# Causes and consequences of a complex recombinational landscape in the ant *Cardiocondyla obscurior*

Mohammed Errbii, Jürgen Gadau, Kerstin Becker, et al.

*Genome Res.* published online June 5, 2024
Access the most recent version at doi:10.1101/gr.278392.123

# Causes and consequences of a complex recombinational landscape in the ant *Cardiocondyla obscurior*

**Running title:** meiotic recombination in *Cardiocondyla obscurior*

Mohammed Errbii[1], Jürgen Gadau[1], Kerstin Becker[2], Lukas Schrader[1*], Jan Oettler[3*]

[1]Institute for Evolution and Biodiversity, University of Münster, 48149, Münster, Germany

[2]Cologne Center for Genomics (CCG), Medical Faculty, University of Cologne, 50931 Cologne, Germany

[3]Lehrstuhl für Zoologie/Evolutionsbiologie, University Regensburg, 93053 Regensburg, Germany

*Correspondence: Lukas Schrader & Jan Oettler

**Email:** lukas.schrader@uni-muenster.de and joettler@gmail.com

## Abstract

Eusocial Hymenoptera have the highest recombination rates among all multicellular animals studied so far, but it is unclear why this is and how this affects the biology of individual species. A high-resolution linkage map for the ant *Cardiocondyla obscurior* corroborates genome-wide high recombination rates reported for ants (8.1 cM/Mb). However, recombination is locally suppressed in regions either enriched with TEs, with strong haplotype divergence, or showing signatures of epistatic selection in *C. obscurior*. The results do not support the hypotheses that high recombination rates are linked to phenotypic plasticity or to modulating selection efficiency. Instead, genetic diversity and the frequency of structural variants correlate positively with local recombination rates, potentially compensating for the low levels of genetic variation expected in haplodiploid social Hymenoptera with low effective population size. Ultimately, the data show that recombination contributes to within-population polymorphism and to the divergence of the lineages within *C. obscurior*.

## Introduction

Meiotic recombination is a key feature of sexually reproducing organisms (Crow 1994; Bell 1982; Maynard Smith 1978), where homologous chromosomes form a bivalent structure through physical linkage involving one to two crossovers per chromosome (Roeder 1997; Kleckner 1996; White 1973). These crossovers are essential for chromosome stability and proper segregation, and can promote genotypic variation, thereby enhancing the adaptive potential of organisms (Roeder 1997; Kleckner 1996; Jones and Franklin 2006; Ritz et al. 2017). Recombination rates vary between and within species, and are strongly influenced by genome architecture (Lynn et al. 2004; Stapley et al. 2017). Recombination itself can drive genome evolution, both through its mutagenic effects (Webster and Hurst 2012; Halldorsson et al. 2019; Arbeithuber et al. 2015; Langley et al. 1988; Montgomery et al. 1991) and by counteracting the diversity-reducing effects of linked selection (Cutter and Payseur 2013; Ellegren and Galtier 2016).

Genetic linkage maps across several Apocrita species (ants, wasps, and bees) (Hunt and Page 1995; Liu et al. 2015; Rueppell et al. 2016; Waiker et al. 2021; Sirviö et al. 2011b, 2006, 2011a; Wilfert et al. 2006; Shi et al. 2013) indicate that social Hymenoptera have particularly high recombination rates compared to other animals studied so far (Wilfert et al. 2007; Stapley et al. 2017). Studies in *Nasonia*, a haplodiploid solitary parasitoid wasp, revealed low recombination rates (~1.5 – 2.5 cM/Mb) (Niehuis et al. 2010; Gadau et al. 1999), indicating that the haplodiploid system alone does not explain the high recombination rates observed in social Hymenoptera. The prevailing hypotheses suggesting that the extreme recombination rates in social insects are related to the complex needs of social life has found only ambiguous support so far, and exclusively from studies on the honey bee *Apis mellifera*.

First, by increasing the efficiency of natural selection (Comeron et al. 2008; Felsenstein 1974; Felsenstein and Yokoyama 1976), high recombination rates could promote the independent evolution of caste-specific genes (Kent et al. 2012; Kent and Zayed 2013). However, studies have shown conflicting results regarding the association between recombination rates and selection efficiency (Kent et al. 2012; Jones et al. 2019). Second, by increasing colony-level genotypic diversity, these elevated rates could facilitate social life (Oldroyd and Fewell 2007; Kozomara et al. 2019; Liu et al. 2015). For example, elevated recombination rates around genes with a caste-biased expression pattern suggest a link with division of labor (Liu et al. 2015; Wallberg et al. 2015; Kent et al. 2012). However, this pattern was also observed for orthologous sequences in a solitary bee, suggesting that the high recombination rates in the vicinity of these genes preceded the evolution of sociality in the bee lineage (Jones et al. 2019). Lastly, increasing diversity could also help resist the high pathogen pressure of social life (Hughes and Boomsma 2004; Kraus and Page 1998; Schmid-Hempel 1998; Liu et al. 2015), but immune-related genes did not coincide with regions of high recombination rates (Liu et al. 2015). In summary, there is no consensus regarding possible causes and consequences of high recombination in honey bees, let alone in other social Hymenoptera. In ants, two studies conducted in *Acromyrmex echinatior* (Sirviö et al. 2006) and *Pogonomyrmex rugosus* (Sirviö et al. 2011b) using a limited number of markers (145–215 markers) corroborated elevated recombination rates in social insects. However, these studies did not explore the genomic features associated with the genome-wide variation in recombination rates.

Here, we use genetic linkage mapping, a chromosome-level genome assembly, and extensive population-genomic data of the model species *Cardiocondyla obscurior* (Oettler 2020), for first analyses of the causes and consequences of recombination rate variation in an ant. In particular,

73 we explore putative factors associated with local recombination rate variation in the *C. obscurior*
74 genome. Furthermore, using published gene expression data of larvae of queens, workers,
75 winged and wingless males (Schrader et al. 2017, 2015) and the ratio between nonsynonymous
76 and synonymous substitutions ($d_N/d_S$), we critically address the hypotheses linking high
77 recombination rates to the evolution or elaboration of sociality in Hymenoptera.

## Results

*Initial linkage mapping and genome assembly refinement and annotation*

To improve the current genome assembly (Cobs2.1) to chromosome resolution, we produced an initial ordering of 57,557 markers (out of 342,508 raw biallelic SNPs) into 29 LGs (Supplemental Fig. S1) using whole-genome data from 100 F2 offspring of a single F1 mother obtained from a cross between a male and a queen belonging to two Brazilian populations of *C. obscurior*. These populations were specifically chosen to maximize the number of heterozygous markers and minimize runs of homozygosity in the experimental cross. Although some small LGs may be physically connected to each other or to larger LGs, they were not combined due to large genetic distances (~50 cM), indicating that they segregate independently from other LGs.

The new genome (Cobs3.1) spans 193.05 Mb and is an improvement over the previous version (Cobs2.1) with ~96% (184.4 Mb) of the physical assembly placed into pseudomolecules and a scaffold N50 of 7.38 Mb compared to 6.29 Mb of Cobs2.1 (Supplemental Table S1). Using the liftoff tool (v.1.6.3) (Shumate and Salzberg 2021), we converted the coordinates of 20,944 genes (36,599 transcripts) annotated in Cobs2.1 to Cobs3.1.

Transposable elements (TEs) exhibit a bimodal distribution in the genome of *C. obscurior* with regions of high TE content ("TE islands") in a genomic background of low TE content (low-density regions, "LDRs") (Supplemental Fig. S2A), as described previously (Errbii et al. 2021; Schrader et al. 2014). A total of 30 TE islands are located at the center (e.g., LGs 1, 2, 3, and 5) and the ends (e.g., LGs 4, 8, and 9) of LGs. We also observed secondary TE accumulations on several LGs (e.g., 1, 5, and 10).

Tandem repeat sequences are commonly associated with centromeres across various species, including ants (Melters et al. 2013; Huang et al. 2016). Upon annotating tandem repeats in *C. obscurior*, we observed an overlap of these sequences with TE islands (Supplemental Fig. S2B). Candidate centromeric repeats from five different ant species (Melters et al. 2013; Huang et al. 2016) were not enriched in TE islands, suggesting a distinct repertoire of repeats in centromeres of *C. obscurior*, consistent with prior findings of rapid divergence among centromeric tandem repeats in ants (Huang et al. 2016). Together, these results suggests that TE islands may serve as potential centromeres for the chromosomes of *C. obscurior*.

*The recombination landscape of C. obscurior*

To estimate recombination rates, we used the updated assembly (Cobs3.1) for a second round of linkage mapping with 57,309 high-quality markers (out of 342,616 raw biallelic SNPs) covering 84% (162.9 Mb) of the physical assembly and revealing 1,061 effectively recombining positions. It is important to note that the number of recombining positions is independent of marker density and is instead determined by the occurrences of crossover events in the experiment. The final map (1,315.4 cM in 29 LGs; Supplemental Fig. S3), showed a positive correlation between genetic and physical positions of markers, indicating a high-quality map (Supplemental Fig. S4; correlation coefficients are given on the figure). Based on the proportion of the physical assembly covered by markers, the average recombination rate in *C. obscurior* is 8.1 cM/Mb. While higher than observed in other non-social insects and animal species (Supplemental Table S2), this estimate falls within the range of other ants (Sirviö et al. 2011b, 2006) and is lower than those of the honey bee (Beye et al. 2006; Liu et al. 2015; Rueppell et al. 2016; Shi et al. 2013).

Recombination rate varied considerably among the 29 LGs (Table 1) with LGs smaller than 5 Mb having a higher recombination rate (median = 14.2 cM/Mb (95% CI [8.52, 18.06])) compared to

LGs larger than 10 Mb (median = 7.5 cM/Mb (95% CI [3.65, 10.29])) and intermediate LGs (≥5 Mb and ≤ 10 Mb; median = 7.5 cM/Mb (95% CI [4.60, 9.55])). Physical and genetic distances correlated positively (Spearman's $\rho$ = 0.57, p = 0.0014) (Supplemental Fig. S5A), while average recombination rate per LG showed a significant negative correlation (Spearman's $\rho$ = -0.39, p = 0.039) with physical lengths of LGs (Supplemental Fig. S5B).

128 **Table 1** Physical length (Mb), map length (cM), recombination rate (cM/Mb) and number of
129 crossovers per LG in *C. obscurior*. Numbers in brackets indicate distances (Mb) between the first
130 and last marker on each LG, used to calculate average recombination rate

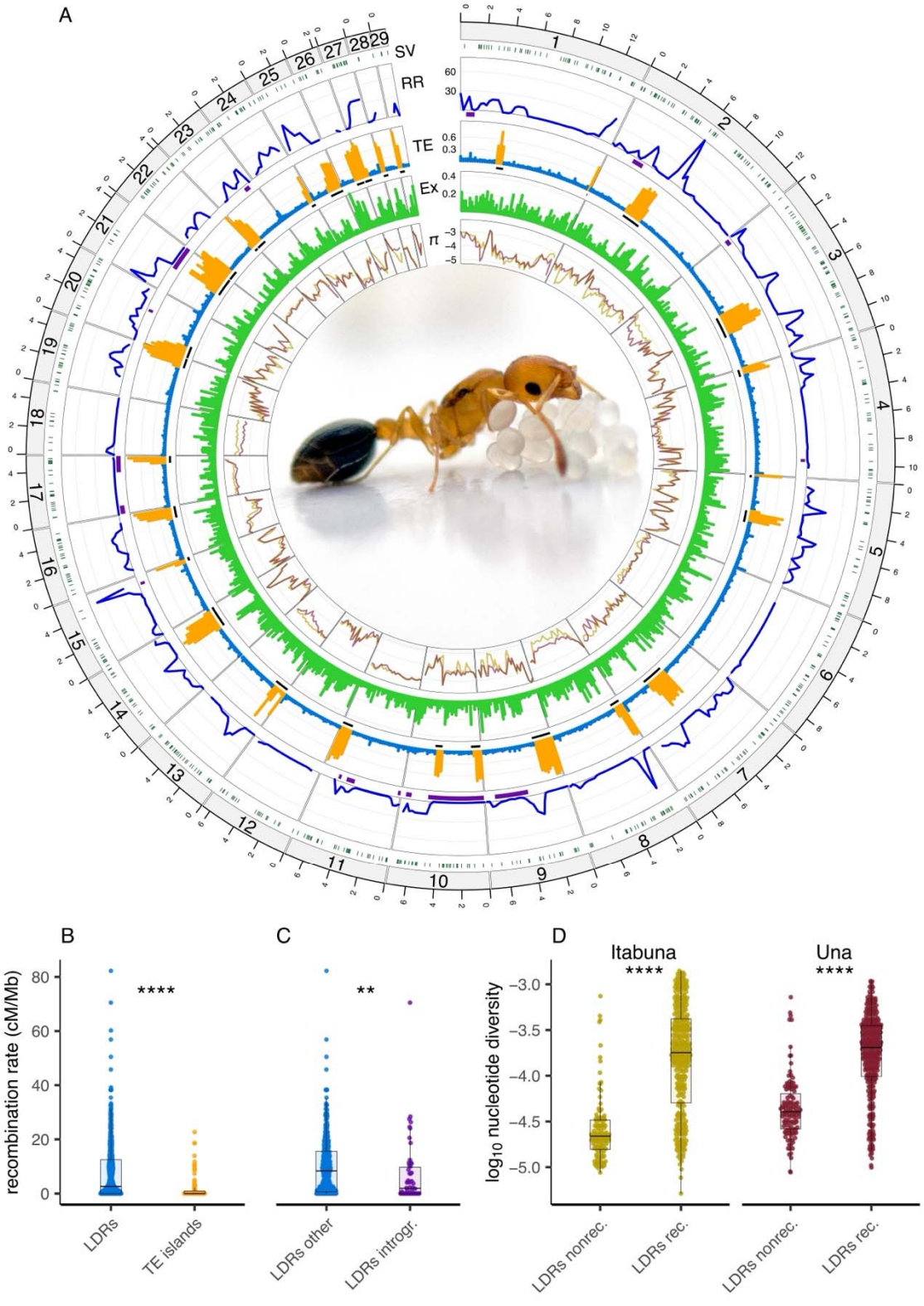
| Linkage group | Physical length (Mb) | Map length (cM) | Recombination rate (cM/Mb) | Crossover events |
|---|---|---|---|---|
| LG1 | 13.82 (13.28) | 91.57 | 6.90 | 89 |
| LG2 | 13.23 (12.87) | 104.21 | 8.09 | 96 |
| LG3 | 11.56 (11.10) | 97.77 | 8.81 | 96 |
| LG4 | 11.12 (10.29) | 41.89 | 4.07 | 41 |
| LG5 | 9.85 (8.94) | 93.37 | 10.44 | 94 |
| LG6 | 9.25 (7.55) | 2.00 | 0.27 | 1 |
| LG7 | 9.20 (8.34) | 111.04 | 13.31 | 109 |
| LG8 | 7.92 (7.88) | 60.82 | 7.72 | 56 |
| LG9 | 7.39 (7.25) | 50.34 | 6.95 | 49 |
| LG10 | 7.38 (7.09) | 45.62 | 6.43 | 44 |
| LG11 | 7.36 (6.10) | 42.17 | 6.91 | 41 |
| LG12 | 6.60 (5.37) | 5.99 | 1.12 | 7 |
| LG13 | 5.99 (5.02) | 37.76 | 7.52 | 38 |
| LG14 | 5.90 (5.70) | 46.27 | 8.12 | 46 |
| LG15 | 5.63 (5.04) | 50.03 | 9.94 | 46 |
| LG16 | 5.48 (4.72) | 45.43 | 9.62 | 45 |
| LG17 | 5.21 (5.07) | 5.01 | 0.99 | 4 |
| LG18 | 5.20 (4.90) | 8.02 | 1.64 | 8 |
| LG19 | 5.03 (2.97) | 44.99 | 15.13 | 44 |
| LG20 | 4.86 (4.40) | 49.62 | 11.28 | 47 |
| LG21 | 4.44 (4.32) | 54.65 | 12.64 | 54 |
| LG22 | 4.34 (2.76) | 33.62 | 12.18 | 33 |
| LG23 | 3.81 (3.12) | 49.24 | 15.78 | 47 |
| LG24 | 3.70 (3.01) | 49.21 | 16.34 | 47 |
| LG25 | 3.14 (2.66) | 45.56 | 17.14 | 44 |
| LG26 | 2.30 (0.33) | 1.00 | 3.06 | 1 |
| LG27 | 1.93 (1.70) | 39.65 | 23.26 | 36 |
| LG28 | 1.68 (0.72) | 1.50 | 2.08 | 1 |
| LG29 | 1.08 (0.36) | 7.00 | 19.19 | 7 |
| Total | 184.4 (162.87) | 1315.33 | 8.08 | 1271 |

131

Eight LGs showed very little or no recombination because of a reduced number of crossovers (LGs 4, 6, 12, 17, 18, 26, 28, 29) (Table 1). Although the number of markers varied between LGs (Supplemental Fig. S6), there was no correlation between average recombination rates and marker count (Spearman's $\rho$ = -0.087, p = 0.65) per LG (Supplemental Fig. S5C). In the following, we investigated potential factors contributing to the reduced number of crossovers and, consequently, recombination rate in *C. obscurior*.

*Recombination rate distribution and genome architecture in C. obscurior*

To investigate the genomic determinants of recombination in *C. obscurior*, we estimated local recombination rate in 250 kb nonoverlapping windows and analyzed these with respect to various sequence parameters (Fig. 1A; Supplemental Figs. S7, S8). Since crossovers can only be detected between two markers, recombination rate estimates were restricted to regions between the first and last markers on each chromosome, excluding the chromosome ends.




144

**Figure 1. Recombination rate variation and genome architecture in *C. obscurior*.** (A) Circos plot showing inwards: (SV) the location of identified structural variants (dark green), (RR) recombination rates across the genome (cM/Mb) with introgressions depicted with purple bars, (TE) TE content ("LDRs" in blue; ("TE islands" in orange) with the location of TE islands shown as black bars, (Ex) exon content and ($\pi$) the $\log_{10}$-transformed nucleotide diversity in two natural populations for *C. obscurior* from Itabuna (gold) and Una (maroon). *C. obscurior* queen with brood (Photo: L. Schrader). (B) Recombination rates between TE-poor regions ("LDRs" in blue) and TE-rich regions ("TE islands" in orange). (C) Recombination rates in introgressed regions ("LDRs introgr." in purple) compared to the remainder of LDRs ("LDRs other" in blue). (d) Levels of nucleotide diversity (π) in nonrecombining ("LDRs nonrec.") compared to recombining regions LDRs ("LDRs rec.") in two populations of *C. obscurior* from Itabuna and Una (****<0.0001; **<0.01).

The recombination rate in *C. obscurior* ranged from 0 to 82.2 cM/Mb, with approximately 38% (~70 Mb) of the analyzed sequence showing minimal to no recombination (e.g., on LGs 6, 12, 17 and 18, Fig. 1A; Supplemental Fig. S7A). Note, the recombination rate estimates were not affected by marker density. While a correlation was observed between marker density and recombination rate in 250 kb windows (Supplemental Fig. S9A; Spearman's $\rho$ = 0.075, p = 0.049), this correlation disappeared when considering only recombining windows (Supplemental Fig. S9B; Spearman's $\rho$ = 0.045, p = 0.37).

Markers were unevenly distributed across the LGs, with 23 out of the 29 LGs having median inter-marker distances below 0.64 kb (range: 0.12 to 256.97 kb; Supplemental Table S3). Three LGs 6, 12, and 18 showed minimal recombination and displayed particularly large inter-marker distances (medians: 210.27, 37.84 and 256.97 kb, respectively), 90% of which were below 750 kb (Supplemental Table S3). Due to chiasma interference (Muller 1916; Sturtevant 1915), where the occurrence of one crossover disrupts the formation of another nearby, these intervals should still be sufficient to capture crossovers. Unless, exceptionally rare double crossovers occur within these short intervals, we expect an average of one crossover every 2.7 Mb (Liu et

al. 2015) to 7.14 Mb (Sirviö et al. 2011b), according to studies in the honey bee and harvester ants, respectively.

Recombination rates are significantly reduced in TE islands (median = 0 cM/Mb), which are gene-poor, compared to LDRs (median = 3.11 cM/Mb) where gene density is higher (Fig. 1B; Wilcoxon test: W = 43644, $p = 5.25 \times 10^{-13}$). Note, the decrease in recombination rate within TE islands is not due to a lack of markers (Supplemental Fig. S6; Supplemental Table S4). LDRs on their own show no significant correlation between recombination rate and TE content (Spearman's $\rho$ = 0.016, p = 0.74) or exon content (Spearman's $\rho = -4 \times 10^{-3}$, p = 0.93) (Supplemental Fig. S8).

Beside TE islands, several other regions in LDRs also showed little or no recombination (Fig. 1A). Some of these nonrecombining regions are characterized by high sequence divergence between grandparental haplotypes (gp-$d_{XY}$) (Supplemental Fig. S7B). In these regions, one of the two grandparental haplotypes showed high sequence similarity to Asian and European populations of *C. obscurior* (Supplemental Fig. S7C), indicating introgression from the Old World lineage (Errbii et al. 2021). Based on a cutoff of threefold mean gp-$d_{XY}$ (~0.001), these introgressions occupy a large fraction (> 40%) of LGs 9, 10, 17, and 21, and smaller regions on LGs 1 and 2. After accounting for differences in the number of introgressed (53) and non-introgressed (330) windows, we found that regions with high gp-$d_{XY}$ recombine significantly less (median = 1.8 cM/Mb) than the rest of the genome (median = 8.52 cM/Mb) (Wilcoxon test: W = 1,043.024, p = 0.046) (Fig. 1C). Several inversions overlap with these introgressions (on LGs 10, 17, and 21) (Supplemental Table S5). Together, this indicates that strong sequence divergence among homologous chromosomes impairs recombination in *C. obscurior*, at least in part caused by polymorphic inversions.

Other nonrecombining regions on LGs 1, 4, 6, 8 and 12 showed no signs of introgression and contained no large-scale inversions that could explain local suppression of recombination (Fig. 1A; Supplemental Figs. S7A, S7B; Supplemental Table S5). Considering that meiotic recombination disrupts the linkage between segregating loci, we thus hypothesized that the absence of recombination in these regions could result from epistatic selection acting to reduce recombination in regions harboring co-adapted loci (Charlesworth 2016; Turner 1967). To explore this possibility, we screened these regions for signs of reduced genetic variation and Tajima's *D*, measures commonly used in population genetics to infer signatures of selection (Nielsen 2005). The nonrecombining regions on LGs 1, 4, 6, 8 and 12 showed significantly reduced levels of nucleotide diversity in natural populations (Itabuna $_{median} = 2.18 \times 10^{-5}$; Una $_{median} = 4.04 \times 10^{-5}$) compared to the remainder of the LDRs (excluding introgressions) (Itabuna $_{median} = 1.79 \times 10^{-4}$; Una $_{median} = 2.03 \times 10^{-4}$) (Fig. 1D; Wilcoxon test; Itabuna: $W = 5559.5$, $p < 2.2 \times 10^{-16}$; Una: $W = 6289.5$, $p < 2.2 \times 10^{-16}$). Similarly, Tajima's *D* was also significantly reduced in these nonrecombining regions (Itabuna $_{median} = -1.86$; Una $_{median} = -2.1$) compared to the remainder of the LDRs (excluding introgressions) (Itabuna $_{median} = -0.99$; Una $_{median} = -1.37$) (Supplemental Figs. S7D, S10; Wilcoxon test; Itabuna: $W = 9547$, $p = 1.32 \times 10^{-15}$; Una: $W = 9860$, $p = 1.07 \times 10^{-14}$).

We further applied a generalized linear model with recombination as the response variable and found that TE content, sequence divergence and epistasis were all significant explanatory variables (Supplemental Table S6).

Finally, there was no correlation between recombination rate and GC content (Supplemental Figs. S7E, S8; overall: Spearman's $\rho = 0.044$; $p = 0.35$; intergenic: Spearman's $\rho = 0.025$; $p = 0.6$) in *C. obscurior*.

*Meiotic recombination contributes to genetic variation in natural populations and promotes intraspecific divergence*

In support of a substantial effect of recombination on patterns of genetic variation, we found a positive correlation of recombination rates with measures of nucleotide diversity (Itabuna: Spearman's $\rho$ = 0.28; p = 2.1 × $10^{-9}$; Una: Spearman's $\rho$ = 0.25; p = 9.9 × $10^{-8}$) and Tajima's *D* (Itabuna: Spearman's $\rho$ = 0.21; p = 7.1 × $10^{-6}$; Una: Spearman's $\rho$ = 0.3; p = 5.6 × $10^{-11}$) calculated in 250 kb windows in LDRs (Fig. 2).

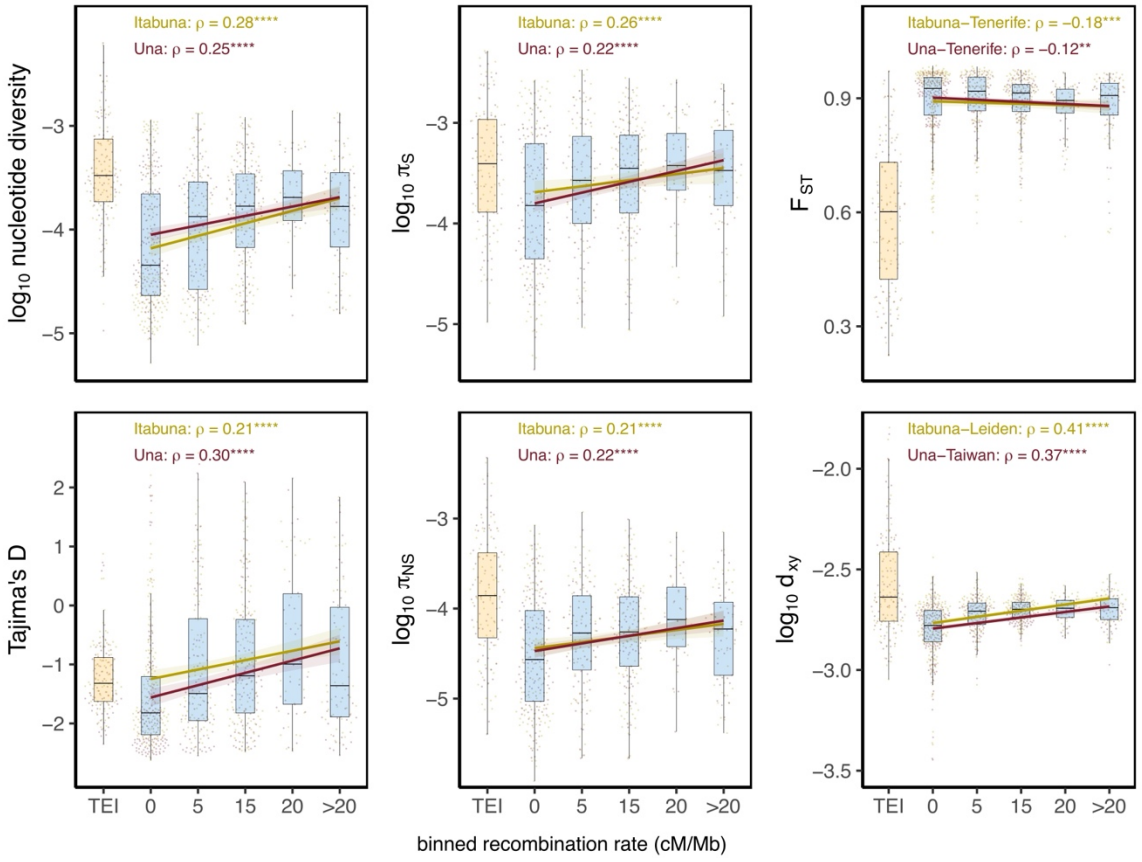



**Figure 2. Relationship between recombination rates and measures of diversity in populations of *C. obscurior*.** First column: correlation between recombination rates and nucleotide diversity and Tajima's *D*. Second column: correlation between recombination rates and nucleotide diversity at synonymous ($\pi_S$) and nonsynonymous ($\pi_{NS}$) sites. Third column: correlation between recombination rates and levels of genetic differentiation ($F_{ST}$) and absolute divergence ($d_{XY}$) between lineages of *C. obscurior*. The x-axis represents binned recombination rates, with the

upper limit of each bin indicated and "TEI" referring to TE islands (orange boxplot). $\rho$: Spearman's rank correlation coefficient, with significance shown by asterisks (****<0.0001; ***<0.001; **<0.01). All correlations were performed using the raw data. Lines and shaded areas are linear regressions and 95% confidence intervals.

Recombination rates were also positively correlated with levels of diversity at both non-synonymous ($\pi_{NS}$) (Itabuna: Spearman's $\rho$ = 0.21; p = $1 \times 10^{-5}$; Una: Spearman's $\rho$ = 0.22; p = $3.7 \times 10^{-6}$) and synonymous sites ($\pi_S$) (Itabuna: Spearman's $\rho$ = 0.26; p = $3.9 \times 10^{-8}$; Una: Spearman's $\rho$ = 0.22; p = $2.3 \times 10^{-6}$) (Fig. 2), but not with their ratio ($\pi_{NS}/\pi_S$) (Itabuna: Spearman's $\rho$ = 0.0062, p = 0.91; Una: Spearman's $\rho$ = 0.0018, p = 0.97) (Supplemental Fig. S11), indicating that meiotic recombination does not correlate with positive or purifying selection.

Consistent with models of linked selection where diversity is reduced in regions of low recombination (Burri 2017; Cutter and Payseur 2013), there was a negative correlation between $F_{ST}$ and recombination in two different population pairs ($F_{ST}$ [Itabuna-Tenerife]: Spearman's $\rho$ = -0.18; p = $1.7 \times 10^{-4}$; $F_{ST}$ [Una-Tenerife]: Spearman's $\rho$ = -0.12, p = $9.9 \times 10^{-3}$) (Fig. 2). However, there was also a significant positive association between recombination and $d_{XY}$ in two additional population pairs ($d_{XY}$ [Itabuna-Leiden]: Spearman's $\rho$ = 0.41; p < $2.2 \times 10^{-16}$; $d_{XY}$ [Una-Taiwan]: Spearman's $\rho$ = 0.37, p = $8.6 \times 10^{-16}$) (Fig. 2). Together, this indicates that recombination contributes to within-population polymorphism and to the divergence of *C. obscurior* lineages. Future studies should explore whether the same patterns persist when using recombination rates estimated for a population of the Old World lineage.

*Recombination is associated with small and large genome rearrangements*

Ectopic recombination (i.e., crossovers between non-allelic homologous loci), particularly in repeat-rich regions, has been shown to produce structural variants (SVs) (Savocco and Piazza

2021; Liu et al. 2012). To investigate to what extent the high recombination rates in *C. obscurior* are correlated with SVs, we used whole genome sequencing data generated previously for 16 individuals from five populations (Errbii et al. 2021).

Analysis of 481 SVs (129 deletions, 54 duplications, and 298 inversions) showed a positive correlation between recombination rates and SV frequencies and their lengths in LDRs (SVs length: Spearman's $\rho$ = 0.14; p = 0.0025; SVs frequency: Spearman's $\rho$ = 0.15; p = 8.1 × $10^{-4}$) (Figs. 3A and 3B). When analyzed separately, none of the individual SV types correlates significantly with recombination (Supplemental Fig. S12). Additionally, genetic diversity at 95,841 short insertions and deletions (indels) showed a strong and significant positive association with recombination (Spearman's $\rho$ = 0.54; p < 2.2× $10^{-16}$) (Fig. 3C). A similar pattern was observed when genic and intergenic indels were analyzed separately (Supplemental Fig. S13).

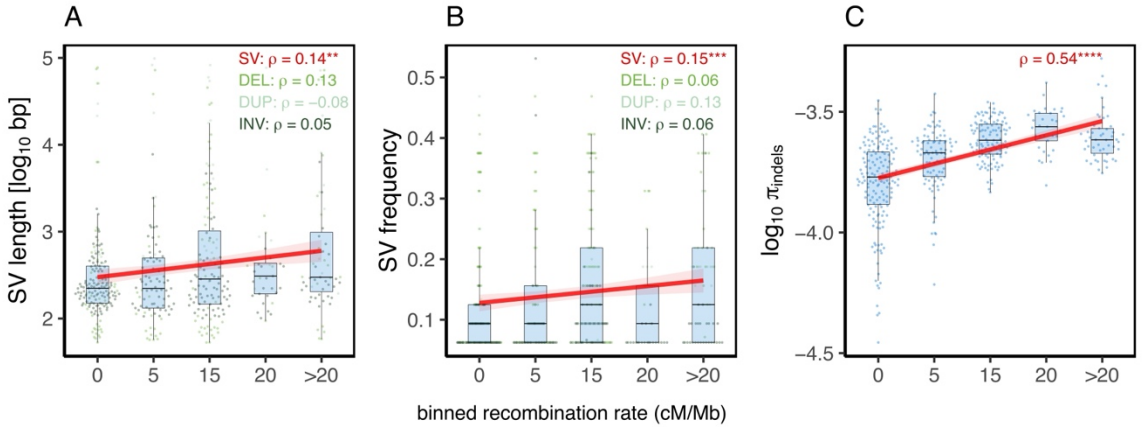



**Figure 3. Association between recombination rates and genome dynamics in *C. obscurior*.** Correlations between recombination rates and length (A) and frequency (B) of SVs and indels polymorphism (C) in 16 individual workers. The x-axis represents binned recombination rates, with the upper limit of each bin indicated. $\rho$: Spearman's rank correlation coefficient is indicated for all structural variants (SV) and separately for deletions (DEL), duplications (DUP) and inversions (INV), with significance shown by asterisks (****<0.0001; ***<0.001; **<0.01). All correlations were performed using the raw data. Red lines and shaded areas are linear regressions and 95% confidence intervals.

278

279 *No association between recombination and evolutionary rates*

280 Meiotic recombination is expected to enhance the efficacy of selection by uncoupling linkage
281 between targets of selection and mitigating the impact of Hill-Robertson interference (HRI)
282 (Betancourt et al. 2009; Campos et al. 2014; Hill and Robertson 1966). To examine this
283 relationship we used omega (ω) as a proxy, i.e., the ratio of nonsynonymous ($d_N$) to synonymous
284 ($d_S$) substitutions (Webster and Hurst 2012). In regions of low recombination, increased genetic
285 interference should compromise both positive selection for advantageous variants and purifying
286 selection against deleterious ones (Betancourt et al. 2009). Given that the majority of
287 nonsynonymous mutations are deleterious, a negative correlation between recombination rate
288 and ω is expected (Bolívar et al. 2016; Haddrill et al. 2007; Betancourt et al. 2009). In contrast,
289 we found no significant correlation between recombination and $d_N$ (Spearman's $\rho$ = 0.008, p =
290 0.61), $d_S$ (Spearman's $\rho$ = -0.015, p = 0.35), or ω (Spearman's $\rho$ = 0.019, p = 0.24) for 3,916
291 single-copy ortholog genes in LDRs (Supplemental Fig. S14).

292 *No association between recombination and gene expression bias*

293 In honey bees, genes with caste-biased expression in brains were enriched in regions of high
294 recombination, suggesting that caste-specific transcriptional plasticity is facilitated by the effects
295 of recombination (Kent et al. 2012; Wallberg et al. 2015; Jones et al. 2019; Liu et al. 2015). In *C.*
296 *obscurior* genes showing caste-biased brain expression are so far unknown. However, plastic
297 gene expression has been studied in this species in the context of caste differentiation (Schrader
298 et al. 2015, 2017). To assess a potential association between recombination and developmental
299 gene expression plasticity, we reanalyzed this RNA-seq dataset of 3$^{rd}$ instar larvae of different

castes and sexes (Supplemental Fig. S15). Based on 6,807 expressed genes, we did not find a significant association between recombination and gene expression plasticity over multiple contrasts (Spearman's $\rho = 0.014$, $p = 0.26$) (Supplemental Fig. S16).

*Functional analysis of genes in regions of high recombination*

To explore whether regions of high recombination rates in *C. obscurior* show signatures of functional specialization, we conducted Gene Ontology (GO) enrichment analyses on genes contained in the 10% most frequently recombining regions in the genome. These regions showing recombination rates over 22.3 cM/Mb and containing 1,541 genes were significantly enriched for fatty acid elongase and reductase genes (Table 2; Supplemental Table S7). While this implies a potential association between recombination and cuticular hydrocarbon (CHC) biosynthesis, it is important to acknowledge the limitations of these findings, particularly the failure to determine the specific genes directly affected by recombination.

**Table 2** GO terms enriched in regions of high recombination rates in *C. obscurior* genome (top 10%; recombination rate > 22.3 cM/Mb). *P* values are false discovery rate (FDR)-corrected

| GO.ID | Term | Annotated | Significant | Expected | *P* value |
|---|---|---|---|---|---|
| GO:0009922 | fatty acid elongase activity | 15 | 10 | 1.22 | $4.6 \times 10^{-6}$ |
| GO:0080019 | fatty-acyl-CoA reductase (alcohol-forming) activity | 19 | 9 | 1.54 | $6.17 \times 10^{-4}$ |

## Discussion

*Linkage mapping and recombination rate for C. obscurior*

The high recombination rates reported for the social Hymenoptera are not easily explained and have not been thoroughly investigated in ants. To start filling this gap we produced a high resolution genetic map. At least 26 of the 29 LGs represent full chromosomes, while the remaining small LGs are most likely unlinked fragments of larger chromosomes or extra ones in the mapping population. The updated chromosome-level genome of *C. obscurior* (Cobs3.1) confirms the architecture of the previous assembly-based versions and that it is the smallest ant genome described so far (Errbii et al. 2021; Schrader et al. 2014).

Average recombination rate in *C. obscurior* is 8.1 cM/Mb which is within rates reported for two ants (Sirviö et al. 2011b, 2006), and is lower compared to several honey bee species (Liu et al. 2015; Beye et al. 2006; Rueppell et al. 2016; Hunt and Page 1995). This value is likely an underestimate, due to two reasons. First, we excluded unlikely double crossover events as false positives. Although such double crossovers are highly improbable due to chiasma interference (Sturtevant 1915; Muller 1916), they may occur at low frequencies. Second, recombination in our experiment was suppressed in a large fraction of the genome (~38%), partly due to introgressed haplotypes in the grandparental cross. These regions are likely to recombine freely in the absence of introgression, which should increase average recombination rates quite significantly.

Two factors can affect the resolution of the map. First, the number of markers and their distribution across the genome. Despite the clustering of nonrecombining markers, which reduced the effective number of markers available for recombination rate estimation (from 57,309 markers to 1,061 clusters), the map maintained its density. The average inter-marker

338 distance in this study stands at 1.24 cM, and is relatively lower than the other two available ant
339 maps (*A. echinatior*: 145 markers or an inter-marker distance of 14.32 cM (Sirviö et al. 2006) and
340 *P. rugosus*: 215 markers or an inter-marker distance of 16.55 cM (Sirviö et al. 2011b)). A second
341 factor that affects map resolution is the size of the mapping population. Here, we used data
342 from 100 F2 individuals from a single queen, comparable to previous maps in ants (Sirviö et al.
343 2011b, 2006) and bees (Liu et al. 2017, 2015; Rueppell et al. 2016), and generally considered
344 sufficient for recombination rate estimation and the construction of a well resolved linkage map.

345 Recombination rates varied among the 29 LGs, showing a negative correlation with
346 chromosome size, a feature shared among many species (Pessia et al. 2012; Hunt and Page
347 1995; Jensen-Seaman et al. 2004; Haenel et al. 2018). Several LGs in our map were
348 approximately 50 cM long, aligning with the expectation under the one obligate crossover rule.
349 This raises the question whether the high recombination rates observed in eusocial species can
350 be attributed to an increase in the number of chromosomes with obligate crossovers. However,
351 chromosome numbers of eusocial and solitary species of Hymenoptera do not differ significantly
352 (Ross et al. 2015). Furthermore, these high rates can also not be linked to genome size
353 reductions in social Hymenoptera, as smaller genomes do not generally exhibit higher
354 recombination rates in animals (Stapley et al. 2017). Thus, neither chromosome number nor
355 genome size are sufficient to explain the high recombination rates of social Hymenoptera.

356 *The factors affecting recombination rates in C. obscurior*

357 Recombination rates, calculated in 250 kb windows, varied considerably (0 to 82.2 cM/Mb),
358 resulting in a complex recombinational landscape. Such regions of extremely high
359 recombination rates have also been observed in the honey bee (Beye et al. 2006; Liu et al.
360 2015). Consistent with previous observations indicating higher recombination rates in gene-

dense regions and open chromatin (Tiley and Burleigh 2015; Marand et al. 2017), gene-rich LDRs displayed higher recombination rates compared to gene-poor TE islands (Errbii et al. 2021; Schrader et al. 2014). This supports the interpretation, that TE islands are (peri-)centromeric regions, where recombination is usually suppressed (Beadle 1932; Talbert and Henikoff 2010; Lambie and Roeder 1986). TE insertions can accumulate in regions with reduced recombination, because ectopic recombination as a mechanism to remove TEs, is not available (Kent et al. 2017; Dolgin and Charlesworth 2008). In accordance, centromere structure and location in heterochromatin can locally suppress recombination (Talbert and Henikoff 2010), allowing accumulations of TEs in these regions to evolve secondarily. Future cytological examinations are necessary to confirm that these regions are indeed centromeres.

Sequence divergence resulting from introgression between distant populations of *C. obscurior* (Errbii et al. 2021) also inhibit recombination, in line with studies in other taxa (Liharska et al. 1996; Li et al. 2006; Opperman et al. 2004; Delame et al. 2019). Particularly, inversions that occur in some of these introgressions, can significantly impact recombination (Stevison et al. 2011; Rieseberg 2001; Noor et al. 2001) by disrupting the pairing of homologous chromosomes and synapsis formation during meiosis (Chakraborty et al. 2019).

Recombination was also suppressed in regions that showed low TE content and no signs of introgression, and significantly reduced levels of genetic diversity in two populations. Such pattern can emerge under epistatic selection (Navarro-Domínguez et al. 2022; Takahasi 2009; Hadany and Comeron 2008; Fuller et al. 2020; Otto and Lenormand 2002) which acts to reinforce linkage between co-adapted and/or essential loci through recruitment of recombination modifiers (Otto and Feldman 1997). These modifiers become associated with high-fitness allele combinations favored by selection, rendering any alternative allelic

384 combination deleterious and making recombinant offspring less viable (Barton 1995; Altenberg
385 and Feldman 1987; Singhal et al. 2019). Future studies examining the functional significance of
386 these regions are necessary to resolve whether they indeed evolved because of epistatic
387 selection of co-evolved loci.

388 Associations between GC content and recombination were described in the honey bee (Jones et
389 al. 2019; Wallberg et al. 2015) and other organisms (Pessia et al. 2012; Meunier and Duret
390 2004), most likely as a result of GC-biased gene conversion (Webster and Hurst 2012). However,
391 our investigation in *C. obscurior*, consistent with recent findings in butterflies (Torres et al.
392 2023), revealed no significant correlation between recombination rates and GC content. This
393 mirrors patterns seen in selfing plants like *Arabidopsis thaliana*, where it has been suggested to
394 result from increased inbreeding and a low $N_e$ (Marais et al. 2004).

395 *No support for hypotheses regarding the role of recombination in social evolution*

396 It has been suggested that recombination contributes to diversity around genes involved with
397 social life. By increasing the efficiency of natural selection (Comeron et al. 2008; Felsenstein
398 1974), high recombination rates may promote the evolution of caste-specific genes (Kent et al.
399 2012; Kent and Zayed 2013). Studies in *Drosophila* have suggested a link between
400 recombination and the efficacy of selection (Betancourt and Presgraves 2002; Campos et al.
401 2014; Haddrill et al. 2007), while other reports found no compelling evidence for such an
402 association (Webster and Hurst 2012), in line with data from humans (Bullaughey et al. 2008).
403 Consistent with a recent study in honey bees (Jones et al. 2019), in *C. obscurior* we found no
404 association between recombination rate and rates of $d_N/d_S$, used as a proxy for selection
405 efficiency. Hence, our finding do not support the hypothesis that high recombination rates

406 promoted the evolution of sociality by reducing HRI effects (Hill and Robertson 1966) and
407 allowing the evolution of genes important for social life.

408 Alternative models suggested that elevated rates of recombination in social Hymenoptera may
409 increase resistance to pathogens and facilitate behavioral and morphological differentiation by
410 increasing genetic diversity at the colony level (Hughes and Boomsma 2004; Shykoff and
411 Schmid-Hempel 1991; Wilfert et al. 2007; Liu et al. 2015; Wallberg et al. 2015; Jones et al. 2019).
412 We found no enrichment of genes involved in immunity or behavioral plasticity in regions of
413 high recombination. Additionally, there was no correlation between recombination rates and
414 genes showing plastic expression in larvae of *C. obscurior*, indicating that high recombination
415 rates are also not related to caste differentiation in this species.

416 GO term enrichment analyses revealed a significant enrichment of elongase and reductase
417 genes in regions of high recombination rates. These genes are primarily known for their roles in
418 regulating fatty acids elongation and their conversion into fatty alcohols during the biosynthesis
419 of CHCs (Holze et al. 2021), which play an important role in nestmate recognition in ants (Hefetz
420 2007; Sprenger and Menzel 2020), including *C. obscurior* (Drakula et al. 2023). Gene family
421 expansions of elongases and reductases have been suggested to contribute to the evolution of
422 CHC complexity (Finck et al. 2016; Hartke et al. 2019) and pheromone communication (Tupec et
423 al. 2019) in ants. To what extent high recombination rates could contribute to the function or
424 evolution of this key component of ant communication has so far not been explored. While the
425 direct impact of recombination on these genes or their surrounding regulatory sequences
426 remains unclear, it is plausible that frequent recombination at these loci could facilitate the
427 production of unique colony odors by increasing genotypic variation between colonies in genes
428 involved in the CHC biosynthesis pathway (Holze et al. 2021). Future studies should investigate

429 whether this pattern can be found in other ants and whether ectopic recombination might have
430 facilitated the expansion of these gene families.

431 *Recombination promotes genetic diversity of haplodiploid social insects with low $N_e$*

432 In haplodiploids, $N_e$ is expected to be ¾ that of diploid species, theoretically reducing genetic
433 variation and constraining adaptation to environmental variation (Kimura and Crow 1964;
434 Wright 1931). In eusocial haplodiploids, $N_e$ and consequently genetic variation are reduced
435 further (Romiguier et al. 2014; Pamilo et al. 1978; Metcalf et al. 1975; Pamilo and Crozier 1997),
436 as only a fraction of the individuals of a population reproduces (but see Weyna and Romiguier
437 2021 for other potential determinants of $N_e$ in Hymenoptera). Studies in taxa beyond
438 Hymenoptera, such as social shrimp (Chak et al. 2022) and termites (Roux et al. 2024),
439 corroborates the association between eusociality and reduced $N_e$. So, how is genetic variation as
440 the substrate for natural selection and species evolution maintained in eusocial haplodiploids?

441 Our study reveals a positive correlation between genetic diversity in populations of *C. obscurior*
442 and recombination rates, consistent with previous findings (Begun and Aquadro 1992; Spencer
443 et al. 2006; Roesti et al. 2013; Jones et al. 2019; Sirviö et al. 2006). Recombination rates also
444 correlated positively with SVs frequency and indel polymorphisms, likely due to ectopic
445 recombination—a significant mechanism driving SV emergence in various organisms, including
446 honey bees (Rueppell et al. 2016), *Drosophila* (Delprat et al. 2009; Montgomery et al. 1991) and
447 humans (Startek et al. 2015; Robberecht et al. 2013; Liu et al. 2012). Thus, high recombination
448 rates may function as a compensatory mechanism for the reduced genetic diversity observed in
449 haplodiploid social Hymenoptera with low $N_e$. While there is some reasoning that increased
450 recombination may evolve in small populations (Otto and Barton 2001), further investigations

451 are necessary to elucidate the relationship between $N_e$ and recombination rate evolution in
452 social and nonsocial haplodiploids, particularly using empirical approaches.

453 *Recombination is associated with genome dynamics and divergence in C. obscurior*

454 The correlation between genetic diversity and recombination in *C. obscurior* has important
455 implications for its genome dynamics and evolutionary processes. Through hitchhiking and
456 background selection, linked selection in regions with low recombination leads to lower
457 diversity and a skew towards rare alleles (Charlesworth et al. 1993; Smukowski and Noor 2011;
458 Ellegren and Galtier 2016; Comeron et al. 2008; Sella et al. 2009). Conversely, the divergence of
459 populations by genetic drift is expected to be stronger in regions where recombination rates and
460 thus genetic diversity are high (Ellegren and Galtier 2016; Hellmann et al. 2003). Accordingly, in
461 *C. obscurior*, recombination rates are positively correlated with genetic diversity and levels of
462 $d_{XY}$. However, there was no correlation with ω, consistent with a role of recombination in driving
463 within-species divergence, but not between-species divergence. Moreover, *C. obscurior* is a
464 highly inbred tramp species and its populations experience genetic bottlenecks, reducing
465 diversity in this species (Heinze 2017). Along with genetic hybridization and an increased TE
466 activity (Errbii et al. 2021), introduced populations of this species may benefit from the effects
467 of recombination on genetic diversity, providing yet another possible resolution to the "genetic
468 paradox of invasive species" (Frankham 2005).

## Methods

### *Mapping population*

*Cardiocondyla obscurior* originates from Southeast Asia, and colonies across the globe belong to two (or more) genetically distinct lineages (Errbii et al. 2021). In this study, we performed a cross involving a male and a queen collected from two New World populations separated by ca. 50km (Itabuna and Una, respectively, Brazil). These samples were collected in 2018 under permit 63371-1 issued by the Brazilian Ministério do Meio Ambiente. In haplodiploid species, haploid males develop from unfertilized eggs that represent the direct product of meiosis in females—the only recombining sex—making them ideal for linkage mapping. However, *C. obscurior* queens are not very productive (Jaimes-Nino et al. 2022) and cannot produce sufficient males for linkage mapping. To compensate, the F1 queen was mated with her brother (herein F1 male) to sample a total of 100 F2 individuals (12 males and 88 workers) at the pupal stage.

### *DNA extraction and sequencing*

DNA was extracted from the two grandparents, the F1 parents, and 100 F2 individuals using a modified CTAB (0.75 M NaCl, 50 mM Tris/HCl (pH = 8.0), 10 mM EDTA, 1% Hexadecyltrimethylammonium bromide) protocol (Doyle and Doyle 1987). Libraries were prepared using the Illumina DNA Prep kit with 5-6 PCR cycles, and 150-bp paired-end read sequencing was performed on an Illumina NovaSeq platform at the Cologne Center for Genomics. The grandmother and the F1 queen were sequenced to an average coverage of ~60x, the males to an average coverage of ~27x, and the workers to an average coverage of ~42x.

### *Read mapping and variant calling*

The quality of the raw reads was assessed using FastQC (v.0.11.7) (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trimmomatic (v.0.38) (Bolger et al. 2014) was used to remove short and low-quality reads as well as adapter sequences (options: ILLUMINACLIP:NexteraPE-PE.fa:2:30:10 SLIDINGWINDOW:4:20 MINLEN:40). The resulting paired reads were mapped to the Cobs2.1 assembly (Errbii et al. 2021) using BWA-MEM (v.0.7.17) (Li and Durbin 2009) with default parameters. The mapping quality was checked using Qualimap (v.2.2.2-dev) (Okonechnikov et al. 2016) (Supplemental Table S8), and duplicate reads were marked using Picard's MarkDuplicates (v.2.20.0). Joint-variant calling was performed using GATK's HaplotypeCaller (v.4.1.2.0) (McKenna et al. 2010) with default parameters.

*Variant filtering and phasing*

After removing indels, the remaining 342,508 biallelic single nucleotide polymorphisms (SNPs) were screened for high-quality markers using the following criteria: (i) Markers must be genotyped in both F1 parents and the grandparents. (ii) In the F1 queen, markers must be heterozygous, while in the grandparents, they must be homozygous for different alleles or heterozygous in the grandmother. Since males are haploid, each SNP is expected to be hemizygous in males. (iii) Consequently, heterozygous markers in at least one male, likely due to copy number variations, were discarded. (iv) Using BCFtools (v.1.14) (Danecek et al. 2021), genotypes with Mendelian errors were removed. Markers were further filtered using VCFtools (v.0.1.16) (Danecek et al. 2011) to only include those with a minor allele frequency (MAF) >0.2, genotyped across >80% of samples, a mean coverage between 19x and 50x, and at least 5 reads per genotype. In total, we identified 57,558 high-quality markers that were heterozygous in the F1 queen and were used as the basis for phasing (Supplemental Fig. S17).

The mapping population included 88 diploid F2 workers, each carrying the F1 father's haplotype and one of the two recombining haplotypes from the F1 queen (Supplemental Fig. S17). Using a custom script (Supplemental Code), we excluded the F1 male contribution from each of the F2 worker, retaining only the maternal contribution.

*Scaffold anchoring into chromosomes*

We improved the current version of the genome (Cobs2.1) by correcting misassembled (i.e., chimeric) and misoriented scaffolds. We used MSTmap (v.1.0) (Wu et al. 2008) with parameters specified in Supplemental Table S9, to order all 57,558 markers into LGs according to Gadau (2009). We then generated a BED file with the corrected genomic coordinates and orientation of scaffolds (Supplemental Table S10) based on the initial genetic map (Supplemental Fig. S1). Using BEDTools' *getfasta* (v.2.30.0) (Quinlan and Hall 2010) and Cobs2.1, we extracted and concatenated genomic sequences belonging to the same LG.

To accurately split chimeric scaffolds, we first checked for the existence of gaps separating the parts belonging to different LGs. When no gap was found, scaffolds were split after the last marker of the first part. For misoriented scaffolds, we inspected the correlation between the physical and genetic positions of markers. When the correlation was positive, the original orientation was kept, and when negative, the reverse complementary sequence was used. Scaffolds lacking markers and those that could not be mapped to any of the LGs were kept unplaced. The final chromosome-level assembly contained 29 LGs and 84 unplaced sequences.

*Gene and (tandem) repeat annotation*

For gene annotation, liftoff (v.1.6.3) (Shumate and Salzberg 2021) was used to transfer the Cobs2.1 annotation (Errbii et al. 2021) to Cobs3.1 with default parameters and the *-polish*

option. For repeat annotation, we used RepeatMasker (v.4.0.7) (http://www.repeatmasker.org) and a previous TE library (Errbii et al. 2021). TE islands were defined as previously described (Errbii et al. 2021).

The Tandem repeats finder tool (v.4.09.1) (Benson 1999) was used with recommended parameters (2 5 7 80 10 50 2000) to identify tandem repeats (TRs) in the genome. Additionally, TRs from five ant species (Melters et al. 2013; Huang et al. 2016) were used with RepeatMasker to explore their enrichment in the genome of *C. obscurior*. Further details on annotation procedures are provided in the Supplemental Methods.

*Estimation of recombination rate*

To calculate local recombination rates, Cobs3.1 was used as a reference following the pipeline described above for read mapping, variant calling, filtering, and phasing.

We visually inspected all markers and discarded erroneous genotypes and those that indicated unlikely double crossover events. These events were primarily identified as single-locus markers as well as blocks of two to three adjacent markers, tens to hundreds of base pairs long (Supplemental Figs. S18-S25), showing genotypic differences from immediately adjacent markers. Our reasoning to exclude these markers, was based on the expectations that double crossovers should be exceedingly rare due the suppression via chiasma interference (Muller 1916; Sturtevant 1915), where the occurrence of one crossover disrupts the formation of another on the same chromosome. The effect of this interference remains robust over long distances, ranging from hundreds of kilobases to tens of megabases (Hillers 2004; Fowler et al. 2018). We also excluded regions where we could not identify crossovers unambiguously due to low mapping accuracy and/or copy number variations (e.g., Supplemental Fig. S26).

Since all 84 unplaced scaffolds had insufficient markers (≤ 2), we only used markers on the 29 refined LGs. The final map was based on 57,532 high-quality markers that were clustered into 29 LGs (Supplemental Fig. S3) using MSTmap (Supplemental Table S9). Crossovers were defined as genotype changes between the maternal haplotypes, with most crossover tracks spanning over 100 kb (Supplemental Figs. S22, S24 and S25).

For local recombination rates, markers on each LG were ordered according to their physical distance and we removed 223 markers showing discrepancies between their physical and genetic positions. Subsequently, nonrecombining markers were clustered according to their shared genetic positions, resulting in 1,061 effectively recombining marker clusters. Local recombination rates were then inferred by calculating the ratio of the genetic (in cM) to the physical (Mb) distances between adjacent clusters. Average distance between crossover events was 149 kb (95% CI [122, 177]). To ensure that genomic regions had a proportional influence on estimates of recombination rate and to account for the non-uniform distribution of crossovers along the LGs, we calculated a weighted mean of recombination rate in 250 kb nonoverlapping windows. To this end, each window was assigned the recombination rate of the corresponding inter-cluster interval. When a window matched multiple intervals, the mean recombination rate of those intervals, weighted by their physical length, was assigned (Supplemental Code).

*Estimating genetic diversity and differentiation*

Nucleotide diversity ($\pi$) and Tajima's *D* were calculated in 250 kb nonoverlapping windows using two pool-seq datasets of *C. obscurior*, a pool of 30 workers from Itabuna, Brazil (Errbii et al. 2021), and a pool of 50 workers from ca. 50km distant location in Una, Brazil. Following Errbii et al. (2021), we filtered paired reads, mapped them to Cobs3.1, and generated an indel-free mpileup file for each population using SAMtools (v.1.15.1) (Li et al. 2009). We used the

PoPoolation (Kofler et al. 2011a) *Variance-sliding.pl* script to calculate $\pi$ and Tajima's *D*. For $\pi$ at synonymous ($\pi_S$) and nonsynonymous ($\pi_{NS}$) sites, we used the PoPoolation *Syn-nonsyn-sliding.pl* script, the mpileup files and the codon and nonsynonymous codon length tables from PoPoolation.

For genetic differentiation ($F_{ST}$) between lineages of *C. obscurior*, we combined the alignment files of the two pools with a third alignment file of a pool of 16 workers from Tenerife, Spain (Errbii et al. 2021), to create a single mpileup file using SAMtools. We then converted this file into a synchronized file and used the *fst-sliding.pl* script under PoPoolation2 (Kofler et al. 2011b) to calculate $F_{ST}$ in 250 kb nonoverlapping windows.

*Calculation of absolute divergence*

To explore genetic divergence between lineages of *C. obscurior* ($d_{XY}$) and the grandparental haplotypes (gp-$d_{XY}$), we calculated absolute divergence in 250 kb windows using scripts available at https://github.com/simonhmartin/genomics_general.

Published paired-end short reads data from single workers of the Old World (Taiwan and Leiden) and New World (Itabuna and Una) lineages of *C. obscurior* (Errbii et al. 2021) were filtered and mapped to Cobs3.1 as described above. Note that reads from both lineages can be aligned to the reference genome with high confidence (Supplemental Fig. S27). Further details on SNP calling and filtering are given in the Supplemental Methods.

*Structural variant identification*

To identify SVs, we adopted an approach by Mérot et al. (Mérot et al. 2023) using published paired-end short read data from 16 workers from five populations (Errbii et al. 2021). Reads were filtered and mapped to Cobs3.1, and read duplicates were removed using Picard's

MarkDuplicates. SVs were identified using three different tools: Manta (v.1.6.0) (Chen et al. 2016), smoove (v.0.2.8) (https://github.com/brentp/smoove) which is based on LUMPY (Layer et al. 2014), and DELLY (v.0.8.1) (Rausch et al. 2012), all run with default parameters. To identify genomic inversions potentially affecting recombination in the F1 queen, we used the three tools and followed the same approach. Subsequently, both sets of variants were filtered following protocols detailed in the Supplemental Methods.

For indels, we used the raw variant calls generated by GATK's HaplotypeCaller for the 16 workers. Indels were extracted and subjected to GATK's recommended hard filtering criteria (QD < 2.0; QUAL < 30.0; FS > 200 and ReadPosRankSum < -20.0). See Supplemental Methods for further details on the filtering process and diversity calculation.

*Genome parameters estimation*

BEDTools and BEDOPS (v.2.4.37) (Neph et al. 2012), along with the gene annotation were used to calculate gene content. Overall and intergenic GC contents were calculated using the BEDTools *nuc* function in 250 kb nonoverlapping windows.

*Estimating substitution rates*

To explore the correlation between recombination and the efficiency of selection, we calculated omega (ω) for single-copy ortholog genes between *C. obscurior* and three ant species (*Monomorium pharaonis* (GCA013373865v2), *Solenopsis invicta* (UNIL_Sinv_3.0) and *Ooceraea biroi* (Obir_v5.4)) and the parasitoid wasp *Nasonia vitripennis* (Nvit_psr_1.1). Genomic and protein sequences, and gene annotations, were downloaded from Ensembl (release-54). To estimate the branch-wide rate ω and rates of synonymous ($d_N$) and synonymous mutations ($d_S$), we applied HyPhy's aBSREL (v.2.5.48) (adaptive Branch-Site Random Effects Likelihood) (Smith

et al. 2015) and the FitMG94 model, implemented in HyPhy. Further details on single-copy ortholog retrieval and ω calculation are given in the Supplemental Methods.

*Gene expression bias*

To investigate the relationship between local recombination rate and developmental gene expression bias, we calculated the plasticity index (Schrader et al. 2017), capturing gene expression bias across multiple caste/morph contrasts. High values indicate caste-specific gene expression, while low values suggest uniform expression across castes. RNA-seq data of worker, queen, wingless male, and winged male third instar larvae (n = 7 individuals each) were retrieved from NCBI (BioProject: PRJNA237579) (Schrader et al. 2015, 2017). See Supplemental Methods for detailed information on read filtering and mapping, differential expression analysis, and computation of plasticity indices.

*GO term enrichment analysis*

To test whether genes found in regions of high recombination rates are enriched for specific functions, we ran a GO enrichment analysis. We defined these regions as windows with recombination rates falling within the top 10% (recombination rate > 22.3 cM/Mb). We used BEDTools' *intersect* to extract 1,541 genes located in these regions and ran topGO (Alexa and Rahnenführer 2020) to identify biological processes enriched in these regions. TE islands and regions affected by introgression were excluded from the analyses.

*Correlation analyses and plotting*

All correlations using the raw data and plots were produced in R (v.4.1.3) (R Core Team 2020). To explore the associations of TE content, sequence divergence and epistasis with recombination rates per 250 kb, a generalized linear model was fitted using the glm function in

646 R. We used Circos (Krzywinski et al. 2009) for graphical visualization in R. LGs 9, 10, 11, 17, and
647 21 were excluded from all correlation analyses, as introgression is evident in a significant
648 fraction of these LGs. Further, LGs 26 and 29 were excluded because of low marker coverage
649 (Supplemental Fig. S6).

## Data access

The genome assembly, annotation and raw sequencing data generated in this study have been submitted to the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under accession number PRJNA934066. Intermediate data (Errbii 2023) required to reproduce the analyses is provided as Supplemental Data and is also available on figshare at https://doi.org/10.6084/m9.figshare.22178717.v2. A detailed description of the bioinformatic analysis pipelines is available as Supplemental Code and on GitHub at https://github.com/merrbii/CobsLinkagemapping.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 403813881 with grants to L.S. (SCHR 1554/2-1) and J.O. (OE 549/4-1) under the priority program "Rapid evolutionary adaptation: Potential and constraints" (SPP 1819). We thank the DFG Research Training Group 2220 "Evolutionary Processes in Adaptation and Disease" and the CORE environment (cardiocondyla.org) for supporting M.E. J.G. acknowledges support from the DFG as part of the SFB TRR 212 ($NC^3$) – TP C04 project numbers 316099922 and 396780988. We further thank Tobias van Elst, Jacques Delabie and Eva Schultner for help collecting ants in the field. We thank Xim Cerda for his support to collect ants in Spain under permission of Real Decreto 124/2017. We thank Simon H. Martin for his valuable advice on the analysis of sequence divergence. We are very grateful to three anonymous reviewers for their constructive comments.

## Authors' Contributions

J.O. and L.S. conceptualized the study and acquired funding. M.E., J.G., J.O. and L.S. designed the experiments. M.E. produced the data and performed the analyses. K.B. was responsible for Illumina sequencing. M.E., L.S., and J.O. wrote the manuscript, all authors contributed to the final version.

## References


Alexa A, Rahnenführer J. 2020. Gene set enrichment analysis with topGO. R package version 2.46.0

Altenberg L, Feldman MW. 1987. Selection, generalized transmission and the evolution of modifier genes. I. The reduction principle. *Genetics* **117**: 559–572.

Arbeithuber B, Betancourt AJ, Ebner T, Tiemann-Boege I. 2015. Crossovers are associated with mutation and biased gene conversion at recombination hotspots. *Proc Natl Acad Sci U S A* **112**: 2109–2114.

Barton NH. 1995. A general model for the evolution of recombination. *Genet Res (Camb)* **65**: 123–144.

Beadle GW. 1932. A Possible Influence of the Spindle Fibre on Crossing-Over in Drosophila. *Proc Natl Acad Sci* **18**: 160–165.

Begun DJ, Aquadro CF. 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in D. melanogaster. *Nature* **356**: 519–520.

Bell G. 1982. *The masterpiece of nature: the evolution and genetics of sexuality*. 1st ed. University of California Press, Berkeley.

Benson G. 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res* **27**: 573–580.

Betancourt AJ, Presgraves DC. 2002. Linkage limits the power of natural selection in Drosophila. *Proc Natl Acad Sci U S A* **99**: 13616–13620.

Betancourt AJ, Welch JJ, Charlesworth B. 2009. Reduced Effectiveness of Selection Caused by a Lack of Recombination. *Curr Biol* **19**: 655–660.

Beye M, Gattermeier I, Hasselmann M, Gempe T, Schioett M, Baines JF, Schlipalius D, Mougel F, Emore C, Rueppell O, et al. 2006. Exceptionally high levels of recombination across the honey bee genome. *Genome Res* **16**: 1339–1344.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Boĺivar P, Mugal CF, Nater A, Ellegren H. 2016. Recombination rate variation modulates gene sequence evolution mainly via GC-Biased gene conversion, not Hill-Robertson interference, in an avian system. *Mol Biol Evol* **33**: 216–227.

Bullaughey K, Przeworski M, Coop G. 2008. No effect of recombination on the efficacy of natural selection in primates. *Genome Res* **18**: 544–554.

Burri R. 2017. Interpreting differentiation landscapes in the light of long-term linked selection. *Evol Lett* **1**: 118–131.

Campos JL, Halligan DL, Haddrill PR, Charlesworth B. 2014. The relation between recombination rate and patterns of molecular evolution and variation in drosophila melanogaster. *Mol Biol Evol* **31**: 1010–1028.

Chak STC, Harris SE, Hultgren KM, Duffy JE, Rubenstein DR. 2022. Demographic Inference Provides Insights into the Extirpation and Ecological Dominance of Eusocial Snapping Shrimps. *J Hered* **113**: 552–562.

Chakraborty U, Mackenroth B, Shalloway D, Alani E. 2019. Chromatin modifiers alter recombination between divergent DNA sequences. *Genetics* **212**: 1147–1162.

Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**: 1289–1303.

Charlesworth D. 2016. The status of supergenes in the 21st century: Recombination suppression

in Batesian mimicry and sex chromosomes and other complex adaptations. *Evol Appl* **9**: 74–90.

Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, Cox AJ, Kruglyak S, Saunders CT. 2016. Manta: Rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **32**: 1220–1222.

Comeron JM, Williford A, Kliman RM. 2008. The Hill-Robertson effect: Evolutionary consequences of weak selection and linkage in finite populations. *Heredity (Edinb)* **100**: 19–31.

Crow JF. 1994. Advantages of sexual reproduction. *Dev Genet* **15**: 205–213.

Cutter AD, Payseur BA. 2013. Genomic signatures of selection at linked sites: Unifying the disparity among species. *Nat Rev Genet* **14**: 262–274.

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.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* **10**.

Delame M, Prado E, Blanc S, Robert-Siegwald G, Schneider C, Mestre P, Rustenholz C, Merdinoglu D. 2019. Introgression reshapes recombination distribution in grapevine interspecific hybrids. *Theor Appl Genet* **132**: 1073–1087.

Delprat A, Negre B, Puig M, Ruiz A. 2009. The transposon Galileo generates natural chromosomal inversions in Drosophila by ectopic recombination. *PLoS One* **4**: e7883.

Dolgin ES, Charlesworth B. 2008. The effects of recombination rate on the distribution and abundance of transposable elements. *Genetics* **178**: 2169–77.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* **19**: 11–15.

Drakula M, Buellesbach J, Schrader L. 2023. The role of cuticular hydrocarbons in intraspecific aggression in the invasive ant Cardiocondyla obscurior. *Myrmecol News* **33**: 187–196.

Ellegren H, Galtier N. 2016. Determinants of genetic diversity. *Nat Rev Genet* **17**: 422–433.

Errbii M. 2023. Recombination rate variation in Cardiocondyla obscurior. https://doi.org/10.6084/m9.figshare.22178717.v2.

Errbii M, Keilwagen J, Hoff KJ, Steffen R, Altmüller J, Oettler J, Schrader L. 2021. Transposable elements and introgression introduce genetic variation in the invasive ant Cardiocondyla obscurior. In *Molecular Ecology*, Vol. 30 of, pp. 6211–6228, John Wiley & Sons, Ltd.

Felsenstein J. 1974. The evolution advantage of recombination. *Genetics* **78**: 737–756.

Felsenstein J, Yokoyama S. 1976. The evolutionary advantage of recombination. II. Individual selection for recombination. *Genetics* **83**: 845–859.

Finck J, Berdan EL, Mayer F, Ronacher B, Geiselhardt S. 2016. Divergence of cuticular hydrocarbons in two sympatric grasshopper species and the evolution of fatty acid synthases and elongases across insects. *Sci Rep* **6**: 1–13.

Fowler KR, Hyppa RW, Cromie GA, Smith GR. 2018. Physical basis for long-distance communication along meiotic chromosomes. *Proc Natl Acad Sci U S A* **115**: E9333–E9342.

Frankham R. 2005. Resolving the genetic paradox in invasive species. *Heredity (Edinb)* **94**: 385–385.

Fuller ZL, Koury SA, Leonard CJ, Young RE, Ikegami K, Westlake J, Richards S, Schaeffer SW, Phadnis N. 2020. Extensive recombination suppression and epistatic selection causes chromosome-wide differentiation of a selfish sex chromosome in drosophila

768 pseudoobscura. *Genetics* **216**: 205–226.
769 Gadau J. 2009. Phase-unknown linkage mapping in ants. *Cold Spring Harb Protoc* **4**.
770 Gadau J, Page RE, Werren JH. 1999. Mapping of hybrid incompatibility loci in Nasonia. *Genetics*
771 **153**: 1731–1741.
772 Hadany L, Comeron JM. 2008. Why Are Sex and Recombination So Common? *Ann N Y Acad Sci*
773 **1133**: 26–43.
774 Haddrill PR, Halligan DL, Tomaras D, Charlesworth B. 2007. Reduced efficacy of selection in
775 regions of the Drosophila genome that lack crossing over. *Genome Biol* **8**: 1–9.
776 Haenel Q, Laurentino TG, Roesti M, Berner D. 2018. Meta-analysis of chromosome-scale
777 crossover rate variation in eukaryotes and its significance to evolutionary genomics. *Mol*
778 *Ecol* **27**: 2477–2497.
779 Halldorsson B V., Palsson G, Stefansson OA, Jonsson H, Hardarson MT, Eggertsson HP,
780 Gunnarsson B, Oddsson A, Halldorsson GH, Zink F, et al. 2019. Characterizing mutagenic
781 effects of recombination through a sequence-level genetic map. *Science (80- )* **363**.
782 Hartke J, Schell T, Jongepier E, Schmidt H, Sprenger PP, Paule J, Bornberg-Bauer E, Schmitt T,
783 Menzel F, Pfenninger M, et al. 2019. Hybrid Genome Assembly of a Neotropical Mutualistic
784 Ant. *Genome Biol Evol* **11**: 2306–2311.
785 Hefetz A. 2007. The evolution of hydrocarbon pheromone parsimony in ants (Hymenoptera:
786 Formicidae) - interplay of colony odor uniformity and odor idiosyncrasy. A review.
787 *Myrmecological News* **10**: 59–68.
788 Heinze J. 2017. Life-history evolution in ants: the case of Cardiocondyla. *Proc R Soc B Biol Sci*
789 **284**: 20161406.
790 Hellmann I, Ebersberger I, Ptak SE, Pääbo S, Przeworski M. 2003. A neutral explanation for the
791 correlation of diversity with recombination rates in humans. *Am J Hum Genet* **72**: 1527–
792 1535.
793 Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genet Res* **8**:
794 269–294.
795 Hillers KJ. 2004. Crossover interference. *Curr Biol* **14**: R1036–R1037.
796 Holze H, Schrader L, Buellesbach J. 2021. Advances in deciphering the genetic basis of insect
797 cuticular hydrocarbon biosynthesis and variation. *Heredity (Edinb)* **126**: 219–234.
798 Huang YC, Lee CC, Kao CY, Chang NC, Lin CC, Shoemaker D, Wang J. 2016. Evolution of long
799 centromeres in fire ants. *BMC Evol Biol* **16**.
800 Hughes WOH, Boomsma JJ. 2004. Genetic diversity and disease resistance in leaf-cutting ant
801 societies. *Evolution (N Y)* **58**: 1251–1260.
802 Hunt GJ, Page RE. 1995. Linkage map of the honey bee, Apis mellifera, based on RAPD markers.
803 *Genetics* **139**: 1371–1382.
804 Jaimes-Nino LM, Heinze J, Oettler J. 2022. Late-life fitness gains and reproductive death in
805 Cardiocondyla obscurior ants. *Elife* **11**.
806 Jensen-Seaman MI, Furey TS, Payseur BA, Lu Y, Roskin KM, Chen CF, Thomas MA, Haussler D,
807 Jacob HJ. 2004. Comparative recombination rates in the rat, mouse, and human genomes.
808 *Genome Res* **14**: 528–538.
809 Jones GH, Franklin FCH. 2006. Meiotic Crossing-over: Obligation and Interference. *Cell* **126**: 246–
810 248.
811 Jones JC, Wallberg A, Christmas MJ, Kapheim KM, Webster MT. 2019. Extreme Differences in
812 Recombination Rate between the Genomes of a Solitary and a Social Bee. *Mol Biol Evol* **36**:
813 2277–2291.

814 Kent CF, Minaei S, Harpur BA, Zayed A. 2012. Recombination is associated with the evolution of
815 genome structure and worker behavior in honey bees. *Proc Natl Acad Sci U S A* **109**:
816 18012–18017.
817 Kent CF, Zayed A. 2013. Evolution of recombination and genome structure in eusocial insects.
818 *Commun Integr Biol* **6**: e22919.
819 Kent T V, Uzunović J, Wright SI. 2017. Coevolution between transposable elements and
820 recombination. *Philos Trans R Soc B Biol Sci* **372**.
821 Kimura M, Crow JF. 1964. THE NUMBER OF ALLELES THAT CAN BE MAINTAINED IN A FINITE
822 POPULATION. *Genetics* **49**: 725–738.
823 Kleckner N. 1996. Meiosis: How could it work? *Proc Natl Acad Sci U S A* **93**: 8167–8174.
824 Kofler R, Orozco-terWengel P, de Maio N, Pandey RV, Nolte V, Futschik A, Kosiol C, Schlötterer C.
825 2011a. Popoolation: A toolbox for population genetic analysis of next generation
826 sequencing data from pooled individuals. *PLoS One* **6**: 15925.
827 Kofler R, Pandey RV, Schlötterer C. 2011b. PoPoolation2: Identifying differentiation between
828 populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* **27**: 3435–
829 3436.
830 Kozomara A, Birgaoanu M, Griffiths-Jones S. 2019. MiRBase: From microRNA sequences to
831 function. *Nucleic Acids Res* **47**: D155–D162.
832 Kraus B, Page RE. 1998. Parasites, pathogens, and polyandry in social insects. *Am Nat* **151**: 383–
833 391.
834 Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009.
835 Circos: An information aesthetic for comparative genomics. *Genome Res* **19**: 1639–1645.
836 Lambie EJ, Roeder GS. 1986. Repression of meiotic crossing over by a centromere (CEN3) in
837 Saccharomyces cerevisiae. *Genetics* **114**: 769–789.
838 Langley CH, Montgomery E, Hudson R, Kaplan N, Charlesworth B. 1988. On the role of unequal
839 exchange in the containment of transposable element copy number. *Genet Res (Camb)* **52**:
840 223–235.
841 Layer RM, Chiang C, Quinlan AR, Hall IM. 2014. LUMPY: A probabilistic framework for structural
842 variant discovery. *Genome Biol* **15**: 1–19.
843 Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform.
844 *Bioinformatics* **25**: 1754–1760.
845 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009.
846 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
847 Li L, Jean M, Belzile F. 2006. The impact of sequence divergence and DNA mismatch repair on
848 homeologous recombination in Arabidopsis. *Plant J* **45**: 908–916.
849 Liharska T, Koornneef M, Van Wordragen M, Van Kammen A, Zabel P. 1996. Tomato
850 chromosome 6: Effect of alien chromosomal segments on recombinant frequencies.
851 *Genome* **39**: 485–491.
852 Liu H, Jia Y, Sun X, Tian D, Hurst LD, Yang S. 2017. Direct determination of the mutation rate in
853 the bumblebee reveals evidence for weak recombination-associated mutation and an
854 approximate rate constancy in insects. *Mol Biol Evol* **34**: 119–130.
855 Liu H, Zhang X, Huang J, Chen J-Q, Tian D, Hurst LD, Yang S. 2015. Causes and consequences of
856 crossing-over evidenced via a high-resolution recombinational landscape of the honey bee.
857 *Genome Biol* **16**: 15.
858 Liu P, Carvalho CMB, Hastings PJ, Lupski JR. 2012. Mechanisms for recurrent and complex
859 human genomic rearrangements. *Curr Opin Genet Dev* **22**: 211–220.

Lynn A, Ashley T, Hassold T. 2004. Variation in human meiotic recombination. *Annu Rev Genomics Hum Genet* **5**: 317–349.
Marais G, Charlesworth B, Wright SI. 2004. Recombination and base composition: the case of the highly self-fertilizing plant Arabidopsis thaliana. *Genome Biol* **5**: 1–9.
Marand AP, Jansky SH, Zhao H, Leisner CP, Zhu X, Zeng Z, Crisovan E, Newton L, Hamernik AJ, Veilleux RE, et al. 2017. Meiotic crossovers are associated with open chromatin and enriched with Stowaway transposons in potato. *Genome Biol* **18**: 1–16.
Maynard Smith J. 1978. *The evolution of sex*. Cambridge University Press, Cambridge.
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**: 1297–1303.
Melters DP, Bradnam KR, Young HA, Telis N, May MR, Ruby JG, Sebra R, Peluso P, Eid J, Rank D, et al. 2013. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol* **14**: R10.
Mérot C, Stenløkk KSR, Venney C, Laporte M, Moser M, Normandeau E, Árnyasi M, Kent M, Rougeux C, Flynn JM, et al. 2023. Genome assembly, structural variants, and genetic differentiation between lake whitefish young species pairs (Coregonus sp.) with long and short reads. *Mol Ecol* **32**: 1458–1477.
Metcalf RA, Marlin JC, Whitt GS. 1975. Low levels of genetic heterozygosity in Hymenoptera. *Nature* **257**: 792–794.
Meunier J, Duret L. 2004. Recombination drives the evolution of GC-content in the human genome. *Mol Biol Evol* **21**: 984–990.
Montgomery EA, Huang SM, Langley CH, Judd BH. 1991. Chromosome rearrangement by ectopic recombination in Drosophila melanogaster: genome structure and evolution. *Genetics* **129**: 1085–1098.
Muller HJ. 1916. The Mechanism of Crossing-Over. *Am Nat* **50**: 193–221.
Navarro-Domínguez B, Chang CH, Brand CL, Muirhead CA, Presgraves DC, Larracuente AM. 2022. Epistatic selection on a selfish Segregation Distorter supergene: drive, recombination, and genetic load. *Elife* **11**: 78981.
Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, Rynes E, Maurano MT, Vierstra J, Thomas S, et al. 2012. BEDOPS: High-performance genomic feature operations. *Bioinformatics* **28**: 1919–1920.
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.
Nielsen R. 2005. Molecular Signatures of Natural Selection. *Annu Rev Genet* **39**: 197–218.
Noor MAF, Gratos KL, Bertucci LA, Reiland J. 2001. Chromosomal inversions and the reproductive isolation of species. *Proc Natl Acad Sci U S A* **98**: 12084–12088.
Oettler J. 2020. Cardiocondyla: Heart Node Ants. In *Encyclopedia of Social Insects*, pp. 1–3, Springer International Publishing.
Okonechnikov K, Conesa A, García-Alcalde F. 2016. Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**: 292–294.
Oldroyd BP, Fewell JH. 2007. Genetic diversity promotes homeostasis in insect colonies. *Trends Ecol Evol* **22**: 408–413.
Opperman R, Emmanuel E, Levy AA. 2004. The effect of sequence divergence on recombination between direct repeats in arabidopsis. *Genetics* **168**: 2207–2215.

Otto SP, Barton NH. 2001. Selection for recombination in small populations. *Evolution (N Y)* **55**: 1921–1931.

Otto SP, Feldman MW. 1997. Deleterious mutations, variable epistatic interactions, and the evolution of recombination. *Theor Popul Biol* **51**: 134–147.

Otto SP, Lenormand T. 2002. Resolving the paradox of sex and recombination. *Nat Rev Genet* **3**: 252–261.

Pamilo P, Crozier RH. 1997. Population biology of social insect conservation. *Mem Museum Victoria* **56**: 411–419.

Pamilo P, Varvio-Aho S, Pekkarinen A. 1978. Low enzyme gene variability in Hymenoptera as a consequence of haplodiploidy. *Hereditas* **88**: 93–99.

Pessia E, Popa A, Mousset S, Rezvoy C, Duret L, Marais GAB. 2012. Evidence for widespread GC-biased gene conversion in eukaryotes. *Genome Biol Evol* **4**: 675–682.

Quinlan AR, Hall IM. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.

R Core Team. 2020. *A Language and Environment for Statistical Computing*. 4.0.2. Vienna.

Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. 2012. DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**: i333–i339.

Rieseberg LH. 2001. Chromosomal rearrangements and speciation. *Trends Ecol Evol* **16**: 351–358.

Ritz KR, Noor MAF, Singh ND. 2017. Variation in Recombination Rate: Adaptive or Not? *Trends Genet* **33**: 364–374.

Robberecht C, Voet T, Esteki MZ, Nowakowska BA, Vermeesch JR. 2013. Nonallelic homologous recombination between retrotransposable elements is a driver of de novo unbalanced translocations. *Genome Res* **23**: 411–418.

Roeder GS. 1997. Meiotic chromosomes: It takes two to tango. *Genes Dev* **11**: 2600–2621.

Roesti M, Moser D, Berner D. 2013. Recombination in the threespine stickleback genome - Patterns and consequences. *Mol Ecol* **22**: 3014–3027.

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.

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.

Roux C, Weyna A, Lodé M, Romiguier J. 2024. Social evolution in termites reduces natural selection efficacy. *bioRxiv* 2024.04.26.591327.

Rueppell O, Kuster R, Miller K, Fouks B, Correa SR, 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.

Savocco J, Piazza A. 2021. Recombination-mediated genome rearrangements. *Curr Opin Genet Dev* **71**: 63–71.

Schmid-Hempel P. 1998. *Parasites in social insects*. Princeton University Press.

Schrader L, Helantera H, Oettler J. 2017. Accelerated evolution of developmentally biased genes in the tetraphenic ant cardiocondyla obscurior. *Mol Biol Evol* **34**: 535–544.

Schrader L, Kim JW, Ence D, Zimin A, Klein A, Wyschetzki K, Weichselgartner T, Kemena C, Stökl J, Schultner E, et al. 2014. Transposable element islands facilitate adaptation to novel environments in an invasive species. *Nat Commun* **5**: 5495.

952 Schrader L, Simola DF, Heinze J, Oettler J. 2015. Sphingolipids, transcription factors, and
953 conserved toolkit genes: Developmental plasticity in the ant cardiocondyla obscurior. *Mol*
954 *Biol Evol* **32**: 1474–1486.
955 Sella G, Petrov DA, Przeworski M, Andolfatto P. 2009. Pervasive natural selection in the
956 Drosophila genome? *PLoS Genet* **5**: e1000495.
957 Shi YY, Sun LX, Huang ZY, Wu XB, Zhu YQ, Zheng HJ, Zeng ZJ. 2013. A SNP Based High-Density
958 Linkage Map of Apis cerana Reveals a High Recombination Rate Similar to Apis mellifera.
959 *PLoS One* **8**: 8–13.
960 Shumate A, Salzberg SL. 2021. Liftoff: accurate mapping of gene annotations. *Bioinformatics* **37**:
961 1639–1643.
962 Shykoff JA, Schmid-Hempel P. 1991. Parasites and the advantage of genetic variability within
963 social insect colonies. *Proc R Soc London Ser B Biol Sci* **243**: 55–58.
964 Singhal S, Gomez SM, Burch CL. 2019. Recombination drives the evolution of mutational
965 robustness. *Curr Opin Syst Biol* **13**: 142–149.
966 Sirviö A, Gadau J, Rueppell O, Lamatsch D, Boomsma JJ, Pamilo P, Page RE. 2006. High
967 recombination frequency creates genotypic diversity in colonies of the leaf-cutting ant
968 Acromyrmex echinatior. *J Evol Biol* **19**: 1475–1485.
969 Sirviö A, Johnston JS, Wenseleers T, Pamilo P. 2011a. A high recombination rate in eusocial
970 Hymenoptera: Evidence from the common wasp Vespula vulgaris. *BMC Genet* **12**: 1–7.
971 Sirviö A, Pamilo P, Johnson RA, Page RE, Gadau J. 2011b. Origin and evolution of the dependent
972 lineages in the genetic caste determination system of Pogonomyrmex ants. *Evolution (N Y)*
973 **65**: 869–84.
974 Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. 2015. Less is
975 more: An adaptive branch-site random effects model for efficient detection of episodic
976 diversifying selection. *Mol Biol Evol* **32**: 1342–1353.
977 Smukowski CS, Noor MAF. 2011. Recombination rate variation in closely related species.
978 *Heredity (Edinb)* **107**: 496–508.
979 Spencer CCA, Deloukas P, Hunt S, Mullikin J, Myers S, Silverman B, Donnelly P, Bentley D,
980 McVean G. 2006. The influence of recombination on human genetic diversity. *PLoS Genet*
981 **2**: 1375–1385.
982 Sprenger PP, Menzel F. 2020. Cuticular hydrocarbons in ants (Hymeno ptera: Formicidae) and
983 other insects: how and why they differ among individuals, colonies, and species. *Myrmecol*
984 *News* **30**: 1–26.
985 Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. 2017. Variation in recombination
986 frequency and distribution across eukaryotes: Patterns and processes. *Philos Trans R Soc B*
987 *Biol Sci* **372**: 20160455.
988 Startek M, Szafranski P, Gambin T, Campbell IM, Hixson P, Shaw CA, Stankiewicz P, Gambin A.
989 2015. Genome-wide analyses of LINE-LINE-mediated nonallelic homologous
990 recombination. *Nucleic Acids Res* **43**: 2188–2198.
991 Stevison LS, Hoehn KB, Noor MAF. 2011. Effects of inversions on within- and between-species
992 recombination and divergence. *Genome Biol Evol* **3**: 830–841.
993 Sturtevant AH. 1915. The behavior of the chromosomes as studied through linkage. *Z Indukt*
994 *Abstamm Vererbungsl* **13**: 234–287.
995 Takahasi KR. 2009. Coalescent under the evolution of coadaptation. *Mol Ecol* **18**: 5018–5029.
996 Talbert PB, Henikoff S. 2010. Centromeres convert but don't cross. *PLoS Biol* **8**: e1000326.
997 Tiley GP, Burleigh G. 2015. The relationship of recombination rate, genome structure, and

patterns of molecular evolution across angiosperms. *BMC Evol Biol* **15**.
Torres AP, Höök L, Näsvall K, Shipilina D, Wiklund C, Vila R, Pruisscher P, Backström N. 2023. The fine-scale recombination rate variation and associations with genomic features in a butterfly. *Genome Res* **33**: 810–823.
Tupec M, Buček A, Janoušek V, Vogel H, Prchalová D, Kindl J, Pavlíčková T, Wenzelová P, Jahn U, Valterová I, et al. 2019. Expansion of the fatty acyl reductase gene family shaped pheromone communication in Hymenoptera. *Elife* **8**.
Turner JRG. 1967. On Supergenes. I. The Evolution of Supergenes. *Am Nat* **101**: 195–221.
Waiker P, de Abreu FCP, Luna-Lucena D, Freitas FCP, Simões ZLP, Rueppell O. 2021. Recombination mapping of the Brazilian stingless bee Frieseomelitta varia confirms high recombination rates in social hymenoptera. *BMC Genomics* **22**.
Wallberg A, Glémin S, Webster MT. 2015. Extreme Recombination Frequencies Shape Genome Variation and Evolution in the Honeybee, Apis mellifera ed. N.H. Barton. *PLoS Genet* **11**: e1005189.
Webster MT, Hurst LD. 2012. Direct and indirect consequences of meiotic recombination: implications for genome evolution. *Trends Genet* **28**: 101–109.
Weyna A, Romiguier J. 2021. Relaxation of purifying selection suggests low effective population size in eusocial Hymenoptera and solitary pollinating bees. *Peer Community J* **1**.
White MJD. 1973. *Animal cytology and evolution*. Cambridge University Press, London.
Wilfert L, Gadau J, Schmid-Hempel P. 2006. A core linkage map of the bumblebee Bombus terrestris. *Genome* **49**: 1215–1226.
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.
Wright S. 1931. Evolution in Mendelian Populations. *Genetics* **16**: 97–159.
Wu Y, Bhat PR, Close TJ, Lonardi S. 2008. Efficient and Accurate Construction of Genetic Linkage Maps from the Minimum Spanning Tree of a Graph. *PLoS Genet* **4**: 1000212.

A

SV

RR

TE

Ex

π

B

C

D

recombination rate (cM/Mb)

LDRs

TE islands

LDRs other

LDRs introgr.

****

**

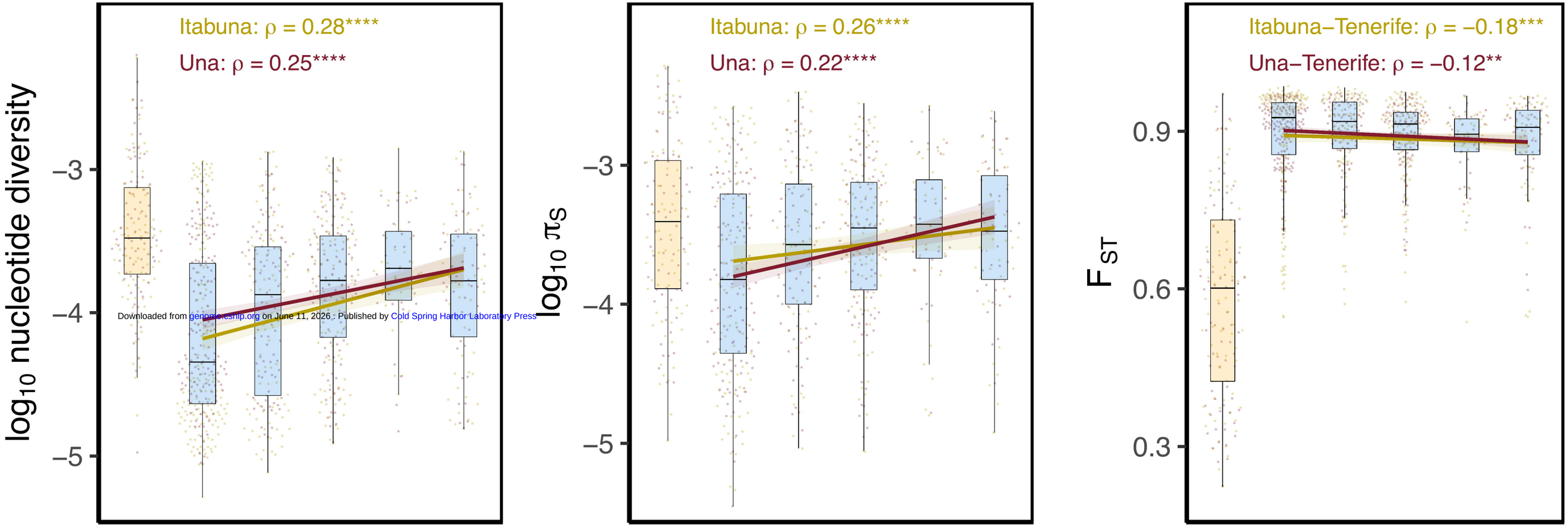
Itabuna

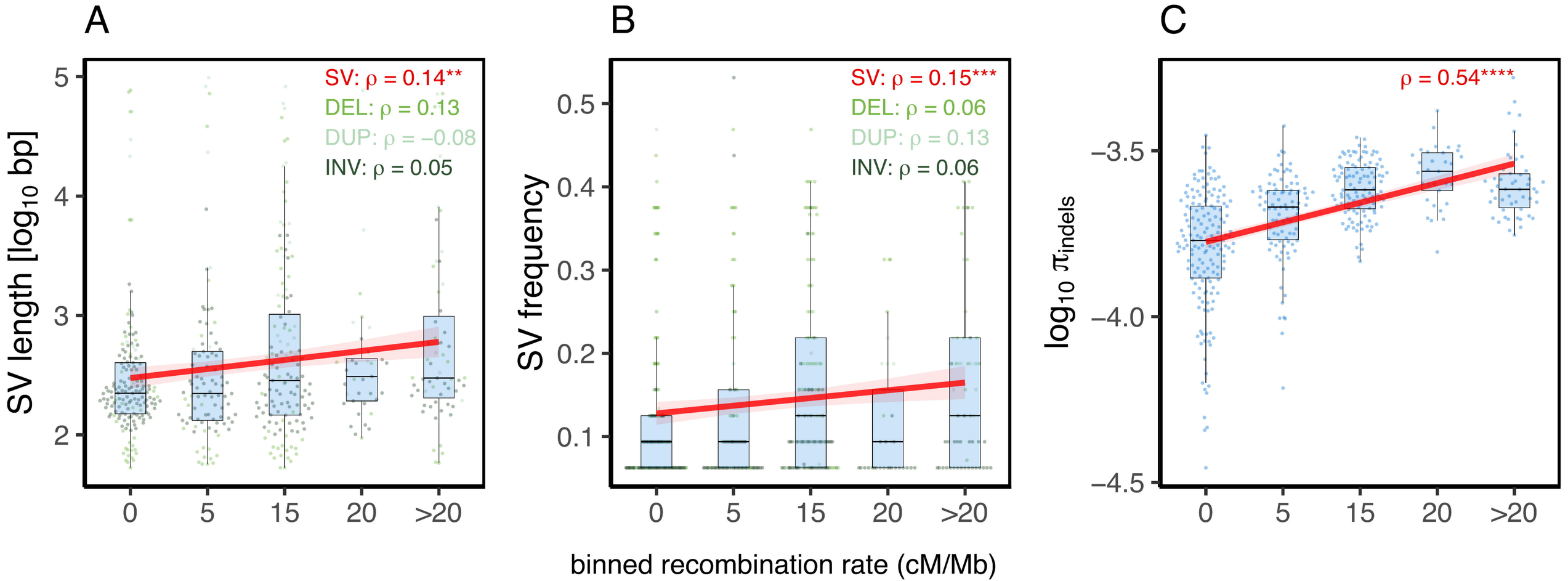
Una

$\log_{10}$ nucleotide diversity

LDRs nonrec.

LDRs rec.




A
B
C
SV length [log10 bp]
SV frequency
log10 πindels
SV: ρ = 0.14**
DEL: ρ = 0.13
DUP: ρ = −0.08
INV: ρ = 0.05
SV: ρ = 0.15***
DEL: ρ = 0.06
DUP: ρ = 0.13
INV: ρ = 0.06
ρ = 0.54****
5
4
3
2
0.5
0.4
0.3
0.2
0.1
−3.5
−4.0
−4.5
0
5
15
20
>20
binned recombination rate (cM/Mb)