



Unexpectedly low recombination rates and presence of hotspots in termite genomes

Turid Everitt, Tilman Rönneburg, Daniel Elsner, et al.

Genome Res. published online March 20, 2025

Access the most recent version at doi:[10.1101/gr.279180.124](https://doi.org/10.1101/gr.279180.124)

P<P Published online March 20, 2025 in advance of the print journal.

Creative Commons License

This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <https://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Research

Unexpectedly low recombination rates and presence of hotspots in termite genomes

Turid Everitt,^{1,5} Tilman Rönneburg,^{1,5} Daniel Elsner,² Anna Olsson,¹ Yuanzhen Liu,¹ Tuuli Larva,¹ Judith Korb,^{2,3} and Matthew T. Webster^{1,4}

¹Medical Biochemistry and Microbiology, Uppsala University, 751 23 Uppsala, Sweden; ²Evolutionary Biology and Ecology, University of Freiburg, D-79104 Freiburg, Germany; ³Research Institute for the Environment and Livelihoods, Charles Darwin University, Casuarina Campus, Darwin, Casuarina NT 0909, Australia; ⁴Science for Life Laboratory, Uppsala University, 752 37 Uppsala, Sweden

Meiotic recombination is a fundamental evolutionary process that facilitates adaptation and the removal of deleterious genetic variation. Social Hymenoptera exhibit some of the highest recombination rates among metazoans, whereas high recombination rates have not been found among nonsocial species from this insect order. It is unknown whether elevated recombination rates are a ubiquitous feature of all social insects. In many metazoan taxa, recombination is mainly restricted to hotspots a few kilobases in length. However, little is known about the prevalence of recombination hotspots in insect genomes. Here we infer recombination rate and its fine-scale variation across the genomes of two social species from the insect order Blattodea: the termites *Macrotermes bellicosus* and *Cryptotermes secundus*. We used linkage disequilibrium-based methods to infer recombination rate. We infer that recombination rates are close to 1 cM/Mb in both species, similar to the average metazoan rate. We also observe a highly punctate distribution of recombination in both termite genomes, indicative of the presence of recombination hotspots. We infer the presence of full-length *PRDM9* genes in the genomes of both species, which suggests recombination hotspots in termites might be determined by *PRDM9*, as they are in mammals. We also find that recombination rates in genes are correlated with inferred levels of germline DNA methylation. The finding of low recombination rates in termites indicates that eusociality is not universally connected to elevated recombination rate. We speculate that the elevated recombination rates in social Hymenoptera are instead promoted by intense selection among haploid males.

[Supplemental material is available for this article.]

Meiotic recombination has two fundamental roles: It is essential for correct disjunction of chromosomes during cell division, and it generates new combinations of genetic variants that form the raw material of evolution (Hartfield and Keightley 2012). However, recombination can also be deleterious, as it can generate structural mutations and break up favorable combinations of alleles. Recombination varies among species, with optimal recombination rate in a genome likely determined by the optimal trade-off of its positive and negative evolutionary and mechanistic effects (Stapley et al. 2017).

The highest recombination rates in metazoans observed so far are found in social insects (Wilfert et al. 2007; Stapley et al. 2017). However, until now, the recombination rate has only been studied in social insects from the order Hymenoptera, to which the majority of social insects belong. Among Hymenoptera, the highest rates are found in the genus *Apis*: the Western honey bee *Apis mellifera* (20.8 cM/Mb) and its relatives *Apis cerana* (17.4 cM/Mb), *Apis florea* (20.8 cM/Mb), and *Apis dorsata* (25.1 cM/Mb) (Beye et al. 2006; Shi et al. 2013; Liu et al. 2015; Wallberg et al. 2015; Rueppell et al. 2016; Kawakami et al. 2019). Other social Hymenoptera also have high rates, including the bumblebee *Bombus terrestris* (8.9 cM/Mb) (Liu et al. 2017; Kawakami et al. 2019), the stingless bee *Frieseomelitta varia* (9.3–12.5 cM/Mb) (Waiker et al. 2021), the

wasp *Vespa vulgaris* (9.7 cM/Mb) (Sirviö et al. 2011a), and the ants *Pogonomyrmex rugosus* (11.1 cM/Mb) (Sirviö et al. 2011b) and *Acromyrmex echinator* (6.1 cM/Mb) (Sirviö et al. 2006). In contrast, the solitary bee *Megachile rotundata* and the solitary wasp *Nasonia* have relatively low rates (1.0 and 1.5 cM/Mb, respectively) (Niehuis et al. 2010; Jones et al. 2019). For comparison, the average recombination rate of 15 insect species outside of Hymenoptera is 2.2 cM/Mb (Wilfert et al. 2007).

We can learn about the forces shaping the evolution of recombination rates in general by understanding why high recombination rates have evolved in social Hymenoptera. Several hypotheses have been proposed to explain this observation. One set of explanations focuses on the effects of recombination on intra-colony genetic diversity. For example, increasing the genetic diversity of a colony could promote task specialization among workers (Kent et al. 2012; Kent and Zayed 2013). Similarly, elevated genetic variation in a colony could prevent invasion by pathogens or parasites owing to diversification of immune genes (Fischer and Schmid-Hempel 2005). Recombination could also reduce variance in relatedness between nestmates, thereby reducing potential kin conflict within colonies (Sherman 1979; Templeton 1979; Wilfert et al. 2007). Additionally, it has been proposed

⁵These authors contributed equally to this work.

Corresponding author: matthew.webster@imbim.uu.se

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.279180.124>.

© 2025 Everitt et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <https://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

that high recombination rates could have facilitated the evolution of eusociality over a longer evolutionary timescale, because recombination between genes that take on functions in different castes permits them to evolve more independently (Kent and Zayed 2013). Several studies have addressed these hypotheses (Kent et al. 2012; Liu et al. 2015, 2017; Wallberg et al. 2015; Rueppell et al. 2016; Jones et al. 2019; Kawakami et al. 2019; Waiker et al. 2021), but so far, none are strongly supported. It is also not clear whether eusociality per se selects for higher recombination rates or whether high rates are a specific feature of eusocial Hymenoptera.

The main characteristics of eusocial insect species are the presence of castes that forgo reproduction in order to care for brood or defend other colony members (workers and soldiers). Eusocial insects are found mainly in Hymenoptera, among bees, wasps and ants, and also Blattodea, in which all termites (Infraclass: Isoptera) are eusocial but exhibit varying levels of social complexity. These insect orders likely diverged ~400 million years ago (Misof et al. 2014), and there are substantial differences between social insects from the two orders (Korb 2008; Korb and Thorne 2017). Hymenoptera belong to the superorder Holometabola, which is the most diverse insect superorder and contains 11 orders, including Lepidoptera (butterflies, moths), Coleoptera (beetles), and Diptera (true flies), which all undergo complete metamorphosis. Blattodea is a hemimetabolous order, which undergoes partial metamorphosis. Hymenoptera have haplodiploid sex determination, in which males develop from haploid, unfertilized eggs and females derive from diploid, fertilized eggs. In Blattodea, both males and females are diploid. Members of the worker caste in eusocial Hymenoptera are all female, whereas eusocial Blattodea workers are of both sexes. Hymenoptera colonies are headed by one or a small number of queens, whereas Blattodea colonies contain both a king and a queen.

The hypotheses proposed to explain the high recombination rates observed in eusocial Hymenoptera predict that recombination should be elevated in all social insects. However, so far it is unknown whether social insects from other orders also have high rates. As described above, there are fundamental differences in the sex determination system and colony compositions of social insects from Hymenoptera and Blattodea. The robustness of the association between eusociality and high recombination rates can therefore be addressed by assessing mean genomic recombination rates in social Blattodea (termites).

In addition to variation between species, the rate of recombination also varies along chromosomes. In mammals, crossover events are mainly restricted to regions that contain specific motifs that act as binding sites for the PR/SET domain 9 (PRDM9) protein that marks regions for double-stranded breaks, which are repaired by recombination during meiosis (Baudat et al. 2010; Myers et al. 2010; Parvanov et al. 2010). Species with an active PRDM9 protein include nearly all mammals and many vertebrate taxa (Baker et al. 2017). In species without PRDM9, recombination can be directed to hotspots defined by other features, such as CpG islands and promoters that have open chromatin. Such hotspots have been characterized in the genomes of diverse taxa, including yeast, birds, and dogs (Axelsson et al. 2012; Berglund et al. 2015; Lam and Keeney 2015; Singhal et al. 2015). In insects, the prevalence of recombination hotspots is not well studied. Fine-scale recombination maps in the fruit fly *Drosophila melanogaster* and honey bee *A. mellifera* indicate a paucity of recombination hotspots (Chan et al. 2012; Smukowski Heil et al. 2015; Wallberg et al. 2015), whereas hotspots appear

to be present in the genome of the butterfly *Leptidea sinapis* (Torres et al. 2023).

Several other factors may also modulate recombination rate along chromosomes, such as germline DNA methylation. Many vertebrate species have elevated recombination in CpG islands, which are usually unmethylated in the germline (Axelsson et al. 2012; Berglund et al. 2015; Singhal et al. 2015). In insect genomes, methylation is not always present but is commonly restricted to gene bodies (Lyko et al. 2010; Wang et al. 2013; Harrison et al. 2018; Bewick et al. 2019; Ventós-Alfonso et al. 2020; Arsalá et al. 2022). In the honey bee *A. mellifera* and solitary bee *M. rotundata*, recombination is reduced in genes, particularly those inferred to have the highest levels of germline methylation (Wallberg et al. 2015; Jones et al. 2019). In addition, some studies have identified correlations between caste biases in gene expression and recombination. In honey bees, genes with worker-biased gene expression have been found to have elevated recombination rates (Kent et al. 2012), but this is likely an indirect result of differences in germline methylation between sets of genes (Wallberg et al. 2015).

Here, we estimate recombination rate and its fine-scale variation across the genomes of two distantly related termite species with contrasting social complexities (Korb and Hartfelder 2008; Korb and Thorne 2017). *Macrotermes bellicosus* (Termitidae: Macrotermitinae) is a fungus-growing termite with a high level of social complexity, with colonies of several millions of individuals and sterile workers. It belongs to the foraging (multiple-pieces nesting) termites, in which workers leave the nest to forage for food. The drywood termite *Cryptotermes secundus* (Kalotermitidae) has a low level of social complexity. Its colonies consist of a few hundred individuals and totipotent workers from which queens and kings develop (Hoffmann et al. 2012). It belongs to the wood-dwelling (one-piece nesting) termites that nest in a single piece of wood that also serves as their food source.

The study has three main aims. First, we aim to test the robustness of the association between sociality and high recombination rate. If sociality is a universal driver of high recombination rates, we would expect both termite species to have elevated recombination rates and that it is particularly elevated in *M. bellicosus*, which has a higher level of social complexity. Second, we aim to determine whether recombination hotspots exist in the two termite genomes and to infer whether active full-length *PRDM9* genes are present in their genomes. Third, we aim to analyze the factors that govern recombination rate variation across the genome, in particular whether it is associated with patterns of methylation and gene expression, as has been reported in other studies.

Results

Genetic variation in two termite species is typical for social insects

We utilized genome assemblies of two termite species: *M. bellicosus* and *C. secundus*. The genome assembly of *M. bellicosus* (Qiu et al. 2023) is 1.3 Gbp in total length across 275 scaffolds with scaffold N50 of 22.4 Mbp. The genome assembly of *C. secundus* (Csec_1.0) (Harrison et al. 2018) is divided into 55,483 scaffolds with a scaffold N50 of 1.2 Mbp and a total length of 1.0 Gbp.

We sequenced 10 unrelated individuals each of *M. bellicosus* and *C. secundus* to a mean depth of 34× and 31× per sample, respectively (Supplemental Table S1). We mapped reads to the appropriate genome assemblies and called variants in each species. In *M. bellicosus*, we called 5.6 million SNPs; in *C. secundus*, 15.1 million SNPs (Table 1). One individual of *C. secundus* (CS_8) was excluded

Table 1. Estimates of levels of genetic variation, effective population size, and recombination rate in two termite species

Species	Assembly length (Mb)	No. of SNPs	θ_W/bp	$\mu/\text{bp/gen}$	N_E	LDhat		LDhelmet	
						ρ/kbp	r (cM/Mb)	ρ/kbp	r (cM/Mb)
<i>M. bellicosus</i>	1139	5,581,662	0.14%	2.7×10^{-10}	1,278,982	1.69	0.033	4.28	0.084
				4.5×10^{-9}	76,876		0.550		1.392
				8.1×10^{-9}	42,898		0.986		2.494
<i>C. secundus</i>	992	15,037,468	0.44%	2.7×10^{-10}	4,080,725	4.85	0.030	16.39	0.100
				4.5×10^{-9}	245,280		0.494		1.671
				8.1×10^{-9}	136,869		0.885		2.994

(θ_W/bp) Watterson's θ per base pair, (ρ/kbp) population recombination parameter, (ρ) per kilobase, ($\mu/\text{bp/gen}$) mutation rate per base pair per generation taken from literature, (N_E) effective population size, and (r [cM/Mb]) recombination rate in centimorgans per megabase. Estimates correspond to *Acyrtosiphon pisum* (2.7×10^{-10}), *Drosophila melanogaster* (4.5×10^{-9}), and *Drosophila pseudoobscura* (8.1×10^{-9}) (Lynch et al. 2023).

owing to a low proportion of mapped reads, poor mapping quality, and strand bias in mapping. Estimates of variation based on θ_W per base pair (Watterson 1975) are 0.14% and 0.44% in *M. bellicosus* and *C. secundus*, respectively, which are broadly comparable to levels of variation in the honey bee *A. mellifera* (0.3%–0.8%) (Wallberg et al. 2014).

We used these estimates of genetic variation to estimate effective population size (N_E) for each of the sampled populations (Table 1). These rely on an estimate of mutation rate, μ , in each species. As there is no estimate of μ for Blattodea, we considered a range of estimates from insects compiled by Lynch et al. (2023) based on experimental assays that includes estimates from Diptera (*Drosophila*) (Keightley et al. 2009, 2014), Hymenoptera (Yang et al. 2015; Liu et al. 2017), and Lepidoptera (Keightley et al. 2015). These estimates range between a minimum of $2.7 \times 10^{-10}/\text{bp}$ per generation in *Acyrtosiphon pisum* (Fazalova and Nevado 2020) to a maximum of $8.1 \times 10^{-9}/\text{bp}$ per generation in *Drosophila pseudoobscura* (Krasovec 2021). The average of multiple estimates in *D. melanogaster* is $4.5 \times 10^{-9}/\text{bp}$ per generation, which is typical of insects (Lynch et al. 2023).

A mutation rate of $\mu = 1.85 \times 10^{-9}/\text{bp}$ per generation has been estimated in the orchid mantis, *Hymenopus coronatus* (Huang et al. 2023) based on inference from levels of interspecific divergence. This species is a member of the order Mantodea, which contains the closest relatives of Blattodea.

Our estimates of N_E based on θ_W and the various mutation rates are about 43,000–1,279,000 in *M. bellicosus*, which has high social complexity, and about 137,000–4,081,000 in *C. secundus*, which has a lower level of social complexity (Table 1). Estimates of N_E assuming the typical mutation rate of $4.5 \times 10^{-9}/\text{bp}$ per generation are 77,000 and 245,000, respectively. Assuming the mutation rate of *H. coronatus* gives an N_E of 187,000 for *M. bellicosus* and 596,000 for *C. secundus*. These estimates are consistent with previous evidence that eusocial species with large colonies and high social complexity tend to have lower effective population sizes, and the values

are broadly comparable to levels of N_E found in social and solitary Hymenoptera (Leffler et al. 2012; Romiguier et al. 2014).

Low average recombination rate in termite genomes

We next examined the decay of linkage disequilibrium (LD) based on the statistic r^2 . Both termite species show more extensive LD compared with *A. mellifera*, which has an extremely high recombination rate (Fig. 1; Wallberg et al. 2015). As N_E is broadly similar in all these species, these differences are expected to largely represent differences in recombination rates and indicate that the two termite species do not exhibit the elevated genome average recombination rate that is observed in honey bees.

We estimated variation in the population recombination rate, ρ , in both data sets using LDhelmet (Chan et al. 2012) and LDhat (Table 1; Auton and McVean 2007). Average rates of ρ/kbp in *M. bellicosus* and *C. secundus* are 1.7 and 4.8, respectively, from LDhat and 4.28 and 16.39, respectively, from LDhelmet. These estimates can be converted to cM/Mb using estimates of N_E , which

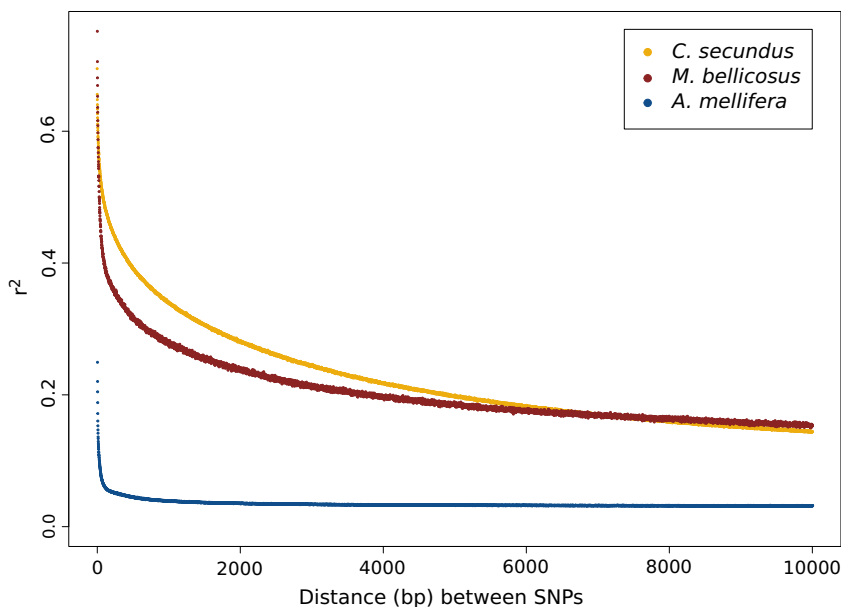


Figure 1. Genome-wide decay of LD measured as r^2 versus distance (base pairs) between SNPs for the two termite species (*M. bellicosus*, dark red; *C. secundus*, yellow) and honey bees (*A. mellifera scutellata*, blue) (data from Wallberg et al. 2017).

gives average recombination rates of 0.033–0.986 cM/Mb in *M. bellicosus* and 0.030–0.885 cM/Mb in *C. secundus* from LDhat and 0.084–2.49 cM/Mb for *M. bellicosus* and 0.1–2.99 cM/Mb for *C. secundus* from LDhelmet. Using a mutation rate close to the average for insects, the recombination rates in *M. bellicosus* and *C. secundus* are 0.550 and 0.494 cM/Mb, respectively, from LDhat and 1.39 and 1.67 cM/Mb, respectively, from LDhelmet. Using the mutation rate from *H. coronatus*, the recombination rates are 0.23 and 0.20 cM/Mb from LDhat and 0.57 and 0.69 cM/Mb from LDhelmet for *M. bellicosus* and *C. secundus*, respectively. Even though these estimates vary depending on the algorithm used for estimating ρ and the assumed mutation rate, they are all substantially lower than estimates from eusocial Hymenoptera (6–25 cM/Mb) (Sirviö et al. 2006) and more similar to rates found in other insects (Wilfert et al. 2007). This represents the first estimation of recombination rate in social insects outside of Hymenoptera and demonstrates that eusociality is not a universal driver of high recombination rates. The Spearman's correlation between the results from LDhat and LDhelmet, using 10 kbp windows, was $\rho=0.946$ and $\rho=0.947$ for *M. bellicosus* and *C. secundus*, respectively, with $P<2.2\times 10^{-16}$ for both species. For the remaining analyses, results from the software LDhelmet were used.

Evidence for active recombination hotspots in termite genomes

Recombination rates are highly variable across the genomes of *M. bellicosus* (Fig. 2A) and *C. secundus* (Fig. 2B). We characterized the distribution of recombination events across the two termite genomes using a cumulative distribution plot (Fig. 3). We found that 50% of the recombination occurs in 0.4% of the genome in *M. bellicosus* and 0.2% of the genome in *C. secundus*. This indicates that the majority of recombination events are restricted to a much smaller portion of the genome than observed in *A. mellifera*, in which 50% of recombination events occur in 32% of the genome (Wallberg et al. 2015).

We further investigated whether recombination hotspots are present in the two termite genomes by looking at the distribution of ρ in windows of 1 kbp, 10 kbp, and 100 kbp across the genome (Supplemental Fig. S1). For all window sizes, a small subset of windows was observed with ρ greatly exceeding the mean in both termite species. For example, the percentage of 1 kbp windows with a recombination rate at least four standard deviations above the genome-wide mean value is 0.2% for *M. bellicosus* and 0.6% for *C. secundus*. In contrast, *A. mellifera* has no windows above this limit for any window size, consistent with a lack of recombination hotspots.

We next defined hotspots as 2 kbp regions with a more than fivefold higher recombination rate than the surround-

ing 100 kbp and used STREME (Bailey 2021) to identify 8- to 20-mers that were enriched in hotspots compared with the rest of the genome. This yielded significantly enriched motifs in both species (four motifs for *M. bellicosus* and one motif for *C. secundus* after compensating for multiple hypothesis testing) (Supplemental Tables S2, S3). However, none of the motifs we identified are present in >5% of hotspots or exhibit more than a twofold difference in frequency of occurrence relative to the background, suggesting they are not viable candidate motifs for promoting recombination in hotspots.

Identification and characterization of the *PRDM9* gene in termites

We searched the *M. bellicosus* and *C. secundus* genomes for evidence of complete *PRDM9* genes (Fig. 4A), which could potentially govern the presence of recombination hotspots in these genomes, as it has been shown to do in mammalian genomes (Baudat et al. 2010; Myers et al. 2010; Parvanov et al. 2010). The *C. secundus* genome contains a *PRDM9* homolog (XP_023708049.2) with annotated KRAB, SSXR, SET, and zinc-finger domains (Fig. 4B). There is no complete homolog of *PRDM9* in the *M. bellicosus* annotation. We therefore used BLAST to search for sequences homologous to the conserved functional domains from *C. secundus*. We identified a putative *PRDM9* ortholog on Scaffold 42 of the *M. bellicosus* genome assembly, containing all domains in the correct order. We

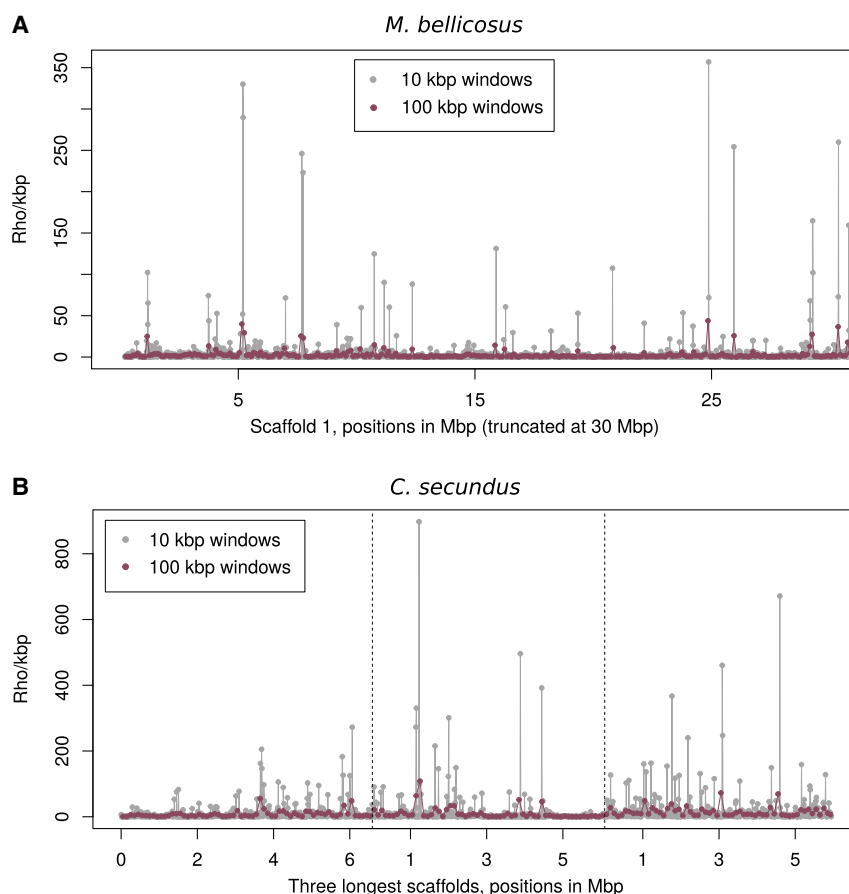


Figure 2. Variation in recombination rates estimated as ρ /kbp across representative genomic scaffolds in *M. bellicosus* (A) and *C. secundus* (B). Estimates of ρ /kbp are calculated in 10 kbp and 100 kbp windows.

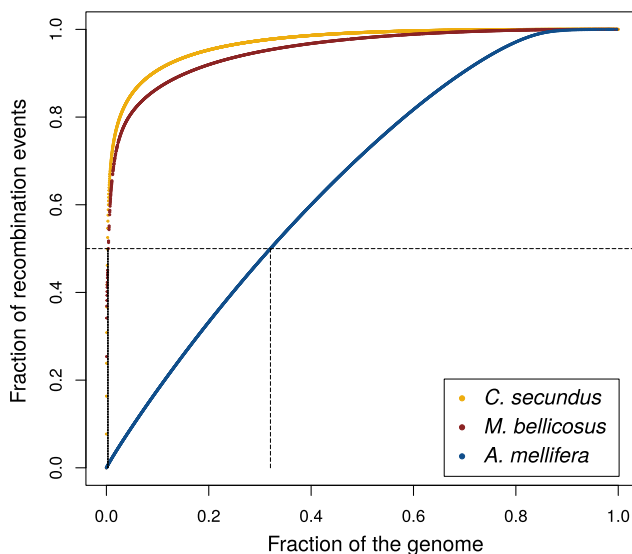


Figure 3. Cumulative plot of proportion of recombination events versus proportion of the genome for the two termite species *M. bellicosus* and *C. secundus* compared with the honey bee *A. mellifera* (Wallberg et al. 2015). Dashed lines show the proportion of the genome in which 50% of the recombination occurs, which for *M. bellicosus* is 0.4%, *C. secundus* is 0.2%, and *A. mellifera* is 32%.

identified retroviral conserved domains in an intron 5' of the zinc-finger domains in this gene, which are not predicted to alter the transcript. We confirmed the presence of a full-length transcript as well as the presence and correct order of conserved domains by alignment of the *Zootermopsis nevadensis* PRDM9 protein sequence to the *M. bellicosus* genome using GeneWise (Fig. 4C; Supplemental Table S4; Madeira et al. 2024).

We used a zinc-finger prediction algorithm (Persikov and Singh 2014) to predict DNA-binding motifs from the zinc-finger domain present in PRDM9 in the *M. bellicosus* and *C. secundus* ge-

nome assemblies. The most probable consensus motifs were TATGGAACGAC AGGAACAACGACATCATCAGCCGCCT AAT and TAATAAGTAGAATCGTTTAG for *M. bellicosus* (Fig. 4D) and *C. secundus* (Fig. 4E), respectively. Base probability matrices for these motifs are shown in Supplemental Table S5. We tested whether the presence of these motifs corresponded to elevated recombination considering 1 kbp around each motif. However, the recombination rate was not significantly elevated around these predicted motifs for either species, either when searching based on the most likely motif or when searching based on the probability matrices.

Genomic correlates of recombination rate

We analyzed genomic correlates of recombination rate in 10 kbp windows. There is a significant correlation between recombination and CpG_{O/E} in both species with Spearman's $\rho=0.24$, $P < 1 \times 10^{-4}$ for *M. bellicosus* and Spearman's $\rho=0.10$, $P < 1 \times 10^{-4}$ for *C. secundus* (Supplemental Fig. S2). A reduction of recombination rates in regions of low CpG content is consistent with interference of germline methylation with recombination. However, as the correlation coefficients are low, the explanatory power is weak. Correlations of similar magnitudes were observed between recombination and GC content (Spearman's $\rho=0.23$ for *M. bellicosus*, $P < 1 \times 10^{-4}$; Spearman's $\rho=0.017$ for *C. secundus*, $P < 1 \times 10^{-4}$). This correlation has been associated with an effect of recombination in driving GC through GC-biased gene conversion, but the relatively low correlation coefficients could indicate that this force is not strong in termites. In contrast, the correlation between GC content and recombination rate is much stronger in the honey bee ($R^2=0.506$), which is likely caused by high recombination rates and intense GC-biased gene conversion (Wallberg et al. 2015).

There is a weak but significant correlation between overall repeat element density and recombination rate in *C. secundus* (Spearman's $\rho=0.02$, $P < 1 \times 10^{-4}$) but not in *M. bellicosus*. There is a weak but significant correlation between simple repeats and

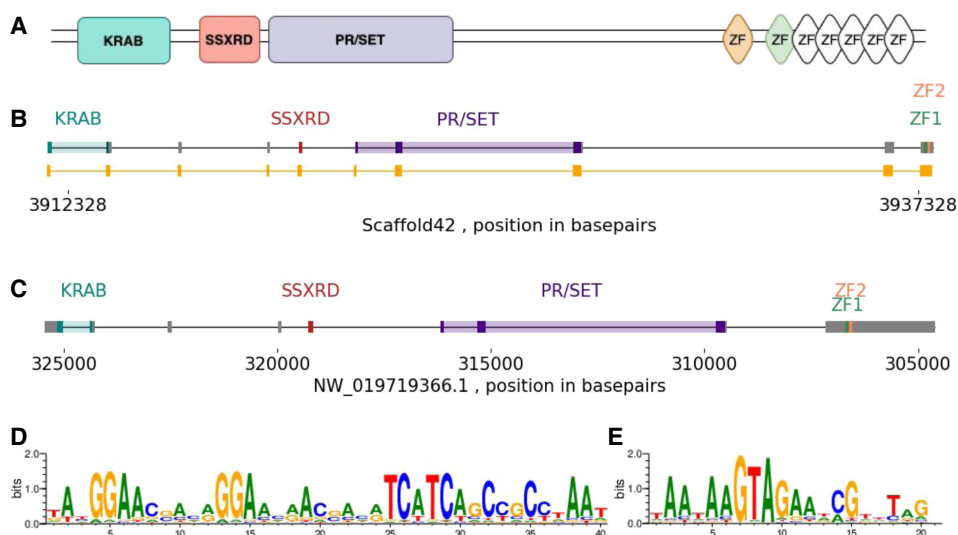


Figure 4. PRDM9 structure and predicted zinc-finger motifs in the *M. bellicosus* and *C. secundus* genomes. (A) Predicted structure of the PRDM9 protein sequence in *C. secundus*. (B) Predicted structure of the PRDM9 gene in *M. bellicosus* identified by homology searches. Exons and conserved functional domains are marked. The mapped exons from *Zootermopsis nevadensis* are also shown. (C) Structure of PRDM9 gene in *C. secundus*. Exons and conserved functional domains are marked. (D) Predicted zinc-finger binding motif in *M. bellicosus*. (E) Predicted zinc-finger binding motif in *C. secundus*.

recombination in both genomes (Spearman's $\rho=0.051$, $P<2.22\times 10^{-16}$ for *M. bellicosus*; Spearman's $\rho=0.055$, $P<2.22\times 10^{-16}$ for *C. secundus*). SINE and LINE elements are also only weakly associated with recombination rate in both genomes (SINEs: Spearman's $\rho=-0.012$, not significant for *M. bellicosus*, and Spearman's $\rho=0.02$, $P<1\times 10^{-4}$ for *C. secundus*; LINEs: Spearman's $\rho=-0.02$, $P<1\times 10^{-4}$ for *M. bellicosus*, and Spearman's $\rho=0.002$ not significant for *C. secundus*). There is a significant negative correlation between gene density and recombination for the *M. bellicosus* genome (Spearman's $\rho=-0.185$, $P<2.22\times 10^{-16}$) but no significant correlation for *C. secundus*. We also tested whether hotspots differ from the background with respect to GC content. The GC content does not differ substantially between hotspots and the rest of the genome in either species (means of 40.1% and 41.0% for *M. bellicosus* and 40.2% and 40.3% for *C. secundus*).

We found a significant correlation between nucleotide diversity (π) and recombination in both species (Supplemental Fig. S2), with Spearman's $\rho=0.48$, $P<1\times 10^{-4}$ for *M. bellicosus* and Spearman's $\rho=0.31$, $P<1\times 10^{-4}$ for *C. secundus*. This correlation is found across a wide range of eukaryotic taxa and is generally accepted to be caused by the interaction between recombination and linked selection (Begun and Aquadro 1992; Nachman 2001).

Recombination and gene expression patterns correlate with CpG_{O/E}

Analysis of the distribution of CpG_{O/E} in both termite species showed that it is bimodally distributed, with substantially lower values in introns and exons compared with flanking noncoding regions (Supplemental Fig. S3). This indicates that germline DNA methylation is biased toward gene bodies in the two termite genomes, a pattern that is also found in honey bees and other insect genomes with functional DNA methylation (Elango et al. 2009; Sarda et al. 2012; Harrison et al. 2018). We found that exons have lower CpG_{O/E} compared with introns, which are both lower than flanking regions (all comparisons $P<0.001$ after Bonferroni correction) (Fig. 5, top row) consistent with higher levels of germline methylation in exons (Supplemental Table S6). We next analyzed how the recombination rate varies among genomic features. We found that ρ /kbp is also significantly reduced in genes compared with flanking regions for *C. secundus*, with exons also show-

ing reduced ρ /kbp compared with introns (all P -values <0.005 after Bonferroni correction) (Fig. 5). The same trends are observed in *M. bellicosus* although they are not significant. The correlation between CpG_{O/E} and recombination rate among genes is higher than for the genome overall (Spearman's $\rho=0.315$, $P<2.22\times 10^{-16}$ for *M. bellicosus*; Spearman's $\rho=0.196$, $P<2.22\times 10^{-16}$ for *C. secundus*).

We next investigated whether patterns of gene expression between castes and sexes (hereafter, "caste-biased gene expression") were correlated with CpG_{O/E} and recombination rate using gene expression data from two studies (Elsner et al. 2018; Lin et al. 2021). It has been proposed that it is evolutionarily advantageous for genes with worker-biased expression to be located in regions with elevated recombination (Kent et al. 2012). Comparisons of CpG_{O/E} and recombination rate for the original gene expression categories reported by Elsner et al. (2018) for *M. bellicosus* are shown in Supplemental Figure S4. We also reclassified the genes into the following expression categories: queen-biased, king-biased, worked-biased, male-biased, female-biased, reproduction-biased, and differentially expressed (for details, see Methods). For *C. secundus*, only worker-biased and queen-biased genes were identified (Lin et al. 2021). We find that mean CpG_{O/E} in gene bodies shows significant variation between gene expression categories in both *M. bellicosus* and *C. secundus* (Fig. 6). However, this variation is not consistent between species. In *M. bellicosus*, queen-biased and female-biased genes have low CpG_{O/E}, and king-biased and male-biased genes have high CpG_{O/E}. In *C. secundus*, queen-biased genes have higher than average CpG_{O/E}.

The differences in CpG_{O/E} between caste-biased gene expression categories are mirrored by differences in ρ /kbp (Fig. 6). Gene expression categories with elevated CpG_{O/E} consistently show elevated ρ /kbp, and this pattern is found in both species. Studies in the honey bee have found that worker-biased genes have lower CpG and higher recombination rates (Kent et al. 2012; Wallberg et al. 2015). In *M. bellicosus*, we find that worker-biased genes have slightly elevated CpG_{O/E} and ρ /kbp. However, in *C. secundus*, these values are reduced in worker-biased genes, whereas queen-biased genes have a slightly elevated CpG_{O/E} and ρ /kbp in this species.

We investigated the relative roles of CpG_{O/E} and gene expression in determining ρ in genes using linear models, with CpG_{O/E}/ ρ

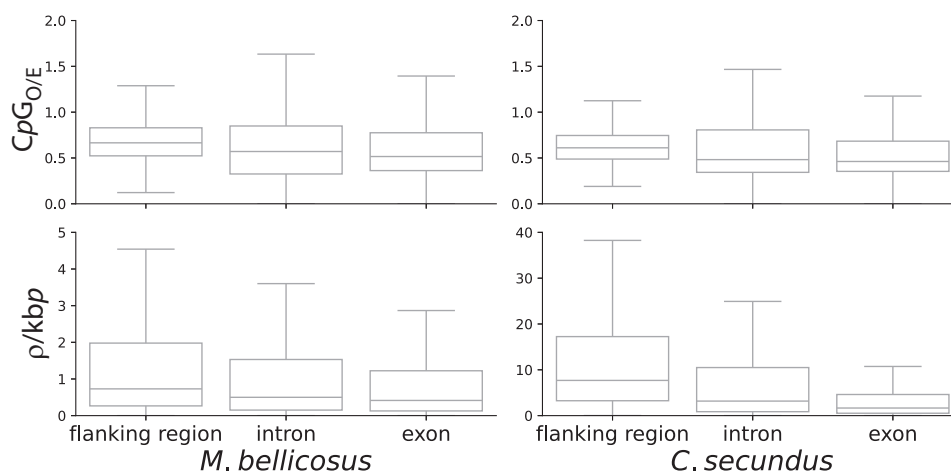


Figure 5. Boxplots showing variation in levels of CpG_{O/E} and recombination rate (ρ /kbp; rho) in genic and flanking regions in the *M. bellicosus* and *C. secundus* genomes. The differences between all categories are significant.

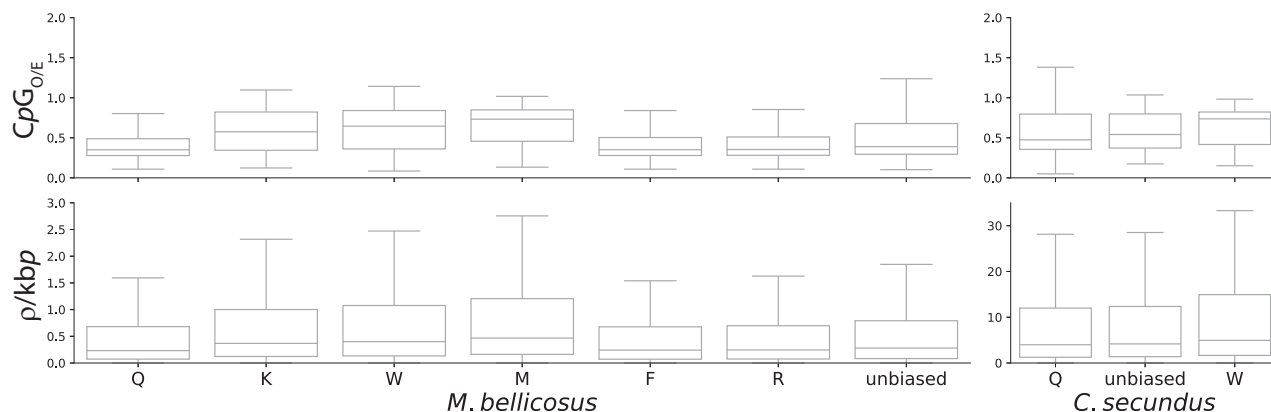


Figure 6. Boxplots showing variation in levels of $\text{CpG}_{\text{O/E}}$ and recombination rate (ρ/kbp) in the coding region of genes in the *M. bellicosus* and *C. secundus* genomes classified according to their patterns of gene expression: (Q) queen-biased, (K) king-biased, (W) worker-biased, (F) female-biased, and (R) reproductive-biased.

in flanking region and expression categories as explanatory variables. In both *M. bellicosus* and *C. secundus*, we found that $\text{CpG}_{\text{O/E}}$ was the strongest predictor of recombination rate variation among genes (Supplemental Tables S7, S8) in line with the analyses above. We found that caste-biased expression had no significant impact on the recombination rate. $\text{CpG}_{\text{O/E}}$ is a better predictor of ρ/kbp in genes than ρ/kbp in the flanking region, and variation in ρ/kbp among expression categories mainly reflects differences in $\text{CpG}_{\text{O/E}}$.

Discussion

The main findings we present here are (1) two distantly related termite species with varying social complexity both have relatively low genomic average rates of meiotic recombination; (2) both genomes possess recombination hotspots and likely contain full-length copies of *PRDM9*; and (3) there are reduced levels of recombination in genomic regions depleted in CpG sites. Our findings contrast with the extremely elevated recombination rates observed in eusocial Hymenoptera, which are likely to be generated by a feature that is specific to these taxa. The finding of recombination hotspots is one of the first in insects. This could reflect a conserved function of *PRDM9* in initiating recombination events in both vertebrates and invertebrates. Our results also support a role for germline methylation in suppressing recombination.

Low recombination rates in termites

Data from a wide range of taxa indicate that at least one crossover per chromosome tetrad is essential for correct chromosomal segregation during meiosis (Fernandes et al. 2018). The chromosome numbers of *M. bellicosus* and *C. secundus* are $2n = 42$ and $2n = 40$, respectively (Jankásek et al. 2021). Considering their assembly lengths, this indicates that a minimum average genomic recombination rate of 0.92 cM/Mb and 1.01 cM/Mb, respectively, would be necessary for accurate meiosis to occur in the two termite species. Our estimates of recombination rate per meiosis based on analysis of LD depend on several factors, including the algorithm to estimate ρ and the estimate of mutation rate. As no estimates of the mutation rate in termites were available, we considered a range of estimates from other insects. These estimates vary by more than an order of magnitude between different insect species (Lynch et al. 2023), and higher estimates of the mutation rate

give higher estimates of the recombination rate. Our estimates based on LDhat are all <1 cM/Mb for both species, even when considering the highest estimate of mutation rate. Using estimates from LDhelmet and the average of experimentally determined mutation rate estimates in insects gives estimates of recombination rate of 1–2 cM/Mb. The species most closely related to termites for which a mutation rate estimate was available was *H. coronatus* (Huang et al. 2023). Using this estimate and ρ from LDhelmet, the recombination rate per meiosis was estimated to 0.57 cM/Mb for *M. bellicosus* and 0.69 cM/Mb for *C. secundus*, namely, both slightly lower than expected based on the requirement of one crossover per chromosome. Taken together, these results indicate that recombination rates in these two termite species are substantially lower than the elevated values found in eusocial Hymenoptera (6–25 cM/Mb) (Wilfert et al. 2007).

It should be noted that our estimates of recombination rate are sex-averaged rates. It is possible that recombination is absent in one of the termite sexes, which would make a lower sex-averaged recombination rate more feasible. In addition, sex-linked contigs have not been identified in the termite genome assemblies utilized here. *M. bellicosus* and *C. secundus* have X1X2/Y1Y2 and X/Y sex-determining systems, respectively (Jankásek et al. 2021). The inclusion of sex chromosomes in the analysis leads to errors in determining the genome-wide recombination rate per meiosis because they have a lower N_E than the rest of the genome. In addition, sex chromosomes are more prone to assembly and read mapping errors, owing to reads aligning with the incorrect copy of the chromosome. However, considering that sex chromosomes only comprise a minor portion of the genome, this is unlikely to substantially alter our results. It has also been observed that some chromosomes form ring or chain structures during meiosis in certain termite species (Bergamaschi et al. 2007), which could plausibly relax the requirement for one crossover per chromosome per meiosis.

Eusociality is not a universal driver of high rates of meiotic recombination

Elevated genomic rates of recombination are characteristic of all eusocial Hymenoptera so far investigated. Several hypotheses have been advanced to explain this phenomenon, which are all based on an evolutionary advantage of recombination in eusocial species (Wilfert et al. 2007). For instance, it has been suggested that

elevated recombination could increase genetic diversity in a colony. This could promote diversity in the workforce, enabling more efficient task specialization among workers (Kent et al. 2012). Alternatively, it could render the colony less susceptible to invasion by parasites and pathogens (Fischer and Schmid-Hempel 2005). Elevated recombination might also favor rapid evolution of caste-specific expressed genes (Kent and Zayed 2013). In addition, increased recombination could lead to reduced variance in relatedness, which could prevent kin conflict in colonies (Sherman 1979; Templeton 1979).

All these hypotheses predict that recombination rate should be elevated in all eusocial insects. Here, however, in the first report of recombination rate in social insects outside of Hymenoptera, we find overall low levels of recombination. Both termite species studied here are eusocial, and in particular, *M. bellicosus* has extremely large and complex colonies characteristic of eusociality (Korb and Thorne 2017). The low recombination rate inferred in the two termite genomes prompts re-evaluation of hypotheses connecting eusociality with high recombination rates. Although both termites and social Hymenoptera share many traits, they differ in others that reflect their different ancestries.

The chromosome numbers of both termite species (*M. bellicosus*, $2n = 42$; *C. secundus*, $2n = 40$) are typical for termites, although variation exists between taxa (Jankásek et al. 2021). These values do not differ from those observed in Hymenoptera, which have a similar range (e.g., *A. mellifera*, $2n = 32$) and do not vary between solitary and eusocial taxa (Ross et al. 2015; Cardoso et al. 2018; Cunha et al. 2021). This suggests that differences in chromosome number are unlikely to be relevant in explaining the differences in recombination rates among these taxa.

One proposed advantage of sex and recombination is that it is associated with sexual selection and a higher intensity of selection on males. This could be advantageous because differential male mating success can reduce mutational load in sexual populations by causing deleterious mutations to segregate at a lower equilibrium frequency on average (Agrawal 2001; Siller 2001). Recombination is essential for effective purging of deleterious alleles as it unlinks deleterious variants from nearby beneficial ones, thus preventing selection interference (Hartfield and Keightley 2012). Selection has been found to be stronger in males than in females across animal species (Janicke et al. 2016; Winkler et al. 2021). Species with a male-biased sex ratio are expected to experience particularly intense selection on males. It is therefore possible that the recombination rate could be particularly elevated in species in which males face intense competition as this increases the evolutionary advantage of recombination.

Another proposed advantage of sex and recombination in general is that they are associated with the intensity of sperm competition. Maynard Smith (1976) presented a model that explains how sib competition can produce an immediate evolutionary advantage to sex and recombination, which occurs when several offspring of a single female compete within a patch. This model was extended by Manning and Chamberlain (1997), who argued that it is analogous to intra-ejaculate competition among sperm, in which recombination increases the variance of fitness. Under these conditions, if inter-ejaculate competition also exists, it is predicted that there will be selection for increased recombination rate. This model predicts a significant correlation between recombination rate and gamete redundancy (excess sperm production), which is observed in mammals (Manning and Chamberlain 1997).

Many social Hymenoptera produce highly male-biased reproductive offspring, and thousands of haploid male drones compete

to fertilize a small number of female queens (Winston 1991; Baer 2005; Koeniger et al. 2011). This situation is similar to sperm competition as haploid males can be considered analogous to gametes. However, as many more genes are expressed in adult haploid males compared with gametes, there are more targets of selection. Selection against deleterious alleles is particularly effective in haploid males because recessive deleterious alleles are exposed (Hedrick and Parker 1997). It is therefore possible that the male-biased sex ratio observed in many eusocial Hymenoptera could favor elevated recombination rates in the species so far studied, owing to elevated competition between males.

Male-biased sex ratios are particularly strong in the genus *Apis*, which is also the genus in which the most extreme recombination rates have been inferred (Beye et al. 2006; Shi et al. 2013; Liu et al. 2015; Wallberg et al. 2015; Rueppell et al. 2016; Kawakami et al. 2019). Male-biased sex ratios are also observed in other social bees and wasps. Both male- and female-biased operational sex ratios have been observed in ants but are difficult to measure because workers invest more resources in female offspring and may manipulate sex ratios. In nonsocial Hymenoptera, females lay roughly equal numbers of haploid and diploid eggs (Trivers and Hare 1976). The sex ratio of reproductive males and females is also relatively even in most termite species (Roisin 2001).

Further research is needed to determine the cause of variation in recombination rates among social and nonsocial insects. For example, simulations could be used to determine the effect of selection on haploid males on recombination rates, considering different sex ratios. In addition, it is necessary to estimate recombination rate in a wider range of social and nonsocial species to determine whether the strength of selection on males is a key factor in determining these rates.

Full-length *PRDM9* gene and recombination hotspots in termite genomes

Our results are also indicative of the presence of recombination hotspots in the two termite genomes. The fine-scale distribution of recombination rate in both termite genomes shows that 50% of the recombination happens in <0.5% of the genome, which is even more extreme than for humans, in whom this corresponds to ~6% of the genome (Myers et al. 2005). This contrasts to the honey bee and solitary bees, in which fine-scale maps are available and recombination is relatively uniformly distributed in the genome (Wallberg et al. 2015; Jones et al. 2019).

Two main mechanisms have been shown to lead to recombination hotspots. The first, which is best understood in humans and the mouse, is governed by the *PRDM9* protein. A zinc-finger domain recognizes a specific DNA sequence motif and then performs a histone modification in the vicinity, which marks the sequence for a DNA double-stranded break that is repaired by recombination during meiosis (Baudat et al. 2010; Myers et al. 2010; Parvanov et al. 2010). In species with this mechanism, both the *PRDM9* zinc-finger domain and the sequences present in hotspots are fast-evolving, which results in a rapid turnover of hotspot locations (Myers et al. 2010). Full-length *PRDM9* orthologs are present among a range of highly diverged vertebrates, including many fish, reptiles, and mammals, but many full or partial losses of the gene have also been inferred. In a study of 225 vertebrate genomes, a minimum of six partial and three complete losses were inferred (Baker et al. 2017). However, it is unclear whether *PRDM9* hotspots are common in invertebrates.

A second mechanism is found in vertebrates that lack PRDM9, in which recombination hotspots are localized to regions of open chromatin such as CpG islands and promoters. This is found in dogs and in most or all birds, in which PRDM9 has been lost, and results in hotspots with stable rather than rapidly evolving locations (Axelsson et al. 2012; Singhal et al. 2015). A similar pattern is also observed in the yeast *Saccharomyces cerevisiae* in which recombination preferentially occurs in promoter regions (Wu and Lichten 1994). PRDM9-independent hotspots can be identified in CpG islands in some vertebrate genomes even when PRDM9 is present (Joseph et al. 2024). In invertebrates such as *Drosophila* and *Caenorhabditis elegans*, recombination hotspots appear to be absent (Coop and Przeworski 2007; Chan et al. 2012; Smukowski Heil et al. 2015). In insects, fine-scale maps of social and solitary bees have not revealed recombination hotspots (Wallberg et al. 2015; Jones et al. 2019), but there is evidence for hotspots in a butterfly genome (Torres et al. 2023).

We find evidence that recombination is directed toward hotspots in both termite species. We also identify a full-length PRDM9 gene in both species, containing all of the domains needed to initiate recombination. However, we are unable to demonstrate that PRDM9 is responsible for the presence of recombination hotspots in the two genomes. We used a motif prediction algorithm to predict the zinc-finger binding motif in *M. bellicosus* and *C. secundus*, but we did not observe elevated recombination rate around instances of the motif in either of the genomes. Furthermore, we did not identify any common sequence motifs that were strongly enriched in recombination hotspots in either of the termite genomes. A plausible interpretation of these observations is that the PRDM9 protein indeed initiates recombination in hotspots in both species but that positive selection leads to frequent shifts in the target motifs, which results in a lack of a strong association between a specific motif and LD-based hotspots (Myers et al. 2010). It is also possible that the sequences of the zinc-finger motifs are not correctly represented in the genome assemblies, which could result from problems with assembly around minisatellites. However, an unknown mechanism could also be responsible for hotspots in termites.

Evidence that germline DNA methylation suppresses recombination

In insects with a functioning DNA methyltransferase machinery, the CpG_{O/E} statistic shows substantial variation along the genome, which mainly reflects variation in levels of germline DNA methylation (Elango et al. 2009). In contrast to vertebrate genomes, methylation is strongly skewed toward gene bodies in the genomes of insects (Glastad et al. 2011). In the two termite genomes studied here, we find lowest values of CpG_{O/E} in exons, whereas noncoding flanking regions are biased toward higher values, indicative of lower levels of DNA methylation (Fig. 5; Supplemental Fig. S3). Exons, introns, and noncoding flanking regions all display a predominantly bimodal distribution of CpG_{O/E}, likely reflective of the presence of methylated and unmethylated regions. We find that recombination rate is reduced in gene bodies compared with flanking regions in one of the two termite genomes, but the overall genomic correlation between CpG_{O/E} and recombination rate, although significant, is weak. This could reflect the fact that the majority of the genomes are not methylated, so that much of the variation in CpG_{O/E} on the genome scale is not strongly influenced by methylation. Our inference of reduced recombination rates in exons could also be influenced by lower N_E

in these regions owing to the effects of linked selection, which could result in more extensive LD (Charlesworth 2009). However, the observation that differences in CpG_{O/E} among genes are associated with recombination rate indicates that linked selection is unlikely to be a major factor driving the observed differences in recombination rate and that germline DNA methylation is more important.

We identified CpG_{O/E} as the factor with the strongest influence on recombination rate variation among genes. There were no significant effects of differences in gene expression patterns across castes, a result that was also observed in the honey bee genome (Wallberg et al. 2015). This indicates that recombination rate variation is not modulated by differences in gene expression in these genomes, which might be expected if natural selection favors increased recombination rates in certain genes. For example, it has been proposed that natural selection could favor elevated recombination rates in genes with roles in worker behavior or in immune function owing to their important roles in colony function (Wilfert et al. 2007; Kent et al. 2012). However, the results presented here suggest that differences in recombination rate between genes with caste-biased patterns of gene expression are mainly because of underlying differences in methylation levels among genes.

A direct effect of germline DNA methylation in suppressing recombination could influence variation in recombination landscapes among vertebrates and invertebrates. Vertebrate genomes are usually highly methylated with the exception of CpG islands. In vertebrate genomes that lack a functional PRDM9 gene, recombination events tend to localize to CpG islands (e.g., in dogs and birds) (Axelsson et al. 2012; Singhal et al. 2015). The effect of DNA methylation on recombination rate variation in invertebrates is less clear, but the results presented here also support a role of germline methylation in suppressing recombination.

Methods

Sample collection

M. bellicosus samples were collected from colonies in Kakpin, next to the Comoé National Park in Côte d'Ivoire (coordinates 8°39'N 3°46'W) (Elsner et al. 2018). *C. secundus* colonies were collected from dead *Cerriops tagal* mangrove trees near Palmerston-Channel Island (Darwin Harbor, Northern Territory, Australia; 12°30'S 131°00'E). Colonies were then kept in *Pinus radiata* wood blocks in climate rooms in Germany, at 28°C, 70% relative humidity, and a 12 h day/night cycle (Lin et al. 2021). For each species, each of the 10 collected samples came from a different colony.

Population sequencing

The termites were cut in two, approximately between head and thorax. DNA was extracted from both parts using the Qiagen blood and tissue kit following the standard protocol, including treatment with 4 μ L RNase and 25 μ L Proteinase K. The DNA extraction product was cleaned and concentrated with ZymoResearch genomic DNA Clean & Concentrator kit. For most samples, the RNase treatment was repeated prior to cleaning and concentrating. One DNA sample from each individual was chosen, based on the DNA amount and quality, to be sequenced. Sequencing libraries were produced using the Illumina DNA PCR-free library preparation kit according to the manufacturer's protocol, using an input DNA quantity of 200 ng per sample. This protocol yields insert sizes of ~350 bp. We performed Illumina short-read sequencing on the libraries using NovaSeq 6000 S4 flow cell with 2 \times 150 bp paired-end reads.

Read mapping and variant calling

The sequence data from the two termite species were mapped to the reference genomes for *M. bellicosus* (Qiu et al. 2023) and *C. secundus* (Csec_1.0) (Harrison et al. 2018) respectively, using the Burrows–Wheeler alignment tool BWA version 0.7.17 (Li and Durbin 2009) with the BWA-MEM algorithm. Sorting and indexing of the BAM-files was done with the SAMtools package version 1.17 (Li et al. 2009), followed by adding read groups and marking duplicate reads with Picard toolkit version 1.118 (<http://broadinstitute.github.io/picard/>). After mapping, we excluded all scaffolds <1 kbp from the *C. secundus* genome, which in total correspond to 2% of the genome size, as they are potentially of lower quality and not informative for analysis of LD or recombination. For *M. bellicosus*, there are no scaffolds <1 kbp, but we excluded the following five scaffolds, which together correspond to 3.78% of the genome length, owing to unexpectedly high heterozygosity: 93, 62, 41, 35, and 32.

One of the *C. secundus* individuals did not map well to the reference genome (71% mapped reads, unequal distribution between forward and reverse strand, and an average mapping quality of 35, whereas all the other samples had an average mapping quality greater than 45). This individual was excluded from all further analyses, and hence, the total number of termite individuals analyzed in this study is 19: 10 *M. bellicosus* and nine *C. secundus*.

Variant calling was done using the tools HaplotypeCaller, GenomicsDBImport, and GenotypeGVCFs from GATK version 4.3.0.0 as described in their best-practices workflow (<https://gatk.broadinstitute.org/hc/en-us/articles/360035535932-Germline-short-variant-discovery-SNPs-Indels->, last accessed 2023-07-11). The variants were filtered in multiple steps, first removing extended regions of low mapping quality or deviating read depth as those regions are likely to be unreliable. This was done based on statistics from the SAMtools mpileup and depth tools, respectively, removing 100 kbp windows with mean mapping quality below 70 or mean read depth more than two standard deviations from the genome-wide mean. Indels were removed using VCFtools version 0.1.16 (Danecek et al. 2011) before GATK VariantFiltration was run with the following limits, which are all equal to or stricter than the ones recommended (<https://gatk.broadinstitute.org/hc/en-us/articles/360035890471-Hard-filtering-germline-short-variants>, last accessed March 11, 2023) and adapted to the observed distributions of the corresponding statistics: QD < 2; FS > 50; MQ < 40; ReadPosRankSum < -4; ReadPosRankSum > 4; SOR > 3; ExcessHet > 5 for both species, as well as MQRankSum with limits -5 and 5 for *C. secundus* and limits -6 and 4 for *M. bellicosus*. This was followed by additional filtering with VCFtools to select only biallelic sites with a minimum quality score of 30 and maximum of 40% missing genotypes. The SNPs with mean read depth in the lowest or highest 2.5% of the depth distribution were also filtered out, which translates to a depth below 10 or above 34 for *C. secundus* and below five or above 40 for *M. bellicosus*. For the recombination rate calculations, the SNPs were also filtered for a minor allele count of two or greater, as rare variants appearing on only one chromosome have a relatively high false-positive rate and are not informative regarding LD. The average proportion of missing genotypes per individual was 4.5×10^{-5} for *C. secundus* and 1.2×10^{-4} for *M. bellicosus*. After variant calling and filtering, the data were phased and imputed using Beagle version 5.1 (Browning and Browning 2007; Browning et al. 2018).

Estimation of LD decay

The decay of LD, r^2 , with increasing physical distance was estimated with the software PopLDDecay version 3.42 (Zhang et al. 2019) over a maximum distance of 10 kbp. For this analysis, sequence

data from *A. mellifera scutellata* were included for comparison (Wallberg et al. 2017). The *A. mellifera scutellata* sequences were mapped to the Amel_HAv3.1 reference genome (Wallberg et al. 2019) and processed and filtered in the same way as described above for the termite data (GATK VariantFiltration with the limits QD < 2, FS > 25, MQ < 40, MQRankSum < -4, MQRankSum > 4, ReadPosRankSum < -4, ReadPosRankSum > 4, SOR > 3, ExcessHet > 5, and VCFtools to select biallelic sites with mean read depth between five and 13, quality score of 30 or more, minor allele count of two or more, and $\leq 40\%$ missing genotypes). For *M. bellicosus*, the reference genome contains some gaps of unknown size, and to avoid estimating r^2 across those gaps, scaffold breaks were introduced at the corresponding positions.

Estimation of population recombination rate

We estimated the population-scaled recombination rate, ρ , using two methods, LDhat (Auton and McVean 2007) and LDhelmet (Chan et al. 2012), which are both based on a Bayesian reversible-jump Markov chain Monte Carlo algorithm (rjMCMC).

For LDhat (<https://github.com/auton1/Ldhat>; downloaded April 19, 2023), the input files were generated from the filtered vcf-file with VCFtools and the `--ldhat` option. Then, the LDhat function complete was run in order to create a lookup table, with the estimated value of θ_W (see below) and a maximum ρ -value of 100 and 101 grid points, as recommended. This was followed by the main function, interval, to perform the rjMCMC with 10 million iterations, sampling every 5000th iteration and a block penalty of one as this has been used previously with LDhat in similar studies (Wallberg et al. 2015; Jones et al. 2019). To summarize the output from interval, the stat function was used, discarding the first 20 samples as burn-in.

The recombination rate was also estimated using the software LDhelmet version 1.10 (Chan et al. 2012). To prepare the input files for LDhelmet, the filtered VCF files were converted into multiple-sequence FASTA files using the `vcf2fasta` function from `vcflib` version 2017-04-04 (Garrison et al. 2022) and to the related file formats SNPs and POS with the `--ldhelmet` option from VCFtools (version 0.1.16). The first step of the LDhelmet workflow is to create haplotype configuration files with the `find_confs` tool, followed by generation of likelihood lookup tables with the tool `table_gen` and computation of Padé coefficients with the tool `pade`. These tools were run with the recommended parameters, meaning that `find_confs` was run with a window size of 50 SNPs, `table_gen` was run with a grid of ρ -values ranging from zero to 100 with increments of 0.1 up to 10 and increments of one for the remaining grid, and `pade` was run to generate 11 Padé coefficients. Scaffolds shorter than or equal to the window size of 50 SNPs were removed, corresponding to 1% of the sequence length in *C. secundus* and even less for *M. bellicosus*. The tool `rjmc`, which contains the main algorithm, was run with a burn-in of 100,000 followed by 1,000,000 iterations with a block penalty of 50 as recommended (Chan et al. 2012). For each species, the `rjmc` tool was run three times with different seeds, and the convergence between the replicates was evaluated in terms of Spearman's rank correlation coefficient, which was > 0.95 with $P < 2.2 \times 10^{-16}$ for all pairwise comparisons of both species. The binary output from `rjmc` was converted to text, extracting the mean ρ -value between each pair of SNPs, using the tool `post_to_text`.

The outputs from LDhat and LDhelmet give the value of the parameter ρ between each pair of consecutive SNPs. Those values were converted to equally sized windows across each scaffold, as well as the coordinates of individual genomic features, by taking a weighted average of values from overlapping intervals. The *M.*

bellicosus scaffolds contain a few gaps of unknown size, and the values between SNPs across such gaps were removed.

The parameter ρ is proportional to the recombination rate r and the effective population size N_E , which in turn can be estimated based on the mutation rate μ and the number of segregating sites, K . To convert ρ to r , the following equations were used, where n is the number of haploid sequences. The number of segregating sites (K) was calculated before the filter on minor allele count or two or more was applied, as this filter likely removes many true SNPs as well, which could affect the estimates.

θ_W was estimated for each sample using the equation below from Watterson (1975):

$$\theta_W = K \left/ \sum_{i=1}^{n-1} \frac{1}{i} \right.$$

We used the value of θ_W to estimate N_E using

$$N_E = \frac{\theta_W}{4\mu}.$$

We estimated N_E using a range of mutation rates estimated from other insect species (Keightley et al. 2009, 2014, 2015; Yang et al. 2015; Liu et al. 2017; Lynch et al. 2023). We then estimated the recombination rate using the following equation, assuming a range of values of N_E :

$$r = \frac{\rho}{4N_E}.$$

Cumulative plot of ρ /kbp

A cumulative plot of the proportion of recombination events versus the proportion of the physical distance along the genome was constructed based on estimates of ρ /bp for the two termite species and previous estimates for *A. mellifera* (obtained from the NCBI BioProject database [<https://www.ncbi.nlm.nih.gov/bioproject/>] under accession number PRJNA236426) (Wallberg et al. 2015). The ρ /bp estimates between markers were placed in decreasing order before multiplication with the respective physical distances between the markers. Then cumulative sums were calculated for the ρ -values, as well as the corresponding physical distances, and divided by the total genome-wide sums.

k -mer enrichment in recombination hotspots

To identify k -mers associated with elevated recombination rate, we defined recombination hotspots as 2 kbp segments with at least a fivefold elevated recombination rate compared with the local background, as defined by 50 kbp flanking upstream of and downstream from the segment, resulting in 39,837 regions for *M. bellicosus* and 40,894 regions for *C. secundus*.

Then, we used these hotspots to search for motifs associated with elevated recombination rate by comparing them against an equivalent number of randomly chosen 2 kbp regions from the rest of the genome. This was done using STREME 5.5.4 (Bailey 2021) using a minimum and maximum size of k -mer of eight and 20 (--minw 8; --maxw 20) and otherwise default options. From all motifs with significant enrichment after adjusting for multiple testing, we then considered only motifs present in >5% of all hotspots and an at least twofold enrichment compared with the background as credible candidates.

Identifying *PRDM9* orthologs in *M. bellicosus* and *C. secundus*

A full-length copy of *PRDM9* is present in the *C. secundus* genome annotation but not in *M. bellicosus*. Using OrthoFinder (Emms and Kelly 2019) and the genome annotation files for both species, we

identified a single ortholog to the *C. secundus PRDM9* in *M. bellicosus*. We analyzed this region using the NCBI conserved domain search (Wang et al. 2023). We also searched the *M. bellicosus* assembly for the PR/SET domains derived from *C. secundus* and *Z. nevadensis* using BLAST (Camacho et al. 2009) and subsequently for all other *PRDM9* domains (KRAB, SSXRD, ZF-Casette) to identify hits that contain all other *PRDM9* domains in the vicinity in the correct order. We used GeneWise (Madeira et al. 2024) with default options to perform gapped alignment of a the *PRDM9* protein sequence from *Z. nevadensis* against the target region to predict exon–intron borders.

Predicting binding sites from zinc-finger motifs

Using the zinc-finger motifs from *C. secundus*, we obtained letter probability matrices for predicted binding motifs using a web-server (Persikov and Singh 2014; <http://zf.princeton.edu/index.php>), which were used to search for binding sites across the genome using FIMO (Grant et al. 2011) with default options. Subsequently, the mean ρ for the surrounding 1 kbp bin was used to test for a difference in mean ρ between bins associated with a suspected binding site, as well as a similar number of randomly chosen bins without association. For significance testing, we repeated this draw 10,000 times to compare the associated bins against the 95% confidence interval of the resulting distribution.

Analysis of genomic correlates of recombination

We divided the genome into 10 kbp windows in which the average ρ and other genomic features were estimated. Windows <5 kbp (half of the specified window size), which appear in short scaffolds or at the ends at longer scaffolds, were excluded from the analyses (excluding <1% of the sequence for each species). Correlation coefficients were estimated with Spearman's ρ , with significance estimated by permutation tests with 10,000 iterations.

To determine the repeat content across the *M. bellicosus* and *C. secundus* genome assemblies, we first generated custom repeat libraries for each species with RepeatModeler (Flynn et al. 2020) using default options. This was then combined with Rebase 29 (Bao et al. 2015) and Dfam 3.8 (Storer et al. 2021) and used to run RepeatMasker with default options on each genome. We estimated the proportion of sequence in each repeat class in 10 kbp windows and used Spearman's rank correlation to correlate this with estimates of recombination rate.

Per gene recombination rate and CpG_{O/E} ratio

The observed/expected frequency of CpG sites (CpG_{O/E}) and mean ρ were calculated on a per-gene basis, as well as for exons and introns and 50 kbp flanking regions 10 kbp upstream of and downstream from the gene, akin to the method of Wallberg et al. (2015). For recombination rate, estimates of ρ between markers were used, utilizing a weighted mean when the element spanned multiple markers. For visualization, extreme CpG_{O/E} values owing to low GC content were set to a maximum of four.

We tested for significant difference in mean between ρ and CpG_{O/E} in flanking regions, exons, and introns using paired t -tests corrected for multiple testing using a Bonferroni threshold.

Differential expression data

We classified genes in both species according to their caste-biased patterns of expression from two studies (Elsner et al. 2018; Lin et al. 2021). Elsner et al. (2018) analyzed four castes of *M. bellicosus*, with two age classes of each: major and minor workers, queens, and

kings. They analyzed differences in gene expression among these classes using RNA-seq data. We reclassified the differential expression data by grouping the direct caste-versus-caste comparisons as “queen-biased,” “king-biased,” “worker-biased,” “reproduction-biased,” “male-biased,” or “female-biased” if they were differentially upregulated for this caste (or group) in comparison to other castes (or groups). Because Elsner et al. (2018) initially mapped the *M. bellicosus* reads against the *Macrotermes natalensis* genome, we used OrthoFinder (Emms and Kelly 2019) to identify the *M. bellicosus* orthologs corresponding to the differentially expressed genes. Only single-copy genes were considered for this analysis.

Lin et al. (2021) performed RNA-seq on samples of workers and queens. Here, we used the data from untreated queens and workers. Differentially expressed *C. secundus* genes were classified as “worker-biased” or “queen-biased” if they were differentially expressed with a bias to the respective caste, or classified as unbiased if they were not, and were subsequently compared between classes. The lists of genes we identified in the caste-biased categories in both species are provided as Supplemental Tables S9 and S10.

Correlates of recombination rate variation among genes

To assess the relative importance of CpG_{O/E} and differential expression patterns on recombination rate in genes (genic ρ /kbp), two ordinary least squares models were built using statsmodels (Seabold and Perktold 2010) for each termite species. Genic CpG_{O/E}, differential expression categories, and flanking region ρ /kbp were used as exogenous variables. In the first model, we included flanking region ρ /kbp and CpG_{O/E}. In the second model, we included all available biased expression categories for each species.

Data access

The Illumina whole-genome sequencing data generated in this study have been submitted to the NCBI BioProject database (<https://www.ncbi.nlm.nih.gov/bioproject/>) under accession number PRJNA1021607. Custom scripts are available on GitHub (<https://github.com/troe27/termite-recombination-rate>) and as Supplemental Code.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We acknowledge grant 2018-03896 from The Swedish Research Council to M.T.W. and from the Deutsche Forschungsgemeinschaft (DFG) KO-1895/26-1: 417980976 to J.K. We acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation, and the Swedish Research Council. The computations were enabled by resources in project NAISS 2023/23-402 provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at UPPMAX, funded by the Swedish Research Council through grant agreement no. 2022-06725. We thank Göran Arnqvist for helpful discussions on the evolution of sex and recombination and Marie Raynaud for assistance running LDhelmet. The Parks and Wildlife Commission, Northern Territory, and the Department of the Environment, Water, Heritage and the Arts, Australia, gave permission to collect (permit number 59044) and export (PWS2016-001559) the *C. secundus* termites. The Office Ivoirien des Parcs et Réserves (OIPR) provided sampling

and export permits for the *M. bellicosus* samples. The study was conducted according to the Nagoya protocol.

Author contributions: M.T.W. designed and led the research; T.E. and T.R. conducted the analysis; M.T.W., T.E., T.R., and J.K. wrote the paper; D.E. and J.K. contributed data and samples; and A.O., Y.L., and T.L. performed laboratory work.

References

- Agrawal AF. 2001. Sexual selection and the maintenance of sexual reproduction. *Nature* **411**: 692–695. doi:10.1038/35079590
- Arsala D, Wu X, Yi SV, Lynch JA. 2022. Dnmt1a is essential for gene body methylation and the regulation of the zygotic genome in a wasp. *PLoS Genet* **18**: e1010181. doi:10.1371/journal.pgen.1010181
- Auton A, McVean G. 2007. Recombination rate estimation in the presence of hotspots. *Genome Res* **17**: 1219–1227. doi:10.1101/gr.6386707
- Axelsson E, Webster MT, Ratnakumar A, Consortium L, Ponting CP, Lindblad-Toh K. 2012. Death of PRDM9 coincides with stabilization of the recombination landscape in the dog genome. *Genome Res* **22**: 51–63. doi:10.1101/gr.124123.111
- Baer B. 2005. Sexual selection in *Apis* bees. *Apidologie (Celle)* **36**: 187–200. doi:10.1051/apido:2005013
- Bailey TL. 2021. STREME: accurate and versatile sequence motif discovery. *Bioinformatics* **37**: 2834–2840. doi:10.1093/bioinformatics/btab203
- Baker Z, Schumer M, Haba Y, Bashkirova L, Holland C, Rosenthal GG, Przeworski M. 2017. Repeated losses of PRDM9-directed recombination despite the conservation of PRDM9 across vertebrates. *eLife* **6**: e24133. doi:10.7554/eLife.24133
- Bao W, Kojima KK, Kohany O. 2015. Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob DNA* **6**: 11. doi:10.1186/s13100-015-0041-9
- Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B. 2010. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327**: 836–840. doi:10.1126/science.1183439
- Begun DJ, Aquadro CF. 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356**: 519–520. doi:10.1038/356519a0
- Bergamaschi S, Dawes-Gromadzki TZ, Scali V, Marini M, Mantovani B. 2007. Karyology, mitochondrial DNA and the phylogeny of Australian termites. *Chromosome Res* **15**: 735–753. doi:10.1007/s10577-007-1158-6
- Berglund J, Quilez J, Arndt PF, Webster MT. 2015. Germline methylation patterns determine the distribution of recombination events in the dog genome. *Genome Biol Evol* **7**: 522–530. doi:10.1093/gbe/evu282
- Bewick AJ, Sanchez Z, McKinney EC, Moore AJ, Moore PJ, Schmitz RJ. 2019. Dnmt1 is essential for egg production and embryo viability in the large milkweed bug, *Oncopeltus fasciatus*. *Epigenetics Chromatin* **12**: 6. doi:10.1186/s13072-018-0246-5
- Beye M, Gattermeier I, Hasselmann M, Gempe T, Schioett M, Baines JF, Schlipalius D, Mougell F, Emore C, Rueppell O, et al. 2006. Exceptionally high levels of recombination across the honey bee genome. *Genome Res* **16**: 1339–1344. doi:10.1101/gr.5680406
- Browning SR, Browning BL. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* **81**: 1084–1097. doi:10.1086/521987
- Browning BL, Zhou Y, Browning SR. 2018. A one-penny imputed genome from next-generation reference panels. *Am J Hum Genet* **103**: 338–348. doi:10.1016/j.ajhg.2018.07.015
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421. doi:10.1186/1471-2105-10-421
- Cardoso DC, Santos HG, Cristiano MP. 2018. The Ant Chromosome Database—ACdb: an online resource for ant (Hymenoptera: Formicidae) chromosome researchers. *Myrmecol News* **27**: 87–91.
- Chan AH, Jenkins PA, Song YS. 2012. Genome-wide fine-scale recombination rate variation in *Drosophila melanogaster*. *PLoS Genet* **8**: e1003090. doi:10.1371/journal.pgen.1003090
- Charlesworth B. 2009. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* **10**: 195–205. doi:10.1038/nrg2526
- Coop G, Przeworski M. 2007. An evolutionary view of human recombination. *Nat Rev Genet* **8**: 23–34. doi:10.1038/nrg1947
- Cunha MS, Cardoso DC, Cristiano MP, de Oliveira Campos LA, Lopes DM. 2021. The Bee Chromosome database (Hymenoptera: Apidae). *Apidologie (Celle)* **52**: 493–502. doi:10.1007/s13592-020-00838-2

- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158. doi:10.1093/bioinformatics/btr330
- Elango N, Hunt BG, Goodisman MA, Yi SV. 2009. DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proc Natl Acad Sci* **106**: 11206–11211. doi:10.1073/pnas.0900301106
- Elsner D, Meusemann K, Korb J. 2018. Longevity and transposon defense, the case of termite reproductives. *Proc Natl Acad Sci* **115**: 5504–5509. doi:10.1073/pnas.1804046115
- Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* **20**: 238. doi:10.1186/s13059-019-1832-y
- Fazalova V, Nevado B. 2020. Low spontaneous mutation rate and Pleistocene radiation of pea aphids. *Mol Biol Evol* **37**: 2045–2051. doi:10.1093/molbev/msaa066
- Fernandes JB, Séguéla-Arnaud M, Larchevêque C, Lloyd AH, Mercier R. 2018. Unleashing meiotic crossovers in hybrid plants. *Proc Natl Acad Sci* **115**: 2431–2436. doi:10.1073/pnas.1713078114
- Fischer O, Schmid-Hempel P. 2005. Selection by parasites may increase host recombination frequency. *Biol Lett* **1**: 193–195. doi:10.1098/rsbl.2005.0296
- Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci* **117**: 9451–9457. doi:10.1073/pnas.1921046117
- Garrison E, Kronenberg ZN, Dawson ET, Pedersen BS, Prins P. 2022. A spectrum of free software tools for processing the VCF variant call format: vcfliib, bio-vcf, cyvcf2, hts-nim and slivar. *PLoS Comput Biol* **18**: e1009123. doi:10.1371/journal.pcbi.1009123
- Glastad KM, Hunt BG, Yi SV, Goodisman MAD. 2011. DNA methylation in insects: on the brink of the epigenomic era. *Insect Mol Biol* **20**: 553–565. doi:10.1111/j.1365-2583.2011.01092.x
- Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**: 1017–1018. doi:10.1093/bioinformatics/btr064
- Harrison MC, Jongepier E, Robertson HM, Arming N, Bitard-Feildel T, Chao H, Childers CP, Dinh H, Doddapaneni H, Dugan S, et al. 2018. Hemimetabolous genomes reveal molecular basis of termite eusociality. *Nat Ecol Evol* **2**: 557–566. doi:10.1038/s41559-017-0459-1
- Hartfield M, Keightley PD. 2012. Current hypotheses for the evolution of sex and recombination. *Integr Zool* **7**: 192–209. doi:10.1111/j.1749-4877.2012.00284.x
- Hedrick PW, Parker JD. 1997. Evolutionary genetics and genetic variation of haplodiploids and X-linked genes. *Annu Rev Ecol Syst* **28**: 55–83. doi:10.1146/annurev.ecolsys.28.1.55
- Hoffmann K, Foster KR, Korb J. 2012. Nest value mediates reproductive decision making within termite societies. *Behav Ecol* **23**: 1203–1208. doi:10.1093/beheco/ars103
- Huang G, Song L, Du X, Huang X, Wei F. 2023. Evolutionary genomics of camouflage innovation in the orchid mantis. *Nat Commun* **14**: 4821. doi:10.1038/s41467-023-40355-1
- Janicke T, Häderer IK, Lajeunesse MJ, Anthes N. 2016. Darwinian sex roles confirmed across the animal kingdom. *Sci Adv* **2**: e1500983. doi:10.1126/sciadv.1500983
- Jankásek M, Kotyková Varadinová Z, Štráhlavský F. 2021. Blattodea karyotype database. *Eur J Entomol* **118**: 192–199. doi:10.14411/eje.2021.020
- Jones JC, Wallberg A, Christmas MJ, Kapheim KM, Webster MT. 2019. Extreme differences in recombination rate between the genomes of a solitary and a social bee. *Mol Biol Evol* **36**: 2277–2291. doi:10.1093/molbev/msz130
- Joseph J, Prentout D, Laverré A, Tricou T, Duret L. 2024. High prevalence of PRDM9-independent recombination hotspots in placental mammals. *Proc Natl Acad Sci* **121**: e2401973121. doi:10.1073/pnas.2401973121
- Kawakami T, Wallberg A, Olsson A, Wintermantel D, de Miranda JR, Allsopp M, Rundlöf M, Webster MT. 2019. Substantial heritable variation in recombination rate on multiple scales in honeybees and bumblebees. *Genetics* **212**: 1101–1119. doi:10.1534/genetics.119.302008
- Keightley PD, Trivedi U, Thomson M, Oliver F, Kumar S, Blaxter ML. 2009. Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Res* **19**: 1195–1201. doi:10.1101/gr.091231.109
- Keightley PD, Ness RW, Halligan DL, Haddrill PR. 2014. Estimation of the spontaneous mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib family. *Genetics* **196**: 313–320. doi:10.1534/genetics.113.158758
- Keightley PD, Pinharanda A, Ness RW, Simpson F, Dasmahapatra KK, Mallet J, Davey JW, Jiggins CD. 2015. Estimation of the spontaneous mutation rate in *Heliconius melpomene*. *Mol Biol Evol* **32**: 239–243. doi:10.1093/molbev/msu302
- Kent CF, Zayed A. 2013. Evolution of recombination and genome structure in eusocial insects. *Commun Integr Biol* **6**: e22919. doi:10.4161/cib.22919
- Kent CF, Minaei S, Harpur BA, Zayed A. 2012. Recombination is associated with the evolution of genome structure and worker behavior in honey bees. *Proc Natl Acad Sci* **109**: 18012–18017. doi:10.1073/pnas.1208094109
- Koeniger G, Koeniger N, Phiancharoen M. 2011. Comparative reproductive biology of honeybees. In *Honeybees of Asia* (ed. Hepburn HR, Radloff SE), pp. 159–206. Springer, Berlin.
- Korb J. 2008. Termites, hemimetabolous diploid white ants? *Front Zool* **5**: 15. doi:10.1186/1742-9994-5-15
- Korb J, Hartfelder K. 2008. Life history and development: a framework for understanding developmental plasticity in lower termites. *Biol Rev* **83**: 295–313. doi:10.1111/j.1469-185X.2008.00044.x
- Korb J, Thorne B. 2017. Sociality in termites. In *Comparative social evolution* (ed. Rubenstein DR, Abbot P), pp. 124–153. Cambridge University Press, Cambridge, UK.
- Krasovec M. 2021. The spontaneous mutation rate of *Drosophila pseudoobscura*. *G3 (Bethesda)* **11**: jkab151. doi:10.1093/g3journal/jkab151
- Lam I, Keeney S. 2015. Nonparadoxical evolutionary stability of the recombination initiation landscape in yeast. *Science* **350**: 932–937. doi:10.1126/science.aad0814
- Leffler EM, Bullaughey K, Matute DR, Meyer WK, Séguérel L, Venkat A, Andolfatto P, Przeworski M. 2012. Revisiting an old riddle: What determines genetic diversity levels within species? *PLoS Biol* **10**: e1001388. doi:10.1371/journal.pbio.1001388
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760. doi:10.1093/bioinformatics/btp324
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079. doi:10.1093/bioinformatics/btp352
- Lin S, Werle J, Korb J. 2021. Transcriptomic analyses of the termite, *Cryptotermes secundus*, reveal a gene network underlying a long lifespan and high fecundity. *Commun Biol* **4**: 384. doi:10.1038/s42003-021-01892-x
- Liu H, Zhang X, Huang J, Chen J-Q, Tian D, Hurst LD, Yang S. 2015. Causes and consequences of crossing-over evidenced via a high-resolution recombination landscape of the honey bee. *Genome Biol* **16**: 15. doi:10.1186/s13059-014-0566-0
- Liu H, Jia Y, Sun X, Tian D, Hurst LD, Yang S. 2017. Direct determination of the mutation rate in the bumblebee reveals evidence for weak recombination-associated mutation and an approximate rate constancy in insects. *Mol Biol Evol* **34**: 119–130. doi:10.1093/molbev/msw226
- Lyko F, Foret S, Kucharski R, Wolf S, Falckenhayn C, Maleszka R. 2010. The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol* **8**: e1000506. doi:10.1371/journal.pbio.1000506
- Lynch M, Ali F, Lin T, Wang Y, Ni J, Long H. 2023. The divergence of mutation rates and spectra across the tree of life. *EMBO Rep* **24**: e57561. doi:10.15252/embr.202357561
- Madeira F, Madhusoodanan N, Lee J, Eusebi A, Niewielska A, Tivey ARN, Lopez R, Butcher S. 2024. The EMBL-EBI job dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res* **52**: W521–W525. doi:10.1093/nar/gkae241
- Manning JT, Chamberlain AT. 1997. Sib competition and sperm competitiveness: an answer to ‘Why so many sperms?’ and the recombination/sperm number correlation. *Proc Biol Sci* **256**: 177–182. doi:10.1098/rspb.1994.0067
- Maynard Smith J. 1976. A short-term advantage for sex and recombination through sib-competition. *J Theor Biol* **63**: 245–258. doi:10.1016/0022-5193(76)90033-3
- Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J, Flouri T, Beutler RG, et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**: 763–767. doi:10.1126/science.1257570
- Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. 2005. A fine-scale map of recombination rates and hotspots across the human genome. *Science* **310**: 321–324. doi:10.1126/science.1117196
- Myers S, Bowden R, Tumian A, Bontrop RE, Freeman C, MacFie TS, McVean G, Donnelly P. 2010. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. *Science* **327**: 876–879. doi:10.1126/science.1182363
- Nachman MW. 2001. Single nucleotide polymorphisms and recombination rate in humans. *Trends Genet* **17**: 481–485. doi:10.1016/S0168-9525(01)02409-X
- Niehuis O, Gibson JD, Rosenberg MS, Pannebakker BA, Koevoets T, Judson AK, Desjardins CA, Kennedy K, Duggan D, Beukeboom LW, et al. 2010. Recombination and its impact on the genome of the haplodiploid

- parasitoid wasp *Nasonia*. *PLoS One* **5**: e8597. doi:10.1371/journal.pone.0008597
- Parvanov ED, Petkov PM, Paigen K. 2010. *Prdm9* controls activation of mammalian recombination hotspots. *Science* **327**: 835. doi:10.1126/science.1181495
- Persikov AV, Singh M. 2014. De novo prediction of DNA-binding specificities for Cys2His2 zinc finger proteins. *Nucleic Acids Res* **42**: 97–108. doi:10.1093/nar/gkt890
- Qiu B, Elsner D, Korb J. 2023. High-quality long-read genome assemblies reveal evolutionary patterns of transposable elements and DNA methylation in termites. *bioRxiv* doi:10.1101/2023.10.31.564968
- Roisin Y. 2001. Caste sex ratios, sex linkage, and reproductive strategies in termites. *Insectes Soc* **48**: 224–230. doi:10.1007/PL00001770
- Romiguier J, Lourenco J, Gayral P, Faivre N, Weinert LA, Ravel S, Ballenghien M, Cahais V, Bernard A, Loire E, et al. 2014. Population genomics of eusocial insects: the costs of a vertebrate-like effective population size. *J Evol Biol* **27**: 593–603. doi:10.1111/jeb.12331
- Ross L, Blackmon H, Lorite P, Gokhman VE, Hardy NB. 2015. Recombination, chromosome number and eusociality in the Hymenoptera. *J Evol Biol* **28**: 105–116. doi:10.1111/jeb.12543
- Rueppell O, Kuster R, Miller K, Fouks B, Rubio Correa S, Collazo J, Phaincharoen M, Tingek S, Koeniger N. 2016. A new Metazoan recombination rate record and consistently high recombination rates in the honey bee genus *Apis* accompanied by frequent inversions but not translocations. *Genome Biol Evol* **8**: 3653–3660. doi:10.1093/gbe/evw269
- Sarda S, Zeng J, Hunt BG, Yi SV. 2012. The evolution of invertebrate gene body methylation. *Mol Biol Evol* **29**: 1907–1916. doi:10.1093/molbev/mss062
- Seabold S, Perktold J. 2010. Statsmodels: econometric and statistical modeling with Python. In *Proceedings of the 9th Python in Science Conference*, Austin, TX (ed. Van der Walt S, Millman J), pp. 92–96.
- Sherman PW. 1979. Insect chromosome numbers and eusociality. *Am Nat* **113**: 925–935. doi:10.1086/283445
- Shi YY, Sun LX, Huang ZY, Wu XB, Zhu YQ, Zheng HJ, Zeng ZJ. 2013. A SNP based high-density linkage map of *Apis cerana* reveals a high recombination rate similar to *Apis mellifera*. *PLoS One* **8**: e76459. doi:10.1371/journal.pone.0076459
- Siller S. 2001. Sexual selection and the maintenance of sex. *Nature* **411**: 689–692. doi:10.1038/35079578
- Singhal S, Leffler EM, Sannareddy K, Turner I, Venn O, Hooper DM, Strand AI, Li Q, Raney B, Balakrishnan CN, et al. 2015. Stable recombination hotspots in birds. *Science* **350**: 928–932. doi:10.1126/science.aad0843
- Sirviö A, Gadau J, Rueppell O, Lamatsch D, Boomsma JJ, Pamilo P, Page RE Jr. 2006. High recombination frequency creates genotypic diversity in colonies of the leaf-cutting ant *Acromyrmex echinatior*. *J Evol Biol* **19**: 1475–1485. doi:10.1111/j.1420-9101.2006.01131.x
- Sirviö A, Johnston JS, Wenseleers T, Pamilo P. 2011a. A high recombination rate in eusocial Hymenoptera: evidence from the common wasp *Vespa vulgaris*. *BMC Genet* **12**: 95. doi:10.1186/1471-2156-12-95
- Sirviö A, Pamilo P, Johnson RA, Page RE Jr, Gadau J. 2011b. Origin and evolution of the dependent lineages in the genetic caste determination system of *Pogonomyrmex* ants. *Evolution* **65**: 869–884. doi:10.1111/j.1558-5646.2010.01170.x
- Smukowski Heil CS, Ellison C, Dubin M, Noor MAF. 2015. Recombining without hotspots: a comprehensive evolutionary portrait of recombination in two closely related species of *Drosophila*. *Genome Biol Evol* **7**: 2829–2842. doi:10.1093/gbe/evv182
- Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. 2017. Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Philos Trans R Soc Lond B Biol Sci* **372**: 20160455. doi:10.1098/rstb.2016.0455
- Storer J, Hubley R, Rosen J, Wheeler TJ, Smit AF. 2021. The Dfam community resource of transposable element families, sequence models, and genome annotations. *Mob DNA* **12**: 2. doi:10.1186/s13100-020-00230-y
- Templeton AR. 1979. Chromosome number, quantitative genetics and eusociality. *Am Nat* **113**: 937–941. doi:10.1086/283446
- Torres API, Höök L, Näsvalk K, Shipilina D, Wiklund C, Vila R, Pruischer P, Backström N. 2023. The fine-scale recombination rate variation and associations with genomic features in a butterfly. *Genome Res* **33**: 810–823. doi:10.1101/gr.277414.122
- Trivers RL, Hare H. 1976. Haplodiploidy and the evolution of the social insect. *Science* **191**: 249–263. doi:10.1126/science.1108197
- Ventós-Alfonso A, Ylla G, Montañes J-C, Belles X. 2020. DNMT1 promotes genome methylation and early embryo development in cockroaches. *iScience* **23**: 101778. doi:10.1016/j.isci.2020.101778
- Waiker P, de Abreu FCP, Luna-Lucena D, Freitas FCP, Simões ZLP, Rueppell O. 2021. Recombination mapping of the Brazilian stingless bee *Friesomelitta varia* confirms high recombination rates in social hymenoptera. *BMC Genomics* **22**: 673. doi:10.1186/s12864-021-07987-3
- Wallberg A, Han F, Wellhagen G, Dahle B, Kawata M, Haddad N, Simões ZLP, Allsopp MH, Kandemir I, De la Rúa P, et al. 2014. A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honeybee *Apis mellifera*. *Nat Genet* **46**: 1081–1088. doi:10.1038/ng.3077
- Wallberg A, Glémin S, Webster MT. 2015. Extreme recombination frequencies shape genome variation and evolution in the honeybee, *Apis mellifera*. *PLoS Genet* **11**: e1005189. doi:10.1371/journal.pgen.1005189
- Wallberg A, Schöning C, Webster MT, Hasselmann M. 2017. Two extended haplotype blocks are associated with adaptation to high altitude habitats in East African honey bees. *PLoS Genet* **13**: e1006792. doi:10.1371/journal.pgen.1006792
- Wallberg A, Bunikis I, Pettersson OV, Mosbech M-B, Childers AK, Evans JD, Mikheyev AS, Robertson HM, Robinson GE, Webster MT. 2019. A hybrid de novo genome assembly of the honeybee, *Apis mellifera*, with chromosome-length scaffolds. *BMC Genomics* **20**: 275. doi:10.1186/s12864-019-5642-0
- Wang X, Wheeler D, Avery A, Rago A, Choi J-H, Colbourne JK, Clark AG, Werren JH. 2013. Function and evolution of DNA methylation in *Nasonia vitripennis*. *PLoS Genet* **9**: e1003872. doi:10.1371/journal.pgen.1003872
- Wang J, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, et al. 2023. The conserved domain database in 2023. *Nucleic Acids Res* **51**: D384–D388. doi:10.1093/nar/gkac1096
- Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* **7**: 256–276. doi:10.1016/0040-5809(75)90020-9
- Wilfert L, Gadau J, Schmid-Hempel P. 2007. Variation in genomic recombination rates among animal taxa and the case of social insects. *Heredity (Edinb)* **98**: 189–197. doi:10.1038/sj.hdy.6800950
- Winkler L, Moiron M, Morrow EH, Janicke T. 2021. Stronger net selection on males across animals. *eLife* **10**: e68316. doi:10.7554/eLife.68316
- Winston ML. 1991. *The biology of the honey bee*, revised ed. Harvard University Press, Cambridge, MA.
- Wu T-C, Lichten M. 1994. Meiosis-induced double-strand break sites determined by yeast chromatin structure. *Science* **263**: 515–518. doi:10.1126/science.8290959
- Yang S, Wang L, Huang J, Zhang X, Yuan Y, Chen J-Q, Hurst LD, Tian D. 2015. Parent-progeny sequencing indicates higher mutation rates in heterozygotes. *Nature* **523**: 463–467. doi:10.1038/nature14649
- Zhang C, Dong S-S, Xu J-Y, He W-M, Yang T-L. 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **35**: 1786–1788. doi:10.1093/bioinformatics/bty875

Received March 22, 2024; accepted in revised form February 18, 2025.