



OPEN ACCESS

EDITED BY
Alison G Nazareno,
Federal University of Minas Gerais, Brazil

REVIEWED BY
Thais Ribeiro Pfeilsticker,
University of Tasmania, Australia
Meaghan Leigh Pimsler,
American Association For The
Advancement of Science, United States

*CORRESPONDENCE
Kevin Maebe,
kevin.maebe@ugent.be

SPECIALTY SECTION
This article was submitted to
Evolutionary and Population Genetics,
a section of the journal
Frontiers in Genetics

RECEIVED 13 July 2022
ACCEPTED 21 September 2022
PUBLISHED 05 October 2022

CITATION
Hart AF, Verbeeck J, Ariza D, Cejas D,
Ghisbain G, Honchar H, Radchenko VG,
Straka J, Ljubomirov T, Lecocq T,
Dániel-Ferreira J, Flaminio S,
Bortolotti L, Karise R, Meeus I,
Smagghe G, Vereecken N,
Vandamme P, Michez D and Maebe K
(2022), Signals of adaptation to
agricultural stress in the genomes of
two European bumblebees.
Front. Genet. 13:993416.
doi: 10.3389/fgene.2022.993416

COPYRIGHT
© 2022 Hart, Verbeeck, Ariza, Cejas,
Ghisbain, Honchar, Radchenko, Straka,
Ljubomirov, Lecocq, Dániel-Ferreira,
Flaminio, Bortolotti, Karise, Meeus,
Smagghe, Vereecken, Vandamme,
Michez and Maebe. This is an open-
access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Signals of adaptation to agricultural stress in the genomes of two European bumblebees

Alex F. Hart¹, Jaro Verbeeck¹, Daniel Ariza¹, Diego Cejas²,
Guillaume Ghisbain^{2,3}, Hanna Honchar⁴,
Vladimir G. Radchenko⁴, Jakub Straka⁵, Toshko Ljubomirov⁶,
Thomas Lecocq⁷, Juliana Dániel-Ferreira⁸, Simone Flaminio⁹,
Laura Bortolotti⁹, Reet Karise¹⁰, Ivan Meeus¹, Guy Smagghe¹,
Nicolas Vereecken¹¹, Peter Vandamme¹², Denis Michez² and
Kevin Maebe^{1*}

¹Ghent University, Faculty of Bioscience Engineering, Department of Plants and Crops, Lab of Agrozoology, Ghent, Belgium, ²Laboratory of Zoology, Research Institute for Biosciences, University of Mons, Mons, Belgium, ³Smithsonian Tropical Research Institute, Gamboa, Panama, ⁴Institute for Evolutionary Ecology, National Academy of Sciences of Ukraine, Kyiv, Ukraine, ⁵Charles University, Faculty of Science, Department of Zoology, Praha, Czech Republic, ⁶Institute of Biodiversity and Ecosystem Research—Bulgarian Academy of Sciences, Sofia, Bulgaria, ⁷Université de Lorraine, INRAE, URAFPA, Nancy, France, ⁸Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden, ⁹Council for Agricultural Research and Economics, Research Centre for Agriculture and Environment, Bologna, Italy, ¹⁰Estonian University of Life Sciences, Institute of Agricultural and Environmental Sciences, Tartu, Estonia, ¹¹Agroecology Lab, Université Libre de Bruxelles (ULB), Brussels, Belgium, ¹²Laboratory of Microbiology, Department of Biochemistry and Microbiology, Ghent University, Ghent, Belgium

Human-induced environmental impacts on wildlife are widespread, causing major biodiversity losses. One major threat is agricultural intensification, typically characterised by large areas of monoculture, mechanical tillage, and the use of agrochemicals. Intensification leads to the fragmentation and loss of natural habitats, native vegetation, and nesting and breeding sites. Understanding the adaptability of insects to these changing environmental conditions is critical to predicting their survival. Bumblebees, key pollinators of wild and cultivated plants, are used as model species to assess insect adaptation to anthropogenic stressors. We investigated the effects of agricultural pressures on two common European bumblebees, *Bombus pascuorum* and *B. lapidarius*. Restriction-site Associated DNA Sequencing was used to identify loci under selective pressure across agricultural-natural gradients over 97 locations in Europe. 191 unique loci in *B. pascuorum* and 260 in *B. lapidarius* were identified as under selective pressure, and associated with agricultural stressors. Further investigation suggested several candidate proteins including several neurodevelopment, muscle, and detoxification proteins, but these have yet to be validated. These results provide insights into agriculture as a stressor for bumblebees, and signal for conservation action in light of ongoing anthropogenic changes.

KEYWORDS

Bombus, global change, RADseq, population genomics, agricultural intensification, bee decline, Anthropocene

Introduction

Humans have been transforming their surrounding landscapes, the oceans, and the atmosphere, in a time period known as the Anthropocene (Crutzen and Stoermer, 2000; Ruddiman, 2013; Lewis and Maslin, 2015). This epoch is characterised by an increase in atmospheric greenhouse gases, changes in land surface structure and composition *via* deforestation, urbanisation, industrial and agricultural development, and a massive reduction in global biodiversity (e.g., Dirzo et al., 2014; Lewis and Maslin, 2015; Sage, 2020). Insects are not immune to the pressures of the Anthropocene, and are also suffering major declines in both abundance and diversity (e.g., Sánchez-Bayo and Wyckhuys, 2019; Wagner et al., 2021). Considerable habitat losses due to deforestation, urban development, intensive agriculture, and changing climatic conditions that are often largely different from the environmental conditions to which they are inherently adapted, are examples of anthropogenic threats (e.g., Sánchez-Bayo and Wyckhuys, 2019; Wagner et al., 2021). In Europe, agricultural practices were identified as a major driver of negative population trends (Van Swaay and Warren, 2012; Nieto et al., 2014).

The wide range of effects of modern intensive agriculture on insect populations have been well-studied (Raven & Wagner, 2021), especially in the contexts of agrochemicals and crop pests (reviewed in Sánchez-Bayo, 2011; Raven and Wagner, 2021; Gunstone et al., 2021). The impact of agricultural practices on insect populations is mostly studied using indirect approaches, studying the by-effects of agricultural intensification and subsequent habitat fragmentation and habitat loss, rather than a direct approach, studying how conspecific populations develop in both natural and agricultural environments (Goulson et al., 2005; Kosior et al., 2007; Goulson et al., 2008). However, direct approaches have already been used to: 1) identify genetic differentiation between populations, associated with habitat differences (Hereward et al., 2013; Andrews et al., 2020); for example, the hemipteran agricultural pest *Creonias dilutes* shows local adaptation to natural arid environments and agricultural sites, thought to be linked to host plant specialisation (Hereward et al., 2013). Similarly, Andrews et al. (2020) found evidence of adaptation to agricultural environments in populations of Elaterid beetles of the genus *Limonius* which was likely correlated with pesticide regimes; 2) or even adaptive responses to insecticide exposure; for instance in the Colorado potato beetle *Leptinotarsa decemlineata* and mosquito *Anopheles arabiensis*, which show significant upregulation of detoxifying cytochrome P450s after agrochemical exposure (Zhu et al., 2016; Oliver and Brooke,

2018), and in *Spodoptera exigua* and *Nilaparvata lugens* which show upregulation of Heat Shock Protein 70 in response to insecticides (You et al., 2016; Tarnawska et al., 2019).

Bumblebees are considered a model species to study anthropogenic influence on insects (Gérard et al., 2021; Maebe et al., 2021; Martinet et al., 2021; Theodorou et al., 2021). Both managed and wild bumblebee species are economically important keystone pollinators of mostly temperate climates, and have relatively generalist diets (Velthuis and Van Doorn, 2006; Klein et al., 2007, but see Wood et al., 2021). A significant proportion of *Bombus* species show declining population trends (Cameron and Sadd, 2020; Rasmont et al., 2021), while a few remain stable or are currently expanding their distribution range (Biella et al., 2020; Ghisbain et al., 2021).

The decline of bumblebee populations has been especially well-studied in Europe and North-America (e.g. Cameron et al., 2011; Kerr et al., 2015; Rasmont et al., 2015; Cameron & Sadd, 2020), and although declines are caused by many interacting factors (Williams and Osborne, 2009), they are mainly attributed to agricultural intensification, especially the loss and fragmentation of natural habitats (Kosior et al., 2007; Goulson et al., 2008; Marshall et al., 2018; Vray et al., 2019). A diverse range of biotic and abiotic stressors impacting bumblebees have been associated with agricultural intensification. Pesticides, such as neonicotinoids, are known to negatively impact bumblebee populations through mortality and sublethal effects (Thompson, 2001; Goulson et al., 2015; Anderson and Harmon-Threatt, 2019). Intensive agriculture is also known to reduce nesting site availability, both directly *via* the removal of hollows such as trees and leaf litter, and indirectly *via* a reduction in rodent burrow abundance (Alford, 1975; Goulson et al., 2008; Liczner and Colla, 2019). Agricultural conditions can generate nutritional stress, providing inadequate nutrient diversity (Vaudo et al., 2020), and/or on a shorter period than the activity period of the species (Goulson et al., 2008; Vanderplanck et al., 2019a). Furthermore, the introduction of non-native or managed bee taxa in agricultural areas can spread both endemic and introduced pathogens to native bee populations (Goulson et al., 2008; Meeus et al., 2011; Graystock et al., 2013; Yañez et al., 2020). Finally, stressors may work multiplicatively to an overall stronger negative effect; for example, pathogen exposure combined with nutritional stress in agricultural environments was found to be a contributor to the regression of bumblebee populations (McNeil et al., 2020).

Although intensive agriculture and habitat fragmentation are a relatively recent phenomenon on an evolutionary scale, recent research has paid a lot of attention to their anthropogenic impact on bumblebees and their resilience to these new

stressors (Maebe et al., 2021). Numerous studies have investigated several morphological traits in bumblebees that may be under selective pressure in agricultural environments, and primarily, body size has been associated with species-specific responses to agricultural intensification (Gérard et al., 2020, 2021) and urbanization (Theodorou et al., 2021; Tommasi et al., 2022). However, the increase in body size might be mainly or solely explained by the increased level of habitat fragmentation in urban and agricultural areas (Theodorou et al., 2021 and Gérard et al., 2021, respectively).

The loss and fragmentation of habitats have been shown to negatively affect both European and North American bumblebee species, leading to smaller, more isolated populations with lower genetic diversity due to genetic drift and inbreeding, which causes vulnerability to additional threats (e.g. Goulson et al., 2008; Jha and Kremen, 2013; Jha, 2015). Furthermore, Jha (2015) found that agricultural areas can act as significant barriers to gene flow in bumblebees. By comparing *B. lapidarius* collected from several urban and rural locations in Germany, Theodorou et al. (2018) found evidence of directional selection associated with urban land, identifying genes related to heat stress, oxidative stress, and metabolism as under selective pressure. Agricultural areas are therefore known to act as barriers to gene flow and to cause habitat fragmentation, which in turn can lead to differential population structuring and local adaptation. To date however, no study has taken a direct approach comparing wild bumblebees from natural and agricultural landscapes to explore possible genomic adaptations to the associated stressors.

The aim of this study was to identify genomic regions and/or underlying genes reflecting the different selective pressures on bumblebees between agricultural and natural areas. We used Restriction-site Associated DNA sequencing (henceforth RADSeq) analyses on specimens of two bumblebee species that are widespread and abundant in Europe: *B. pascuorum* (Scopoli, 1763) and *B. lapidarius* (Linnaeus, 1758). RADSeq is a powerful and cost-efficient technique for massive SNP discovery in large populations (Baird et al., 2008; Narum et al., 2013). Such reduced representation approaches allow for thorough coverage of the genome without the high costs associated with whole genome sequencing (WGS) (Baird et al., 2008; Davey et al., 2011; Narum et al., 2013), and are being used to highlight population structure, detect loci under selective pressure, and adaptation, among others, in bumblebees (e.g., Lozier et al., 2016; Jackson et al., 2018, 2020; Theodorou et al., 2018; Silva et al., 2020). Furthermore, we explored and ruled out population differentiation and demographic history as interfering factors when testing for signatures of selection (Nielsen, 2001; but see Theodorou et al., 2018). For this study, we hypothesize that the combination of several agricultural pressures (such as pesticide use, pathogens, and habitat fragmentation) will be detectable in both bumblebee species as a consequence of selection, including in genes related to detoxification, immunity and/or flight muscle development.

Materials and methods

Model species and sampling

Bombus pascuorum (Figures 1A,B) and *B. lapidarius* (Figures 1D,E) are two common and widespread bumblebee species found across Europe in a wide range of environmental conditions (Rasmont et al., 2015, 2021). Both are generalist foragers found in almost all habitats, and for which high intraspecific variation has been described (Lecocq et al., 2015, 2020; Williams et al., 2020; Rasmont et al., 2021). In terms of nesting, both species prefer relatively sheltered areas, although *B. lapidarius* usually settle underground in different cavities, while *B. pascuorum* builds its nests more often aboveground on the soil surface among the thick grass or in abandoned bird nests or tree holes (Alford, 1975; Radchenko & Pesenko, 1994). These two species were chosen as they belong to the two groups of the genus *Bombus*, the “short-faced” “pollen-storers” (*B. lapidarius*; Figure 1C) and the “long-faced” “pocket-makers” (*B. pascuorum*; Figure 1F) (Cameron et al., 2007; Goulson, 2010; Peeters et al., 2012).

Specimens were collected from 103 sampling sites in 16 countries across Europe during the bumblebee foraging seasons of 2018–2020 (Figure 2). Where possible, a paired sampling took place in both an agriculturally intensive landscape and a nearby low anthropogenic impact area (so-called natural landscape), such as forest, nature reserve, or unmanaged meadow by the same collector (see Supplementary Table S1 for sampling description and Supplementary Table S2 for full site coordinates). Most specimens were frozen and shipped on ice; although some collectors preferred to pin the recent specimens before shipment. Upon reception, specimens were immediately stored at -20°C . We selected only female specimens for further processing. After excluding 10 samples with no geographic data available, a total of 118 candidate *B. pascuorum* and 88 candidate *B. lapidarius* workers were collected, leaving 15 countries and 97 sites remaining (Table 1).

DNA extraction

Before extraction, the abdomen of each sample was removed under sterile conditions to reduce contamination from gut contents and microbiota (Vanderplanck et al., 2019b). DNA was extracted from the remaining tissue using an Invisorb Spin Tissue Mini Kit (Invitex Molecular GmbH, Germany) according to manufacturer’s instructions. To maximise yield three changes were made to the protocol: silica sand and stainless steel beads (Qiagen) were added to the lysis step using a TissueLyser II (Qiagen), 500–600 μl RLS buffer was used to increase DNA concentration, and samples were split across multiple spin filters when possible. DNA was cleaned and concentrated using Amicon Ultra Centrifugal

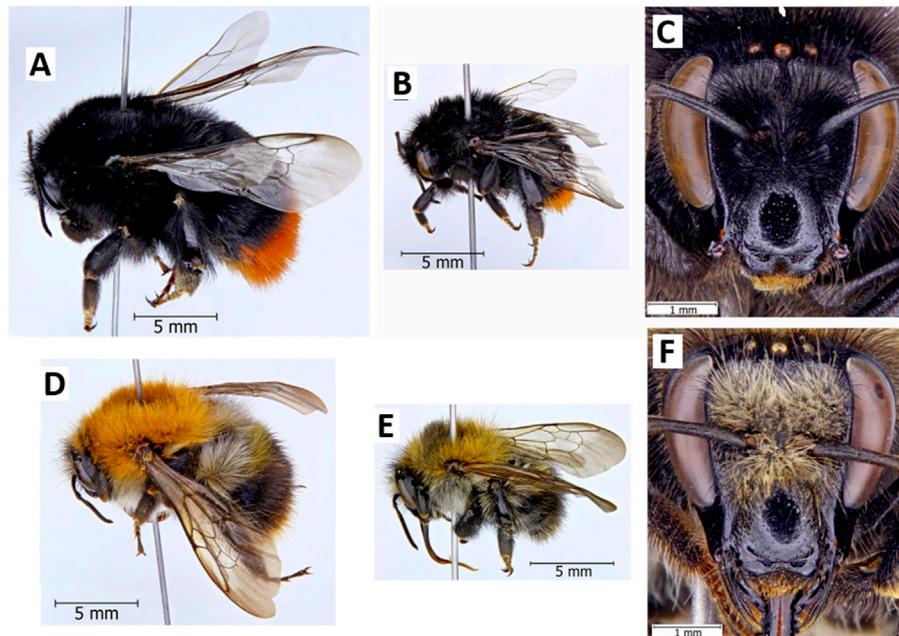


FIGURE 1

Queens and workers of *Bombus lapidarius* (A–C) and *B. pascuorum* (D–F). (A), (D)—Queen; (B), (E) Worker; (C), (F) Head of worker in frontal view. Photo credit Vladimir G. Radchenko⁴.

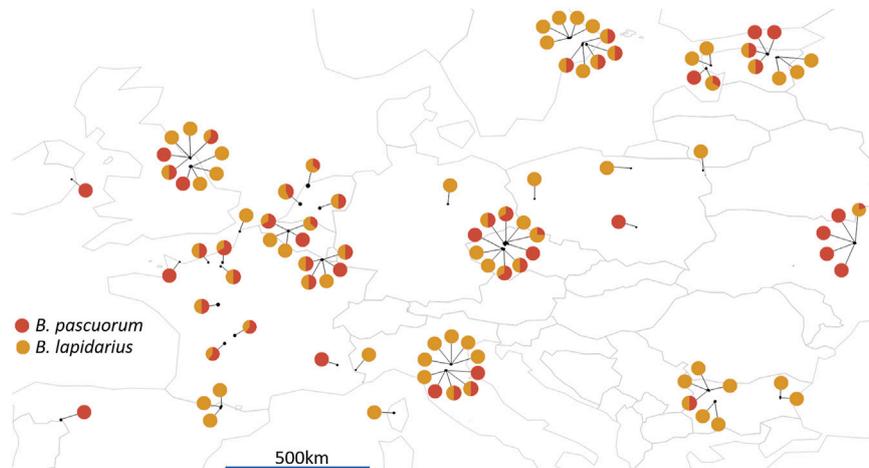


FIGURE 2

Map of the 97 sampling site locations across Europe. The pie charts represents the number of *B. pascuorum* (in red) and *B. lapidarius* (in orange) specimens sequenced per site. Data visualised with R packages 'worldmap' (South 2011) to draw the map and 'ggplot2' (Wickham 2009) to draw the pie charts.

Filters (Merck Millipore, Belgium). Excess RNA was apparent when the DNA quality was visually assessed using agarose gels, so samples were treated with 1 μ l of RNase A (Thermo Fisher Scientific, Belgium) at 2 mg/ml per 20–30 μ g of DNA.

Concentration was measured using a TECAN platform and Quant-it PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific, Belgium) before normalisation to 50 μ l of 20 ng/ μ l for library preparation.

TABLE 1 The total number of specimens by country and species that have been RAD sequenced and remained after filtering, with the number of sampling sites in parentheses, and total numbers in *italic*.

Country	No. of specimens	
	<i>B. pascuorum</i>	<i>B. lapidarius</i>
Belgium	11 (8)	9 (7)
Bulgaria	8 (7)	—
Czechia	11 (9)	10 (7)
England	8 (7)	10 (4)
Estonia	10 (8)	5 (5)
France	11 (6)	15 (7)
Germany	1 (1)	—
Italy	5 (5)	3 (3)
Netherlands	13 (3)	9 (3)
Poland	3 (3)	—
Portugal	—	4 (4)
Spain	5 (3)	—
Sweden	10 (9)	3 (3)
Ukraine	1 (1)	5 (5)
Total	<i>97 (70)</i>	<i>73 (48)</i>

Library preparation and RAD-Sequencing

Samples were outsourced to Floragenex Inc. (Oregon, United States) for library preparation and RAD-Sequencing. Samples underwent classic single-enzyme digestion with PstI (Baird et al., 2008), before shearing and single-end sequencing on an Illumina HiSeq 4000 with a target read length of 125 bp or 150 bp. We opted to use single-end sequencing as Jackson et al. (2018) found that it was not significantly different in reliably genotyping samples when compared with paired-end sequencing.

Data pre-processing

Sequencing data from Floragenex underwent a series of pre-processing steps before analysis. Finding appropriate thresholds and parameters for QC and read/locus filtering can be challenging. Overconservative handling of sequencing data may exclude ‘true’ positive outlier loci, while inappropriately liberal treatments may result in the inclusion of false positives; finding a balance in settings and parameters is key to producing robust but reliable data from noise and artefacts (Shafer et al., 2015; Paris et al., 2017; O’Leary et al., 2018). To mitigate such issues in our own dataset, close attention was paid to similar literature (e.g., Theodorou et al., 2018; Jackson et al., 2020; Silva et al., 2020) and to published guidance on parameter optimisation (Andrews et al., 2016; Paris et al., 2017; Rochette and Catchen, 2017; O’Leary et al., 2018).

Firstly, residual sequencing adaptors were removed by Floragenex prior to data delivery. Afterwards, samples were demultiplexed using Stacks (2.0) *process_radtags* using options *c*, *r*, and *q*, which remove reads with uncalled bases, ‘rescue’ barcodes and RADtags, and discard reads with low quality scores, respectively. Barcode and RADtag rescue will correct errors that are within a set distance of a ‘true’ barcode or cut site (Catchen et al., 2013). After demultiplexing, reads were further trimmed for quality using Trimmomatic (0.38) with the following settings: *slidingwindow:5:28* to ensure the average base PHRED quality never drops below 28 over a sliding window of 5 bp, *minlen:80* to remove reads shorter than 80 bp, and finally the standard recommended *leading:3* and *trailing:3* to remove low quality and ambiguous base calls at the leading and trailing ends of a read (Bolger et al., 2014). The Bowtie2 (2.4.2) aligner was used to first generate indexes for the reference genomes of *Bombus opulentus* (GenBank Accession: GCA_014737405.1) and *B. pyrosoma* (GCA_014825855.1), the closest available relatives to *B. pascuorum* and *B. lapidarius*, respectively, using default settings (Landmead and Salzberg, 2012; Sun et al., 2021). After index generation, reads were aligned to the genome using end-to-end mode under default settings. The genome alignments produced by Bowtie2 require further processing before they can be used for analysis. SAMtools (1.11) was used to filter out any unmapped reads using *view*, sort the files by their genome coordinates using *sort*, and create an index for each alignment (which helps speed up read time) using *index* (Li et al., 2009). Finally, Stacks was then used to identify the SNP variants present in the dataset, using the *ref_map* wrapper under default settings.

Quality control

VCFtools (0.1.16) was used to report the number of RAD tags (as called by Stacks *ref_map*) and average read depth for each sample using *-depth* (Danacek et al., 2011). *B. pascuorum* and *B. lapidarius* samples with fewer than 100,000 or 90,000 sites respectively, or an average depth below 10x, were removed. VCFtools was also used to check for heterozygosity using *-het*; samples with an inbreeding coefficient (F_{IS}) higher than 0.5 were removed. Additionally, VCFtools was used to generate a list of loci with a maximum mean depth of twice the average read depth and maximum proportion of missing data of 25%. The maximum read depth is capped at twice the average to prevent paralogous regions being counted twice. Loci that failed either of these thresholds were added to a blacklist.

Species identification and kinship analysis

After QC, libraries were initially mapped to the *B. terrestris* reference genome using Bowtie2, and run through Stacks with no



FIGURE 3

Example of sampling sites, as seen using satellite imagery (Map data: Google Maps 2022, Maxar Technologies). A set of sites in Bulgaria: (A) shows an area of intensive farmland, close to the south of Devene, categorised as 'high agriculture'; (B) shows an 'intermediate' area of mixed natural and agricultural land, to the west of Etropole, while (C) shows the heavily forested 'low agriculture' area to the south of Boykovets. Scale bar represents 500 m.

blacklist, to act as a basis for species determination (Sadd et al., 2015). A combination of approaches were used to determine species, as initial identifications were not always performed by experts in the field, which can cause misidentification due to visual similarity, e.g. *B. lapidarius* being highly similar in appearance to *B. ruderarius* and *B. pascuorum* being similar to *B. humilis* and *B. muscorum* (Rasmont et al., 2021). Bioinformatic methods of species identification relied primarily on the inference of admixture coefficients in R-package sNMF run from $K = 1:10$ with 10 reps, followed by cross-entropy and admixture visualisation to identify the number of groups within the population, with additional use of relatedness indices calculated by VCFtools, *-relatedness* and *-relatedness2*, to highlight outliers. *B. pascuorum* with a *-relatedness* or *-relatedness2* result of below -0.4 were considered off-target species. For *B. lapidarius*, the thresholds to be considered off-target species were below -0.3 for *-relatedness* and below -0.5 for *-relatedness2*.

Relatives were identified by grouping samples by natural population using VCFtools *--relatedness* and *-relatedness2*. For both species, samples were considered siblings if they had a *--relatedness* result of above 0, or a *-relatedness2* result of above 0.1. Standard thresholds (Jackson et al., 2018) were made slightly stricter to avoid second-degree relationships (Manichaikul et al., 2010).

Environmental data

CORINE land cover proportion was calculated for a 0.5 km radius around each sampling site using a custom R-script

containing the following packages: rgdal 1.5–23 (Bivand et al., 2021), raster 3.4–13 (Hijmans, 2020), plyr 1.8.6 (Wickham, 2011), dplyr 1.0.7 (Wickham et al., 2021), sf 1.0–2 (Pebesma, 2018), fasterize 1.0.3 (Ross, 2020), and data.table 1.14.0 (Dowle and Srinivasan, 2021). A radius of 0.5 km was chosen since estimates of bumblebee foraging range average at around 300m, with a maximum of 600–800 m, although this does vary with species and the method used (Osborne et al., 1999; Wolf and Moritz, 2008). In particular, the mean worker foraging distances differs significantly between species: *B. pascuorum* exhibited significantly smaller mean foraging distances (average 272 m) than *B. lapidarius* (average 536 m) (Redhead et al., 2016).

CORINE data, generated in 2018, classifies land as agricultural or natural based on usage, e.g. farmland, pasture, and forest. The proportion of each sampling site belonging to agricultural, urban, or natural (after water was removed) was calculated. Sampling sites were arbitrary ordered into three categories; $<33.3\%$ (= low agricultural), $33.3\text{--}66.6\%$ (= intermediate); and $>66.6\%$ (= high agricultural). Figure 3 illustrates the terrain of the high agricultural (Figure 3A), intermediate (Figure 3B), and the low agricultural (Figure 3C) categories. As this analysis was performed after the sampling - many of the paired sampling locations did not hold up after CORINE categorization. Analysis is not based on the initial pairs, only on the categorization following the CORINE data.

Specimens from neighbouring sites (<7 km) were grouped into populations. Populations were a minimum of 20 km

TABLE 2 Grouping of all remaining specimens (after filtering) into populations, their land cover categorisation, and the number of specimens in each population, per species.

Population	Land cover category	<i>B. pascuorum</i>	<i>B. lapidarius</i>
Belgium north	High agriculture	7	5
Belgium south	High agriculture	4	4
Bulgaria north	High agriculture	3	0
Bulgaria south	Low agriculture	3	0
Czechia north	High agriculture	6	9
Czechia south	Intermediate	3	0
England Huggate	High agriculture	4	9
England Moors	Low agriculture	4	0
Estonia perressare	Low agriculture	3	0
Estonia Vinni	Intermediate	3	4
Italy rural	High agriculture	3	0
Italy wilds	Intermediate	0	3
Netherlands dog park	High agriculture	3	3
Netherlands Leiden	High agriculture	4	3
Netherlands Texel	Intermediate	6	3
Portugal	Low agriculture	0	4
Spain mountain	Low agriculture	5	0
Sweden forest	Low agriculture	6	0
Sweden uppland	Intermediate	4	3
Ukraine Kiev	Intermediate	0	5

apart. CORINE land cover proportions were also analysed within the 7 km radius. Only populations consisting of three or more samples were used for population genetic estimations. 20 such populations were identified, retaining 71 *B. pascuorum* (73.2% of samples) and 55 *B. lapidarius* (75.3%) (Table 2). The thresholds for division were based on the flight distances of mated queens; *B. pascuorum* and *B. lapidarius* queens fly up to three and 5 km respectively (Lepais et al., 2010), but not more than 10 km (Kraus et al., 2008).

Population structure and isolation by distance

To explore possible interference of structure and isolation across populations, we tested for significant relationships between geographic distance and genetic divergence by Mantel testing (Mantel, 1967). Mantel testing was performed using R-package ape4 (Dray and Dufour, 2007) on the bumblebee populations described above. Mantel testing attempts to derive the strength and direction of correlation between two matrices; in our case, a matrix of pairwise genetic distances (F_{ST})—generated by Stacks, linearized using the formula ($F_{ST}/1-F_{ST}$), and haversine geographic distance between populations, calculated using R-package geosphere (Hijmans, 2019). The number of repetitions was set to 10,000.

Demographic history

To infer demographic history, *dadi* version (2.1.1) was used to generate folded Site Frequency Spectra (SFS) (Gutenkunst et al., 2009). Given the absence of population structure (see the results on population structuring), we decided to here group the samples of each species into one cluster each. The VCF-file generated by Stacks:ref_map was converted into an SFS, folded, and used as a basis for 1D-demographic models of ‘growth,’ ‘bottlegrowth,’ ‘two_epoch,’ and ‘standard neutral model,’ using 100 optimisation runs. Bootstrapping was used to estimate confidence intervals for each model parameter; AIC was calculated to determine which model best fitted the available data; Likelihood Ratio Tests (LRTs) were used to compare the neutral model with the other models.

Scans for loci under selective pressure

Three different methods were used to detect loci under selective pressure and were performed on all specimens that passed the above described QC-steps. The first genome scan method we used was BayeScan (Foll and Gaggiotti, 2008). BayeScan uses differences in allele frequencies between discrete populations and the multinomial-Dirichlet model to find evidence of local adaptation. BayeScan (2.1) was run with

default settings. Secondly, BayeScEnv (a similar method to BayeScan) was used, which takes a gradient approach to populations and includes environmental variables in calculations to identify local adaptation (de Villemereuil and Gandaggiotti, 2015). BayeScEnv (1.1) was run with default settings except *-nbp* which was set to 10, *-pilot* which was set to 2000, and *-burnin* which was set to 10,000. The average proportion of agricultural land cover for each of the categories, normalised to between 1 and 0, was used as environmental variable. Finally, the *lfmm* function in R-package *LEA* was used to perform regression analysis on genomic data and environmental variables to associate allele frequencies with environmental factors (Frichot and Francois, 2015). Here, land cover data was inputted as a list of agricultural land cover proportion for each sample. *lfmm* was run with $K = 1$ (given the absence of population structure) and five repetitions. The z-scores of each run were combined, averaged (median) and adjusted for genomic inflation factor lambda; λ was set to 0.6 for *B. pascuorum* and 0.5 for *B. lapidarius* based on the histogram visualisation as directed by Francois's *LEA* tutorial. *p*-values were then adjusted for multiple testing using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995).

Annotation

As both genomes used for read alignment were published recently (Sun et al., 2021), levels of genome annotation were insufficient to provide detailed functional information on the loci highlighted by previous programs. We tried to bypass these shortcomings by using a custom R-script and the package *seqinr* (Charif and Lobry, 2007). The reference genome sequence for 100 bp either side of a highlighted SNP was extracted and saved into a data frame, and then investigated for matches in the NCBI database, using *megablast* (version 2.12.0) under default settings, filtering for matches to Arthropoda exclusively (Altschul et al., 1990). BLAST results were manually searched for 'informative' descriptions—i.e., the name of a specific gene or function, not just a hit to an unannotated stretch of genome.

Given that bumblebees are not model organisms, performing Gene Ontology (GO) analysis is a substantial challenge. However, using BioDBnet's *dbReport* function, it was possible to perform an elementary analysis on the available data (Mudunuri et al., 2009). Gene IDs (e.g., LOC100749465) were extracted from the BLAST results for each species and processed with *dpReport*, which searches several genomic databases and returns all available information for a given identifier. From these results, the GO data for each locus was retrieved, covering the three standard GO categories, "cellular component", "biological process," and "molecular function" (Ashburner et al., 2000).

To determine the likelihood of non-synonymous mutation in the highlighted SNPs, each unique locus with an informative

TABLE 3 Per species locus counts for loci generated and filtered by Stacks, and blacklisted by VCFTools.

	<i>B. pascuorum</i>	<i>B. lapidarius</i>
Raw locus count	115,199	151,037
VCFTools blacklist	62,054	51,433
Percent blacklisted	53.9%	34.1%
Stacks internal filter	39,156	85,624
Percent filtered by Stacks	34.0%	56.7%
Final locus count	13,989	13,980
Percent retained from raw	12.1%	9.3%

BLAST result was investigated to determine whether the difference in major and minor alleles would have a functional consequence. For each locus, the reference sequence was extracted, but the SNP location in the reference sequence was replaced with the major and minor allele from our own sequencing data, as the reference genome and our study species sequences may differ. The major allele sequence was BLASTed (version 2.13.0), the results filtered for matches to Arthropods, and searched for the highest quality informative match. If the query sequence was matched to a coding region (CDS), ExPASy's web 'translate' tool (Gasteiger et al., 2003) was used to determine the open reading frame (ORF) of the sequences, and whether there was a difference in amino acid sequence between the major and minor alleles.

Results

Sequencing

118 *B. pascuorum* and 88 *B. lapidarius* candidates were distributed across five lanes of Illumina HiSeq 4000 single-end sequencing. A total of 1.67 billion reads were produced across all plates with an average of 334 million per plate. After demultiplexing and initial low-quality filtering, 92.8% of reads were retained, with 4.28% dropped due to ambiguous barcodes, 2.79% dropped due to ambiguous RAD-tags, and 0.08% dropped due to low quality, leaving 1.55 billion reads remaining.

Pre-processing, QC and kinship analysis

Individual sample libraries underwent further processing with TRIMMOMATIC. The average sample library input 3.24 million reads, and retained 2.78 million of them (85.5%). Initial species identification, found 118 candidate *B. pascuorum* and 88 candidate *B. lapidarius*, which were then re-mapped to the genomes of *B. opulentus* and *B. pyrosoma*, respectively, and resulted in an average read retention of 91.4%, and 79.2%,

TABLE 4 Demographic model Log Likelihood (LL) and Akaike Information Criterion (AIC) generated by *dadi*. The best-fitting model for each species has been highlighted in bold. For all models, the parameter T refers to the time at which change began, in units of $1.5 N_a$ (ancestral population size) generations. Nu refers to the ratio of modern to ancient population size; NuB is the ratio of the bottleneck population size to ancient population size; NuF is also the ratio of modern to ancient population size.

Species	Model	Log likelihood	AIC	Best fit parameters
<i>B. pascuorum</i>	Neutral model	-6313.55	—	—
	Growth	-259.12	522.24	Nu = 22.43, T = 0.55
	Bottlegrowth	-259.21	524.42	NuB = 1.04, NuF = 22.51, T = 0.53
	Two Epoch	-374.74	753.48	Nu = 12.44, T = 0.29
<i>B. lapidarius</i>	Neutral model	-6386.24	—	—
	Growth	-176.32	356.64	Nu = 18.94, T = 0.59
	Bottlegrowth	-176.31	358.62	NuB = 1.01, NuF = 18.99, T = 0.58
	Two epoch	-260.75	525.50	Nu = 10.74, T = 0.33

respectively. After the remapping, more specimens were removed due to several additional QC steps (see [Supplementary Table S3](#)), leaving 97 *B. pascuorum* and 73 *B. lapidarius* samples remaining.

Loci filtering

Stacks identified 115,199 loci in the *B. pascuorum* and 151,037 loci in the *B. lapidarius* samples, with a mean number of 131.4 and 129.3 sites per locus respectively ([Table 3](#)). Effective per-sample coverage was 47.7x in *B. pascuorum* and 35.4x in *B. lapidarius*. After several filtering steps a total of 13,989 loci in *B. pascuorum* (12.14%) and 13,980 in *B. lapidarius* (9.26%) remained for analysis ([Table 3](#)).

Land cover categories

Samples were separated into three categories based on the proportion of 'natural' land cover at their collection site, based on CORINE analysis ([Supplementary Table S4](#)). The *B. pascuorum* dataset contained 62 samples in the 'high agriculture' category, 19 samples from the 'low agriculture', and 16 from the 'intermediate' category. In the *B. lapidarius* dataset, 52 samples were classed as 'high agriculture', 10 as 'low agriculture', and 11 as 'intermediate' ([Supplementary Table S4](#)).

Demographic history

Using *dadi* to test several models of demographic history, and comparing their AIC for goodness-of-fit purposes, we found that in *B. pascuorum*, the 'growth' model (AIC = 522.24) fit our data the best followed very closely by 'bottlegrowth' (AIC = 524.42; [Table 4](#)). In *B. lapidarius*, results were similar, with 'bottlegrowth' being the best fitted model (AIC = 358.62), followed closely by 'growth' (AIC = 356.64; [Table 4](#)). All three growth-based models

TABLE 5 Mantel test results for both species, comparing geographic distance (Haversine distance in km) against genetic distance (F_{ST} calculated by Stacks).

Species	r	P	Std. Obs	Expectation	Variance
<i>B. pascuorum</i>	0.353	0.004	23067	-0.0028	0.0238
<i>B. lapidarius</i>	0.392	0.058	16064	0.0002	0.0595

fit the data better than the neutral model, in both species ($p < 0.001$). Models were rerun after filtering out outlier loci identified with BayeScan, BayeScEnv and LEA (see below) but were not meaningfully different ([Supplementary Table S5](#)). Populations of *B. pascuorum* increased in sizes about 70,600 years ago, while population growth for *B. lapidarius* started 98,400 years ago ([Supplementary Table S6](#)).

Population structure

In *B. pascuorum*, the average F_{ST} between populations was 0.103, ranging from 0.174 (Estonia Vinni—Italy Rural) to 0.058 (Belgium North—Czechia North) ([Supplementary Table S7](#)). In *B. lapidarius*, the F_{ST} average was 0.120, with a maximum of 0.224 (Sweden Uppland—Portugal) and a minimum of 0.053 (Ukraine—Czechia North) ([Supplementary Table S8](#)). A Mantel test was performed to compare geographic distance and genetic distance ([Table 5](#)) and with outlier loci removed ([Supplementary Table S9](#)). In *B. pascuorum*, the correlation between geographic and genetic distance was moderate, $r = 0.35$, and significant, $p = 0.004$. In *B. lapidarius* the correlation was also moderate ($r = 0.39$) but not significant ($p = 0.058$). sNMF clustering analysis was also performed on the full dataset (including samples that did not fall into populations) and found $K = 1$ as the most likely scenario for both *B. pascuorum* and *B. lapidarius* ([Supplementary Table S10](#)). Overall these results indicate a

general lack of population structure across Europe for both species, with mild isolation by distance in *B. pascuorum*.

Outlier detection

Across all programs, we detected 191 unique loci under selective pressure in the *B. pascuorum* specimens, and 260 unique loci in the *B. lapidarius* specimens (Supplementary Tables S11, S12). LEA detected 179 loci in *B. pascuorum*, and 196 in *B. lapidarius*, BayeScan found 12 and 55 loci, and BayeScEnv three and 16, respectively. In *B. pascuorum*, all three loci detected by BayeScEnv were also highlighted by BayeScan, while in *B. lapidarius*, only seven were common to both programs. No overlap was found between LEA and BayeScan or BayeScEnv.

Annotation results

To better understand the function of the genes under selective pressure, sequences associated with each SNP highlighted by BayeScan, BayeScEnv, and LEA were BLASTed for similar sequences (Supplementary Tables S13, S14). The majority of hits were matching uninformative sequences, so results were manually studied for functional information. *B. pascuorum* had informative hits for 26 or 13.6% of loci, while *B. lapidarius* had hits for 53 loci or 19.2% (Supplementary Tables S13, S14).

In terms of GO analysis, very few queries (loci with gene IDs such as LOC100749465; unique loci may have multiple queries) returned terms (Supplementary Table S15). For *B. pascuorum*, 26 unique loci with informative results lead to 97 input queries, of which 13 (13.4%) returned “Cellular Component” descriptors, 6 (6.2%) returned “Biological Process” descriptors, and 16 (16.5%) returned “Molecular Function” descriptors (Supplementary Table S15). The most common “Cellular Component” category was GO:0005634 (nucleus); three “Biological Process” categories showed equal distribution (GO:0006629 lipid metabolic process, GO:0006397 mRNA processing, GO:0006139 nucleobase-containing metabolic process); and “Molecular Function” had three categories tie for most common with GO:0003677 DNA binding, GO:0005524 ATP binding, and GO:0003723 RNA binding. For *B. lapidarius*, 53 unique loci with informative results generating in 337 queries, of which 36 (10.7%) had “Cellular Component” descriptors, 26 (7.7%) had “Biological Process” descriptors, and 42 (12.5%) had “Molecular Function” descriptors (Supplementary Table S15). The most common “Cellular Component” category was GO:0016021 integral component of membrane; for “Biological Process”, GO:0005975 carbohydrate metabolism and GO:0051301 cell division were both most common; Five

categories for “Molecular Function” showed equally high prevalence: GO:0005524 ATP binding, GO:0003677 DNA binding, GO:0046872 Metal ion binding, GO:0046983 Protein dimerisation activity, and GO:0003700 Sequence-specific DNA binding/transcription factor activity.

The same unique loci were studied to find evidence of non-synonymous mutations. In *B. pascuorum*, 19.23% of the loci investigated were found to have non-synonymous differences between the major and minor alleles present in the sequencing data; for *B. lapidarius*, the percentage was 26.42 (Supplementary Table S16). Of the non-synonymous SNPs in coding regions, several associated genes with interesting functions were identified (Supplementary Table S16); In *B. pascuorum*, transcriptional regulators (such as an ATRX homolog), lipid metabolism-related enzymes (such as triacylglycerol lipase), titin (an important muscle structure protein), and neurodevelopment proteins (such as notch) were prominent. In *B. lapidarius*, genes highlighted include several neurodevelopment and behaviour-related proteins (such as ataxin, SOX, and tyrosine-protein phosphatase 99A-like), as well as transcription factors (Forkhead transcription factor FD4), and TipE, a sodium-channel structural protein with strong associations with insecticide resistance.

Discussion

Demographic history

From our demographic analysis we found that the ‘growth’ and ‘bottlegrowth’ were the best models for our data of *B. pascuorum* and *B. lapidarius*, respectively. In both cases, the other growth model—i.e. ‘bottlegrowth’ for *B. pascuorum* and ‘growth’ for *B. lapidarius*—also fit the data incredibly closely and was not far behind in terms of log likelihood and AIC. This pattern is similar to that of another study detecting clear signals of past population growth for *B. lapidarius*; with ‘growth’ as best fitting model, followed closely by ‘bottlegrowth’ (Theodorou et al., 2018). Studies focused on understanding the evolution and demographic history of both *B. pascuorum* and *B. lapidarius* indicate that both species likely retreated to various ‘refugia’ across Europe during the last glacial period (Pirounakis et al., 1998; Widmer and Schmid-Hempel, 1999; Lecocq et al., 2013). The last glacial period is estimated to have taken place between 110,000–12,500 years ago; our own calculations suggest that growth began approximately 70,600 years ago for *B. pascuorum* and 98,400 years ago for *B. lapidarius*. Overall, our results are congruent with the idea that both species had smaller populations, likely during the ice age as other studies have indicated, and have since grown considerably (Theodorou et al., 2018).

Population structure and isolation by distance

Our sNMF results indicated a lack of significant population structure across Europe for both species. This is in line with previous research using microsatellite markers detecting no or slight population structuring for the bumblebees on the local and even continental scale (Dreier et al., 2014; Maebe et al., 2015). The only population of *B. pascuorum* in our dataset that occurred beneath the Alps (Italy_wilds) showed much higher F_{ST} -values compared to the other pairwise estimates. This result might correspond to the two described gene pools of *B. pascuorum* that are separated by the Alps as suggested by Widmer and Schmid-Hempel (1999). Of the 4 *B. lapidarius* subspecies that are described in Europe, two are relevant to our sampling; *B.l. lapidarius* from the European mainland, and *B.l. decipiens*, from the Iberian Peninsula (Lecocq et al., 2013). Although sNMF did not suggest a second population, the pairwise F_{ST} of samples from Portugal was considerably higher than average (0.111 without Portuguese samples), suggesting a pattern of differentiation similar to the Iberian subspecies (Lecocq et al., 2013, 2020). Mantel testing results found mild but significant IBD in *B. pascuorum* but not in *B. lapidarius*, implying strong gene flow between populations, and agreeing with the sNMF results. Dreier et al. (2014) found similar results using microsatellite markers; weak but significant IBD in *B. pascuorum* and insignificant results in *B. lapidarius*. Theodorou et al. (2018) also found a lack of IBD in *B. lapidarius*, which contrasts with the weak but significant IBD across Europe described for *B. lapidarius* (Lecocq et al., 2013, 2017; Theodorou et al., 2018). Overall, these mixed results paint a similar picture; weak IBD is possibly present in both species but strength and presence vary by sampling and method.

Local adaptation

By comparing bumblebees from natural, agricultural, and intermediate areas we were able to find genes associated to agriculture over a large geographic range within two bumblebee species. Furthermore, we were able to exclude the population differentiation from neutral population structuring or historical geographic isolation. In general, this main result implies that bees of agricultural areas differ genetically from bees from more natural areas. That bees of agricultural areas tend to experience different strengths of certain pressures than bees from natural areas would be logical given the difference in, among others things, pesticide use,

habitat fragmentation, and availability of food resources between these different habitats (e.g., Goulson, 2010; Cameron and Sadd, 2020). However, in our study, we detected no population structuring at putatively neutral loci for both species across a large part of their distribution in Europe. With this in mind, one might think that due to the likely high levels of gene flow supporting the lack of differentiation on a continental scale, the observed differences may not appear between bees from agricultural and natural locations, and especially not between locations within such close proximities (<50 km). However, that we were still able to detect outlier loci here, suggests that these loci have been selected and maintained in agricultural areas.

Although we were able to associate genes with agricultural stressors, we cannot exclude that these genes might also be correlated with other causes of (anthropogenic) stress, such as: urbanization. Indeed, urbanization can stress bees in a similar way to agriculture. This includes heat stress due to the heat island effect (e.g., Chapman et al., 2017; Levermore et al., 2018), nutritional stress due to decrease in floral diversity, disturbance of bee foraging patterns, and increased impervious surface and habitat fragmentation (e.g., Fortel et al., 2014; Zhao et al., 2016), and chemical stress through all kinds of environmental pollution (e.g., Hladun et al., 2015; Sivakoff and Gardiner, 2017; Grubisic et al., 2018). Furthermore, previous research found several of the same gene ontology categories as outlier loci being correlated to urbanization (Theodorou et al., 2018) or to a bioclimatic landscape (Jackson et al., 2018, 2020). Further work/analysis needs to be done on other features of anthropogenic land use to exclude them as contributors to the patterns we observed.

Primary candidate genes

Based on the BLAST search, we were able to link some of the detected candidate genes to ongoing agricultural stressors. In both *B. pascuorum* and *B. lapidarius*, we detected several SNPs in the coding regions of genes (some synonymous and some not) linked with neurodevelopment, transcriptional regulation, and metabolism, as well as genes related to muscle development and lipid metabolism in *B. pascuorum*, and genes implicated in detoxification and cell signalling in *B. lapidarius* (Supplementary Tables S13-S16). To us, these would be the primary candidate genes for further investigation or validation. Although the presence of non-synonymous differences between the major and minor alleles provides the strongest evidence of selection, silent mutations, such as those we detected in titin in *B. pascuorum* and TipE in *B. lapidarius* should not be discounted; synonymous mutations can, for example,

greatly affect codon usage/bias, mRNA splicing, and mRNA stability, with tangible consequences for gene expression (Hunt et al., 2014). The presence of SNPs in muscle-related and lipid-metabolism related genes could be particularly interesting as several studies found a strong correlation between levels of agricultural intensification and larger body size in bumblebees (Gérard et al., 2020, 2021; Nooten and Rehan, 2020). The association of genes which may relate to body size changes in our genetic data may indicate corroboration of these findings. One hypothesis is that larger-bodied bumblebees have the ability to forage at greater distances, thus overcoming habitat fragmentation (but see Gérard et al., 2020; 2021); the detection of selective pressure on titin in *B. pascuorum* may be related to muscle changes resulting from the need to fly greater distances, also related to habitat fragmentation; further study is necessary to find empirical evidence to support this. Both species had genes highlighted which are related to the nervous system and neurodevelopment, such as ataxin and notch, which may have some significance in localised behaviour, such as optimising landscape-specific foraging patterns. Ataxin has been related to circadian rhythms in *D. melanogaster* (Lee et al., 2017); Notch is involved in a range of neurogenic and regulatory processes, also in *D. melanogaster* (Lehmann et al., 1983; Xu et al., 1990; Lai, 2004). Together these may add up to complex behavioural changes in response to the environment. Additionally, in *B. lapidarius* we found associations with the heat stress and detoxification-related gene TipE. TipE, a sodium-channel structural protein has been linked to pyrethroid resistance (Warmke et al., 1997; Derst et al., 2006; Silver et al., 2014). Heat stress related genes were found to be upregulated in the face of pesticide exposure in *Spodoptera* and *Nilaparvata* (You et al., 2016; Tarnawska et al., 2019), and also identified as under selective pressure across an urbanisation gradient (Theodorou et al., 2018). This result might imply the need for *B. lapidarius* to adapt to temperature differences in natural and agricultural landscapes, or an alternative mechanism for xenobiotic resistance. Given the wide disparity in expected use of pesticides between natural and agricultural areas, it is not surprising that detoxification-related genes could be associated with agricultural-based selective pressure.

Genome assemblies and annotation

At the time of analysis, no reference genomes for *B. pascuorum* and *B. lapidarius* were available, only *de novo* assembled genomes for relatives in the same subgenus (*Thoracobombus* and *Melanobombus*, respectively) (Sun et al., 2021). While alignment of reads to closer relatives than the initial mapping to *B. terrestris* resulted in a considerable increase in the proportion of reads that mapped to the genome and number of loci detected, bumblebee genomes have been shown to have considerable rearrangements and even differing chromosome

numbers within subgenera (Owen et al., 1995; Sun et al., 2021). Thus, having species-specific high-quality reference genomes may allow for the detection of loci which are not present in close relatives. Similarly, in the absence of manual and comprehensive annotation data, many loci in exons detected in outlier analysis will be overlooked in favour of loci with concrete functional labels. Indeed, a high proportion of SNPs were present in non-coding or unannotated regions (~81% in both species). While high-quality, curated annotation data is available for several *Bombus* species, the genus is still relatively understudied compared to other model insect species such as *Drosophila melanogaster*, and consequently the annotation data is comparatively more limited. Consequently, our Gene Ontology investigations suffer from the same data limitations. With incomplete genome annotation and no available GO annotation services for *Bombus* or better-studied close relative *Apis mellifera*, the results of the limited GO analysis were inconclusive at best, with seemingly no obvious patterns. Likewise, Gene Set Enrichment Analysis (GSEA), which looks for over- and under-representation of functions in lists of genes, cannot be performed without a reference dataset. With further study of *Bombus* and related species, sufficient GO annotation and GSEA databases will allow for understanding overall patterns and themes present in such datasets with statistical support. Because of our reduced-representation approach, certainly not all nor the strongest associations of genomic regions and loci related to agricultural stress were detected for both species. Although retrieving the genome of both bumblebee species and additional annotation data would have increased the power of this, and will increase that of other future genomic studies, they were not necessary for the outcome of our study. Indeed, we were able to reach the major target of this study which was to find any loci as signals of adaptation to agricultural stress, in both bumblebee species.

Possible caveats

In this study, we explored and ruled out population differentiation and demographic history as interfering factors when testing for signatures of selection. Most outlier tests will identify loci as under selection when having higher (divergent selection) or lower (balancing selection) F_{ST} -values than expected under the neutral model (Foll and Gaggiotti, 2008; de Villemereuil et al., 2014; Lotterhos and Whitlock, 2014). When significant population structuring or isolation by distance is present, this will conflict with the assumption of such neutral model (Foll and Gaggiotti, 2008; de Villemereuil et al., 2014; Lotterhos and Whitlock, 2014). Therefore, it is crucial to identify a species genetic structure to account for possible risk of identifying false positives or negatives when attempting to identify adaptive loci (as reviewed by Ahrens et al., 2018). This is why most outlier loci detection approaches already require a

priori definition of genetic structure to inform calculations and probability tests (e.g., de Villemereuil and Gaggiotti, 2015; Frichot and Francois, 2015; Ahrens et al., 2018). In a similar manner, some approaches also account for species demographic history, such as population contraction or expansion, which can affect the species' genetic structure and may thus also result in false positives (Lotterhos and Whitlock, 2014). In our study, we combined three different programs, which are mainly based on two different approaches: two F_{ST} -based methods, one with and one without a direct association to an environmental parameter (BayeScEnv; de Villemereuil and Gaggiotti, 2015 and BayeScan; Foll & Gaggiotti, 2008, respectively), and an approach using latent factor mixed models (LEA) to test for the correlations between allele counts and an environmental variable while accounting for demographic structure using K latent factors (Frichot and Francois, 2015). Combining such outlier scanning approaches increases the chance of detecting candidate loci while reducing the false positive rate (François et al., 2016). Unfortunately, in our results, only a few loci were confirmed between the F_{ST} -based methods, while none were cross-validated with LEA. This does not mean that the detected loci are false-positives. The separate results of each of these outlier scanning approaches are useful and may differ due to the methods behind them, for instance, most of our identified loci were detected with LEA, which accounted for demographic structure. In addition, the credibility of the identified genes could be strengthened by some kind of validation, for instance by functional validation through RNAi knockout, or by performing comparative studies across bumblebee species, to determine whether these identified candidate genes are actually involved in local adaptation. This would be an important avenue for future research.

A second potential caveat to our study may be the rather limited population sizes ($n \geq 3$). While performing a population genetic study with populations represented by only three specimens was previously and usually considered too small, recent studies have shown that accurate population parameters such as F_{ST} -values or genetic diversity measurements can be obtained when a large number of SNPs are used to compensate for such low population sizes (Nazareno et al., 2017; McLaughlin and Winker, 2020). Here, by conducting a genomic study with approximately 14,000 SNPs per species after filtering, we met the latter criteria. Therefore, we do not consider population size to be a major issue in our study. In line with this, we would nevertheless like to emphasize that we could have chosen to RAD sequence more specimens per location, but that would have reduced the strength of our current sampling which focused on specimens sampled from multiple locations across a large geographic range in Europe.

Another possible caveat, which is related to previous one, is the difference in number of specimens per land cover category. The

overrepresentation of specimens in the high agricultural category might have biased the detection of outlier loci. However, and as stated before, the combined use of different methods and approaches to identify loci under environmentally-associated selection pressure should have not excluded but reduced the impact of such bias.

Importance for conservation and management

Understanding the precise consequences of each environmental stressor to bees would be key to preventing, mitigating, or reversing anthropogenic population declines. The aim of this study was therefore to improve our understanding of anthropogenic influences on European bumblebees by searching for genomic evidence of adaptation. Such an understanding of the loci under selection and their functions may help pinpoint the exact variations in such populations which convey advantages. The preservation of such adaptive genomic variation might then be key for resilience to anthropogenic stressors, as was discussed by Jackson et al. (2020). Furthermore, understanding which traits are necessary for positive population trends in specific environmental conditions will provide insights into which populations are more or less vulnerable to anthropogenic stressors. Thus, identifying whether genetic adaptations are present or absent within populations will allow conservationists to allocate attention and conservation resources accordingly. This is particularly important for ecological and economical key pollinators like bumblebees (Klein et al., 2007; Garibaldi et al., 2013).

Conclusion

Through the RADSeq of 97 *B. pascuorum* and 73 *B. lapidarius*, from sites across a European gradient of agricultural-to-natural landscapes, we were able to discover signs of local adaptation and selective pressure exerted by associated stressors. Through the use of BLAST, we associated several candidate genes involved in neurological development, muscle structure, and detoxification, among others. Further study, such as whole genome sequencing, using better genome annotation, or functional studies, would first allow for the validation of these candidate genes, and when confirmed, will also help: 1) to define the identity of the drivers of selection, 2) to understand precisely how variance in these genes conveys adaptation to different environments, and 2) how this information be applied for conservation purposes.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://doi.org/10.5061/dryad.00000005j>, <https://github.com/AlexFHart/RADSeq-Pipeline>.

Author contributions

KM and IM conceived and designed the experiments; AH and JV performed the genomic analysis; DA performed the land cover analysis; AH, KM, DC, HH, VR, JS, TL, ToL, JD-F, SF, LB, and RK collected the specimens; GS provided the resources; AH and KM interpreted the data and drafted the manuscript. All authors provided intellectual feedback on the manuscript and approved the final version of the manuscript.

Funding

This work was supported through funding from the Research Foundation—Flanders (FWO) as part of the FWO-research project [3G042618], and by the FWO and F.R.S.-FNRS under the Excellence of Science (EOS) programme for the project “Climate change and its impact on pollination services” (CLiPS, n°3094785). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Ahrens, C. W., Rymer, P. D., Stow, A., Bragg, J., Dillon, S., Umbers, K. D. L., et al. (2018). The search for loci under selection: Trends, biases and progress. *Mol. Ecol.* 27, 1342–1356. doi:10.1111/mec.14549
- Alford, D. V. (1975). *Bumblebees*. London: Davis-Poynter.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi:10.1016/S0022-2836(05)80360-2
- Anderson, N. L., and Harmon-Threatt, A. N. (2019). Chronic contact with realistic soil concentrations of imidacloprid affects the mass, immature development speed, and adult longevity of solitary bees. *Sci. Rep.* 9, 3724. doi:10.1038/s41598-019-40031-9
- Andrews, K. R., Gerritsen, A., Rashed, A., Crowder, D. W., Rondon, S. I., van Herk, W. G., et al. (2020). Wireworm (Coleoptera: Elateridae) genomic analysis reveals putative cryptic species, population structure, and adaptation to pest control. *Commun. Biol.* 3, 489. doi:10.1038/s42003-020-01169-9
- Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., and Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* 17, 81–92. doi:10.1038/nrg.2015.28
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, M. J., et al. (2000). Gene ontology: Tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* 25, 25–29. doi:10.1038/75556
- Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3, e3376. doi:10.1371/journal.pone.0003376
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57 (1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Biella, P., Četković, A., Gogala, A., Neumayer, J., Sárospatakis, M., Šima, P., et al. (2020). North-westward range expansion of the bumblebee *Bombus haematurus* into Central Europe is associated with warmer winters and niche conservatism. *Insect Sci.* 28, 861–872. doi:10.1111/1744-7917.12800
- Bivand, R., Keitt, T., and Rowlingson, B. (2021). rgdal: Bindings for the ‘Geospatial’ data abstraction library. R package version 1.5-23. Available at: <https://CRAN.R-project.org/package=rgdal>.
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170
- 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 bumblebees. *Proc. Natl. Acad. Sci. U. S. A.* 108, 662–667. doi:10.1073/pnas.1014743108
- Cameron, S. A., and Sadd, B. M. (2020). Global trends in bumblebee health. *Annu. Rev. Entomol.* 69, 209–232. doi:10.1146/annurev-ento-011118-111847
- Cameron, S., Hines, H. M., and Williams, P. H. (2007). A comprehensive phylogeny of the bumble bees (*Bombus*). *Biol. J. Linn. Soc. Lond.* 91, 161–188. doi:10.1111/j.1095-8312.2007.00784.x
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., and Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–3140. doi:10.1111/mec.12354
- Chapman, S., Watson, J. E. M., Salazar, A., Thatcher, M., and McAlpine, C. A. (2017). The impact of urbanization and climate change on urban temperatures: A systematic review. *Landsc. Ecol.* 32, 1921–1935. doi:10.1007/s10980-017-0561-4
- Charif, D., and Lobry, J. (2007). “SeqinR 1.0-2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis,” in *Structural approaches to sequence evolution: Molecules, networks,*

Acknowledgments

The authors would like to thank two reviewers for their constructive comments which helped to improve our manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.993416/full#supplementary-material>

- populations, series biological and medical physics, biomedical engineering. Editors U. Bastolla, M. Porto, H. Roman, and M. Vendruscolo (New York: Springer-Verlag), 207–232. 978-3-540-35305-8, 2007).
- Crutzen, P., and Stoermer, E. F. (2000). The "Anthropocene." *IGBP Newsl.* 41, 17–18. Available at: <http://www.igbp.net/download/18.316f18321323470177580001401/1376383088452/NL41.pdf#page=17>.
- Danacek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi:10.1093/bioinformatics/btr330
- Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., and Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12, 499–510. doi:10.1038/nrg3012
- de Villemereuil, P., Frichot, É., Bazin, É., François, O., and Gaggiotti, O. E. (2014). Genome scan methods against more complex models: When and how much should we trust them? *Mol. Ecol.* 23, 2006–2019. doi:10.1111/mec.12705
- de Villemereuil, P., and Gaggiotti, O. E. (2015). A new FST-based method to uncover local adaptation using environmental variables. *Methods Ecol. Evol.* 6, 1248–1258. doi:10.1111/2041-210X.12418
- Derst, C., Walther, C., Veh, R. W., Wicher, D., and Heinemann, S. H. (2006). Four novel sequences in *Drosophila melanogaster* homologous to the auxiliary Para sodium channel subunit TipE. *Biochem. Biophys. Res. Commun.* 339, 939–948. doi:10.1016/j.bbrc.2005.11.096
- Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., and Collen, B. (2014). Defaunation in the Anthropocene. *Science* 345, 401–406. doi:10.1126/science.1251817
- Dowel, M., and Srinivasan, A. (2021). data.table: Extension of 'data.frame'. R package version 1.14.0. Available at: <https://CRAN.R-project.org/package=data.table>.
- Dray, S., and Dufour, A. (2007). The ade4 package: Implementing the duality diagram for ecologists. *J. Stat. Softw.* 22 (4), 1–20. doi:10.18637/jss.v022.i04
- Dreier, S., Redhead, J. W., Warren, I. A., Bourke, A. F. G., Heard, M. S., Jordan, W. C., et al. (2014). Fine-scale spatial genetic structure of common and declining bumble bees across an agricultural landscape. *Mol. Ecol.* 23, 3384–3395. doi:10.1111/mec.12823
- Foll, M., and 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. doi:10.1534/genetics.108.092221
- Fortel, L., Henry, M., Guilbaud, L., Guirao, A. L., Kuhlmann, M., Mouret, H., et al. (2014). Decreasing abundance, increasing diversity and changing structure of the wild bee community (Hymenoptera: Anthophila) along an urbanization gradient. *PLoS ONE* 9 (8), e104679. doi:10.1371/journal.pone.0104679
- François, O., Martins, H., Caye, K., and Schoville, S. D. (2016). Controlling false discoveries in genome scans for selection. *Mol. Ecol.* 25, 454–469. doi:10.1111/mec.13513
- Frichot, E., and François, O. (2015). Lea: An R package for landscape and ecological association studies. *Methods Ecol. Evol.* 6, 925–929. doi:10.1111/2041-210X.12382
- Garibaldi, L. A., Steffan-Dewenter, I., Winfree, R., Aizen, M. A., Bommarco, C., Cunningham, S. A., et al. (2013). Wild pollinators enhance fruit set of crops regardless of honey bee abundance. *Science* 339, 1608–1611. doi:10.1126/science.1230200
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExpASY: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. doi:10.1093/nar/gkg563
- Gérard, M., Marshal, L., Martinet, B., and Michez, D. (2021). Impact of landscape fragmentation and climate change on body size variation of bumblebees during the last century. *Ecography* 44 (2), 255–264. doi:10.1111/ecog.05310
- Gérard, M., Martinet, B., Maebe, K., Marshall, L., Smagge, G., Vereecken, N. J., et al. (2020). Shift in size of bumblebee queens over the last century. *Glob. Chang. Biol.* 26, 1185–1195. doi:10.1111/gcb.14890
- Ghisbain, G., Gérard, M., Wood, T. J., Hines, H. M., and Michez, D. (2021). Expanding insect pollinators in the Anthropocene. *Biol. Rev. Camb. Philos. Soc.* 96, 2755–2770. doi:10.1111/brv.12777
- Goulson, D. (2010). *Bumblebees: Their behaviour and ecology*. 2nd edition. Oxford: Oxford University Press.
- Goulson, D., Hanley, M. E., Darvill, B., Ellis, J. S., and Knight, M. E. (2005). Causes of rarity in bumblebees. *Biol. Conserv.* 122, 1–8. doi:10.1016/j.biocon.2004.06.017
- Goulson, D., Lye, G. C., and Darvill, B. (2008). Decline and conservation of bumblebees. *Annu. Rev. Entomol.* 53, 191–208. doi:10.1146/annurev.ento.53.103106.093454
- Goulson, D., Nicholls, E., Botías, C., and Rotheray, E. L. (2015). Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. *Science* 347 (6229), 1255957. doi:10.1126/science.1255957
- Graystock, P., Yates, K., Evison, S., Darvill, B., Goulson, D., and Hughes, W. (2013). The trojan hives: Pollinator pathogens, imported and distributed in bumblebee colonies. *J. Appl. Ecol.* 50, 1207–1215. doi:10.1111/1365-2664.12134
- Grubisic, M., van Grunsven, R., Kyba, C., Manfrin, A., and Hölker, F. (2018). Insect declines and agroecosystems: Does light pollution matter? *Ann. Appl. Biol.* 173, 180–189. doi:10.1111/aab.12440
- Gunstone, T., Cornelisse, T., Klein, K., Dubey, A., and Donley, N. (2021). Pesticides and soil invertebrates: A hazard assessment. *Front. Environ. Sci.* 9, 643847. doi:10.3389/fevs.2021.643847
- Gutenkunst, R., Hernandez, R., Williamson, S., and Bustamante, C. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* 5 (10), e1000695. doi:10.1371/journal.pgen.1000695
- Hereward, J. P., Walter, G. H., DeBarro, A. J., Lowe, A. J., and Riginos, C. (2013). Gene flow in the green mirid, *Creontiades dilutus* (Hemiptera: Miridae), across arid and agricultural environments with different host plant species. *Ecol. Evol.* 3, 807–821. doi:10.1002/ece3.510
- Hijmans, R. J. (2019). geosphere: Spherical trigonometry. R package version 1.5-10. Available at: <https://CRAN.R-project.org/package=geosphere>.
- Hijmans, R. J. (2020). raster: Geographic data analysis and modeling. R package version 3.4-13 Available at: <https://CRAN.R-project.org/package=raster>.
- Hladun, K. R., Parker, D. R., and Trumble, J. T. (2015). Cadmium, copper, and lead accumulation and bioconcentration in the vegetative and reproductive organs of *Raphanus sativus*: Implications for plant performance and pollination. *J. Chem. Ecol.* 41, 386–395. doi:10.1007/s10886-015-0569-7
- Hunt, R. C., Simhadri, V. L., Iandoli, M., Sauna, Z. E., and Kimchi-Sarfaty, C. (2014). Exposing synonymous mutations. *Trends Genet.* 30, 308–321. doi:10.1016/j.tig.2014.04.006
- 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 bumblebees across latitude and altitude. *Mol. Ecol.* 27, 2926–2942. doi:10.1111/mec.14735
- Jackson, J., Pimsler, M., Oyen, K., Strange, J., Dillon, M., and Lozier, J. (2020). Local adaptation across a complex bioclimatic landscape in two montane bumble bee species. *Mol. Ecol.* 29, 920–939. doi:10.1111/mec.15376
- Jha, S. (2015). Contemporary human-altered landscapes and oceanic barriers reduce bumblebee gene flow. *Mol. Ecol.* 24, 993–1006. doi:10.1111/mec.13090
- Jha, S., and Kremen, C. (2013). Urban land use limits regional bumble bee gene flow. *Mol. Ecol.* 22, 2483–2495. doi:10.1111/mec.12275
- Kerr, J. T., Pindar, A., Galpern, P., Packer, L., Potts, S. G., Roberts, S. M., et al. (2015). CLIMATE CHANGE. Climate change impacts on bumblebees converge across continents. *Science* 349, 177–180. doi:10.1126/science.aaa7031
- Klein, A. M., Vaissière, Cane, B. E., J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., Tscharntke, T., et al. (2007). Importance of pollinators in changing landscapes for world crops. *Proc. Biol. Sci.* 274, 303–313. doi:10.1098/rspb.2006.3721
- Kosior, A., Celary, W., Olejniczak, P., Fijał, J., Król, W., Solarz, W., et al. (2007). The decline of the bumblebees and cuckoo bees (hymenoptera: Apidae: Bombini) of western and central Europe. *Oryx* 41, 79–88. doi:10.1017/S0030605307001597
- Kraus, F., Wolf, S., and Moritz, R. (2008). Male flight distance and population substructure in the bumblebee *Bombus terrestris*. *J. Anim. Ecol.* 78 (1), 247–252. doi:10.1111/j.1365-2656.2008.01479.x
- Lai, E. C. (2004). Notch signaling: Control of cell communication and cell fate. *Development* 131, 965–973. doi:10.1242/dev.01074
- Langmead, B., and Salzberg, S. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi:10.1038/nmeth.1923
- Lecocq, T., Gérard, M., Michez, D., and Dellicour, S. (2017). Conservation genetics of European bees: New insights from the continental scale. *Conserv. Genet.* 18, 585–596. doi:10.1007/s10592-016-0917-3
- Lecocq, T., Biella, P., Martinet, B., and Rasmont, P. (2020). Too strict or too loose? Integrative taxonomic assessment of *Bombus lapidarius* complex (hymenoptera: Apidae). *Zool. Scr.* 9, 187–196. doi:10.1111/zsc.12402
- Lecocq, T., Brasero, N., Martinet, B., Valterová, I., and Rasmont, P. (2015). Highly polytypic taxon complex: Interspecific and intraspecific integrative taxonomic assessment of the widespread pollinator *Bombus pascuorum* scopoli 1763 (hymenoptera: Apidae). *Syst. Entomol.* 40, 881–890. doi:10.1111/syen.12137
- Lecocq, T., Dellicour, S., Michez, D., Lhomme, P., Vanderplanck, M., Valterová, I., et al. (2013). Scent of a break-up: Phylogeography and reproductive trait

- divergences in the red-tailed bumblebee (*Bombus lapidarius*). *BMC Evol. Biol.* 13, 263. doi:10.1186/1471-2148-13-263
- Lee, J., Yoo, E., Lee, H., Park, K., Hur, J., and Lim, C. (2017). LSM12 and ME31B/DDX6 define distinct modes of posttranscriptional regulation by ATAXIN-2 protein complex in *Drosophila* circadian pacemaker neurons. *Mol. Cell* 66 (1), 129–140. doi:10.1016/j.molcel.2017.03.004
- Lehmann, R., Jiménez, F., Dietrich, U., and Campos-Ortega, J. A. (1983). On the phenotype and development of mutants of early neurogenesis in *Drosophila melanogaster*. *Wilehm. Roux. Arch. Dev. Biol.* 192, 62–74. doi:10.1007/BF00848482
- Lepais, O., Darvill, B., O'Connor, S., Osborne, J., Sanderson, R., Cussans, J., et al. (2010). Estimation of bumblebee queen dispersal distances using sibship reconstruction method. *Mol. Ecol.* 19, 819–831. doi:10.1111/j.1365-294X.2009.04500.x
- Levermore, G., Parkinson, J., Lee, K., Laycock, P., and Lindley, S. (2018). The increasing trend of the urban heat island intensity. *Urban Clim.* 24, 360–368. doi:10.1016/j.uclim.2017.02.004
- Lewis, S., and Maslin, M. (2015). Defining the Anthropocene. *Nature* 519, 171–180. doi:10.1038/nature14258
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25–20782079. doi:10.1093/bioinformatics/btp352
- Liczner, A. R., and Colla, S. R. (2019). A systematic review of the nesting and overwintering habitat of bumble bees globally. *J. Insect Conserv.* 23, 787–801. doi:10.1007/s10841-019-00173-7
- Lotterhos, K. E., and Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Mol. Ecol.* 23, 2178–2192. doi:10.1111/mec.12725
- Lozier, J. D., Jackson, J., Dillon, M., and Strange, J. (2016). Population genomics of divergence among extreme and intermediate color forms in a polymorphic insect. *Ecol. Evol.* 6, 1075–1091. doi:10.1002/ece3.1928
- Maebe, K., Hart, A. F., Marshall, L., Vandamme, P., Vereecken, N. J., Michez, D., et al. (2021). Bumblebee resilience to climate change, through plastic and adaptive responses. *Glob. Chang. Biol.* 27, 4223–4237. doi:10.1111/gcb.15751
- Maebe, K., Meeus, I., Ganne, M., De Meulemeester, T., Biesmeijer, K., and Smaghe, G. (2015). Microsatellite analysis of museum specimens reveals historical differences in genetic diversity between declining and more stable *Bombus* species. *PLoS ONE* 10, e0127870. doi:10.1371/journal.pone.0127870
- Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., and Chen, W. M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics* 26 (22), 2867–2873. doi:10.1093/bioinformatics/btq559
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27 (2), 209–220. PMID: 6018555.
- Marshall, L., Biesmeijer, J. C., Rasmont, P., Vereecken, N. J., Dvorak, L., Fitzpatrick, U., et al. (2018). The interplay of climate and land use change affects the distribution of EU bumblebees. *Glob. Chang. Biol.* 24, 101–116. doi:10.1111/gcb.13867
- Martinet, B., Zambra, E., Przybyla, K., Lecocq, T., Nonclercq, D., Rasmont, P., et al. (2021). Mating under climate change: Impact of simulated heatwaves on the reproduction of model pollinators. *Funct. Ecol.* 35 (3), 739–752. doi:10.1111/1365-2435.13738
- McLaughlin, J. F., and Winker, K. (2020). An empirical examination of sample size effects on population demographic estimates in birds using single nucleotide polymorphism (SNP) data. *PeerJ* 8, e9939. doi:10.7717/peerj.9939
- McNeil, D. J., McCormick, E., Heimann, A. C., Kammerer, M., Douglas, M. R., Goslee, S. C., et al. (2020). Bumblebees in landscapes with abundant floral resources have lower pathogen loads. *Sci. Rep.* 10, 22306. doi:10.1038/s41598-020-78119-2
- Meeus, I., Brown, M. J. F., De Graaf, D. C., and Smaghe, G. (2011). Effects of invasive parasites on bumblebee declines. *Conserv. Biol.* 25, 662–671. doi:10.1111/j.1523-1739.2011.01707.x
- Mudunuri, U., Che, A., Yi, M., and Stephens, R. M. (2009). bioDBnet: the biological database network. *Bioinformatics* 25, 555–556. doi:10.1093/bioinformatics/btn654
- Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., and Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Mol. Ecol.* 22, 2841–2847. doi:10.1111/mec.12350
- Nazareno, A. G., Bemmels, J. B., Dick, C. W., and Lohmann, L. G. (2017). Minimum sample sizes for population genomics: An empirical study from an amazonian plant species. *Mol. Ecol. Resour.* 17 (6), 1136–1147. doi:10.1111/1755-0998.12654
- Nielsen, R. (2001). Statistical tests of selective neutrality in the age of genomics. *Heredity* 86, 641–647. doi:10.1046/j.1365-2540.2001.00895.x
- Nieto, A., Roberts, S. P., Kemp, J., Rasmont, P., Kuhlmann, M., Criado, M. G., et al. (2014). *European red list of bees*. Luxembourg: Publication Office of the European Union, 98.
- Nooten, S. S., and Rehan, S. M. (2019). Historical changes in bumble bee body size and range shift of declining species. *Biodivers. Conserv.* 29, 451–467. doi:10.1007/s10531-019-01893-7
- O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., and Portnoy, D. S. (2018). These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. *Mol. Ecol.* 27, 3193–3206. doi:10.1111/mec.14792
- Oliver, S. V., and Brooke, B. D. (2018). The effect of metal pollution on the life history and insecticide resistance phenotype of the major malaria vector *Anopheles arabiensis* (Diptera: Culicidae). *Anopheles arabiensis* (Diptera: Culicidae). *PLoS ONE* 13 (2), e0192551. doi:10.1371/journal.pone.0192551
- Osborne, J., Clark, S., Morris, R., Williams, I., Riley, J., Smith, A., et al. (1999). A landscape-scale study of bumble bee foraging range and constancy, using harmonic radar. *J. Appl. Ecol.* 36, 519–533. doi:10.1046/j.1365-2664.1999.00428.x
- Owen, R. E., Richards, K. W., and Wilkes, A. (1995). Chromosome numbers and karyotypic variation in bumble bees (hymenoptera: Apidae: bombini). *J. Kans. Entomol. Soc.* 68 (3), 290–302. Available at: <https://www.jstor.org/stable/25085597>.
- Paris, J. R., Stevens, J. R., and Catchen, J. M. (2017). Lost in parameter space: A road map for stacks. *Methods Ecol. Evol.* 8, 1360–1373. doi:10.1111/2041-210X.12775
- Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. *R J.* 10 (1), 439–446. doi:10.32614/RJ-2018-009
- Peeters, T. M. J., Nieuwenhuijsen, H., Smit, J., van der Meer, F., Raemakers, I. P., Heitmans, W. R. B., et al. (2012). *De nederlandse bijen (hymenoptera: Apidae s.l.)*. Leiden, Netherlands: - Natuur van Nederland.
- Pirounakis, K., Koulianos, S., and Schmid-Hempel, P. (1998). Genetic variation across European populations of *Bombus pascuorum* (Hymenoptera: Apidae) from mitochondrial DNA sequence data. *Eur. J. Entomol.* 95, 27–33.
- Radchenko, V. G., and Pesenko, Y. A. (1994). *Biology of bees (hymenoptera, apoidea)*. St. Petersburg: Zoological Institute of the Russian Academy of Sciences, 350. doi:10.13140/2.1.3938.6242
- Rasmont, P., Franzén, M., Lecocq, T., Harpke, A., Roberts, S., Biesmeijer, J. C., et al. (2015). Climatic risk and distribution atlas of European bumblebees. *BioRisk* 10, 1–236. doi:10.3897/biorisk.10.4749
- Rasmont, P., Ghisbain, G., and Terzo, M. (2021). *Bumblebees of Europe and neighbouring regions*. Verrières-le-Buisson, France: NAP Editions.
- Raven, P. H., and Wagner, D. L. (2021). Agricultural intensification and climate change are rapidly decreasing insect biodiversity. *Proc. Natl. Acad. Sci. U. S. A.* 118 (2), e2002548117. doi:10.1073/pnas.2002548117
- Redhead, J. W., Dreier, S., Bourke, A. F. G., Heard, M. S., Jordan, W. C., Sumner, S., et al. (2016). Effects of habitat composition and landscape structure on worker foraging distances of five bumble bee species. *Ecol. Appl.* 26 (3), 726–739. doi:10.1890/15-0546
- Rochette, N., and Catchen, J. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nat. Protoc.* 12, 2640–2659. doi:10.1038/nprot.2017.123
- Ross, N. (2020). fasterize: Fast polygon to raster conversion. R package version 1.0.3. Available at: <https://CRAN.R-project.org/package=fasterize>.
- Ruddiman, W. F. (2013). The Anthropocene. *Annu. Rev. Earth Planet. Sci.* 41 (1), 45–68. doi:10.1146/annurev-earth-050212-123944
- Sadd, B. M., Barribeau, S. M., Bloch, G., de Graaf, D. C., Dearden, P., Elsik, C. G., et al. (2015). The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biol.* 16, 76. doi:10.1186/s13059-015-0623-3
- Sage, R. F. (2020). Global change biology: A primer. *Glob. Chang. Biol.* 26 (1), 3–30. doi:10.1111/gcb.14893
- Sánchez-Bayo, F. (2011). *Impacts of agricultural pesticides on terrestrial ecosystems*. Bussum, Netherlands: Bentham Science Publisher, 63–87. Chapter 4.
- Sánchez-Bayo, F., and Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biol. Conserv.* 232, 8–27. doi:10.1016/j.biocon.2019.01.020
- Shafer, A. B., Wolf, J. B., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., et al. (2015). Genomics and the challenging translation into conservation practice. *Trends Ecol. Evol.* 30, 78–87. doi:10.1016/j.tree.2014.11.009

- Silva, S. E., Seabra, S. G., Carvalheiro, L. G., Nunes, V. L., Marabuto, E., Mendes, R., et al. (2020). Population genomics of *Bombus terrestris* reveals high but unstructured genetic diversity in a potential glacial refugium. *Biol. J. Linn. Soc. Lond.* 129, 259–272. doi:10.1093/biolinnean/blz182
- Silver, K. S., Du, Y., Nomura, Y., Oliveira, E. E., Salgado, V. L., Zhorov, B. S., et al. (2014). Voltage-gated sodium channels as insecticide targets. *Adv. Insect Phys.* 46, 389–433. doi:10.1016/B978-0-12-417010-0.00005-7
- Sivakoff, F. S., and Gardiner, M. M. (2017). Soil lead contamination decreases bee visit duration at sunflowers. *Urban Ecosyst.* 20, 1221–1228. doi:10.1007/s11252-017-0674-1
- South, A. (2011). rworldmap: a new R package for mapping global data. *R. J.* 31 (1), 35–43. doi:10.32614/RJ-2011-006
- Sun, C., Huang, J., Wang, Y., Zhao, X., Su, L., Thomas, G. W. C., et al. (2021). Genus-wide characterization of bumblebee genomes provides insights into their evolution and variation in ecological and behavioral traits. *Mol. Biol. Evol.* 38, 486–501. doi:10.1093/molbev/msaa240
- Tarnawska, M., Kafel, A., Augustyniak, M., Rost-Roszkowska, M., and Babczyńska, A. (2019). Microevolution or wide tolerance? Level of stress proteins in the beet armyworm *Spodoptera exigua* hübner (Lepidoptera: Noctuidae) exposed to cadmium for over 150 generations. *Ecotoxicol. Environ. Saf.* 30 (178), 1–8. doi:10.1016/j.ecoenv.2019.04.017
- Theodorou, P., Baltz, L. M., Paxton, R. J., and Soro, A. (2021). Urbanization is associated with shifts in bumblebee body size, with cascading effects on pollination. *Evol. Appl.* 14, 53–68. doi:10.1111/eva.13087
- Theodorou, P., Radzevičiūtė, R., Kahnt, B., Soro, A., Grosse, I., and Paxton, R. J. (2018). Genome-wide single nucleotide polymorphism scan suggests adaptation to urbanization in an important pollinator, the red-tailed bumblebee (*Bombus lapidarius* L.). *Proc. Biol. Sci.* 285, 20172806. doi:10.1098/rspb.2017.2806
- Thompson, H. M. (2001). Assessing the exposure and toxicity of pesticides to bumblebees (*Bombus* sp.). *Apidologie* 32, 305–321. doi:10.1051/apido:2001131
- Tommasi, N., Pioltelli, E., Biella, P., Labra, M., Casiraghi, M., and Galimberti, A. (2022). Effect of urbanization and its environmental stressors on the intraspecific variation of flight functional traits in two bumblebee species. *Oecologia* 199, 289–299. doi:10.1007/s00442-022-05184-x
- Van Swaay, C. A. M., and Warren, M. S. (2012). “Developing butterflies as indicators in Europe: Current situation and future options,” in *De vlinderstichting/ Dutch butterfly conservation* (Butterfly Conservation UKWageningen: Butterfly Conservation Europe). reportnr. VS2012.012.
- Vanderplanck, M., Martinet, B., Carvalheiro, L., Rasmont, P., Barraud, A., Renaudeau, C., et al. (2019a). Ensuring access to high-quality resources reduces the impacts of heat stress on bees. *Sci. Rep.* 9, 12596. doi:10.1038/s41598-019-49025-z
- Vanderplanck, M., Roger, N., Moerman, R., Ghisbain, G., Gérard, M., Popowski, D., et al. (2019b). Bumble bee parasite prevalence but not genetic diversity impacted by the invasive plant *Impatiens glandulifera*. *Ecosphere* 10 (7), e02804. doi:10.1002/ecs2.2804
- Vaudo, A., Tooker, J., Patch, H., Biddinger, D., Coccia, M., Crone, M., et al. (2020). Pollen protein: Lipid macronutrient ratios may guide broad patterns of bee species floral preferences. *Insects* 11 (2), 132. doi:10.3390/insects11020132
- Velthuis, H. H. W., and van Doorn, A. (2006). A century of advances in bumblebee domestication and the economic and environmental aspects of its commercialization for pollination. *Apidologie* 37, 421–451. doi:10.1051/apido:2006019
- Vray, S., Rollin, O., Rasmont, P., Dufrière, M., Michez, D., and Dendoncker, N. (2019). A century of local changes in bumblebee communities and landscape composition in Belgium. *J. Insect Conserv.* 23, 489–501. doi:10.1007/s10841-019-00139-9
- Wagner, D. L., Grames, E. M., Forister, M. L., Berenbaum, M. R., and Stopak, D. (2021). Insect decline in the Anthropocene: Death by a thousand cuts. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2023989118. doi:10.1073/pnas.2023989118
- Warmke, J. W., Reenan, R. A. G., Wang, P., Qian, S., Arena, J. P., Wang, J., et al. (1997). Functional expression of *Drosophila para* sodium channels: Modulation by the membrane protein TipE and toxin pharmacology. *J. Gen. Physiol.* 110, 119–133. doi:10.1085/jgp.110.2.119
- Wickham, H., François, R., Henry, L., and Müller, K. (2021). dplyr: A grammar of data manipulation. R package version 1.0.7. Available at: <https://CRAN.R-project.org/package=dplyr>.
- Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer.
- Wickham, H. (2011). The split-apply-combine strategy for data analysis. *J. Stat. Softw.* 40 (1), 1–29. doi:10.18637/jss.v040.i01
- Widmer, A., and Schmid-Hempel, P. (1999). The population genetic structure of a large temperate pollinator species, *Bombus pascuorum* (Scopoli) (Hymenoptera: Apidae). *Mol. Ecol.* 8, 387–398. doi:10.1046/j.1365-294X.1999.00584.x
- Williams, P. H., Altanchimeg, D., Byvaltsev, A., De Jonghe, R., Jaffar, S., Japoshvili, G., et al. (2020). Widespread polytypic species or complexes of local species? Revising bumblebees of the subgenus *Melanobombus* world-wide (hymenoptera, apidae, *Bombus*). *Eur. J. Taxon.* 719 (1), 1–120. doi:10.5852/ejt.2020.719.1107
- Williams, P. H., and Osborne, J. L. (2009). Bumblebee vulnerability and conservation world-wide. *Apidologie* 40, 367–387. doi:10.1051/apido/2009025
- Wolf, S., and Moritz, R. F. A. (2008). Foraging distance in *Bombus terrestris* L. (Hymenoptera: Apidae). *Apidologie* 39, 419–427. doi:10.1051/apido:2008020
- Wood, T., Ghisbain, G., Rasmont, P., Kleijn, D., Raemakers, I., Praz, C., et al. (2021). Global patterns in bumble bee pollen collection show phylogenetic conservation of diet. *J. Anim. Ecol.* 90 (10), 2421–2430. doi:10.1111/1365-2656.13553
- Xu, T., Rebay, I., Fleming, R. J., Scottgale, T. N., and Artavanis-Tsakonsa, S. (1990). The Notch locus and the genetic circuitry involved in early *Drosophila* neurogenesis. *Genes Dev.* 4, 464–475. doi:10.1101/gad.4.3.464
- Yañez, O., Piot, N., Dalmon, A., de Miranda, J. R., Chantawannakul, P., Panziera, D., et al. (2020). Bee viruses: Routes of infection in hymenoptera. *Front. Microbiol.* 11, 943. doi:10.3389/fmicb.2020.00943
- You, L. L., Wu, Y., Xu, B., Ding, J., Ge, L. Q., Yang, G. Q., et al. (2016). Driving pest insect populations: Agricultural chemicals lead to an adaptive syndrome in *Nilaparvata lugens* stål (Hemiptera: Delphacidae). *Sci. Rep.* 23, 37430. doi:10.1038/srep37430
- Zhao, S., Liu, S., and Zhou, D. (2016). Prevalent vegetation growth enhancement in urban environment. *Proc. Natl. Acad. Sci. U. S. A.* 113 (22), 6313–6318. doi:10.1073/pnas.1602312113
- Zhu, F., Moural, T. W., Nelson, D. R., and Palli, S. R. (2016). A specialist herbivore pest adaptation to xenobiotics through up-regulation of multiple Cytochrome P450s. *Sci. Rep.* 6, 20421. doi:10.1038/srep20421