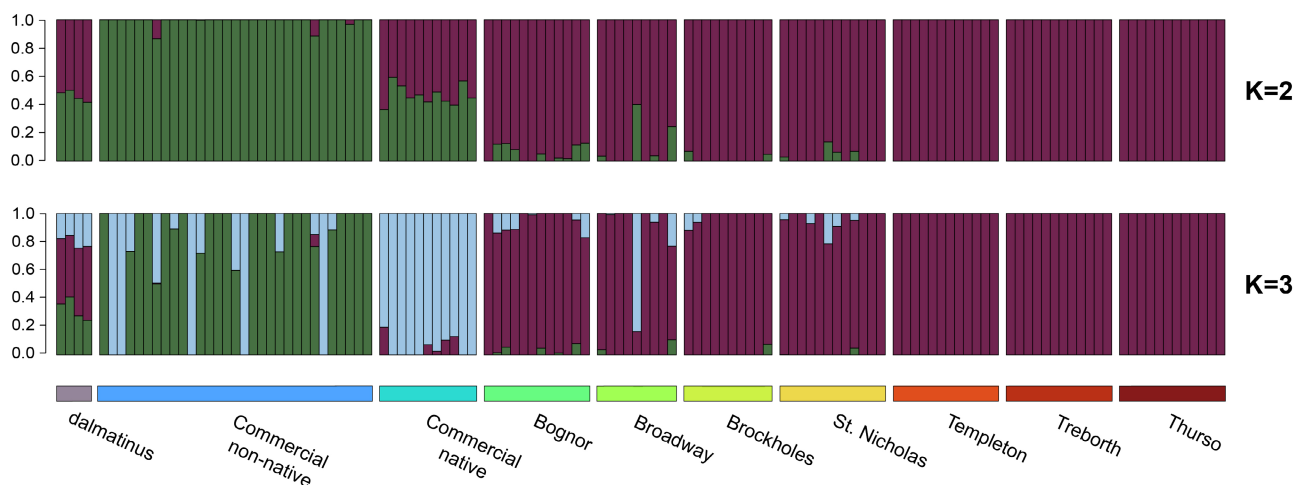


FIGURE 2 Genetic clustering and individual ancestry at $K=2$ and 3, the most supported numbers of clusters in an Admixture analysis (see cross-validation error plot in Figure S7) that revealed evidence of recent limited admixture between wild and commercial populations.

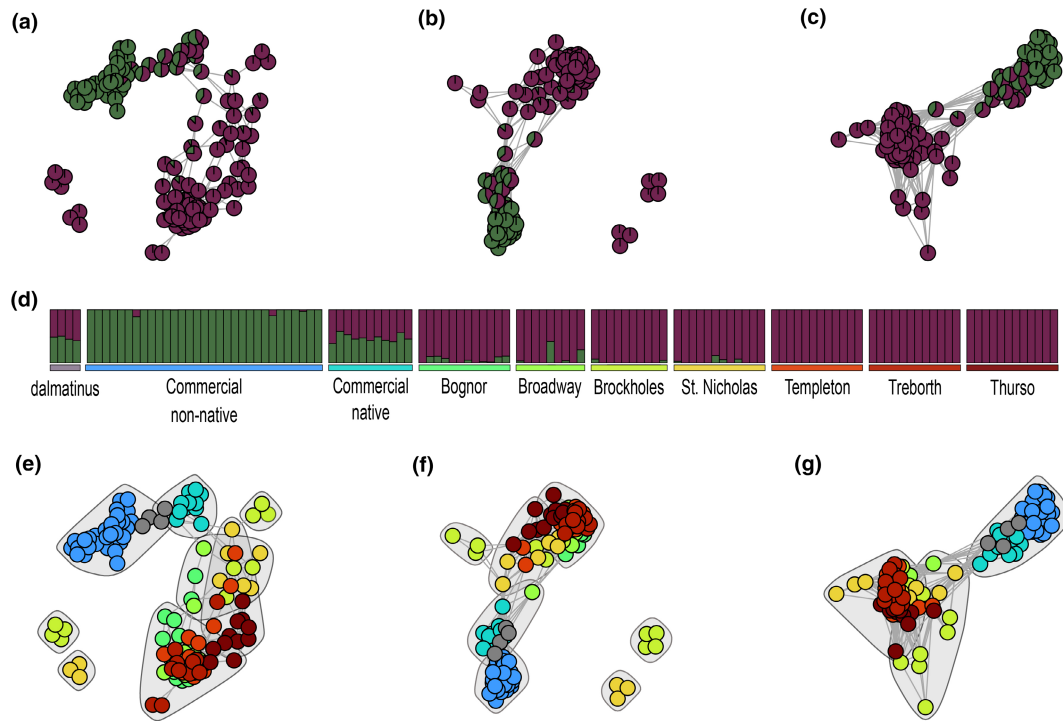


FIGURE 3 NetView networks constructed at k-NN values of 20 (a and e), 40 (b and f) and 60 (c and g). Individuals are coloured based on sampling locality (e, f and g) and according to the Admixture assignment proportions at $K=2$ (a, b and c). Admixture plot (d) is the same as Figure 2.

corresponding to 45 unique windows in the bumblebee genome (we note that the same window can be 'high F_{ST} ' in multiple comparisons). Across the 67 windows deemed as 'high F_{ST} ', 40 are also considered 'high genetic diversity' in at least one of the two populations making up each comparison, four are also considered 'low genetic diversity' in at least one of the two populations in the comparison, and 24 are neither low nor high genetic diversity for both the two populations making up each comparison (Table S5). Out of the 67 high F_{ST} windows, two (1 unique) are also low d_{xy} and low π in both populations being compared (regions under putative global adaptation). Conversely, 50 windows (31 unique) had high d_{xy} and were deemed as regions under putative local adaptation. Notably, 39 of these 50 windows (25 unique) were also 'high π ' in at least one population being compared.

In terms of overlap for regions with high or low nucleotide diversity, we found contrasting results. In the case of regions of high nucleotide diversity, there was always a significant ($p < .01$; Table S6) overlap in regions between pairs of populations (i.e. different populations tend to share the same regions of high nucleotide diversity). Even though all tests were significant, the Z scores tended to be higher when testing the overlap between wild populations, and lower in comparisons involving the commercial samples (Table S6). The Z scores in this analysis are the number of standard deviations between the observed number of overlaps and the average number of overlaps obtained under the null hypothesis that overlaps only occur by chance. Thus, a higher Z score in a given comparison can be interpreted as stronger evidence for overlap and a stronger

departure from overlap just due to chance. Conversely, tests of overlaps for low nucleotide diversity windows were almost always non-significant (Table S6), except for a few comparisons involving the populations of Bognor, Broadway, Templeton and Treborth.

The tests of overlaps comparing high F_{ST} windows under putative global adaptation for wild samples with low π regions in commercial samples were always non-significant (1 overlap with commercial non-native, $p = .49$, Z score 0.64; 1 overlap with commercial native, $p = .3$, Z score 1.28). That is, regions under putative global adaptation in the UK (which are characterized, among other things, by low π) are not typically the same regions which are low π in commercial samples. Finally, the test of overlap comparing windows in wild samples showing putative signals of local adaptation, with the location of SNPs containing private alleles for the wild samples, was also not significant (25 overlaps, $p = -.77$, Z score -0.74).

The genomic regions likely under global and local adaptation contain 21 and 182 protein-coding genes respectively (Tables S7 and S8). While we found no functional enrichment in the first gene set (global adaptation), the genes of the second set (local adaptation) showed enrichment for 12 GO terms related to, among others, taste receptor, oxidoreductase activity, fatty acid and lipid biosynthetic processes, and for one KEGG pathway (biosynthesis of unsaturated fatty acids; Table S9).

Using RADpainter, we aimed to identify genetic divergence and signatures of hybridization among populations. The RADpainter results mirrored the main patterns detected in the results of the other analyses. We find support for a relatedness scenario with two fully

supported clusters: one that includes all the commercially reared native and non-native individuals, the *B. t. dalmatinus* bees, and a single wild individual from the locality of Broadway; a second cluster that includes the wild individuals (Figure 4). In the first main cluster, RADpainter identified two well-resolved sub-clusters: the first grouping all the commercially reared non-native individuals, with high values of coancestry; the second grouping the commercially reared native individuals with the wild individual nested into, and the *B. t. dalmatinus* bees. In the second main cluster, only the individuals from the far north-eastern wild population, Thurso were grouped into a monophyletic clade (high ancestry values), while the remaining wild bees were assigned to several sub-clusters which do not necessarily correspond to the populations of origin (Figure 4).

Similar to what was highlighted by RADpainter, the hierarchical Bayesian model implemented in the program Entropy recovered a clear distinction between the wild and commercially reared non-native samples, while, on average, the commercially reared native individuals and the *B. t. dalmatinus* bees showed higher levels of admixture (Figure 5a). Notably, at $K=2$ all individuals from the wild populations further from farms were assigned to the same genetic

cluster with a very high assignment proportion (Q ranging from 0.9903 to 0.9998, average $Q=0.9984$; Figure 5a,b), while wild individuals closer to farms showed a slightly higher level of admixture with the cluster corresponding to the commercially reared non-native bees (Q ranging from 0.6572 to 0.9998, average $Q=0.9687$). Similar patterns were observed at $K=3$, where individuals from populations closer to the farms showed higher level of admixture with the two genetic clusters formed by commercially reared bees (Figure 5a,c). Further sub-structuring at higher levels of K (4–10) was observed suggesting on the one hand a low level of divergence among the wild populations, and on the other hand a non-homogenous genetic background of the commercial bees, either natives or non-native subspecies (Figure 5; Figure S8).

4 | DISCUSSION

We assessed the implications of utilizing commercially reared non-native and native bumblebee subspecies in the UK for the genetic integrity of their wild native counterparts, and we asked whether

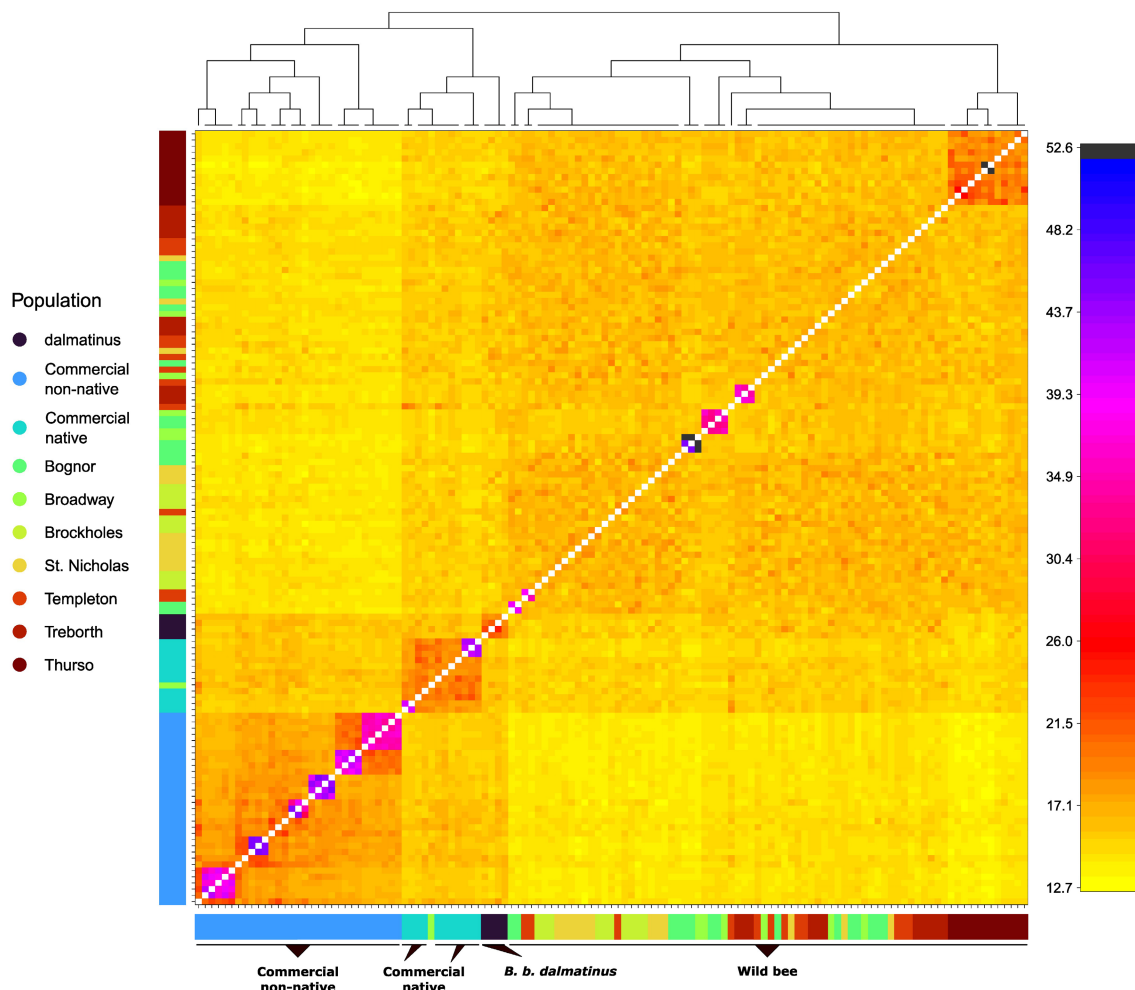


FIGURE 4 Clustered RADpainter coancestry matrix of the 125 individuals genotyped in this study. The darker the square, the higher the genetic similarity between a pair of individuals.

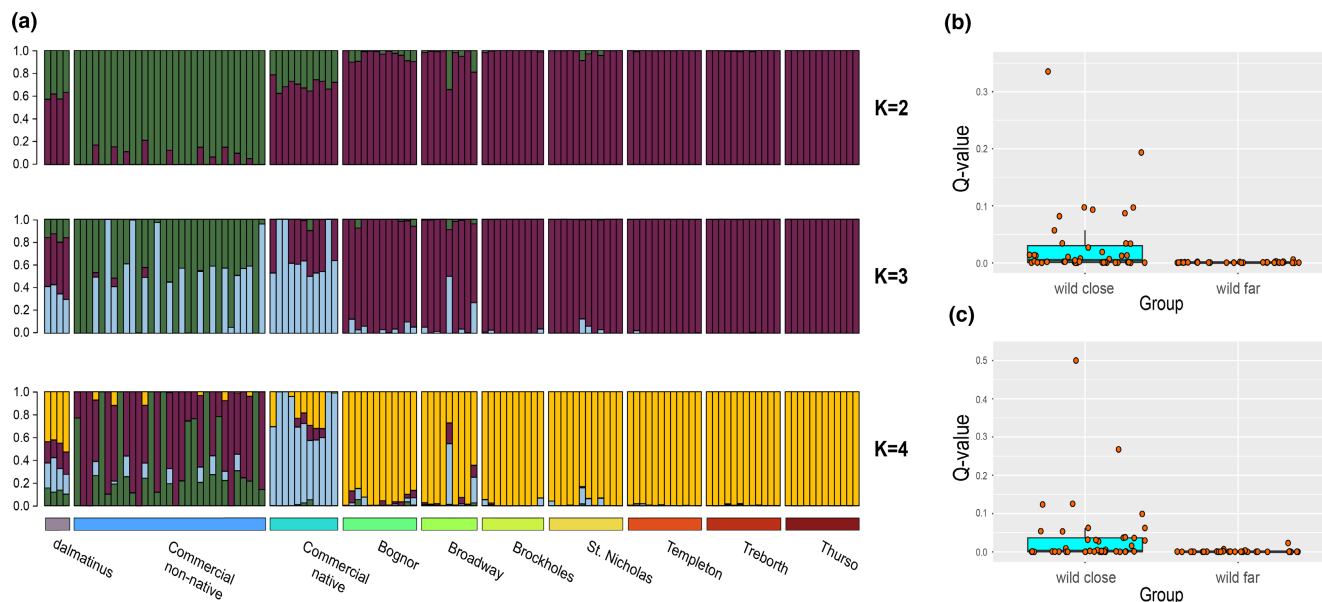


FIGURE 5 (a) Individual coancestry inferred by the hierarchical Bayesian model implemented in the program Entropy for $K=2$, 3 and 4. (b) Assignment scores (Q) to the commercial non-native cluster (inferred from Entropy at $K=2$ and represented by green bars in a) of the wild bees close and far to the farms. (c) Assignment scores (Q) to the commercial native cluster ($K=3$, blue bars in a) of the wild bees close and far to the farms.

these wild bumblebees showed signs of local and global adaptation. This was achieved by investigating the population genomics of seven wild populations of *Bombus terrestris* throughout the UK as well as populations of commercially reared non-native and native subspecies. Minimal genetic structuring was detected in wild populations across the UK, and all populations were found to be genetically distinct from populations of the commercial subspecies, with low levels of introgression detected between commercial and wild populations. Wild bees occurring close to farms that utilized commercial subspecies for pollination were genetically assigned, typically in low proportions, to the commercial subspecies. We detected private alleles in the wild bees that can be used as genomic markers to help distinguish the native and non-native subspecies of *B. terrestris*. In addition, we identified genomic regions potentially involved in adaptation.

Using different lines of investigation into population genetic structuring, we find that overall, wild populations of *B. terrestris* throughout the UK are not genetically distinct from one another, suggesting that there are no real barriers to gene flow. However, one exception is the population from far northern Scotland, Thurso, where individuals are relatively more strongly differentiated from the other wild populations. This suggests that there may be an element of isolation by distance in populations occurring at geographic extremes. The general results found here are mirrored by earlier studies where, for instance, a relatively homogenous populations structure was found across mainland Europe populations, but where high differentiation was detected between islands, such as between Mediterranean islands, and between Ireland and Britain, as well as between the islands and continental Europe (Estoup et al., 1996; Moreira et al., 2015; Widmer et al., 1998).

The wild and commercial populations of *B. terrestris* were found to be genetically differentiated, with generally only low levels of introgression observed between commercially reared populations and wild populations, and limited evidence for established populations of the commercial subspecies being detected. However, even limited introgression is likely associated with long-term risks for the conservation of wild bees. Importantly, these population genomic patterns are evident from several lines of investigation, including Admixture analysis, genotype relationships using NetView, estimates of F_{ST} , genetic ancestry and implementing the hierarchical Bayesian model in Entropy. Interestingly, these patterns of limited introgression are evident despite seemingly ample opportunity for invasion and hybridization of non-native subspecies due to a history of importation of commercial subspecies from the 1980s until 2015 in the UK (Department for the Environment, Food and Rural Affairs, 2014; Goulson, 2010; Moreira et al., 2015). Further, this provides evidence contrary to predictions that *B. terrestris* in the UK is a single homogenous population in which native and imported subspecies have introgressed (including the native *B. terrestris audax* and non-native *B. terrestris dalmatinus* originating from Greece and Turkey; Goulson, 2010). In addition, the lack of established populations of the commercial subspecies suggests that recent observations of winter-active *B. terrestris* in the southern UK (Hart et al., 2021; Robertson, 1991; Stelzer et al., 2010) may instead be native bees that are adapting to increasing temperatures and the availability of non-native garden flowers for foraging.

Where introgression was detected was between the commercial subspecies and wild populations close to agricultural regions where greenhouses utilize commercial colonies. These wild populations include those in the regions of Bognor Regis, Broadway, Brockholes

Nature Preserve and Deeping St Nicholas. Perhaps unsurprisingly, the overall direction of gene flow was found to be from commercial populations towards wild populations, which is similar to the findings of Seabra et al. (2019) in their analysis of introgression between commercial and native *B. terrestris* in the western Iberian Peninsula. Interestingly, one individual from Broadway stood out as being highly assigned to the commercial native bee cluster (at $K=3$) using Admixture and Entropy. This individual is thus potentially either highly introgressed with native commercial bees, an escaped bee from farms utilizing native commercial colonies, genetically similar to the bees used to establish the commercial native stock, or a bee from established wild colonies of commercial bees. The geographic location of Broadway in the central region of the United Kingdom means that at the time of sampling there were farms using both non-native and native commercial colonies close to, or within foraging distance of the sampling location (Natural England), which could facilitate the different scenarios of introgression or escape. Indeed escapees from commercial colonies have been detected in other studies of *B. terrestris* in the Iberian peninsula and southern Sweden (Kardum Hjort et al., 2022; Seabra et al., 2019), and *B. impatiens* in New England, USA (Suni et al., 2017). The finding that individuals from the commercial native subspecies are genetically assigned in similar proportions to the wild individuals cluster and the commercial non-native bees cluster (at $K=2$) may be due to the use of *B. terrestris* queens from several different origins in commercial breeding programs (Velthuis & van Doorn, 2006), and/or interbreeding in breeding facilities.

There are several factors that may be acting to limit introgression between the different subspecies of *B. terrestris*, despite the use of commercial colonies having presented opportunities for interaction between the groups. There may be mating barriers between the different subspecies, even though they reportedly can mate and produce viable offspring under laboratory conditions (De Jonghe, 1986; Ings et al., 2005; Lecocq et al., 2015), as well as under natural conditions where different subspecies co-exist, such as reported in southern France (Ings et al., 2005). For instance, male bumblebee sex pheromones, which act as an attractant for unmated queens (Kullenberg et al., 1973), have been found to vary distinctly in composition between *dalmatinus*, *terrestris* and *sassaricus*, whereas the subspecies *terrestris* and *lusitanicus* were found to be similar in composition (Coppee et al., 2008). These sex pheromones are thought to be crucial in the bumblebee mate recognition system, with males of several bumblebee species performing patrolling behaviour and scent marking objects that attracts conspecific virgin females (reviewed in Baer, 2003). Further, it could be possible that the timing of the production of the queens and drones may vary between the different subspecies, and the phenology of the different subspecies may be less well adapted to the UK climate (e.g. Owen et al., 2016).

We find that, while no population in our analyses has distinctively higher genetic diversity than the other populations, commercial populations showed a tendency towards higher genetic diversity compared to wild populations. This higher level of genetic diversity may again be influenced by the breeding program techniques, including

the use of queens from different origins (Velthuis & van Doorn, 2006) and crossing between lineages in breeding facilities. Other studies have found no large difference in genetic diversity between wild and commercial *B. terrestris* (Kardum Hjort et al., 2022; Seabra et al., 2019). Interestingly, although Kardum Hjort et al. (2022) find that both wild and commercial groups show moderate genetic diversity, they find that wild bumblebees have higher nucleotide diversity (π), while the number of segregating sites (θ_w) was higher in commercial bumblebees. These authors show that this discrepancy is reflected in a significantly positive Tajima's D that may be indicative of balancing selection in wild bumblebees in Southern Sweden. Additional studies have found somewhat higher genetic diversity in wild compared to commercial bumblebee populations (Hart et al., 2021; Moreira et al., 2015; Suni et al., 2017). These different patterns in diversity may be due to differences in commercial bumblebee rearing techniques, as well as differences in population dynamics, such as population contractions and declines between different geographic regions. Declines in diversity and abundance, as well as population contractions in wild bees, have been reported in several studies (e.g. Goulson & Hughes, 2015; Marshman et al., 2019).

Interestingly, we find putative evidence for both global and local genomic adaptation. We caution, however, that it remains a challenge, in general, to accurately disentangle the contributions of the different evolutionary processes (e.g. Charlesworth & Jensen, 2021, 2022). Charlesworth and Jensen (2022), for instance, highlight several neutral genetic processes, as well as demographic and selective forces, that likely play a role in restricting observed levels of DNA sequence variation in natural populations to a smaller range that would be expected from differences in their population sizes alone (Lewontin's Paradox). Their evaluations suggest that population size, as well as selective sweeps and mutational bias, are of high relative importance in resolving this paradox.

Here, we find that regions of high nucleotide diversity tend to overlap across all comparisons, including between commercial and wild populations, suggesting that regions of the genome harbouring high diversity tend to be the same between populations of the different subspecies. The existence of typically high diversity regions across subspecies may be related to variation in recombination rates across regions of the genome – which we did not study here. However, regions of low nucleotide diversity did not significantly overlap. Several processes can drive low nucleotide diversity, and if more processes are acting there will be a higher likelihood that different factors are driving each population. Further, the finding that there is a lack of overlap between regions under putative global adaptation in the UK (characterized by low π , low d_{xy} and high F_{ST}) in wild individuals, and low nucleotide diversity regions in commercial individuals, provides support for these low π regions of the genome being specific to the wild bees rather than being a general characteristic of the genome of both groups.

We found 21 genes in the regions of the genome putatively under global adaptation, and 182 genes in the regions under local adaptation. Notably, the genes found to be under local adaptation

were enriched for functions related to taste receptor activity, oxidoreductase activity, fatty acid and lipid biosynthetic processes. Taste receptor functions are likely important for foraging in bees as the taste system is known to function in animals more generally to help distinguish between harmful and nutritious foods which contains sugars or fats and amino acids (Lim et al., 2019). Honey bees, for instance, have been shown to prefer pollen that is richer in the most essential amino acids (Cook et al., 2003). Honey bee mouth parts have been shown to respond to amino acids, and amino acid receptors have been identified in honey bees (Lim et al., 2019). Such gustatory systems are key mechanisms for animals, from mammals to insects, for sensing amino acids in the environment (Lim et al., 2019). In addition, we find enrichment for one KEGG pathway, the biosynthesis of unsaturated fatty acids, which has been shown to have a role in thermoregulation in *Drosophila* (Suito et al., 2020).

5 | CONCLUSIONS

Using genome-wide SNPs, we document limited levels of introgression of commercial non-native and native subspecies into wild populations of the bumblebee *B. terrestris* in the UK. However, this does not imply that the long-term risks associated with such introgression are negligible for the conservation of wild bees. Such risks include – but are not limited to – loss of genetic diversity, particularly in genomic regions adapted to the ecological conditions found either across the UK (possible 'global adaptation' regions in this study), or more specific locations within the country (possible 'local adaptation' regions in this study). Genetic variation in these regions may be critical for the long-term resilience of these pollinators. Only the discovery of very substantial barriers to the long-term establishment of non-native subspecies, and to hybridization or introgression (e.g. extremely reduced fitness in hybrids), or alternatively determining that introgression between subspecies is highly adaptive, would alleviate concern about this potential threat to biodiversity. In the meantime, action needs to be taken to reduce the risks of compromising the genetic integrity of local populations. The main strategies going forward need to be the use of native bees for pollination where possible, and the use of native subspecies in commercial facilities where these bees are produced in a way that does not endanger local genetic diversity. Genetic testing of bees reared for pollination, such as using the private alleles detected in wild populations here, can serve as a check of industry procedures. Additionally, other threats of commercial trade such as the spread of pathogens need to be investigated further and mitigated.

AUTHOR CONTRIBUTIONS

P.F., W.O.H.H. and J.C.J. designed the research. All authors performed the research. P.F., C.F., V.S. and J.C.J. performed analyses. P.F., C.F. and J.C.J. wrote the manuscript in collaboration with all authors.

ACKNOWLEDGEMENTS

This research was funded by a Marie Curie Fellowship to J.C.J. We thank Anna Olsson for assistance with DNA extractions of the *B. t. dalmatinus* samples. We thank tomato farmers in England for providing access to their greenhouses and commercial *B. terrestris* populations.

CONFLICT OF INTEREST STATEMENT

There are no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

All sequence data presented here (*fastq* format) and the file containing individual genotypes (*vcf* format) have been deposited into the NCBI SRA database (PRJNA1015651) and the Dryad Digital Repository (doi: [10.5061/dryad.5tb2rbp9r](https://doi.org/10.5061/dryad.5tb2rbp9r)) respectively.

BENEFIT-SHARING STATEMENT

Benefits generated from this research come from the sharing of our data and results on public databases as described above.

ORCID

Paolo Franchini  <https://orcid.org/0000-0002-8184-1463>

REFERENCES

- Acosta, A. L., Giannini, T. C., Imperatriz-Fonseca, V. L., & Saraiva, A. M. (2016). Worldwide alien invasion: A methodological approach to forecast the potential spread of a highly invasive pollinator. *PLoS One*, 11(2), e0148295. <https://doi.org/10.1371/journal.pone.0148295>
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Anderson, M. J. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, 62(1), 245–253. <https://doi.org/10.1111/j.1541-0420.2005.00440.x>
- Baer, B. (2003). Bumblebees as model organisms to study male sexual selection in social insects. *Behavioral Ecology and Sociobiology*, 54(6), 521–533. <https://doi.org/10.1007/s00265-003-0673-5>
- Bartomeus, I., Molina, F. P., Hidalgo-Galiana, A., & Ortego, J. (2020). Safeguarding the genetic integrity of native pollinators requires stronger regulations on commercial lines. *Ecological Solutions and Evidence*, 1(1), e12012. <https://doi.org/10.1002/2688-8319.12012>
- Biesmeijer, J. C., Roberts, S. P., Reemer, M., Ohlemüller, R., Edwards, M., Peeters, T., Schaffers, A. P., Potts, S. G., Kleukers, R., Thomas, C. D., Settele, J., & Kunin, W. E. (2006). Parallel declines in pollinators and insect-pollinated plants in Britain and The Netherlands. *Science*, 313, 351–354.
- Booker, T. R., Yeaman, S., & Whitlock, M. C. (2021). Global adaptation complicates the interpretation of genome scans for local adaptation. *Evolution Letters*, 5(1), 4–15. <https://doi.org/10.1002/evl3.208>
- Bookstein, F. L. (2017). A newly noticed formula enforces fundamental limits on geometric morphometric analyses. *Evolutionary Biology*, 44(4), 522–541. <https://doi.org/10.1007/s11692-017-9424-9>
- Buchmann, S. (1983). Buzz pollination in angiosperms. In C. Jones & R. Little (Eds.), *Handbook of experimental pollination biology* (pp. 73–113). Van Nostrand-Rheinhold.

- Cameron, S. A., & Sadd, B. M. (2020). Global trends in bumble bee health. *Annual Review of Entomology*, 65, 209–232. <https://doi.org/10.1146/annurev-ento-011118-111847>
- Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3 (Bethesda, Md.)*, 1(3), 171–182. <https://doi.org/10.1534/g3.111.000240>
- Cejas, D., De la Rúa, P., Ornos, C., Michez, D., & Muñoz, I. (2021). Spatial and temporal patterns of genetic diversity in *Bombus terrestris* populations of the Iberian Peninsula and their conservation implications. *Scientific Reports*, 11(1), 22471.
- Cejas, D., Lopez-Lopez, A., Munoz, I., Ornos, C., & De la Rúa, P. (2020). Unveiling introgression in bumblebee (*Bombus terrestris*) populations through mitogenome-based markers. *Animal Genetics*, 51(1), 70–77. <https://doi.org/10.1111/age.12874>
- Cejas, D., Ornos, C., Muñoz, I., & De la Rúa, P. (2018). Searching for molecular markers to differentiate *Bombus terrestris* (Linnaeus) subspecies in the Iberian Peninsula. *Sociobiology*, 65(4), 558. <https://doi.org/10.13102/sociobiology.v65i4.3442>
- Charlesworth, B., & Jensen, J. D. (2021). Effects of selection at linked sites on patterns of genetic variability. *Annual Review of Ecology, Evolution, and Systematics*, 52(1), 177–197. <https://doi.org/10.1146/annurev-ecolsys-010621-044528>
- Charlesworth, B., & Jensen, J. D. (2022). How can we resolve Lewontin's Paradox? *Genome Biology and Evolution*, 14(7), evac096. <https://doi.org/10.1093/gbe/evac096>
- Cook, S. M., Awmack, C. S., Murray, D. A., & Williams, I. H. (2003). Are honey bees' foraging preferences affected by pollen amino acid composition? *Ecological Entomology*, 28, 622–627.
- Cooley, H., & Vallejo-Marin, M. (2021). Buzz-pollinated crops: A global review and meta-analysis of the effects of supplemental bee pollination in tomato. *Journal of Economic Entomology*, 114, 505–519.
- Coppee, A., Terzo, M., Valterova, I., & Rasmont, P. (2008). Intraspecific variation of the cephalic labial gland secretions in *Bombus terrestris* (L.) (Hymenoptera: Apidae). *Chemistry & Biodiversity*, 5, 2654–2661.
- Dafni, A., Kevan, P., Gross, C. L., & Goka, K. (2010). *Bombus terrestris*, pollinator, invasive and pest: An assessment of problems associated with its widespread introductions for commercial purposes. *Applied Entomology and Zoology*, 45(1), 101–113. <https://doi.org/10.1303/aez.2010.101>
- Dafni, A., & Shmida, A. (1996). The possible ecological implications of the invasion of *Bombus terrestris* (L.) (Apidae) at Mt Carmel, Israel. In A. Matheson (Ed.), *The conservation of bees* (pp. 183–200). The Linnean Society of London and The International Bee Research Association.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., De Pisto, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & Group, G. P. A. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCftools. *GigaScience*, 10(2), giab008. <https://doi.org/10.1093/gigascience/giab008>
- De Jonghe, R. (1986). Male parentage in a bumblebee: *Bombus terrestris* (Linnaeus, 1758) (Hymenoptera: Apoidea). *Phegea*, 14, 49–54.
- Department for the Environment, Food and Rural Affairs. (2014). Consultation on class licence WML-CL22: To permit there lease of non-native subspecies of the bumblebee (*Bombus terrestris*) in commercial glass-houses or poly-tunnels for crop pollination and research [pdf]. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/341110/wml-cl22-consultation-notes.pdf
- Estoup, A., Solignac, M., Cornuet, J. M., Goudet, J., & Scholl, A. (1996). Genetic differentiation of continental and Island populations of *Bombus terrestris* (Hymenoptera: Apidae) in Europe. *Molecular Ecology*, 5(1), 19–31.
- Facon, B., Crespin, L., Loiseau, A., Lombaert, E., Magro, A., & Estoup, A. (2011). Can things get worse when an invasive species hybridizes? The harlequin ladybird *Harmonia axyridis* in France as a case study. *Evolutionary Applications*, 4(1), 71–88. <https://doi.org/10.1111/j.1752-4571.2010.00134.x>
- Field, J., Paxton, R. J., Soro, A., & Bridge, C. (2010). Cryptic plasticity underlies a major evolutionary transition. *Current Biology*, 20, 2028–2031.
- Franchini, P., Jones, J. C., Xiong, P., Kneitz, S., Gompert, Z., Warren, W. C., Walter, R. B., Meyer, A., & Schartl, M. (2018). Long-term experimental hybridisation results in the evolution of a new sex chromosome in swordtail fish. *Nature Communications*, 9(1), 5136. <https://doi.org/10.1038/s41467-018-07648-2>
- Franchini, P., Monné Parera, D., Kautt, A. F., & Meyer, A. (2017). quadRAD: A new high-multiplexing and PCR duplicate removal ddRAD protocol produces novel evolutionary insights in a nonradiating cichlid lineage. *Molecular Ecology*, 26(10), 2783–2795. <https://doi.org/10.1111/mec.14077>
- Fruciano, C. (2018). *GeometricMorphometricsMix: Miscellaneous functions useful for geometric morphometrics*. version 0.0.3. <https://github.com/fruciano/GeometricMorphometricsMix>
- Fruciano, C., Pappalardo, A. M., Tigano, C., & Ferrito, V. (2014). Phylogeographical relationships of Sicilian brown trout and the effects of genetic introgression on morphospace occupation. *Biological Journal of the Linnean Society*, 112, 387–398.
- Gel, B., Diez-Villanueva, A., Serra, E., Buschbeck, M., Peinado, M. A., & Malinverni, R. (2016). regioneR: An R/Bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics*, 32(2), 289–291. <https://doi.org/10.1093/bioinformatics/btv562>
- Gill, R. J., Baldock, K. C. R., Brown, M. J. F., Cresswell, J. E., Dicks, L. V., Fountain, M. T., Garratt, M. P. D., Gough, L. A., Heard, M. S., Holland, J. M., Ollerton, J., Stone, G. N., Tang, C. Q., Vanbergen, A. J., Volger, A. P., Woodward, G., Arce, A. N., Boatman, N. D., Brand-Hardy, R., ... Potts, S. G. (2016). Protecting an ecosystem service: Approaches to understanding and mitigating threats to wild insect pollinators. *Advances in Ecological Research*, 54, 135–206. <https://doi.org/10.1016/bs.aecr.2015.10.007>
- Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5, 184–186. <https://doi.org/10.1111/j.1471-8278>
- Goulson, D. (2010). Impacts of non-native bumblebees in Western Europe and North America. *Applied Entomology and Zoology*, 45(1), 7–12. <https://doi.org/10.1303/aez.2010.7>
- Goulson, D., & Hughes, W. O. H. (2015). Mitigating the anthropogenic spread of bee parasites to protect wild pollinators. *Biological Conservation*, 191, 10–19. <https://doi.org/10.1016/j.biocon.2015.06.023>
- Goulson, D., Nicholls, E., Botías, C., & Rotheray, E. L. (2015). Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. *Science*, 347(6229), 1255957. <https://doi.org/10.1126/science.1255957>
- Grixti, J. C., Wong, L. T., Cameron, S. A., & Favret, C. (2009). Decline of bumble bees (*Bombus*) in the North American Midwest. *Biological Conservation*, 142(1), 75–84. <https://doi.org/10.1016/j.biocon.2008.09.027>
- Hart, A. F., Maebe, K., Brown, G., Smagghe, G., & Ings, T. (2021). Winter activity unrelated to introgression in British bumblebee *Bombus terrestris* *audax*. *Apidologie*, 52(2), 315–327. <https://doi.org/10.1007/s13592-020-00822-w>
- Hingston, A., Marsden-Smedley, J., Driscoll, D., Corbett, S., Fenton, J., Anderson, R., Plowman, C., Mowling, F., Jenkin, M., Matsui, K., Bonham, K. J., Ilowski, M., Mcquillan, P. B., Yaxley, B., Reid, T., Storey, D., Poole, L., Mallick, S. A., Fitzgerald, N., ... Desmarchelier, J. M. (2002). Extent of invasion of Tasmanian native vegetation by

- the exotic bumblebee *Bombus terrestris* (Apoidea: Apidae). *Austral Ecology*, 27, 162–172.
- Hingston, A., & McQuillan, P. (1998). Does the recently introduced bumblebee *Bombus terrestris* (Apidae) threaten Australian ecosystems? *Australian Journal of Ecology*, 23, 539–549.
- Howlett, B. G., & Donovan, B. J. (2010). A review of New Zealand's deliberately introduced bee fauna: Current status and potential impacts. *New Zealand Entomologist*, 33(1), 92–101. <https://doi.org/10.1080/00779962.2010.9722196>
- Inari, N., Nagamitsu, T., Kenta, T., Goka, K., & Hiura, T. (2005). Spatial and temporal pattern of introduced *Bombus terrestris* abundance in Hokkaido, Japan, and its potential impact on native bumblebees. *Population Ecology*, 47(1), 77–82. <https://doi.org/10.1007/s10144-004-0205-9>
- Ings, T., Raine, N. E., & Chittka, L. (2005). Mating preference in the commercially imported bumblebee species *Bombus terrestris* in Britain (Hymenoptera: Apidae). *Entomologia Generalis*, 28, 233–238.
- Ings, T. C., Ward, N. L., & Chittka, L. (2006). Can commercially imported bumble bees out-compete their native conspecifics? *Journal of Applied Ecology*, 43(5), 940–948. <https://doi.org/10.1111/j.1365-2664.2006.01199.x>
- Jaffe, R., Pope, N., Acosta, A. L., Alves, D. A., Arias, M. C., De la Rua, P., Francisco, F. O., Giannini, T. C., Gonzalez-Chaves, A., Imperatriz-Fonseca, V. L., Tavares, M. G., Jha, S., & Carvalheiro, L. G. (2016). Beekeeping practices and geographic distance, not land use, drive gene flow across tropical bees. *Molecular Ecology*, 25(21), 5345–5358. <https://doi.org/10.1111/mec.13852>
- Jombart, T., Pontier, D., & Dufour, A. B. (2009). Genetic markers in the playground of multivariate analysis. *Heredity (Edinb)*, 102(4), 330–341. <https://doi.org/10.1038/hdy.2008.130>
- Kardum Hjort, C., Paris, J. R., Olsson, P., Herbertsson, L., de Miranda, J. R., Dudaniec, R. Y., & Smith, H. G. (2022). Genomic divergence and a lack of recent introgression between commercial and wild bumblebees (*Bombus terrestris*). *Evolutionary Applications*, 15(3), 365–382. <https://doi.org/10.1111/eva.13346>
- Klein, A. M., Vaissiere, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. *Proceedings of the Biological Sciences*, 274(1608), 303–313. <https://doi.org/10.1098/rspb.2006.3721>
- Knaus, B. J., & Grunwald, N. J. (2017). vcfr: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, 17(1), 44–53. <https://doi.org/10.1111/1755-0998.12549>
- Korunes, K. L., & Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, 21(4), 1359–1368. <https://doi.org/10.1111/1755-0998.13326>
- Kraus, F. B., Wolf, S., & Moritz, R. F. A. (2009). Male flight distance and population substructure in the bumblebee *Bombus terrestris*. *Journal of Animal Ecology*, 78, 247–252. <https://doi.org/10.1111/j.1365-2656.2008.01479.x>
- Kullenberg, B., Bergström, G., Bringer, B., Carlberg, B., & Cederberg, B. (1973). Observations on scent marking by *Bombus Latr.* and *Psithyrus Lep.* males (Hym., Apidae) and localization of site of production of the secretion. *Zoon Suppl*, 1, 23–30.
- Lecocq, T., Coppée, A., Mathy, T., Lhomme, P., Cammaerts-Tricot, M.-C., Urbanová, K., Valterova, I., & Rasmont, P. (2015). Subspecific differentiation in male reproductive traits and virgin queen preferences, in *Bombus terrestris*. *Apidologie*, 46(5), 595–605. <https://doi.org/10.1007/s13592-015-0349-y>
- Lecocq, T., Rasmont, P., Harpke, A., & Schweiger, O. (2016). Improving international trade regulation by considering intraspecific variation for invasion risk assessment of commercially traded species: The *Bombus terrestris* case. *Conservation Letters*, 9(4), 281–289. <https://doi.org/10.1111/conl.12215>
- Lepais, O., Darvill, B., O'Connor, S., Osbourne, J. L., Sanderson, R. A., Cussans, J., Goffe, L., & Goulson, D. (2010). Estimation of bumblebee queen dispersal distances using sibship reconstruction method. *Molecular Ecology*, 19(4), 819–831.
- Lever, J., Krzywinski, M., & Altman, N. (2017). Principal component analysis. *Nature Methods*, 14(7), 641–642. <https://doi.org/10.1038/nmeth.4346>
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21), 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Lim, S., Jung, J., Yunusbaev, U., Ilyasov, R., & Kwon, H. W. (2019). Characterization and its implication of a novel taste receptor detecting nutrients in the honey bee, *Apis mellifera*. *Scientific Reports*, 9(1), 11620. <https://doi.org/10.1038/s41598-019-46738-z>
- Løken, A. (1973). Studies on Scandinavian bumblebees (Hymenoptera, Apidae). *Norsk Entomologisk Tidsskrift*, 20, 1–219.
- Lye, G. C., Lepais, O., & Goulson, D. (2011). Reconstructing demographic events from population genetic data: The introduction of bumblebees to New Zealand. *Molecular Ecology*, 20(14), 2888–2900. <https://doi.org/10.1111/j.1365-294X.2011.05139.x>
- Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population inference from RADseq data. *Molecular Biology and Evolution*, 35(5), 1284–1290. <https://doi.org/10.1093/molbev/msy023>
- Mallinger, R. E., Gaines-Day, H. R., & Gratton, C. (2017). Do managed bees have negative effects on wild bees?: A systematic review of the literature. *PLoS One*, 12(12), e0189268. <https://doi.org/10.1371/journal.pone.0189268>
- Marshman, J., Blay-Palmer, A., & Landman, K. (2019). Anthropocene crisis: Climate change, pollinators, and food security. *Environments*, 6(2), 22. <https://doi.org/10.3390/environments6020022>
- Matsumura, C., Yokoyama, J., & Washitani, I. (2004). Invasion status and potential ecological impacts of an invasive alien bumblebee, *Bombus terrestris* L. (Hymenoptera: Apidae) naturalized in southern Hokkaido, Japan. *Global Environmental Research*, 8, 51–66.
- Moreira, A. S., Horgan, F. G., Murray, T. E., & Kakouli Duarte, T. (2015). Population genetic structure of *Bombus terrestris* in Europe: Isolation and genetic differentiation of Irish and British populations. *Molecular Ecology*, 24(13), 3257–3268. <https://doi.org/10.1111/mec.13235>
- Nayak, R. K., Rana, K., Bairwa, V. K., & Bharti, V. D. (2020). A review on role of bumblebee pollination in fruits and vegetables. *Phytopathology*, 9(3), 1328–1334. <https://doi.org/10.22271/phyto.2020.v9.i3v.11494>
- Nyerki, E., Kalmar, T., Schutz, O., Lima, R. M., Neparaczki, E., Torok, T., & Maroti, Z. (2023). correctKin: An optimized method to infer relatedness up to the 4th degree from low-coverage ancient human genomes. *Genome Biology*, 24(1), 38. <https://doi.org/10.1186/s13059-023-02882-4>
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Michin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., & Wagner, H. (2016). *vegan: Community ecology package (version R package version 2.3-3)*.
- Ollerton, J., Erenler, H., Edwards, M., & Crockett, R. (2014). Extinctions of aculeate pollinators in Britain and the role of large-scale agricultural changes. *Science*, 346, 1360–1362.
- Owen, E. L., Bale, J. S., & Hayward, S. A. (2016). Establishment risk of the commercially imported bumblebee *Bombus terrestris dalmatinus* – Can they survive UK winters? *Apidologie*, 47, 66–75. <https://doi.org/10.1007/s13592-015-0376-8>
- Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). Global pollinator declines: Trends, impacts and

- drivers. *Trends in Ecology & Evolution*, 25(6), 345–353. <https://doi.org/10.1016/j.tree.2010.01.007>
- Powney, G. D., Carvell, C., Edwards, M., Morris, R. K. A., Roy, H. E., Woodcock, B. A., & Isaac, N. J. B. (2019). Widespread losses of pollinating insects in Britain. *Nature Communications*, 10(1), 1018. <https://doi.org/10.1038/s41467-019-08974-9>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- Rasmont, P., Coppee, A., Michez, D., & De Meulemeester, T. (2008). An overview of the *Bombus terrestris* (L. 1758) subspecies (Hymenoptera: Apidae). *Annales de la Société Entomologique de France*, 44(2), 243–250. <https://doi.org/10.1080/00379271.2008.10697559>
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, 47(W1), W191–W198. <https://doi.org/10.1093/nar/gkz369>
- Robertson, A. (1991). A mid-winter colony of *Bombus terrestris* L. (Hym., Apidae) in Devon. *Entomologist's Monthly Magazine*, 127, 165–166.
- Russ, J., & Jasper, S. Y. (2020). Local adaptation can cause both peaks and troughs in nucleotide diversity within populations. *bioRxiv* 2020.06.03.132662.
- Schmid-Hempel, R., Eckhardt, M., Goulson, D., Heinzmann, D., Lange, C., Plischuk, S., Escudero, L. R., Salathe, R., Scriven, J. J., & Schmid-Hempel, P. (2014). The invasion of southern South America by imported bumblebees and associated parasites. *Journal of Animal Ecology*, 83(4), 823–837. <https://doi.org/10.1111/1365-2656.12185>
- Seabra, S. G., Silva, S. E., Nunes, V. L., Sousa, V. C., Martins, J., Marabuto, E., Rodrigues, A. S. B., Pina-Martins, F., Laurentino, T. G., Rebelo, M. T., Figueiredo, E., & Paulo, O. S. (2019). Genomic signatures of introgression between commercial and native bumblebees, *Bombus terrestris*, in western Iberian Peninsula-implications for conservation and trade regulation. *Evolutionary Applications*, 12(4), 679–691. <https://doi.org/10.1111/eva.12732>
- Semmens, T. D., Turner, E., & Buttermore, R. (1993). *Bombus terrestris* (L.) (Hymenoptera: Apidae) now established in Tasmania. *Australian Journal of Entomology*, 32(4), 346. <https://doi.org/10.1111/j.1440-6055.1993.tb00598.x>
- Shastri, V., Adams, P. E., Lindtke, D., Mandeville, E. G., Parchman, T. L., Gompert, Z., & Buerkle, C. A. (2021). Model-based genotype and ancestry estimation for potential hybrids with mixed-ploidy. *Molecular Ecology Resources*, 21(5), 1434–1451. <https://doi.org/10.1111/1755-0998.13330>
- Sjöstrand, A. E., Sjödin, P., & Jakobsson, M. (2014). Private haplotypes can reveal local adaptation. *BMC Genetics*, 15, 61.
- Soroye, P., Newbold, T., & Kerr, J. (2020). Climate change contributes to widespread declines among bumble bees across continents. *Science*, 367, 685–688.
- Steinig, E. J., Neuditschko, M., Khatkar, M. S., Raadsma, H. W., & Zenger, K. R. (2016). netview p: A network visualization tool to unravel complex population structure using genome-wide SNPs. *Molecular Ecology Resources*, 16(1), 216–227. <https://doi.org/10.1111/1755-0998.12442>
- Stelzer, R. J., Chittka, L., Carlton, M., & Ings, T. C. (2010). Winter active bumblebees (*Bombus terrestris*) achieve high foraging rates in urban Britain. *PLoS One*, 5(3), e9559. <https://doi.org/10.1371/journal.pone.0009559>
- Suito, T., Nagao, K., Takeuchi, K., Juni, N., Hara, Y., & Umeda, M. (2020). Functional expression of $\Delta 12$ fatty acid desaturase modulates thermoregulatory behaviour in *Drosophila*. *Scientific Reports*, 10(1), 11798. <https://doi.org/10.1038/s41598-020-68601-2>
- Suni, S. S., Scott, Z., Averill, A., & Whiteley, A. (2017). Population genetics of wild and managed pollinators: Implications for crop pollination and the genetic integrity of wild bees. *Conservation Genetics*, 18(3), 667–677. <https://doi.org/10.1007/s10592-017-0955-5>
- Sutherland, W. J., Barnard, P., Broad, S., Clout, M., Connor, B., Côté, I. M., Dicks, L. V., Doran, H., Entwistle, A. C., Fleishman, E., Fox, M., Gaston, K. J., Gibbons, D. W., Jiang, Z., Keim, B., Lickorish, F. A., Markillie, P., Monk, K. A., Pearce-Higgins, J. W., ... O'Connor, B. (2017). A 2017 horizon scan of emerging issues for global conservation and biological diversity. *Trends in Ecology & Evolution*, 32(1), 31–40. <https://doi.org/10.1016/j.tree.2016.11.005>
- Van der Auwera, G., & O'Connor, B. (2020). *Genomics in the cloud: Using Docker, GATK, and WDL* in O'Reilly Media Inc.
- Velthuis, H. H. W. (2002). The historical background of the domestication of the bumble-bee, *Bombus terrestris*, and its introduction in agriculture. In P. Kevan & V. L. Imperatriz-Fonseca (Eds.), *Pollinating bees: The conservation link between agriculture and nature* (pp. 177–184). Ministry of Environment.
- Velthuis, H. H. W., & van Doorn, A. (2006). A century of advances in bumblebee domestication and the economic and environmental aspects of its commercialization for pollination. *Apidologie*, 37(4), 421–451. <https://doi.org/10.1051/apido:2006019>
- Wahengbam, J., Raut, A. M., Pal, S., & Najitha Banu, A. (2019). Role of bumble bee in pollination. *Annals of Biology*, 35(2), 290–295.
- Walther-Hellwig, K., & Frankl, R. (2000). Foraging habitats and foraging distances of bumblebees, *Bombus* spp. (Hym., Apidae), in an agricultural landscape. *Journal of Applied Entomology*, 124(7–8), 299–306. <https://doi.org/10.1046/j.1439-0418.2000.00484.x>
- Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571–3594.
- Widmer, A., Schmid-Hempel, P., Estoup, A., & Scholl, A. (1998). Population genetic structure and colonization history of *Bombus terrestris* sl (Hymenoptera: Apidae) from the Canary Islands and Madeira. *Heredity*, 81(5), 563–572. <https://doi.org/10.1046/j.1365-2540.1998.00407.x>
- Williams, P. H., & Osborne, J. L. (2009). Bumblebee vulnerability and conservation world-wide. *Apidologie*, 40(3), 367–387. <https://doi.org/10.1051/apido/2009025>

SUPPORTING INFORMATION

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

How to cite this article: Franchini, P., Fruciano, C., Wood, T. J., Shastri, V., Goulson, D., Hughes, W. O. H., & Jones, J. C. (2023). Limited introgression from non-native commercial strains and signatures of adaptation in the key pollinator *Bombus terrestris*. *Molecular Ecology*, 32, 5709–5723. <https://doi.org/10.1111/mec.17151>