Journal of Heredity, 2026, **117**, 611–622
https://doi.org/10.1093/jhered/esaf109
Advance access publication 22 December 2025
Genome Resources

American Genetic Association

OXFORD

Genome Resources

Genome assemblies of two bumble bee species: *Bombus sonorus* and *Bombus vosnesenskii*

Kimberly M. Ballare[1,2,*,†], Merly Escalona[3,4,†], Kaleigh Fisher[5,6], Blanca M. Guillén[6], Leif Richardson[7], William Seligmann[1], Noravit Chumchim[8], Mohan P.A. Marimuthu[8], Oanh H. Nguyen[8], Neal M. Williams[9], Beth Shapiro[1,10], S. Hollis Woodard[6]

[1]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA, United States, [2]World Wildlife Fund, Great Plains Program, Bozeman, MT, United States, [3]Department of Biomolecular Engineering, University of California, Santa Cruz, CA United States, [4]Institute for Systems Genomics, University of Connecticut, Storrs, CT, United States, [5]Department of Evolution and Ecology, University of California, Davis, CA, United States, [6]Department of Entomology, University of California, Riverside, CA, United States, [7]Xerces Society for Invertebrate Conservation, Riverside, CA, United States, [8]DNA Technologies and Expression Analysis Core Laboratory, Genome Center, University of California, Davis, CA, United States, [9]Department of Entomology and Nematology, University of California, Davis, CA, United States, [10]Colossal Biosciences, Santa Cruz, CA, United States

*Corresponding author: Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA, United States. Email: kim.ballare@gmail.com

†Authors Ballare and Escalona contributed equally to this work and can prioritize their own names when listing this reference on their CVs.

Abstract

Globally, many bumble bee species are declining in abundance due to diverse factors that include habitat loss, pesticide exposure, disease, and climate change. Some species' populations, however, appear to be relatively stable or even increasing, with reasons for this disparity unclear. Increased genomic resources for different bumble bee species will aid in identifying any genetic causes of these differences. We assembled highly contiguous and complete de novo genomes of two species with different levels of population imperilment: the Sonoran bumble bee (*Bombus sonorus* Say 1837) and the Vosnesenky or yellow-faced bumble bee (*Bombus vosnesenskii* Radoszkowski 1862). Here we present genomes for both species assembled with Pacific Biosciences HiFi long reads and Dovetail Omni-C data. These genomes are at the scaffold level, but contain multiple chromosome-length scaffolds, and have similar or higher levels of contiguity of other published chromosome-level *Bombus* genomes. These genomic resources support conservation planning for both important pollinator species by facilitating future resequencing efforts to understand genetic variation within the species.

Key words: California conservation genomics project, CCGP, Sonoran bumble bee, Vosnesensky bumble bee, yellow-faced bumble bee

Introduction

Bumble bees, comprising the genus *Bombus* Latreille 1802 (Hymenoptera: Apidae), are a group of social insects that are vital pollinators for many wild plants (Goulson et al. 2008; Jha et al. 2013) and economically important crops (Garibaldi et al. 2013; Strange 2015). Recent widespread declines in bumble bee abundance and range restrictions are driven by diverse factors including habitat loss, climate change, disease, and pesticide exposure (Potts et al. 2010; Cameron et al. 2011; Jha and Kremen 2013; Hatfield et al. 2014, 2021; Goulson et al. 2015; Maxfield-Taylor et al. 2015; Soroye et al. 2020; IUCN Bumblebee Specialist Group 2023). The global decline of bumble bees raises concerns for the conservation of this insect group, as well as for plant reproduction, human food security, and ecosystem stability.

At least one quarter of bumble bee species are imperiled, which is a higher proportion of declining species than many other bee groups whose extinction risk has been evaluated (Cameron and Sadd 2020). Some bumble bee species, however, appear to be relatively stable, and a small number are even exhibiting population increases and range expansions (Cameron et al. 2011). The reasons for the disparity between declining and stable species remain largely unclear (but see Goulson et al. 2005; Williams et al., 2009; Williams and Osborne, 2009; Arbetman et al. 2017). Increased genomic resources for different bumble bee species will aid in determining potential genetic causes of these differences, such as environmental signatures of selection, as well as help identify distinct conservation units based on genetic data. The latter is a key component of bumble bee species protection and recovery efforts (e.g. U.S. Fish and Wildlife Service 2021; Mola et al. 2024). Here we present the genomes of one declining species, the Sonoran bumble bee (*Bombus sonorus* Say 1837), and the most widespread and abundant bumble bee species in Western North America, the Vosnesensky bumble bee (Thorp

Received on 24 January 2025; revised on 14 November 2025; accepted on 18 December 2025

et al. 1983; Fisher et al. 2022), also known as the yellow-faced bumble bee (*Bombus vosnesenskii* Radoszkowski 1862).

B. sonorus is primarily found in open grasslands, shrublands, and farmlands across its range spanning from the California Central Valley south to Baja California and Central Mexico, and east to West Texas (Thorp et al. 1983; Warriner 2012; California Bumble Bee Atlas 2025; Fig. 1A). Although it is currently recognized within the Integrated Taxonomic Information System (ITIS) as a valid species, some experts consider it to be conspecific with the American bumble bee *Bombus pensylvanicus* De Geer 1773 (Williams 1998; Williams et al. 2014). The primary physical distinction between the species is that *B. sonorus* has more yellow-colored hairs on its abdomen, and recent molecular work supports its status as a distinct species (Beckham et al. 2024; but see Williams et al. 2024). Much of the evidence for the decline of *B. sonorus*, however, comes from assessments of *B. pensylvanicus* that consider *B. sonorus* a synonym (Hatfield et al. 2015). There are well-documented declines of *B. pensylvanicus* in the eastern U.S. and Canada (Colla and Packer 2008; Cameron et al. 2011; Warriner 2012; MacPhail et al. 2019) as well as declines in suitable habitat across the *B. pensylvanicus* range (Carrasco et al. 2021). *Bombus pensylvanicus* sensu lato is currently under consideration for protection under the U.S. Endangered Species Act (2021), and this species is currently listed as Vulnerable by the IUCN and NatureServe (G3 ranking) with occurrence data for *B. sonorus* included in the analysis of *B. pensylvanicus* (Hatfield et al. 2015). Within California, *B. sonorus* was historically (before 2005) found throughout much of the state, but since this time it has become extremely geographically restricted to the far southern coastal part of the state (Fig. 1A). There are currently no published genomic resources for *B. sonorus*. Further genetic resources for *B. sonorus* may help resolve its taxonomic status as distinct or as a conspecific of *B. pensylvanicus* and improve conservation planning by revealing geographic patterns of dispersal, relatedness, and genetic structure.

In contrast, *B. vonsenesenskii* populations appear to have remained stable, and possibly increased in relative abundance in comparison to many other bumble bee species (Cameron et al. 2011; Fraser et al. 2012; Fisher et al. 2022). *B. vosnesenskii* inhabits a wide variety of ecoregions, including high montane areas and low elevation regions, such as dry coastal and inland sage scrub. Its range spans from British Columbia in the north to Baja California in the south, and it occupies much of California (Fig. 1B). Despite its relatively large population size compared to other bumble bee species, *B. vosnesenskii* exhibits high levels of male diploidy (Schenau and Jha 2016), a marker of inbreeding for haplodiploid species such as bees (Zayed et al. 2005), which raises concerns for its genetic and population health. Existing genomic resources for *B. vosnesenskii* are limited, with one draft genome available (Heraghty et al. 2020).

We present highly contiguous and complete de novo genome assemblies for the Sonoran and yellow-faced bumble bees, generated from wild-caught individuals, using Pacific Biosciences (PacBio, CA) HiFi long reads and Omni-C data. The final genome for *B. sonorus* is 0.37 Gb across 282 scaffolds, with a scaffold N50 of 16.2 Mb and a BUSCO score of 97.6%. The genome for *B. vosnesenskii* is 0.46 Gb across 433 scaffolds, scaffold N50 of 12.5 Mb, and a BUSCO score of 97.5%. These genomes were generated as part of a larger effort to assemble genomes for diverse taxa across the state of California (The California Conservation Genomics Project—CCGP; Shaffer et al. 2022).

Methods

Biological materials

Two male specimens of *B. vosnesenskii* (UCR museum accession IDs 888-SC and 887-D) and two female specimens of *B. sonorus* (UCR museum accession IDs: 021S and 022S) were collected in the summer of 2021 via opportunistic hand

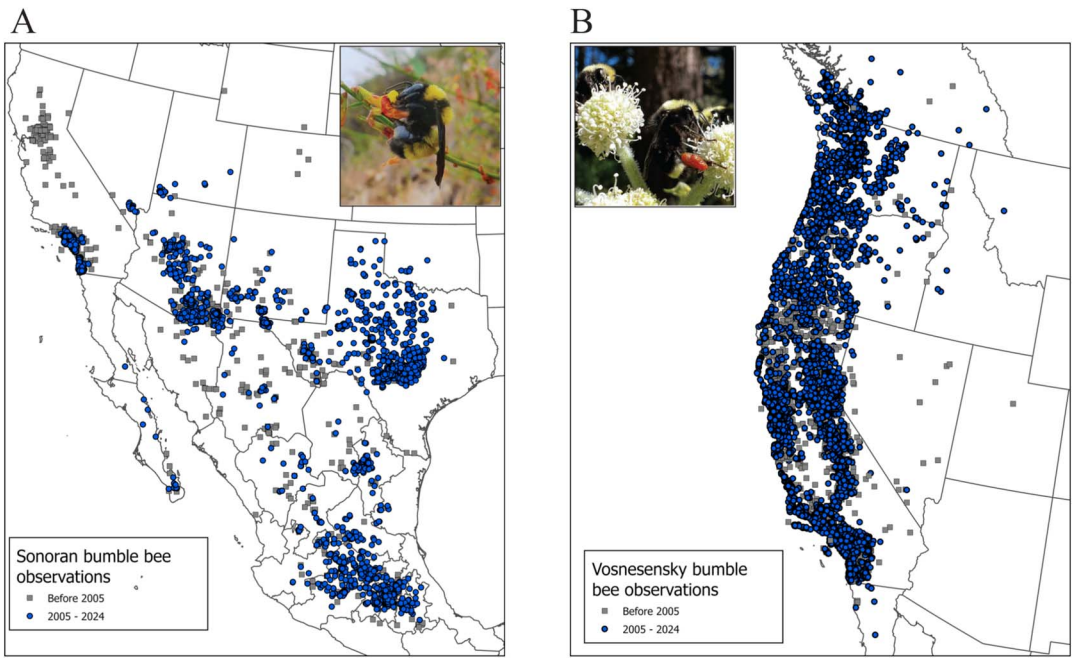

Fig. 1. Historic (< 2005, gray) and modern (≥2005, blue) distributions of two species of bumble bees, based on unpublished data from bumble bees of North America project and aggregated from specimen record and online observation databases. Data source:www.leifrichardson.org/bbna. (A) Distribution of *B. sonorus*. Inset photo © K. James Hung. (B) Distribution of *B. vosnesenskii*. Inset photo © Leif Richardson.

netting and placed directly onto dry ice. *B. vosnesenskii* 888-SC was collected at the San Diego Botanic Garden (33.085, −117 142) and *B. sonorus* 0022S was collected at the Niguel Botanical Preserve (33.531, −117.709). Specimens were stored frozen until DNA extraction.

High molecular weight DNA extraction

B. sonorus. We extracted high molecular weight (HWM) genomic DNA (gDNA) from the abdominal tissue (Specimen 021S) using the Nanobind Tissue Big DNA kit following all manufacturer's instructions (Pacific Biosciences [PacBio], Menlo Park, CA). Subsequently, the extracted HMW DNA underwent additional purifications through the phenol-chloroform method (PacBio). DNA purity was estimated using absorbance ratios on the NanoDrop ND-1000 spectrophotometer (260/280 = 1.83 and 260/230 = 4.79), and the yield (786 ng) was quantified using the Quantus Fluorometer with the Quantiflour ONE dsDNA Dye Assay (Promega, Madison, WI). Size distribution analysis of the HMW DNA was performed using the Femto Pulse System (Agilent, Santa Clara, CA), and found that 71% of the fragments were 50 kb or longer.

B. vosnesenskii. Because the initial extraction from *B. soronus* resulted in a relatively low yield, we employed an alternative method for *B. vosnesenskii*. Thoracic tissue (Specimen 887-D) was homogenized in 750 μL of homogenization buffer (10 mM Tris–HCl, pH 8.0, and 25 mM EDTA) using TissueRuptor II (Qiagen Cat # 9002755). The homogenate was incubated overnight at room temperature with an equal volume of lysis buffer (10 mM Tris, 25 mM EDTA, 200 mM NaCl, and 1% SDS) and proteinase K (100 μg/ml). The lysate was then treated with RNAse A (20 μg/ml) at 37°C for 30 min and subsequently cleaned with equal volumes of phenol/chloroform using phase-lock gels (Quantabio, Beverly, MA; Cat # 2302830). The DNA was precipitated by adding 0.4× volume of 5 M ammonium acetate and 3× volume of ice-cold ethanol. The DNA pellets were washed twice with 70% ethanol and resuspended in elution buffer (10 mM Tris, pH 8.0). The purity of the DNA was estimated using absorbance ratios (260/280 = 1.91 and 260/230 = 1.96) on the NanoDrop ND-1000 spectrophotometer. The final DNA yield (2.2 μg) was quantified using the Quantus Fluorometer (QuantiFluor ONE dsDNA Dye assay; Promega). The size distribution of the DNA was estimated using the Femto Pulse system (Agilent), and it was found that 63% of the fragments were 50 kb or longer.

PacBio HiFi library preparation and sequencing

HiFi low input library preparation. For each species, HiFi SMRTbell libraries were constructed using the SMRTbell Express Template Prep Kit v2.0 (PacBio, Cat. #100–938-900) according to the manufacturer's instructions. HMW gDNA was sheared to a target DNA size distribution between 12–20 kb using Diagenode's Megaruptor 3 system (Diagenode, Belgium; cat. B06010003). The sheared gDNA was concentrated using 1.8× of AMPure PB beads (PacBio, Cat. #100–265-900) for the removal of single-strand overhangs at 37°C for 15 min, followed by further enzymatic steps of DNA damage repair at 37°C for 30 min, end repair and A-tailing at 20°C for 10 min and 65°C for 30 min, and ligation of overhang adapters v3 at 20°C for 60 min. The SMRTbell libraries were purified and concentrated with 0.45× AMPure PB beads for size selection with 2.2× of 40% v/v diluted AMPure PB beads to remove short SMRTbell templates, < 3 kb. Each 12–20 kb average HiFi SMRTbell library was sequenced at UC Davis DNA Technologies Core (Davis, CA) using one 8 M SMRT cell, Sequel II sequencing chemistry 2.0, and 30-hour movies each on a PacBio Sequel IIe sequencer.

Omni-C library preparation and sequencing

Omni-C libraries were also prepared for each species using the Dovetail™ Omni-C™ Kit (Dovetail Genomics, Scotts Valley, CA) according to the manufacturer's protocol with slight modifications. First, each specimen tissue (whole insect, *B. vosnesenskii*—888-SC; *B. sonorus*—022S) was thoroughly ground with a mortar and pestle while cooled with liquid nitrogen. Subsequently, chromatin was fixed in place in the nucleus. The suspended chromatin solution was then passed through 100 μm and 40 μm cell strainers to remove large debris. Fixed chromatin was digested under various conditions of DNase I in situ until a suitable fragment length distribution of DNA molecules was obtained. After digestion, the cells were lysed with sodium dodecyl sulfate (SDS), and chromatin fragments were bound to chromatin capture beads. Chromatin ends were repaired and ligated to a biotinylated bridge adapter followed by proximity ligation of adapter containing ends. After proximity ligation, crosslinks were reversed and the DNA was purified from proteins. Purified DNA was treated to remove biotin that was not internal to ligated fragments. A next-generation sequencing (NGS) library was generated using an NEB Ultra II DNA Library Prep kit (New England Biolabs, Ipswich, MA) with an Illumina compatible y-adaptor. Biotin-containing fragments were then captured using streptavidin beads. The post-capture product was split into two replicates prior to PCR enrichment to preserve library complexity with each replicate receiving unique dual indices. The libraries were sequenced at Vincent J. Coates Genomics Sequencing Lab (Berkeley, CA) on an Illumina NovaSeq 6000 platform (Illumina, CA) to generate ∼100 million 2 × 150 bp read pairs per GB genome size per species.

Nuclear genome assembly

We assembled the *B. sonorus* and *B. vosnesenskii* genomes following the CCGP assembly protocol Version 5.0, as outlined in Table 1, using PacBio HiFi reads and Omni-C data to generate high-quality, highly contiguous nuclear genome assemblies. The bees used in this study were opportunistically collected according to availability, and the best haplotype was used as the final genome assembly. Bees, as with all insects in the Hymenoptera, follow a haplo-diploid chromosomal system where males are haploid and females are diploid. While for both species, we removed remnant adapter sequences from the PacBio HiFi datasets using HiFiAdapterFilt (Sim et al. 2022), the initial contig-level genome assembly varies according to the ploidy of the sampled individual per species.

For *B. sonorus*, which was a diploid female specimen, we obtained the initial diploid phased assembly using HiFiasm with the filtered PacBio HiFi reads and the Omni-C dataset (Cheng et al. 2021) in HiC mode. For *B. vosnesenskii,* a haploid male, we generated an initial haploid assembly using HiFiasm, with the filtered PacBio HiFi reads and specifying no purging and the ploidy corresponding to the sequenced haploid male. From the generated output, we kept the file corresponding to the primary assembly file and removed sequences corresponding to haplotypic duplications and contig overlaps with purge_dups (Guan et al. 2020).

Table 1. Assembly pipeline and software used.

Assembly step	Software and non-default options[§]	Version	Reference
Initial assembly			
Filtering PacBio HiFi adapters	HiFiAdapterFilt	Commit 64d1c7b	Sim et al. 2022
K-mer counting	Meryl (k = 21)	1	https://github.com/marbl/meryl
Estimation of genome size and heterozygosity	GenomeScope	2	Ranallo-Benavidez et al. 2020
De novo assembly (contig-level assembly)	*B. sonorus: HiFiasm (Hi-C Mode, –primary, output hic.hap1.p_ctg, hic.hap2.p_ctg)* *B. vosnesenskii: HiFiasm (–l0 –n-hap 1, output bp.p_ctg)*	0.16.1-r375	Cheng et al. 2022
Remove low-coverage, duplicated contigs	*B. vosnesenskii: Purge_dups (cutoffs: 5, 14,14,28,46,84)*	1.2.5	Guan et al. 2020
Scaffolding			
Omni-C data alignment	*Arima Genomics Mapping Pipeline (AGMP)*	Commit 2e74ea4	https://github.com/ArimaGenomics/mapping_pipeline
Arima Genomics Mapping Pipeline (AGMP)	BWA-MEM	0.7.17-r1188	Li 2013
	samtools	1.11	Danecek et al. 2021
	filter_five_end.pl (AGMP)	Commit 2e74ea4	https://github.com/ArimaGenomics/mapping_pipeline
	two_read_bam_combiner.pl (AGMP)	Commit 2e74ea4	https://github.com/ArimaGenomics/mapping_pipeline
	picard	2.27.5	https://broadinstitute.github.io/picard/
Omni-C Scaffolding	SALSA (-DNASE, –i 20, –p yes)	2	Ghurye et al. 2017, Ghurye et al. 2019
Omni-C contact map generation			
Short-read alignment	BWA-MEM (-5SP)	0.7.17-r1188	Li 2013
SAM/BAM processing	samtools	1.11	Danecek et al. 2021
SAM/BAM filtering	pairtools	0.3.0	Open2C et al. 2024
Pairs indexing	pairix	0.3.7	Lee et al. 2022
Matrix generation	cooler	0.8.10	Abdennur and Mirny 2020
Matrix balancing	hicExplorer (hicCorrectmatrix correct –filterThreshold –2 4)	3.6	Ramírez et al. 2018
Contact map visualization	HiGlass	2.1.11	Kerpedjiev et al. 2018
	PretextMap	0.1.4	https://github.com/wtsi-hpag/PretextView
	PretextView	0.1.5	https://github.com/wtsi-hpag/PretextMap
	PretextSnapshot	0.0.3	https://github.com/wtsi-hpag/PretextSnapshot
Manual curation tools	Rapid curation pipeline (Wellcome Trust Sanger Institute, Genome Reference Informatics Team)	Commit 7acf220c	https://gitlab.com/wtsi-grit/rapid-curation
Genome quality assessment			
Basic assembly metrics	QUAST (–-est-ref-size)	5.0.2	Gurevich et al. 2013
Assembly completeness	BUSCO (–m geno, –l hymenoptera)	5.0.0	Manni et al. 2021
	Merqury	2020 January 29	Rhie et al. 2020
Contamination screening			
Local alignment tool	BLAST+ (–db nt, –outfmt '6 qseqid staxids bitscore std', –max_target_seqs 1, –max_hsps 1, –evalue 1e-25)	2.10	Camacho et al. 2009
General contamination screening	BlobToolKit (PacBIo HiFi Coverage, BUSCODB = hymenoptera, NCBI Taxa ID = 207 650 for *B. vosnesenskii* and Taxa ID = 203 818 for *B. sonorus*)	2.3.3	Challis et al. 2020
Mitochondrial assembly			
Mitochondrial genome assembly	MitoHiFi (–r, –p 90, –o 1) Reference: *Bombus schrencki* (NCBI:LC702705.1)	2.2	Uliano-Silva et al. 2023
MitoHiFi: Mitochondrial genome annotation	MitoFinder (MitoHiFi pipeline parameters)	1.4	Allio et al. 2020
Genome assembly comparison			
Pairwise sequence alignment tool	mummer (nucmer)	4.0.0rc1	Marçais et al. 2018
Nucmer output to Dot! preparation	DotPrep.py (–-overview 10 000)	Commit b18fed0	https://github.com/marianattestad/dot
Dot plot visualization	DOT	Last accessed September 8th, 2020	https://dot.sandbox.bio/

[§]Options detailed for non-default parameters.

Then, for each species, we aligned the Omni-C data to the corresponding assembly following the Arima Genomics Mapping Pipeline (https://github.com/ArimaGenomics/mapping_pipeline) and scaffolded it with SALSA (Ghurye et al. 2017, 2019). Later, each assembly was manually curated by iteratively generating and analyzing the corresponding Omni-C contact maps. The Omni-C contact maps were generated by aligning the Omni-C data against the corresponding assembly with BWA-MEM (Li 2013), identified ligation junctions, and generated Omni-C pairs (Lee et al. 2022) using pairtools (Open2C et al. 2024). We generated a multi-resolution Omni-C matrix with cooler (Abdennur and Mirny 2020) and balanced it with hicExplorer (Ramírez et al. 2018). We used HiGlass (Kerpedjiev et al. 2018) and the PretextSuite (https://github.com/wtsi-hpag/PretextView; https://github.com/wtsi-hpag/PretextMap; https://github.com/wtsi-hpag/PretextSnapshot) to visualize the contact maps where we identified misassemblies and misjoins, and finally modified the assemblies using the Rapid Curation pipeline from the Wellcome Trust Sanger Institute, Genome Reference Informatics Team (https://gitlab.com/wtsi-grit/rapid-curation). Some of the remaining gaps were closed using the PacBio HiFi reads and YAGCloser (https://github.com/merlyescalona/yagcloser). Finally, we checked for contamination using the BlobToolKit Framework (Challis et al. 2020).

Mitochondrial genome assembly

We assembled the mitochondrial genomes from the PacBio HiFi reads using the reference-guided pipeline MitoHiFi (https://github.com/marcelauliano/MitoHiFi; Allio et al. 2020). We used the mitochondrial sequence of *Bombus schrencki* (NCBI: LC702705.1) as the starting reference. After completing the nuclear genome, we searched the nuclear genome assembly for matches to the resulting mitochondrial assembly sequence using BLAST+ (Camacho et al. 2009) and filtered out contigs and scaffolds with a sequence identity $>99\%$ and a size smaller than the mitochondrial assembly sequence.

Genome size estimation and quality assessment

We generated k-mer counts (k = 21) from the PacBio HiFi reads using meryl (https://github.com/marbl/meryl). The k-mer database was then used in GenomeScope2.0 (Ranallo-Benavidez et al. 2020) to estimate genome features, including genome size, heterozygosity, and repeat content. To obtain general contiguity metrics, we ran QUAST [Version 5.0.2] (Gurevich et al. 2013). To evaluate genome quality and completeness, we used BUSCO [Version 5.0.0] (Manni et al. 2021) with the OrthoDB v.10 (Kriventseva et al. 2019) Hymenopteran dataset ($n = 5991$ genes). Assessment of base level accuracy (QV) and k-mer completeness was performed using the previously generated meryl database and merqury (Rhie et al. 2020). We further estimated genome assembly accuracy using the BUSCO gene-set frameshift analysis pipeline described in Korlach et al. (2017). Measurements of the size of the phased blocks are based on the size of the contigs generated by HiFiasm in HiC mode. We follow the quality metrics nomenclature established by Rhie et al. (2021), we will use the derived genome quality notation x.y.P.Q.C, where, x = log10[contig NG50]; y = log10[scaffold NG50]; P = log10 [phased block NG50]; Q = Phred base accuracy QV (quality value); C = % genome represented by the first 'n' scaffolds, following a karyotype of $n = 18$ inferred from published karyotypes of *B. vosnesenskii* and *B. pensylvanicus* (Owen et al. 1995). Quality metrics for the notation were calculated for both species on the primary assembly.

Comparisons to closely related Bombus genome assemblies

We compared the genomes in this manuscript with other published closely related *Bombus* genomes in three different ways: first, we compared *B. vosnesenskii* genome to another published *B. vosnesenskii* genome (NCBI:GCA_011952255.1); second, and we compared the *B. sonorus* genome with the genome of *B. pensylvanicus*, the most closely related genome available on NCBI (NCBI:GCA_051853225; Lozier et al. 2025); and finally, we contextualize the assembled genomes in with the rest of the available Bombus genomes available on NCBI.

For the *B. vosnesenskii* comparison, we aligned the *B. vosnesenskii* genome generated to another published *B. vosnesenskii* genome (scaffold-level, NCBI:GCA_011952255.1) with the nucmer module from mummer (Marçais et al. 2018) and used Dot (https://dot.sandbox.bio/) to visualize forward, reverse, and repetitive alignments across all genome comparisons.

For the *B. sonorus* comparison, we ran a synteny analysis of the *B. sonorus* genome in our study, alongside the genome assembly for *B. pensylvanicus* (subgenus Thoracobombus; chromosome-level, NCBIGCA_051853225;). We estimated sequence divergence among genomes using mash v2.3 (Ondov et al. 2016, 2019). Then, we used the estimated sequence divergence to analyze genome synteny across all species with ntSynt v1.0.2 (−d 5,-b 100 000; Coombe et al. 2024), and visualized the output using ntSynt_viz v1.0.0 (Coombe, Warren, and Birol 2025).

To compare contiguity with existing bumble bee genomes, we identified *Bombus* genomes on NCBI and selected those that were representative of the species. When multiple assemblies were available for a species and did not represent different haplotypes for a single individual, we selected the one with better resolution. If the assemblies corresponded to different haplotypes of an individual, we selected the principal or primary haplotype assemblies. Then, we downloaded the genomes using the datasets (O'Leary et al., 2024), obtained contiguity metrics with QUAST, and functional completeness metrics with BUSCO (using the hymenoptera gene set).

Results

All genome sizes were similar to the estimated value from GenomeScope 2.0. NCBI BioProject numbers and assembly identifiers are listed in Supplementary Table 1. Detailed assembly statistics for all genomes are reported in Table 2, and graphical representations of the primary assemblies are shown in Figs. 2 and 3 (see Supplementary Fig. 1 for visualization of the alternate assembly of *B. sonorus*).

Bombus sonorus

The final assembly (iyBomSono1) consists of 2 phased haplotypes, and the assemblies are considered primary and alternate haplotypes throughout the manuscript based on the overall completeness of the assemblies. The primary haplotype assembly (iyBomSono1.0.hap1) consists of 282 scaffolds spanning 378.8 Mb with a contig N50 of 10.5 Mb, scaffold N50

Table 2. Assembly statistics.

		Bombus sonorus		Bombus vosnesenskii
Assembly identifier (Quality code[a])		iyBomSono1. (7.7.P7.Q65.C79)		iyBomVosn1(6.7.P7.Q65.C58)
HiFi Read coverage[§]		93.59×		37.82×
		Primary	**Alternate**	**Primary**
Number of contigs		314	222	444
Contig N50 (bp)		10 531 476	9 658 121	8 871 735
Contig NG50[§]		14 268 688	10 894 028	12 190 813
Longest Contigs		19 563 964	15 925 711	20 255 864
Number of scaffolds		282	190	433
Scaffold N50		16 265 523	14 188 697	12 536 986
Scaffold NG50[§]		19 108 446	15 148 397	13 818 281
Largest scaffold		25 366 763	25 396 551	24 536 892
Size of final assembly		376 851 466	328 596 561	478 844 667
Phased block NG50[§]		14 747 493	10 894 028	12 190 813
Gaps per Gbp (# Gaps)		85(32)	97(32)	23(11)
Indel QV (Frame shift)		47.82759193	48.31887061	47.92219806
Base pair QV		65.4264	65.4264	65.6608
		Full assembly = 65.1247		
k-mer completeness		98.3785	97.59	99.3283
		Full assembly = 99.5626		
BUSCO completeness[b]	**C**	97.60%	97.20%	97.50%
(hymenoptera gene set) $n = 5991$	**S**	97.30%	96.80%	97.10%
	D	0.30%	0.40%	0.40%
	F	0.50%	0.50%	0.50%
	M	1.90%	2.30%	2.00%
Organelles		# Partial mitochondrial sequence: JARWAU010000282.1		# Partial mitochondrial sequence: JAOPFL010000433.1

[§]Read coverage and NGx statistics have been calculated based on the estimated genome size of 286 Mb for *B. sonorus* and 399.2 Mb for *B. vosnesenskii*. [a]Assembly quality code x.y.P.Q.C derived notation, from (Rhie et al. 2021). x = log10[contig NG50]; y = log10[scaffold NG50]; P = log10 [phased block NG50]; Q = Phred base accuracy QV (Quality value); C = % genome represented by the first 'n' scaffolds, following a known karyotype for *B. sonorus* and *B. vosnesenskii* of $n = 18$ (Owen et al. 1995). Bp: base pairs. Quality code for assemblies denoted by primary assembly. [b]BUSCO scores: (C)omplete and (S)ingle; (C)omplete and (D)uplicated; (F)ragmented and (M)issing BUSCO genes. n, number of BUSCO genes in the set/database.

of 16.2 Mb, largest contig of 19.5 Mb, and largest scaffold of 25.3 Mb. The alternate haplotype assembly (iyBomSono1.0.hap2) consists of 190 scaffolds, spanning 328.5 Mb with contig N50 of 9.6 Mb, scaffold N50 of 14.8 Mb, largest contig 15.9 Mb, and largest scaffold of 23.4 Mb. The primary assembly has a BUSCO completeness score of 97.6% using the hymenoptera gene set, a per-base quality (QV) of 65.4, a k-mer completeness of 98.37, and a frameshift indel QV of 47.82; while the alternate assembly has a BUSCO completeness score of 97.2% using the same gene set, a per-base quality (QV) of 65.4, a k-mer completeness of 97.59, and a frameshift indel QV of 48.3. We manually generated 61 joins (28 on the primary assembly and 33 on the alternate) and 4 breaks (1 on the primary assembly and 3 on the alternate assembly). We were able to close a total of 9 gaps, 5 on the primary assembly and 4 on the alternate. Finally, we filtered out 27 contigs during contamination screening with Blobtools (detailed taxonomic assignments of removed contigs in Supplementary Table 2) and 4 contigs corresponding to mitochondrial matches (2 per assembly). No further contigs were removed.

The Omni-C contact maps show that both primary and alternate assemblies are highly contiguous, suggesting that the *B. sonorus* genome is organized in 19 chromosomes based on the number of major bins along the diagonal axis of the plot (Fig. 2B and C). We observe some rearrangements compared to *B. pensylvanicus*, which may require further analysis to confirm (Supplementary Fig. 2). We named the chromosome-scaffolds based on their ordered length (Supplementary Table 3). In addition, we observed that 19 of the *B. sonorus* chromosome scaffolds had forward alignments with at least one chromosome in the *B. pensylvanicus* genome.

We assembled a partial mitochondrial genome with MitoHiFi. The final mitochondrial genome size was 16 406 bp. The base composition of the final assembly version is A = 44.08%, C = 5.388%, G = 7.357%, T = 43.17%, and consists of 20 unique tRNAs, 1 rRNA, and 12 protein coding genes.

Bombus vosnesenskii

The final assembly comprises 433 scaffolds spanning 378.8 Mb, with a contig N50 of 8.8 Mb, scaffold N50 of 12.5 Mb, largest contig of 20.2 Mb, and largest scaffold of 24.5 Mb. The final BUSCO completeness score is 97.5%, a QV of 65.6, a k-mer completeness score of 99.3, and a frameshift indel QV of 47.9. We manually joined 9 scaffolds and closed a single gap on the assembly. Finally, we filtered out three contigs from the assembly, two corresponding to mitochondrial contamination and one corresponding to *Apicystis bombi*. No further contigs were removed. The Omni-C contact map shows that the assembly is highly contiguous, containing at least 18 chromosome-scaffolds (Fig. 2D and F). We named the chromosome-scaffolds based on their ordered length (Supplementary Table 4). We observed that 17 of the major *B. vosnesenskii* scaffolds had forward alignments with at least one scaffold in another published *B. vosnesenskii* genome, also at the scaffold level (Supplementary Fig. 3).

The partial mitochondrial genome size was 17 353 bp. The base composition of the final assembly version is A = 41.69%, C = 3.043%, G = 9.883%, T = 45.38%, and consists of 15 unique tRNAs, 1 rRNA, and 8 protein coding genes.

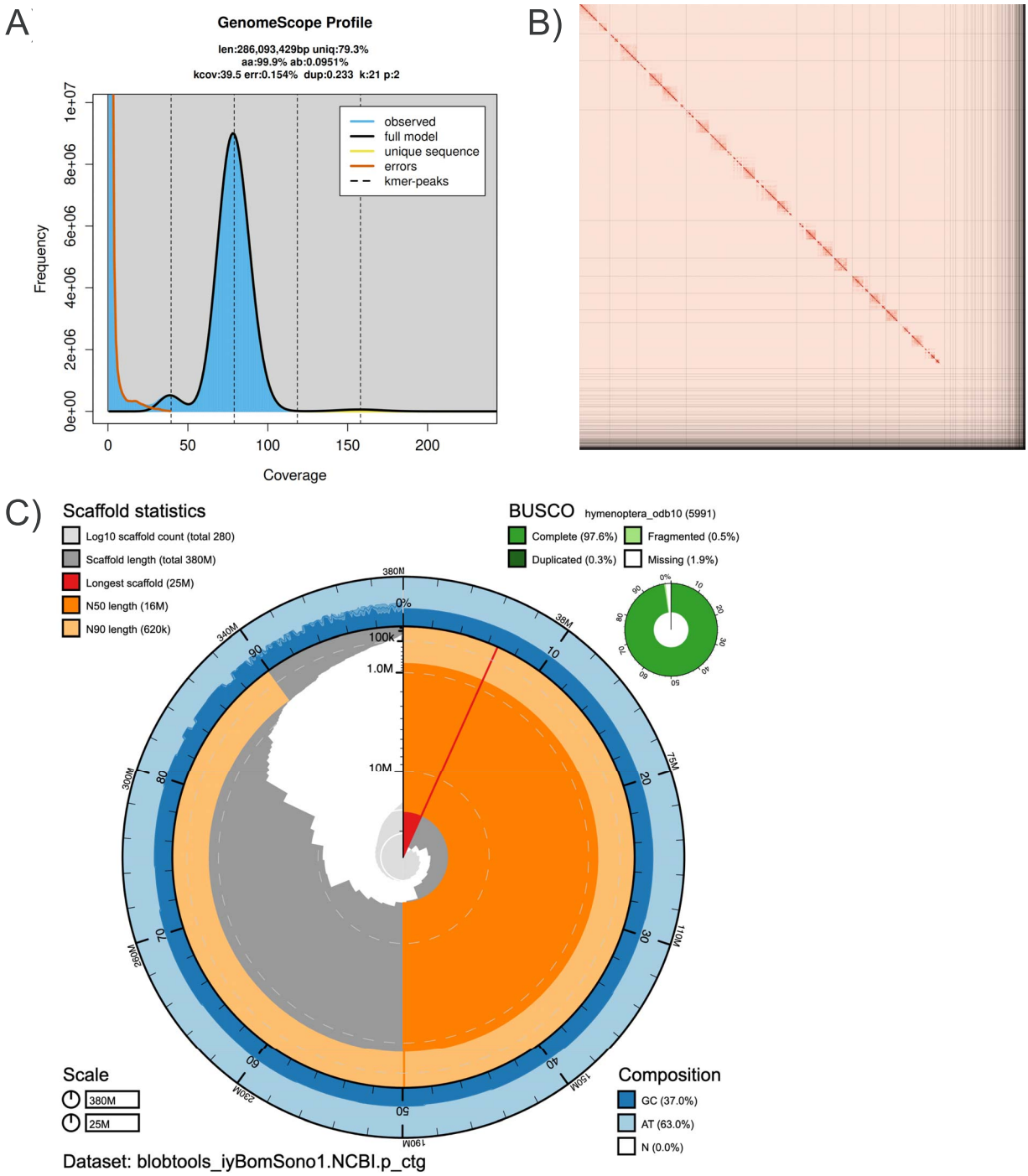

Fig. 2. Visual overview of *Bombus sonorus* genome assembly metrics. (A) K-mer spectrum output generated from PacBio HiFi data without adapters using GenomeScope 2.0. The bimodal pattern observed corresponds to a diploid genome. (B) Omni-C contact map for the *Bombus sonorus* primary genome assembly generated with PretextSnapshot. Omni-C contact maps translate proximity of genomic regions in 3D space to contiguous linear organization. Each cell in the contact map corresponds to sequencing data supporting the linkage (or join) between 2 of such regions. Scaffolds are separated by dark lines and higher density corresponds to higher levels of fragmentation. (C) BlobToolKit snail plots showing a graphical representation of the quality metrics presented in Table 2 for the primary assembly of *B. sonorus* (iyBomSono1.0.hap1.). The plot circle represents the full size of the assembly. From the inside-out, the central plot covers length-related metrics. The red line represents the size of the longest scaffold; all other scaffolds are arranged in size order moving clockwise around the plot and drawn in gray starting from the outside of the central plot. Dark and light orange arcs show the scaffold N50 and scaffold N90 values. The central light gray spiral shows the cumulative scaffold count with a white line at each order of magnitude. White regions in this area reflect the proportion of ns in the assembly the dark versus light blue area around it shows mean, maximum and minimum GC versus AT content at 0.1% intervals (Challis et al. 2020).

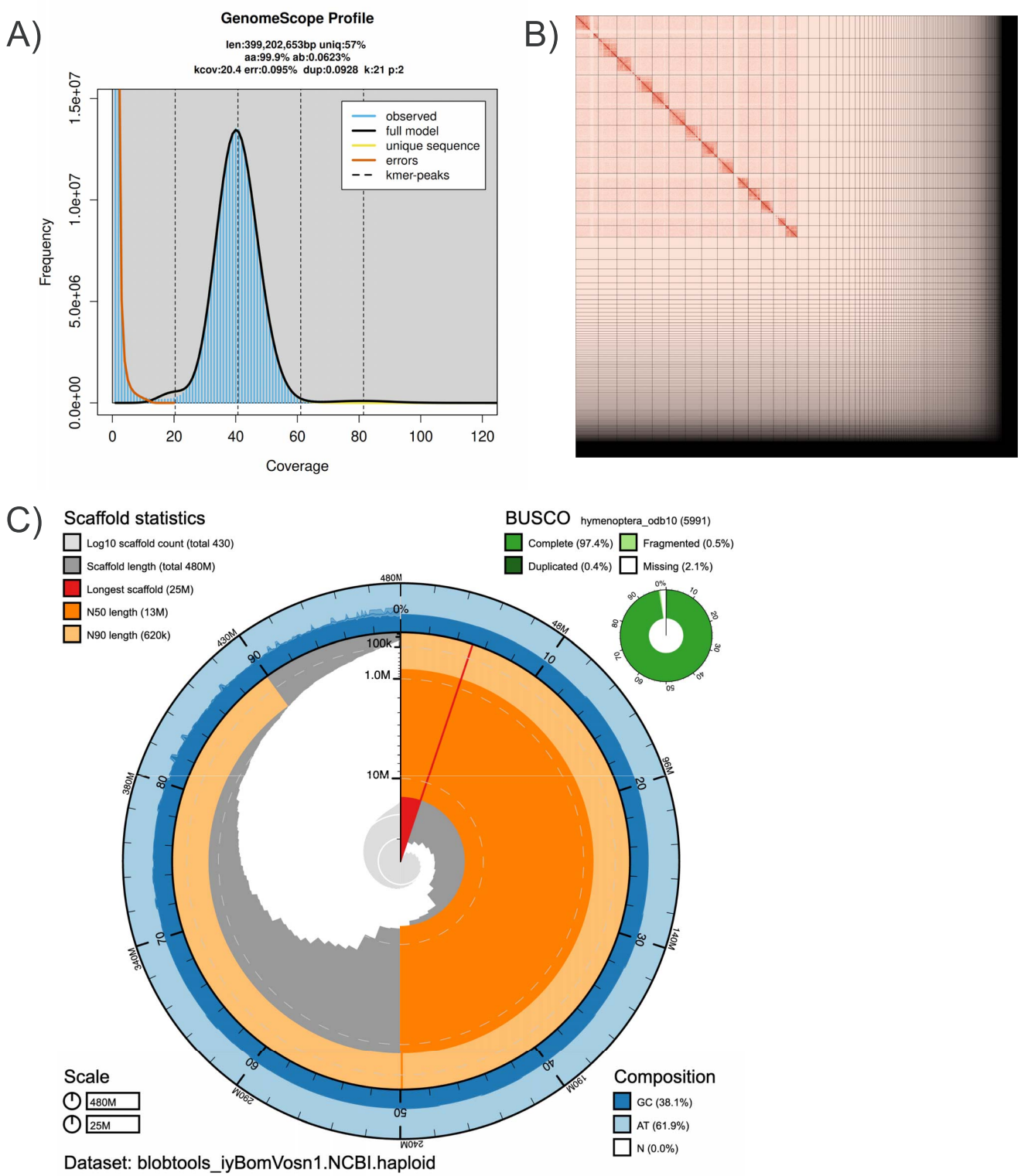

Fig. 3. Visual overview of *Bombus vosnesenskii* genome assembly metrics. (A) K-mer spectrum output generated from PacBio HiFi data without adapters using GenomeScope 2.0. The monomodal pattern observed corresponds to a haploid genome. (B) Omni-C contact map for the *B. vosnesenkii* genome assembly generated with PretextSnapshot. Omni-C contact maps translate proximity of genomic regions in 3D space to contiguous linear organization. Each cell in the contact map corresponds to sequencing data supporting the linkage (or join) between 2 of such regions. Scaffolds are separated by dark lines and higher density corresponds to higher levels of fragmentation. (C) BlobToolKit snail plots showing a graphical representation of the quality metrics presented in Table 2 for the sole assembly for *B. vosnesnskii* (iyBomVosn1.0). The plot circle represents the full size of the assembly. From the inside-out, the central plot covers length-related metrics. The red line represents the size of the longest scaffold; all other scaffolds are arranged in size order moving clockwise around the plot and drawn in gray starting from the outside of the central plot. Dark and light orange arcs show the scaffold N50 and scaffold N90 values. The central light gray spiral shows the cumulative scaffold count with a white line at each order of magnitude. White regions in this area reflect the proportion of ns in the assembly the dark versus light blue area around it shows mean, maximum and minimum GC versus AT content at 0.1% intervals (Challis et al. 2020).

Discussion

Here we present complete draft genome assemblies for the Sonoran and yellow-faced bumble bees, assembled using long reads and chromosome-scale sequencing data. These assemblies are among the most contiguous yet published for *Bombus*, showing similar or greater contiguity to chromosome-level assemblies on NCBI (Supplementary Table 5). There are currently over 50 *Bombus* species with genomes available on NCBI, 13 of which are North American species. The Sonoran bumble bee genome is the first genome available for this species and contributes crucial genomic information within the *B. pensylvanicus* s.l. complex of bumble bee species—a notoriously difficult group to distinguish taxonomically (Williams et al. 2024). Our preliminary comparison with the recently published *B. pensylvanicus* genome (Lozier et al. 2025) suggests that both species contain 19 haploid chromosomes. Although within the observed karyotype range, 19 chromosomes is somewhat unusual among *Bombus* species, most of which have a karyotype number of 18 (Owen et al. 1995; Sun et al. 2021). We also observed some potential mismatches and rearrangements between the two genomes (Supplementary Fig. 2). Further analysis, outside of the scope of this announcement, would be useful in confirming genomic similarities and discordance between the two species. Our yellow-faced bumble bee genome is nearly four times as contiguous as the other scaffold-level *B. vosnesenkii* genome available (Heraghty et al. 2020) and is one of the most contiguous genomes within the subgenus Pyrobombus.

Many North American bumble bee species are threatened, with recent analyses indicating up to one quarter of all North American species are at risk of extinction (Cameron and Sadd 2020). High-resolution genomic data will enable critical research informing species-specific conservation efforts. This includes assessing contemporary and historic patterns of gene flow and genetic structure across their respective ranges and exploring adaptive capacity and signatures of local-scale adaptation across the ranges of these species. For Sonoran bumble bees, high-quality genomic resources are also needed to determine taxonomic status with its close relative and potential conspecific *B. pensylvanicus*. Given the documented declines of *B. pensylvanicus* across the U.S. and Canada (Hatfield et al. 2015; Carrasco et al. 2021), which often include *B. sonorus* in analyses, a more fully resolved species status will enable accurate species assessments and conservation management plans at both state and federal levels. In contrast, *B. vosnesenskii* is considered widespread and stable, and even potentially expanding in parts of its range across Western North America. In part due to its abundance and ease of laboratory rearing in comparison to other *Bombus* species, *B. vosnesenkii* has been used in numerous studies as a model system for behavioral (Heinrich 1972; Malfi et al. 2019) and molecular ecology research (Lozier et al. 2011; Jha 2015; Schenau and Jha 2016; Jackson et al. 2022). Full genomic resources for this species, however, have previously been limited (but see Heraghty et al. 2020), and additional genomic information will aid important ongoing basic research of this species as well as management efforts for wild populations.

Worldwide, bumble bee declines are driven by habitat loss, climate change, pathogens, and pesticides, among other factors (Potts et al. 2010; Cameron et al. 2011; Jha and Kremen 2013; Hatfield et al. 2014, 2021; Goulson et al. 2015; Maxfield-Taylor et al. 2015; Soroye et al. 2020; IUCN Bumblebee Specialist Group 2023). Increased genomic information for diverse *Bombus* species, such as we present here, provides foundational information about the biology of this imperiled insect group. Moreover, these resources support conservation management, as well as enable comparative research to pinpoint genomic drivers of species sensitivity to individual and interacting stressors.

Acknowledgments

PacBio Sequel II library prep and sequencing were carried out at the DNA Technologies and Expression Analysis Cores at the UC Davis Genome Center, supported by National Institute of Health (NIH) Shared Instrumentation Grant 1S10OD010786-01. Deep sequencing of Omni-C libraries used the Novaseq S4 sequencing platforms at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. We thank the staff at the UC Davis DNA Technologies and Expression Analysis Cores and the UC Santa Cruz Paleogenomics Laboratory for their diligence and dedication to generating high-quality sequence data. Thanks to Dr Oscar Martinéz-López for providing additional feedback on the manuscript.

Author contributions

Kimberly M. Ballare (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Validation, Visualization, Writing—original draft, Writing—review & editing), Merly Escalona (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing—original draft, Writing—review & editing), Kaleigh Fisher (Data curation, Methodology, Writing—original draft, Writing—review & editing), Blanca M. Guillén (Data curation, Investigation, Methodology, Writing—review & editing), Leif Richardson (Data curation, Formal analysis, Visualization, Writing—review & editing), William Seligmann (Data curation, Investigation, Methodology, Writing—original draft, Writing—review & editing), Noravit Chumchim (Data curation, Investigation, Methodology, Validation, Writing—original draft, Writing—review & editing), Mohan P.A. Marimuthu (Data curation, Investigation, Methodology, Validation, Writing—review & editing), Oanh Nguyen (Data curation, Investigation, Methodology, Validation, Writing—review & editing), Neal M. Williams (Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing—review & editing), Beth Shapiro (Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing—review & editing), and S. Hollis Woodard (Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing—original draft, Writing—review & editing)

Funding

This work was supported by the California Conservation Genomics Project, with funding provided to the University of California by the State of California, State Budget Act of 2019 [UC Award ID RSI-19-690224]. Additional funding support was provided by The Howard Hughes Medical Institute Professors Program grant no. GT10483.

Data availability

Data generated for this study are available under NCBI BioProject PRJNA720569. Raw sequencing data for all samples (NCBI BioSamples SAMN33567128, SAMN33567129, SAMN30567649, SAMN30567651) are deposited in the NCBI Short Read Archive (SRA)

under SRX21226991 and SRR25468310 for PacBio HiFi sequencing data, and SRX21226992–3, SRR25468308, and SRR25468309 for the Omni-C Illumina sequencing data. GenBank accessions for both primary and alternate assemblies of *B. sonorus* are JARWAU000000000 and JARWAV000000000; and for genome sequences GCA_029958995.1 and GCA_029959005.1. Accessions for the *B. vosnesenskii* assembly and sequence are JAOPFL000000000 and GCA_026744215.1. The GenBank organelle genome assembly sequences for the partial mitochondrial genomes are JARWAU010000282.1 *(B. sonorus*) and JAOPFL010000433.1 (*B. vosnesenskii*). Assembly scripts and other data for the analyses presented can be found at the following GitHub repository: www.github.com/ccgproject/ccgp_assembly

References

Abdennur N, Mirny LA. Cooler: scalable storage for hi-C data and other genomically labeled arrays (J wren, ed.). *Bioinformatics*. 2020;**36**:311–316. https://doi.org/10.1093/bioinformatics/btz540

Allio R, Schomaker-Bastos A, Romiguier J, Prosdocimi F, Nabholz B, Delsuc F. MitoFinder: efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Mol Ecol Resour*. 2020;**20**:892–905. https://doi.org/10.1111/1755-0998.13160

Arbetman MP, Gleiser G, Morales CL, Williams P, Aizen MA. Global decline of bumblebees is phylogenetically structured and inversely related to species range size and pathogen incidence. *Proc R Soc B Biol Sci*. 2017;**284**:20170204. https://doi.org/10.1098/rspb.2017.0204

Beckham JL, Johnson JA, Pfau RS. Molecular data support *Bombus sonorus* and *Bombus pensylvanicus* (Hymenoptera, Apidae) as distinct species. *J Hymenopt Res*. 2024;**97**:895–914. https://doi.org/10.3897/jhr.97.132937

California bumble bee atlas. <cabumblebeeatlas.org> Accessed 15 Jan. 2025.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;**10**:421. https://doi.org/10.1186/1471-2105-10-421

Cameron SA, Sadd BM. Global trends in bumble bee health. *Annu Rev Entomol*. 2020;**65**:209–232. https://doi.org/10.1146/annurev-ento-011118-111847

Cameron SA, Lozier JD, Strange JP, Koch JB, Cordes N, Solter LF, Griswold TL. Patterns of widespread decline in north American bumble bees. *Proc Natl Acad Sci USA*. 2011;**108**:662–667. https://doi.org/10.1073/pnas.1014743108

Carrasco L, Papeş M, Lochner EN, Ruiz BC, Williams AG, Wiggins GJ. Potential regional declines in species richness of tomato pollinators in North America under climate change. *Ecol Appl*. 2021;**31**:e02259. https://doi.org/10.1002/eap.2259

Challis R, Richards E, Rajan J, Cochrane G, Blaxter M. BlobToolKit – interactive quality assessment of genome assemblies. *G3 Genes|Genomes|Genetics*. 2020;**10**:1361–1374. https://doi.org/10.1534/g3.119.400908

Cheng H, Jarvis ED, Fedrigo O, Koepfli K-P, Urban L, Gemmell NJ, Li H. Haplotype-resolved assembly of diploid genomes without parental data. *Nature Biotechnology*. 2022;**40**:1332–1335. https://doi.org/10.1038/s41587-022-01261-x

Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods*. 2021;**18**:170–175. https://doi.org/10.1038/s41592-020-01056-5

Colla SR, Packer L. Evidence for decline in eastern north American bumblebees (Hymenoptera: Apidae), with special focus on Bombus affinis Cresson. *Biodivers Conserv*. 2008;**17**:1379–1391. https://doi.org/10.1007/s10531-008-9340-5

Coombe L, Kazemi P, Wong J, Birol I, Warren RL. Multi-genome synteny detection using minimizer graph mappings." (p. 2024.02.07.579356 bioRxiv. 2024. https://doi.org/10.1101/2024.02.07.579356

Coombe L, Warren RL, Birol I. ntSynt-viz: Visualizing synteny patterns across multiple genomes (p. 2025.01.15.633221). *bioRxiv*. 2025. https://doi.org/10.1101/2025.01.15.633221

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. Twelve years of SAMtools and BCFtools. *GigaScience*. 2021;**10**:giab008. https://doi.org/10.1093/gigascience/giab008

Fisher K, Watrous KM, Williams NM, Richardson LL, Woodard SH. A contemporary survey of bumble bee diversity across the state of California. *Ecol Evolution*. 2022;**12**:e8505. https://doi.org/10.1002/ece3.8505

Fraser DF, Copley CR, Elle E, Cannings RA. Changes in the status and distribution of the yellow-faced bumble bee. *J Entomol Soc British Columbia*. 2012;**109**:31–37.

Garibaldi LA, Steffan-Dewenter I, Winfree R, Aizen MA, Bommarco R, Cunningham SA, Kremen C, Carvalheiro LG, Harder LD, Afik O, et al. Wild pollinators enhance fruit set of crops regardless of honey bee abundance. *Science*. 2013;**339**:1608–1611. https://doi.org/10.1126/science.1230200

Ghurye J, Pop M, Koren S, Bickhart D, Chin C-S. Scaffolding of long read assemblies using long range contact information. *BMC Genomics*. 2017;**18**:527. https://doi.org/10.1186/s12864-017-3879-z

Ghurye J, Rhie A, Walenz BP, Schmitt A, Selvaraj S, Pop M, Phillippy AM, Koren S. Integrating hi-C links with assembly graphs for chromosome-scale assembly. *PLoS Comput Biol*. 2019;**15**:e1007273. https://doi.org/10.1371/journal.pcbi.1007273

Goulson D, Hanley ME, Darvill B, Ellis JS, Knight ME. Causes of rarity in bumblebees. *Biol Conserv*. 2005;**122**:1–8. https://doi.org/10.1016/j.biocon.2004.06.017

Goulson D, Lye GC, Darvill B. Decline and conservation of bumble bees. *Annu Rev Entomol*. 2008;**53**:191–208. https://doi.org/10.1146/annurev.ento.53.103106.093454

Goulson D, Nicholls E, Botias C, Rotheray EL. Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. *Science*. 2015;**347**:1255957–1255957. https://doi.org/10.1126/science.1255957

Guan D, McCarthy SA, Wood J, Howe K, Wang Y, Durbin R. Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*. 2020;**36**:2896–2898. https://doi.org/10.1093/bioinformatics/btaa025

Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013;**29**:1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hatfield R, Colla S, Jepsen S, Richardson L, Thorp R, Foltz S. IUCN assessments for north American Bombus spp. for the North American IUCN Bumble Bee Specialist Group; 2014.

Hatfield R, Jepsen S, Thorp R, Richardson L, Colla SR, Foltz JS. *Bombus pensylvanicus*, American Bumblebee. *The IUCN Red List of Threatened Species*. 2015;**e.T2121517**:1–15.

Hatfield RG, Strange JP, Koch JB, Jepsen S, Stapleton I. Neonicotinoid pesticides cause mass fatalities of native bumble bees: a case study from Wilsonville, Oregon, United States. *Environ Entomol*. 2021;**50**:1095–1104. https://doi.org/10.1093/ee/nvab059

Heinrich B. Physiology of brood incubation in the Bumblebee queen, Bombus *vosnesenskii*. *Nature*. 1972;**239**:223–225. https://doi.org/10.1038/239223a0

Heraghty SD, Sutton JM, Pimsler ML, Fierst JL, Strange JP, Lozier JD. De Novo Genome Assemblies for Three North American Bumble Bee Species: *Bombus bifarius*, Bombus vancouverensis, and *Bombus vosnesenskii*G3 (Bethesda). 2020;**10**:2585–2592.

IUCN Bumblebee Specialist Group. BumbleBee Sub Group Annual Report 2023 of the IUCN Wild Bee Specialist Group. *IUCN*; 2023

Jackson HM, Johnson SA, Morandin LA, Richardson LL, Guzman LM, M'Gonigle LK. Climate change winners and losers among north American bumblebees. *Biol Lett*. 2022;**18**:20210551. https://doi.org/10.1098/rsbl.2021.0551

Jha S. Contemporary human-altered landscapes and oceanic barriers reduce bumble bee gene flow. *Mol Ecol*. 2015;**24**:993–1006. https://doi.org/10.1111/mec.13090

Jha S, Kremen C. Urban land use limits regional bumble bee gene flow. *Mol Ecol*. 2013;**22**:2483–2495. https://doi.org/10.1111/mec.12275

Jha S, Stefanovich L, Kremen C. Bumble bee pollen use and preference across spatial scales in human-altered landscapes. *Ecol Entomol*. 2013;**38**:570–579. https://doi.org/10.1111/een.12056

Kerpedjiev P, Abdennur N, Lekschas F, McCallum C, Dinkla K, Strobelt H, Luber JM, Ouellette SB, Azhir A, Kumar N, et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol*. 2018;**19**:125. https://doi.org/10.1186/s13059-018-1486-1

Korlach J, Gedman G, Kingan SB, Chin C-S, Howard JT, Audet J-N, Cantin L, Jarvis ED. De novo PacBio long-read and phased avian genome assemblies correct and add to reference genes generated with intermediate and short reads. *GigaScience*. 2017;**6**:1–16. https://doi.org/10.1093/gigascience/gix085

Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Res*. 2019;**47**:D807–D811. https://doi.org/10.1093/nar/gky1053

Lee S, Bakker CR, Vitzthum C, Alver BH, Park PJ. Pairs and Pairix: a file format and a tool for efficient storage and retrieval for hi-C read pairs. *Bioinformatics*. 2022;**38**:1729–1731. https://doi.org/10.1093/bioinformatics/btab870

Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013;arXiv:1303.3997.

Lozier JD, Strange JP, Stewart IJ, Cameron SA. Patterns of range-wide genetic variation in six north American bumble bee (Apidae: *Bombus*) species. *Mol Ecol*. 2011;**20**:4870–4888. https://doi.org/10.1111/j.1365-294X.2011.05314.x

Lozier JD, Schweizer RM, Sim SB, Koch JBU, Branstetter MG, Benavides LR, Geib SM, Evans JD. New reference genome assembly for the declining American bumble bee, Bombus pensylvanicus, *G3 genes|genomes|genetics,* jkaf181. 2025;**15**. https://doi.org/10.1093/g3journal/jkaf181

MacPhail VJ, Richardson LL, Colla SR. Incorporating citizen science, museum specimens, and field work into the assessment of extinction risk of the American bumble bee (*Bombus pensylvanicus* De Geer 1773) in Canada. *J Insect Conserv*. 2019;**23**:597–611. https://doi.org/10.1007/s10841-019-00152-y

Malfi RL, Crone E, Williams N. Demographic benefits of early season resources for bumble bee (*B. Vosnesenskii*) colonies. *Oecologia*. 2019;**191**:377–388. https://doi.org/10.1007/s00442-019-04472-3

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 2021;**38**:4647–4654. https://doi.org/10.1093/molbev/msab199

Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. MUMmer4: a fast and versatile genome alignment system. *PLoS Comput Biol*. 2018;**14**:e1005944. https://doi.org/10.1371/journal.pcbi.1005944

Maxfield-Taylor SA, Mujic AB, Rao S. First detection of the larval chalkbrood disease pathogen Ascosphaera apis (Ascomycota: Eurotiomycetes: Ascosphaerales) in adult bumble bees. *PLoS One*. 2015;**10**:e0124868. https://doi.org/10.1371/journal.pone.0124868

Mola JM, Pearse IS, Boone ML, Evans E, Hepner MJ, Jean RP, Kochanski JM, Nordmeyer C, Runquist E, Smith TA, et al. Range-wide genetic analysis of an endangered bumble bee (*Bombus affinis*, Hymenoptera: Apidae) reveals population structure, isolation by distance, and low colony abundance. *J Insect Sci*. 2024;**24**:19. https://doi.org/10.1093/jisesa/ieae041

O'Leary NA, Cox E, Holmes JB, Anderson WR, Falk R, Hem V, Tsuchiya MTN, Schuler GD, Zhang X, Torcivia J, et al. Exploring and retrieving sequence and metadata for species across the tree of life with NCBI datasets. *Sci Data*. 2024;**11**:732. https://doi.org/10.1038/s41597-024-03571-y

Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, & Phillippy AM. Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biology*. 2016;**17**:132. https://doi.org/10.1186/s13059-016-0997-x

Ondov BD, Starrett GJ, Sappington A, Kostic A, Koren S, Buck CB, & Phillippy AM. Mash Screen: High-throughput sequence containment estimation for genome discovery. *Genome Biology*. 2019;**20**:232. https://doi.org/10.1186/s13059-019-1841-x

Open2C, Abdennur N, Fudenberg G, Flyamer IM, Galitsyna AA, Goloborodko A, et al. Pairtools: from sequencing data to chromosome contacts. *PLoS Comput Biol*. 2024;**20**:e1012164.

Owen RE, Richards KW, Wilkes A. Chromosome numbers and karyotypic variation in bumble bees (Hymenoptera: Apidae; Bombini). *J Kansas Entomol Soc*. 1995;**68**:290–302.

Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, Kunin WE. Global pollinator declines: trends, impacts and drivers. *Trends Ecol Evol*. 2010;**25**:345–353. https://doi.org/10.1016/j.tree.2010.01.007

Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, Habermann B, Akhtar A, Manke T. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat Commun*. 2018;**9**:189. https://doi.org/10.1038/s41467-017-02525-w

Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nature Communications*. 2020;**11**:1. https://doi.org/10.1038/s41467-020-14998-3

Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol*. 2020;**21**:245. https://doi.org/10.1186/s13059-020-02134-9

Rhie A, McCarthy SA, Fedrigo O, Damas J, Formenti G, Koren S, Uliano-Silva M, Chow W, Fungtammasan A, Kim J, et al. Towards complete and error-free genome assemblies of all vertebrate species. *Nature*. 2021;**592**:737–746. https://doi.org/10.1038/s41586-021-03451-0

Schenau E, Jha S. High levels of male diploidy but low levels of genetic structure characterize *Bombus vosnesenskii* populations across the western US. *Conserv Genet*. 2016;**18**:597–605.

Shaffer HB, Toffelmier E, Corbett-Detig RB, Escalona M, Erickson B, Fiedler PGold M, Harrigan RJ, Hodges S, Luckau TK, et al. Landscape genomics to enable conservation actions: the California conservation genomics project. *J Hered*. 2022;**113**:577–588. https://doi.org/10.1093/jhered/esac020

Sim SB, Corpuz RL, Simmonds TJ, Geib SM. HiFiAdapterFilt, a memory efficient read processing pipeline, prevents occurrence of adapter sequence in PacBio HiFi reads and their negative impacts on genome assembly. *BMC Genomics*. 2022;**23**:157. https://doi.org/10.1186/s12864-022-08375-1

Soroye P, Newbold T, Kerr J. Climate change contributes to widespread declines among bumble bees across continents. *Science*. 2020;**367**:685–688. https://doi.org/10.1126/science.aax8591

Strange JP. Bombus huntii, Bombus impatiens, and Bombus vosnesenskii (Hymenoptera: Apidae) pollinate greenhouse-grown tomatoes in western North America. *J Econ Entomol*. 2015;**108**:873–879. https://doi.org/10.1093/jee/tov078

Sun C, Jiaxing H, Yun W, et al. Genus-wide characterization of Bumblebee genomes provides insights into their evolution and variation in ecological and Behavioral traits. *Mol Biol Evol*. 2021;**38**:486–501. https://doi.org/10.1093/molbev/msaa240

Thorp RW, Horning DS, Dunning LL. *Bumble bees and cuckoo bumble bees of California (Hymenoptera. Apidae)*. University of California Press; Berkeley and Los Angeles, California, 1983.

Uliano-Silva M, Ferreira JGRN, Krasheninnikova K, Blaxter M, Mieszkowska N, Hall N, Holland P, Durbin R, Richards T, Kersey P, et al. MitoHiFi: A python pipeline for mitochondrial genome assembly from PacBio high fidelity reads. *BMC Bioinformatics*. 2023;**24**:288. https://doi.org/10.1186/s12859-023-05385-y

U.S. Fish and Wildlife Service. *Recovery plan for rusty patched bumble bee*. Fish and Wildlife Service, Midwest Regional Office: U.S; 2021.

Warriner MD. Bumble bees (Hymenoptera: Apidae) of Texas: historical distributions. *swna*. 2012;**57**:442–445. https://doi.org/10.1894/0038-4909-57.4.442

Williams PH. An annotated checklist of bumble bees with an analysis of patterns of description (Hymenoptera: Apidae, Bombini). *Bulletin-Natural History Museum Entomology Series*. 1998;**67**: 79–152.

Williams P, Colla S, Xie Z. Bumblebee Vulnerability: Common Correlates of Winners and Losers across Three Continents. *Conservation Biology*. 2009;**23**:931-940. Portico. https://doi.org/10.1111/j.1523-1739.2009.01176.x

Williams PH Osborne JL. Bumblebee vulnerability and conservation world-wide. *Apidologie*. 2009;**40**:367–387. https://doi.org/10.1051/apido/2009025

Williams PH, Thorp RW, Richardson LL, Colla SR (2014). *Bumble bees of North America: an identification guide*. Princeton University Press, Princeton, New Jersey. https://doi.org/10.2307/j.ctt6wpzr9

Williams PH, Sagot P, Martínez-López O, Barajas RA, Mérida-Rivas JA, Vandame R. A new wave of Mesoamerican bumblebees? Revising the weisi-complex to reject numts and pseudospecies (Apidae: Bombus). *Zootaxa*. 2024;**5514**:301–318.

Zayed A, Packer L, Grixti JC, Ruz L, Owen RE, Toro H. Increased genetic differentiation in a specialist versus a generalist bee: implications for conservation. *Conserv Genet*. 2005;**6**: 1017–1026.

Journal of Heredity, 2026, 117, 611–622
https://doi.org/10.1093/jhered/esaf109
Genome Resources