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

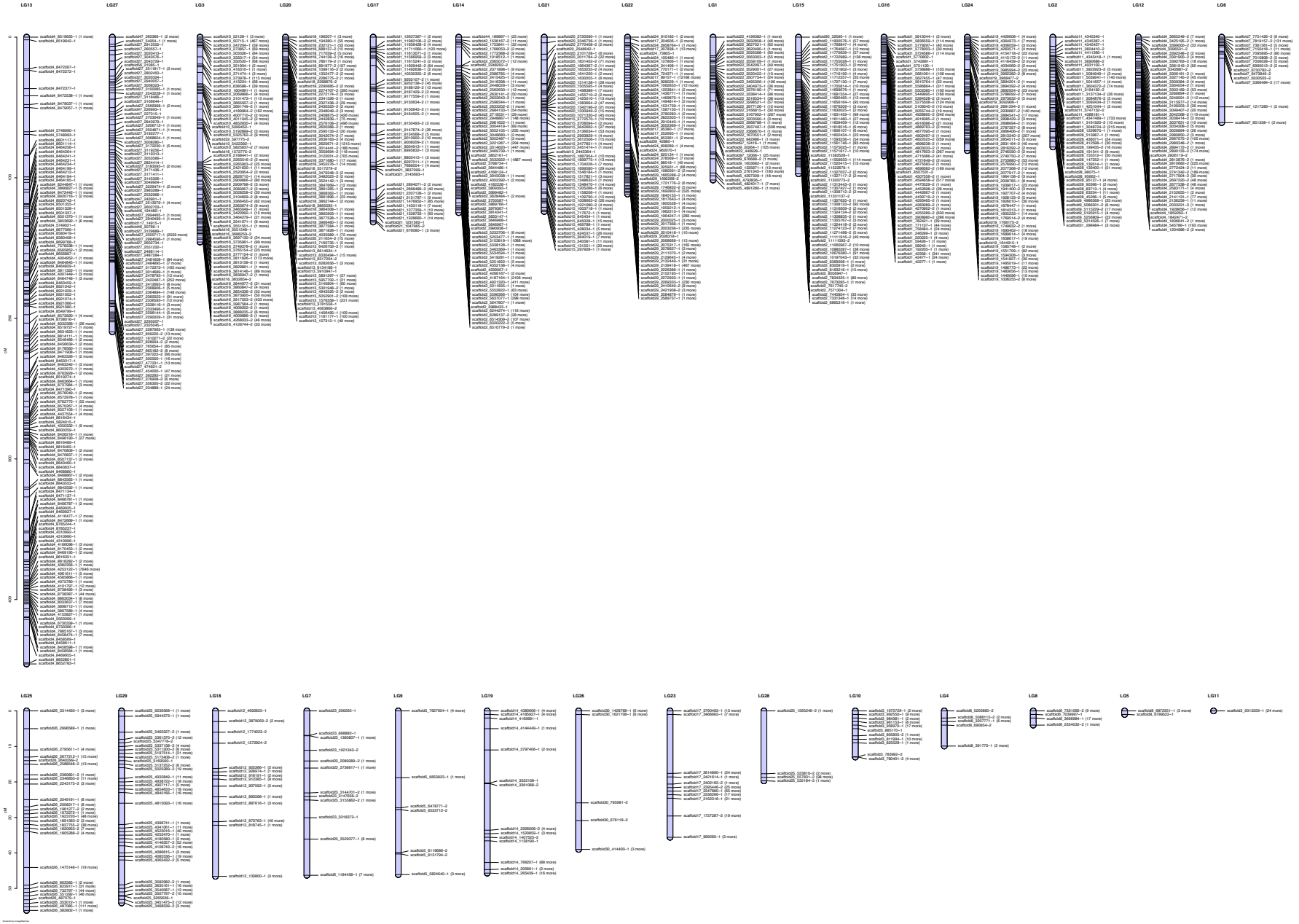
[3]Lehrstuhl für Zoologie/Evolutionsbiologie, University Regensburg, 93053 Regensburg, Germany

*Correspondence: Jan Oettler & Lukas Schrader.

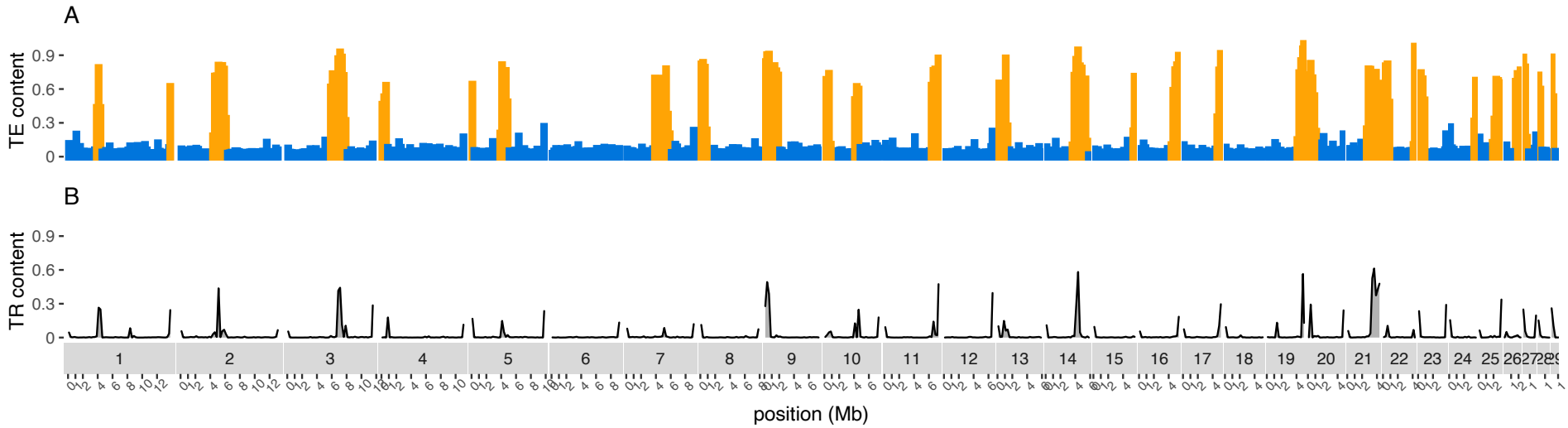
Email: joettler@gmail.com and lukas.schrader@wwu.de

This Supplemental Material includes:

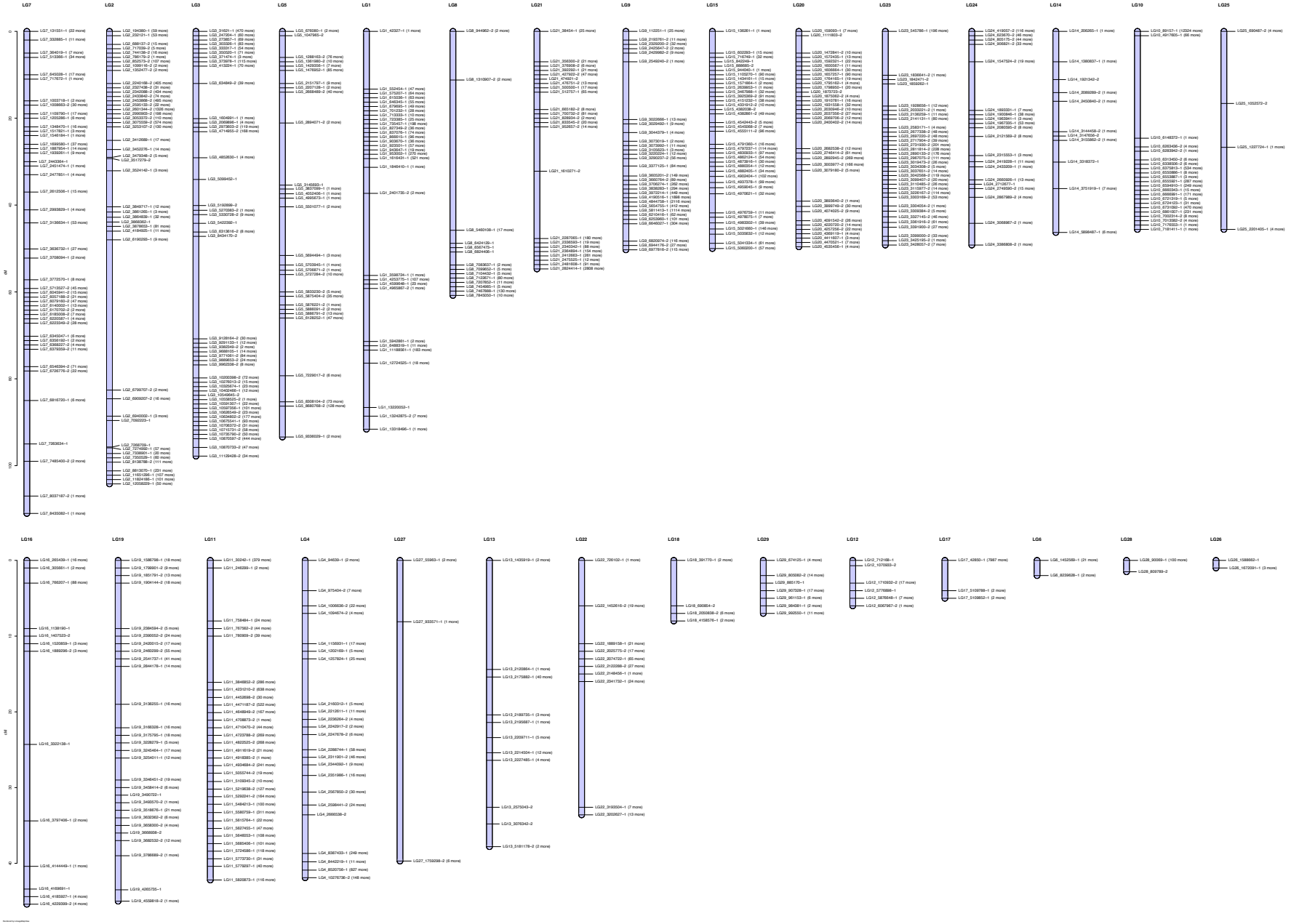
Supplemental Figures 1 to 27
Supplemental Tables 1 to 10
Supplemental References



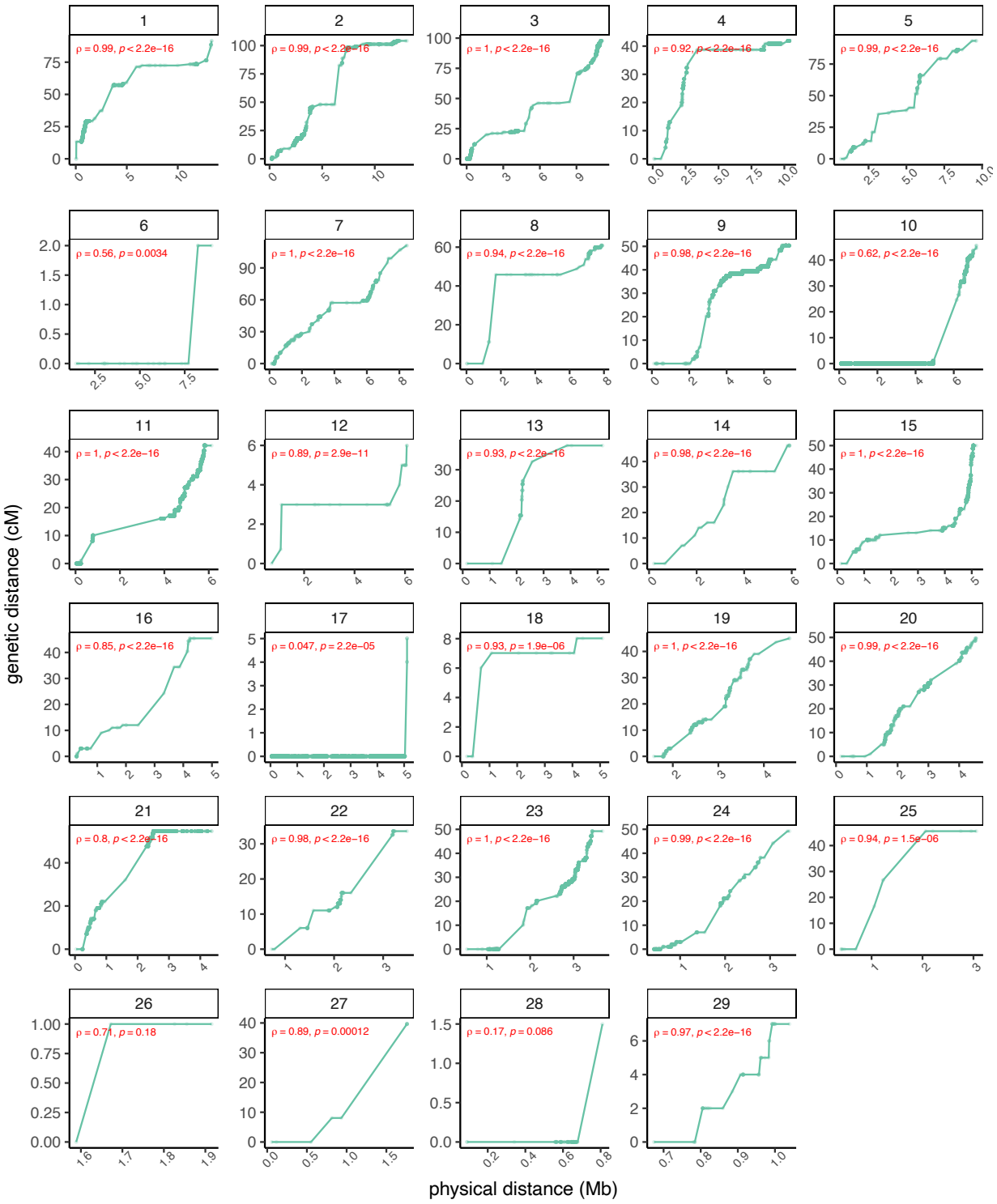
Supplemental Figure 1. Initial genetic linkage map of *C. obscurior* used to refine the Cobs2.1 genome assembly. Left axis shows genetic distances in centimorgans (cM) and marker names are given next to each linkage group (LG). When markers with the same map position are present, only the first marker is displayed followed by the number of duplicated markers between brackets. Plot produced using LinkageMapView (Ouellette et al. 2018).



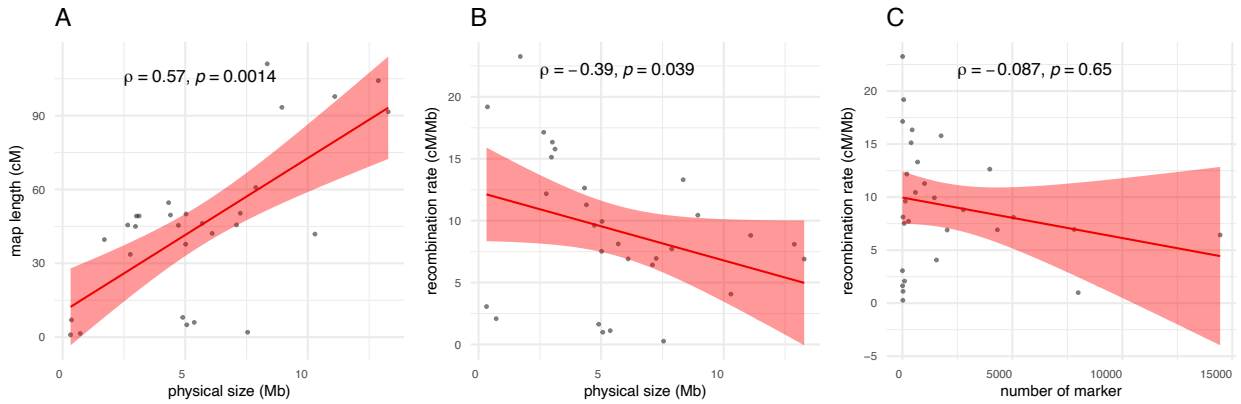
Supplemental Figure 2. (A) Relative content of transposable element (TE)- as well as (B) tandem repeat (TR)-derived sequences along the 29 LGs of the *C. obscurior* genome assembly (Cobs3.1). In (A), blue represents TE-poor low density regions (e.g., "LDRs") while orange depicts TE-rich "TE islands". The relative TE and TR content of these TE islands suggest that they represent potential (peri-) centromeres of *C. obscurior*.



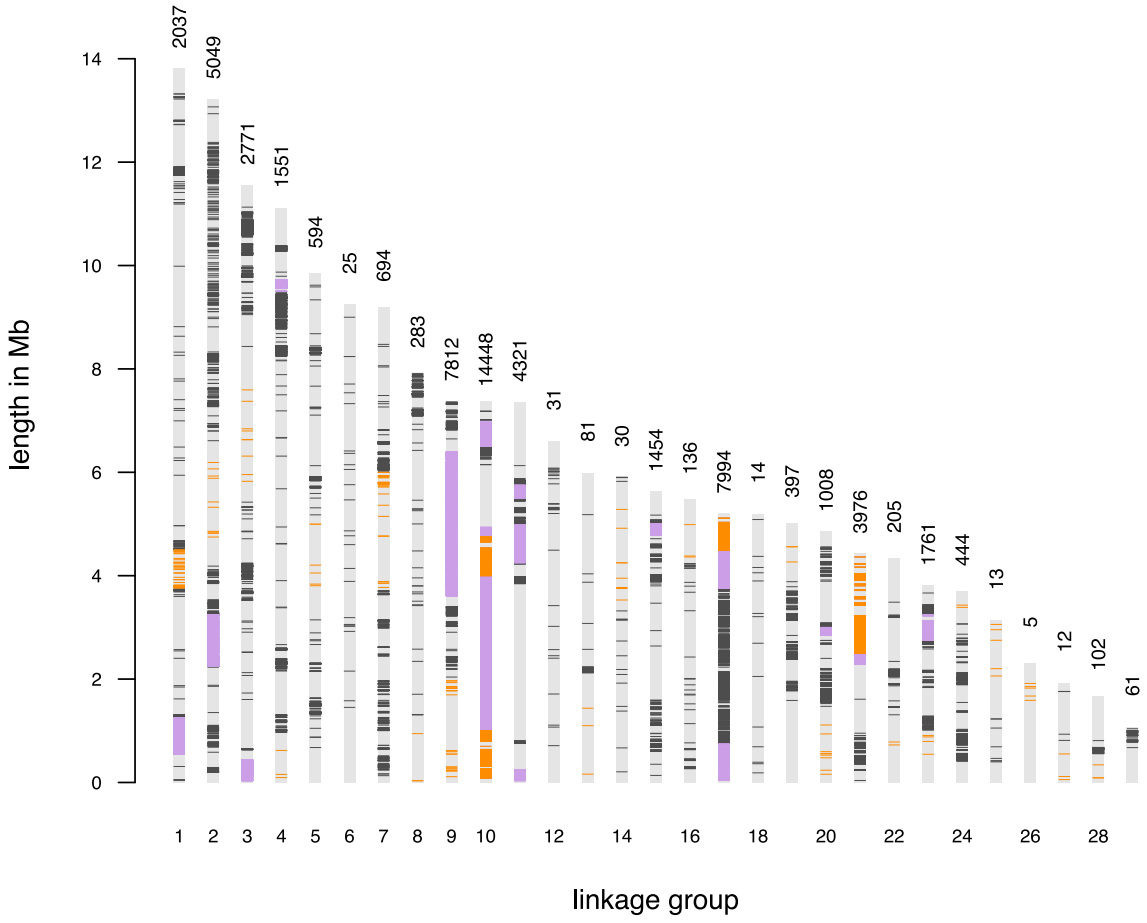
Supplemental Figure 3. Reference genetic linkage map of *C. obscurior*. Axis on the left side shows genetic distances in cM and marker names are given next to each LG. When markers with the same map position are present, only the first marker is displayed followed by the number of duplicated markers between brackets. Plot produced using LinkageMapView.



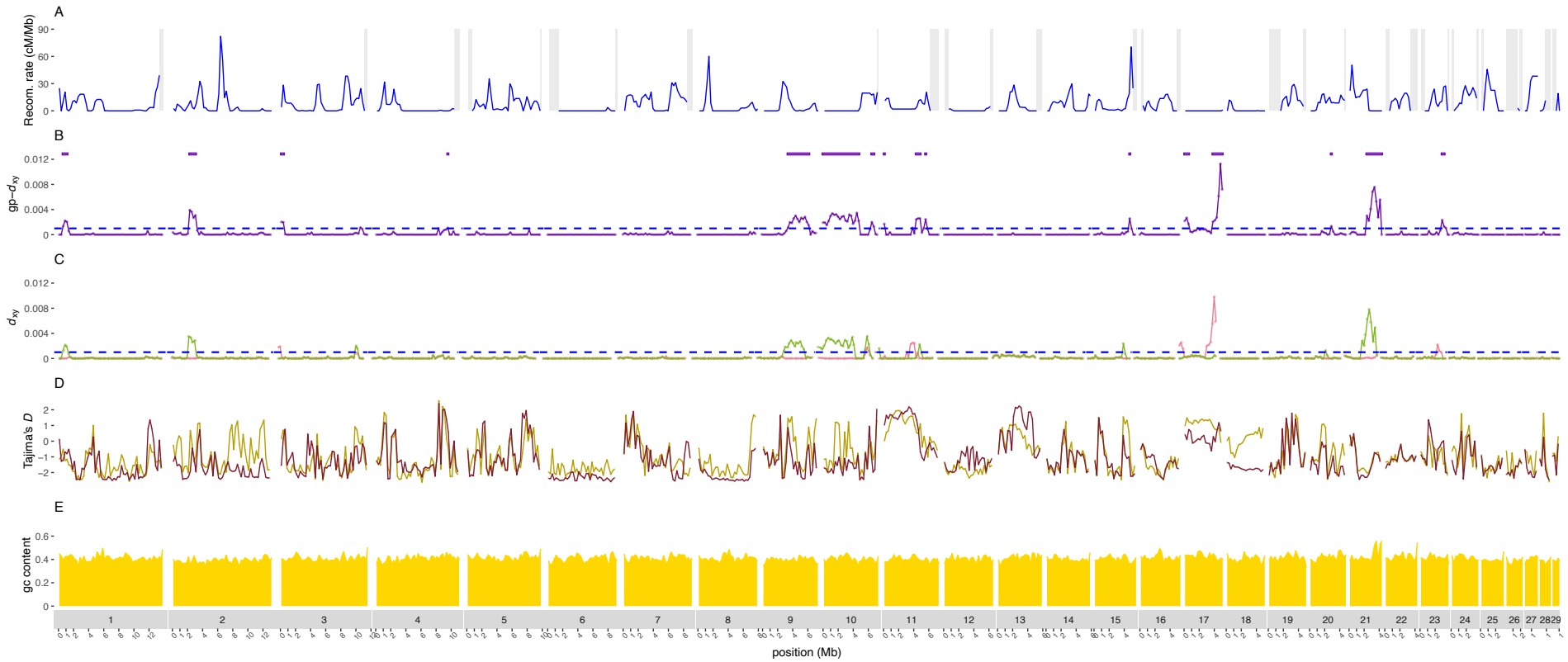
Supplemental Figure 4. Relationship between genetic (cM; y-axis) and physical (Mb; x-axis) distances for LGs of *C. obscurior*. Steep slopes (e.g. LGs 2 &3) indicate regions with high rates of recombination, while flat slopes (e.g. LGs 6, 17, 28) correspond to regions with suppressed recombination. ρ: Spearman's rank correlation coefficient and associated p-values are given for each LG.



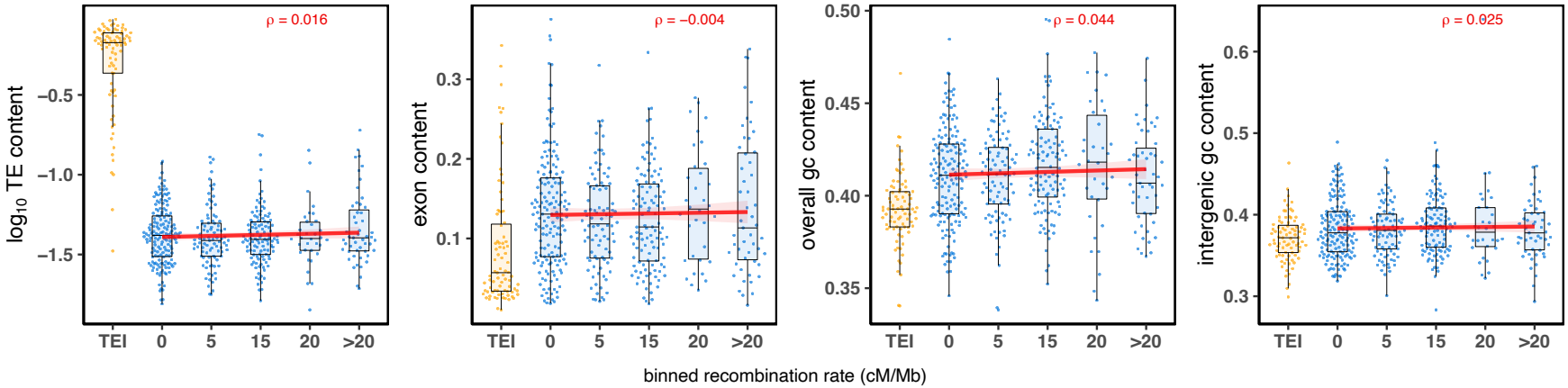
Supplemental Figure 5. Relationship between physical length (Mb; x-axis) and (A) map length (cM) and (B) recombination rate (cM/Mb) for each LG. (C) Association between recombination rate (cM/Mb; y-axis) and the number of markers (x-axis) for each LG. ρ: Spearman's rank correlation coefficient. Red line and shaded areas are linear regressions and 95% confidence intervals.



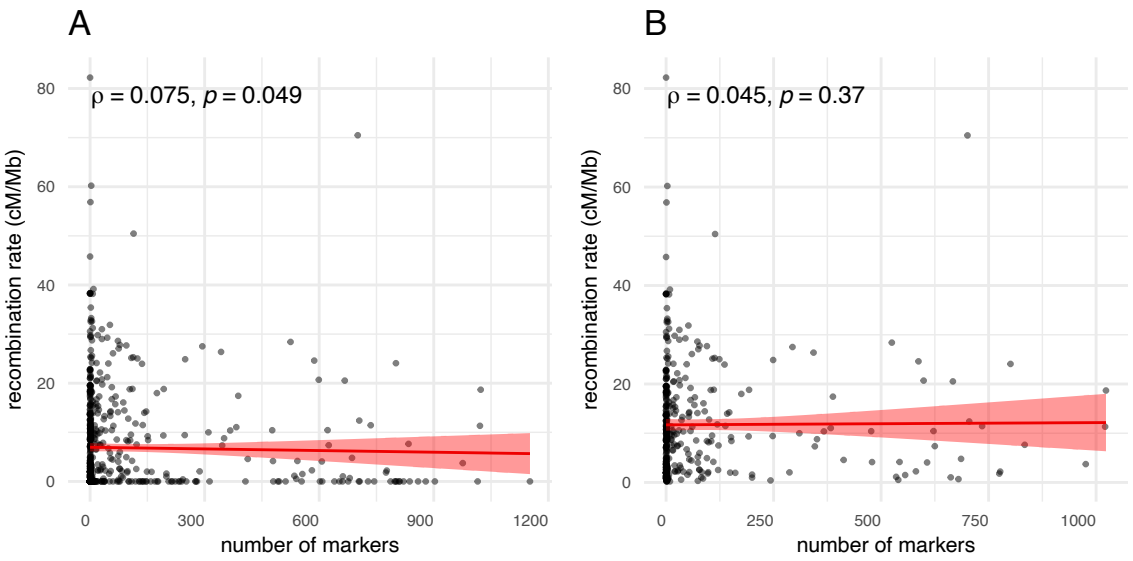
Supplemental Figure 6. Marker distribution across the 29 LGs of *C. obscurior* genetic linkage map. Grey bars represent the different LGs, ordered according to decreasing physical length (Mb; y-axis). Ticks on each LG indicate marker locations, while numbers on top of each LG give the total number of markers for a given LG. Marker colors indicate location in TE islands (orange), introgressions (purple), or LDRs (dark). High marker density regions appear as solid blocks due to tight marker spacing. Absence of ticks (markers) indicates regions where the grandparents used in this study exhibit runs of homozygosity for identical alleles, or where the raw markers in these regions did not meet the filtering criteria outlined in the Methods section.



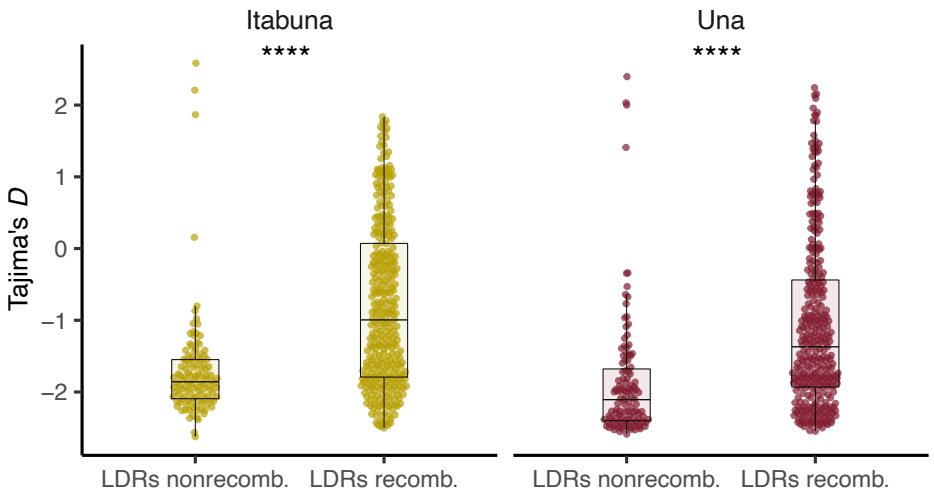
Supplemental Figure 7. (A) Recombination rate (cM/Mb), (B) level of sequence divergence between grandparental haplotypes (gp-d_{XY}), (C) level of sequence divergence between grandparental haplotypes (d_{XY}; male in green; queen in red) and representatives of the Old World lineage from Leiden. Unlike the male's haplotype (green), divergence between the queen's haplotype (red) and the Old World lineage on e.g., LGs 9 and 10 is low, consistent with the queen carrying introgressions from the Old World lineage. (D) Tajima's *D* in populations from Itabuna (gold) and Una (maroon), and (E) gc content across the 29 LGs of the *C. obscurior* genome assembly (Cobs3.1). In (A), grey vertical bars show regions lacking markers while in (B), purple horizontal bars indicate regions with substantially elevated levels of sequence divergence or threefold mean gp-d_{XY} (~0.001).



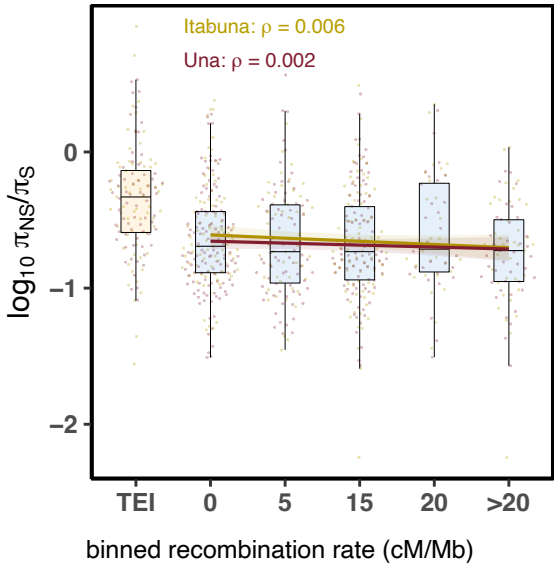
Supplemental Figure 8. Relationship between recombination rates (x-axis) and sequence (TE, exon and gc content) parameters in *C. obscurior*. The x-axis represents binned recombination rates, with the upper limit of each bin indicated and TEI referring to TE islands. ρ: Spearman's rank correlation coefficient, with none of the comparisons being significant. All correlations were performed using the raw data. Red line and shaded areas are linear regressions and 95% confidence intervals.



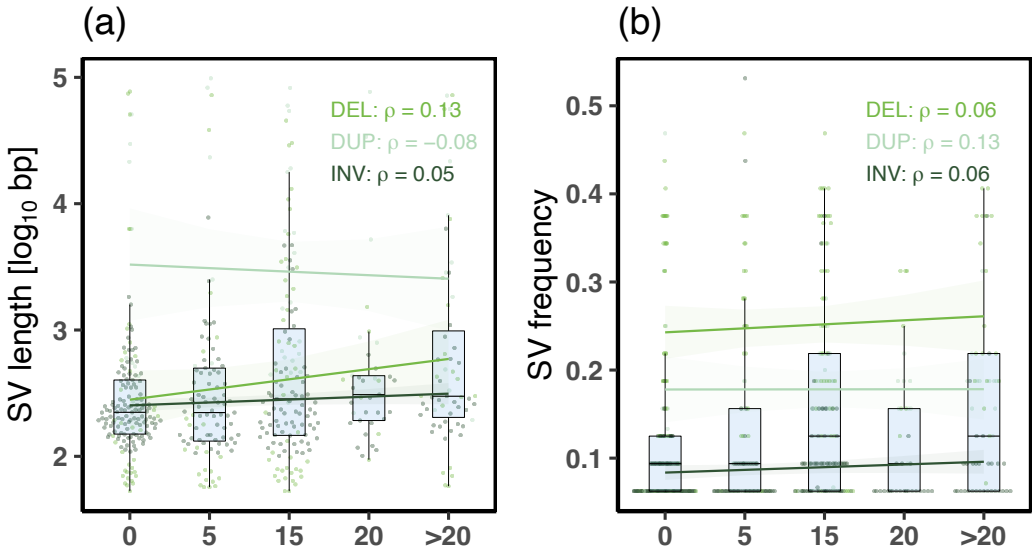
Supplemental Figure 9. Relationship between marker count (x-axis) and recombination rate (cM/Mb) for each 250 kb window when considering all (A) or only recombining windows (B). ρ: Spearman's rank correlation coefficient. Red line and shaded areas are linear regressions and 95% confidence intervals.



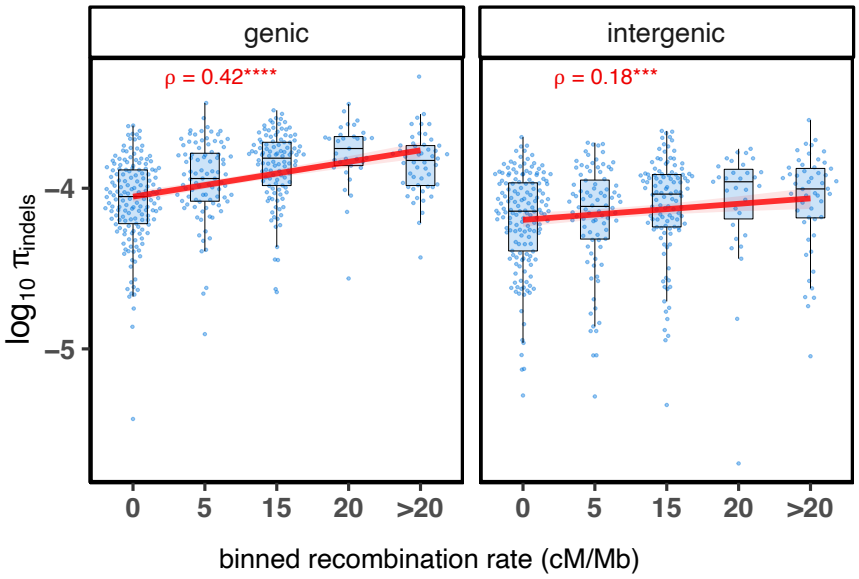
Supplemental Figure 10. Tajima's *D* estimates in nonrecombining (LDRs nonrecomb.) compared to the recombining regions LDRs (LDRs recomb.) excluding introgressions in two populations of *C. obscurior* from Itabuna and Una (****<0.0001).



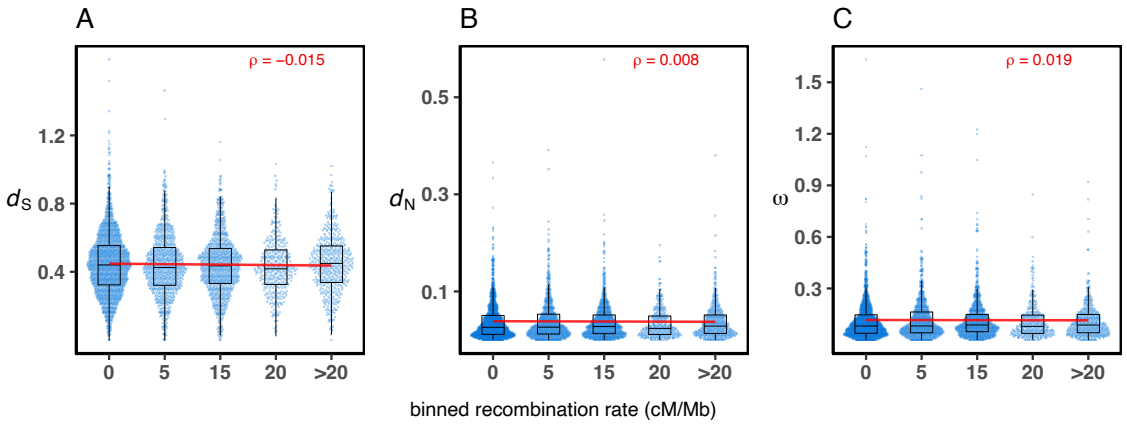
Supplemental Figure 11. Relationship between recombination rates and the ratio of nonsynonymous (π_{NS}) to synonymous (π_S) nucleotide diversity, in two populations from Itabuna (gold) and Una (maroon). The x-axis represents binned recombination rates, with the upper limit of each bin indicated. TEI: TE islands (orange boxplot). ρ: Spearman's rank correlation coefficient showing no significant association. All correlations were performed using the raw data. Lines and shaded areas are linear regressions and 95% confidence intervals.



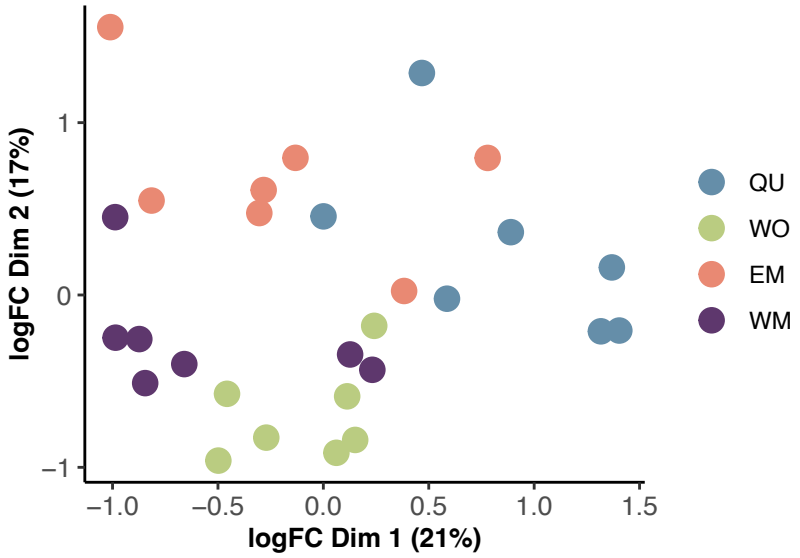
Supplemental Figure 12. Association between recombination rates and length (A) and frequency (B) of structural variants in 16 individual workers of *C. obscurior*. The x-axis represents binned recombination rates, with the upper limit of each bin indicated. ρ: Spearman's rank correlation coefficient is indicated separately for deletions (DEL), duplications (DUP) and inversions (INV), with none of the comparisons being significant. All correlations were performed using the raw data. Lines and shaded areas are linear regressions and 95% confidence intervals.



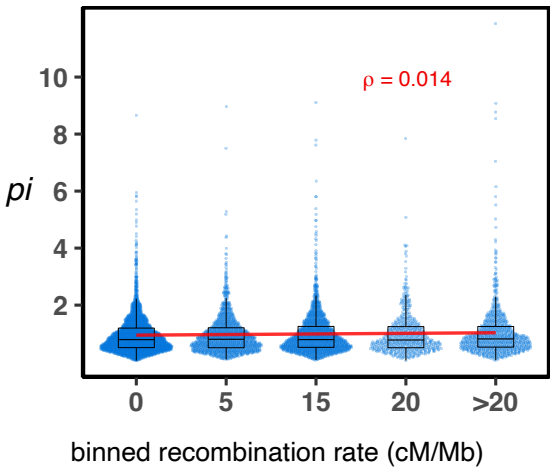
Supplemental Figure 13. Correlation between recombination rates and InDels polymorphism in 16 individual workers when considering genic or intergenic InDels separately. The x-axis represents binned recombination rates, with the upper limit of each bin indicated. ρ: Spearman's rank correlation coefficient. All correlations were performed using the raw data. Red line and shaded areas are linear regressions and 95% confidence intervals.



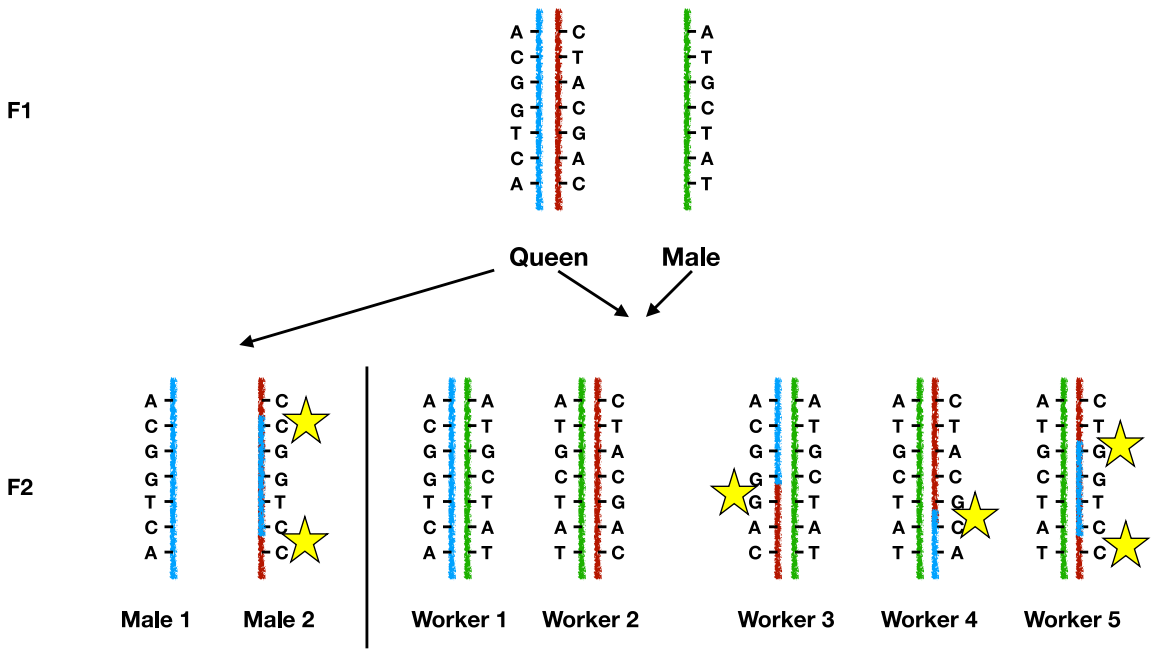
Supplemental Figure 14. Association between recombination rates and the rates of (A) synonymous (d_S) and (B) nonsynonymous (d_N) as well as (C) the d_N/d_S ratio in *C. obscurior*. The x-axis represents binned recombination rates, with the upper limit of each bin indicated. ρ: Spearman's rank correlation coefficient, with none of the comparisons being significant. All correlations were performed using the raw data. Red line and shaded areas are linear regressions and 95% confidence intervals.



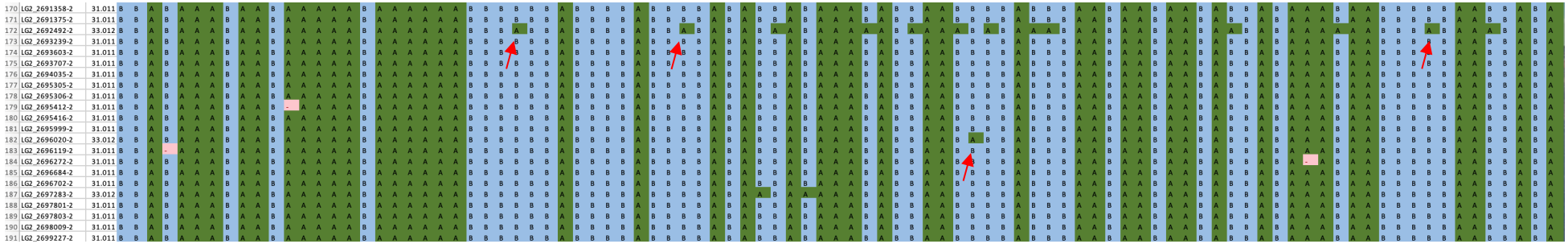
Supplemental Figure 15. Multidimensional scaling (MDS) plot of gene expression profiles of worker (WO), queen (QU), wingless male (WM), and winged male (EM) third instar larvae (n = 7 individuals each). Distances represent leading $\log_2$-fold changes between samples.



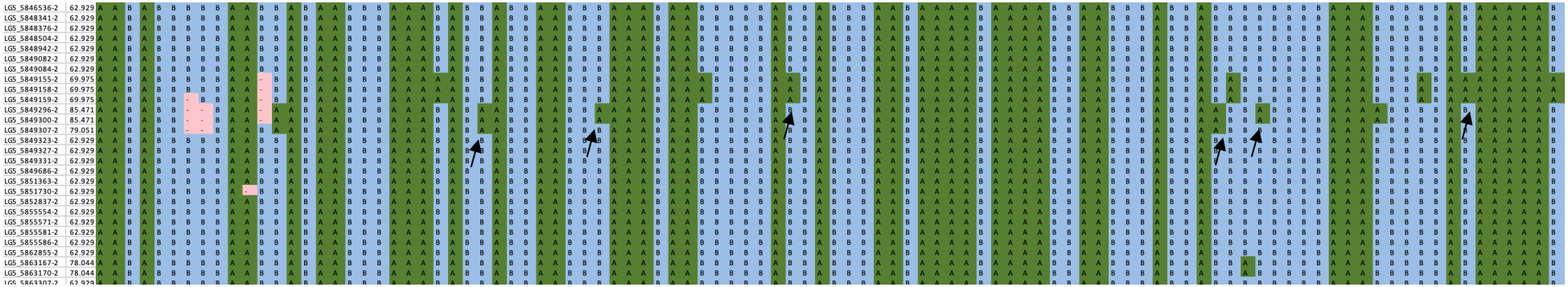
Supplemental Figure 16. Association between recombination rates and *pi*, a measure of gene expression plasticity among developing 3rd instar larvae of *C. obscurior* queens, workers, wingless males and winged males. The x-axis represents binned recombination rates, with the upper limit of each bin indicated. ρ: Spearman's rank correlation coefficient showing no significant association. All correlations were performed using the raw data. Red line and shaded areas are linear regressions and 95% confidence intervals.



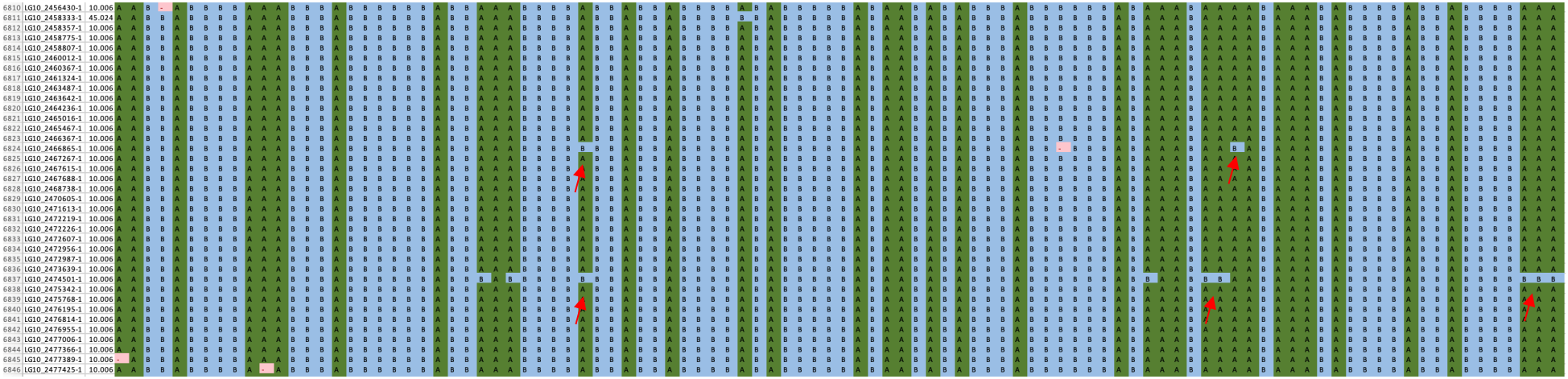
Supplemental Figure 17. Schematic representation of the phasing pipeline used in the study. Following the screening of raw single nucleotide polymorphisms (SNPs) for high-quality markers, a total of ~57k markers, which were heterozygous in the F1 queen where recombination occurs and genotyped in the F1 male and over 80% of the F2 offspring, were phased to retrieve the maternal haplotype in the 88 F2 female workers included in the mapping population. In haplodiploid ants, males are haploid (hence the F1 male possesses only one chromosome represented in green), while females are diploid (thus both the F1 queen and the F2 workers have two chromosomes). The F2 males are the direct outcome of meiosis in the F1 queen, and their genomes can exhibit either recombination events (male 2; highlighted with a yellow star) or lack thereof (male 1). Each F2 worker on the other hand carries the F1 father's contribution and one of the two chromosomes from the F1 queen. The maternal chromosome inherited by the F2 workers may feature one or more recombination events (workers 3, 4, and 5; highlighted with a yellow star) or none (workers 1 and 2). Using the F1 male genotype information, a custom script (Supplementary Code) was employed to extract the maternal contribution from each F2 worker. Subsequently, 88 sequences from the F2 workers (i.e., maternal haplotype) and 12 males were used for linkage mapping analysis using MSTmap (v.1.0) (Wu et al. 2008), with parameters specified in Supplemental Table S9, to order all ~57k markers.



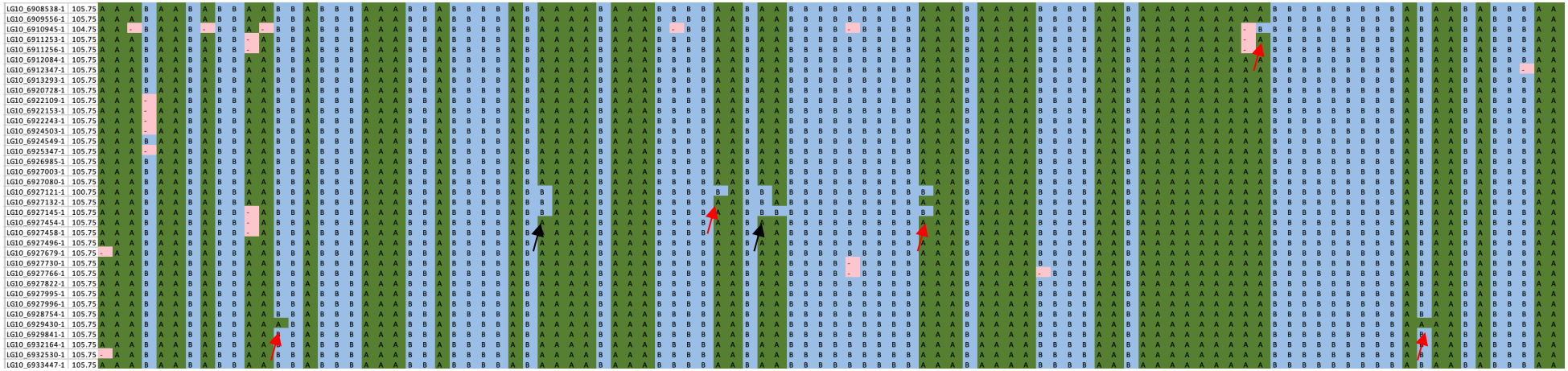
Supplemental Figure 18. Screenshot displaying genotypes of different F2 individuals (columns) at various marker positions (rows) on LG 2, coded as allele A (green) or B (blue). Missing genotypes are represented as "–" (colored in red). The screenshot shows examples of singletons (red arrows), indicating a double crossover event where a single marker locus is flanked upstream and downstream by different alleles. These events were excluded from our analysis because of their high improbability due to chiasma interference, likely representing genotyping errors.



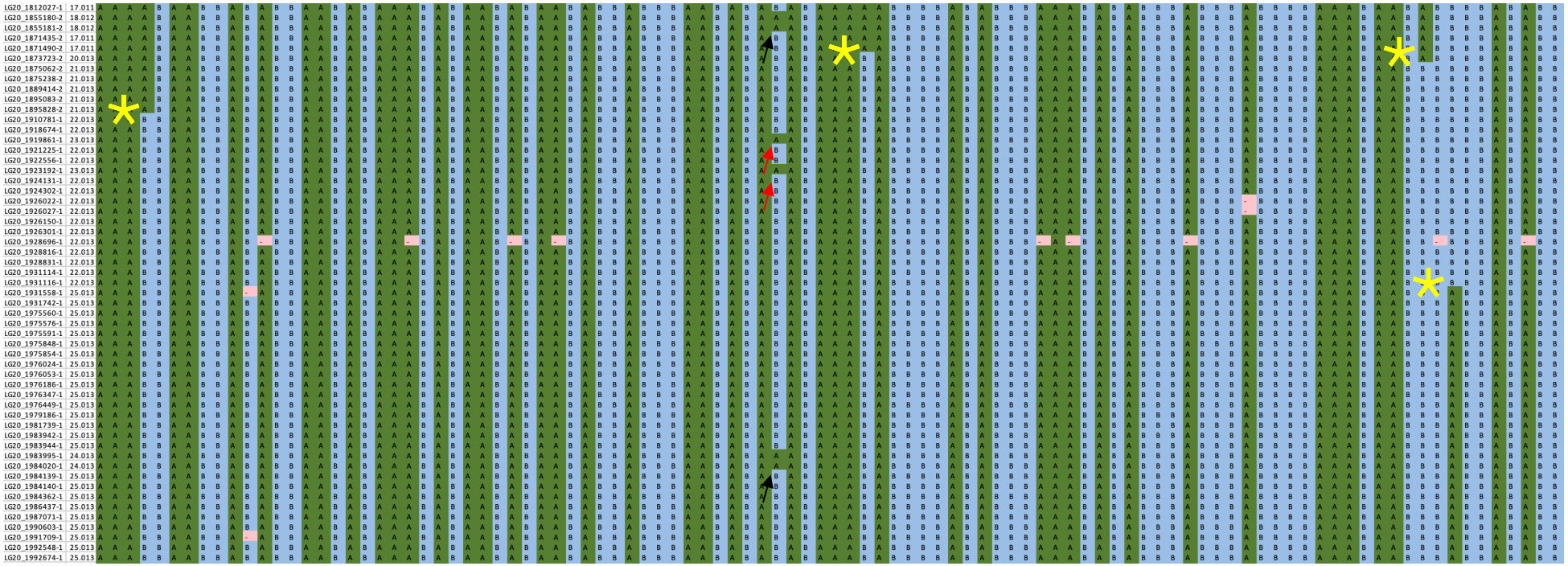
Supplemental Figure 19. Screenshot displaying genotypes of different F2 individuals (columns) at various marker positions (rows) on LG 5, coded as allele A (green) or B (blue). Missing genotypes are represented as "–" (colored in red). The screenshot shows examples of blocks (black arrows), indicating a double crossover event where few adjacent markers are flanked upstream and downstream by different alleles. These events were excluded from our analysis because of their high improbability due to chiasma interference, likely representing genotyping errors.



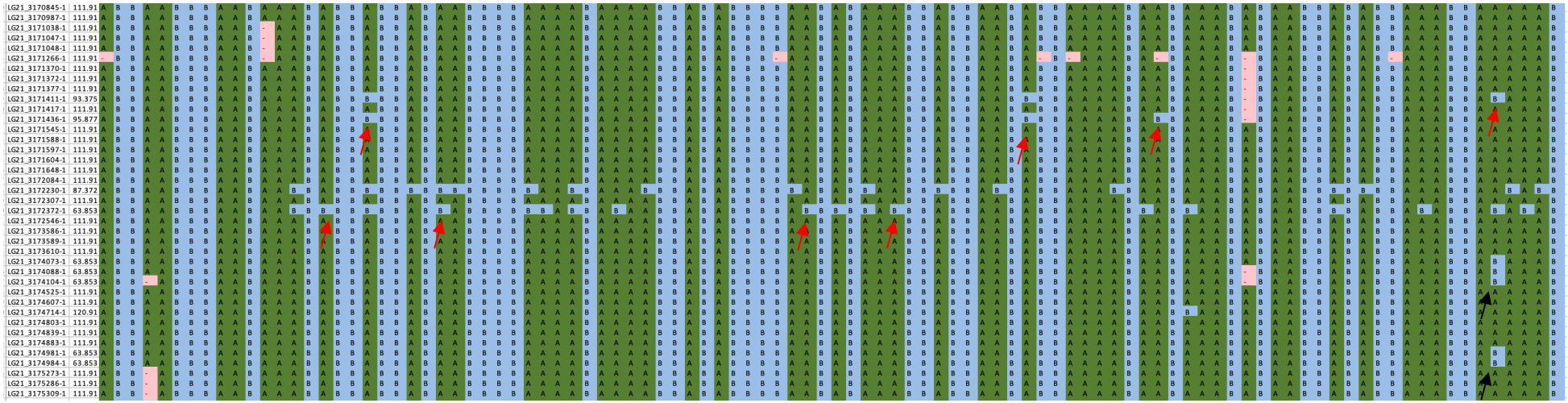
Supplemental Figure 20. Screenshot displaying genotypes of different F2 individuals (columns) at various marker positions (rows) on LG 10, coded as allele A (green) or B (blue). Missing genotypes are represented as "–" (colored in red). The screenshot shows examples of singletons (red arrows), indicating a double crossover event where a single marker locus is flanked upstream and downstream by different alleles. These events were excluded from our analysis because of their high improbability due to chiasma interference, likely representing genotyping errors.



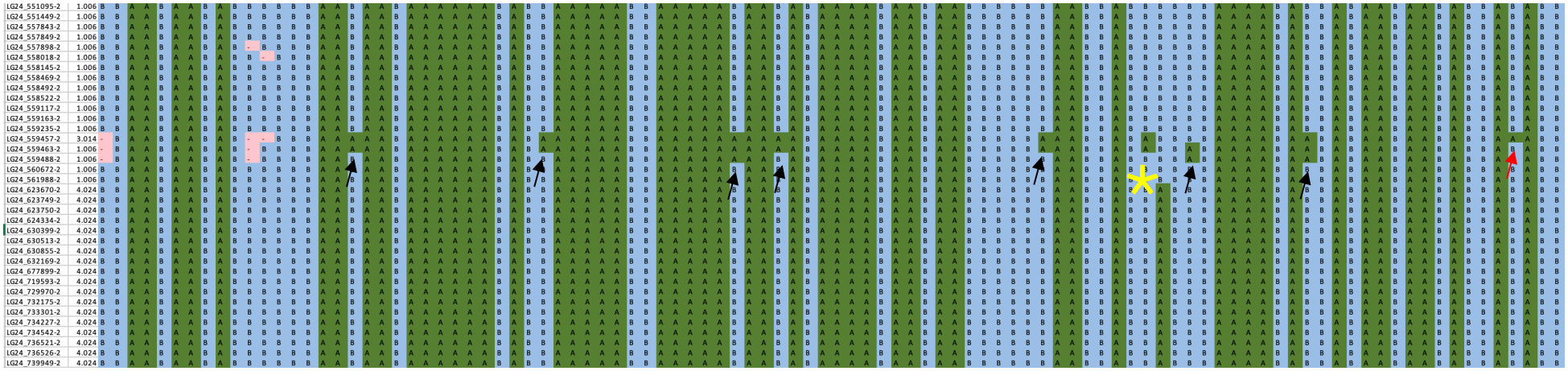
Supplemental Figure 21. Screenshot displaying genotypes of different F2 individuals (columns) at various marker positions (rows) on LG 10, coded as allele A (green) or B (blue). Missing genotypes are represented as "–" (colored in red). The screenshot shows examples of singletons (red arrows), indicating a double crossover event where a single marker locus is flanked upstream and downstream by different alleles. Additionally, the screenshot illustrates blocks of adjacent markers (black arrows) indicating a double crossover event. These events were excluded from our analysis because of their high improbability due to chiasma interference, likely representing genotyping errors.



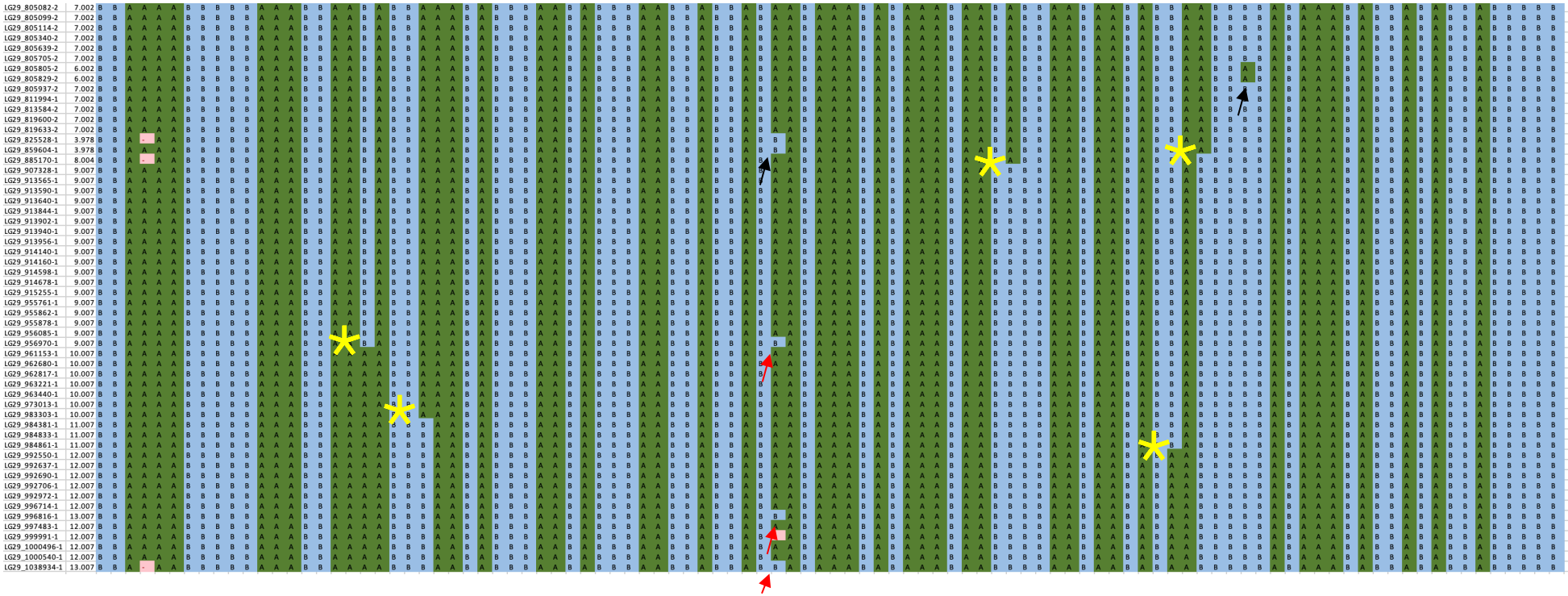
Supplemental Figure 22. Screenshot displaying genotypes of different F2 individuals (columns) at various marker positions (rows) on LG 20, coded as allele A (green) or B (blue). Missing genotypes are represented as "–" (colored in red). The screenshot shows two singletons (red arrows), indicating a double crossover event where a single marker locus is flanked upstream and downstream by different alleles. Additionally, the screenshot illustrates two blocks of adjacent markers (black arrows) indicating a double crossover event. These events were excluded from our analysis because of their high improbability due to chiasma interference, likely representing genotyping errors. Yellow stars mark crossover events, demonstrating switches between A and B alleles over long distances.



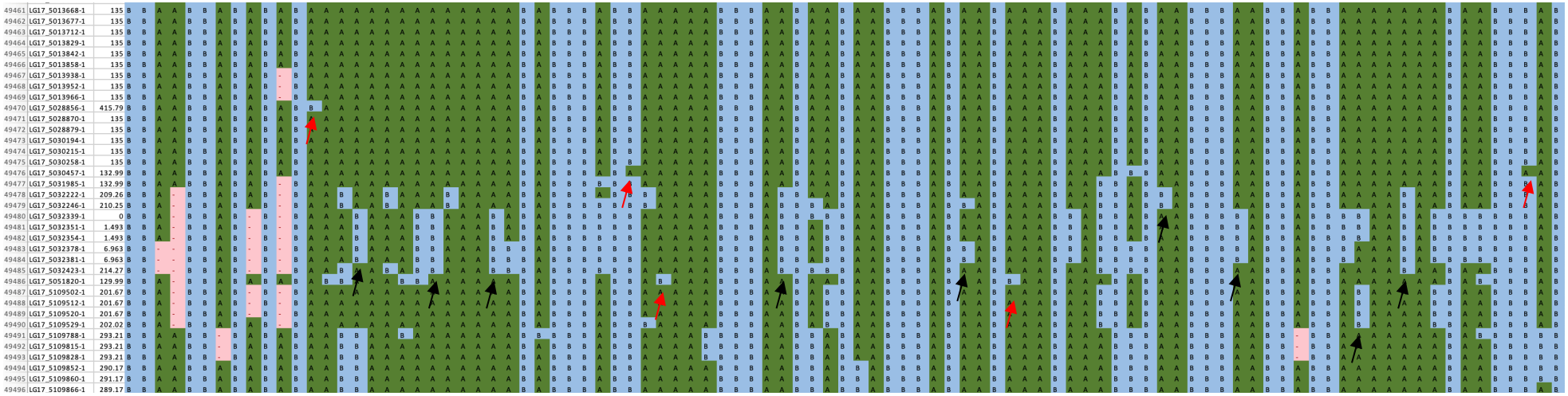
Supplemental Figure 23. Screenshot displaying genotypes of different F2 individuals (columns) at various marker positions (rows) on LG 21, coded as allele A (green) or B (blue). Missing genotypes are represented as "–" (colored in red). The screenshot shows multiple singletons (red arrows), indicating a double crossover event where a single marker locus is flanked upstream and downstream by different alleles. Additionally, the screenshot illustrates blocks of adjacent markers (black arrows) indicating a double crossover event. These events were excluded from our analysis because of their high improbability due to chiasma interference, likely representing genotyping errors.



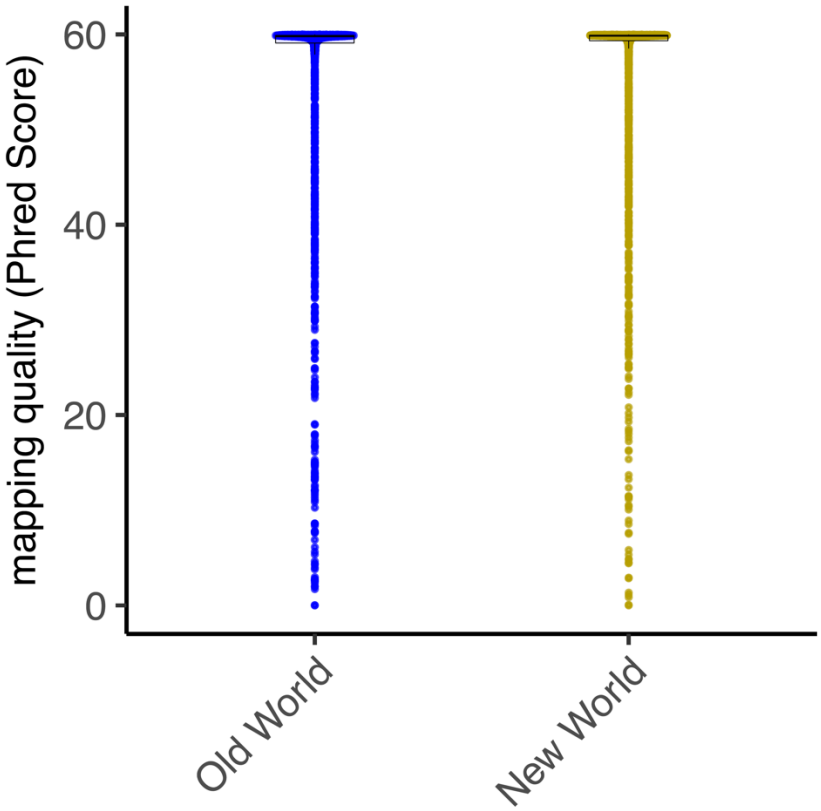
Supplemental Figure 24. Screenshot displaying genotypes of different F2 individuals (columns) at various marker positions (rows) on LG 24, coded as allele A (green) or B (blue). Missing genotypes are represented as "–" (colored in red). The screenshot shows a singleton (red arrows), indicating a double crossover event where a single marker locus is flanked upstream and downstream by different alleles. Additionally, the screenshot illustrates blocks of adjacent markers (black arrows) indicating a double crossover event. These events were excluded from our analysis due to their high improbability due to chiasma interference, likely representing genotyping errors. The yellow star marks a crossover event, demonstrating a switch between A and B alleles over a long distance.



Supplemental Figure 25. Screenshot displaying genotypes of different F2 individuals (columns) at various marker positions (rows) on LG 29, coded as allele A (green) or B (blue). Missing genotypes are represented as "–" (colored in red). The screenshot shows multiple singletons (red arrows), indicating a double crossover event where a single marker locus is flanked upstream and downstream by different alleles. Additionally, the screenshot illustrates blocks of two adjacent markers (black arrows) indicating a double crossover event. These events were excluded from our analysis because of their high improbability due to chiasma interference, likely representing genotyping errors. Yellow stars mark crossover events, demonstrating switches between A and B alleles over long distances.



Supplemental Figure 26. Screenshot displaying genotypes of different F2 individuals (columns) at various marker positions (rows) on LG 17, coded as allele A (green) or B (blue). Missing genotypes are represented as "–" (colored in red). The screenshot shows multiple singletons (red arrows), indicating a double crossover event where a single marker locus is flanked upstream and downstream by different alleles. Additionally, the screenshot illustrates blocks of adjacent markers (black arrows) indicating a double crossover event. These events were excluded from our analysis because of their high improbability due to chiasma interference, likely representing genotyping errors.



	n	mean	sd	median
New World	1861	56.34	9.51	59.86
Old World	1861	55.48	11.15	59.80

Supplemental Figure 27. Boxplot depicts mapping qualities computed in 100 kb windows for samples from both the Old World (Tenerife) and the New World (Itabuna) lineages of *C. obscurior*. Irrespective of their source, reads from both lineages can be aligned to the reference genome with high confidence.

Supplemental Table 1. QUAST based statistics calculated for the Cobs3.1 in comparison to the Cobs2.1 genome assembly.

Assembly	Cobs3.1	Cobs2.1
# contigs (>= 0 bp)	113	127
# contigs (>= 1000 bp)	113	127
# contigs (>= 5000 bp)	111	125
# contigs (>= 10000 bp)	107	121
# contigs (>= 25000 bp)	91	105
# contigs (>= 50000 bp)	76	91
Total length (>= 0 bp)	193.05	193.05
Total length (>= 1000 bp)	193.05	193.05
Total length (>= 5000 bp)	193.05	193.05
Total length (>= 10000 bp)	193.03	193.03
Total length (>= 25000 bp)	192.76	192.76
Total length (>= 50000 bp)	192.22	192.24
# contigs	113	127
# placed contigs (i.e. linkage groups)	29	0
# unplaced contigs	84	127
Largest contig	13.82	13.15
Total length	193.05	193.05
GC (%)	41.02	41.02
N50	7.38	6.29
N75	5.20	4.49
L50	10	11
L75	18	21
# N's per 100 kbp	94.71	94.76

Supplemental Table 2. Meiotic recombination rate estimates in *Cardiocondyla obscurior* and other animal species. For a more comprehensive catalogue of recombination rate estimates across taxa, please refer to Stapley et al. (Stapley et al. 2017).

Species	Recombination rates (cM/Mb)	References
Social hymenopterans:		
Cardiocondyla obscurior	8.1	This study
Acromyrmex echinatior	6.20	(Sirviö et al. 2006)
Pogonomyrmex californicus	14.1	Errbii M and Gadau J, unpublished
P. rugosus	14	(Sirviö et al. 2011b)
Apis mellifera	19-37	(Beye et al. 2006; Liu et al. 2015; Hunt and Page 1995)
A. florea	20.8	(Rueppell et al. 2016)
A. dorsata	25.1	(Rueppell et al. 2016)
A. cerana	17.4	(Shi et al. 2013)
Bombus terrestris	4-8.7	(Wilfert et al. 2006; Niehuis et al. 2008; Gadau et al. 2001; Stolle et al. 2011; Liu et al. 2017)
Vespula vulgaris	9.7	(Sirviö et al. 2011a)
Parasitoid hymenopterans:		
Nasonia vitripennis	1.4-1.6	(Niehuis et al. 2010)
Nasonia vitripennis x giraulti	2.5	(Gadau et al. 1999)
Lysiphlebus fabarum	5.8	(Ulrich et al. 2021)
Other insects:		
Drosophila melanogaster	1.59	(Comeron et al. 2012)
Heliconius melpomene	5.5	(Jiggins et al. 2005)
Vanessa cardui	3.81–4.05	(Shipilina et al. 2022)
Bombyx mori	2.97-4.00	(Yasukochi 1998; Yamamoto et al. 2008)
Vertebrates:		
Homo sapiens	1.1	(Kong et al. 2002; Dumont and Payseur 2008)
Mus musculus	0.5	(Dumont and Payseur 2008; Dietrich et al. 1996)
Other animals:		
Meloidogyne hapla	20	(Opperman et al. 2008)
Ciona intestinalis	13.37-27.00	(Kano et al. 2006)

Supplemental Table 3. Inter-marker distances across the different linkage groups in kilobases. the q90 column shows the 90th percentile, i.e. 90% of all inter-marker distances on a given linkage group are equal to or lower than that value.

Linkage group	Median	Mean	q90
LG1	0.29	6.52	1.73
LG2	0.16	2.55	1.27
LG3	0.19	4.00	1.42
LG4	0.36	6.63	3.15
LG5	0.20	15.06	5.99
LG6	210.27	301.89	736.46
LG7	0.63	12.02	12.19
LG8	0.42	27.84	23.43
LG9	0.19	0.93	1.09
LG10	0.18	0.49	0.86
LG11	0.20	1.41	0.97
LG12	37.84	173.19	602.16
LG13	0.56	61.98	14.02
LG14	128.26	189.88	470.63
LG15	0.18	3.46	1.52
LG16	0.39	34.73	110.82
LG17	0.14	0.63	1.12
LG18	256.97	350.02	691.50
LG19	0.48	7.49	8.26
LG20	0.24	4.36	2.33
LG21	0.12	1.09	0.77
LG22	0.30	13.47	4.96
LG23	0.23	1.77	1.41
LG24	0.38	6.78	7.42
LG25	141.00	204.50	511.66
LG26	59.87	65.29	125.62
LG27	0.39	142.04	418.62
LG28	0.32	7.06	3.71
LG29	0.23	5.98	21.01

Supplemental Table 4. Inter-marker distances within TE islands across the different LGs in kilobases. These distances are on average below 450 kb indicating that we have sufficient markers at relatively short distances from each other, which should be adequate to capture crossovers if they occur within TE-rich regions.

Linkage group	Median	Mean
LG1	2.4	12.8
LG2	113.8	200.5
LG3	220.2	241.3
LG4	61.1	175
LG5	183.9	308.4
LG7	1.7	46.7
LG8	1.7	304.1
LG9	11.9	74.2
LG10	0.2	0.7
LG13	337.2	425
LG14	216.8	245.4
LG16	132.2	253
LG17	0	0.3
LG19	294.1	225.6
LG20	51	105.9
LG21	0.1	0.7
LG22	28.4	28.4
LG23	57.7	90.8
LG24	180.5	180.5
LG25	205	365.4
LG26	59.9	65.3
LG27	60	165.3
LG28	0	83.7

Supplemental Table 5. List of inversions identified using two different tools, specifying individual genotypes as follows: 0/0 for homozygous reference, 0/1 for a heterozygous inversion, and 1/1 for a homozygous alternate allele. Inversions that appear as heterozygous in the F1 queen can potentially affect recombination between the haplotype carrying the inversion and the unaffected one.

Linkage group	Position	Length (bp)	Tool	F1 Male	F1 Queen	P Male	P Queen
LG3	6594552	185558	delly	0/0	0/1	0/0	0/1
LG4	256407	131	delly	0/0	0/1	0/0	0/1
LG4	256435	103	delly	0/0	0/1	0/0	0/1
LG5	3980872	793	delly	0/0	0/1	0/0	0/1
LG10	6759325	2442	delly	1/1	0/1	0/0	0/1
LG10	6759328	1937	delly; smoove	1/1	0/1	0/0	0/1
LG10	6765022	1609	delly	1/1	0/1	0/0	0/1
LG10	6766856	2347	delly	1/1	0/1	0/0	0/1
LG10	7294090	363	delly	0/0	0/1	0/0	0/1
LG11	649131	2111	delly	0/0	0/1	0/0	0/1
LG14	5008854	239	delly	0/0	0/1	0/0	0/1
LG17	4596294	18365	delly	0/0	0/1	1/1	0/0
LG19	4850491	6541	delly	0/0	0/1	0/0	0/1
LG20	200199	7736	delly	0/0	0/1	0/0	0/1
LG21	3148566	328392	delly	1/1	0/1	0/0	1/1
LG22	944085	122	delly	0/0	0/1	0/0	0/1
LG23	299549	1845	delly	0/0	0/1	0/0	0/1
LG26	198852	117	delly; smoove	0/0	0/1	0/0	0/1
LG4	256363	173	smoove	0/0	0/1	0/0	0/1
LG7	4357657	218	smoove	0/0	0/1	0/0	0/1
LG11	1147539	76	smoove	0/0	0/1	0/0	0/1
LG20	2948131	56	smoove	1/1	0/1	0/0	1/1
LG30	641526	668	smoove	0/0	0/1	0/0	0/1

Supplemental Table 6. Summary of the results from a generalized linear regression model with three explanatory variables and recombination as the response variation.

	Estimate	Std. Error	*t*	*P*-value
(Intercept)	2.3131	0.0519	44.571	< 2e-16 ***
TEislands	-1.7016	0.2428	-7.007	5.90e-12 ***
introgression	-0.4114	0.1714	-2.4	0.0167 *
epistasis	-3.756	0.6545	-5.739	1.44e-08 ***

Supplemental Table 7. Results of Gene Ontology enrichment analyses using topGO of genes contained in high recombination regions in *C. obscurior* genome (top 10%; gwRR > 22.3 cM/Mb).

GO.ID	Term	Annotated	Significant	Expected	FDR	Ontology
GO:0009922	fatty acid elongase activity	15	10	1.22	4.60E-06	MF
GO:0080019	fatty-acyl-CoA reductase (alcohol-forming) activity	19	9	1.54	6.17E-04	MF
GO:0004114	3',5'-cyclic-nucleotide phosphodiesterase activity	8	4	0.65	1.53E-01	MF
GO:0005085	guanyl-nucleotide exchange factor activity	55	10	4.46	3.14E-01	MF
GO:0019901	protein kinase binding	7	3	0.57	3.14E-01	MF
GO:0005247	voltage-gated chloride channel activity	3	2	0.24	3.14E-01	MF
GO:0008641	ubiquitin-like modifier activating enzyme activity	8	3	0.65	3.14E-01	MF
GO:0004970	ionotropic glutamate receptor activity	21	5	1.7	3.14E-01	MF
GO:0008061	chitin binding	54	9	4.37	3.14E-01	MF
GO:0019789	SUMO transferase activity	4	2	0.32	3.14E-01	MF
GO:0030527	structural constituent of chromatin	17	4	1.38	3.14E-01	MF
GO:0007165	signal transduction	444	41	32.06	1.24E-01	BP
GO:0006888	endoplasmic reticulum to Golgi vesicle-mediated transport	18	5	1.3	3.98E-01	BP
GO:0030198	extracellular matrix organization	7	3	0.51	3.98E-01	BP
GO:0006275	regulation of DNA replication	5	2	0.36	3.98E-01	BP
GO:0000786	nucleosome	23	5	1.59	1.00E+00	CC

Supplemental Table 8. Quality control metrics for the sequencing data used in this study.

Sample Name	Sample ID	Mean coverage	Aligned reads	Sex
PQueen	150645	54.3	90.80%	female
F1Queen	150647	66.8	89.80%	female
W6	150668	29.1	99.50%	female
W7	150670	31.3	99.50%	female
W8	150672	35.9	99.60%	female
W9	150674	34.3	99.70%	female
W13	150676	30.6	99.50%	female
W14	150678	30.1	99.60%	female
W15	150680	33.8	99.70%	female
W16	150682	27.6	99.50%	female
W17	150684	28.7	99.60%	female
W18	150686	43.6	99.50%	female
W19	150688	48.1	99.40%	female
W23	150690	49.2	99.40%	female
W25	150692	43.8	99.00%	female
W26	150694	38.4	99.70%	female
W27	150696	46.9	99.10%	female
W29	150698	47.3	99.50%	female
W30	150700	46.3	98.90%	female
W32	150702	51.9	99.10%	female
W33	150704	41.9	99.10%	female
W34	150706	55.2	99.30%	female
W40	150708	49.00	99.80%	female
W41	150710	52.7	99.80%	female
W42	150712	45.5	99.30%	female
W43	150714	49.3	99.50%	female
W44	150716	42.6	99.10%	female
W45	150718	50.2	99.60%	female
W46	150720	40.3	99.50%	female
W49	150722	45.8	99.10%	female
W50	150724	50.8	99.50%	female
W51	150726	42.3	99.40%	female
W52	150728	42.3	99.50%	female
W53	150730	50.9	99.10%	female
W56	150732	57.5	99.60%	female
W58	150734	51.3	99.60%	female

W60	150736	48.5	98.90%	female
W61	150738	56.2	98.60%	female
W62	150925	43.7	99.00%	female
W63	150928	44.1	99.10%	female
W68	150930	53.7	99.30%	female
W69	150932	49.1	98.90%	female
W70	150934	43.7	98.90%	female
W71	150936	48.7	99.50%	female
W73	150938	47.8	99.30%	female
W74	150940	41.5	99.40%	female
W75	150942	47.8	99.30%	female
W76	150944	45.2	98.80%	female
W80	150946	38.7	99.40%	female
W81	150948	41.2	98.90%	female
W82	150950	41.4	98.50%	female
W83	150952	37.3	99.60%	female
W84	150954	29.9	99.50%	female
W86	150956	35.3	99.60%	female
W88	150958	42.7	99.40%	female
W89	150960	42.3	99.50%	female
W90	150962	39.8	99.40%	female
W91	150964	42.2	99.60%	female
W77	150966	44.8	99.30%	female
W78	150968	41.7	98.80%	female
W92	150970	39.1	98.60%	female
W97	150972	41	98.90%	female
W103	150974	46.9	99.00%	female
W101	150976	37.4	99.30%	female
W106	150978	42.4	99.00%	female
W107	150980	45.6	99.60%	female
W109	150982	39.8	99.60%	female
W110	150984	41.7	99.40%	female
W111	150986	39.9	99.70%	female
W200	150988	41.9	99.60%	female
W201	150990	35.3	99.70%	female
W202	150992	43.1	99.20%	female
W204	150994	41.4	99.40%	female
W205	150996	37.6	99.50%	female

W206	150998	32.7	99.40%	female
W207	151000	38.3	99.60%	female
W211	151002	29.1	99.10%	female
W212	151004	39.2	99.70%	female
W221	151006	40.8	99.50%	female
W300	151008	36.8	99.30%	female
W302	151010	39.1	99.40%	female
W303	151012	45.4	99.60%	female
W304	151014	36.3	99.50%	female
W55	151473	36.8	99.50%	female
W57	151475	48.4	99.10%	female
W59	151477	39.5	99.70%	female
W64	151479	50.9	99.00%	female
W72	151481	40.9	97.50%	female
W85	151483	46.2	99.40%	female
W301	151485	38.5	99.50%	female
EM2	150650	44.00	98.80%	male
EM4	150652	41.00	98.90%	male
EM35	150654	22.9	99.30%	male
EM36	150656	18.9	99.40%	male
EM37	150658	19.00	99.10%	male
EM38	150660	14.2	99.00%	male
EM39	150662	36.7	99.40%	male
EM78	150664	28.5	99.30%	male
EM112	150666	23.8	99.50%	male
PMale	151464	28.1	99.60%	male
EM1	151467	19.1	99.40%	male
EM3	151469	25.3	99.20%	male
EM5	151471	28.8	99.40%	male
F1Male	151487	29.1	99.80%	male

Supplemental Table 9. Parameters used for running MSTmap.

Parameter	Value
population_type	DH
population_name	Cobs
distance_function	kosambi
cut_off_p_value	0.0000001
no_map_dist	16
no_map_size	2
missing_threshold	0.2
estimation_before_clustering	no
detect_bad_data	yes
objective_function	COUNT

Supplemental Table 10. Genomic coordinates and orientation of scaffolds (Cobs2.1) in LGs of the new assembly (Cobs3.1). If present, gaps were used to split chimeric scaffolds (see Methods).

Scaffold	Start	End	Orientation	Gap	Cobs3.1
scaffold22	0	4241709	-	-	LG1
scaffold97	0	47470	+	-	LG1
scaffold9	0	4138037	+	-	LG1
scaffold5	0	5392496	+	5392497 to 5426199	LG1
scaffold18	0	4606763	+	-	LG2
scaffold35	0	540112	+	-	LG2
scaffold34	0	580016	+	-	LG2
scaffold33	0	214576	+	214577 to 214616	LG2
scaffold13	0	7284943	-	-	LG2
scaffold10	0	6032196	+	-	LG3
scaffold32	0	724972	+	-	LG3
scaffold67	0	104227	+	-	LG3
scaffold16	0	4698819	+	85721 to 87912	LG3
scaffold9	4138037	6290588	-	-	LG4
scaffold28	0	3097748	+	-	LG4
scaffold11	0	5868402	+	-	LG4
scaffold21	0	4384715	+	-	LG5
scaffold1	7686053	13148674	+	7361188 to 7383474	LG5
scaffold3	1075729	10323382	+	-	LG6
scaffold15	0	5070358	+	-	LG7
scaffold20	355366	4487289	+	355366 to 355366	LG7
scaffold70	0	93624	+	-	LG8
scaffold7	0	7825401	+	-	LG8
scaffold40	0	359143	+	-	LG9
scaffold44	0	298451	+	-	LG9
scaffold2	0	6729055	+	-	LG9
scaffold17	0	273158	+	-	LG10
scaffold39	0	486599	+	-	LG10
scaffold24	0	3831917	+	-	LG10
scaffold29	0	2791255	+	-	LG10
scaffold1	0	7361187	+	7361188 to 7383474	LG11
scaffold6	1263099	7865674	+	-	LG12
scaffold19	0	239252	+	-	LG13
scaffold12	4692625	5755492	+	-	LG13
scaffold12	0	4692625	+	-	LG13

scaffold23	0	4217338	+	-	LG14
scaffold49	0	422459	+	-	LG14
scaffold6	0	1263099	+	-	LG14
scaffold2	6729055	12360777	+	-	LG15
scaffold14	0	5477151	+	-	LG16
scaffold4	3811214	9024684	+	3811115 to 3811214	LG17
scaffold8	0	5200860	+	-	LG18
scaffold25	1678838	6706168	+	-	LG19
scaffold42	0	344800	+	-	LG20
scaffold19	239252	4751815	+	-	LG20
scaffold27	0	3383218	+	-	LG21
scaffold46	0	391713	+	-	LG21
scaffold47	0	367889	+	-	LG21
scaffold56	0	188944	+	-	LG21
scaffold65	0	105511	+	-	LG21
scaffold17	273158	4616616	+	-	LG22
scaffold4	0	3811114	+	3811115 to 3811214	LG23
scaffold80	0	72378	+	-	LG24
scaffold26	0	3629807	+	-	LG24
scaffold5	5426199	8562622	+	-	LG25
scaffold8	5200860	7505125	+	-	LG26
scaffold58	0	137590	+	-	LG27
scaffold30	0	1791517	+	-	LG27
scaffold25	1565245	1599841	+	-	LG28
scaffold25	0	1565245	+	-	LG28
scaffold25	1599841	1678838	+	-	LG28
scaffold3	0	1075729	+	-	LG29

Supplemental References

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.

Comeron JM, Ratnappan R, Bailin S. 2012. The Many Landscapes of Recombination in Drosophila melanogaster. *PLoS Genet* **8**: e1002905.

Dietrich WF, Miller J, Steen R, Merchant MA, Damron-Boles D, Husain Z, Dredge R, Daly MJ, Ingalls KA, O'Connor TJ, et al. 1996. A comprehensive genetic map of the mouse genome. *Nature* **380**: 149–152.

Dumont BL, Payseur BA. 2008. Evolution of the genomic rate of recombination in mammals. *Evolution (N Y)* **62**: 276–294.

Gadau J, Gerloff CU, Krüger N, Chan H, Schmid-Hempel P, Wille A, Page RE. 2001. A linkage analysis of sex determination in Bombus terrestris (L.) (Hymenoptera: Apidae). *Heredity (Edinb)* **87**: 234–242.

Gadau J, Page RE, Werren JH. 1999. Mapping of hybrid incompatibility loci in Nasonia. *Genetics* **153**: 1731–1741.

Hunt GJ, Page RE. 1995. Linkage map of the honey bee, Apis mellifera, based on RAPD markers. *Genetics* **139**: 1371–1382.

Jiggins CD, Mavarez J, Beltrán M, McMillan WO, Johnston JS, Bermingham E. 2005. A genetic linkage map of the mimetic butterfly Heliconius melpomene. *Genetics* **171**: 557–570.

Kano S, Satoh N, Sordino P. 2006. Primary genetic linkage maps of the ascidian, Ciona intestinalis. *Zoolog Sci* **23**: 31–39.

Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, et al. 2002. A high-resolution recombination map of the human genome. *Nat Genet* **31**: 241–247.

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.

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 recombinational landscape of the honey bee. *Genome Biol* **16**: 15.

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.

Niehuis O, Judson AK, Gadau J. 2008. Cytonuclear Genic Incompatibilities Cause Increased Mortality in Male F2 Hybrids of Nasonia giraulti and N. vitripennis. *Genetics* **178**: 413–426.

Opperman CH, Bird DM, Williamson VM, Rokhsar DS, Burke M, Cohn J, Cromer J, Diener S, Gajan J, Graham S, et al. 2008. Sequence and genetic map of Meloidogyne hapla: A compact nematode genome for plant parasitism. *Proc Natl Acad Sci U S A* **105**: 14802–14807.

Ouellette LA, Reid RW, Blanchard SG, Brouwer CR. 2018. LinkageMapView-rendering high-resolution linkage and QTL maps. *Bioinformatics* **34**: 306–307.

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.

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**: 8–13.

Shipilina D, Näsvall K, Höök L, Vila R, Talavera G, Backström N. 2022. Linkage mapping and genome annotation give novel insights into gene family expansions and regional recombination rate variation in the painted lady (Vanessa cardui) butterfly. *Genomics* **114**: 110481.

Sirviö A, Gadau J, Rueppell O, Lamatsch D, Boomsma JJ, Pamilo P, Page RE. 2006. High recombination frequency creates genotypic diversity in colonies of the leaf-cutting ant Acromyrmex echinatior. *J*

Evol Biol **19**: 1475–1485.
Sirviö A, Johnston JS, Wenseleers T, Pamilo P. 2011a. A high recombination rate in eusocial Hymenoptera: Evidence from the common wasp Vespula vulgaris. *BMC Genet* **12**: 1–7.
Sirviö A, Pamilo P, Johnson RA, Page RE, Gadau J. 2011b. Origin and evolution of the dependent lineages in the genetic caste determination system of Pogonomyrmex ants. *Evolution (N Y)* **65**: 869–84.
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 B Biol Sci* **372**: 20160455.
Stolle E, Wilfert L, Schmid-Hempel R, Schmid-Hempel P, Kube M, Reinhardt R, Moritz RFA. 2011. A second generation genetic map of the bumblebee Bombus terrestris (Linnaeus, 1758) reveals slow genome and chromosome evolution in the Apidae. *BMC Genomics* **12**: 1–17.
Ulrich GF, Zemp N, Vorburger C, Boulain H. 2021. Quantitative trait locus analysis of parasitoid counteradaptation to symbiont-conferred resistance. *Heredity (Edinb)* **127**: 219–232.
Wilfert L, Gadau J, Schmid-Hempel P. 2006. A core linkage map of the bumblebee Bombus terrestris. *Genome* **49**: 1215–1226.
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.
Yamamoto K, Nohata J, Kadono-Okuda K, Narukawa J, Sasanuma M, Sasanuma SI, Minami H, Shimomura M, Suetsugu Y, Banno Y, et al. 2008. A BAC-based integrated linkage map of the silkworm Bombyx mori. *Genome Biol* **9**: 1–14.
Yasukochi Y. 1998. A dense genetic map of the silkworm, Bombyx mori, covering all chromosomes based on 1018 molecular markers. *Genetics* **150**: 1513–1525.