G3, 2026, **16(3),** jkaf305
https://doi.org/10.1093/g3journal/jkaf305
Advance Access Publication Date: 23 December 2025
Genome Report

First genome assemblies of Neotropical Thoracobombus bumblebees *Bombus pauloensis* and *Bombus pullatus*

Andres Felipe Lizcano-Salas [1], Jesús Camilo Jacome-García,[2] Diego Riaño-Jiménez,[2,*] Marcela Guevara-Suarez[1,*]

[1]Applied Genomics Research Group, Universidad de los Andes, Bogotá 111711, Colombia
[2]Investigación Traslacional Exploratorium Shaio, Fundación Abood Shaio, Bogotá 111321, Colombia

*Corresponding authors: Marcela Guevara-Suarez, Applied Genomics Research Group, Universidad de los Andes, Bogotá 111711, Colombia. Email: mi.guevara34@uniandes.edu.co; Diego Riaño-Jiménez, Investigación Traslacional Exploratorium Shaio, Fundación Abood Shaio, Bogotá 111321, Colombia. Email: diego.riano@shaio.org

Bumblebees (*Bombus*) are considered to be essential pollinators of a wide range of flowering plants, within both agricultural and natural ecosystems. *Bombus pauloensis* and *Bombus pullatus* are 2 closely related Neotropical species with a wide altitudinal and latitudinal distribution that belong to the Thoracobombus genus. To the best of our knowledge, there is no genome assembly available for any species of Neotropical *Bombus*. Therefore, the goal of this study is to produce high-quality genomes of *B. pauloensis* and *B. pullatus*. In order to achieve this objective, we obtained long-read sequences using the Oxford Nanopore Technologies platform. We then proceeded to assemble the genomes and annotate these assemblies. As a result, we obtained assemblies of ~240 Mb represented in 72 contigs with an N50 of ~9.08 Mb for *B. pullatus* and ~239 Mb represented in 66 contigs with an N50 of ~9 Mb for *B. pauloensis*. The completeness evaluated by compleasm returned a score of >99% for both species. It is hoped that these genomes will facilitate a more profound comprehension of the biology of Neotropical bumblebees.

Keywords: Animalia; *Bombus pullatus*; *Bombus pauloensis*; insecticide resistance; synteny; genome assembly

Introduction

Bumblebees (*Bombus*) are essential pollinators of a wide range of flowering plants, within both agricultural and natural ecosystems. Despite its importance, bumblebees are declining worldwide due to multiple threats such as habitat loss, climate change, parasites, diseases, and pesticide use (Goulson et al. 2008; Sánchez-Bayo and Wyckhuys 2019). The subgenus Thoracobombus comprises around 50 species from the Old and New World (Williams et al. 2008). All members of this subgenus are characterized by the construction of ground nests, which are typically covered only by herbaceous plant material, such as grass stems (Carder bumblebees) (Williams et al. 2008). In this subgenus, *Bombus pullatus* and *Bombus pauloensis* are 2 common species of Neotropical ecosystems in Central and South America, respectively (Abrahamovich and Díaz 2002). According to Santos Júnior et al. (2022), *B. pauloensis* and *B. pullatus* are closely related species that diverged approximately 5 million years ago following the uplift of the Andes. This geological event likely contributed to their current distinct distribution patterns. Unlike other species, colonies of *B. pauloensis* and *B. pullatus* tend to be larger, often reaching 400 workers, and perennial (Cameron and Williams 2003; Hines et al. 2007; Riaño-Jiménez et al. 2020). Although these species are not currently classified as vulnerable or endangered, their populations may still be declining because of ongoing human-driven habitat alteration.

B. pauloensis is native to South American ecosystems, with a wide distribution across several countries, including Argentina, Bolivia, Brazil, Colombia, Paraguay, Peru, Uruguay, and Venezuela (Abrahamovich and Díaz 2002; Pinilla-Gallego et al. 2017). While it can be found at elevations ranging from 150 to 3,500 meters above sea level (masl), it is most observed between 1,800 and 2,800 masl (Abrahamovich and Díaz 2002; Pinilla-Gallego et al. 2017). This species exhibits a wide range of color variations, from melanic (completely black) to flavinic (yellow bands on the thorax and abdomen) and ferruginous (reddish bands on terga IV–VI) (Ospina et al. 1987; Pinilla-Gallego et al. 2017). *B. pauloensis* is the most studied South American bumblebee, with research encompassing both basic aspects (biology and ecology) and applications (breeding and crop pollination) (da Silva-Matos and Garofalo 1995; Cameron and Jost 1998; da Silva-Matos and Garófalo 2000; Gonzalez et al. 2004; Gamboa et al. 2015; Riaño-Jiménez et al. 2020; Salvarrey et al. 2020; Cruz et al. 2024). It is a common species that can thrive in a variety of habitats, including highly disturbed environments such as urban gardens, parks, and pastures used for cattle. A remarkable biological characteristic of *B. pauloensis* is its reproductive plasticity. Colonies can exhibit various social structures, including monogyny, polygyny, and competition, which allows for perennial colonies that can last for years, unlike most bumblebee species that have annual colonies (da Silva-Matos and Garofalo 1995; Cameron and Jost 1998; da Silva-Matos and Garófalo 2000). Furthermore, *B. pauloensis* is a highly effective pollinator in high Andean ecosystems and agroecosystems. This has led to the development of industrial captive rearing and commercialization of colonies for crop pollination purposes in several countries, such as Argentina (Almanza et al. 2005; Riaño et al. 2015; Salvarrey et al. 2020; Nery et al. 2024). However, little is known about the genome structure and the genetic basis of their unique characteristics.

B. pullatus is a little-known species recorded from lowlands of Central and Northern South America (Colombia, Costa Rica,

Received on 12 October 2025; accepted on 25 November 2025

Guatemala, Honduras, Nicaragua, Panama, and Venezuela) (Abrahamovich and Díaz 2002). Although it exhibits a wide altitudinal distribution (0 to 3,900 masl), *B. pullatus* is most found between 0 and 800 masl (Abrahamovich and Díaz 2002). *B. pullatus* is characterized by its melanic coloration, with short, dense black hairs and dark wings (Pinilla-Gallego et al. 2017). Little is known about the biology and ecology of *B. pullatus*, having described nest architecture and foraging activity in Costa Rica (Janzen 1971; Chavarria 1996; Hines et al. 2007). Nests of *B. pullatus* are like those described in other Thoracobombus such as *B. pauloensis* and *B. transversalis*, constructed over soil, mostly with small cut pieces of dried grass (Hines et al. 2007).

Genetic studies are an essential tool for understanding the vulnerability of bumblebee populations or species to factors linked to their decline, including climatic change, habitat loss, parasites and pathogens, pesticide use, and microplastics (Tsvetkov et al. 2021; Boeing et al. 2024; Eldem et al. 2025). Several bumblebee genomes have been sequenced, including those of *Bombus dahlbomii* (Martínez et al. 2024), *Bombus huntii* (Koch et al. 2024), *Bombus impatiens* (Sadd et al. 2015), and *Bombus terrestris* (Sadd et al. 2015). In addition, the genome of *Bombus pensylvanicus*, which is part of the same clade as the New World Thoracobombus species examined in this study, has recently been annotated (Lozier et al. 2025 Oct 8). The availability of this high-quality assembly will support broader comparative genomic analyses as more genomes from this lineage become available. However, the genetic diversity of Neotropical bumblebee species remains largely unknown, as, to the best of our knowledge, no genomes from this region have been sequenced to date.

Therefore, the goal of this study is to generate high-quality genomes of *B. pauloensis* and *B. pullatus*. These genomes represent the first sequenced genomes of Neotropical bumblebees. Given the ecological and biological significance of bumblebees, particularly as key pollinators in Neotropical ecosystems and agroecosystems (Abrahamovich and Díaz 2002), the sequencing of these genomes will enable the identification of gene families involved in detoxification, adaptation, and metabolic processes. This will serve as a valuable tool for understanding the impacts of pesticides, global warming, and diet on bumblebee health. Moreover, this endeavor will contribute to the advancement of knowledge in the fields of evolutionary dynamics and plant–pollinator interactions (Clare et al. 2013), with potential applications in fields such as conservation, evolutionary biology, and agricultural research.

Materials and methods

Sample collection and processing

B. pauloensis specimens (12 workers) were obtained from 3 colonies reared in captivity from queens from the municipality of Sopo, Cundinamarca, Colombia (4.908, −73.944). *B. pullatus* specimens (12 males) were obtained from a wild colony located in the municipality of Nilo, Cundinamarca, Colombia (4.34920, −74.65430).

The specimens were transported in a living state to the Sequencing Core Facility - GenCore (Universidad de los Andes, Bogotá, Colombia) and subsequently frozen in a refrigerator at −5 °C for 20 min. Then, the specimens were dissected in PBS buffer, removing the brain and thoracic muscle tissue (Farris et al. 1999). The tissues were immediately processed for DNA extraction to avoid DNA degradation during storage.

DNA extraction and sequencing

Three pools per species were processed: 2 pools of 6 brains and 1 pool of 2 thoraxes. DNA extraction was performed with a modified protocol of the QIAGEN DNeasy Blood & Tissue Kit. Briefly, each pool was ground using a sterile pestle. Next, each ground tissue was homogenized with 600 µL of lysis mix (540 µL Buffer ATL and 40 µL Proteinase K). Then, lysis was performed with the following cycles: 5 min disruption at 800 rpm, 25 min incubation at 56 °C, 5 min disruption at 800 rpm, 45 min incubation at 56 °C with 5 s vortexing every 10 min, and a final 10 s final vortexing. The disruption steps were performed in a BeadBlaster 96 Ball Mill Homogenizer. After, 600 µL Buffer AL and 600 µL absolute ethanol were added to each pool. Next, the solution was transferred to a column in 3 consecutive transfers. The washing steps were performed as specified by the manufacturer. Finally, the elution was performed with 100 µL of DNAse-free water preheated at 37 °C. All columns per each species were mixed during the elution. The concentration of the eluted DNA was verified using Qubit 4.0 with the Qubit dsDNA HS Assay Kit.

Nanopore sequencing libraries were prepared according to the genomic DNA Ligation Sequencing Kit V14 (SQK-LSK114) protocol of Oxford Nanopore Technologies (ONT). Prepared libraries were loaded on PromethION flow cells (R10.4.1) and sequenced with the PromethION 2 (P2) solo device. Finally, basecalling of raw ONT signal data was completed using Dorado v0.7.3 (https://github.com/nanoporetech/dorado) with sup model version 5.0.0.

Assembly

First, the adapters were trimmed using porechop v0.2.4 with the "-discard_middle" option (https://github.com/rrwick/Porechop). Reads were filtered into 2 datasets: dataset 1, reads of minimum length of 200 bp and mean quality of 20, and dataset 2, reads of minimum length of 10 Kb and mean quality of 20 (Table 1). Dataset 1 was used to estimate genome size with *k*-mer frequency analysis using KAT v2.4.2 (Mapleson et al. 2017) with *k*-mer lengths of 21, 27, and 31. Dataset 2 was used to assemble the genome with Flye v2.9.4-b1799 (Kolmogorov et al. 2019) with 2 polishing iterations and without alternative contigs. Then, the draft assembly was polished with medaka v1.12.1 (https://github.com/nanoporetech/medaka) using the reads of dataset 1. Contigs with lengths less than 50 Kb were removed because they were likely assembly artifacts. These artifacts often align to longest contigs, exhibit low sequencing depth, and are probably related to issues introduced by the complex pool of DNA used for assembly (these contigs represent less than 0.005% of the assembly). After, the assembly was checked for contamination using the NCBI Foreign Contamination Screen (FCS, https://github.com/ncbi/fcs) tool suite using the FCS-GX function on the Galaxy platform (https://usegalaxy.org/) with the "Animals (Metazoa) - insects" GX-division. A second verification step was performed in BlobToolKit v4.3.11 (Challis et al. 2020) using coverage and hit data. Briefly, reads from dataset 1 were mapped to the assembly using minimap2 v2.28 (Li 2018), and the mapping data were sorted with SAMtools v1.16.1 (Li et al. 2009). Contigs were searches against the core_nt database (accessed 2024 August 30)

Table 1. Statistics for raw and filtered reads.

Species	Dataset	Size (Gbp)	Number of reads (M)	Mean quality	N50
B. pullatus	Basecalled (raw)	53.5	17	14.1	6,245
	Dataset 1	31	10.15	23.4	6,109
	Dataset 2	9.6	0.62	23.1	15,455
B. pauloensis	Basecalled (raw)	47.4	15	13.7	7,202
	Dataset 1	27.5	8.7	23.6	6,949
	Dataset 2	10.2	0.62	23.2	16,701

with BLASTn v2.16.0 (Camacho et al. 2009) and against the UniProt reference proteome database (accessed 2024 August 30) with DIAMOND v2.1.9 (Buchfink et al. 2015) following the blobtools2 manual (https://blobtoolkit.genomehubs.org/). For BLAST and DIAMOND results, the assignments were made at a genus level. Next, duplicate core genes were identified by compleasm v0.2.6 (Huang and Li 2023) using the lineage Hymenoptera OrthoDB v10 database. Duplications of these genes between different contigs were manually checked in Gepard v2.1 (Krumsiek et al. 2007), and contigs identified as alternative haplotypes of a longer contig or assembling artifacts were removed (in each assembly, 2 contigs were removed). Then, a round of ntLink v1.10.3 (Coombe et al. 2023) was performed with a *k*-mer size of 40 and window of 500, including the *gap_filling* option, to improve the contiguity of the assembly. An additional polishing step was performed after the gap-filling. Briefly, reads of dataset 1 were mapped to the assembly with minimap2 to run 1 round of polishing in racon v1.5.0 (Vaser et al. 2017) with default parameters. Finally, a round of medaka with the reads of dataset 1 was performed.

Quality assessment and species verification

The quality of the final assembly was checked with BlobToolKit (Challis et al. 2020) (including coverage data as previously described) that includes metrics like the number of contigs, N50, depth, GC content and presence/absence of contaminant contigs, and compleasm (Huang and Li 2023) using the lineage Hymenoptera OrthoDB v10 database to assess the completeness of the assembly in terms of the presence of core single copy orthologs.

In order to verify species identity, 3 nuclear genes were analyzed: *arginine kinase* (*Argk*), *long-wavelength rhodopsin gene* (*Opsin*), and *phosphoenolpyruvate carboxykinase* (*PEPCK*); these genes had previously been used for phylogenetic analysis (Santos Júnior et al. 2022). The selection of these genes was based on the availability of sequences for both species (*B. pauloensis* and *B. pullatus*) in NCBI. Genomic regions of these genes were identified using BLASTn (Camacho et al. 2009), extracted using SAMtools (Li et al. 2009), and, if necessary, the reverse complement was identified using the revseq function in EMBOSS v6.6.0 (Rice et al. 2000). Then, each gene was aligned using MAFFT v7.525 (Katoh and Standley 2013) with the "–genafpair –maxiterate 1000" parameters. Finally, phylogenetic reconstruction was performed with *Bombus funerarius* as the outgroup using IQ-TREE2 v2.2.5 with 1,000 UFboostrap replicates (Hoang et al. 2018; Minh et al. 2020). The best-fit evolutionary model of each gene was selected using ModelFinder (Kalyaanamoorthy et al. 2017) under the AICc (corrected Akaike Information Criterion) criteria.

Genome annotation

The identification and masking of transposable elements (TEs) and repetitive sequences were conducted utilizing the TransposonUltimate pipeline (Riehl et al. 2022). Briefly, transposons were predicted employing the following software: HelitronScanner v1.0 (Xiong et al. 2014), LTRharvest v1.6.2 (Ellinghaus et al. 2008), LTRpred (Drost 2020), MiteFinderII v1.0.006 (Hu et al. 2018), MITE-Tracker v1.0.1 (Crescente et al. 2018), Must v2.4.001 (Ge et al. 2017), NCBI CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml), RepeatModeler v2.0.5 (Flynn et al. 2020), RepeatMasker v4.1.6 (https://www.repeatmasker.org/RepeatMasker/), SINE-Finder v1.0.1 (Wenke et al. 2011), SINE-Scan v1.1.2 (Mao and Wang 2017), TIRvish v1.6.2 (Gremme et al. 2013), and TransposonPSI (https://transposonpsi.sourceforge.net/). RepeatMasker was used with the parameters "-species hymenoptera -e abblast" against the Dfam database release 3.8 (accessed 2024 September 17). The results of each software were parsed, and duplicates were filtered. Additional transposon copies were identified by searching against software results using CD-HIT v4.8.1 (Fu et al. 2012) and BLASTn (Camacho et al. 2009). The final copies were filtered and annotated using random forest selective binary classifier (RFSB) (Riehl et al. 2022). Based on the identified transposons, the genomes were masked using bedtools v2.30.0.

Gene prediction was performed on the masked assembly using BRAKER3 v3.0.6 (Gabriel et al. 2024) with the OrthoDB v10 (Kriventseva et al. 2019) database of Hymenoptera as the protein database. Available RNA-seq data from *Bombus* species within the Thoracobombus subgenus (*B. dahlbomii* [SRR28005379], *Bombus muscorum* [ERR11837462], *Bombus pascuorum* [SRR6148372], *B. pascuorum* [SRR6148369, SRR6148376, and SRR6148366], and *Bombus opulentus* [SRR12527964]) were incorporated into the gene prediction process. Then, genes with incomplete models or ORFs shorter than 100 amino acids were filtered out prior to downstream analysis with AGAT v1.4.1 (Dainat et al. 2020).

After, genes were annotated using the Trinotate pipeline (https://github.com/Trinotate/Trinotate/wiki). Briefly, cDNA and protein sequences of each gene model were searched against the Swiss-Prot reference databases (The Uniprot Consortium 2024) with BLAST (Camacho et al. 2009). Protein sequences were then analyzed using HMMER v3.4 (https://www.ebi.ac.uk/Tools/hmmer/home) against the Pfam-A database (Mistry et al. 2021) to identify protein domains. Signal peptides were predicted with Signalp v6 (Teufel et al. 2022). TMHMM v2.0c (Krogh et al. 2001) was used to identify putative transmembrane regions, and eggNOG-mapper v2.1.8 (Cantalapiedra et al. 2021) with the eggNOG database v5.0.2 (Huerta-Cepas et al. 2019) was used for orthology prediction.

Finally, we identified tRNA using tRNAscan-SE v 2.0.12 (Chan et al. 2021) with the default parameters. For other noncoding RNA, cmscan function of Infernal v1.1.5 (Nawrocki and Eddy 2013) was used against the Rfam database (Kalvari et al. 2021) (accessed 2024 September 27) using the "–rfam –cut_ga –nohmmonly" parameters. Then, results were parsed and summarized.

Whole genome comparison

The available *B. pascuorum* genome assembly (GCF_905332965.1) was utilized to identify contigs associated with known chromosomes. Synteny analysis was performed using NGSEP v5.0.0 (Tello et al. 2023) with the *B. pascuorum* genome and annotation as reference. A synteny graph was generated using RIdeogram (Hao et al. 2020) to visualize synteny between *B. pascuorum* chromosomes and corresponding contigs in our assemblies. Additionally, paired comparisons were performed between our assemblies and assemblies of other Thoracobombus species, focusing on contigs/scaffolds associated with known chromosomes (*B. pascuorum* [GCF_905332965.1], *B. opulentus* [GCA_034509555.1], *B. muscorum* [GCA_963971185.1], and *B. dahlbomii* [GCA_037178635.1]). Dot plots were generated using D-GENIES v1.5.0 (Cabanettes and Klopp 2018) with minimap2 as mapping tools and the "Many repeats" option. For dot plot visualization, matches were sorted, short matches were filtered, and the "strong precision" option was enabled.

Results

High-quality ONT sequencing data provided sufficient coverage to facilitate the successful de novo assembly of the genomes of 2 *Bombus* species (Table 1). *k*-mer analysis estimated genome sizes between 239 to 246 Mb and 234 to 241 Mb for *B. pullatus* and *B. pauloensis*, respectively. The final assembly for *B. pullatus* comprised 72 contigs with an estimated genome size of approximately 240 Mb, while the *B. pauloensis* assembly consisted of 66 contigs and an estimated genome size of 239 Mb (Fig. 1a and b). The

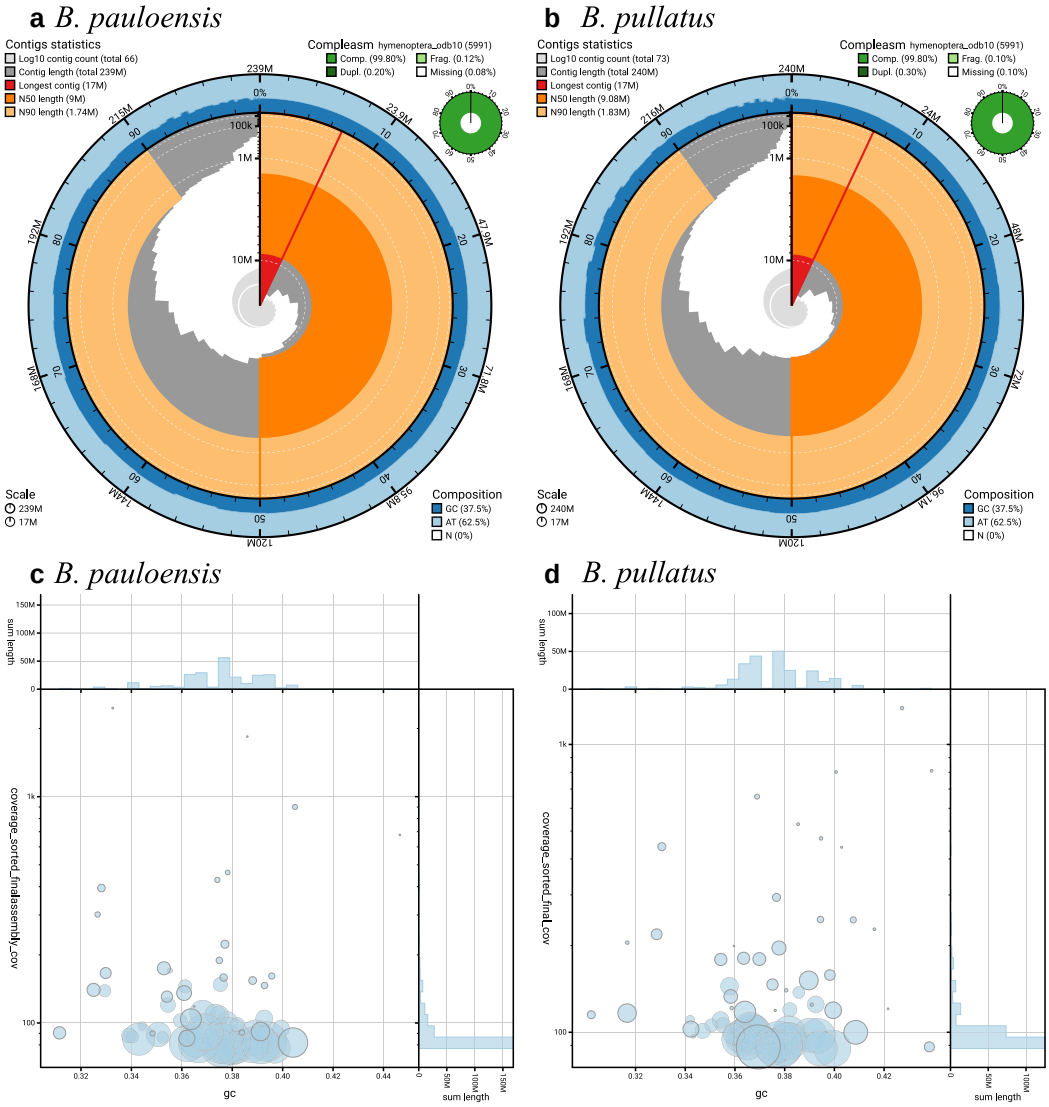

Fig. 1. Snail plot visualization of genome assembly statistics for a) *B. pauloensis* and b) *B. pullatus*. Blob plots showing the read depth and GC content of each contig for c) *B. pauloensis* and d) *B. pullatus*.

mean depth of each contig for the second dataset (used for polishing) was approximately 88× for *B. pullatus* and 91× for *B. pauloensis* (Fig. 1c and d). The N50 for each assembly was approximately 9 Mb for both species, and the completeness evaluated by compleasm was higher than 99%. Additionally, no contamination was detected in either genome assembly using the FCS and BlobToolKit pipelines. The genomes were verified using 3 nuclear genes and phylogenetically are grouped with previously known sequences of *B. pauloensis* and *B. pullatus* (Fig. 2).

TE analysis shows that approximately 15% of each genome is composed of TEs. The Zator superfamily of DNA transposons exhibited the highest copy number in both assemblies (Table 2). However, the hAT superfamily of DNA transposons has the highest representation in both genomes in terms of base pairs represented by these TEs (Table 2), accounting for approximately 4.2% to 4.5% of the genome in both species.

Protein-coding gene annotation in *B. pauloensis* resulted in 10,662 genes with 14,941 transcripts, while *B. pullatus* had 10,681 genes with 14,890 transcripts (Table 3). The average gene length was 8,798 bp in *B. pauloensis* (range: 306 to 753,695 bp) and 8,704 bp in *B. pullatus* (range: 306 to 759,035 bp). Both species exhibited an average of 1.4 transcripts per gene. During the annotation, 1,987 and 1,985 genes in *B. pullatus* and *B. pauloensis*, respectively, could not be annotated, lacking associations with known functions, domains, biological processes, or cellular localizations.

Synteny analysis revealed that 45 and 47 contigs in *B. pauloensis* and *B. pullatus*, respectively, could be associated with known chromosomes of *B. pascuorum* (Fig. 3), representing approximately 96% of each genome. This analysis indicated high synteny between our assemblies and the *B. pascuorum* assembly, with evidence of few translocation and inversion events (Fig. 3).

Comparison of chromosome-associated contigs between *B. pauloensis* and *B. pullatus* revealed high synteny with limited inversions and translocations (Fig. 4a). Additionally, comparisons with previously reported assemblies of other Thoracobombus species revealed some deletions in our assemblies, likely due to the smaller genome sizes of these species (Fig. 4b to i). Furthermore, our assemblies exhibited fewer inversions compared to the *B. opulentus* assembly (Fig. 4d and g).

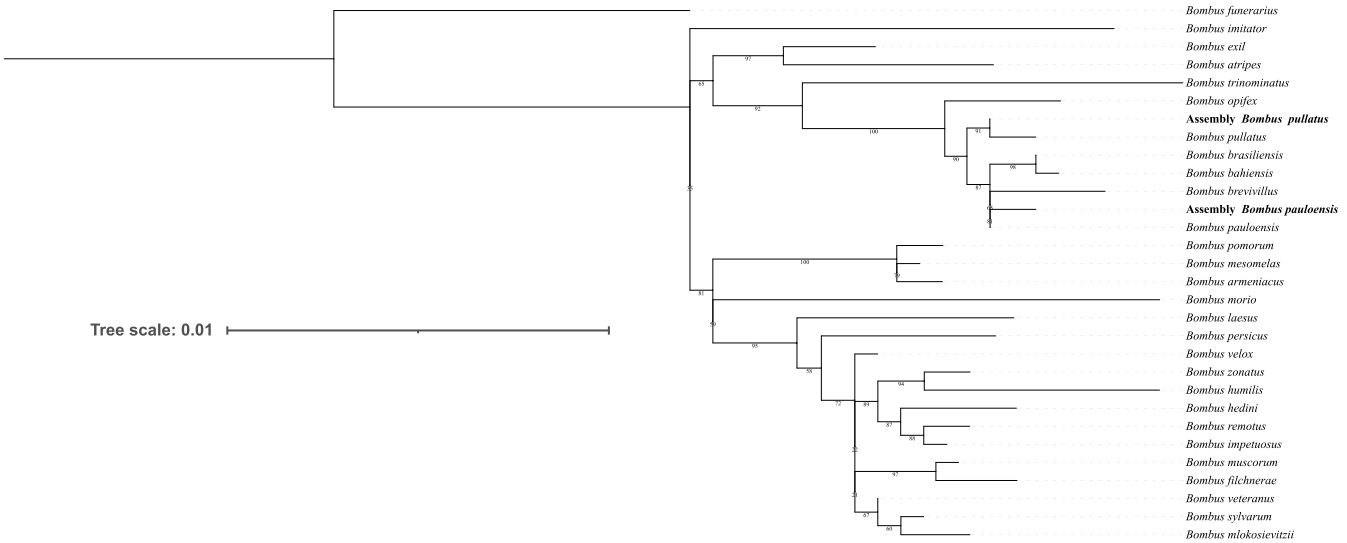

Fig. 2. Phylogenetic tree inferred from a concatenated set of genes including *Argk* (521 bp), *Opsin* (648 bp), and *PEPCK* (490 bp) of 27 Thoracobombus species. Maximum-likelihood analysis with 1,000 ultrafast bootstrap replicates performed in IQ-Tree. *B. funerarius* was used as the outgroup. Support values are shown below branches and represent bootstrap values.

Table 2. Distribution of TE superfamilies.

Species	Class	Superfamily	Copy number	Base pair representation (bp)
B. pauloensis	Retrotransposon	Copia	187	849,643
		Gypsy	1,190	3,103,272
		ERV	2	2,870
		LINE	103	914,918
		SINE	0	0
	DNA transposon	Tc1-Mariner	2,508	8,821,212
		hAT	3,518	9,988,851
		CMC	1,290	9,581,690
		Sola	860	333,438
		Zator	4,603	1,406,201
		Novosib	43	275,165
		Helitron	25	162,339
		MITE	1,466	898,266
B. pullatus	Retrotransposon	Copia	221	443,799
		Gypsy	1,264	4,253,208
		ERV	3	1,053
		LINE	168	1,188,672
		SINE	0	0
	DNA transposon	Tc1-Mariner	2,383	8,835,882
		hAT	3,687	10,915,817
		CMC	1,393	8,400,228
		Sola	827	272,705
		Zator	4,424	1,260,866
		Novosib	57	252,602
		Helitron	47	205,034
		MITE	814	479,891

Discussion

In this study, we present the first genome assembly of *B. pauloensis* and *B. pullatus*, 2 Neotropical species present in Colombia. The genome sizes obtained for *B. pullatus* (240 Mb) and for *B. pauloensis* (239 Mb) are consistent with previously estimated genome sizes for *Bombus* species, which range from 230 to 393 Mb (Sun et al. 2021). These estimates also fall within the range reported for other species within the Thoracobombus subgenus (234 to 317 Mb). As expected, the genome sizes of *B. pauloensis* and *B. pullatus* are similar, given their close phylogenetic relationship (Santos Júnior et al. 2022). The results of the *k*-mer analysis closely predicted the expected genome size for both species, as previously reported in other studies (Koch et al. 2023). With regard to TEs, the proportion of the genome occupied by TEs falls within the expected range for the genus *Bombus* (9% to 17%) as reported by Sun et al. (2021).

During our analysis, we observed that chromosome naming in previously published genome assemblies of species within the Thoracobombus subgenus was often based on scaffold length (Crowley et al. 2023; Broad and Barnes 2024; Martínez et al. 2024; Sang et al. 2024). To maintain consistency in chromosome

nomenclature with previously published assemblies, we organized and associated our contigs with the known chromosomes of *B. pascuorum*. This species was selected because it is currently the only member of this subgenus with a reference genome available in the RefSeq database (GCF_905332965.1) (Crowley et al. 2023). During the synteny analysis, we identified a few rearrangements within the selected species in the Thoracobombus subgenus (Fig. 4). These results suggest a degree of synteny conservation within the Thoracobombus subgenus. However, as previous studies have established, species within the genus *Bombus* do not exhibit a fixed chromosome number (Owen et al. 1995). The reported range varies from 12 to 19, with most species possessing 18 chromosomes (Owen et al. 1995). More recent genomic data have even reported up to 25 chromosomes (Crowley 2025). In the context of our comparative analysis, 2 of the species used, *B. pascuorum* and *B. muscorum*, have 17 reported chromosomes (Crowley et al. 2023; Broad and Barnes 2024), while 2 others, *B. opulentus* and *B. dahlbomii*, have 18 reported chromosomes (Martínez et al. 2024; Sang et al. 2024). Given this variability and the absence of Hi-C data in our project, future genomic work on *B. pauloensis* and *B. pullatus* should include efforts to accurately determine their respective chromosome numbers.

Table 3. Annotation statistics of assemblies.

		B. pauloensis	*B. pullatus*
Genes	Total	11,202	11,179
Protein-coding genes	Total	10,662	10,681
Noncoding genes	Total	540	498
	Cis-reg	13	11
	miRNA	65	64
	Ribozymes	3	4
	rRNA	100	81
	snoRNA	15	14
	snRNA	62	70
	SRP RNA	4	4
	tRNA	278	250

The number of protein-coding genes identified in our study falls within the previously reported range for *Bombus* species, which is between 10,000 and 17,000 genes (Heraghty et al. 2020; Sun et al. 2021; Koch et al. 2023, 2024; Martínez et al. 2024). However, the relatively lower number of genes identified in the present study compared to previous ones (Heraghty et al. 2020; Koch et al. 2023, 2024; Martínez et al. 2024) may be attributed to the lack of identification of long noncoding RNAs. Long noncoding RNAs represent the larger group of noncoding genes in *Bombus* (Sun et al. 2021). It is therefore recommended that future studies concentrate on the identification of long noncoding RNAs in these species of *Bombus*, with a view to achieving a more comprehensive understanding of their gene repertoire and their potential roles in the biology of bumblebees.

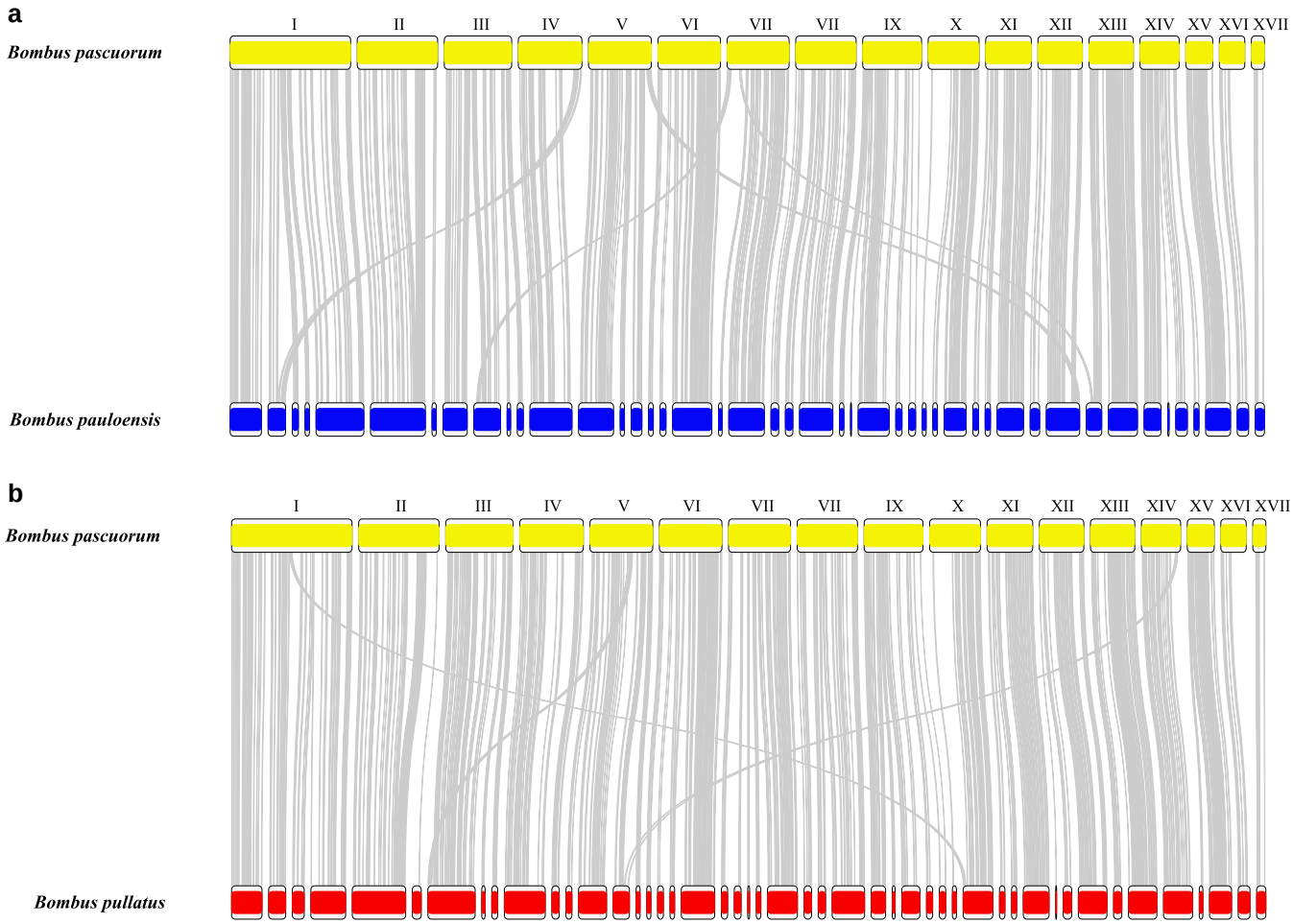

Fig. 3. Synteny plot of *B. pascuorum* with a) *B. pauloensis* and b) *B. pullatus*. Roman numerals represent the chromosome number for *B. pascuorum*.

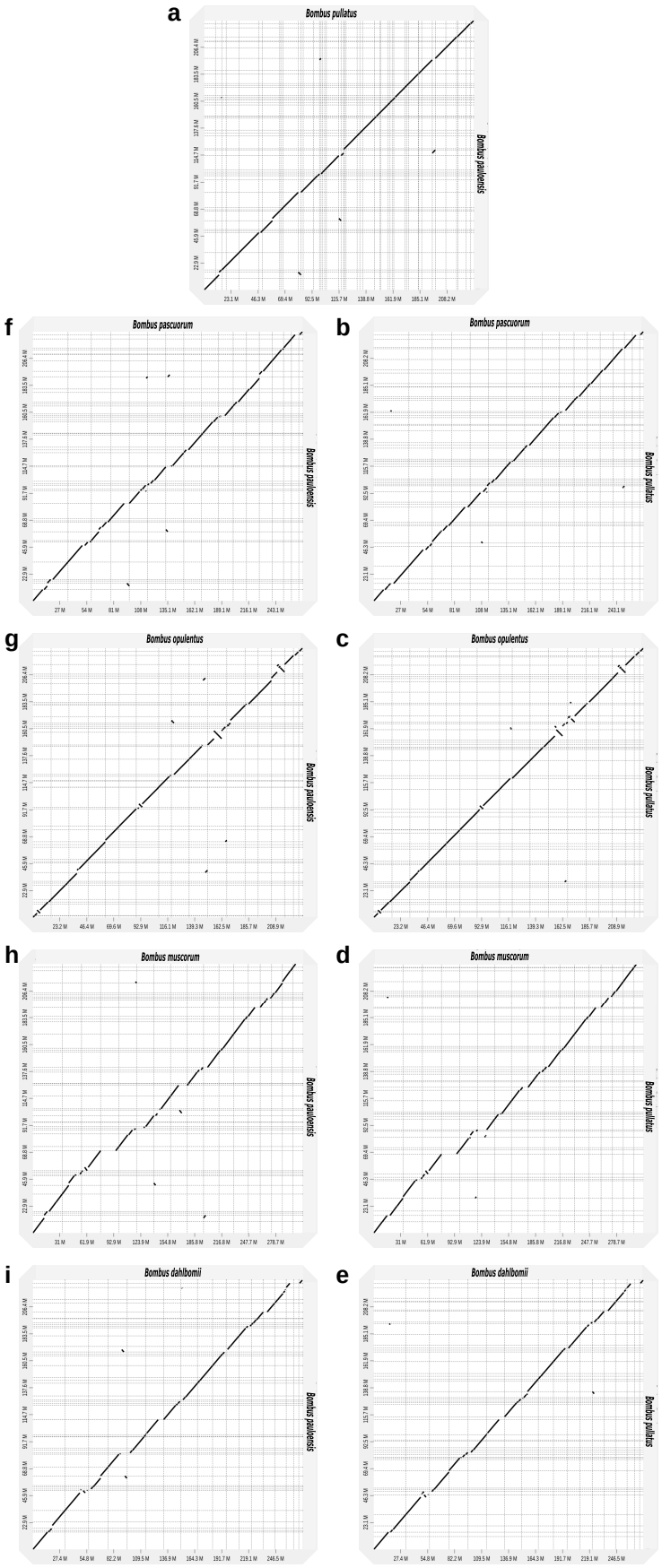

Fig. 4. Dot plot of *B. pullatus* against a) *B. pauloensis*, b) *B. pascuorum*, c) *B. opulentus*, d) *B. muscorum*, and e) *B. dahlbomii*. Dot plots of *B. pauloensis* against f) *B. pascuorum*, g) *B. opulentus*, h) *B. muscorum*, and i) *B. dahlbomii*.

In conclusion, the present study presents the first high-quality genome assemblies of 2 closely related Neotropical *Bombus* species in Colombia. These genomes exhibit a high degree of contiguity; however, they have not yet been delineated at the chromosome level, a characteristic that distinguishes them from some previously described genomes within this genus. It is evident that further efforts are required in order to complete these assemblies and to improve genome annotation. This will facilitate a more profound comprehension of insecticide resistance, evolutionary dynamics, and adaptations in these and other Neotropical species. These assemblies will be important for identifying genes associated with various traits of interest in these species and ultimately contribute to their conservation. Finally, we hope that this work represents only the first step toward the genetic conservation of Neotropical bumblebees.

Data availability

The sequencing data were deposited under BioProject PRJNA1242843. The reads were deposited in the sequence read archive (SRA) under accession numbers SRR32927162 and SRR32927163. The genome assemblies were deposited in GenBank under accession numbers JBNQWY000000000 and JBNQWZ000000000. The scripts used in this study, along with select results pertaining to transposon identification, gene prediction, and annotation, are deposited in a GitHub repository (https://github.com/andres2901/NeotropicalBumbleBeesAssemblies).

Acknowledgments

To Constanza Mendoza, manager of the Mana Dulce Nature Reserve, Nilo, Cundinamarca, for allowing the collection of *B. pullatus* samples. The present study would not have been possible without the sequencing and bioinformatics analysis support provided by the Sequencing Core Facility - GenCore at Universidad de los Andes. The High-Performance Computing (HPC) Service at Universidad de los Andes is also acknowledged for providing HPC resources necessary to obtain these results in this study. Furthermore, gratitude is extended to the National Environmental Licensing Authority (ANLA) for the approval of the Marco permit for sampling (contract 02492).

Funding

This work was supported by a grant from the Ministerio de Ciencia, Tecnología e Innovación-Minciencias, the Abood Shaio Foundation, and the Universidad de los Andes (contract 036-2024).

Conflicts of interest. None declared.

Author contributions

A.F.L.-S.: design, data collection, DNA extraction, sequencing, bioinformatics analysis, and manuscript writing. J.C.J.-G.: design, *B. pauloensis* rearing, data collection, and manuscript writing. D.R.-J.: design, *B. pauloensis* rearing, data collection, manuscript writing, and project administration. M.G.-S.: design and manuscript writing.

Literature cited

Abrahamovich AH, Díaz NB. 2002. Bumble bees of the Neotropical Region (Hymenoptera: Apidae). Biota Colomb. 3:199–214. https://revistas.humboldt.org.co/index.php/biota/article/view/115.

Almanza MT, Cropscience B, Benavides LA. 2005. Native bumblebees rearing for pollination of crops in the highlands of Colombia. Fao Report: Case Studies on Pollinators and pollination.

Boeing GANS, *et al.* 2024. Spray paint-derived microplastics and incorporated substances as ecotoxicological contaminants in the Neotropical bumblebee *Bombus atratus*. Environ Toxicol Pharmacol. 112:104586. https://doi.org/10.1016/J.ETAP.2024.104586.

Broad GR, Barnes I. 2024. The genome sequence of the moss carder bee, *Bombus muscorum* (Linnaeus, 1758). Wellcome Open Res. 9:397. https://doi.org/10.12688/wellcomeopenres.22739.1.

Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 12:59–60. https://doi.org/10.1038/nmeth.3176.

Cabanettes F, Klopp C. 2018. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. PeerJ 6:e4958. https://doi.org/10.7717/peerj.4958.

Camacho C, *et al.* 2009. BLAST+: architecture and applications. BMC Bioinform. 10:1–9. https://doi.org/10.1186/1471-2105-10-421.

Cameron SA, Jost MC. 1998. Mediators of dominance and reproductive success among queens in the cyclically polygynous Neotropical bumble bee *Bombus atratus* Franklin. Insectes Soc. 45:135–149. https://doi.org/10.1007/s000400050075.

Cameron SA, Williams PH. 2003. Phylogeny of bumble bees in the New World subgenus *Fervidobombus* (Hymenoptera: Apidae): congruence of molecular and morphological data. Mol Phylogenet Evol. 28:552–563. https://doi.org/10.1016/S1055-7903(03)00056-3.

Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 38:5825–5829. https://doi.org/10.1093/molbev/msab293.

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

Chan PP, Lin BY, Mak AJ, Lowe TM. 2021. TRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res. 49:9077–9096. https://doi.org/10.1093/nar/gkab688.

Chavarria G. 1996. Notes on a combined nest of *Bombus pullatus* (Hymenoptera: Apidae) and *Acromyrmex octospinosus* (Hymenoptera: Formicidae). J Kans Entomol Soc. 69:403–405. http://www.jstor.org/stable/25085736.

Clare EL, Schiestl FP, Leitch AR, Chittka L. 2013. The promise of genomics in the study of plant-pollinator interactions. Genome Biol 14:207. https://doi.org/10.1186/gb-2013-14-6-207.

Coombe L, Warren RL, Wong J, Nikolic V, Birol I. 2023. ntLink: a toolkit for de novo genome assembly scaffolding and mapping using long reads. Curr Protoc. 3:e733. https://doi.org/10.1002/CPZ1.733.

Crescente JM, Zavallo D, Helguera M, Vanzetti LS. 2018. MITE Tracker: an accurate approach to identify miniature inverted-repeat transposable elements in large genomes. BMC Bioinform. 19:1–10. https://doi.org/10.1186/s12859-018-2376-y.

Crowley LM. 2025. The genome sequence of the vestal cuckoo bee, *Bombus vestalis* (Geoffroy, 1785). Wellcome Open Res. 10:218. https://doi.org/10.12688/wellcomeopenres.24004.1.

Crowley LM, Sivell O, Sivell D. 2023. The genome sequence of the common carder bee, *Bombus pascuorum* (Scopoli, 1763). Wellcome Open Res. 8:142. https://doi.org/10.12688/wellcomeopenres.19251.1.

Cruz CP, Riaño-Jiménez D, Hakim JRC. 2024. Pollination efficiency and foraging behavior of *Bombus pauloensis* (Hymenoptera: Apidae) on two highbush blueberry cultivars (*Vaccinium corymbosum*). Sociobiology. 71:e9222. https://doi.org/10.13102/sociobiology.v71i1.9222.

Dainat J. 2020. AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF/GFF format. (Version v1.4.1). Zenodo. https://www.doi.org/10.5281/zenodo.3552717.

da Silva-Matos EV, Garofalo CA. 1995. Observations on the development of queenless colonies of *Bombus atratus* (Hymenoptera, Apidae). J Apic Res. 34:177–185. https://doi.org/10.1080/00218839.1995.11100903.

da Silva-Matos EV, Garófalo CA. 2000. Worker life tables, survivorship, and longevity in colonies of *Bombus* (*Fervidobombus*) *atratus* (Hymenoptera: Apidae). Rev Biol Trop. 48:657–664. https://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S0034-77442000000200038.

Drost H-G. 2020. LTRpred: de novo annotation of intact retrotransposons. J Open Source Softw. 5:2170. https://doi.org/10.21105/joss.02170.

Eldem V, *et al.* 2025. De novo genome assembly and annotations of *Bombus lapidarius* and *Bombus niveatus* provide insights into the environmental adaptability. Apidologie. 56:1–25. https://doi.org/10.1007/s13592-024-01140-1.

Ellinghaus D, Kurtz S, Willhoeft U. 2008. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. BMC Bioinform. 9:18. https://doi.org/10.1186/1471-2105-9-18.

Farris SM, Robinson GE, Davis RL, Fahrbach SE. 1999. Larval and pupal development of the mushroom bodies in the honey bee, *Apis mellifera*. J Comp Neurol. 414:97–113. https://doi.org/10.1002/(SICI)1096-9861(19991108)414:1<97::AID-CNE8>3.0.CO;2-Q.

Flynn JM, *et al.* 2020. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A. 117:9451–9457. https://doi.org/10.1073/pnas.1921046117.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 28:3150–3152. https://doi.org/10.1093/bioinformatics/bts565.

Gabriel L, *et al.* 2024. BRAKER3: fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. Genome Res. 34:769–777. https://doi.org/10.1101/gr.278090.123.

Gamboa V, *et al.* 2015. Bee pathogens found in *Bombus atratus* from Colombia: a case study. J Invertebr Pathol. 129:36–39. https://doi.org/10.1016/j.jip.2015.05.013.

Ge R, *et al.* 2017. MUSTv2: an improved de novo detection program for recently active miniature inverted repeat transposable elements (MITEs). J Integr Bioinform. 14:1–7. https://doi.org/10.1515/jib-2017-0029.

Gonzalez VH, Mejia A, Rasmussen C. 2004. Ecology and nesting behavior of *Bombus atratus* Franklin in Andean Highlands (Hymenoptera: Apidae). J Hymenopt Res. 13:234–242. https://www.biodiversitylibrary.org/part/19214.

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

Gremme G, Steinbiss S, Kurtz S. 2013. GenomeTools: a comprehensive software library for efficient processing of structured genome annotations. IEEE/ACM Trans Comput Biol Bioinform. 10:645–656. https://doi.org/10.1109/TCBB.2013.68.

Hao Z, *et al.* 2020. RIdeogram: drawing SVG graphics to visualize and map genome-wide data on the idiograms. PeerJ Comput Sci. 6:e251. https://doi.org/10.7717/peerj-cs.251.

Heraghty SD, *et al.* 2020. De novo genome assemblies for three North American bumble bee species: *Bombus bifarius*, *Bombus vancouverensis*, and *Bombus vosnesenskii*. G3 (Bethesda). 10:2585–2592. https://doi.org/10.1534/g3.120.401437.

Hines HM, Cameron SA, Deans AR. 2007. Nest architecture and foraging behavior in *Bombus pullatus* (Hymenoptera: Apidae), with comparisons to other tropical bumble bees. J Kans Entomol Soc. 80:1–15. https://doi.org/10.2317/0022-8567(2007)80[1:NAAFBI]2.0.CO;2.

Hoang DT, *et al.* 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 35:518–522. https://doi.org/10.1093/molbev/msx281.

Hu J, Zheng Y, Shang X. 2018. MiteFinderII: a novel tool to identify miniature inverted-repeat transposable elements hidden in eukaryotic genomes. BMC Med Genomics. 11:101. https://doi.org/10.1186/s12920-018-0418-y.

Huang N, Li H. 2023. Compleasm: a faster and more accurate reimplementation of BUSCO. Bioinformatics. 39:2023.06.03.543588. https://doi.org/10.1093/BIOINFORMATICS/BTAD595.

Huerta-Cepas J, *et al.* 2019. EggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 47:D309–D314. https://doi.org/10.1093/nar/gky1085.

Janzen DH. 1971. The ecological significance of an arboreal nest of *Bombus pullatus* in Costa Rica. J Kans Entomol Soc. 44:210–216. https://www.jstor.org/stable/25082404.

Kalvari I, *et al.* 2021. Rfam 14: expanded coverage of metagenomic, viral and microRNA families. Nucleic Acids Res. 49(D1):D192–D200. https://doi.org/10.1093/nar/gkaa1047.

Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 14:587–589. https://doi.org/10.1038/nmeth.4285.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30:772–780. https://doi.org/10.1093/molbev/mst010.

Koch JBU, Sim SB, Scheffler B, Geib SM, Smith TA. 2023. Chromosome-scale genome assembly of the rusty patched bumble bee, *Bombus affinis* (Cresson) (Hymenoptera: Apidae), an endangered North American pollinator. G3 (Bethesda). 13:jkad119. https://doi.org/10.1093/g3journal/jkad119.

Koch JBU, Sim SB, Scheffler B, Lozier JD, Geib SM. 2024 May 2. Chromosome-scale genome assembly of the Hunt bumble bee, *Bombus huntii* Greene, 1860, a species of agricultural interest. G3 (Bethesda). 14:jkae160. https://doi.org/10.1101/2024.04.30.591905.

Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.

Kriventseva Ev, *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 Res. 47(D1):D807–D811. https://doi.org/10.1093/nar/gky1053.

Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 305:567–580. https://doi.org/10.1006/jmbi.2000.4315.

Krumsiek J, Arnold R, Rattei T. 2007. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinformatics. 23:1026–1028. https://doi.org/10.1093/BIOINFORMATICS/BTM039.

Li H, *et al.* 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 25:2078–2079. https://doi.org/10.1093/BIOINFORMATICS/BTP352.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 34:3094–3100. https://doi.org/10.1093/BIOINFORMATICS/BTY191.

Lozier JD, *et al.* 2025. New reference genome assembly for the declining American Bumble Bee, *Bombus pensylvanicus*. G3 (Bethesda). 15:jkaf181. https://doi.org/10.1093/g3journal/jkaf181.

Mao H, Wang H. 2017. SINE-scan: an efficient tool to discover short interspersed nuclear elements (SINEs) in large-scale genomic datasets. Bioinformatics. 33:743–745. https://doi.org/10.1093/bioinformatics/btw718.

Mapleson D, Accinelli GG, Kettleborough G, Wright J, Clavijo BJ. 2017. KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. Bioinformatics. 33:574–576. https://doi.org/10.1093/BIOINFORMATICS/BTW663.

Martínez L, *et al.* 2024. Chromosome-level assembly and annotation of the genome of the endangered giant Patagonian bumble bee *Bombus dahlbomii*. Genome Biol Evol. 16:evae146. https://doi.org/10.1093/gbe/evae146.

Minh BQ, *et al.* 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 37:1530–1534. https://doi.org/10.1093/MOLBEV/MSAA015.

Mistry J, *et al.* 2021. Pfam: The protein families database in 2021. Nucleic Acids Res. 49(D1):D412–D419. https://doi.org/10.1093/nar/gkaa913.

Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics. 29:2933–2935. https://doi.org/10.1093/bioinformatics/btt509.

Nery D, Palottini F, Farina WM. 2024. The South American black bumblebee (*Bombus pauloensis*) as a potential pollinator of alfalfa (*Medicago sativa*). Agriculture. 14:2192. https://doi.org/10.3390/agriculture14122192.

Ospina TR, Liévano A, Nates G. 1987. El patrón de coloración del abejorro social *Bombus atratus*, Franklin en Cundinamarca, Colombia: una población diferenciada. Rev Biol Trop. 35:317–324. https://archivo.revistas.ucr.ac.cr/index.php/rbt/article/view/24096.

Owen RE, Richards KW, Wilkes A. 1995. Chromosome numbers and karyotypic variation in bumble bees (Hymenoptera: Apidae: Bombini). J Kans Entomol Soc. 68:290–302. https://www.jstor.org/stable/25085597.

Pinilla-Gallego MS, Osipna R, Cure JR. 2017. Los abejorros de páramo. In: Nates Parra G, editors. Iniciativa Colombiana de Polinizadores. Universidad Nacional de Colombia. p. 129–142.

Riaño-Jiménez D, Guerrero M, Alarcón P, Cure JR. 2020. Effects of climate variability on queen production and pollen preferences of Neotropical Bumblebee *Bombus atratus* in a high Andean suburban condition. Neotrop Entomol. 49:586–594. https://doi.org/10.1007/s13744-019-00758-6.

Riaño JD, Pacateque EJ, Cure JR, Rodríguez D. 2015. Comportamiento y eficiencia de polinización de *Bombus atratus* Franklin en pimentón (*Capsicum annum* L.) sembrado bajo invernadero. Revista Colombiana de Ciencias Hortícolas. 9:259. https://doi.org/10.17584/rcch.2015v9i2.4182.

Rice P, Longden L, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. Trends Genet. 16:276–277. https://doi.org/10.1016/S0168-9525(00)02024-2.

Riehl K, Riccio C, Miska EA, Hemberg M. 2022. TransposonUltimate: software for transposon classification, annotation and detection. Nucleic Acids Res. 50:e64. https://doi.org/10.1093/nar/gkac136.

Sadd BM, *et al.* 2015. The genomes of two key bumblebee species with primitive eusocial organization. Genome Biol. 16:76. https://doi.org/10.1186/s13059-015-0623-3.

Salvarrey S, Santos E, Arbulo N, Gimenéz G, Invernizzi C. 2020. Characteristics of the tomato fruit (*Solanum lycopersicum*) using native bumblebees (*Bombus atratus*) as pollinators in greenhouse. Agrociencia Uruguay 24:101. https://doi.org/10.31285/AGRO.24.101.

Sánchez-Bayo F, Wyckhuys KAG. 2019. Worldwide decline of the entomofauna: a review of its drivers. Biol Conserv. 232:8–27. https://doi.org/10.1016/J.BIOCON.2019.01.020.

Sang H, *et al.* 2024. Conservation genomics analysis reveals recent population decline and possible causes in bumblebee *Bombus opulentus*. Insect Sci. 31:1631–1644. https://doi.org/10.1111/1744-7917.13324.

Santos Júnior JE, *et al.* 2022. Biogeography and diversification of bumblebees (Hymenoptera: Apidae), with emphasis on Neotropical species. Diversity (Basel). 14:238. https://doi.org/10.3390/d14040238.

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

Tello D, *et al.* 2023. NGSEP 4: efficient and accurate identification of orthogroups and whole-genome alignment. Mol Ecol Resour. 23:712–724. https://doi.org/10.1111/1755-0998.13737.

Teufel F, *et al.* 2022. Signalp 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40:1023–1025. https://doi.org/10.1038/s41587-021-01156-3.

The UniProt Consortium. 2024. UniProt: the Universal protein knowledgebase in 2025. Nucleic Acids Res. 53(D1):D609–D617. https://doi.org/10.1093/nar/gkae1010.

Tsvetkov N, MacPhail VJ, Colla SR, Zayed A. 2021. Conservation genomics reveals pesticide and pathogen exposure in the declining bumble bee *Bombus terricola*. Mol Ecol. 30:4220–4230. https://doi.org/10.1111/MEC.16049.

UniProt consortium The. 2025. UniProt: the Universal protein knowledgebase in 2025. Nucleic Acids Res. 53(D1):D609–D617. https://doi.org/10.1093/nar/gkae1010.

Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 27:737–746. https://doi.org/10.1101/gr.214270.116.

Wenke T, *et al.* 2011. Targeted identification of short interspersed nuclear element families shows their widespread existence and extreme heterogeneity in plant genomes. Plant Cell. 23:3117–3128. https://doi.org/10.1105/TPC.111.088682.

Williams PH, Cameron SA, Hines HM, Cederberg B, Rasmont P. 2008. A simplified subgeneric classification of the bumblebees (genus *Bombus*). Apidologie. 39:46–74. https://doi.org/10.1051/apido:2007052.

Xiong W, He L, Lai J, Dooner HK, Du C. 2014. HelitronScanner uncovers a large overlooked cache of Helitron transposons in many plant genomes. Proc Natl Acad Sci U S A. 111:10263–10268. https://doi.org/10.1073/pnas.1410068111.

Editor: S. Ramírez