Annals of the Entomological Society of America, XX(X), 2024, 1–15
https://doi.org/10.1093/aesa/saae023
Special Collection

Special Collection

Genomic analyses of the southern and eastern yellowjacket wasps (Hymenoptera: Vespidae) reveal evolutionary signatures of social life

Michael A. Catto[1,2], Paige B. Caine[2], Sarah E. Orr[2], Brendan G. Hunt[1], Michael A. D. Goodisman[2,*]

[1]Department of Entomology, University of Georgia, Athens, GA 30602, USA, [2]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA *Corresponding author, email: michael.goodisman@biology.gatech.edu

Subject Editor: Lisa Knolhoff

Received on 29 January 2024; revised on 30 July 2024; accepted on 13 August 2024

Insects have evolved remarkably complex social systems. Social wasps are particularly noteworthy because they display gradations in social behaviors. Here, we sequence the genomes of two highly diverged *Vespula* wasps, *V. squamosa* and *V. maculifrons* Buysson (Hymenoptera: Vespidae), to gain greater insight into the evolution of sociality. Both *V. squamosa* and *V. maculifrons* are social wasps that live in large colonies characterized by distinct queen and worker castes. However, *V. squamosa* is a facultative social parasite, and *V. maculifrons* is its frequent host. We found that the genomes of both species were ~200 Mbp in size, similar to the genome sizes of congeneric species. Analyses of gene expression from members of different castes and developmental stages revealed similarities in expression patterns among immature life stages. We also found evidence of DNA methylation within the genome of both species by directly analyzing DNA sequence reads. Moreover, genes that were highly and uniformly expressed were also relatively highly methylated. We further uncovered evidence of differences in patterns of molecular evolution in the two taxa, consistent with *V. squamosa* exhibiting alterations in evolutionary pressures associated with its facultatively parasitic or polygyne life history. Finally, rates of gene evolution were correlated with variation in gene expression between castes and developmental stages, as expected if more highly expressed genes were subject to stronger levels of selection. Overall, this study expands our understanding of how social behavior relates to genome evolution in insects.

Keywords: eusocial, genomics, epigenetics, DNA methylation, gene expression, molecular evolution, i5k

Introduction

Highly social insects are remarkable because society members belong to phenotypically distinct castes. Members of each caste take on different tasks within the colony and often differ widely in shape, size, physiology, and behavior. For example, in hymenopteran social insects, the queen and male castes mate and reproduce while the worker caste undertakes activities related to foraging, nest building, nursing, and colony function (Oster and Wilson 1978, Wilson and Hölldobler 2005). The integration among castes maximizes the success of insect societies.

Social wasps in the family Vespidae represent important subjects for understanding the ecology, evolution, and behavior of social insects, because vespid wasps show great variation in levels of sociality (Hunt and Toth 2017, Taylor et al. 2018, Wyatt et al. 2023). Vespid wasps in the genus *Vespula* form populous colonies and are widespread across the northern hemisphere (Akre and MacDonald 1986, Greene 1991). Further, some *Vespula* species are highly invasive and have been introduced to a variety of locations around the world (Lester and Beggs 2019, Wilson-Rankin 2021).

New *Vespula* colonies are typically initiated by a single queen, which founds a colony independently after emerging from hibernation (Greene 1991). The queen establishes the nest and rears the first set of workers, who then take over all functions of maintaining the nest and rearing the developing young. The colony grows throughout the year and produces new queens and males at the end of the reproductive season. However, some *Vespula* species show variation

in their life cycle and do not found their own nest independently. Instead, some *Vespula* species engage in socially parasitic behaviors (Rabeling 2020).

Social parasites are species that take advantage of existing societies for their own end; this lifestyle has evolved independently at least 88 times and is found in ants, social bees, and social wasps (Buschinger 2009, Rabeling 2020). In hymenopteran social insects, social parasitism may take several forms (Lenoir et al. 2001, Brandt et al. 2005, Cervo 2006). In some of the most extreme cases, foreign queens enter an established colony of another species and dispatch the resident queen. The invading queen effectively takes over the colony and ultimately replaces the remaining resident workers with her own offspring.

The social wasp *Vespula squamosa*, commonly known as the southern yellowjacket, is known to be a facultative social parasite (Taylor 1939, MacDonald and Matthews 1975, 1984). *Vespula squamosa* (Figure 1A-C) is found throughout the eastern part of the United States, extending south to Honduras (Akre et al. 1980, Hunt et al. 2001). *Vespula squamosa* queens can found colonies independently under some circumstances. However, *V. squamosa* also acts as a social parasite in much of its range (Matthews 1982).

Vespula squamosa queens take over the active nests of other species (Taylor 1939, MacDonald and Matthews 1975, 1984, Greene 1991) and are known to parasitize other *Vespula* or *Vespa* taxa (Greene 1991). Typically, *V. squamosa* queens invade established nests and kill the resident queen. Then, the unknowing host workers help the *V. squamosa* queen rear new *V. squamosa* offspring. When the remaining host workers die, the entire colony comes to be inhabited by *V. squamosa* wasps.

In the southeastern part of the United States, a main host of *V. squamosa* is the eastern yellowjacket, *V. maculifrons* (Figure 1D-F; MacDonald and Matthews 1975, 1984). Previous studies have found that, in some locations, up to 40% of *V. maculifrons* colonies may be taken over by *V. squamosa* (MacDonald and Matthews 1981). Moreover, analysis of nest structure indicates that at least 80% of *V. squamosa* colonies originate by parasitism of *V. maculifrons* colonies in this region (MacDonald and Matthews 1975; Figure 1G). The remarkable intra- and interspecific social activities displayed by *V. squamosa* and *V. maculifrons* have made these taxa useful models for understanding advanced social behavior and evolution.

Investigations of genetic variation within and between species hold great promise in helping understand behavioral and social differences. Prior genetic studies in *V. squamosa* and *V. maculifrons* have determined that both taxa, like all *Vespula* species, are moderately polyandrous (Goodisman et al. 2007a, 2007b, Johnson et al. 2009, Loope et al. 2014, 2017). Investigations of long-lived *V. squamosa* nests have also uncovered changes in social and reproductive patterns associated with facultative recruitment of multiple queens (Wilson et al. 2009, Scarparo et al. 2021, Snyder and Loope 2021, Dyson et al. 2022, Sankovitz et al. 2023). Population genetic studies reveal that *Vespula* wasps show high gene flow within their native ranges (Hoffman et al. 2008, Dyson et al. 2021), suggesting that perennial polygyny may be an environmentally responsive trait (Scarparo et al. 2021, Dyson et al. 2022, Sankovitz et al. 2023).

Genomic studies are increasingly being used to understand the evolution of social behavior (Harpur et al. 2014, Rubin and Moreau 2016, Toth and Rehan 2017, Rubenstein et al. 2019, Imrit et al. 2020, Barkdull and Moreau 2023, Orr and Goodisman 2023, Wyatt et al. 2023, Mikhailova et al. 2024). Genomic investigations can provide insight into how life histories and behaviors evolve, as well as how social variation influences genome evolution, both within and between species (Rubenstein et al. 2019, Warners et al. 2019).

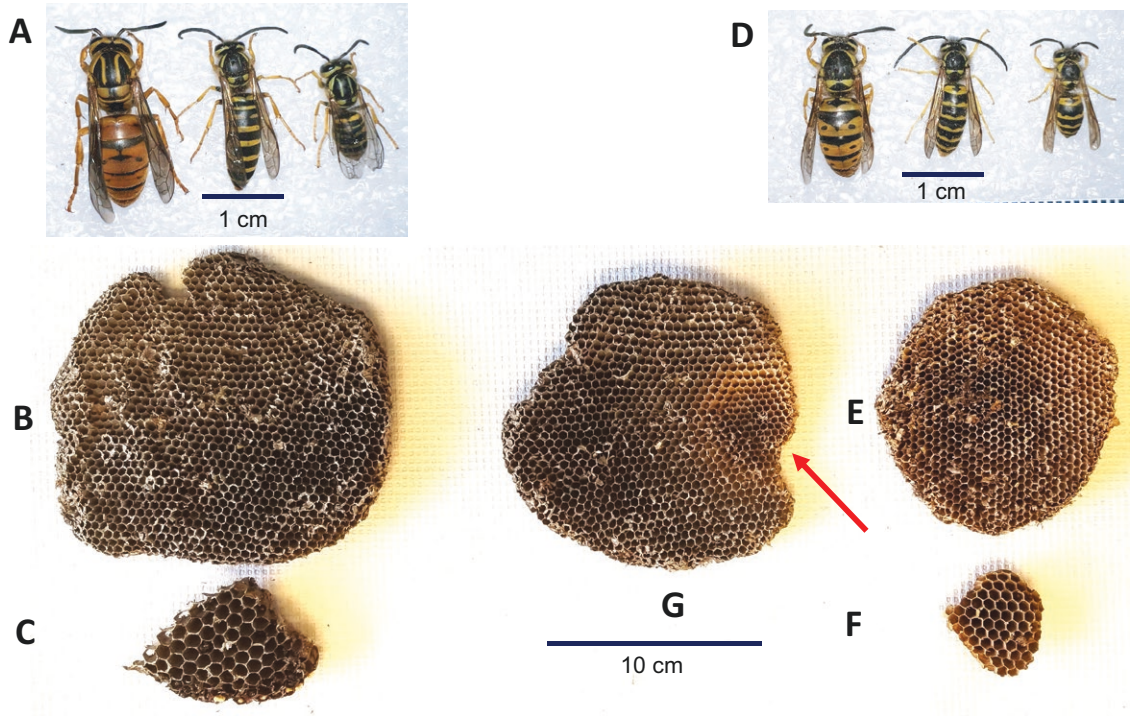

Figure 1. Southern and eastern yellowjacket wasps. (A) *Vespula squamosa* and (D) *Vespula maculifrons* queen, male, and worker castes displayed left to right. *V. squamosa* (B) worker comb and (C) queen comb. *V. maculifrons* (E) worker comb and (F) queen comb. (G) 'Hybrid' *V. squamosa-V. maculifrons* worker comb. Note the right sector of the hybrid comb, indicated by the arrow, was constructed by *V. maculifrons* workers whereas the remainder of the comb was constructed by *V. squamosa* workers after the nest had been parasitized by a *V. squamosa* queen.

For example, recent studies have investigated related *Polistes* and *Vespula* wasps and uncovered important insights related to molecular evolution, epigenetic inheritance, and behavior (Patalano et al. 2015, Standage et al. 2016, Bluher et al. 2020, Harrop et al. 2020, Miller et al. 2020, 2022, Shell et al. 2021, Crowley 2022, Miller and Sheehan 2023). Moreover, genomic analyses have recently been used to detect patterns of molecular evolution that correlate with social complexity (Dogantzis et al. 2018). Such genome-level investigations contribute to our understanding of the development of sociality in insects and provide evolutionary context for these social behaviors.

In this study, we sequenced the genomes of the facultative social parasite *V. squamosa* and its host *V. maculifrons* to further understand molecular and social evolution in social insects. We were specifically interested in determining if the genome of the social parasite *V. squamosa* provided clues to the ultimate and proximate mechanisms governing socially parasitic behaviors. We expected that a social parasite such as *V. squamosa*, which engages in both dependent and independent colony founding, would show stability in genome size, in contrast to other parasites which show genome reduction. We also expected *V. squamosa* to show molecular evolutionary changes associated with signal detection and production. Overall, our aim was to gain further insight into the structure, function, and evolution of social insect genomes in the context of evolutionary variation in social life histories.

Materials and Methods

Biological Samples

We collected live, mature *V. squamosa* and *V. maculifrons* colonies in and around Atlanta, Georgia, USA in November of 2021. Wasp colonies were stored in a 4°C walk-in refrigerator for several days to prevent flying and aggressive behavior. Then, we sorted live individuals into sterile microcentrifuge tubes before submerging them into liquid nitrogen to flash freeze and euthanize each wasp. RNA extractions were performed immediately following flash freezing. Samples were stored at −80 °C for subsequent genomic analysis.

We extracted genomic DNA from the head and thorax of a single, representative haploid male from each species. We used the Zymo Quick DNA HMW MagBead Kit to extract genomic DNA following the manufacturer's suggested protocol but adding 50 µl magbeads to increase DNA yield. We also included the optional RNase treatment steps. We performed all pipetting with wide-bore pipette tips to avoid shredding DNA. To verify our genomic DNA extractions were successful, we measured 260/280 and 260/230 ratios of each sample on a NanaDrop OneC (Thermo Fisher Scientific). All purified DNA was kept in a −80 °C freezer and transported on dry ice.

We extracted total RNA from six different samples from each of both species to study patterns of gene expression across life stages and castes, and aid in gene annotation. Specifically, we extracted RNA from pooled eggs, whole larvae, whole pupae, adult worker thoraces, adult male thoraces, and adult gyne thoraces in *V. squamosa and V. maculifrons*. We used the Zymo Quick RNA Kit to extract RNA according to the manufacturer's suggested protocol including the optional DNAse treatment. To confirm RNA extraction success, we measured the concentration and purity by estimating the 260/280 and 260/230 ratios on a Nanodrop OneC. All purified RNA was kept in a −80 °C freezer and transported on dry ice.

Genome and transcriptome sequencing

The University of Georgia Genomics and Bioinformatics Core (GGBC) first confirmed the DNA quality and created high-fidelity libraries. Next, the GGBC sequenced the genomic DNA using PacBio Sequel II Single Molecule Real-Time (SMRT) technology. The genomic DNA from each species was sequenced on an individual SMRT cell. PacBio HiFi reads were marked for duplicates using SMRT-Link v10.2 pbmarkdup.

RNA sample quality was confirmed using Qubit 4 Fluorometer analysis, and poly(A) enriched stranded RNA-seq libraries were prepared by the GGBC. The 12 samples were then sequenced by the GGBC on one NextSeq2000 P2 PE100 run with ~500 million read pairs, which provided approximately 40 million read pairs per sample. This deep sequencing provided ample data for gene models and predictions during bioinformatics analysis.

We produced genome assemblies using Flye v2.9 (Lin et al. 2016, Kolmogorov et al. 2019), with the parameters set to run two iterations and an expected size of 200 Mb for each genome. Benchmarking Universal Single Copy Orthologs (BUSCO) v4.0.6 was then used to determine the presence of highly conserved genes using Insecta odb10 and Hymenoptera odb10 databases (Simao et al. 2015). The Quality Assessment Tool for Genome Assemblies (QUAST) v5.0.2 (Gurevich et al. 2013) was used to determine the overall assembly statistics. MultiQC v1.11 (Ewels et al. 2016) was then employed to visualize assembly statistics from BUSCO and QUAST.

We used Trimmomatic v0.39 (Bolger et al. 2014) to remove adapter contamination with the following parameters: ILLUMINACLIP: TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36. FastQC v 0.11.9 was performed on raw and trimmed reads. Reports were visualized using the MultiQC tool (Ewels et al. 2016). HiSat2 v2.4.5 (Kim et al. 2019) was used to map the trimmed reads to each respective genome assembly (Supplementary Table S1). The bam files were sorted by SAMtools v1.16.1 (Li et al. 2009, Danecek et al. 2021) and assembled into transcriptome assemblies using Trinity v2.10.0 (Grabherr et al. 2011).

Feature annotation

We used BBMap v38.93 (Bushnell 2014) to split the genome assembly into manageable chunks for computational considerations. MAKER v3.01.03 (Campbell et al. 2014) was run for four iterations with the alternative splicing parameter set to zero (alt_splice = 0). Expressed sequence tag (EST) evidence was provided from the reference guided transcriptome assembly for the species of interest. Alternate EST evidence was obtained from the reference guided transcriptome assembly of the alternate species. Additionally, we included protein information from *V. germanica*, *V. pensylvanica*, and *V. vulgaris* (Harrop et al. 2020) to MAKER. In between each MAKER iteration, training sets were generated using SNAP v2013-11-29 (Li et al. 2007) and AUGUSTUS v3.4.0 (Stanke et al. 2006, 2008). We used BUSCO v4.0.6 to help determine the completeness of the gene models against the Insecta odb10 and Hymenoptera odb10 (Simao et al. 2015). The final round of gene models was filtered using an annotation edit distance (AED) cutoff of < 0.75. Functional annotation was carried out using the OmicsBox pipeline, which includes Blast2GO (Gotz et al. 2008) and InterProScan (Paysan-Lafosse et al. 2023; Supplementary Tables S2 to S5).

Intronic regions were identified using the tool 'Extract-intron-from-gff3'. Promoter sequences were defined by using the tool 'extract-promoter-sequences' with upstream (-u) set to 1.5 kb and the downstream (-d) set to 0. BEDtools v2.30.0 (Quinlan and Hall 2010) was used to sort the genomic feature format 3 (gff3) files prior to using Another GFF Analysis Toolkit v0.8.1 to extract sequence data. All additional genomic features were appended to the gff3 file by using BEDtools and BEDOPS v2.4.39 (Neph et al. 2012).

We identified transposable elements (TEs) within the genomes of 24 arthropod species including *V. squamosa* and *V. maculifrons* using the Extensive de novo TE Annotator (EDTA) v2.1.0 tool (Su et al. 2021). TEs can be broken down into the categories of fragmented and structurally intact and may also be nested within other TEs (Bourgeois and Boissinot 2019). The resulting TE content was categorized by all annotated TEs (anno), only intact TEs (intact), and reconstructed non-redundant TEs (TElib).

Full genome alignments

The online tool D-Genies v1.4.0 (Cabanettes and Klopp 2018) was used for pairwise comparisons of multiple genomes. The alignment for each pairwise comparison was performed using the Minimap2 v2.24 (Li 2018), which is integrated into D-Genies. The D-Genies repeatedness parameter was set to 'some repeats', which corresponds to average precision (-f 0.002) from Minimap2. The pafr tool (https://github.com/dwinter/pafr) was used to visualize the genome alignments.

DNA methylation

We investigated patterns of DNA methylation throughout the *V. squamosa* and *V. maculifrons* genomes. We used two approaches to understanding patterns of DNA methylation. First, we analyzed the prevalence of CpG dinucleotides in genomic features (Simola et al. 2013, Thomas et al. 2020). The CpG observed/expected ratio (CpG o/e) was calculated using Notos (Bulla et al. 2018) with a minimum length (-m) set to 200bp. The GpC o/e values were calculated as a control metric using a modified version of CpG.pl from cpg_tools (https://github.com/swebb1/cpg_tools). Genes less than 200 bp were excluded from downstream analysis. Evidence of past DNA methylation in genetic elements is inferred from CpG o/e values below the expected value of 1.0 for those elements (Yi and Goodisman 2009).

We also obtained evidence of DNA methylation by analyzing the PacBio genomic DNA reads themselves. Information on the methylation status of cytosine bases is derived from the kinetics of DNA polymerase as it adds bases to the extending sequencing strand during HiFi runs (Li et al. 2020). DNA 5-methylcytosine (5mCpG) calls were determined by PacBio primrose tool, which is a component of SMRT-Link v11.1.0 (https://github.com/mattoslmp/primrose). The pbmm2 tool was used to align reads to the reference and the pb-CpG-tool was run with the count pileup mode to determine per-site methylation statistics. The presence of 5mCpG was then determined via the PacBio primrose tool, which has an 85% read-level accuracy (Ni et al. 2023) and a >90% correlation with bisulfite sequencing data (Tse et al. 2021). Visualization of the methylation coverage was performed using SeqMonk v1.48.0 with a probe size of 200 bp (Supplementary Tables S6 and S7).

Gene expression

Trimmed short reads from the samples used in the reference guided transcriptome assemblies were mapped to the respective genomes using RNA-Seq by Expectation-Maximization (RSEM) v1.3.3 (Li and Dewey 2011) with Spliced Transcripts Aligned to Reference (STAR) v2.7.10 (Dobin et al. 2012; Supplementary Table S8). Distance matrices and a metric of specificity of expression among sample types, tau (Yanai et al. 2005; Supplementary Tables S9 and S10), were computed using the normalized transcript per kilobase million (TPM) gene counts (Supplementary Tables S11 and S12). We examined differential gene expression between samples based on the TPM ratios. For the purposes of heatmap visualization using complex heatmap tool (Gu et al. 2016), the TPM reports (Supplementary Tables S11 and S12) were filtered to exclude occurrences in which all samples had zero counts for a given gene.

Ortholog determination and tests of selection

Orthovenn3 (Sun et al. 2023) was used to determine the gene orthologs across the five *Vespula* species with the e-value cutoff of $1e^{-5}$ (Supplementary Table S13). Additional ortholog determination was conducted using OrthoFinder2 v2.5.2 (Emms and Kelly 2019) across all Arthropoda, *Polistes*, *Vespa*, and *Vespula* species (Supplementary Tables S14 and S15). The phylogenetic tree was run through multiple iterations using IQ-TREE v2.2.2 (Minh et al. 2020) and visualized using FigTree v1.4.4 (Rambaut 2010). The tree was pruned for downstream visualization using Analysis of Phylogenetics and Evolution (APE; Popescu et al. 2012). We then used Orthologr v0.4.2 (Drost et al. 2015), which runs a Double Index Alignment Of Next-generation sequencing Data (DIAMOND; Buchfink et al. 2021) reciprocal best hit (RBH), amino acid alignment using Needleman-Wunsch, pal2nal (Suyama et al. 2006) for codon alignments.

We used Comeron's method for dN/dS inference for the pairwise comparison of *V. squamosa* and *V. maculifrons* (Supplementary Table S16). Evidence of positive selection was sought using (i) the five *Vespula* species single copy orthogroups and (ii) *Polistes*, *Vespa*, and *V. squamosa* and *V. maculifrons* single copy orthogroups determined from OrthoFinder2. The nucleotide sequences of the single copy orthogroups were aligned with multiple alignment using fast Fourier transformation (MAFFT) v7.487 (Katoh et al. 2002). Alignment scores were checked with the AliStat tool (https://github.com/thomaskf/AliStat) in both cases (Supplementary Tables S17 and S18). We used the HyPhy v2.5.15 (Pond et al. 2020) aBSREL tool to test for branch specific positive selection of the aligned sequences (Supplementary Table S19). HyPhy RELAX (Wertheim et al. 2015) was then used to test for relaxed and intensified selection; tests were applied to the gene trees of single copy orthologs of the five *Vespula* species, with each *Vespula* species terminal branch designated the 'test' branch when compared to the other four *Vespula* set as 'reference' branches (Supplementary Table S20). Additional HyPhy aBSREL and RELAX tests were run on *V. squamosa* and *V. maculifrons* 'test' branches against 'reference' branches of *Polistes* and *Vespa* without other *Vespula* taxa included (Supplementary Tables S21 and S22). Gene ontology (GO) term enrichment of genes subject to distinct lineage-specific selective pressures was calculated using rrvgo for tests including only *Vespula* species; results were visualized using the Revigo online tool, with single copy orthologs used as the background set of genes (Supek et al. 2011, Sayols 2023).

Results

Genome structure

We sequenced the genomes of the social wasps *V. squamosa* and *V. maculifrons*. We found that the genome sizes were 196.0 and 193.3 Mb, respectively. The N50 metric, which serves as a guide for assessing contiguity of genome assembly, was 6.2 Mbp for *V. squamosa* and 4.9 Mbp for *V. maculifrons*. We note that the genome assembly sizes of *V. squamosa* and *V. maculifrons* were marginally larger than those of related *Vespula* species (Figure 2). These differences in assembly sizes may be partially attributed to the different sequencing technologies employed. For instance, integration of Hi-C data, as utilized by Harrop et al. (2020), can enhance the scaffolding of genome assemblies, providing chromosome-level resolution and further mitigating ambiguities caused by repetitive regions.

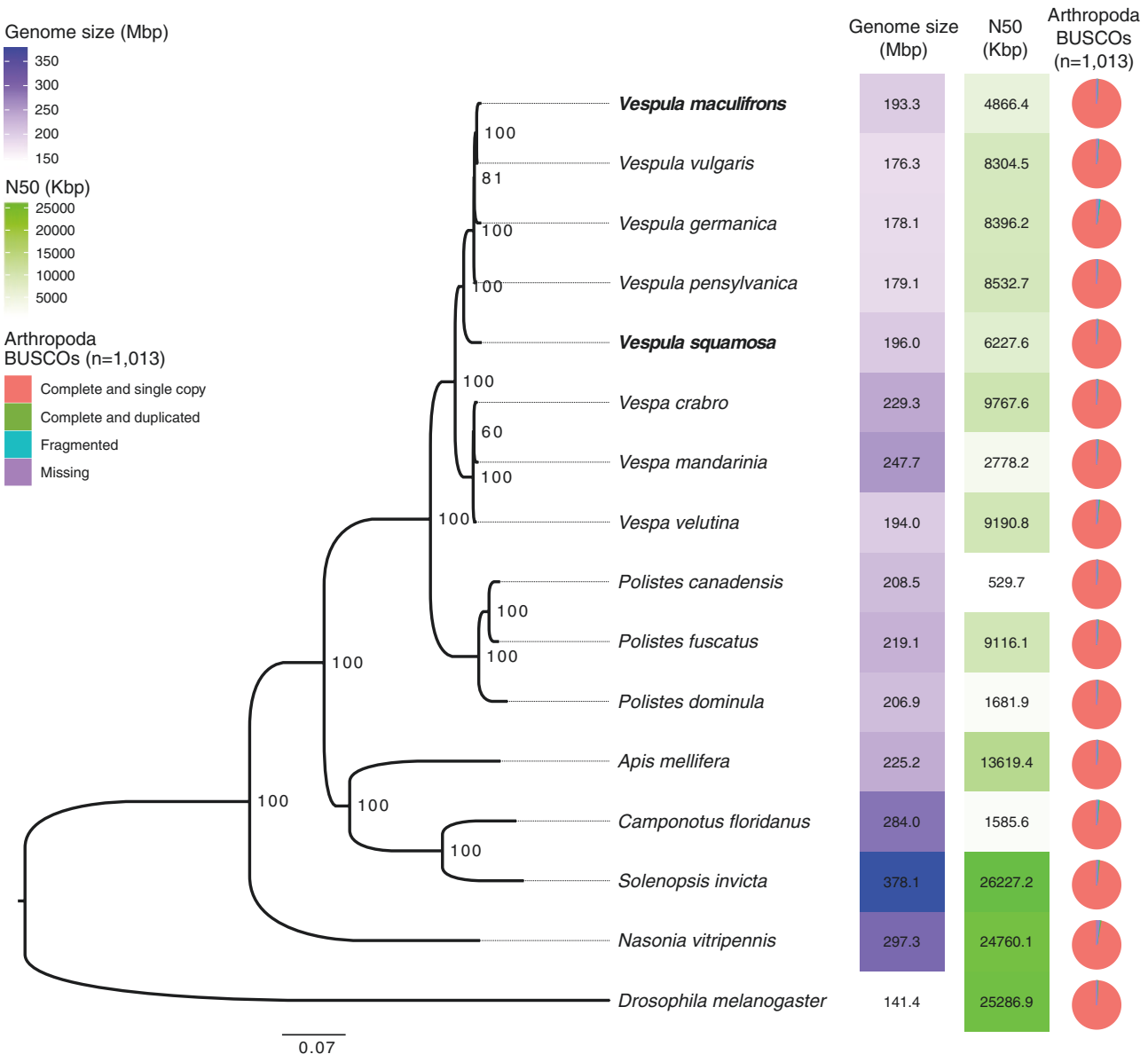

Figure 2. Phylogenetic placement of *V. squamosa* and *V. maculifrons* wasps. Phylogenetic tree of 11 Vespids, 3 outgroup hymenopterans, and a dipteran showing genome size, N50, and Arthropoda BUSCO scores.

BUSCO scores (Simao et al. 2015) describe the completeness and duplication of orthologous genes (Supplementary Table S23). Both *V. squamosa* and *V. maculifrons* exhibited high completeness scores with 933 and 938 complete and single-copy BUSCOs, respectively. Moreover, *V. squamosa* had 17 complete and duplicated BUSCOs, while *V. maculifrons* had 22. We also characterized TEs for each species (Supplementary Figures S1, Supplementary Tables S24, and S25). The *V. squamosa* genome consisted of 8.00% interspersed repeats while the *V. maculifrons* genome contained 7.70% interspersed repeats (Supplementary Table S25). Similar proportions of TEs were found to overlap coding sequences in the genome of each species (Supplementary Figure S1).

Phylogenetics and comparative genomics

We constructed a phylogenetic tree including *V. squamosa*, *V. maculifrons*, and related social wasps and outgroup insects (Figure 2). The tree was constructed based on sets of single copy orthogroups across 16 Insecta species. The phylogenetic tree recovers the monophyly of *Vespula*, specifically, and the Vespidae, in general.

Analysis of genomic synteny (collinearity) within the *Vespula* genus revealed that *V. squamosa* and *V. maculifrons* largely parallel the genomic organization of *V. vulgaris*, *V. germanica*, and *V. pensylvanica* (Supplementary Figure S2A). Further, global alignment of the *V. squamosa* and *V. maculifrons* genome assemblies showed that a few regions of the *Vespula* genomes may have been subject to inversions or translocations following speciation (Supplementary Figure S2B). In particular, we identified 49 inversions out of the 104 collinear regions between *V. maculifrons* and *V. squamosa*, 32 inversions out of the 86 conserved regions within *V. maculifrons*, and 29 inversions out of the 65 conserved regions within *V. squamosa* (Supplementary Table S26).

DNA methylation

Putative patterns of DNA methylation were analyzed using CpG o/e (observed/expected CpG ratio) metrics in 24 Arthropoda species

(Yi and Goodisman 2009). The distribution of CpG o/e values for genetic elements indicated variation in DNA methylation within and between species (Walsh et al. 2010; Supplementary Figure S3). *V. squamosa* and *V. maculifrons* exons, in particular, showed broad distributions, consistent with DNA methylation. These patterns were similar to distributions of CpG o/e in other *Vespula* species. Thus, computational analysis of nucleotide content suggested that DNA methylation had been present historically in the *Vespula* lineage.

Actual DNA methylation patterns of *V. squamosa* and *V. maculifrons* were determined by analyzing the output of the PacBio sequencing reads (Figure 3, Supplementary Figure S4). A putative baseline level of methylation of ~15% throughout the genome of both taxa was observed (Supplementary Figure S4A). Given known patterns and levels of DNA methylation in other insect taxa, this baseline likely represents false positive methylation calls, whereas values above this baseline likely represent actually methylated DNA. DNA methylation patterns in *V. squamosa* and *V. maculifrons* across coding sequences (Figure 3A) showed low levels of methylation upstream of the translation start site (TSS) and downstream of the translation termination site (TTS). However, methylation was substantially higher within the coding sequences themselves. Given the baseline levels of methylation observed in the two taxa, methylation in *V. squamosa* may be slightly higher than in *V. maculifrons*. Moreover, exons showed substantially higher levels of DNA methylation than introns. Methylation levels were substantially higher in the more central exons and introns of the gene (Figure 3B).

We also found a strong, but nonlinear, relationship between gene expression variability among sample types (tau) and the level of DNA methylation of genes in both *V. squamosa* and *V. maculifrons* (Figure 4A and B). We uncovered the inverse pattern when comparing the relationship between gene expression levels among sample types (TPM median) and levels of DNA methylation (Figure 4C and D). In particular, approximately 2/3 of the genes displayed relatively low methylation, low gene expression, and high variance in gene expression among castes and developmental stages. The other 1/3 of the genes displayed relatively high levels of methylation, high gene expression, and low variance in expression among castes and developmental stages.

Gene expression

We sequenced the transcriptomes of *V. squamosa* and *V. maculifrons* adult gyne, worker, and male thoraces, whole larvae and pupae, and pooled whole eggs. The transcriptomes aided in developing gene models to interpret the sequenced genomes and provided initial information about the patterns of gene expression in these taxa. Notably, our gene expression results were limited by sample size and should be interpreted as preliminary data for robust RNA-seq studies in the future.

Samples generally grouped by developmental stage. In both species, the adult gyne, male, and worker transcriptomes were more similar to each other than to the other immature developmental stages (Figure 5A and B). In *V. squamosa*, the egg, larva, and pupa grouped closely, while in *V. maculifrons* the larval sample was slightly more like the grade of adult samples. Notably, however, the patterns of gene expression of adult workers were more similar to those in adult males than they were to adult gynes.

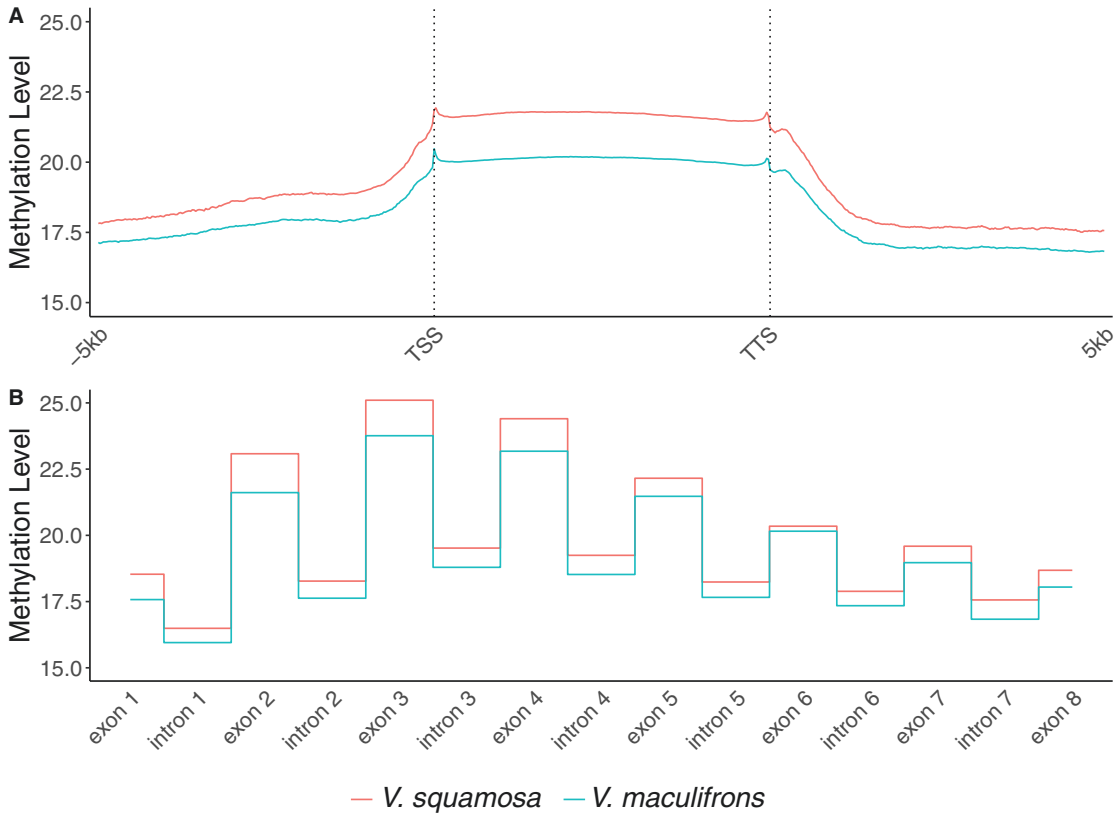

Figure 3. DNA methylation level subdivided by genic features. Average DNA methylation levels (normalized to transcript length) are plotted across (A) coding sequences (translation start site, TSS; translation termination site; TTS) and for (B) exons and introns of genes in *V. squamosa* and *V. maculifrons* wasps.

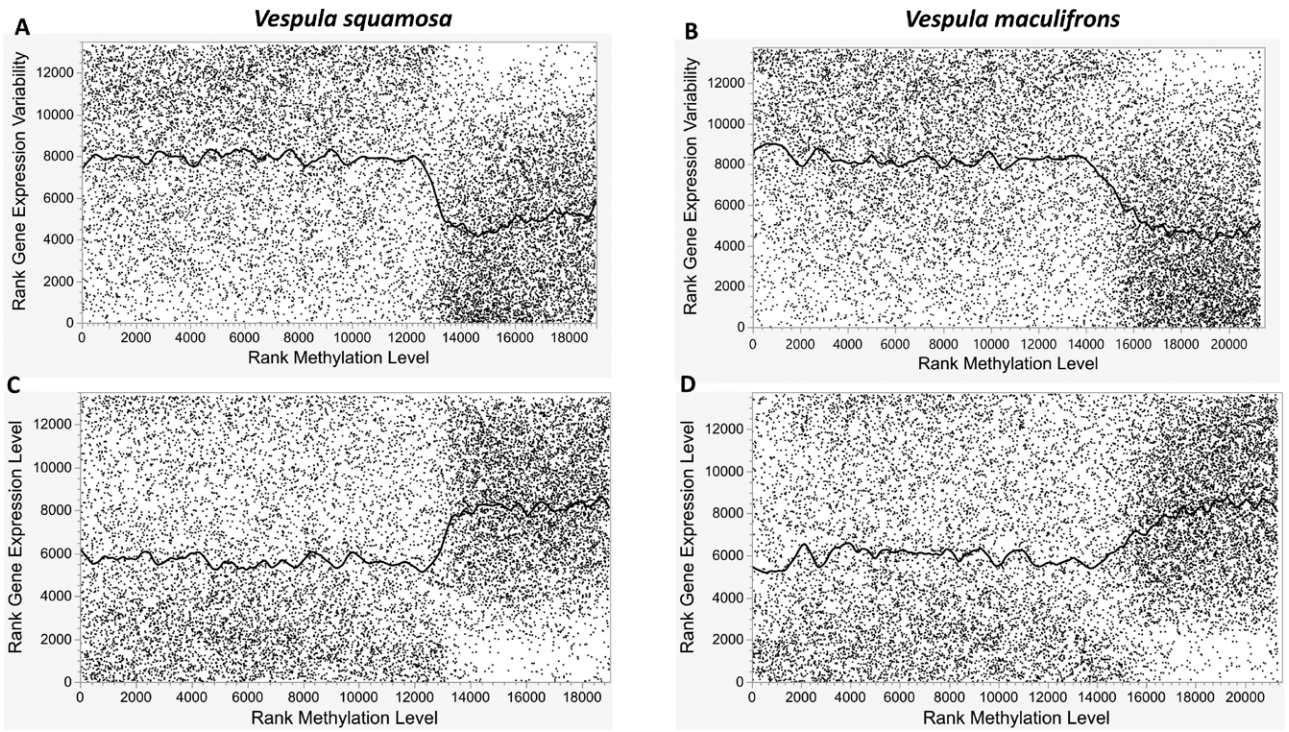

Figure 4. Relationships between DNA methylation and gene expression in *V. squamosa* and *V. maculifrons* wasps. Rank gene expression variability vs rank DNA methylation level in (A) *V. squamosa* genes and (B) *V. maculifrons genes*. Rank gene expression level vs rank methylation level in (C) *V. squamosa* genes and (D) *V. maculifrons genes*. Smoothing curves are displayed to demonstrate the bimodal relationships between methylation and gene expression.

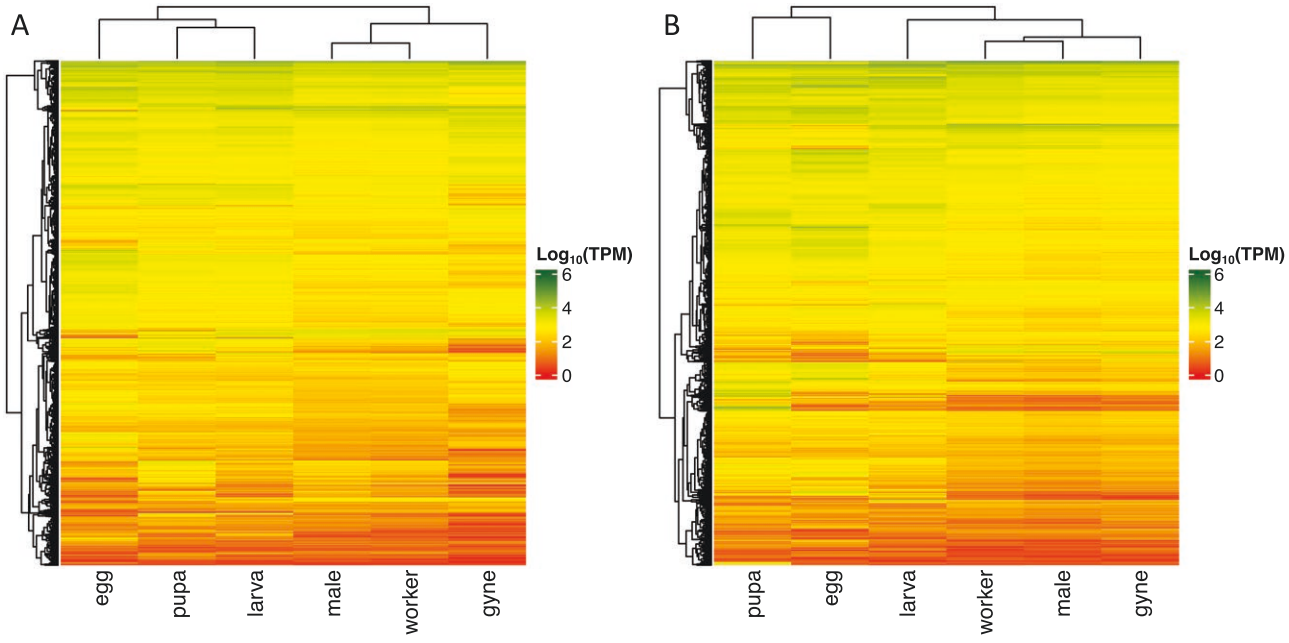

Figure 5. Caste and stage-specific gene expression. Gene expression levels, transcripts per kilobase million (TPM), within each caste of (A) *V. squamosa* ($n = 9546$ genes) and (B) V. *maculifrons* ($n = 9880$ genes) wasps.

We also examined highly differentially expressed genes between our samples in *V. squamosa* (Supplementary Table S27) and *V. maculifrons* (Supplementary Table S28). First, we calculated the mean gene expression level (TPM) for immature life stages (egg, pupa, and larva) and adult life stages (worker, gyne, and male). We found that immature life stages in *V. squamosa* highly expressed genes like *Siwi*, an endoribonuclease. Further, *V. squamosa* adults had high levels of *chitin deacetylase* expression compared to immature life stages. We also found that *V. squamosa* workers exhibited higher expression of *defensin*, an antimicrobial protein, compared to males. In *V. maculifrons*, immature life stages expressed high levels of *fatty acid synthase* compared to adults. We also observed an overexpression of *cytochrome P450 4C1*, an enzyme which plays a role in energy substrate mobilization, in *V. maculifrons* gynes compared to workers.

Table 1. Signatures of selection on terminal gene tree branches among 4,031 *Vespula* single copy orthologs, demonstrating high selection ratio in *V. squamosa*

Test branch	Positive selection *	Relaxation †	Intensification ‡	Selection ratio ₮
V. squamosa	561	572	280	2.04
V. maculifrons	369	216	260	0.83
V. pensylvanica	313	329	325	1.01
V. germanica	491	295	403	0.73
V. vulgaris	473	159	316	0.50

*aBSREL (Branches Under Positive Selection, $P < 0.05$).
†RELAX (K results, $P < 0.05$, Relaxation).
‡RELAX (K results, $P < 0.05$, Intensification).
₮Relaxation/Intensification.

Molecular evolution and selection

We identified genes under positive selection, relaxed selection, and selection intensification in the terminal *V. squamosa* and *V. maculifrons* gene tree branches, relative to four other *Vespula* species (Table 1). Our analysis revealed that *V. squamosa* is experiencing more than twice as much relaxation of selection relative to intensification when compared to any other *Vespula* taxon. However, we note that *V. squamosa* is an outgroup to the other *Vespula* species and thus its terminal gene tree branches are generally longer than for the other taxa, encompassing a greater swath of evolutionary history (Table 1, Figure 2).

In order to provide an alternative view of lineage-specific selection with an equal branch length for *V. squamosa* and *V. maculifrons*, we again assessed lineage-specific selection for *V. squamosa* and *V. maculifrons* in gene trees constructed from single copy orthologs among only these two *Vespula*, three species of *Vespa*, and three species of *Polistes* (Supplementary Figure S5). Among this reduced set of single copy orthologs, we detected 19% more instances of lineage-specific relaxation in *V. squamosa* as compared to *V. maculifrons* (157 and 132 genes, respectively) and 12% fewer instances of lineage-specific intensification of selection in *V. squamosa* relative to *V. maculifrons* (200 and 227 genes, respectively; Supplementary Table S29). Thus, this more conservative approach still detected an increase in relaxation of selection in *V. squamosa* relative to *V. maculifrons*. We further found that the *V. squamosa* genes experiencing relaxed selection were enriched for functions associated with catabolism, autophagy, and intracellular transport (Figure 6C).

The ratio of non-synonymous to synonymous substitution rates (d*N*/d*S*) provides information about the strength and mode of selection pressures in genomes (Jeffares et al. 2015). We found that in both *V. squamosa* and *V. maculifrons*, dN/dS was negatively correlated with overall magnitude of gene expression across castes (TPMmedian; Table 2). In addition, d*N*/d*S* was positively correlated with the specificity of gene expression (tau) across castes and developmental stages in both taxa. Notably, the correlations were roughly similar in magnitude in *V. squamosa* and *V. maculifrons*.

Discussion

We sequenced and analyzed the genomes of two important social wasps, *V. squamosa* and *V. maculifrons* (Jandt and Toth 2015, Sumner et al. 2023). These species are of particular interest because *V. squamosa* is a facultative social parasite of *V. maculifrons* (MacDonald and Matthews 1975, 1984). We were interested in understanding if aspects of the genomes of these two taxa might provide insight into the structure of their societies and into their social behaviors.

Genome structure of *V. squamosa* and *V. maculifrons*

The genome sizes of *V. squamosa* and *V. maculifrons* were similar to the genome sizes of other *Vespula* species (Figure 2), which range from 175 to 200 Mbp. More generally, the genome sizes of *Vespula* and *Polistes* wasps fall roughly within this range. These genera appear to have undergone a genome size reduction relative to more distantly related taxa in the Apidae and Formicidae (Patalano et al. 2015, Standage et al. 2016, Harrop et al. 2020, Crowley 2022, Miller et al. 2022).

We found that the phylogenetic placement of *V. squamosa* and *V. maculifrons* conformed to expectations based on prior studies (Lopez-Osorio et al. 2014). That is, *V. squamosa* and *V. maculifrons* are relatively distantly related within the *Vespula*. Although the two species share superficial morphological similarities, such as the characteristic yellow and black coloration of individuals in this genus, they are quite distinct in phenotype otherwise (Figure 1).

Tranposable elements (TEs) accumulate in the genomes of organisms and influence genetic diversity and adaptation (Friedli and Trono 2015, Harrison et al. 2018, Schrader and Schmitz 2019, Gilbert et al. 2021, Berger et al. 2022). TEs are thought to be fixed predominantly by genetic drift rather than natural selection (Charlesworth and Langley 2021). Consequently, if *V. squamosa*, the social parasite, has a smaller effective population size than *V. maculifrons*, then the genomes of the two taxa might display different levels of TEs. However, *V. squamosa* was found to have 8.00% of its sequence comprised of interspersed repeats while *V. maculifrons* displayed only a marginally lower value of 7.70% (Supplementary Table S25). Thus, these data do not support the hypothesis that the effective population sizes of these two taxa differ dramatically. These low levels of TEs also stand in contrast to the high TE load observed in insects with larger genome sizes, including termites, which have approximately half of their genomes comprised of interspersed repetitive elements (Harrison et al. 2018).

DNA methylation

This study is one of the first to use direct sequencing to analyze patterns of DNA methylation in an invertebrate. PacBio HiFi DNA sequencing reads can be used to identify patterns of DNA methylation because the kinetics of DNA polymerase differs if a methylated or an unmethylated cytosine is incorporated in the sequencing process (Tse et al. 2021, Ni et al. 2023, Nanda et al. 2024, Sigurpalsdottir et al. 2024). Our results show that methylation is present in both *V. squamosa* and *V. maculifrons*. Notably, however, we detected high levels of inferred false-positive DNA methylation (~15%) within the *Vespula* genomes. Therefore, analysis of PacBio HiFi sequencing reads does not detect precise levels of DNA methylation. However, PacBio HiFi sequencing reads are still informative for gaining general information about patterns of DNA methylation across different genetic elements in the genome.

Analysis of patterns of DNA methylation within *V. squamosa* and *V. maculifrons* demonstrated that coding regions of the genomes are enriched for methylated cytosines (Figure 3). DNA methylation levels were relatively low in the most 5' and 3' exons and introns, but methylation reached substantial levels in the middle exons and introns. Moreover, our study found minimal methylation in TEs, as is typical for holometabolous insects (Lewis et al. 2020; Supplementary

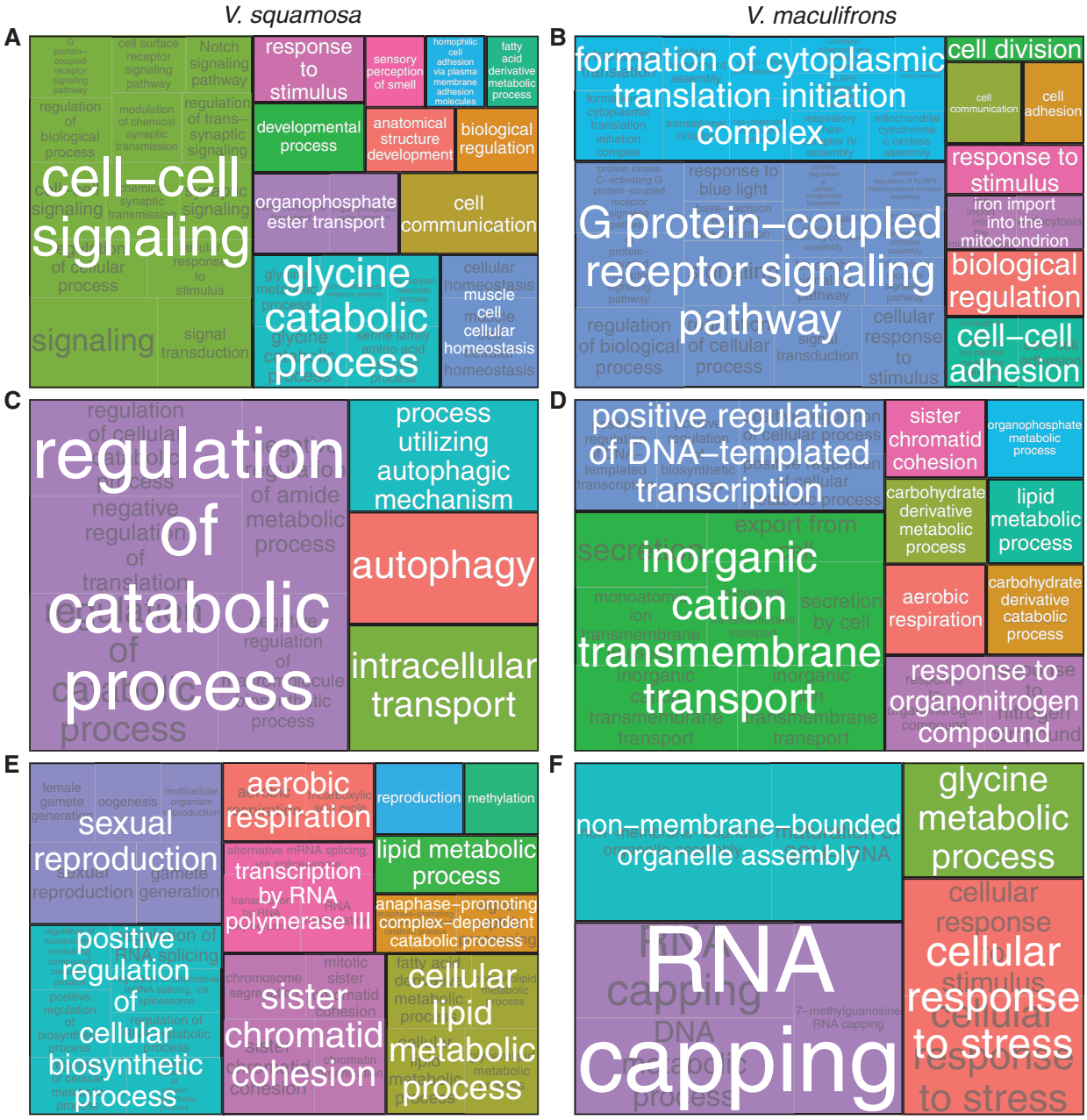

Figure 6. Ontology (GO) term enrichment tree map based on biological processes for genes under (A and B) positive selection, (C and D) relaxation, and (E and F) intensification in *V. squamosa* and *V. maculifrons* within *Vespula*, respectively. Tree map clusters are comprised of connected GO terms to form superclusters. The size of a cluster reflects the $\log_{10}$ transformed p-values from GO term enrichment. Boxes of a larger size represent GO terms with higher significance levels.

Figure S4). Overall, these distributions of DNA methylation have been identified previously in insect genomes and seem fundamental to DNA methylation patterns of the Holometabola (Glastad et al. 2013, 2014a, 2015, 2016, Hunt et al. 2013c, 2013b, Rehan et al. 2016, Rahman and Lozier 2023).

Our analysis of DNA methylation from the PacBio HiFi sequencing reads uncovered relationships between DNA methylation and gene expression. Interestingly, the relationship between gene expression and DNA methylation is not linear or, apparently, even continuous (Figure 4). Instead, genes seem to belong to two distinct classes. The first class of genes shows low methylation, low overall gene expression, high variation in gene expression among phenotypes, and low constraint on gene evolution. The second class of genes shows high methylation, high overall gene expression, low variation in gene expression among phenotypes, and high constraint on gene evolution. This two-class relationship among genes has been found across many insect taxa (Hunt et al. 2013c, Glastad et al. 2014a, 2014b, 2015, Jeong et al. 2018). In general, DNA methylation within genes is found in conserved, housekeeping genes and is correlated with high gene expression and uniform expression across tissues and phenotypes. Thus, the bimodal distribution of gene function in insects seems to be a core aspect of insect genomes.

It is notable that *Vespula* species apparently possess a functional DNA methylation system (Harrop et al. 2020) whereas related *Polistes* species, which have lost DNA Methyltransferase 3 (DNMT3), do not (Weiner et al. 2013, Patalano et al. 2015,

Table 2. Spearman's rank correlation (ρ) for metrics of DNA methylation, gene expression, and molecular evolution for *Vespula* single copy orthologs

		V. squamosa ρ (*n*)	*V. maculifrons* ρ (*n*)
Methylation level	Expression specificity	−0.26*** (3985)	−0.40*** (4015)
Methylation level	Expression level	0.23*** (3985)	0.26*** (4015)
dN/dS[a]	Expression specificity	0.20*** (3986)	0.10*** (4017)
dN/dS[a]	Expression level	−0.10*** (3986)	−0.04* (4017)
dN/dS[a]	Methylation level	−0.24*** (4030)	−0.10*** (4029)

[a]dN/dS values are for terminal species branch in gene trees as calculated by HyPhy RELAX.
*$P < 0.01$, ***$P < 0.0001$.

Standage et al. 2016, Miller et al. 2022). Indeed, the distribution of DNA methylation across insects is highly variable (Bewick et al. 2017, Glastad et al. 2017, Thomas et al. 2020). Holometabolous insects tend to show lower levels of DNA methylation compared to hemimetabolous insects and other basal arthropods (Glastad et al. 2016). Evidence that DNA methylation causes changes in gene expression in insects is controversial and the function of DNA methylation in insects continues to be highly debated.

Patterns of gene expression

We examined patterns of gene expression among different castes and development stages in both *V. squamosa* and *V. maculifrons* (Orr and Goodisman 2023). We note that our ability to assess differential gene expression was limited due to sequencing single replicates. Thus, these findings should be viewed as preliminary and inspire more robust RNA-seq studies exploring gene expression differences between social wasp life stages, castes, and species in the future.

We found that adult male and adult worker gene expression patterns were relatively similar in both *V. squamosa* and *V. maculifrons* (Figure 5). Notably, our results differ from previous findings that *V. squamosa* adult workers and gynes have more similar gene expression patterns than males and workers (Hoffman and Goodisman 2007). However, the age and tissues sampled of individual wasps differed between these studies and such variation is expected to impact patterns of gene expression (Ferreira et al. 2013, Toth et al. 2014, Berens et al. 2015, Jones et al. 2017, Favreau et al. 2023a, b, Taylor et al. 2023). Our study revealed that developmental stage had a strong effect on differential gene expression among samples, as observed previously (Hoffman and Goodisman 2007, Morandin et al. 2015, Ingram et al. 2016, Lucas et al. 2017). Indeed, many important growth and development pathways are required for successful ontogenetic transitions between egg, larva, pupa, and adult stages in insects (Nijhout and McKenna 2018).

We also examined specific differentially expressed genes between immature life stages (egg, pupa, and larva) and adult life stages (worker, gyne, and male thoraces). Generally, we found that immature life stages had elevated expression of genes involved in developmental pathways. For example, the genes *Siwi* and *fatty acid synthase* were elevated in immature life stages in *V. squamosa* and *V. maculifrons*, respectively. Siwi proteins play an essential role in the piRNA pathways of insects and have been associated with RNA transport (Santos et al. 2023). Further, fatty acid synthase has been shown to be essential in metamorphosis of insects (Song et al. 2022). Likewise, a previous study revealed metabolism gene pathways were upregulated in *V. squamosa* larvae (Hoffman and Goodisman 2007). Thus, we generally find that many genes in developmental pathways are highly expressed in egg, pupa, and larva of both *Vespula* species studied.

Previous studies have investigated caste-specific gene expression in social wasps in order to better understand how gene function correlates with caste differences (Hunt and Goodisman 2010, Toth and Robinson 2010, Jandt and Toth 2015). For example, an early analysis of patterns of gene expression in *V. squamosa* and *V. maculifrons* revealed caste-specific expression in many genes including vitellogenin, an important egg yolk protein highly expressed in the queen caste (Hunt and Goodisman 2010). Moreover, metabolism-related genes were upregulated in queen-destined *V. squamosa* larvae (Hoffman and Goodisman 2007). Wyatt et al. (2023) examined caste-biased expression in a diversity of social wasps, including *Vespa* and *Vespula* species, and found some evidence for conserved caste-biased expression patterns across species. More recent studies in *Vespa* have also uncovered differentially expressed genes between queen and worker castes (Favreau et al. 2023). Several studies in the Polistinae, which tend to have simpler societies than the Vespinae, have provided further insight into the nature of caste-biased expression in social wasps. For example, strong expression differences of immunity and metabolism-related genes were uncovered between worker and gyne castes in *Polistes dominula* (Geffre et al. 2017). Other studies in *Polistes* have identified highly variable proportions of genes differentially expressed between behavioral groups or castes (Toth et al. 2010, Ferreira et al. 2013, Patalano et al. 2015). Moreover, analyses of expression in different tissues demonstrated important seasonal and physiological effects on gene expression patterns (Toth et al. 2014, Patalano et al. 2022). Future studies should examine gene expression differences in social wasp castes and life stages more robustly to better understand how gene function varies within and between social wasp species (Berens et al. 2015).

Molecular evolution in *Vespula*

We investigated the strength of natural selection operating on *V. squamosa* and *V. maculifrons* genes, by measuring the ratio of nonsynonymous to synonymous substitutions, d*N*/d*S*. Our analyses revealed that d*N*/d*S* was positively correlated with tau, the measure of variation across gene expression libraries, and negatively correlated with TPMmedian, a metric of overall gene expression level, in both species (Table 2). These correlations demonstrate that highly expressed genes are generally under stronger selective constraint than genes expressed at lower levels. This relationship arises because more highly and uniformly expressed genes are subject to stronger levels of natural selection than genes expressed at lower and variable levels. These correlations have been identified previously and represent near universal relationships between patterns of gene expression and gene evolution (Duret and Mouchiroud 2000, Pal et al. 2001, Gout et al. 2010, Hunt et al. 2013a, Zhang and Yang 2015, Kapheim et al. 2020, Bertorelle et al. 2022, Tosto et al. 2023, Mikhailova et al. 2024).

We also investigated if patterns of genome evolution differed between *V. squamosa* and *V. maculifrons*. Rates of molecular evolution have been found to correlate with social complexity in social insect taxa (Toth and Rehan 2017, Dogantzis et al. 2018, Imrit et al. 2020, Rubin 2022, Mikhailova et al. 2024). Thus, we predicted that the genome of *V. squamosa* might show the evolutionary signature of its unusual social behavior (Kilner and Langmore 2011, Arthur and Romiguier 2021, Schrader et al. 2021, Bousjein et al. 2022).

We found that *V. squamosa,* the facultative social parasite, had the highest number of genes under relaxed selection among the species analyzed. *V. squamosa* also displayed a strikingly higher ratio of genes under relaxed selection relative to genes under intensified selection (Table 1). This genome-wide pattern has been observed in other social parasites previously (Schrader et al. 2021). Relaxed selection could be related to decreased selective pressure operating on processes crucial for free-living ancestors and relatives. Additionally, social parasites often have smaller effective population sizes, which could limit the efficacy of selection (Papkou et al. 2016, Schrader et al. 2017, Weyna and Romiguier 2021). However, we note that we do not have evidence of differences in effective population sizes between *V. squamosa* and *V. maculifrons*.

The genes identified as experiencing relaxed selection in *V. squamosa* were enriched for biological processes related to catabolism and autophagy (Figure 6C, Supplementary Table S30), which could be indicative of relaxed selection pressure during independent colony founding. In several ant species, queen fat reserve content has been associated with colony founding strategy, such that larger queens with greater fat reserves are more likely to participate in independent colony founding, whereas smaller queens with fewer fat reserves tend to participate in dependent colony founding (Keller and Passera 1989, Howard 2006). Therefore, relaxed selection on catabolism and autophagy genes in *V. squamosa* could be partially influenced by facultative social parasitism and facultative polygyny. Thus, *V. squamosa* life history could shape its genome and have important evolutionary consequences.

Genes under intensified selection in *V. squamosa* were enriched for reproductive processes such as gamete generation. This may indicate a higher selective emphasis on reproduction rather than more basic processes, which are experiencing intensified selective pressure in *V. maculifrons* (Figure 6E and F, Supplementary Table S30). The results could also arise from the alternative life-history strategies of social parasitism and polygyny, both exploited by *V. squamosa*. The possibility that selection is targeting reproduction in *V. squamosa* could signal the evolution of even more extreme parasitic behaviors (Lhomme and Hines 2018, Rabeling 2020). Overall, this work suggests that the facultative social parasite *V. squamosa* is experiencing different selective pressures than its frequent host, *V. maculifrons*. This research thus deepens our knowledge of host-parasite evolutionary dynamics and expands our understanding of social evolution.

Funding

This work was funded by National Science Foundation grants DEB: 2105033 and IOS: 2019799.

Author contributions

Michael Catto (Data curation [Lead], Formal analysis [Lead], Investigation [Equal], Methodology [Lead], Software [Lead], Validation [Lead], Visualization [Lead], Writing—original draft [Equal], Writing—review & editing [Equal]), Paige Caine (Investigation [Supporting], Methodology [Supporting], Writing—original draft [Supporting], Writing—review & editing [Supporting]), Sarah Orr (Investigation [Supporting], Methodology [Supporting], Writing—original draft [Supporting], Writing—review & editing [Supporting]), Brendan Hunt (Conceptualization [Supporting], Project administration [Supporting], Resources [Supporting], Supervision [Lead], Writing—original draft [Equal], Writing—review & editing [Equal]), and Michael Goodisman (Conceptualization [Lead], Funding acquisition [Lead], Project administration [Lead], Resources [Lead], Supervision [Equal], Visualization [Supporting], Writing—original draft [Equal], Writing—review & editing [Equal])

Supplementary Data

Supplemental information, software, and data files can be found in the Dryad repository: https://doi.org/10.5061/dryad.w6m905qxj.

Data Availability

All genomic and transcriptomic data associated with this project has been deposited to NCBI under the BioProject PRJNA924791. The raw reads can be found in the SRA under the accessions SRR24983457-SRR24983470. The assembled genomes accessions are JAUDFV000000000 and JAYRBN000000000, which includes the corresponding genes models. The transcriptome assemblies' accessions are GKNK00000000 and GKNB00000000.

References

Akre RD, MacDonald JF. 1986. 'Biology, economic importance and control of yellow jackets'. In: Vinson SB, editor. *Economic Impact and Control of Social Insects*. New York: Praeger; p. 353–412.

Akre RD, Greene A, MacDonald JF, et al. 1980. The yellow jackets of America north of Mexico. Washington DC: U. S. Department of Agriculture.

Arthur W, Romiguier J. 2021. Relaxation of purifying selection suggests low effective population size in eusocial Hymenoptera and solitary pollinating bees. Peer Commun J. 1:e2. https://doi.org/10.24072/pcjournal.3

Barkdull M, Moreau CS. 2023. Worker reproduction and caste polymorphism impact genome evolution and social genes across the ants. Genome Biol. Evol. 15(6):24. https://doi.org/10.1093/gbe/evad095

Berens AJ, Hunt JH, Toth AL. 2015. Comparative transcriptomics of convergent evolution: different genes but conserved pathways underlie caste phenotypes across lineages of eusocial insects. Mol. Biol. Evol. 32(3):690–703. https://doi.org/10.1093/molbev/msu330

Berger J, Legendre F, Zelosko KM, et al. 2022. Eusocial transition in Blattodea: transposable elements and shifts of gene expression. Genes-Basel 13(11):1948. https://doi.org/10.3390/genes13111948

Bertorelle G, Raffini F, Bosse M, et al. 2022. Genetic load: genomic estimates and applications in non-model animals. Nat. Rev. Genet. 23(8):492–503. https://doi.org/10.1038/s41576-022-00448-x

Bewick AJ, Vogel KJ, Moore AJ, et al. 2017. Evolution of DNA methylation across Insects. Mol. Biol. Evol. 34(3):654–665. https://doi.org/10.1093/molbev/msw264

Bluher SE, Miller SE, Sheehan MJ. 2020. Fine-scale population structure but limited genetic differentiation in a cooperatively breeding paper wasp. Genome Biol. Evol. 12(5):701–714. https://doi.org/10.1093/gbe/evaa070

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina Sequence Data. Bioinformatics 30(15):2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bourgeois Y, Boissinot S. 2019. On the population dynamics of junk: a review on the population genomics of transposable elements. Genes-Basel 10(6):419. https://doi.org/10.3390/genes10060419

Bousjein NS, Tierney SM, Gardner MG, et al. 2022. Does effective population size affect rates of molecular evolution: mitochondrial data for host/parasite species pairs in bees suggests not. Ecol. Evol. 12(2):ARTN e8562. https://doi.org/10.1002/ece3.8562

Brandt M, Foitzik S, Fischer-Blass B, et al. 2005. The coevolutionary dynamics of obligate ant social parasite system-between prudence and antagonism. Biol. Rev. 80(2):251–267. https://doi.org/10.1017/s1464793104006669

Buchfink B, Reuter K, Drost HG. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat. Meth. 18(4):366–368. https://doi.org/10.1038/s41592-021-01101-x

Bulla I, Aliaga B, Lacal V, et al. 2018. Notos - a galaxy tool to analyze CpN observed expected ratios for inferring DNA methylation types. BMC Bioinf. 19(1):105. https://doi.org/10.1186/s12859-018-2115-4

Buschinger A. 2009. Social parasitism among ants: a review (Hymenoptera: Formicidae). Myrmecol News 12:219–235.

Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner. Paper Presented at: 9th Annual Genomics of Energy & Environment Meeting. Lawrence Berkeley National Lab.

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

Campbell MS, Law MY, Holt C, et al. 2014. MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations. Plant Physiol. 164(2):513–524. https://doi.org/10.1104/pp.113.230144

Cervo R. 2006. *Polistes* wasps and their social parasites: an overview. Ann. Zool. Fenn. 43(5-6):531–549.

Charlesworth B, Langley CH. 2021. The evolution of self-regulated transposition of transposable elements. Gene. 112(2): 359–383. https://doi.org/10.1093/genetics/112.2.359

Crowley LM; University of Oxford and Wytham Woods Genome Acquisition Lab. 2022. The genome sequence of the German wasp, *Vespula germanica* (Fabricius, 1793). Wellcome Open Res. 7:60. https://doi.org/10.12688/wellcomeopenres.17703.1

Danecek P, Bonfield JK, Liddle J, et al. 2021. Twelve years of SAMtools and BCFtools. GigaScience 10(2):ARTN giab008. https://doi.org/10.1093/gigascience/giab008

Dobin A, Davis CA, Schlesinger F, et al. 2012. STAR: ultrafast universal RNA-seq aligner. Bioinformat. 29(1):15–21. https://doi.org/10.1093/bioinformatics/bts635

Dogantzis KA, Harpur BA, Rodrigues A, et al. 2018. Insects with similar social complexity show convergent patterns of adaptive molecular evolution. Sci. Rep.-UK 8(1):ARTN 10388. https://doi.org/10.1038/s41598-018-28489-5

Drost HG, Gabel A, Grosse I, et al. 2015. Evidence for active maintenance of Phylotranscriptomic hourglass patterns in animal and plant embryogenesis. Mol. Biol. Evol. 32(5):1221–1231. https://doi.org/10.1093/molbev/msv012

Duret L, Mouchiroud D. 2000. Determinants of substitution rates in mammalian genes: Expression pattern affects selection intensity but not mutation rate. Mol. Biol. Evol. 17(1):68–74. https://doi.org/10.1093/oxfordjournals.molbev.a026239

Dyson CJ, Piscano OL, Durham RM, et al. 2021. Temporal analysis of effective population size and mating system in a social wasp. J. Hered. 112(7):626–634. https://doi.org/10.1093/jhered/esab057

Dyson CJ, Crossley HG, Ray CH, et al. 2022. Social structure of perennial *Vespula squamosa* wasp colonies. Ecol. Evol. 12(2):e8569. https://doi.org/10.1002/ece3.8569

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20(1):ARTN 238. https://doi.org/10.1186/s13059-019-1832-y

Ewels P, Magnusson M, Lundin S, et al. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32(19):3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Favreau E, Cini A, Taylor D, et al. 2023a. Putting hornets on the genomic map. Sci. Rep.-UK 13(1):ARTN 6232. https://doi.org/10.1038/s41598-023-31932-x

Favreau E, Geist KS, Wyatt CDR, et al. 2023b. Co-expression gene networks and machine-learning algorithms unveil a core genetic toolkit for reproductive division of labour in Rudimentary Insect Societies. Genome Biol. Evol. 15(1):ARTN evac174. https://doi.org/10.1093/gbe/evac174

Ferreira PG, Patalano S, Chauhan R, et al. 2013. Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. Genome Biol. 14(2):R20. https://doi.org/10.1186/gb-2013-14-2-r20

Friedli M, Trono D. 2015. The developmental control of transposable elements and the evolution of higher species. Annu. Rev. Cell Dev. Biol. 31(429):429–451. https://doi.org/10.1146/annurev-cellbio-100814-125514

Geffre AC, Liu RL, Manfredini F, et al. 2017. Transcriptomics of an extended phenotype: parasite manipulation of wasp social behaviour shifts expression of caste-related genes. Proc. R. Soc. B Biol. Sci. 284(1852):20170029. https://doi.org/10.1098/rspb.2017.0029

Gilbert C, Peccoud J, Cordaux R. 2021. Transposable elements and the evolution of insects. Annu. Rev. Entomol. 66:355–372. https://doi.org/10.1146/annurev-ento-070720-074650

Glastad KM, Hunt BG, Goodisman MAD. 2013. Evidence of a conserved functional role for DNA methylation in termites. Insect. Mol. Biol. 22(2):143–154. https://doi.org/10.1111/imb.12010

Glastad KM, Hunt BG, Goodisman MAD. 2014a. Evolutionary insights into DNA methylation in insects. Curr. Opin. Insect Sci. 1:25–30. https://doi.org/10.1016/j.cois.2014.04.001

Glastad KM, Hunt BG, Yi SV, et al. 2014b. Epigenetic inheritance and genome regulation: is DNA methylation linked to ploidy in haplodiploid insects? Proc. R. Soc. B Biol. Sci. 281(1785):20140411. https://doi.org/10.1098/rspb.2014.0411

Glastad KM, Chau LM, Goodisman MAD. 2015. '7 Epigenetics in social insects'. In: Zayed A, Kent CF, editors. *Advances in Insect Physiology*. Oxford: Academic Press; p. 227–269.

Glastad KM, Gokhale K, Liebig J, et al. 2016. The caste- and sex-specific DNA methylome of the termite *Zootermopsis nevadensis*. Sci. Rep.-UK 6(1):37110.31038/srep37110. https://doi.org/10.1038/srep37110

Glastad KM, Arsenault SV, Vertacnik KL, et al. 2017. Variation in DNA methylation is not consistently reflected by sociality in Hymenoptera. Genome Biol. Evol. 9(6):1687–1698. https://doi.org/10.1093/gbe/evx128

Goodisman MAD, Kovacs JL, Hoffman EA. 2007a. Lack of conflict during queen production in the social wasp *Vespula maculifrons*. Mol. Ecol. 16(12):2589–2595. https://doi.org/10.1111/j.1365-294X.2007.03316.x

Goodisman MAD, Kovacs JL, Hoffman EA. 2007b. The significance of multiple mating in the social wasp *Vespula maculifrons*. Evolution 61(9):2260–2267. https://doi.org/10.1111/j.1558-5646.2007.00175.x

Gotz S, Garcia-Gomez JM, Terol J, et al. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 36(10):3420–3435. https://doi.org/10.1093/nar/gkn176

Gout JF, Kahn D, Duret L; Paramecium Post-Genomics Consortium. 2010. The relationship among gene expression, the evolution of gene dosage, and the rate of protein evolution. PLoS Genet. 6(6):ARTN e1000944. https://doi.org/10.1371/journal.pgen.1000944

Grabherr MG, Haas BJ, Yassour M, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29(7):644–652. https://doi.org/10.1038/nbt.1883

Greene A. 1991. *Dolichovespula and Vespula*'. In: Ross KG, Matthews RW, editors. *The Social Biology of Wasps*. Ithaca, NY: Comstock Publishing Associates; p. 263–305.

Gu ZG, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32(18):2847–2849. https://doi.org/10.1093/bioinformatics/btw313

Gurevich A, Saveliev V, Vyahhi N, et al. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29(8):1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Harpur BA, Kent CF, Molodtsova D, et al. 2014. Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. Proc. Natl. Acad. Sci. USA. 111(7):2614–2619. https://doi.org/10.1073/pnas.1315506111

Harrison MC, Jongepier E, Robertson HM, et al. 2018. Hemimetabolous genomes reveal molecular basis of termite eusociality. Nat. Ecol. Evol. 2(3):557–566. https://doi.org/10.1038/s41559-017-0459-1

Harrop TWR, Guhlin J, McLaughlin GM, et al. 2020. High-quality assemblies for three invasive social wasps from the *Vespula* Genus. G3 (Bethesda, MD) 10(10):3479–3488. https://doi.org/10.1534/g3.120.401579

Hoffman EA, Goodisman MAD. 2007. Gene expression and the evolution of phenotypic diversity in social wasps. BMC Biol. 5(Art. 23):23. https://doi.org/10.1186/1741-7007-5-23

Hoffman EA, Kovacs JL, Goodisman MAD. 2008. Genetic structure and breeding system in a social wasp and its social parasite. BMC Evol. Biol. 8:Artn 239. https://doi.org/10.1186/1471-2148-8-239

Howard KJ. 2006. Three queen morphs with alternative nest-founding behaviors in the ant,. Insectes Soc. 53(4):480–488. https://doi.org/10.1007/s00040-006-0905-6

Hunt BG, Goodisman MAD. 2010. Evolutionary variation in gene expression is associated with dimorphism in eusocial vespid wasps. Insect. Mol. Biol. 19(5):641–652. https://doi.org/10.1111/j.1365-2583.2010.01021.x

Hunt JH, Toth AL. 2017. Sociality in wasps. Comp. Soc. Evol. 84–123. https://doi.org/10.1017/9781107338319

Hunt JH, Cave RD, Borjas GR. 2001. First records from Honduras of a yellowjacket wasp, *Vespula squamosa* (Drury) (Hymenoptera: Vespidae: Vespinae). J. Kans. Entomol. Soc. 74(2):118–119.

Hunt BG, Ometto L, Keller L, et al. 2013a. Evolution at two levels in fire ants: the relationship between patterns of gene expression and protein sequence evolution. Mol. Biol. Evol. 30(2):263–271. https://doi.org/10.1093/molbev/mss234

Hunt BG, Glastad KM, Yi SV, et al. 2013b. The function of intragenic DNA methylation: insights from insect epigenomes. Integr. Comp. Biol. 53(2):319–328. https://doi.org/10.1093/icb/ict003

Hunt BG, Glastad KM, Yi SV, et al. 2013c. Patterning and regulatory associations of DNA methylation are mirrored by histone modifications in insects. Genome Biol. Evol. 5(3):591–598. https://doi.org/10.1093/gbe/evt030

Imrit MA, Dogantzis KA, Harpur BA, et al. 2020. Eusociality influences the strength of negative selection on insect genomes. Proc. R. Soc. B Biol. Sci. 287(1933):ARTN 20201512. https://doi.org/10.1098/rspb.2020.1512

Ingram KK, Gordon DM, Friedman DA, et al. 2016. Context-dependent expression of the foraging gene in field colonies of ants: the interacting roles of age, environment and task. Proc. R. Soc. B Biol. Sci. 283(1837):ARTN 20160841. https://doi.org/10.1098/rspb.2016.0841

Jandt JM, Toth AL. 2015. Physiological and genomic mechanisms of social organization in wasps (Family: Vespidae)‘. In: Zayed A, Kent CF, editors. *Advances Insect Physiology*. Oxford: Academic Press; p. 95–130.

Jeffares DC, Tomiczek B, Sojo V, et al. 2015. A beginners guide to estimating the non-synonymous to synonymous rate ratio of all protein-coding genes in a genome. Methods Mol. Biol. (Clifton, NJ) 1201:65–90. https://doi.org/10.1007/978-1-4939-1438-8_4

Jeong H, Wu X, Smith B, et al. 2018. Genomic landscape of methylation islands in hymenopteran insects. Genome Biol. Evol. 10(10):2766–2776. https://doi.org/10.1093/gbe/evy203

Johnson EL, Cunningham TW, Marriner SM, et al. 2009. Resource allocation in a social wasp: effects of breeding system and life cycle on reproductive decisions. Mol. Ecol. 18(13):2908–2920. https://doi.org/10.1111/j.1365-294X.2009.04240.x

Jones BM, Kingwell CJ, Wcislo WT, et al. 2017. Caste-biased gene expression in a facultatively eusocial bee suggests a role for genetic accommodation in the evolution of eusociality. Proc. R. Soc. B Biol. Sci. 284(1846):20162228. https://doi.org/10.1098/rspb.2016.2228

Kapheim KM, Jones BM, Pan HL, et al. 2020. Developmental plasticity shapes social traits and selection in a facultatively eusocial bee. Proc. Natl. Acad. Sci. USA. 117(24):13615–13625. https://doi.org/10.1073/pnas.2000344117

Katoh H, Miura T, Maekawa K, et al. 2002. Genetic variation of symbiotic fungi cultivated by the macrotermitinae termite *Odontotermes formosanus* (Isoptera: Termitidae) in the Ryukyu Archipelago. Mol. Ecol. 11(8):1565–1572. https://doi.org/10.1046/j.1365-294x.2002.01535.x

Keller L, Passera L. 1989. Size and fat content of gynes in relation to the mode of colony founding in ants (Hymenoptera; Formicidae). Oecologia 80(2):236–240. https://doi.org/10.1007/BF00380157

Kilner RM, Langmore NE. 2011. Cuckoos *versus* hosts in insects and birds: adaptations, counter-adaptations and outcomes. Biol. Rev. Camb. Philos. Soc. 86(4):836–852. https://doi.org/10.1111/j.1469-185X.2010.00173.x

Kim D, Paggi JM, Park C, et al. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37(8):907–915. https://doi.org/10.1038/s41587-019-0201-4

Kolmogorov M, Yuan J, Lin Y, et al. 2019. Assembly of long, error-prone reads using repeat graphs. Nat. Biotechnol. 37(5):540–546. https://doi.org/10.1038/s41587-019-0072-8

Lenoir A, D’Ettorre P, Errard C, et al. 2001. Chemical ecology and social parasitism in ants. Annu. Rev. Entomol. 46:573–599. https://doi.org/10.1146/annurev.ento.46.1.573

Lester PJ, Beggs JR. 2019. Invasion success and management strategies for social *Vespula* Wasps. Annu. Rev. Entomol. 64:51–71. https://doi.org/10.1146/annurev-ento-011118-111812

Lewis SH, Ross L, Bain SA, et al. 2020. Widespread conservation and lineage-specific diversification of genome-wide DNA methylation patterns across arthropods. PLoS Genet. 16(6):e1008864. https://doi.org/10.1371/journal.pgen.1008864

Lhomme P, Hines H. 2018. Reproductive dominance strategies in insect social parasites. J. Chem. Ecol. 44(9):838–850. https://doi.org/10.1007/s10886-018-0971-z

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34(18):3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf. 12. https://doi.org/10.1186/1471-2105-12-323

Li ST, Ma LJ, Li H, et al. 2007. Snap: an integrated SNP annotation platform. Nucleic Acids Res. 35(suppl_1):D707–D710. https://doi.org/10.1093/nar/gkl969

Li H, Handsaker B, Wysoker A, et al; 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li T, Zhang XK, Luo F, et al. 2020. MultiMotifMaker: a multi-thread tool for identifying DNA methylation motifs from Pacbio Reads. IEEE/ACM Trans. Comput. Biol. Bioinf. 17(1):220–225. https://doi.org/10.1109/TCBB.2018.2861399

Lin Y, Yuan J, Kolmogorov M, et al. 2016. Assembly of long error-prone reads using de Bruijn graphs. Proc. Natl. Acad. Sci. USA. 113(52):E8396–E8405. https://doi.org/10.1073/pnas.1604560113

Loope KJ, Chien C, Juhl M. 2014. Colony size is linked to paternity frequency and paternity skew in yellowjacket wasps and hornets. BMC Evol. Biol. 14(1):ARTN 2625. https://doi.org/10.1186/S12862-014-0277-X

Loope KJ, Lopez-Osorio F, Dvorak L. 2017. Convergent reversion to single mating in a wasp social parasite. Am. Nat. 189(6):E138–E151. https://doi.org/10.1086/691405

Lopez-Osorio F, Pickett KM, Carpenter JM, et al. 2014. Phylogenetic relationships of yellowjackets inferred from nine loci (Hymenoptera: Vespidae, Vespinae, *Vespula* and *Dolichovespula*). Mol. Phylogenet. Evol. 73:190–201. https://doi.org/10.1016/j.ympev.2014.01.007

Lucas ER, Romiguier J, Keller L. 2017. Gene expression is more strongly influenced by age than caste in the ant Lasius niger. Mol. Ecol. 26(19):5058–5073. https://doi.org/10.1111/mec.14256

MacDonald JF, Matthews RW. 1975. *Vespula squamosa*: a yellow jacket wasp evolving towards parasitism. Science 190(4218):1003–1004.

MacDonald JF, Matthews RW. 1981. Nesting biology of the Eastern yellowjacket, *Vespula maculifrons* (Hymenoptera: Vespidae). J. Kans. Entomol. Soc. 54(3):433–457.

MacDonald JF, Matthews RW. 1984. Nesting biology of the southern yellowjacket, *Vespula squamosa* (Hymenoptera: Vespidae): social parasitism and independent founding. J. Kans. Entomol. Soc. 57(1):134–151.

Matthews RW. 1982. Social parasitism in yellowjackets’. In: Jaisson P, editor. *Social Insects in the Tropics*. Paris: University of Paris-Nord; p. 193–202.

Mikhailova AA, Rinke S, Harrison MC. 2024. Genomic signatures of eusocial evolution in insects. Curr. Opin. Insect Sci. 61:101136. https://doi.org/10.1016/j.cois.2023.101136

Miller SE, Sheehan MJ. 2023. Sex differences in deleterious genetic variants in a haplodiploid social insect. Mol. Ecol. 32(16):4546–4556. https://doi.org/10.1111/mec.17057

Miller SE, Legan AW, Henshaw MT, et al. 2020. Evolutionary dynamics of recent selection on cognitive abilities. Proc. Natl. Acad. Sci. USA. 117(6):3045–3052. https://doi.org/10.1073/pnas.1918592117

Miller SE, Legan AW, Uy FMK, et al. 2022. Highly contiguous genome assemblies of the Guinea paper wasp (*Polistes exclamans*) and *Mischocyttarus mexicanus*. Genome Biol. Evol. 14(8):evac110. https://doi.org/10.1093/gbe/evac110

Minh BQ, Schmidt HA, Chernomor O, et al. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37(5):1530–1534. https://doi.org/10.1093/molbev/msaa015

Morandin C, Dhaygude K, Paviala J, et al. 2015. Caste-biases in gene expression are specific to developmental stage in the ant Formica exsecta. J. Evol. Biol. 28(9):1705–1718. https://doi.org/10.1111/jeb.12691

Nanda AS, Wu K, Irkliyenko I, et al. 2024. Direct transposition of native DNA for sensitive multimodal single-molecule sequencing. Nat. Genet. 56(6):1300–1309. https://doi.org/10.1038/s41588-024-01748-0

Neph S, Kuehn MS, Reynolds AP, et al. 2012. BEDOPS: high-performance genomic feature operations. Bioinformatics 28(14):1919–1920. https://doi.org/10.1093/bioinformatics/bts277

Ni P, Nie F, Zhong Z, et al. 2023. DNA 5-methylcytosine detection and methylation phasing using PacBio circular consensus sequencing. Nat. Commun. 14(1):4054. https://doi.org/10.1038/s41467-023-39784-9

Nijhout HF, McKenna KZ. 2018. The distinct roles of insulin signaling in polyphenic development. Curr. Opin. Insect Sci. 25:58–64. https://doi.org/10.1016/j.cois.2017.11.011

Orr SE, Goodisman MAD. 2023. Social insect transcriptomics and the molecular basis of caste diversity. Curr. Opin. Insect Sci. 57:8. https://doi.org/10.1016/j.cois.2023.101040

Oster GF, Wilson EO. 1978. Caste and ecology in the social insects. Princeton, NJ: Princeton University Press.

Pal C, Papp B, Hurst LD. 2001. Highly expressed genes in yeast evolve slowly. Genetics 158(2):927–931. https://doi.org/10.1093/genetics/158.2.927

Papkou A, Gokhale CS, Traulsen A, et al. 2016. Host-parasite coevolution: why changing population size matters. Zoology (Jena, Germany) 119(4):330–338. https://doi.org/10.1016/j.zool.2016.02.001

Patalano S, Vlasova A, Wyatt C, et al. 2015. Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies. Proc. Natl. Acad. Sci. USA. 112(45):13970–13975. https://doi.org/10.1073/pnas.1515937112

Patalano S, Alsina A, Gregorio-Rodriguez C, et al. 2022. Self-organization of plasticity and specialization in a primitively social insect. Cell Syst. 13(9):768–779.e4. https://doi.org/10.1016/j.cels.2022.08.002

Paysan-Lafosse T, Blum M, Chuguransky S, et al. 2023. InterPro in 2022. Nucleic Acids Res. 51(D1):D418–D427. https://doi.org/10.1093/nar/gkac993

Pond SLK, Poon AFY, Velazquez R, et al. 2020. HyPhy 2.5-A customizable platform for evolutionary hypothesis testing using phylogenies. Mol. Biol. Evol. 37(1):295–299. https://doi.org/10.1093/molbev/msz197

Popescu AA, Huber KT, Paradis E. 2012. ape 3.0: new tools for distance-based phylogenetics and evolutionary analysis in R. Bioinformatics 28(11):1536–1537. https://doi.org/10.1093/bioinformatics/bts184

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6):841–842. https://doi.org/10.1093/bioinformatics/btq033

Rabeling C. 2020. Social parasitism'. In: Starr CK, editor. *Encyclopedia of Social Insects*. Cham: Springer International Publishing; p. 1–23.

Rahman SR, Lozier JD. 2023. Genome-wide DNA methylation patterns in bumble bee *Bombus vosnesenskii* populations from spatial-environmental range extremes. Sci. Rep-UK 13(1):ARTN 14901 . https://doi.org/10.1038/s41598-023-41896-7

Rambaut A. 2010. 'Chapter FigTree v1.3.1'. Edinburgh: Institute of Evolutionary Biology, University of Edinburgh,.

Rehan SM, Glastad KM, Lawson SP, et al. 2016. The genome and methylome of a subsocial small carpenter bee, Ceratina calcarata. Genome Biol. Evol. 8(5):1401–1410. https://doi.org/10.1093/gbe/evw079

Rubenstein DR, Agren JA, Carbone L, et al. 2019. Coevolution of genome architecture and social behavior. Trends Ecol. Evol. 34(9):844–855. https://doi.org/10.1016/j.tree.2019.04.011

Rubin BER. 2022. Social insect colony size is correlated with rates of molecular evolution. Insectes Soc. 69(2-3):147–157. https://doi.org/10.1007/s00040-022-00859-3

Rubin BER, Moreau CS. 2016. Comparative genomics reveals convergent rates of evolution in ant-plant mutualisms. Nat. Commun. 7(1):ARTN 12679. https://doi.org/10.1038/ncomms12679

Sankovitz M, Loope KJ, Rankin EWE, et al. 2023. Unequal reproduction early in a social transition: insights from invasive wasps. Am. Nat 201(2):241–255. https://doi.org/10.1086/722514

Santos D, Feng M, Kolliopoulou A, et al. 2023. What are the functional roles of Piwi proteins and piRNAs in insects? Insects 14(2):187. https://doi.org/10.3390/insects14020187

Sayols S. 2023. rrvgo: a Bioconductor package for interpreting lists of Gene Ontology terms. Micro. Publ. Biol. https://doi.org/10.17912/micropub.biology.000811

Scarparo G, Sankovitz M, Loope KJ, et al. 2021. Early queen joining and long-term queen associations in polygyne colonies of an invasive wasp revealed by longitudinal genetic analysis. Evol. Appl. 14(12):2901–2914. https://doi.org/10.1111/eva.13324

Schrader L, Schmitz J. 2019. The impact of transposable elements in adaptive evolution. Mol. Ecol. 28(6):1537–1549. https://doi.org/10.1111/mec.14794

Schrader L, Helantera H, Oettler J. 2017. Accelerated evolution of developmentally biased genes in the Tetraphenic Ant *Cardiocondyla obscurior*. Mol. Biol. Evol. 34(3):535–544. https://doi.org/10.1093/molbev/msw240

Schrader L, Pan H, Bollazzi M, et al. 2021. Relaxed selection underlies genome erosion in socially parasitic ant species. Nat. Commun. 12(1):ARTN 2918. https://doi.org/10.1038/s41467-021-23178-w

Shell WA, Steffen MA, Pare HK, et al. 2021. Sociality sculpts similar patterns of molecular evolution in two independently evolved lineages of eusocial bees. Commun. Biol. 4(1):ARTN 253. https://doi.org/10.1038/s42003-021-01770-6

Sigurpalsdottir BD, Stefansson OA, Holley G, et al. 2024. A comparison of methods for detecting DNA methylation from long-read sequencing of human genomes. Genome Biol. 25(1):69. https://doi.org/10.1186/s13059-024-03207-9

Simao FA, Waterhouse RM, Ioannidis P, et al. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31(19):3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Simola DF, Wissler L, Donahue G, et al. 2013. Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. Genome Res. 23(8):1235–1247. https://doi.org/10.1101/gr.155408.113

Snyder AK, Loope KJ. 2021. High rates of polygyny in tropical Mexico within the native range of *Vespula squamosa*. Georgia Southern.

Song Y, Gu FM, Liu ZX, et al. 2022. The key role of fatty acid synthase in lipid metabolism and metamorphic development in a destructive insect pest, *Spodoptera litura* (Lepidoptera: Noctuidae). Int. J. Mol. Sci. 23(16):9064. https://doi.org/10.3390/ijms23169064

Standage DS, Berens AJ, Glastad KM, et al. 2016. Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. Mol. Ecol. 25(8):1769–1784. https://doi.org/10.1111/mec.13578

Stanke M, Schoffmann O, Morgenstern B, et al. 2006. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinf. 7(1). https://doi.org/10.1186/1471-2105-7-62

Stanke M, Diekhans M, Baertsch R, et al. 2008. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics 24(5):637–644. https://doi.org/10.1093/bioinformatics/btn013

Su WJ, Ou SJ, Hufford MB, et al. 2021. A tutorial of EDTA: extensive de novo TE annotator. Plant Transposable Elements 2250(55-67):55–67. https://doi.org/10.1007/978-1-0716-1134-0_4

Sumner S, Favreau E, Geist K, et al. 2023. Molecular patterns and processes in evolving sociality: lessons from insects. Philos. Trans. R. Soc. B 378(1874):10. https://doi.org/10.1098/rstb.2022.0076

Sun JH, Lu F, Luo YJ, et al. 2023. OrthoVenn3: an integrated platform for exploring and visualizing orthologous data across genomes. Nucleic Acids Res. 51(W1):W397–W403. https://doi.org/10.1093/nar/gkad313

Supek F, Gibas C, Bošnjak M, et al. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS ONE 6(7):e21800–e21800. https://doi.org/10.1371/journal.pone.0021800

Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34(Web Server issue):W609–W612. https://doi.org/10.1093/nar/gkl315

Taylor LH. 1939. Observations on social parasitism in the genus *Vespula* Thompson. Ann. Entomol. Soc. Am. 32:304–315.

Taylor D, Bentley MA, Sumner S. 2018. Social wasps as models to study the major evolutionary transition to superorganismality. Curr. Opin. Insect Sci. 28:26–32. https://doi.org/10.1016/j.cois.2018.04.003

Taylor BA, Taylor D, Bodrug-Schepers A, et al. 2023. Molecular signatures of alternative reproductive strategies in a facultatively social hover wasp. Mol. Ecol. 15. https://doi.org/10.1111/mec.17217

Thomas GWC, Dohmen E, Hughes DST, et al. 2020. Gene content evolution in the arthropods. Genome Biol. 21(1):ARTN 15. https://doi.org/10.1186/s13059-019-1925-7

Tosto NM, Beasley ER, Wong BBM, et al. 2023. The roles of sexual selection and sexual conflict in shaping patterns of genome and transcriptome variation. Nat. Ecol. Evol. 7(7):981–993. https://doi.org/10.1038/s41559-023-02019-7

Toth AL, Rehan SM. 2017. Molecular evolution of insect sociality: an Eco-Evo-Devo perspective. Annu. Rev. Entomol. 62:419–442. https://doi.org/10.1146/annurev-ento-031616-035601

Toth AL, Robinson GE. 2010. Evo-Devo and the evolution of social behavior: brain gene expression analyses in social Insects'. Cold Spring Harbor, New York, USA: Cold Spring Harbor Laboratory Press; p. 1–8.

Toth AL, Varala K, Henshaw MT, et al. 2010. Brain transcriptomic analysis in paper wasps identifies genes associated with behaviour across social insect lineages. Proc. Biol. Sci. 277(1691):2139–2148. https://doi.org/10.1098/rspb.2010.0090

Toth AL, Tooker JF, Radhakrishnan S, et al. 2014. Shared genes related to aggression, rather than chemical communication, are associated with reproductive dominance in paper wasps (*Polistes metricus*). BMC Genom. 15(1):75. https://doi.org/10.1186/1471-2164-15-75

Tse OYO, Jiang PY, Cheng SH, et al. 2021. Genome-wide detection of cytosine methylation by single molecule real-time sequencing. Proc. Natl. Acad. Sci. USA. 118(5):ARTN e2019768118. https://doi.org/10.1073/pnas.2019768118

Walsh TK, Brisson JA, Robertson HM, et al. 2010. A functional DNA methylation system in the pea aphid, Acyrthosiphon pisum. Insect. Mol. Biol. 19(Suppl 2):215–228. https://doi.org/10.1111/j.1365-2583.2009.00974.x

Warners MR, Qiu LJ, Holmes MJ, et al. 2019. Convergent eusocial evolution is based on a shared reproductive groundplan plus lineage-specific plastic genes. Nat. Commun. 10(ARTN 2651). https://doi.org/10.1038/s41467-019-10546-w

Weiner SA, Galbraith DA, Adams DC, et al. 2013. A survey of DNA methylation across social insect species, life stages, and castes reveals abundant and caste-associated methylation in a primitively social wasp. Naturwissenschaften 100(8):795–799. https://doi.org/10.1007/s00114-013-1064-z

Wertheim JO, Murrell B, Smith MD, et al. 2015. RELAX: detecting relaxed selection in a phylogenetic framework. Mol. Biol. Evol. 32(3):820–832. https://doi.org/10.1093/molbev/msu400

Weyna A, Romiguier J. 2021. Relaxation of purifying selection suggests low effective population size in eusocial Hymenoptera and solitary pollinating bees. Peer Commun. J. 1:e2. https://doi.org/10.24072/pcjournal.3

Wilson EO, Hölldobler B. 2005. Eusociality: origin and consequences. Proc. Natl. Acad. Sci. USA. 102(38):13367–13371. https://doi.org/10.1073/pnas.0505858102

Wilson EE, Mullen LM, Holway DA. 2009. Life history plasticity magnifies the ecological effects of a social wasp invasion. Proc. Natl. Acad. Sci. USA. 106(31):12809–12813. https://doi.org/10.1073/pnas.0902979106

Wilson-Rankin EE. 2021. Emerging patterns in social wasp invasions. Curr. Opin. Insect Sci. 46:72–77. https://doi.org/10.1016/j.cois.2021.02.014

Wyatt CDR, Bentley MA, Taylor D, et al. 2023. Social complexity, life-history and lineage influence the molecular basis of castes in vespid wasps. Nat. Commun. 14(1):16. https://doi.org/10.1038/s41467-023-36456-6

Yi SV, Goodisman MAD. 2009. Computational approaches for understanding the evolution of DNA methylation in animals. Epigenetics 4(8):551–556. https://doi.org/10.4161/epi.4.8.10345

Yanai I, Benjamin H, Shmoish M, et al. 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. Bioinformatics 21(5):650–659. https://doi.org/10.1093/bioinformatics/bti042

Zhang JZ, Yang JR. 2015. Determinants of the rate of protein sequence evolution. Nat. Rev. Genet. 16(7):409–420. https://doi.org/10.1038/nrg3950