Check for updates

Molecular Ecology

WILEY

MOLECULAR ECOLOGY

ORIGINAL ARTICLE OPEN ACCESS

Genome Variation in Three *Anthophora* Bee Species Reflects Divergent Demographic Histories

Demetris Taliadoros[1] | André E. R. Soares[1] | Guilherme Dias[2] | Ignas Bunikis[3] | Martin Pippel[2] | Anna Olsson[1] | Mai-Britt Mosbech[3] | Julia Heintz[3] | Nina Lager[3] | Anna-Sofi Strand[3] | Mats Pettersson[1] | Olga Vinnere Pettersson[3] | Henrik Lantz[1] | Björn Cederberg[4] | Kerstin Lindblad-Toh[1,5] | Matthew T. Webster[1]

[1]Department of Medical Biochemistry and Microbiology, SciLifeLab, Uppsala University, Uppsala, Sweden | [2]Department of Cell and Molecular Biology, SciLifeLab, Uppsala University, Uppsala, Sweden | [3]Department of Immunology, Genetics and Pathology, SciLifeLab, Uppsala University, Uppsala, Sweden | [4]Swedish Species Information Centre, Swedish University of Agricultural Sciences, Uppsala, Sweden | [5]Broad Institute of MIT and Harvard, Cambridge, USA

Correspondence: Matthew T. Webster (matthew.webster@imbim.uu.se)

Received: 24 April 2025 | **Revised:** 6 November 2025 | **Accepted:** 1 December 2025

Handling Editor: Sean D Schoville

Keywords: conservation genetics | empirical | genomics/proteomics | insects | population dynamics | population genetics

ABSTRACT

Population genomics can reveal trends and drivers of biodiversity loss, but it is still unclear how best to use measures of genome variation to understand population vulnerability in insects. Here we study genome variation in three species of *Anthophora* bees that show contrasting population trends in northern Europe. Two species, *Anthophora plagiata* and *Anthophora retusa* , have experienced declines and recoveries of different magnitudes in the last 50 years, whereas a third species, *Anthophora quadrimaculata* , has relative population stability. We generate highly contiguous genome assemblies and use them to study genome variation in 136 samples of these species collected throughout Sweden. We find exceedingly low genetic variation in *A. plagiata* , which has experienced a severe recent bottleneck, but high genetic variation in *A. retusa* , despite a similar recent population trajectory. Fragmented populations of the threatened species *A. plagiata* appear isolated from each other, but in *A. retusa*, there is a lack of deep population structure among geographically separated subpopulations. We infer population size in the distant past using MSMC2 and recent past using GONE. These methods are remarkably concordant and indicate ancient fluctuations in population size dating back to the Pleistocene, with moderate expansions in the past century in all three species. These results are comparable to some other studies of endangered insects, which have experienced population declines that predate the modern era. We detect long blocks of identity-by-state in *A. plagiata* , indicative of severe recent inbreeding. Translocations between isolated populations of this species could have a positive effect on their resilience.

1 | Background

Our planet is experiencing an unprecedented loss of biodiversity, which impacts all forms of life (IPBES 2019). In particular, several recent studies have identified drastic declines in insect populations on a local scale, with some meta-analyses indicating that declines are occurring globally (Hallmann et al. 2017; Sánchez-Bayo and Wyckhuys 2019; van Klink et al. 2020; Wagner 2020). These reports are concerning due to the vital ecosystem functions performed by insects, including pollination, decomposition, maintaining soil health and serving as a food source for other animals.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2025 The Author(s). *Molecular Ecology* published by John Wiley & Sons Ltd.

Pollinating insects in particular are indispensable for many crops and wild plants, with about 80% of common crop species relying on pollination services (Klein et al. 2006). The contribution of wild bee pollination to crop production in the USA alone is in excess of 1.5 billion USD per year (Reilly et al. 2020). There is, however, clear evidence that pollinating insects have been negatively impacted by factors such as habitat loss, pesticides, climate change and pathogens (Biesmeijer et al. 2006; LeBuhn and Vargas Luna 2021; Potts et al. 2010; Powney et al. 2019; Soroye et al. 2020). Bees are the most important group of pollinators, and the majority of bee species worldwide (90%) are solitary (Danforth et al. 2019). Recent declines in solitary bees are driven by human activities, with habitat loss and degradation being a primary cause (LeBuhn and Vargas Luna 2021).

Population genomic analyses coupled with the availability of high-quality reference genome assemblies have the potential to revolutionise our understanding of population trends and conservation status of vulnerable species (Allendorf et al. 2010; Formenti et al. 2022; Hohenlohe et al. 2021; Ouborg et al. 2010; Supple and Shapiro 2018; Theissinger et al. 2023). Traditional conservation genetic approaches analyse levels of genetic variation at a limited number of loci, such as using short tandem repeats or mitochondrial DNA, which can provide estimates of population structure, inbreeding and effective population size. By contrast, whole-genome sequencing at a population level—population genomics—can prove a much richer and more precise set of inferences that can be used to manage vulnerable populations. These include direct quantification of the occurrence of deleterious mutations and adaptive potential of populations. Population genomics can also uncover runs of homozygosity in the genome due to inbreeding, model past population trends using coalescent-based approaches, and reveal hidden structure among populations.

Genomic approaches have proved particularly valuable for diagnosing the status of threatened terrestrial vertebrates. For example, population-scale sequencing was used to assess the extent of inbreeding and genetic load in fragmented red deer (*Cervus elaphus*) populations in the Netherlands (De Jong et al. 2020), the threatened Iberian lynx (*Lynx pardinus*) in Spain (Abascal et al. 2016) and wolf (*Canis lupus lupus*) populations in Sweden (Smeds and Ellegren 2023). These approaches can also be informative to understand the effects of translocations on population health, in examples such as the Florida panther (Saremi et al. 2019) and the arctic fox (Hasselgren et al. 2024).

The use of population genomics for insect conservation has been limited, but such analyses are very promising to understand population trends and inform conservation (Webster et al. 2023). Compared to vertebrates, insect populations have some specific features that need to be considered in interpreting and applying genomic studies to conservation. In particular, effective population sizes of insect species are typically much larger than vertebrates and can undergo large fluctuations in a short time. Furthermore, even small geographic areas can support high genetic variation in insect species, as shown by the example of an endangered flightless grasshopper with a highly fragmented population, which was still found to have high levels of genetic variation (Hoffmann et al. 2021). Levels of genetic variation found in a species depend strongly on life history in addition to recent population fluctuations (Romiguier, Gayral, et al. 2014; Romiguier, Lourenco, et al. 2014). For these reasons, it is not trivial to determine conservation status from estimates of genetic variation.

Another important question is therefore to understand whether low variation is the result of recent human-mediated declines or long-term patterns. For example, a study on the Xerces blue butterfly, which became extinct in the last century, revealed that high levels of genetic load and low effective population size were likely a long-term feature of this species, which predated the human era (de-Dios et al. 2024). Similarly, population genomics of isolated Apollo butterfly populations in France and endangered Queen Alexandra's Birdwing butterflies in Papua New Guinea both indicate population declines that began several thousand years ago (Kebaïli et al. 2022; Reboud et al. 2023).

In this study, we focus on three solitary bee species of conservation interest in Europe from the genus *Anthophora*. This is one of the largest genera in the family Apidae, containing over 450 species (Danforth et al. 2019), with a common ancestor that lived around 40 million years ago (Henríquez-Piskulich et al. 2024). By choosing multiple species from the same genus, we aim to disentangle the effects of life history, which are broadly similar between species, and population fluctuations on genetic variation. We also aim to inform ongoing conservation efforts targeting our focal species.

We focus on the following species: *Anthophora quadrimaculata* (Panzer, 1798) (Four-banded Flower Bee), *Anthophora retusa* (Linnaeus, 1758) (Potter Flower Bee) and *Anthophora plagiata* (Illiger, 1806). All of the species are distributed widely in Europe and have been observed further east in Eurasia (Figures S1–S3). They are all classified as generalist foragers that are most common in grasslands (Rasmont and Dehon 2014a, 2014b, 2014c). However, they differ in abundance, and different population trends have been reported among species. *A. quadrimaculata* has a large spatial distribution throughout Europe and, although population trends are not well described, there are no reports of declines. *A. retusa* is distributed throughout Europe and the Near East but has been considerably regressing in parts of its range, particularly in western Europe. *A. plagiata* has a similar distribution to *A. retusa* and shows even more extreme regional population declines, having disappeared from several European countries.

The Swedish populations of these three *Anthophora* species have been the subject of ongoing management programs, and a large amount of observation data exists. Similar to the rest of its distribution, *A. quadrimaculata* appears to have a ubiquitous stable population in Sweden, and no declines have been observed. It is assigned the least concern (LC) category in the most recent Swedish red list (SLU Artdatabanken 2020). The population of *A. retusa* has experienced a drastic decline within the last 50 years, followed by a recent subsequent recovery. Between 1980 and 2006, observations of this species were restricted to the islands of Öland and Gotland, with only a handful of observations on the mainland. A nationwide

13652394x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

management program was started in 2007 (Nilsson and Andersson 2007). Since this time, there has been a large increase in observations in Öland, Gotland and southeastern Sweden, in addition to an extensive expansion of its distribution around Stockholm and lake Mälaren. This species was classified as near threatened (NT) in the Swedish red list (SLU Artdatabanken 2020). In the 1940s, *A. plagiata* was spread from the far south of Sweden to Värmland in the north. However, this species also underwent a drastic decline and was believed to have disappeared from Sweden by the end of the last century. However, a small population was discovered in Skåne with only a few tens of individuals in 2004, which grew to more than 150 breeding females by 2014, when a management program was initiated in order to promote its growth (Cederberg 2014). Recently, a new population of this species with only a few tens of individuals was discovered in Värmland, more than 400 km from the existing population, but it is unknown how they are related. *A. plagiata* is currently classified as endangered (EN) in Sweden (SLU Artdatabanken 2020).

Here we generate high-quality genome reference assemblies and annotations for the three *Anthophora* species. We then use these assemblies to study genome variation in 136 samples collected from throughout Sweden. We address several questions. Firstly, what impact have the observed population declines had on genetic variation in the species they have affected? Second, how are geographically separate populations related to each other? Third, what historical fluctuations in population size are apparent in the evolutionary history of these species? Fourth, what is the extent of inbreeding observed in the populations? In all cases, we compare patterns of genome variation in *A. quadrimaculata* , which currently has a large and stable population, with those in *A. retusa* and *A. plagiata* , which have both experienced recent population declines to better understand the effects of recent population fluctuations. Our findings have relevance for management strategies for these threatened species.

2 | Material and Methods

2.1 | Biological Material

To study the population history and inbreeding of three *Anthophora* species in Sweden, a total of 136 bees were collected. The sampling occurred between 2020 and 2022 throughout the known local habitat range of the three species. For *A. quadrimaculata* , 52 individuals (40 females and 12 males) were collected from ten localities within five counties in Sweden, distributed across a large area of the country. We collected 57 samples of *A. retusa* (10 females and 47 males) from 12 localities in six counties. For *A. plagiata* , 27 bees, all males, were collected from five localities, four of which were in Skåne (Skillinge, Bratevik, Lund and Hörjel) and one in Årjäng, Värmland (Figure 1A–C, Tables S1–S3). Samples for long-read sequencing were flash frozen in liquid nitrogen and stored in a −80°C freezer, whereas samples for short-read sequencing were placed in 95% ethanol and stored at −20°C. For each species, two males were used for the genome assembly and annotation. One adult male thorax was used for PacBio HiFi, or in the case of *A. plagiata* , Oxford Nanopore DNA sequencing. The second male was used for the PacBio IsoSeq and Illumina RNA-seq (see below).

2.2 | Long-Read DNA and RNA Sequencing

High-molecular-weight genomic DNA was extracted from snap-frozen thorax tissue of *A. plagiata* , *A. retusa* and *A. quadrimaculata* using SDS lysis (2% N-lauryl Sarcosine, 0.5% SDS, 50 mM Tris pH 8, 10 mM EDTA, 2.5 mg/mL Proteinase K) followed by phenol:chloroform:isoamyl alcohol (25:24:1) extraction using MaXtract High Density phase lock tubes (Qiagen Cat #129056). For *A. plagiata* , 1% 2-Mercaptoethanol was added to the lysis buffer to improve purity. DNA was purified with chloroform:isoamylalcohol (24:1) and a high-salt/low ethanol precipitation using 0.3× vol of 99% ethanol, which precipitates polysaccharides while gDNA remains in solution. DNA was precipitated using 1.7× vol 99% ethanol, washed twice with 70% ethanol and eluted in 200 μL TE-buffer (10 mM Tris, 0.1 mM EDTA). Samples were left at room temperature with gentle mixing at 100 rpm on a platform rocker for several days to increase homogeneity before performing QC.

Sample AP2 of *A. plagiata* was sequenced on the Oxford Nanopore (ONT) PromethION system. Before library prep, size selection was performed using the Circulomics Short Read Eliminator (SRE) XS Kit. Starting material for SRE was 2 μg and was followed by a shearing step using Megarupor 3 (Diagenode), speed 35, which resulted in a 13 kb DNA fragment length. Femto Pulse (Agilent) was used to assess shearing results. 660 ng of size-selected and sheared DNA was taken into the library prep, using the ONT Ligation Sequencing gDNA kit (SQK-LSK110), with end prep times of 30 + 30 min and an adapter ligation time of 25 min. Bead binding and elution times were 20 and 60 min respectively. Quality control of the library was performed using the Qubit dsDNA BR kit. Of the final library, 91.5 ng (~10.6fmol) was loaded on one PromethION flow cell (FLO-PRO002, version R9), with 72 h run time. One additional flow cell was loaded with 130 ng (~15fmol) library, which gave more data.

Libraries for PacBio sequencing were prepared according to the PacBio protocol 'Preparing HiFi SMRTbell Libraries using the SMRTbell Express Template Prep Kit 2.0', PN 101–853-100 Version 05 (August 2021) using the SMRTbell Express Template Prep Kit 2.0.1 μg of each sample was first sheared on the Megaruptor 3 (Diagenode) with speed setting 32, resulting in 15–16 kb fragment lengths. An AMPure bead purification was performed after the shearing, according to the library prep protocol. Size selection of libraries was performed using 3.7X AMPure beads according to the PacBio protocol 'Procedure & Checklist – Using AMPure PB Beads for Size-Selection', PN 101–854-900 Version 02 (January 2020).

Quality control of sheared DNA and SMRTbell libraries was performed on Fragment Analyser (Agilent), using the Large Fragment standard sensitivity 492 kit (DNF-492-33). After library prep, samples were pooled equimolarly. Primer annealing and polymerase binding were performed using the Sequel II binding kit 2.2. The pooled samples were sequenced on three PacBio SMRTcells on the Sequel IIe instrument, using the Sequel II sequencing plate 2.0 and the Sequel II SMRT Cell 8 M,

13652394x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

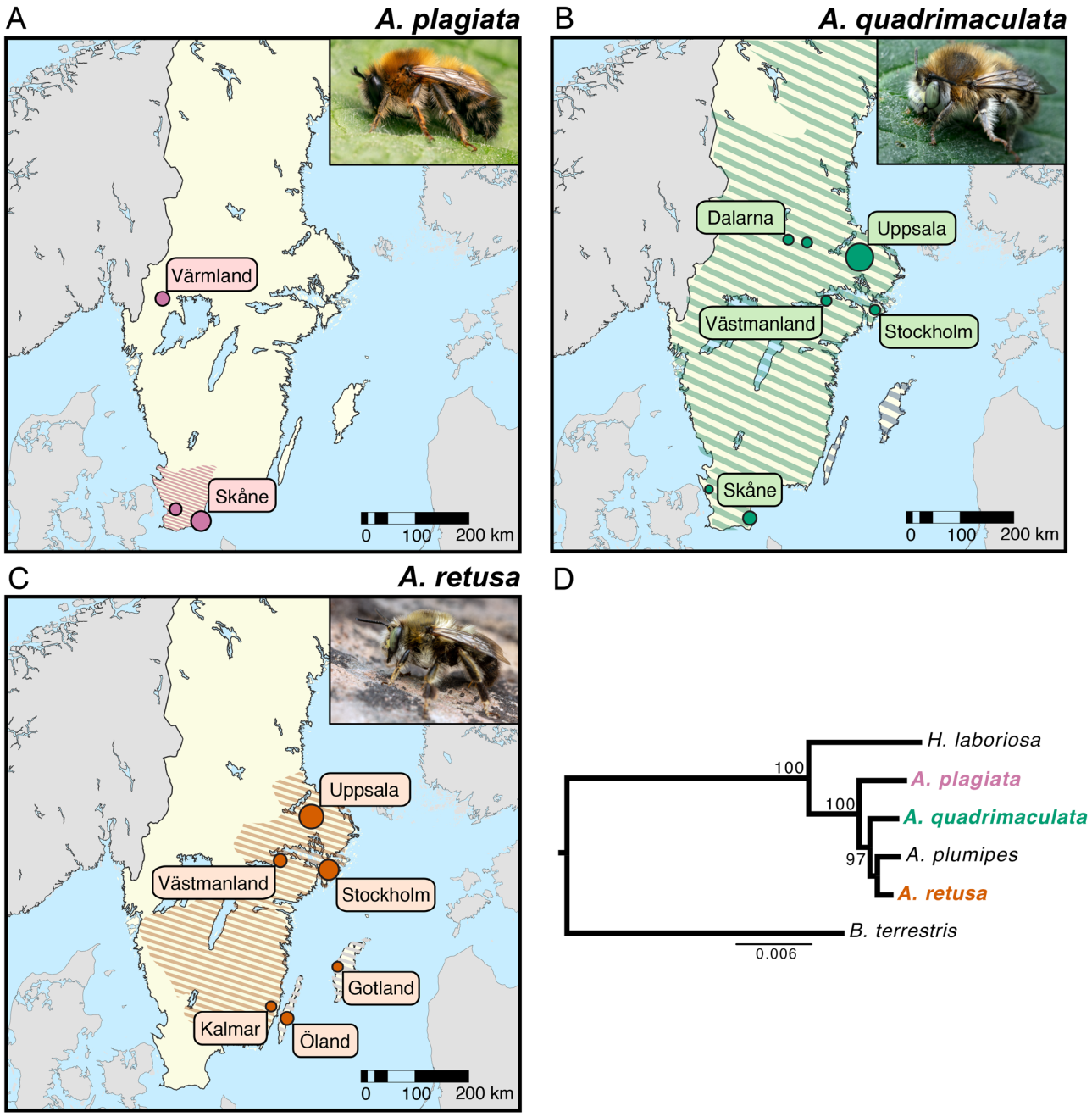

FIGURE 1 | Sampling and phylogenetic relationship of the three *Anthophora* species. (A) Distribution and sampling localities for each species. Photographs by Krister Hall. (B) Phylogenetic tree using IQtree2 (Maximum likelihood inference, log likelihood: −857128.545) to assess the evolutionary relationship between Anthophorini species. Branch numbers reflect maximum likelihood bootstrap values. The tree was rooted using the species *Bombus terrestris* . Evolutionary distances are converted into million years of divergence using the formula T = D/2r (see Section 2).

with 110-150pM on plate loading, 30 h movie time and 2 h pre-extension time.

We generated RNA sequence data using both PacBio iso-seq and Illumina RNA-seq. As with the DNA extraction, RNA was extracted from snap-frozen thorax tissue of *A. plagiata* , the three species using the TRIzol Reagent and Phasemaker Tubes Complete System (Invitrogen Cat #A33250) following the Invitrogen user guide (Pub. No. MAN0016163 Rev. A.0) except for steps 2.a. and 3.d., which were omitted. RNA was resuspended in 87.5 μL RNase-free water and immediately subjected to DNase treatment followed by purification according to the RNeasy Micro Handbook (pages 74 and 53). The eluted RNA was stored at −70°C. For *A. quadrimaculata* , we generated 3.3 Gbp and 45.2 Gbp, respectively. For *A. retusa* , these figures were 2.5 Gbp and 47.9 Gbp, respectively, and for *A. plagiata* , they were 2.6 Gbp and 60.8 Gbp, respectively.

For the isoseq preparation, libraries for PacBio sequencing were prepared as described in 'Preparing Iso-Seq libraries' (PN 102–396-000 REV02 April 2022) using the SMRTbell prep kit 3.0 (PacBio). 300 ng of RNA was used for cDNA synthesis.

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Additional cycles were performed according to step 3.2.16.166 ng of each final SMRTbell was pooled, and primer annealing and polymerase binding were performed using the Sequel II binding kit 3.1. Finally, the pooled samples were sequenced on one Sequel II SMRT Cell 8 M on Sequel II using Sequel II Sequencing Plate 1.0, on a plate loading concentration of 100 pM and 24 h movie time.

2.3 | Genome Assembly and Annotation

The PacBio HiFi sequencing reads from *A. quadrimaculata* and *A. retusa* were assembled using hifiasm v0.16.0 (Cheng et al. 2021). Mitochondrial sequences were identified and removed from the primary contigs with MitoHiFi v3.0.0 (Uliano-Silva et al. 2023) and annotated with MitoFinder v1.4.1 (Allio et al. 2020) with manual adjustments. Contigs shorter than 1 kb were removed from the final assembly before annotation. The Oxford Nanopore sequencing reads from *A. plagiata* were assembled with Flye v2.9 (Kolmogorov et al. 2019). Potential haplotype duplications were removed with Purge_Dups v1.2.5 (Guan et al. 2020). The assembly was then polished with Medaka v1.7.2 (available at https://github.com/nanoporetech/medaka). Potential contaminants were identified and removed using BlobTools v4.2.1 (Laetsch and Blaxter 2017), highly repetitive contigs were identified and removed using Tandem Repeats Finder v4.09.1 (Benson 1999), and mitochondrial fragments were identified and removed using BLAST v2.10.0+ (Altschul et al. 1990). Contigs shorter than 1 kbp were removed from the final assembly prior to annotation. The completeness of each assembly was evaluated with BUSCO v. 5.4.6 using the hymenoptera_odb10 lineage ($n = 5991$).

For each species, a custom repeat library was created with RepeatModeler2 v2.0.2 (Flynn et al. 2020) and masked with RepeatMasker v4.1.5 (https://www.repeatmasker.org/) before annotation. As repeats can be part of actual protein-coding genes, the candidate repeats obtained by RepeatModeler2 were vetted against the Uniprot/Swissprot protein set (minus transposons) to exclude any nucleotide motif stemming from low-complexity protein-coding sequences.

Gene prediction was carried out in three independent steps, using evidence from Illumina RNA-seq data, protein sequences from multiple species, and PacBio Iso-Seq transcripts. First, RNA-seq reads were processed with BRAKER v3.0.3 (Gabriel et al. 2024), which used GeneMark-ET (Lomsadze et al. 2014) and AUGUSTUS (Stanke et al. 2006) to detect splice signals, train models and predict genes. Second, a comprehensive set of arthropod protein sequences from OrthoDB v11 (Kriventseva et al. 2019) was aligned to the genome with miniprot v0.10-r226-dirty (Li 2023). These alignments were used to train and predict genes with AUGUSTUS through the GALBA v1.0.6 pipeline (Brůna et al. 2023). Because BRAKER3 performs best with large protein datasets (Gabriel et al. 2024), we used the full arthropod set rather than a smaller taxon-specific subset. Third, high-quality PacBio Iso-Seq transcripts were aligned to the genome using minimap2 v2.26 (Li 2018), and gene models were predicted with GeneMarkS-T v5.1 (Tang et al. 2015), following the BRAKER long-read protocol. Finally, gene models from BRAKER, GALBA and GeneMarkS-T were merged and filtered with TSEBRA (Gabriel et al. 2021) to produce the final annotation. All commands used are available in the GitHub repository.

The combined gene models were processed with AGAT v1.2.0 (available at: https://zenodo.org/records/817887) to fix overlapping genes and then functionally annotated using the NBIS functional_annotation nextflow pipeline v2.0.0 commit c7f0d2f (https://github.com/NBISweden/pipelines-nextflow). Briefly, this pipeline performs similarity searches using BLAST between the annotated proteins and the UniProtKB/Swiss-Prot database (Magrane, and Consortium, U 2011) (downloaded on 2022–12; 568,363 proteins). Then, it uses InterProScan (Jones et al. 2014) to query the proteins against InterPro v59-91 databases (Paysan-Lafosse et al. 2023) and merges results using AGAT. Single-exon genes without any InterPro annotation were removed to reduce potential false positives. Transfer RNAs (tRNAs) were predicted using tRNAscan-SE v2.0.12 (Lowe and Eddy 1997) with default eukaryote parameters. Further details of the genome annotation procedure are provided in Data S1.

2.4 | Phylogenetic Reconstruction

To infer phylogenetic relationships of the three newly sequenced species within the genus *Anthophora* (family, Apidae; tribe, Anthophorini), we also obtained whole genome assemblies of *Anthophora plumipes* and of the closely related Anthophorini species *Habropoda labiriosa* available on NCBI under the BioProject ID PRJEB62774 and PRJNA279436, respectively (Kapheim et al. 2015). Furthermore, the *Bombus terrestris* genome assembly iyBomTerr1.2 (Crowley et al. 2023), available on NCBI under BioProject PRJEB45694, was obtained to be used as an outgroup. A whole genome alignment was performed with Cactus v. 2.8.2 (Paten et al. 2011) using default parameters. The alignment was then used in a maximum-likelihood (ML) analysis with IQtree version 2.0.3 (Minh et al. 2020). The most appropriate substitution model was chosen based on the Bayesian Information Criterion (BIC) using the ModelFinder algorithm implemented in IQtree version 2.0.3 (Kalyaanamoorthy et al. 2017).

Branch support was obtained using the bootstrap approximation option of IQtree (Hoang et al. 2018), performing 1000 bootstrap replicates. The resulting trees were edited in FigTree 1.4.4. To estimate divergence time between species, we scaled the branch length using an estimate of substitution rate per site per generation of 3.6×10^{-9} (95% confidence intervals: 2.38×10^{-9}–5.4×10^{-9}) derived from estimates of mutation rate in *Bombus terrestris* (Liu et al. 2017).

2.5 | Population Sequencing, Read Mapping and Variant Calling

Samples for population sequencing were dissected to extract wing muscle from the thorax. DNA was extracted using the Qiagen DNeasy Blood & Tissue Kit. We performed library preparation using the Illumina Nextera Flex kit. The libraries were pooled and run on a single S4 flow cell of an Illumina NovaSeq6000 instrument. The Bcl to FastQ conversion was performed using bcl2fastq_v2.20.0.422 from the CASAVA software suite.

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

We filtered and mapped the Illumina reads to the reference genomes to obtain high-quality genomic variants. In brief, we trimmed sequencing adapters and kept bases with a sequencing quality PHRED score over 33 using the software Trimmomatic version 0.38 (Bolger et al. 2014). Additionally, we only kept reads with a minimum length of 30 bp. We used Burrows-Wheeler Aligner (BWA) mem version 0.7.17 (Li and Durbin 2010) to map the reads to the reference genomes. Subsequently, we used the GATK HaplotypeCaller version 4.2.18 and GenotypeGVCF version 4.2.18 (McKenna et al. 2010) to obtain a variant calling format (VCF) file with raw variants. Male *Anthophora* bees were called haploid, while females were called diploid. Following the GATK 'best practice guidelines' we further filtered low-quality SNPs by applying four hard-filtering criteria. We kept variants that (1) had a minimum quality by depth (QD) value of two, (2) had a minimum depth (DP) per genome value of five, (3) had a minimum mapping quality (MQ) per site of 40, (4) had a minimum average per genome coverage of one. These criteria were applied using GATK VariantFiltration version 4.0.11 (McKenna et al. 2010). Subsequently, we filtered positions that appeared to be heterozygous in males using the reference and alternative allele count (AC) and DP encoded in the VCF file. The position was discarded if the AC to DP ratio was between 0.3 and 0.7 in any male. To this end, we used a custom script available on GitHub: https://github.com/Jimi92/PopGenScripts/tree/main/MultiPloidy_VCF_tools. Summary statistics related to the read mapping and variant calling are summarised in Tables S1–S3.

With the application of these criteria, we obtained datasets of 13,461,036 (from which 11,096,707 biallelic SNPs), 4,343,624 (3,628,889) and 109,048 (22,312) variants for *A. quadrimaculata, A. retusa* and *A. plagiata* , respectively. We refer to these datasets as 'full high-quality datasets'. Several downstream analyses assume variants are unlinked; therefore, we filtered based on the linkage disequilibrium (LD) patterns using the Plink tool '--indep-pairwise' (Purcell et al. 2007) to filter the full high-quality datasets for LD considering a pairwise r2 threshold of 0.25. After filtering for LD, 5,315,366, 1,547,231 and 25,438 variants were retained. We further refer to this dataset as the 'independent SNP dataset'.

The number of males (haploid) and females (diploid) in our dataset differed among species, and only male individuals were collected for *A. plagiata* . To avoid potential biases due to differences in variant calling in haploids and diploids, we repeated some analyses using only males for all species. We created a VCF file for each species in which haploid males were combined in pairs and encoded as pseudo-diploid phased individuals. To this end, we used bcftools query with the options -H -f '%CHROM %POS %ID %REF %ALT [%GT]' to extract the genotype of males for each position. A custom script (available in the GitHub repository) was used to merge the male genotypes into pseudo-diploid phased individuals. This dataset is referred to as the 'male-independent SNP dataset.'

2.6 | Population Structure

We characterised the population genetic structure of the three *Anthophora* species based on complementary methods using genome-wide SNP data. Firstly, we investigated the extent of clustering using a principal component analysis (PCA) (Figure 2A). To perform the PCA analysis, we used the R package SNPRelate v. 1.6.4 (Zheng et al. 2012). Subsequently, the results were visualised using the R package ggplot2 (Gómez-Rubio 2017). We further explored population structure by inferring the extent of shared ancestry using a maximum likelihood approach implemented in ADMIXTURE (Alexander et al. 2009). Ten replicate runs were performed for a range of K-values (1–10) to select the most appropriate value for the number of hypothetical ancestral groups (K). The best K value was determined to yield the lowest average cross-validation error over the ten runs (Figure 2B, Tables S4–S6). We employed a Mantel test to assess correlations between genetic and geographic distance. The genetic distance was measured as the number of pairwise differences between the individuals. We ran the Mantel test using the Pearson correlation coefficient with 9999 permutations. The analysis was performed with the R package ape v. 5.8.1 (Paradis and Schliep 2019).

2.7 | Genetic Diversity and Linkage Disequilibrium Decay

We used the 'full high-quality dataset' to compute and compare genetic variation among populations. To this end, we estimated Watterson's theta (θ_w) as the number of segregating sites divided by the harmonic sum of the number of haploid genomes sampled (Watterson 1975). To estimate how θ_w varies across the genome of each species, we estimated θ_w in 1 kbp non-overlapping windows using a custom script. PopLDdecay (Zhang et al. 2019) was used with default parameters to estimate each species' linkage disequilibrium (LD) decay. VCFtools v. 0.1.17 (Danecek et al. 2011) was used to calculate Tajima's D for each population.

2.8 | Inference of Demographic History

To infer changes in the effective population size through time, we first used MSMC2 v2.1.1 (Wang et al. 2020). MSMC2 uses a Sequential Markovian coalescence model to infer coalescence rates between individuals (Wang et al. 2020). The MSMCtools bamCaller script available on GitHub (https://github.com/stschiff/msmc-tools) was used for the preparation of mask files for the low-coverage regions and to 'diploidize' the haploid vcf files. After that, the script generate_multihetsep.py, included in the MSMC2 software package, was used to create the input files for the analysis. We used a mutation rate (μ) of 3.6×10^{-9} (Liu et al. 2017) per base pair per generation estimated from the bumblebee *Bombus terrestris* , assuming one generation per year to scale the MSMC2 results. We used a recombination rate estimated for the solitary bee *Megachile rotundata* of 1.02 cM/Mb (Jones et al. 2019). This value is equivalent to a per base recombination rate of 1.02×10^{-8}. Both *Bombus* and *Megachile* are in the Apidae bee family together with *Anthophora*. Combining these estimates produces a recombination to mutation rate ratio of 2.5.

Sequentially Markovian coalescence methods like MSMC2 are optimised to capture variation in effective population size in the distant, rather than recent past (Nadachowska-Brzyska et al. 2022). We used GONE v. 1 (Santiago et al. 2020) to further understand how the effective population sizes of *Anthophora* species have changed over the past 100 generations. GONE

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

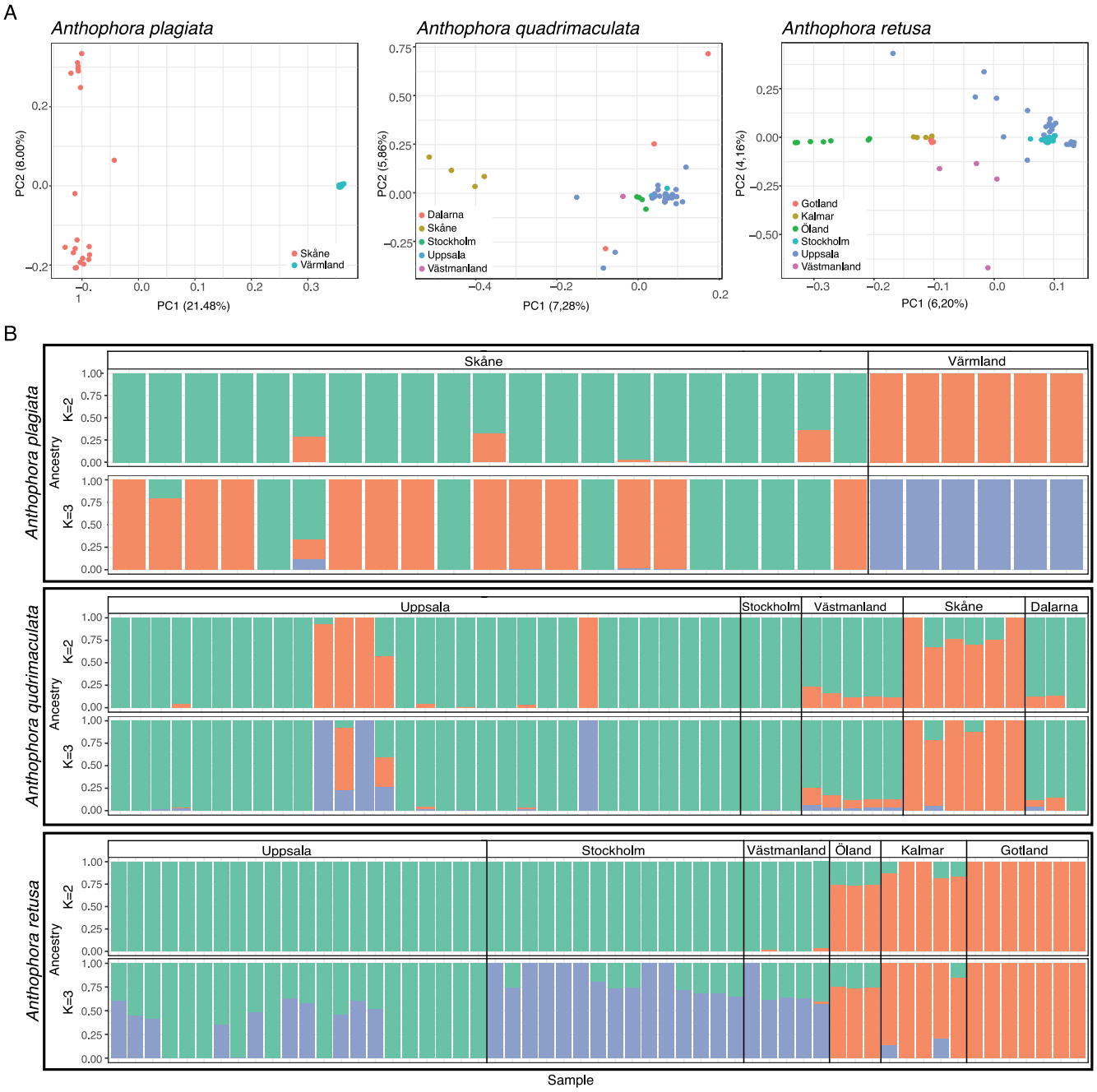

FIGURE 2 | Population structure of the three *Anthophora* species (A) PCA analysis per species. The colours of points reflect the geographic origin of each sample (B) Model-based structure analysis, as implemented in the program ADMIXTURE per species. Based on the cross-validation method, the most fit number of hypothetical ancestral groups was identified as $K=1$ for all species (see Table S2). Here, we present admixture patterns for two and three hypothetical ancestral groups.

implements a genetic algorithm to estimate the recent demographic history based on LD patterns between SNPs along the genome. Due to software limitations, we used the 199 largest scaffolds for each species for these estimations, covering 99.8% and 82.7% of *A. plagiata* , *A. retusa* , respectively. The assembly of *A. quadrimaculata* had less than 199 contigs (39); therefore, no reduction was necessary for that species. We used the software with the following parameters: (1) the data was treated as phased and we used a fixed recombination rate of 1 cM/Mb (Jones et al. 2019).

2.9 | Quantification of Long Runs of Identity-By-State

We estimate the distribution of long runs of identity-by-state (IBS) in genomes by scanning the genome for long homozygous fragments using the 'male-independent SNP dataset' in which males had been combined to generate pseudo-diploids. Prior to the detection of runs of homozygosity (ROH), SNPs with MAF ≥ 0.01 were filtered for each population. The software PLINK (Purcell et al. 2007) was used with the following parameters:

13655294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the minimum number of 50 SNPs in a ROH (–homozyg-snp 50), sliding windows of 50 SNPs (–homozyg-window-snp 50), allowance for not more than five missing SNPs (–homozyg-window-missing 5), and three heterozygous SNPs per window (–homozyg-window-het 3). The minimum length of an ROH segment was 300 kb (–homozyg-kb 300). The minimum SNP density was 1 SNP per 50 kb (–homozyg-density 50), and the maximum gap between two consecutive SNPs was 1000 kb (–homozyg-gap 1000). Only six drones were available for *A. plagiata* 's Värmland population. Therefore, we used every possible combination of males (a total of 15 combinations) to identify long runs of IBS. Sample size could bias the results of this analysis, making comparisons between populations and the other species difficult. Hence, we randomly selected 15 male pairs from all populations of each species, keeping them consistent across populations. Translocation of individuals between the two *A. plagiata* populations could potentially mitigate the effects of inbreeding. To estimate the effect of theoretical matings between individuals from different populations on the percentage of the genome found in IBS, we randomly selected 120 males from each of the two localities and paired them, creating an *A. plagiata* metapopulation. We then ran the IBS analysis as we did for each population. The significance of differences between these hypothetical individuals in the metapopulation and those in the original populations was assessed with an ANOVA analysis paired with a TukeyHSD post hoc test.

3 | Results

3.1 | Highly Contiguous Reference Genome Assemblies of Three *Anthophora* Species

A single haploid drone from each species was sequenced to provide data for genome assemblies. *A. quadrimaculata* and *A. retusa* were sequenced using PacBio HiFi technology, resulting in 8.6 and 6.7 Gbp of sequencing reads, respectively. *A. plagiata* was sequenced using Oxford Nanopore technology, resulting in 39.4 Gbp of sequencing reads. The *A. quadrimaculata* assembly (iyAntQuad1) consisted of 39 contigs, with a total genome length of 358 Mb and an N50 of 22.87 Mb. Genome completeness analysis showed 92.8% of BUSCO genes were complete (Simão et al. 2015) (Table 1). The *A. retusa* assembly (iyAntRet1) had 295 contigs and a total length of 325 Mb, with an N50 of 17.73 Mb. It had a higher BUSCO genome completeness at 97.5%. The *A. plagiata* assembly (iyAntPlag1) had 1422 contigs and a total length of 445 Mb, with an N50 of 12.66 Mb and a BUSCO completeness of 97.3%. In all assemblies, almost all BUSCO genes were present as single copy, with less than 1% in the fragmented and duplicated categories. Repeat analysis using RepeatModeler2 and RepeatMasker revealed high levels of repetitive sequences in all three genomes, with total repeats comprising 41.7% of the *A. quadrimaculata* assembly, 50.3% of the *A. retusa* assembly and 32.9% of the *A. plagiata* assembly (Table 1).

TABLE 1 | Assembly statistics.

Species	*A. plagiata*	*A. quadrimaculata*	*A. retusa*
Assembly	iyAntPlag1	iyAntQuad1	iyAntRet1
Sequencing technology	Nanopore	PacBio HiFi	PacBio HiFi
Total assembly size (Mbp)	325	358	445
Number of contigs	295	39	1422
Contig N50 (Mbp)	17.7	22.87	12.66
Contig N90 (Mbp)	2.45	10.52	0.1
Number of genes	11,064	10,494	11,000
Mean gene length (bp)	9940	10,332	10,622
BUSCO analysis[a]			
Complete BUSCO (%)	97.3	92.8	97.5
Single-copy (%)	97.1	92.7	97.3
Duplicated (%)	0.2	0.1	0.2
Fragmented (%)	0.4	0.6	0.5
Missing (%)	2.3	6.6	2.0
RepeatMasker analysis			
Total repeats (%)	32.9	41.7	50.3
DNA transposons (%)	4.97	4.42	3.88
Retroelements (%)	1.38	1.57	1.53
Unclassified repeats (%)	26.58	35.70	44.86

[a]BUSCO was run with the hymenoptera_odb10 lineage ($n = 5991$).

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Genome annotation across the three *Anthophora* species produced high-quality gene sets supported by multiple complementary evidence sources. Gene prediction combined RNA-seq, Arthropod protein homology and PacBio Iso-Seq evidence via BRAKER, GALBA and GeneMarkS-T, with final models filtered using TSEBRA. The resulting annotations were comparable across species, with 11,064, 10,495 and 11,000 protein-coding genes in *A. plagiata* , *A. quadrimaculata* and *A. retusa* , respectively. The mean gene lengths were similar for all species, ranging from 9.9 to 10.6 kb. Transcript counts and exon/intron length distributions were also similar, reflecting conserved genome architecture among these closely related bees. Overall, these results indicate high-confidence and well-supported gene annotations.

3.2 | Phylogenetic Relationships and the Emergence of the Genus *Anthophora*

We used the software Cactus to perform whole-genome alignment and estimate phylogenetic relationships between the *Anthophora* species (Figure 1D). To identify which substitution model was most appropriate for our data, we used the IQtree ModelFinder algorithm. The ModelFinder analysis indicated a general time reversible model with unequal rates (GTR) with empirical base frequencies (F), allowing for a proportion of invariant sites (I) and gamma-distributed rates with gamma-distributed rate (G4) nucleotide substitution model as the best fit for our dataset. The inclusion of the closely related species *Habropoda laboriosa* , of the *Anthophorini* tribe, allowed us to approximate the time of emergence of the genus *Anthophora* and other splits in the tree (Table S4). Estimations of divergence times were calculated from estimates of substitution rate and their 95% confidence intervals reported for *Bombus terrestris* (Liu et al. 2017). We estimate that the genus *Anthophora* split from *Habropoda* around 42.2 MYA (95% CI: 24 MYA—58 MYA). Among the *Anthophora* species used in this analysis, *A. plagiata* is the most distantly related. The phylogenetic analysis revealed a close relationship between the species *A. retusa* and *A. plumipes*, which diverged ca. 4.5 MYA (95% CI: 2.5 MYA—6.1 MYA). *A. quadrimaculata* diverged from *A. retusa* and *A. plumipes* around 7.6 MYA (95% CI: 4.2 MYA—10.4 MYA). *A. plagiata* and *A. retusa* diverged around 11.4 MYA (95% CI: 6.3 MYA—15.6 MYA).

3.3 | Contrasting Levels of Genome-Wide Variation Among Three *Anthophora* Species

To study the divergence and genetic diversity of *A. plagiata, A. retusa* and *A. quadrimaculata* and to infer the recent history of these species, we generated a population genomic dataset comprising sequence data of 136 individuals from Sweden (Table 2). The genomes were sequenced with Illumina technology, and sequencing reads were mapped to the newly generated reference genomes to identify SNPs. The average read coverage of genomes was 4.3×, 5.1× and 5.3×, respectively for *A. plagiata, A. retusa* and *A. quadrimaculata*.

Levels of genetic variation are highly contrasting between species. We recorded very low levels of diversity for *A. plagiata* ($\theta_w = 0.0002$), which were smaller than the other two *Anthophora* species ($\theta_w = 0.0034$ and $\theta_w = 0.0167$ for *A. retusa* and *A. quadrimaculata* , respectively) (Table 2). *A. plagiata* has experienced a drastic decline and is currently extremely rare, existing in a small number of isolated populations. *A. retusa* has also experienced a recent decline and recovery, but this does not appear to have strongly affected genetic variation. High levels of variation observed in *A. quadrimaculata* are consistent with its status as a widespread species without observed evidence of recent population fluctuation.

3.4 | Population Structure Analyses Revealed Fragmented *A. plagiata* Populations, but Extensive Historical Gene Flow for the Other Two Species in Sweden

We characterised the population genetic structure of the three *Anthophora* species based on complementary methods using genome-wide SNP data. Firstly, we investigated the extent of clustering using a principal component analysis (PCA) (Figure 2A). The PCA clustering mainly reflected the geographical origin of the samples. We further explored population structure by inferring the extent of shared ancestry using an ADMIXTURE analysis (Alexander et al. 2009) (Figure 2B). We performed a cross-validation error (CVE) analysis to select the most appropriate number of clusters (K). The CVE was minimised at $K = 2$ for *A. plagiata* , with clusters corresponding to two geographically structured populations. The Skåne and Värmland

TABLE 2 | Sample information and genetic variation statistics per species.

Species	Individuals	Females	Males	Genome size (Mbp)	SNPs	Waterson's theta (θ_W)	Tajima's *D*	*Ne* ($\theta_W/3\mu$)[a]
A. plagiata	27	0	27	325	22,312	0.0002 (±0.0001)	−0.399 (±1.08)	18,519
A. retusa	57	10	47	445	3,628,889	0.0034 (±0.0025)	0.675 (±1.08)	314,815
A. quadrimaculata	52	40	12	358	11,096,707	0.0167 (±0.0083)	0.782 (±1.09)	1,546,296

[a]Estimate of mutation rate per base per generation: $\mu = 3.6 \times 10^{-9}$ (Liu et al. 2017).

13655294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

populations of *A. plagiata* showed high genetic differentiation, with a genome-wide relative divergence (F_{ST}) value of 0.403, estimated by the ADMIXTURE analysis. The correlation between genetic and geographic distance was also supported by the Mantel test ($r = 0.78$, $p < 10^{-4}$; Figure S4). Although the Skåne population shows a tendency to be separated into two clusters on the second principal component (Figure 2A), these clusters do not correspond to geographical origin. For both *A. quadrimaculata* and *A. retusa*, CVE was minimised at K = 1, suggesting no significant genetic differentiation between the sampling localities (Tables S5–S7).

Despite a lack of significance, a limited degree of clustering related to geography can be observed in *A. retusa* . This species is observed most commonly around lake Mälaren in central Sweden, but is also found in southern Sweden and on the islands of Öland and Gotland. These populations are distinguishable with PCA, and the admixture analysis indicates geographical substructure between the lake Mälaren population and the remaining populations. Further evidence of a significant correlation between geographic and genetic distance was obtained by the Mantel test ($r = 0.63$, $p < 10^{-4}$; Figure S5). No signal of population substructure that corresponds to geography is evident in *A. quadrimaculata* was observed in our PCA and ADMIXTURE analysis, while a slight but significant correlation between genetic and geographic distance was supported by the Mantel test ($r = 0.54$, $p < 10^{-4}$; Figure S6). Taken together, these results highlight a lack of gene flow between the two populations of *A. plagiata* but indicate an absence of long-term barriers to gene flow between populations of *A. quadrimaculata* and *A. retusa* , reflecting historically well-connected sub-populations.

3.5 | Coalescent and LD-Based Methods Reveal Divergent Population Histories Between Species

We applied a Multiple Sequentially Markovian Coalescent (MSMC2) (Wang et al. 2020) approach to investigate population size variation through time. To this end, we considered a mutation rate (μ) of 3.6×10^{-9} (Liu et al. 2017) per base pair and assumed one generation per year. Population size changes were inferred using five haplotypes per species (Figure 3A). We found a decrease in effective population size in all three species since the end of the last interglacial period ~119 KYA. After this time point, we see a population decline in all species. *A. plagiata* populations remain constantly low, from roughly 10 KYA to 1 KYA. We found that the effective population size of *A. retusa* experienced a demographic expansion about 10 KYA. For *A. quadrimaculata* , the population decline followed by a more recent expansion occurred more gradually.

We used GONE v. 1 (Santiago et al. 2020) to infer how the effective population sizes of *Anthophora* species have changed over the past 100 generations (Figure 3B,C). The estimates are consistent with those from MSMC2. We infer that the *A. plagiata* effective population size (Ne) has remained dramatically low (Ne = 8000–10,000 individuals) in the recent past. This pattern was consistent for both populations. In line with observation data, we captured an increase of the *A. retusa* population in the recent past with a 30-fold growth in population size over the past 100 generations. The highest effective population size (Ne = 2–3 million individuals) and growth were observed for *A. quadrimaculata* . Together, our analyses of the effective population size fluctuations in the distant and recent past revealed a dramatically low long-term effective population size for *A. plagiata* populations. Moreover, we infer that Ne in both *A. retusa* and *A. quadrimaculata* populations has increased in the recent past.

3.6 | *A. plagiata* Populations Have Long Blocks of Identity-By-State in Their Genomes

The population histories we inferred are consistent with patterns of LD decay (Figure 4A). Analysis of the two *A. plagiata* populations revealed a greater extent of LD than the other species, with the r^2 statistic remaining above 0.2 in up to 36 and 75 kb for the Skåne and Värmland populations, respectively (Figure 4A). To assess the level of inbreeding for each species, we quantified the average genome proportion present in long blocks of identity-by-state (IBS) between any two chromosome copies. This method gives an estimation of runs of homozygosity in diploids assuming random mating but is also applicable to haploid samples. The total fraction of the genome presenting IBS in genomic regions > 1 Mb is much higher in *A. plagiata* populations (7%–9% of the genome) compared to *A. quadrimaculata* (2.6%) and *A. retusa* (2.4%) (Figure 4B), consistent with a greater degree of inbreeding in the *A. plagiata* populations.

We next compared overall levels of IBS in the two *A. plagiata* populations in Värmland and Skåne, considering only blocks > 1 Mbp. We found that the average level of IBS between haploid individuals in the Värmland population (9.0% of the genome) was significantly higher (ANOVA, $p < 2 \times 10^{-16}$, post hoc TukeyHSD, $p = 0.0002$, Tables S8 and S9) than in the Skåne population (7.2% of the genome). This reflects a higher degree of inbreeding and relatedness in random samples collected from the Värmland population. We next estimated levels of IBS in comparisons of haploid samples taken from the two populations. Levels were on average lower than the within-population comparisons (3.4%) (TukeyHSD, p-value < 0.0001, Table S9). This result suggests that runs of homozygosity in diploids would be reduced by translocation of individuals between populations.

4 | Discussion

The population genomic toolbox contains a powerful set of analyses that can be used to understand population history and guide management strategies. These methods are promising for understanding and ameliorating insect declines. A difficulty in interpreting patterns of genetic variation is in understanding the relative effects of species-specific life-history traits and recent population size changes in determining them. This study is designed to address this difficulty by comparing patterns of genetic variation in three closely related species, which have experienced contrasting recent population histories. We selected three bee species from the genus *Anthophora*, for which population trends in Sweden have been well-documented: *A. retusa* , *A. quadrimaculata* and *A. plagiata* . We present highly contiguous genome assemblies of these three species. We then used these assemblies to study genetic variation by sequencing the genomes of 136 additional samples. We performed a variety of analyses to

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

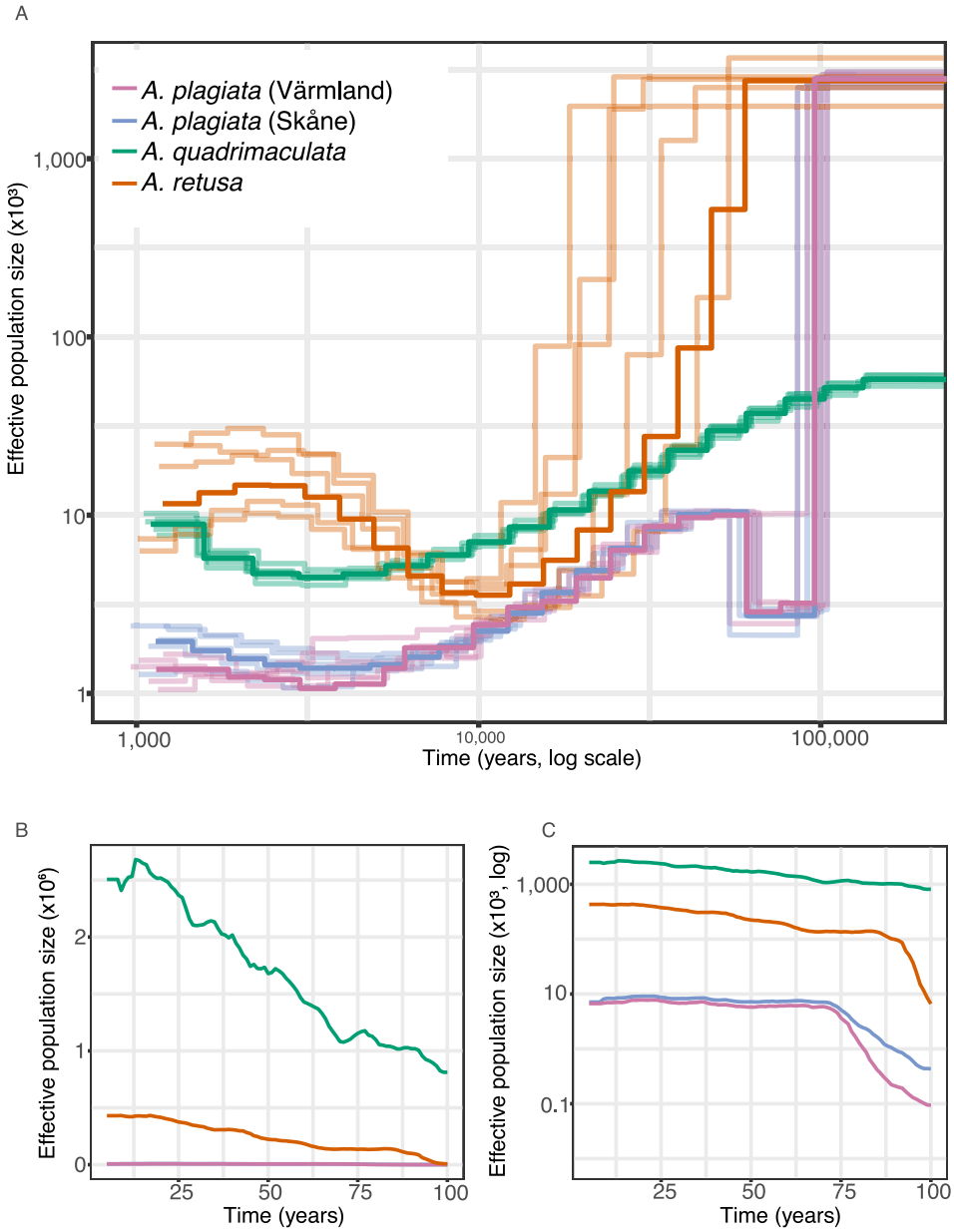

FIGURE 3 | Changes in effective population size (Ne) for the three *Anthophora* species. (A) We estimated changes in Ne using MSMC2. The axes were scaled with a mutation rate of 3.6×10^{-9} per site per generation and one generation per year. (B) Changes in effective population size over the past 100 generations inferred based on LD patterns with the software GONE.

infer both recent and ancient population history and compare our findings among species and with information on species abundance.

4.1 | Genome Assemblies of Three *Anthophora* Species

We used long-read technologies to generate high-quality assemblies of three solitary bee species. The *A. quadrimaculata* and *A. retusa* assemblies show high contiguity and completeness, with low contig counts, high N50 values (22.87 and 17.73 Mb, respectively), and the proportion of complete BUSCO genes (92.8% and 97.5%, respectively). The assembly for *A. plagiata* , sequenced using Oxford Nanopore technology, resulted in a higher number of contigs (1422) and a lower N50 value (12.66 Mb) than the other assemblies, but genome completeness remained high (97.3%), suggesting a near-complete gene set despite higher fragmentation. Repeat analysis using RepeatModeler2 and RepeatMasker revealed substantial levels of repetitive sequences across all three genomes, with repeat content ranging from 32.9% to 50.3%. These values are comparable to those observed in other solitary bees, such as the Mojave poppy bee (*Perdita meconis*), where repetitive elements comprise 37.3% of the genome (Schweizer et al. 2023), and *Tetrapedia diversipes* , in which repeats account for 38.7% of the genome (Santos et al. 2024). In these cases, and in the *Anthophora* species studied here, the majority of repeats are classified as 'unclassified' (Table 1), which means that the genomic regions are recognised as interspersed repeats but do not match any known repeat families. Because the observed differences in repeat content are largely confined to the unclassified fraction, their cause and functional significance are unclear.

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

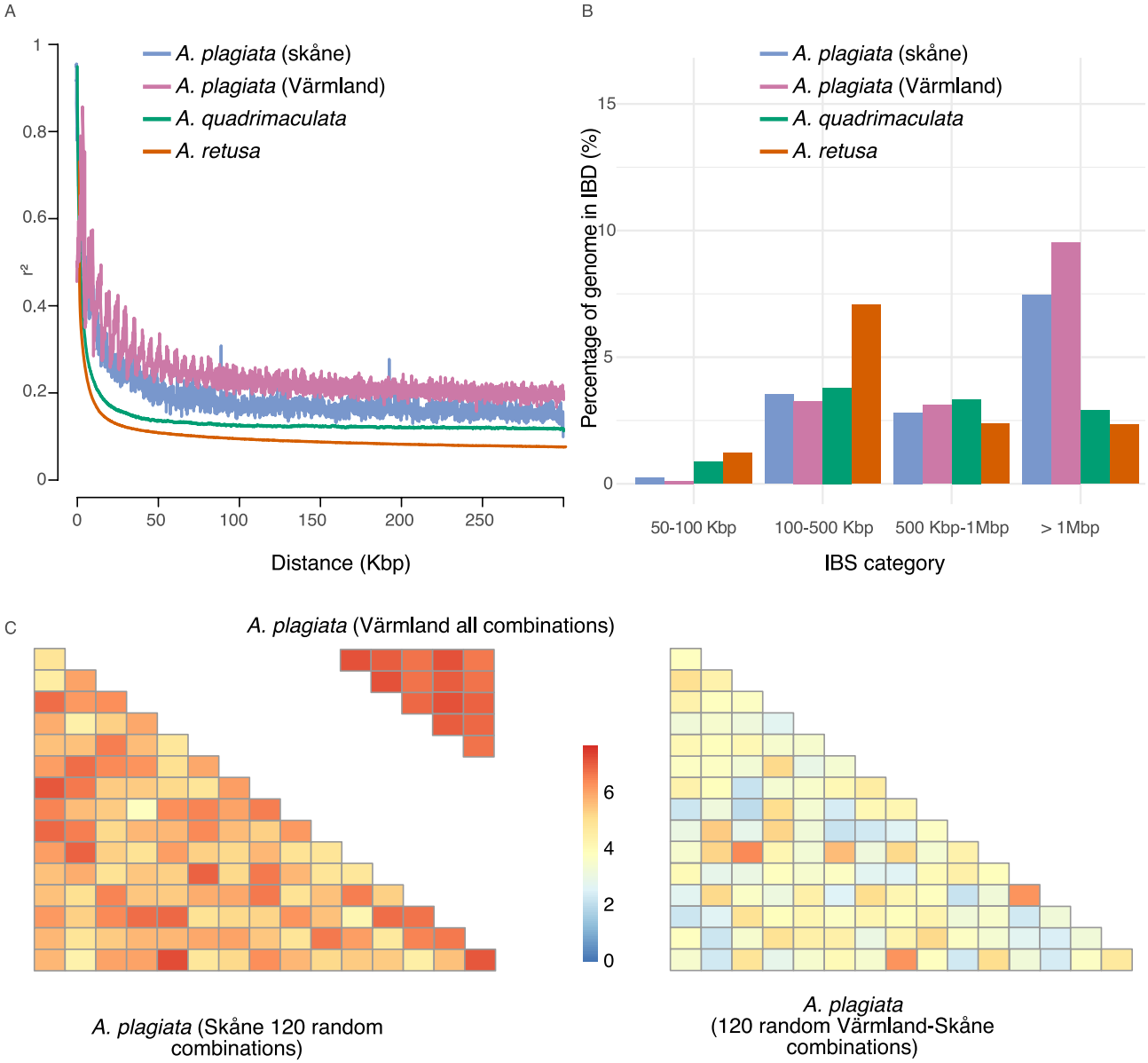

FIGURE 4 | Inbreeding statistics in *Anthophora*. (A) Linkage disequilibrium decay per species. (B) Percentage of the autosomal genome in IBS by size bins: Very short IBS (< 100 Kb), short IBS (100–500 Kb), intermediate IBS (500 Kb—1 Mb) and long (> 1 Mb). Short IBS are indicative of LD patterns, intermediate-size IBS reflect background inbreeding due to genetic drift and long IBS are indicative of recent inbreeding. (C) Pairwise comparisons of total level of IBS in stretches > 1 Mb in the two populations of *A. plagiata* , and in random comparisons of two individuals from different populations (measured in percentage).

By generating a whole-genome alignment of our new genomes and existing ones from the family Apidae, we could estimate the timing of lineage splits. We estimate that the last common ancestor of the *Bombus* and *Anthophora* genera at the base of the family Apidae lived around 104 million years ago. We estimate that the genus *Anthophora* split from its close relative genus *Habropoda* approximately 42.2 Mya. These findings are in line with previous estimations using genomics, individual gene and fossil information to estimate the emergence of the family Apidae and the genus *Anthophora* (Cardinal and Danforth 2013; Henríquez-Piskulich et al. 2024). They are concordant with the interpretation that the diversification of Apidae began in the mid-Cretaceous, and that the genus *Anthophora* emerged in the early Paleogene at a time when bees had already colonised the Northern Hemisphere from an origin in the Southern Hemisphere (Almeida et al. 2023).

Within the *Anthophora* genus, we find *A. retusa* and *A. plumipes* to be the most closely related, with a more divergent split from this pair and *A. quadrimaculata* , followed by *A. plagiata* , being the most distantly related among the studied species. These inferred phylogenetic relationships match relationships inferred based on morphological data (Brooks 1988). The divergence times (4.5–11.4 Mya) of the splits are also concordant with those based on more limited sequence data per species (Henríquez-Piskulich et al. 2024).

13655294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4.2 | Population Structure Consistent With Recent Expansion Scenarios for *A. retusa* and *A. quadrimaculata*

The three *Anthophora* species show contrasting population structure patterns reflecting their different population histories. *A. retusa* and *A. quadrimaculata* show limited population structure, with PCA and ADMIXTURE analyses indicating little differentiation among sampling localities in Sweden, suggesting historically well-connected populations. Despite the limited population structure, our isolation-by-distance analysis showed statistically significant correlations between genetic and geographic distance. The limited population structure is consistent with patterns observed in other insect species, such as the common and German wasps *Vespa vulgaris* and *V. germanica* in New Zealand (Schmack et al. 2019) and the congeneric black bark beetle *Hylugnus ligniperda* and *H. micklitzi* (Yuan et al. 2025). In *A. retusa*, only weak geographic clustering is detectable despite past regional declines, suggesting recent growth from remnant but connected populations rather than recolonization from a distant population. In contrast, *A. plagiata* exhibits deep population structure, with two highly divergent and genetically isolated populations in Skåne and Värmland ($F_{ST}=0.403$), extensive long runs of identity-by-state, and minimal contemporary gene flow, reflecting a severe bottleneck and population fragmentation. There is a slight indication that the Skåne population separates into two clusters in the PCA (Figure 2A), but these clusters do not reflect the geographical origin of the samples.

4.3 | Contrasting Levels of Genetic Variation Among Species Reflect Ancient Population History

Our analyses revealed contrasting levels of genetic diversity among the three species. *A. plagiata* exhibited extremely low diversity of $\theta_w=0.02\%$ per base, which implies a long-term Ne of around 20,000 (Table 2). Estimates in the other two species studied here are orders of magnitude larger than this: *A. retusa*, $\theta_w=0.34\%$ per base indicating Ne of around 300,000, and *A. quadrimaculata*, $\theta_w=1.6\%$ per base indicating Ne of around 1,500,000.

So far, most estimates of levels of genetic variation in bees come from social species. For example, the common bumble bee species *Bombus pascuorum* has $\pi=0.24\%$, and the rarer montane species have π close to 0.3% (Liu et al. 2023; Lozier et al. 2023). Populations of the honeybee *A. mellifera* have $\pi=0.3\%–0.8\%$ (Wallberg et al. 2014). Estimates of genetic variation in other social insects are also in this range, indicating that Ne in social insects is similar to vertebrates (Romiguier, Gayral, et al. 2014; Romiguier, Lourenco, et al. 2014). It is predicted that social species should have lower Ne than solitary ones with similar numbers of individuals due to the smaller number of reproductives. However, both solitary and social bees have been found to exhibit relaxed purifying selection, indicating they have similarly low Ne (Weyna and Romiguier 2021). In support of this, estimates of genetic variation in the solitary *Megachile rotundata* (Jones et al. 2019) and *Nomia melanderi* (Kapheim et al. 2019) are similar to *A. mellifera* . By contrast, much higher levels of heterozygosity are reported in butterfly species with $\pi=1\%–4\%$ (Reboud et al. 2023), indicating much higher Ne, which is typical of insects (Romiguier, Gayral, et al. 2014; Romiguier, Lourenco, et al. 2014).

Compared to currently available estimates of genetic variation in bees, it is therefore clear that the genetic variation we observe in *A. plagiata* is extremely low, consistent with its low abundance. By contrast, genetic variation in *A. retusa* appears typical for a bee species, whereas genetic variation in *A. quadrimaculata* is much higher than so far observed in any solitary bee and more similar to other insect clades, which typically have higher genetic variation than bees. However, there are currently few estimates of genome variation in solitary bees, and it is possible that such high variation is not unusual.

Estimates of Ne over the last century using patterns of LD with GONE (Santiago et al. 2020) suggest that populations of the species studied here have been relatively stable or increasing over this time period. In particular, there is no evidence for a drastic decline in the *A. plagiata* population, indicating that its low level of genetic variation reflects long-term population trends. However, it is possible that declines in recent tens of generations have not been captured by this analysis. The general trend for *A. retusa* over the last century is a population expansion. The observed population decline and recovery in the last 50 years in Sweden (Nilsson and Andersson 2007) are not captured by this analysis. It appears that these events have not had a major impact on overall genetic variation, and the GONE analysis may lack power to detect such recent fluctuations.

Over a longer timescale, our MSMC2 results infer that populations of all three species were likely much larger during the Eemian interglacial period that ended about 115,000 years ago and subsequently experienced declines. All species exhibit expansions of variable magnitude during the Holocene, with the largest fluctuations observed in *A. retusa* . Interestingly, similar patterns of decline since the last interglacial have been inferred using SMC methods in several extinct or endangered butterfly species such as the Xerces blue butterfly (*Glaucopsyche xerces*) (de-Dios et al. 2023), the Queen Alexandra's Birdwing (*Ornithoptera alexandrae*) (Reboud et al. 2023), and *Baronia brevicornis* (Marino et al. 2023). Further research is needed to indicate whether this trend is commonly observed in insects. Both *A. retusa* and *A. plagiata* have experienced bottlenecks in the past. Assessing the accumulation of deleterious mutations in these species would shed light on how these historical bottlenecks have shaped their genetic load.

4.4 | *Anthophora retusa* Has Maintained Large Connected Populations Despite a Recent Decline

Observations of *A. retusa* were extremely rare in Sweden during the period 1980–2006, when it was mainly observed on the Baltic islands of Öland and Gotland (Nilsson and Andersson 2007). Since this time, a substantial expansion has been observed, centered on two main areas: the southeast around Kalmar and the Mälardalen region including Stockholm. Despite evidence of a population bottleneck from observation data, our analysis does not indicate a decline in genetic variation or Ne. Our analysis did not identify significant evidence of fragmentation among the *A.*

13652949x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

retusa populations, indicating they have been historically connected. Surprisingly, the population from the island of Gotland clustered closely together with populations from the island of Öland and from Kalmar, despite Gotland being separated by ~60 km in the Baltic Sea, indicating the presence of some degree of gene flow. This isolation appears to have prevented gene flow in the Gotland population of the bumblebee *B. pascuorum* (Liu et al. 2023).

One question of interest is whether the population recovery was due to the migration of *A. retusa* from further south in Europe. We argue that this scenario is unlikely because substantial source populations of *A. retusa* are not known in Denmark. Furthermore, the weak clustering observed in Sweden suggests that population expansion likely occurred from existing populations rather than due to replacement from a common source. The reasons for the decline and subsequent recovery of these populations in recent decades are unclear; however, the population is currently healthy and has maintained high levels of genetic variation despite the observed recent population fluctuations. It is therefore unlikely that the population bottleneck has led to significant accumulation of deleterious mutations and elevated genetic load.

4.5 | *Anthophora plagiata* Exhibits Extreme Lack of Genetic Diversity in Isolated Populations

A. plagiata appeared to be absent from Sweden for much of the 20th century until it was reported in Skåne in 2004 and, more recently, in Värmland. Our admixture and PCA analyses suggest that these two populations are highly diverged, with an F_{ST} value of 0.403 between them, indicating limited or no gene flow. The populations have extremely low levels of genetic variation. As only haploid males were sampled from these populations, it was not possible to directly quantify the presence of runs of homozygosity. We therefore quantified the extent of the genome found in runs of identity-by-state (IBS) between random comparisons of haploid individuals. We identified extensive runs of IBS regions in *A. plagiata* compared to the other species, consistent with significant levels of inbreeding in these populations, with > 7% of the genome in homozygous blocks over 1 Mb. These runs of IBS were significantly more extensive in Värmland compared to the Skåne population.

The presence of long blocks of homozygosity is indicative of inbreeding and the presence of deleterious variants in homozygous state. This causes individuals to have reduced fitness due to realised genetic load (Bertorelle et al. 2022). A potential way to ameliorate these effects is by genetic rescue, whereby individuals from a more diverse population are translocated in order to restore variability. Our analysis indicates that mixing the two *A. plagiata* populations in this way would substantially reduce the extent of runs of IBS. Care must be taken when designing genetic rescue attempts, because more diverse populations maintain a larger number of masked deleterious variants. These can be introduced into the smaller target population and increase genetic load (Hasselgren et al. 2024; Norén and Hasselgren 2025). However, in this case both of the *A. plagiata* populations have limited variation and have likely experienced purging of highly deleterious variants. It is therefore likely that translocating individuals between populations would reduce inbreeding and genetic load and boost the viability of the populations. Levels of divergence between the Swedish populations and those in the rest of its distribution are not known.

5 | Conclusion

Here we describe new genome assemblies for three species of *Anthophora* solitary bees and use them to study genetic diversity in their populations in Sweden. We find significant interspecies differences in genetic diversity among species, reflecting differences in both ancient and recent demographic history. An observed population decline and recovery in *A. retusa* in the last fifty years appears not to have left a substantial impact on its genetic variation, and populations now do not seem to be genetically vulnerable. In contrast, we find extremely low genetic diversity and high inbreeding in fragmented populations of *A. plagiata* , a species that was absent in Sweden until twenty years ago. These results highlight the extreme vulnerability of these populations and suggest that conservation action such as translocations could improve their viability. Further studies of other species will provide a better picture of genetic variation in pollinators and other insects and guide conservation strategies (Webster et al. 2023).

Author Contributions

M.T.W. and B.C. conceived and designed the study. K. L.-T. and M.T.W. acquired funding for the study. D.T., A.E.R.S., G.D., I.B., M.P., A.O., M-B.M., J.H., N.L., A.-S.S., and M.P. performed experiments and analysis. M.T.W., K.L.-T., H.L., and O.V.P. coordinated the study. D.T. and M.T.W. wrote the manuscript.

Acknowledgements

This project was supported by the Swedish Research Council, project 2020-06174_VR. The computational analysis was enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), projects SNIC 2022/5-333 and SNIC 2023/23-465. We also acknowledge support from the National Genomics Infrastructure (NGI) funded by Science for Life Laboratory and the National Bioinformatics Infrastructure in Sweden (NBIS). We thank Adriana M. Cintrón Santiago for help with DNA extractions. We thank data stewards Yvonne Kallberg and Stephan Nylinder for uploading sequences to data repositories. We thank members of MW's research group for fruitful discussions.

Funding

This study was supported by Science for Life Laboratory; Vetenskapsrådet, 2020-06174_VR; National Academic Infrastructure for Supercomputing in Sweden, SNIC 2022/5-333 and SNIC 2023/23-465.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The genome assemblies, and raw DNA and RNA sequence data are available at ENA under the following accession numbers: A. quadrimaculata (iyAntQuad1), PRJEB72355; A. plagiata (iyAntPlag1), PRJEB72356; A. retusa (iyAntRet1), PRJEB72357. All Illumina DNA sequencing reads were deposited on NCBI under the BioProject PRJNA1255204. Scripts

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

are available at the GitHub repository: https://github.com/Jimi92/Popgen_of_anthophora.

References

Abascal, F., A. Corvelo, F. Cruz, et al. 2016. "Extreme Genomic Erosion After Recurrent Demographic Bottlenecks in the Highly Endangered Iberian Lynx." *Genome Biology* 17, no. 1: 251. https://doi.org/10.1186/s13059-016-1090-1.

Alexander, D. H., J. Novembre, and K. Lange. 2009. "Fast Model-Based Estimation of Ancestry in Unrelated Individuals." *Genome Research* 19, no. 9: 1655–1664. https://doi.org/10.1101/gr.094052.109.

Allendorf, F. W., P. A. Hohenlohe, and G. Luikart. 2010. "Genomics and the Future of Conservation Genetics." *Nature Reviews Genetics* 11, no. 10: 697–709. https://doi.org/10.1038/nrg2844.

Allio, R., A. Schomaker-Bastos, J. Romiguier, F. Prosdocimi, B. Nabholz, and F. Delsuc. 2020. "MitoFinder: Efficient Automated Large-Scale Extraction of Mitogenomic Data in Target Enrichment Phylogenomics." *Molecular Ecology Resources* 20, no. 4: 892–905. https://doi.org/10.1111/1755-0998.13160.

Almeida, E. A. B., S. Bossert, B. N. Danforth, et al. 2023. "The Evolutionary History of Bees in Time and Space." *Current Biology* 33, no. 16: 3409–3422.e6. https://doi.org/10.1016/j.cub.2023.07.005.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. "Basic Local Alignment Search Tool." *Journal of Molecular Biology* 215: 403–410.

Benson, G. 1999. "Tandem Repeats Finder: A Program to Analyze DNA Sequences." *Nucleic Acids Research* 27, no. 2: 573–580. https://doi.org/10.1093/nar/27.2.573.

Bertorelle, G., F. Raffini, M. Bosse, et al. 2022. "Genetic Load: Genomic Estimates and Applications in Non-Model Animals." *Nature Reviews Genetics* 23, no. 8: 492–503. https://doi.org/10.1038/s41576-022-00448-x.

Biesmeijer, J. C., S. P. M. Roberts, M. Reemer, et al. 2006. "Parallel Declines in Pollinators and Insect-Pollinated Plants in Britain and the Netherlands." *Science* 313, no. 5785: 351–354. https://doi.org/10.1126/science.1127863.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30, no. 15: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Brooks, R. W. 1988. "Systematics and Phylogeny of Anthophorine Bees (Hymenoptera; Anthophoridae: Anthophorini)." *University of Kansas Science Bulletin* 53, no. 9: 437–575.

Brůna, T., H. Li, J. Guhlin, et al. 2023. "Galba: Genome Annotation With Miniprot and AUGUSTUS." *BMC Bioinformatics* 24, no. 1: 327. https://doi.org/10.1186/s12859-023-05449-z.

Cardinal, S., and B. N. Danforth. 2013. "Bees Diversified in the Age of Eudicots." *Proceedings of the Royal Society B: Biological Sciences* 280, no. 1755: 20122686. https://doi.org/10.1098/rspb.2012.2686.

Cederberg, B. 2014. *Åtgärdsprogram för humlepälsbi*. Naturvårdsverket. https://www.naturvardsverket.se/publikationer/6600/atgardsprogram-humlepalsbi/.

Cheng, H., G. T. Concepcion, X. Feng, H. Zhang, and H. Li. 2021. "Haplotype-Resolved de Novo Assembly Using Phased Assembly Graphs With Hifiasm." *Nature Methods* 18, no. 2: 170–175. https://doi.org/10.1038/s41592-020-01056-5.

Crowley, L. M., O. Sivell, D. Sivell, et al. 2023. "The Genome Sequence of the Buff-Tailed Bumblebee, *Bombus terrestris* (Linnaeus, 1758)." *Wellcome Open Research* 8: 161. https://doi.org/10.12688/wellcomeopenres.19248.1.

Danecek, P., A. Auton, G. Abecasis, et al. 2011. "The Variant Call Format and VCFtools." *Bioinformatics* 27, no. 15: 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Danforth, B. N., R. L. Minckley, and J. L. Neff. 2019. *The Solitary Bees: Biology, Evolution, Conservation*. Princeton University Press. https://press.princeton.edu/books/hardcover/9780691168982/the-solitary-bees.

De Jong, J. F., P. Van Hooft, H.-J. Megens, et al. 2020. "Fragmentation and Translocation Distort the Genetic Landscape of Ungulates: Red Deer in the Netherlands." *Frontiers in Ecology and Evolution* 8: 535715. https://doi.org/10.3389/fevo.2020.535715.

de-Dios, T., C. Fontsere, P. Renom, et al. 2023. "Whole-Genomes From the Extinct Xerces Blue Butterfly Can Help Identify Declining Insect Species." *eLife* 12: RP87928. https://doi.org/10.7554/eLife.87928.2.

de-Dios, T., C. Fontsere, P. Renom, et al. 2024. "Whole Genomes From the Extinct Xerces Blue Butterfly Can Help Identify Declining Insect Species." *eLife* 12: RP87928. https://doi.org/10.7554/eLife.87928.

Flynn, J. M., R. Hubley, C. Goubert, et al. 2020. "RepeatModeler2 for Automated Genomic Discovery of Transposable Element Families." *Proceedings of the National Academy of Sciences* 117, no. 17: 9451–9457. https://doi.org/10.1073/pnas.1921046117.

Formenti, G., K. Theissinger, C. Fernandes, et al. 2022. "The Era of Reference Genomes in Conservation Genomics." *Trends in Ecology & Evolution* 37, no. 3: 197–202. https://doi.org/10.1016/j.tree.2021.11.008.

Gabriel, L., T. Brůna, K. J. Hoff, et al. 2024. "BRAKER3: Fully Automated Genome Annotation Using RNA-Seq and Protein Evidence With GeneMark-ETP, AUGUSTUS, and TSEBRA." *Genome Research* 34, no. 5: 769–777. https://doi.org/10.1101/gr.278090.123.

Gabriel, L., K. J. Hoff, T. Brůna, M. Borodovsky, and M. Stanke. 2021. "TSEBRA: Transcript Selector for BRAKER." *BMC Bioinformatics* 22, no. 1: 566. https://doi.org/10.1186/s12859-021-04482-0.

Gómez-Rubio, V. 2017. "ggplot2—Elegant Graphics for Data Analysis (2nd Edition)." *Journal of Statistical Software* 77: 1–3. https://doi.org/10.18637/jss.v077.b02.

Guan, D., S. A. McCarthy, J. Wood, K. Howe, Y. Wang, and R. Durbin. 2020. "Identifying and Removing Haplotypic Duplication in Primary Genome Assemblies." *Bioinformatics* 36, no. 9: 2896–2898. https://doi.org/10.1093/bioinformatics/btaa025.

Hallmann, C. A., M. Sorg, E. Jongejans, et al. 2017. "More Than 75 Percent Decline Over 27Years in Total Flying Insect Biomass in Protected Areas." *PLoS One* 12, no. 10: e0185809. https://doi.org/10.1371/journal.pone.0185809.

Hasselgren, M., N. Dussex, J. von Seth, A. Angerbjörn, L. Dalén, and K. Norén. 2024. "Strongly Deleterious Mutations Influence Reproductive Output and Longevity in an Endangered Population." *Nature Communications* 15, no. 1: 8378. https://doi.org/10.1038/s41467-024-52741-4.

Henríquez-Piskulich, P., A. F. Hugall, and D. Stuart-Fox. 2024. "A Supermatrix Phylogeny of the World's Bees (Hymenoptera: Anthophila)." *Molecular Phylogenetics and Evolution* 190: 107963. https://doi.org/10.1016/j.ympev.2023.107963.

Hoang, D. T., O. Chernomor, A. Von Haeseler, B. Q. Minh, and L. S. Vinh. 2018. "UFBoot2: Improving the Ultrafast Bootstrap Approximation." *Molecular Biology and Evolution* 35, no. 2: 518–522. https://doi.org/10.1093/molbev/msx281.

Hoffmann, A. A., V. L. White, M. Jasper, H. Yagui, S. J. Sinclair, and M. R. Kearney. 2021. "An Endangered Flightless Grasshopper With Strong Genetic Structure Maintains Population Genetic Variation Despite Extensive Habitat Loss." *Ecology and Evolution* 11, no. 10: 5364–5380. https://doi.org/10.1002/ece3.7428.

Hohenlohe, P. A., W. C. Funk, and O. P. Rajora. 2021. "Population Genomics for Wildlife Conservation and Management." *Molecular Ecology* 30, no. 1: 62–82. https://doi.org/10.1111/mec.15720.

IPBES. 2019. *Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity*

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and Ecosystem Services. Zenodo. https://doi.org/10.5281/zenodo.6417333.

Jones, J. C., A. Wallberg, M. J. Christmas, K. M. Kapheim, and M. T. Webster. 2019. "Extreme Differences in Recombination Rate Between the Genomes of a Solitary and a Social Bee." *Molecular Biology and Evolution* 36, no. 10: 2277–2291. https://doi.org/10.1093/molbev/msz130.

Jones, P., D. Binns, H.-Y. Chang, et al. 2014. "InterProScan 5: Genome-Scale Protein Function Classification." *Bioinformatics* 30, no. 9: 1236–1240. https://doi.org/10.1093/bioinformatics/btu031.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. Von Haeseler, and L. S. Jermiin. 2017. "ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates." *Nature Methods* 14, no. 6: 587–589. https://doi.org/10.1038/nmeth.4285.

Kapheim, K. M., H. Pan, C. Li, et al. 2015. "Genomic Signatures of Evolutionary Transitions From Solitary to Group Living." *Science* 348, no. 6239: 1139–1143. https://doi.org/10.1126/science.aaa4788.

Kapheim, K. M., H. Pan, C. Li, et al. 2019. "Draft Genome Assembly and Population Genetics of an Agricultural Pollinator, the Solitary Alkali Bee (Halictidae: *Nomia melanderi*)." *G3 (Bethesda, Md.)* 9, no. 3: 625–634. https://doi.org/10.1534/g3.118.200865.

Kebaïli, C., S. Sherpa, D. Rioux, and L. Després. 2022. "Demographic Inferences and Climatic Niche Modelling Shed Light on the Evolutionary History of the Emblematic Cold-Adapted Apollo Butterfly at Regional Scale." *Molecular Ecology* 31, no. 2: 448–466. https://doi.org/10.1111/mec.16244.

Klein, A.-M., B. E. Vaissière, J. H. Cane, et al. 2006. "Importance of Pollinators in Changing Landscapes for World Crops." *Proceedings of the Royal Society B: Biological Sciences* 274, no. 1608: 303–313. https://doi.org/10.1098/rspb.2006.3721.

Kolmogorov, M., J. Yuan, Y. Lin, and P. A. Pevzner. 2019. "Assembly of Long, Error-Prone Reads Using Repeat Graphs." *Nature Biotechnology* 37, no. 5: 540–546. https://doi.org/10.1038/s41587-019-0072-8.

Kriventseva, E. V., D. Kuznetsov, F. Tegenfeldt, et al. 2019. "OrthoDB v10: Sampling the Diversity of Animal, Plant, Fungal, Protist, Bacterial and Viral Genomes for Evolutionary and Functional Annotations of Orthologs." *Nucleic Acids Research* 47, no. D1: D807–D811. https://doi.org/10.1093/nar/gky1053.

Laetsch, D. R., and M. L. Blaxter. 2017. "BlobTools: Interrogation of Genome Assemblies." *F1000Research* 6: 1287. https://doi.org/10.12688/f1000research.12232.1.

LeBuhn, G., and J. Vargas Luna. 2021. "Pollinator Decline: What Do We Know About the Drivers of Solitary Bee Declines?" *Current Opinion in Insect Science* 46: 106–111. https://doi.org/10.1016/j.cois.2021.05.004.

Li, H., and R. Durbin. 2010. "Fast and Accurate Long-Read Alignment With Burrows-Wheeler Transform." *Bioinformatics* 26, no. 5: 589–595. https://doi.org/10.1093/bioinformatics/btp698.

Li, H. 2018. "Minimap2: Pairwise Alignment for Nucleotide Sequences." *Bioinformatics* 34, no. 18: 3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

Li, H. 2023. "Protein-To-Genome Alignment With Miniprot." *Bioinformatics* 39, no. 1: btad014. https://doi.org/10.1093/bioinformatics/btad014.

Liu, H., Y. Jia, X. Sun, D. Tian, L. D. Hurst, and S. Yang. 2017. "Direct Determination of the Mutation Rate in the Bumblebee Reveals Evidence for Weak Recombination-Associated Mutation and an Approximate Rate Constancy in Insects." *Molecular Biology and Evolution* 34, no. 1: 119–130. https://doi.org/10.1093/molbev/msw226.

Liu, Y., A. Olsson, T. Larva, et al. 2023. "Genomic Variation in Montane Bumblebees in Scandinavia: High Levels of Intraspecific Diversity Despite Population Vulnerability." *Molecular Ecology* 33: e17251. https://doi.org/10.1111/mec.17251.

Lomsadze, A., P. D. Burns, and M. Borodovsky. 2014. "Integration of Mapped RNA-Seq Reads Into Automatic Training of Eukaryotic Gene Finding Algorithm." *Nucleic Acids Research* 42, no. 15: e119. https://doi.org/10.1093/nar/gku557.

Lowe, T. M., and S. R. Eddy. 1997. "tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence." *Nucleic Acids Research* 25, no. 5: 955–964.

Lozier, J. D., J. P. Strange, and S. D. Heraghty. 2023. "Whole Genome Demographic Models Indicate Divergent Effective Population Size Histories Shape Contemporary Genetic Diversity Gradients in a Montane Bumble Bee." *Ecology and Evolution* 13, no. 2: e9778. https://doi.org/10.1002/ece3.9778.

Magrane, M., and Consortium, U. 2011. "UniProt Knowledgebase: A Hub of Integrated Protein Data." *Database* 2011: bar009. https://doi.org/10.1093/database/bar009.

Marino, A., E. L. Reboud, E. Chevalier, et al. 2023. "Genomics of the Relict Species Baronia Brevicornis Sheds Light on Its Demographic History and Genome Size Evolution Across Swallowtail Butterflies." *G3 (Bethesda, Md.)* 13, no. 12: jkad239. https://doi.org/10.1093/g3journal/jkad239.

McKenna, A., M. Hanna, E. Banks, et al. 2010. "The Genome Analysis Toolkit: A MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data." *Genome Research* 20: 1297–1303. https://doi.org/10.1101/gr.107524.110.20.

Minh, B. Q., H. A. Schmidt, O. Chernomor, et al. 2020. "IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era." *Molecular Biology and Evolution* 37, no. 5: 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Nadachowska-Brzyska, K., M. Konczal, and W. Babik. 2022. "Navigating the Temporal Continuum of Effective Population Size." *Methods in Ecology and Evolution* 13, no. 1: 22–41. https://doi.org/10.1111/2041-210X.13740.

Nilsson, L. A., and H. Andersson. 2007. *Åtgärdsprogram för svartpälsbi*. Naturvårdsverket. https://www.naturvardsverket.se/publikationer/5700/atgardsprogram-svartpalsbi/.

Norén, K., and M. Hasselgren. 2025. "To Genetic Rescue or Not?" *Trends in Genetics* 41, no. 3: 185–186. https://doi.org/10.1016/j.tig.2024.11.004.

Ouborg, N. J., C. Pertoldi, V. Loeschcke, R. K. Bijlsma, and P. W. Hedrick. 2010. "Conservation Genetics in Transition to Conservation Genomics." *Trends in Genetics: TIG* 26, no. 4: 177–187. https://doi.org/10.1016/j.tig.2010.01.001.

Paradis, E., and K. Schliep. 2019. "Ape 5.0: An Environment for Modern Phylogenetics and Evolutionary Analyses in R." *Bioinformatics* 35, no. 3: 526–528. https://doi.org/10.1093/bioinformatics/bty633.

Paten, B., D. Earl, N. Nguyen, M. Diekhans, D. Zerbino, and D. Haussler. 2011. "Cactus: Algorithms for Genome Multiple Sequence Alignment." *Genome Research* 21, no. 9: 1512–1528. https://doi.org/10.1101/gr.123356.111.

Paysan-Lafosse, T., M. Blum, S. Chuguransky, et al. 2023. "InterPro in 2022." *Nucleic Acids Research* 51, no. D1: D418–D427. https://doi.org/10.1093/nar/gkac993.

Potts, S. G., J. C. Biesmeijer, C. Kremen, P. Neumann, O. Schweiger, and W. E. Kunin. 2010. "Global Pollinator Declines: Trends, Impacts and Drivers." *Trends in Ecology & Evolution* 25: 345–353. https://doi.org/10.1016/j.tree.2010.01.007.

Powney, G. D., C. Carvell, M. Edwards, et al. 2019. "Widespread Losses of Pollinating Insects in Britain." *Nature Communications* 10, no. 1: 1018. https://doi.org/10.1038/s41467-019-08974-9.

Purcell, S., B. Neale, K. Todd-brown, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81: 559–575. https://doi.org/10.1086/519795.

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Rasmont, P., and M. Dehon. 2014a. "IUCN Red List of Threatened Species: *Anthophora plagiata*. IUCN Red List of Threatened Species." https://www.iucnredlist.org/en.

Rasmont, P., and M. Dehon. 2014b. "IUCN Red List of Threatened Species: *Anthophora quadrimaculata*. IUCN Red List of Threatened Species." https://www.iucnredlist.org/en.

Rasmont, P., and M. Dehon. 2014c. "IUCN Red List of Threatened Species: *Anthophora retusa*. IUCN Red List of Threatened Species." https://www.iucnredlist.org/en.

Reboud, E. L., B. Nabholz, E. Chevalier, M. Tilak, D. Bito, and F. L. Condamine. 2023. "Genomics, Population Divergence, and Historical Demography of the World's Largest and Endangered Butterfly, the Queen Alexandra's Birdwing." *Genome Biology and Evolution* 15, no. 4: evad040. https://doi.org/10.1093/gbe/evad040.

Reilly, J. R., D. R. Artz, D. Biddinger, et al. 2020. "Crop Production in the USA Is Frequently Limited by a Lack of Pollinators." *Proceedings of the Royal Society B: Biological Sciences* 287, no. 1931: 20200922. https://doi.org/10.1098/rspb.2020.0922.

Romiguier, J., P. Gayral, M. Ballenghien, et al. 2014. "Comparative Population Genomics in Animals Uncovers the Determinants of Genetic Diversity." *Nature* 515, no. 7526: 261–263. https://doi.org/10.1038/nature13685.

Romiguier, J., J. Lourenco, P. Gayral, et al. 2014. "Population Genomics of Eusocial Insects: The Costs of a Vertebrate-Like Effective Population Size." *Journal of Evolutionary Biology* 27, no. 3: 593–603. https://doi.org/10.1111/jeb.12331.

Sánchez-Bayo, F., and K. A. G. Wyckhuys. 2019. "Worldwide Decline of the Entomofauna: A Review of Its Drivers." *Biological Conservation* 232: 8–27. https://doi.org/10.1016/j.biocon.2019.01.020.

Santiago, E., I. Novo, A. F. Pardiñas, M. Saura, J. Wang, and A. Caballero. 2020. "Recent Demographic History Inferred by High-Resolution Analysis of Linkage Disequilibrium." *Molecular Biology and Evolution* 37, no. 12: 3642–3653. https://doi.org/10.1093/molbev/msaa169.

Santos, P. K. F., N. de Souza Araujo, E. Françoso, J. H. Werren, K. M. Kapheim, and M. C. Arias. 2024. "The Genome of the Solitary Bee *Tetrapedia diversipes* (Hymenoptera, Apidae)." *G3: Genes, Genomes, Genetics* 15, no. 2: jkae264. https://doi.org/10.1093/g3journal/jkae264.

Saremi, N. F., M. A. Supple, A. Byrne, et al. 2019. "Puma Genomes From North and South America Provide Insights Into the Genomic Consequences of Inbreeding." *Nature Communications* 10, no. 1: 4769. https://doi.org/10.1038/s41467-019-12741-1.

Schmack, J. M., E. C. Brenton-Rule, R. Veldtman, et al. 2019. "Lack of Genetic Structuring, Low Effective Population Sizes and Major Bottlenecks Characterise Common and German Wasps in New Zealand." *Biological Invasions* 21, no. 10: 3185–3201. https://doi.org/10.1007/s10530-019-02039-0.

Schweizer, R. M., C. G. Meidt, L. R. Benavides, et al. 2023. "Reference Genome for the Mojave Poppy Bee (*Perdita meconis*), a Specialist Pollinator of Conservation Concern." *Journal of Heredity* 115, no. 4: 470–479. https://doi.org/10.1093/jhered/esad076.

Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov. 2015. "BUSCO: Assessing Genome Assembly and Annotation Completeness With Single-Copy Orthologs." *Bioinformatics* 31, no. 19: 3210–3212. https://doi.org/10.1093/bioinformatics/btv351.

SLU Artdatabanken. 2020. *Rödlistade arter i Sverige*. SLU. https://www.slu.se/artdatabanken/publikationer/rodlistor/rodlista-2020/.

Smeds, L., and H. Ellegren. 2023. "From High Masked to High Realized Genetic Load in Inbred Scandinavian Wolves." *Molecular Ecology* 32, no. 7: 1567–1580. https://doi.org/10.1111/mec.16802.

Soroye, P., T. Newbold, and J. Kerr. 2020. "Climate Change Contributes to Widespread Declines Among Bumble Bees Across Continents." *Science* 367, no. 6478: 685–688. https://doi.org/10.1126/science.aax8591.

Stanke, M., O. Keller, I. Gunduz, A. Hayes, S. Waack, and B. Morgenstern. 2006. "AUGUSTUS: Ab Initio Prediction of Alternative Transcripts." *Nucleic Acids Research* 34: W435–W439. https://doi.org/10.1093/nar/gkl200.

Supple, M. A., and B. Shapiro. 2018. "Conservation of Biodiversity in the Genomics Era." *Genome Biology* 19, no. 1: 131. https://doi.org/10.1186/s13059-018-1520-3.

Tang, S., A. Lomsadze, and M. Borodovsky. 2015. "Identification of Protein Coding Regions in RNA Transcripts." *Nucleic Acids Research* 43, no. 12: e78. https://doi.org/10.1093/nar/gkv227.

Theissinger, K., C. Fernandes, G. Formenti, et al. 2023. "How Genomics Can Help Biodiversity Conservation." *Trends in Genetics* 39, no. 7: 545–559. https://doi.org/10.1016/j.tig.2023.01.005.

Uliano-Silva, M., J. G. R. N. Ferreira, K. Krasheninnikova, et al. 2023. "MitoHiFi: A Python Pipeline for Mitochondrial Genome Assembly From PacBio High Fidelity Reads." *BMC Bioinformatics* 24, no. 1: 288. https://doi.org/10.1186/s12859-023-05385-y.

van Klink, R., D. E. Bowler, K. B. Gongalsky, A. B. Swengel, A. Gentile, and J. M. Chase. 2020. "Meta-Analysis Reveals Declines in Terrestrial but Increases in Freshwater Insect Abundances." *Science* 368, no. 6489: 417–420. https://doi.org/10.1126/science.aax9931.

Wagner, D. L. 2020. "Insect Declines in the Anthropocene." *Annual Review of Entomology* 65: 457–480. https://doi.org/10.1146/annurev-ento-011019-025151.

Wallberg, A., F. Han, G. Wellhagen, et al. 2014. "A Worldwide Survey of Genome Sequence Variation Provides Insight Into the Evolutionary History of the Honeybee *Apis mellifera*." *Nature Genetics* 46, no. 10: 1081–1088. https://doi.org/10.1038/ng.3077.

Wang, K., I. Mathieson, J. O'Connell, and S. Schiffels. 2020. "Tracking Human Population Structure Through Time From Whole Genome Sequences." *PLoS Genetics* 16, no. 3: e1008552. https://doi.org/10.1371/journal.pgen.1008552.

Watterson, G. A. 1975. "On the Number of Segregating Sites in Genetical Models Without Recombination." *Theoretical Population Biology* 7: 256–276.

Webster, M. T., A. Beaurepaire, P. Neumann, and E. Stolle. 2023. "Population Genomics for Insect Conservation." *Annual Review of Animal Biosciences* 11, no. 1: 115–140. https://doi.org/10.1146/annurev-animal-122221-075025.

Weyna, A., and J. Romiguier. 2021. "Relaxation of Purifying Selection Suggests Low Effective Population Size in Eusocial Hymenoptera and Solitary Pollinating Bees." *Peer Community Journal* 1: e2. https://doi.org/10.24072/pcjournal.3.

Yuan, Y., A. Roques, E. Magnoux, et al. 2025. "One Is a Worldwide Invader, the Other Never Established in Other Continents: Contrasted Phylogeographic Patterns of Two Congeneric Hylurgus Conifer Bark Beetles." *NeoBiota* 101: 91–117. https://doi.org/10.3897/neobiota.101.150337.

Zhang, C., S. S. Dong, J. Y. Xu, W. M. He, and T. L. Yang. 2019. "PopLDdecay: A Fast and Effective Tool for Linkage Disequilibrium Decay Analysis Based on Variant Call Format Files." *Bioinformatics* 35, no. 10: 1786–1788. https://doi.org/10.1093/bioinformatics/bty875.

Zheng, X., D. Levine, J. Shen, S. M. Gogarten, C. Laurie, and B. S. Weir. 2012. "A High-Performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data." *Bioinformatics* 28, no. 24: 3326–3328. https://doi.org/10.1093/bioinformatics/bts606.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** mec70204-sup-0001-DataS1.pdf. **Figure S1:** Global distribution of *A. plagiata* . Information on global observations of the species were obtained from the Global

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Biodiversity Information facility (GBIF) database. The data used for the creation of this map can be downloaded using this link: https://doi.org/10.15468/dl.ss8rp7. **Figure S2:** Global distribution of *A. quadrimaculata* . Information on global observations of the species were obtained from the Global Biodiversity Information facility (GBIF) database. The data used for the creation of this map can be downloaded using this link: https://doi.org/10.15468/dl.pgsc9q. **Figure S3:** Global distribution of *A. retusa* . Information on global observations of the species were obtained from the Global Biodiversity Information facility (GBIF) database. The data used for the creation of this map can be downloaded using this link: https://doi.org/10.15468/dl.syqd28. **Figure S4:** Mantel test shows significant correlation between geographic and genetic distance in *A. plagiata* ($r=0.78$, $p<10^{-4}$). **Figure S5:** Mantel test shows significant correlation between geographic and genetic distance in *A. retusa* ($r=0.63$, $p<10^{-4}$). **Figure S6:** Mantel test shows significant correlation between geographic and genetic distance in *A. quadrimaculata* ($r=0.54$, $p<10^{-4}$). **Table S1:** Sampling and alignment information for *A. plagiata* . **Table S2:** Sampling and alignment information for *A. quadrimaqulata*. **Table S3:** Sampling and alignment information for *A. retusa* . **Table S4:** Divergence times of *Anthophora* species. **Table S5:** Admixture cross-validation error for *A. plagiata* . **Table S6:** Admixture cross-validation error for *A. quadrimaculata* . **Table S7:** Admixture cross-validation error for *A. retusa* . **Table S8:** ANOVA test on the percentage of genome in IBS in different *A. plagiata* populations. **Table S9:** TukeyHSD post hoc test for pairwise comparisons of the percentage of IBS in *A. plagiata* populations.

1365294x, 2025, 24, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70204 by Swedish University Of Agricultural Sciences, Wiley Online Library on [15/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License