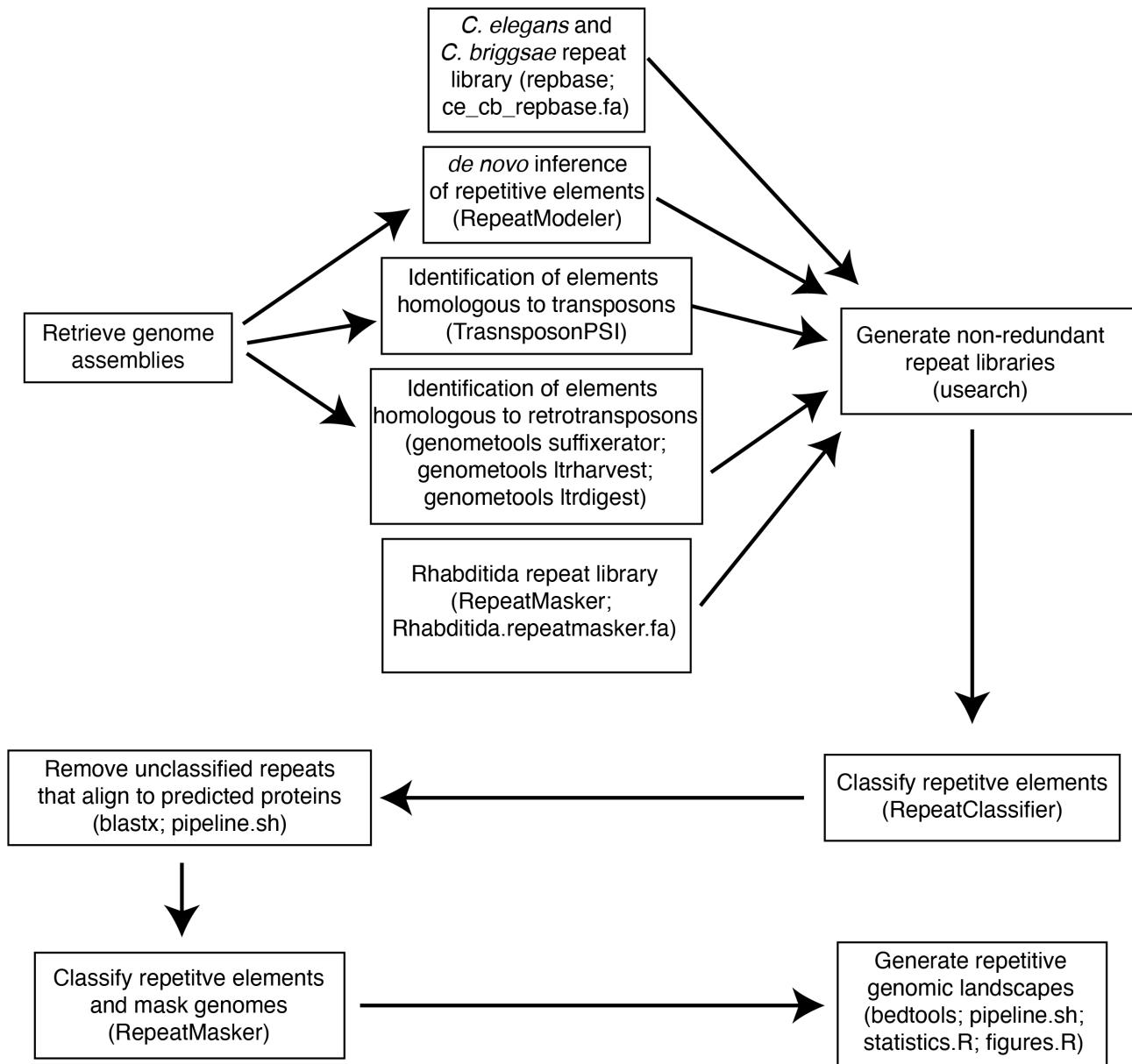
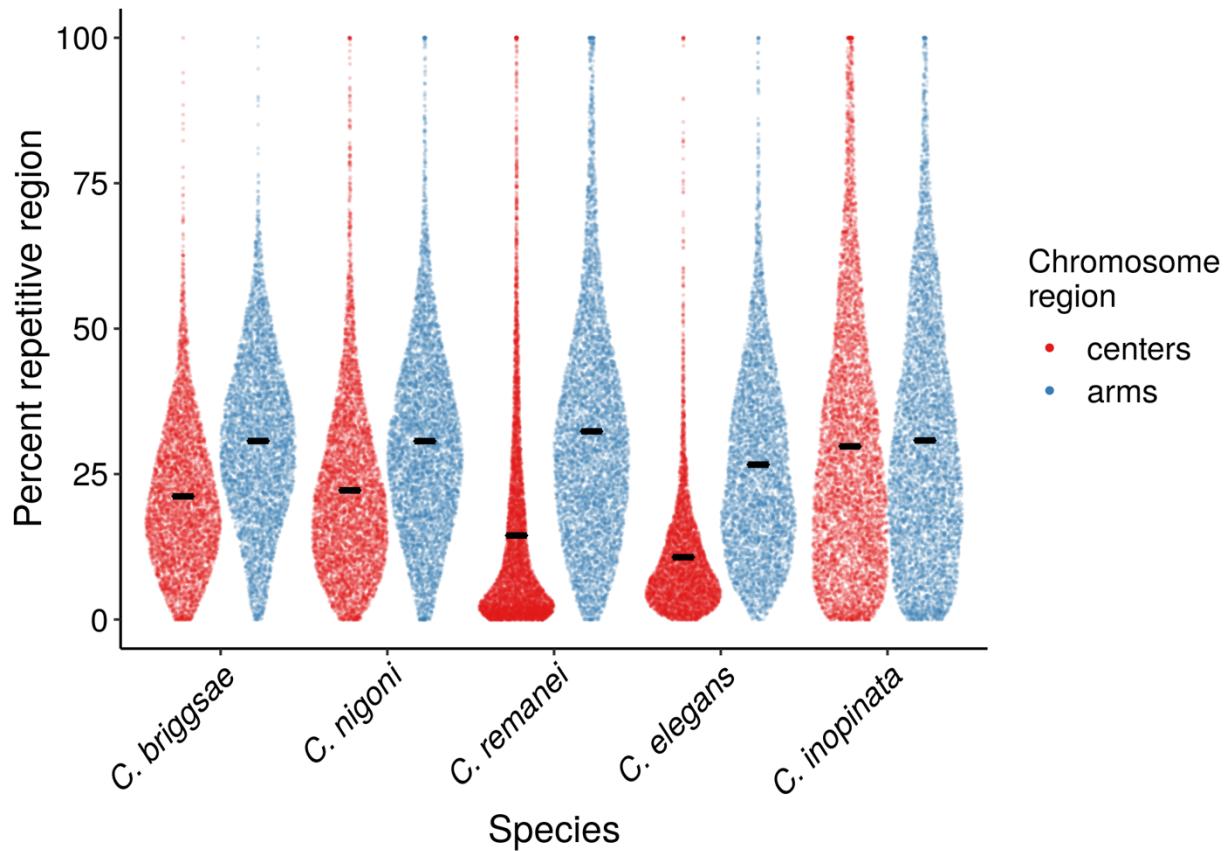


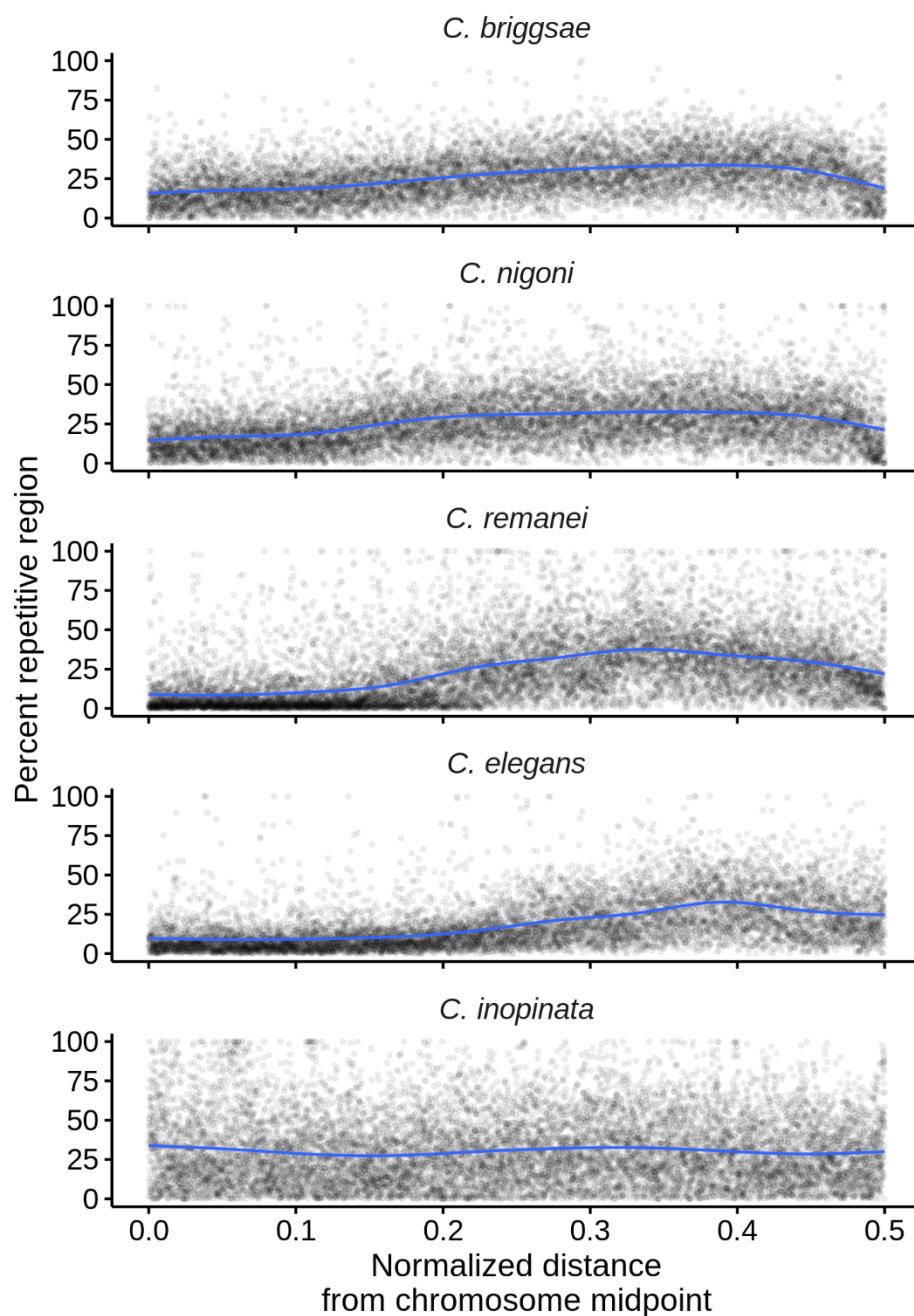
Supplemental figure 1. Relationship between genome assembly statistics and repeat content. Repeat content is plotted at the percentage of the genome assembly annotated as repetitive by the workflow described in the methods and in Supplemental Figure 2. These are plotted against various measures of genome quality on the axis in six facets. These measures and additional assembly statistics can be found in supplemental_tables.xls (assembly statistics, Sheet 1).



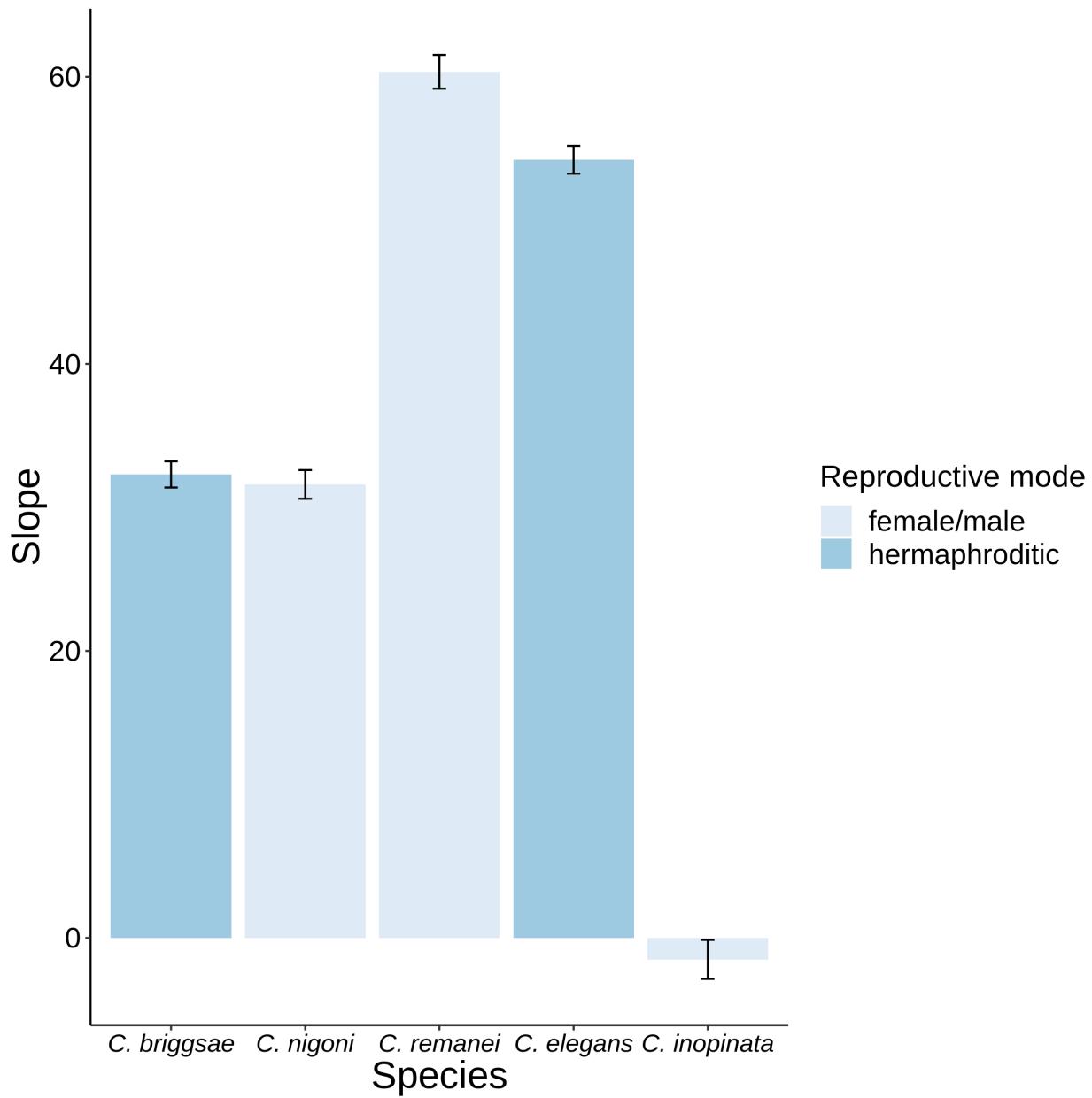
Supplemental figure 2. Repetitive element inference computational workflow. Libraries were generated independently for each species; genomes were masked with species-specific libraries (that is, they were not combined before the final masking).



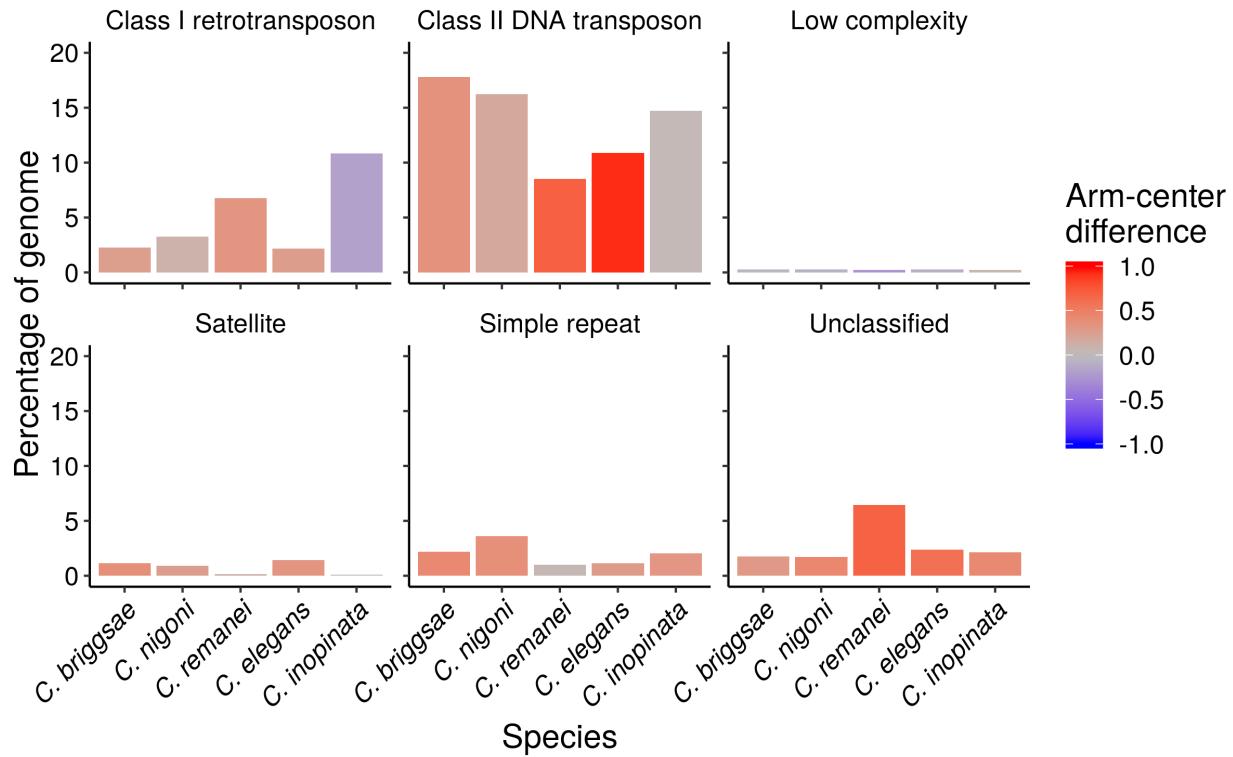
Supplemental Figure 3. Distribution of repeat densities on chromosome arms and centers. Each point represents one 10 kb genomic window. After normalizing chromosome positions by distance to chromosome midpoint, windows were classified as being in chromosome “centers” (middle half of chromosome) or “arms” (outer half of chromosome). Sina plots (strip charts with points taking the contours of a violin plot) reveal the distribution of repeat densities of these windows in chromosome centers and arms for all species. Black horizontal lines, means.



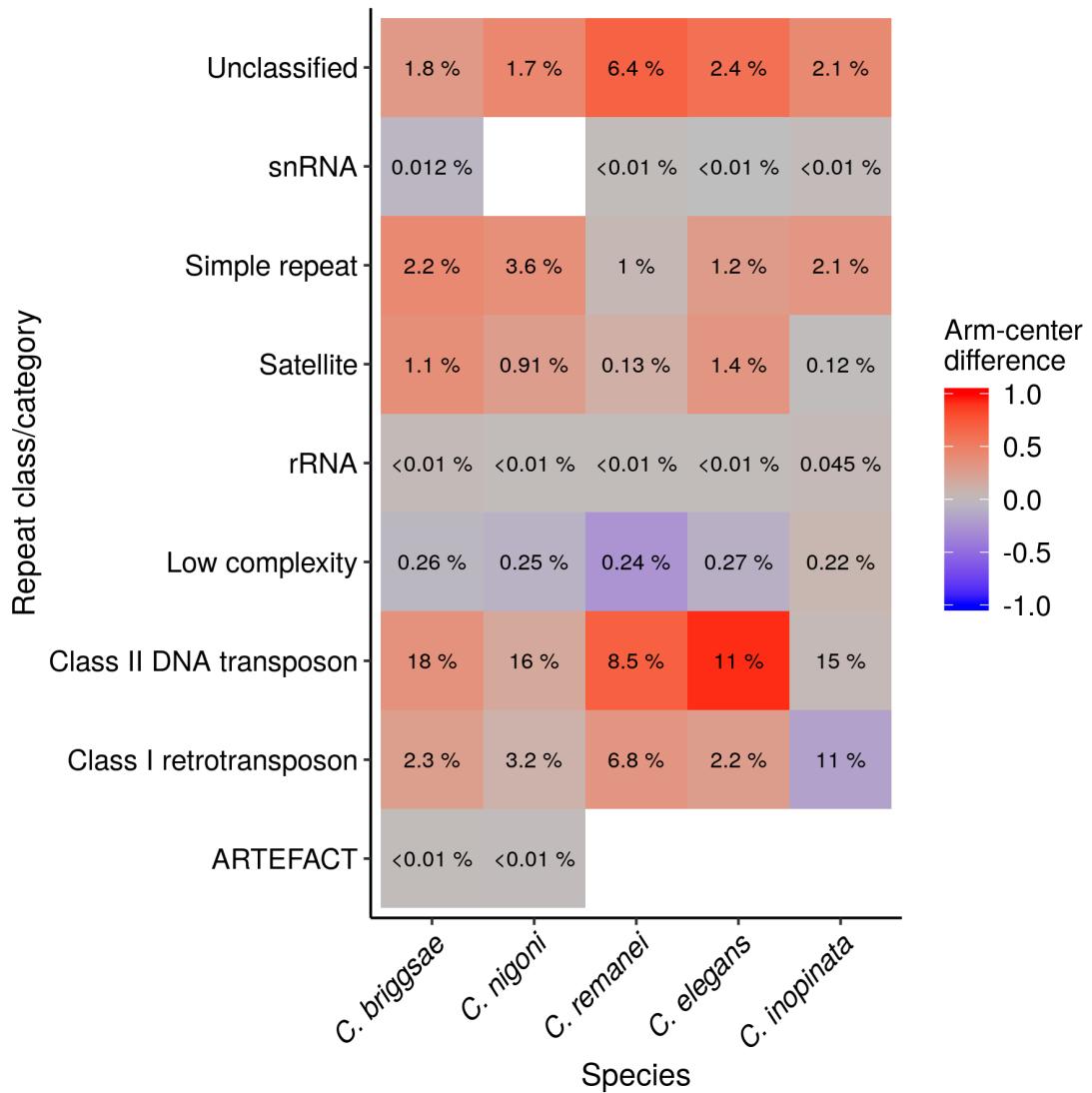
Supplemental Figure 4. The genomic landscape of repetitive elements in *Caenorhabditis* after normalizing genomic position by distance from chromosome midpoints. Here, all chromosomes are plotted, and “0” on the x-axis represents chromosome midpoints. Columns represent the six chromosomes; rows are the species ordered phylogenetically as in Figure 1. Plotted are the percentages of 10 kb windows that contain repetitive regions by genomic position. Blue lines are fit by generalized additive models.



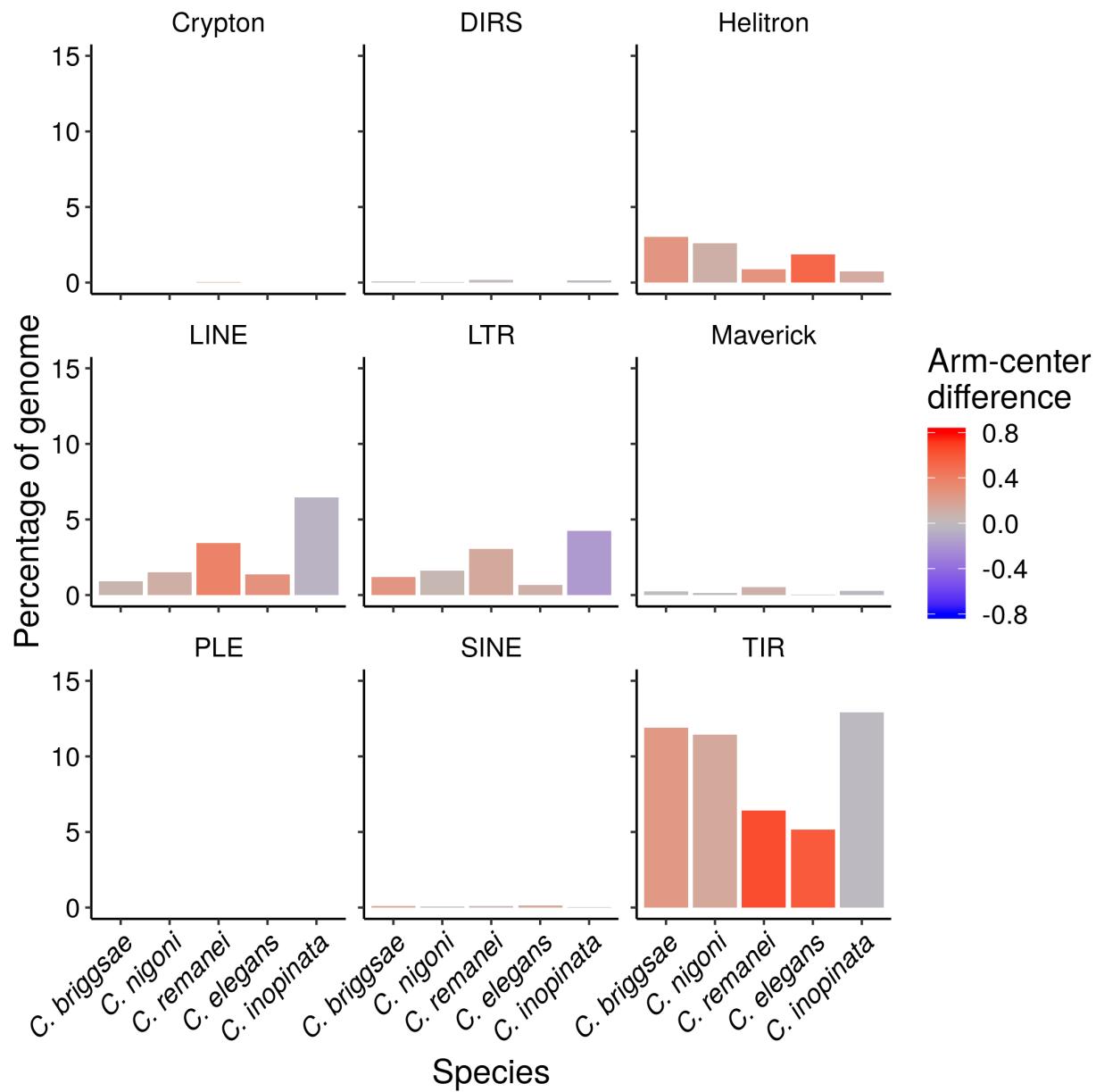
Supplemental Figure 5. Coefficients (slopes) of linear models of the relationship between repeat density (percentage of window repetitive) and normalized distance from chromosome center (range=0-0.5; “0” is the chromosomal midpoint). All model $p<0.0001$ except for *C. inopinata* ($p=0.27$). See Supplemental Material (supplemental_statistics.xls) for model statistics.



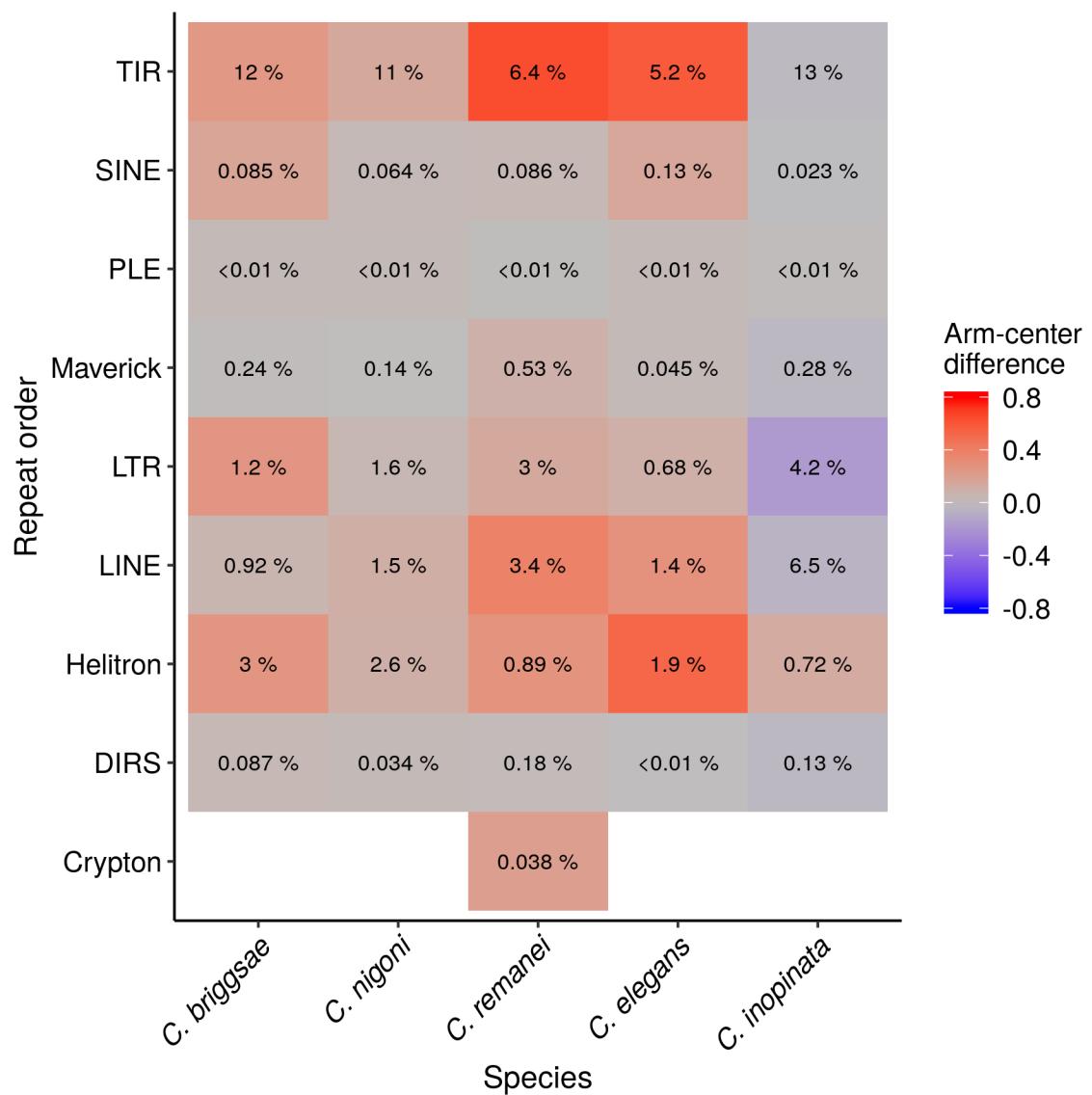
Supplemental Figure 6. The abundance and genomic landscape of the dominant repeat categories and the two transposable element classes in all species. The y-axis represents the percentage of the genome covered by that repeat category. The color gradient follows the Cohen's d effect size of chromosome position on repeat density as illustrated in Figure 4b. "Low complexity" refers to A-, GA-, or G-rich regions. "Simple repeat" refers to tandem repeats. Repetitive elements classified as "low complexity," "simple repeats" and "satellite" are included here as a matter of completeness but should be interpreted with caution because of their difficulty to assemble and annotate.



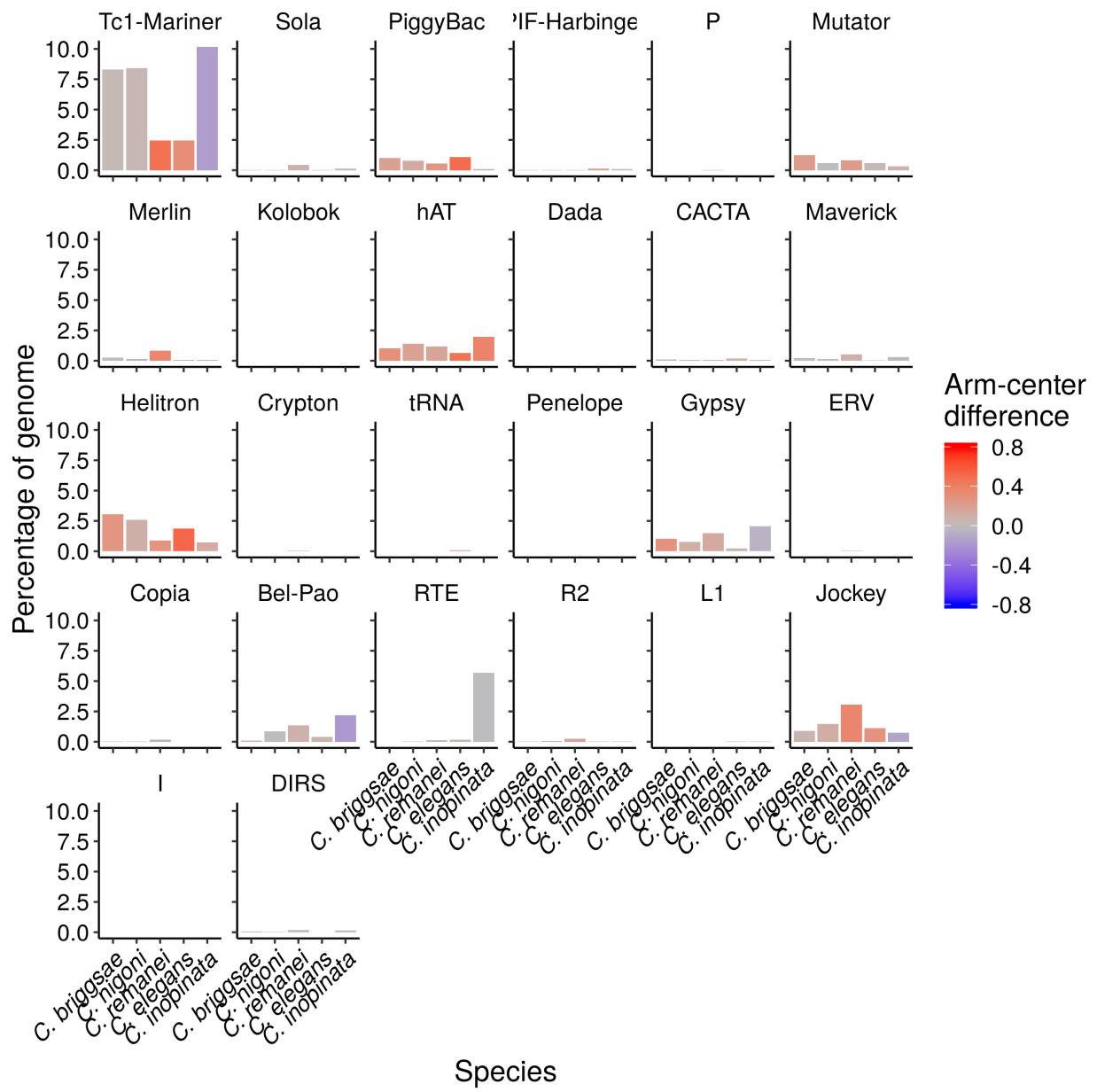
Supplemental Figure 7. Heat map of the genomic landscape of all repeat categories and the two transposable element classes in all species. The same data as in Supplemental Figure 6 are plotted. The color gradient follows the Cohen's d effect size of chromosome position on repeat density as illustrated in Figure 4b. The numbers in the tiles note the percentage of the genome occupied by the given repeat category. Uncolored tiles represent taxa or categories that were not detected in that species. "Unknown" are repetitive sequences that were not classified by RepeatClassifier. "ARTEFACT" refers to one repetitive element in the *C. briggsae* and *C. nigoni* genomes that aligned to *E. coli* in a BLAST search. "Low complexity" refers to A-, GA-, or G-rich regions. "Simple repeat" refers to tandem repeats. Repetitive elements classified as "low complexity," "simple repeats" and "satellite" are included here as a matter of completeness but should be interpreted with caution because of their difficulty to assemble and annotate.



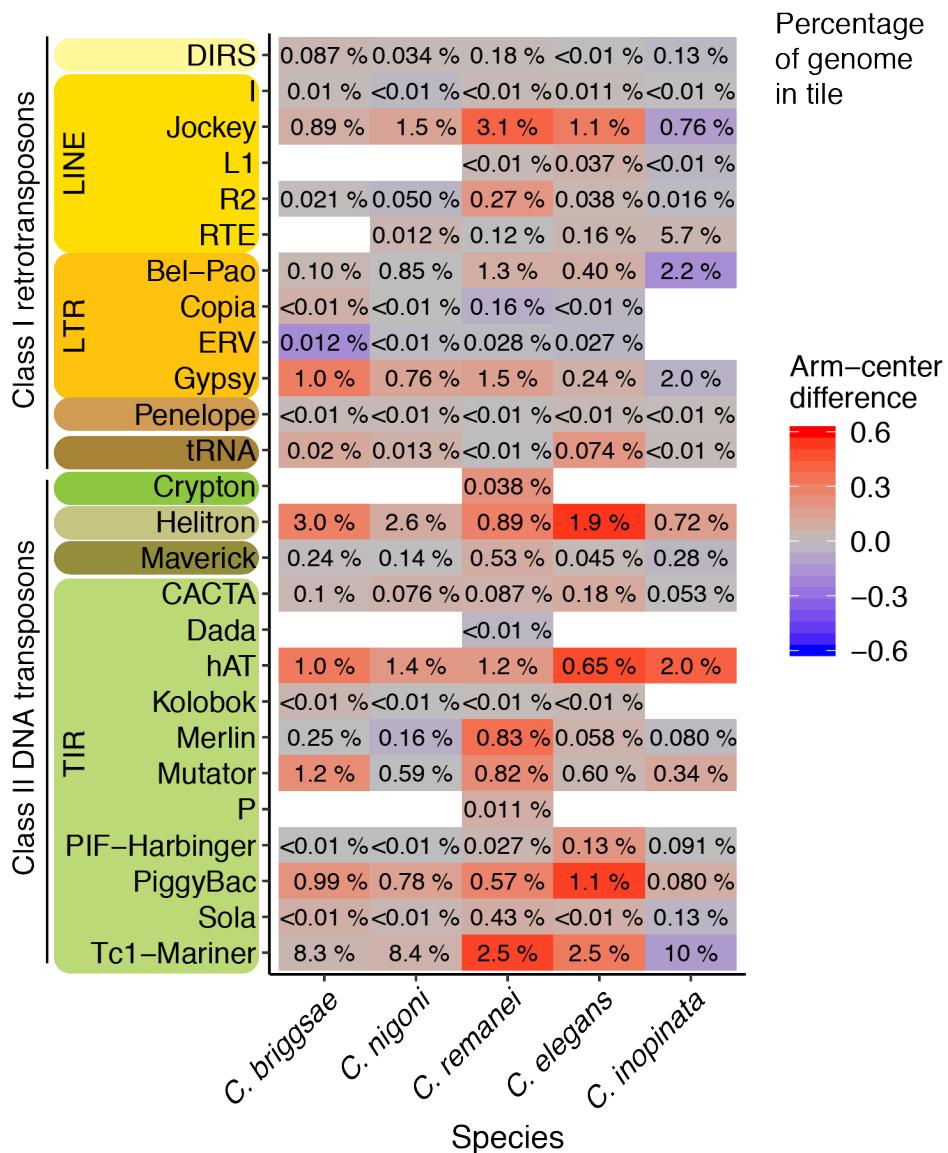
Supplemental Figure 8. The abundance and genomic landscape of all transposable element orders in all species. The y-axis represents the percentage of the genome covered by that repeat category. The color gradient follows the Cohen's d effect size of chromosome position on repeat density as illustrated in Figure 4b. Some elements in higher taxa were unable to be categorized into lower taxonomic ranks.



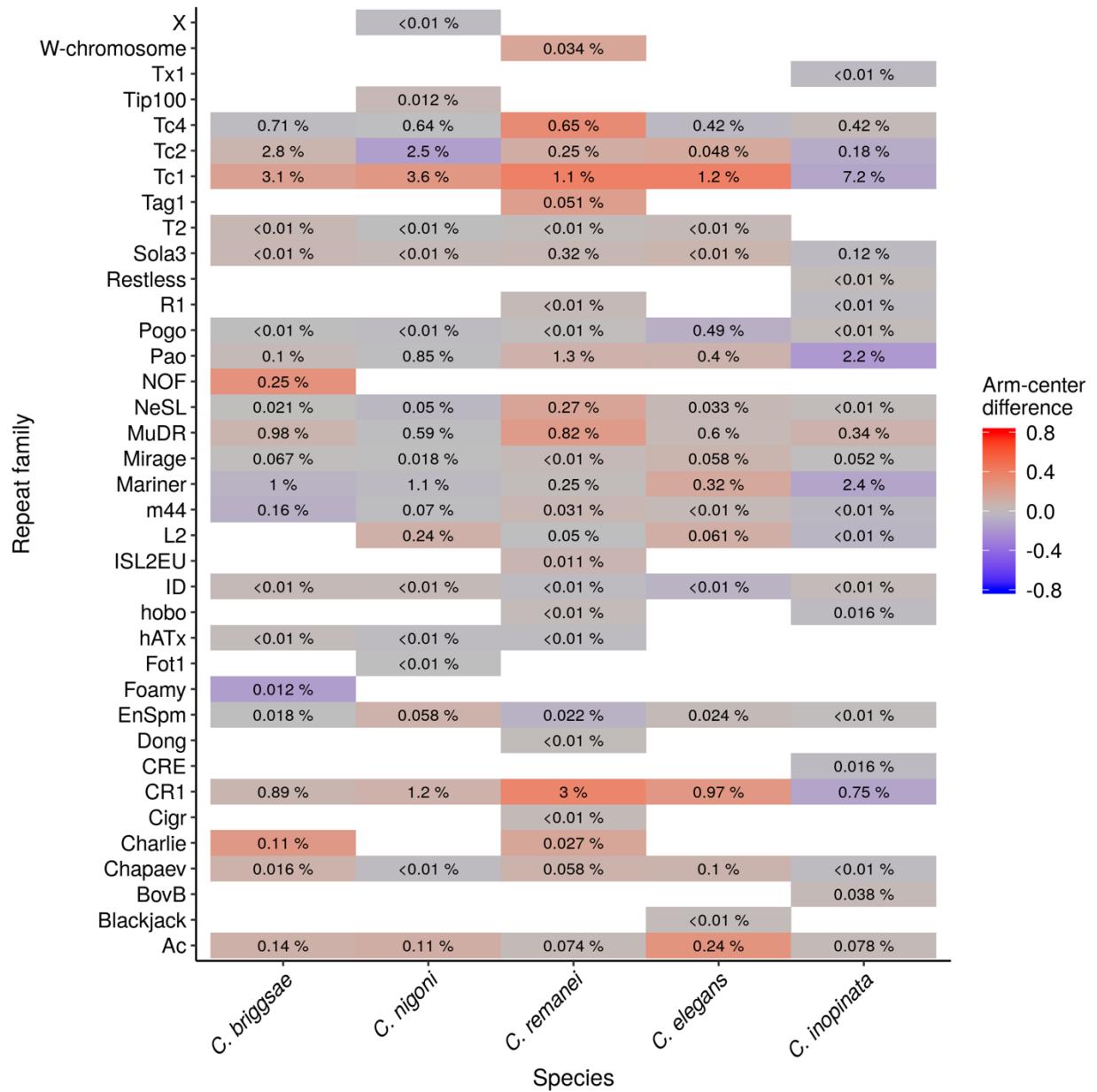
Supplemental Figure 9. Heat map of the genomic landscape all transposable element orders in all species. The same data as in Supplemental Figure 6 are plotted. The color gradient follows the Cohen's d effect size of chromosome position on repeat density as illustrated in Figure 4b. The numbers in the tiles note the percentage of the genome occupied by the given transposable element order. Uncolored tiles represent taxa that were not detected in that species. Some elements in higher taxa were unable to be categorized into lower taxonomic ranks.



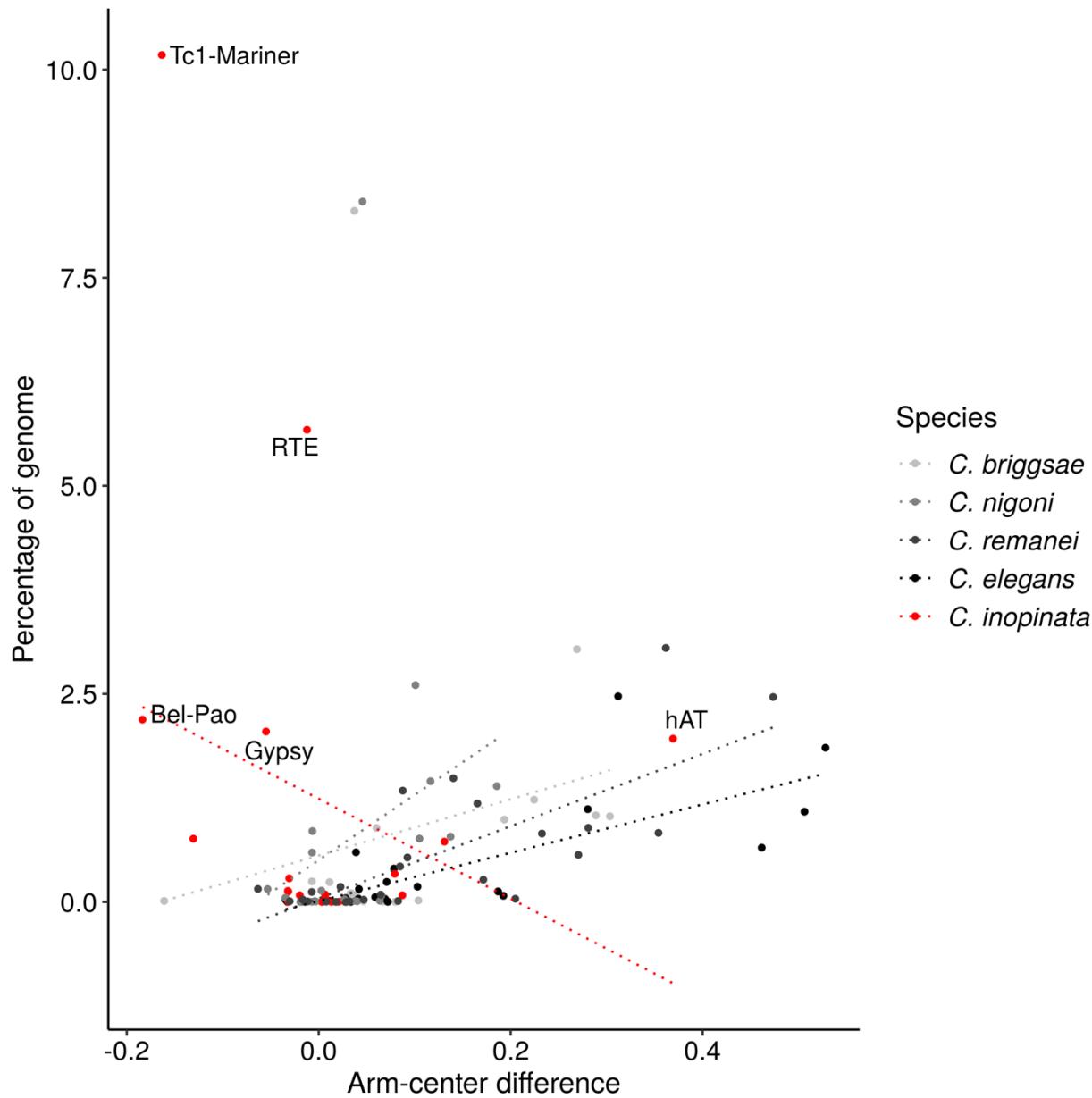
Supplemental Figure 10. The abundance and genomic landscape of all transposable element superfamilies in all species. The y-axis represents the percentage of the genome covered by that repeat category. The color gradient follows the Cohen's d effect size of chromosome position on repeat density as illustrated in Figure 4b. Some elements in higher taxa were unable to be categorized into lower taxonomic ranks.



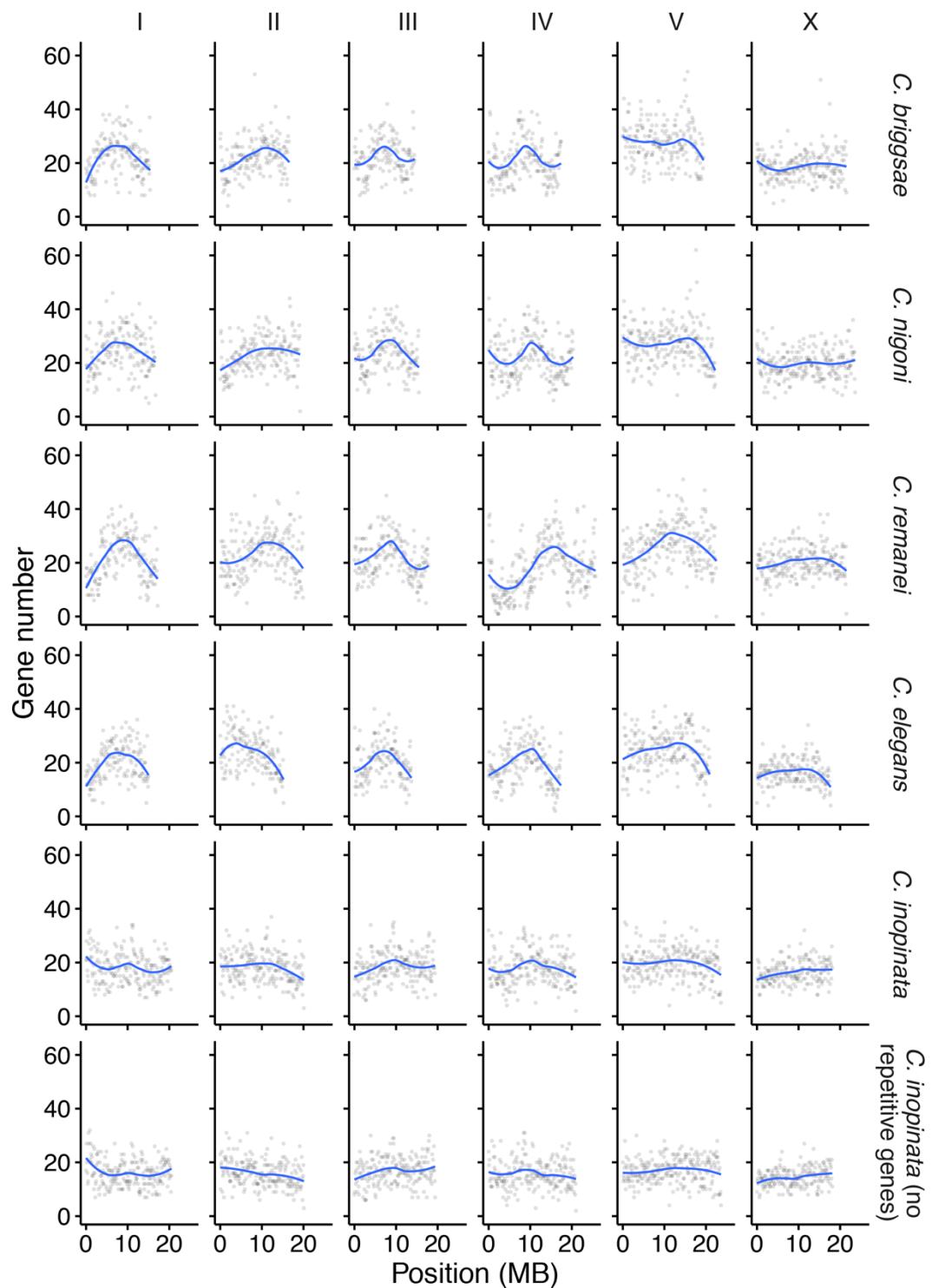
Supplemental Figure 11. Heat map of the genomic landscape of all transposable element superfamilies in all species. The same data as in Supplemental Figure 8 are plotted. The color gradient follows the Cohen's d effect size of chromosome position on repeat density as illustrated in Figure 4b. The numbers in the tiles note the percentage of the genome occupied by the given repeat superfamily. Superfamilies are clustered by more-encompassing taxonomic groups, and they are colored by repeat orders. Uncolored tiles represent superfamilies that were not detected in that species. Some elements in higher taxa were unable to be categorized into lower taxonomic ranks.



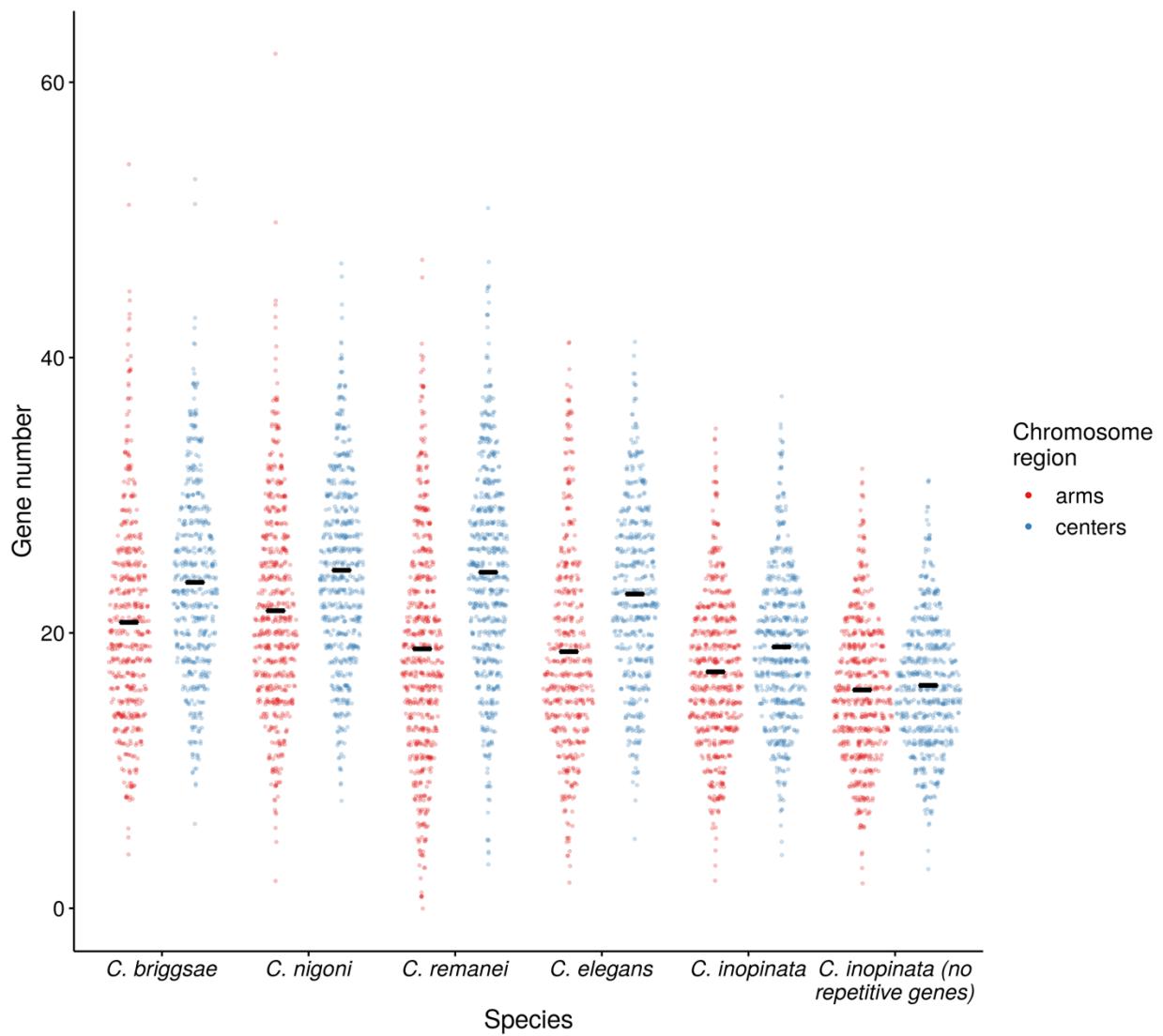
Supplemental Figure 12. Heat map of the genomic landscape of all transposable element families in all species. The color gradient follows the Cohen's d effect size of chromosome position on repeat density as illustrated in Figure 4b. The numbers in the tiles note the percentage of the genome occupied by the given transposable element family. Uncolored tiles represent families that were not detected in that species. Some elements in higher taxa were unable to be categorized into lower taxonomic ranks.



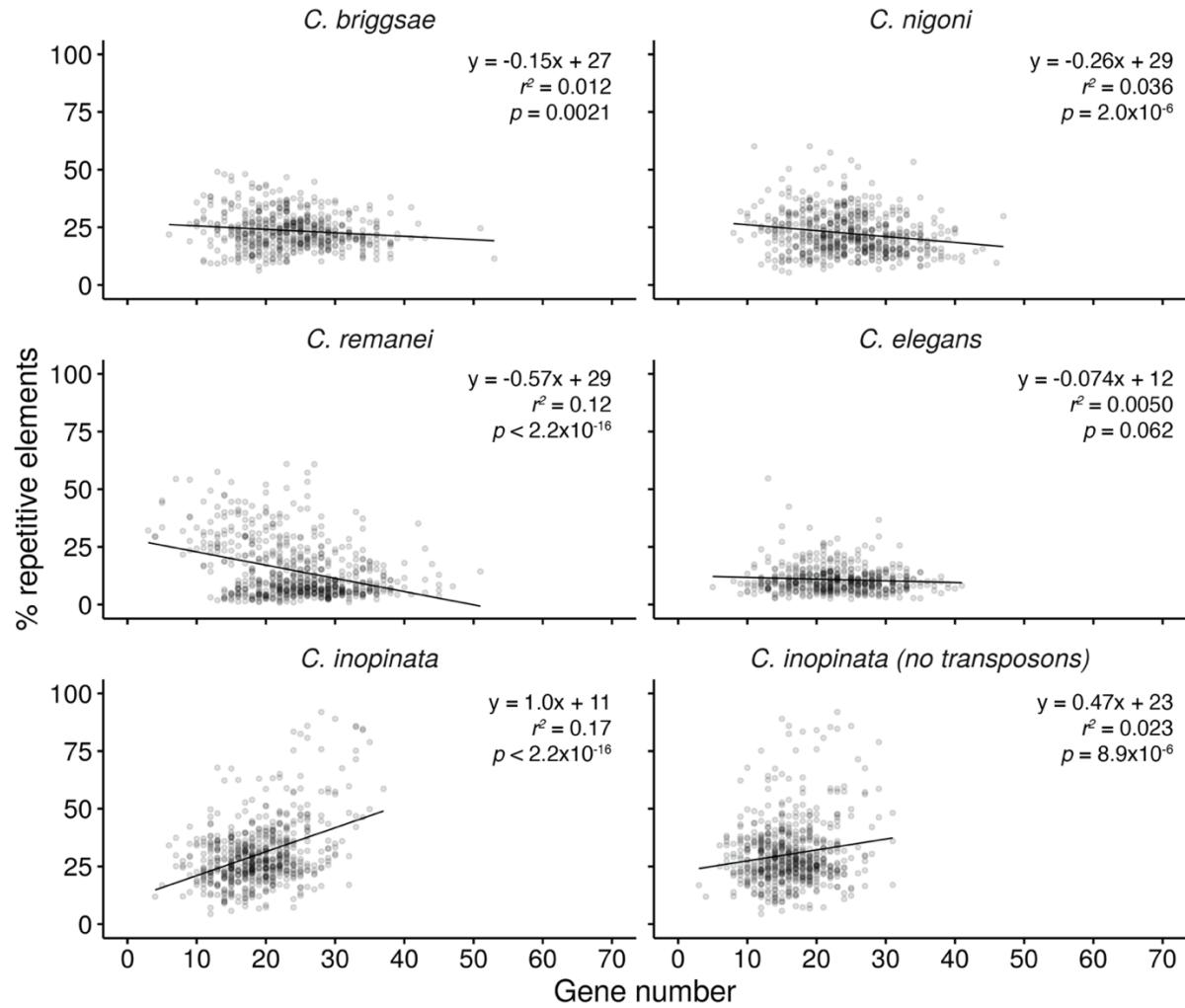
Supplemental Figure 13. The relationship between repeat chromosomal structure and total genomic repeat content among repeat superfamilies in five *Caenorhabditis* species (not log-transformed). The “arm-center difference” is the Cohen’s d effect size of chromosome position on repeat density as described in Figure 4b; this is the same data as Figure 4c but not log-transformed. The five most abundant repeat superfamilies in *C. inopinata* are labeled. Regression statistics can be found in the text and in the supplemental material.



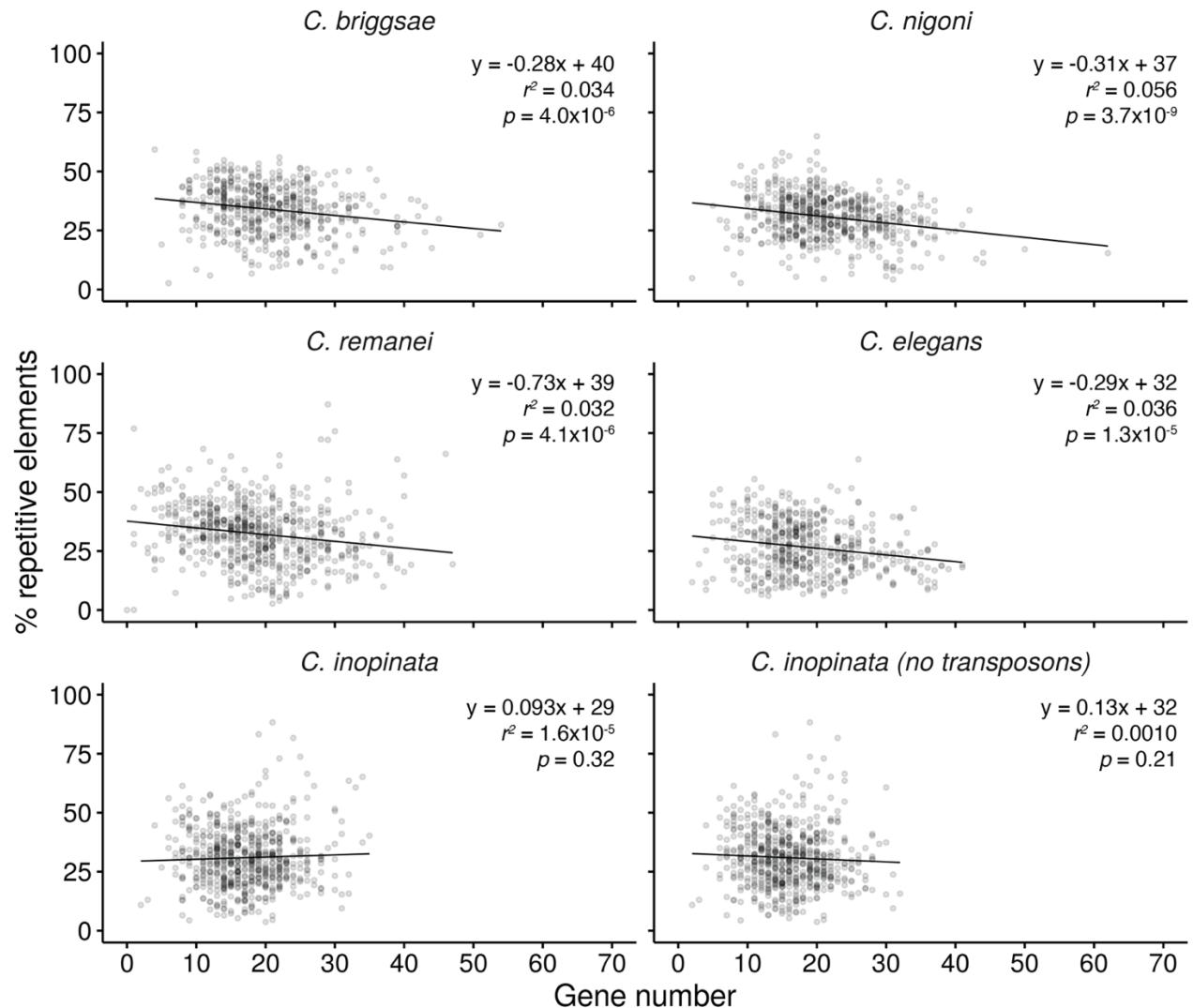
Supplemental Figure 14. Gene number among 100kb genomic windows in five *Caenorhabditis* species. Here, the gene densities of *C. inopinata* when 2,489 transposon-aligning genes are excluded are also plotted. The blue line was fit by LOESS local regression.



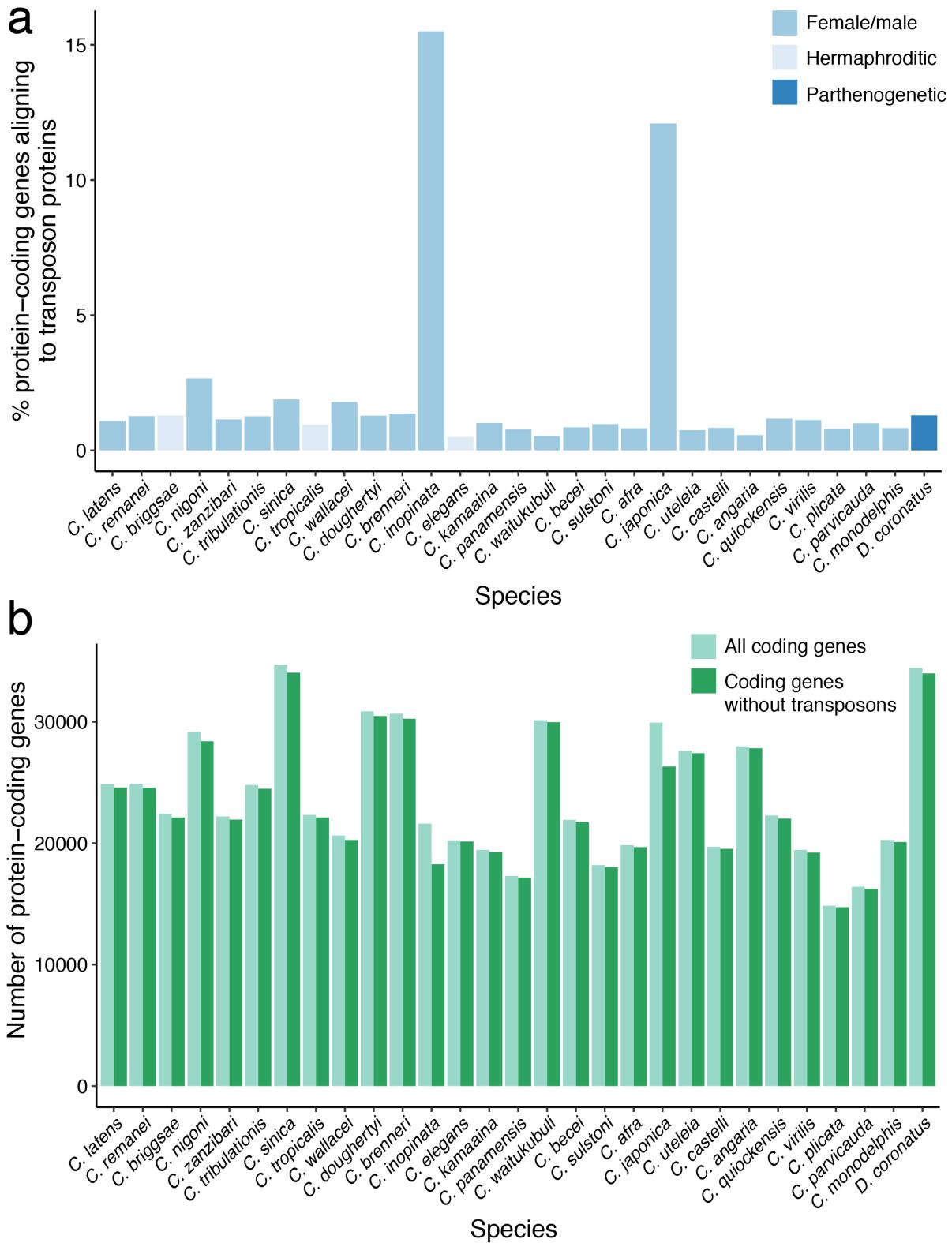
Supplemental Figure 15. Distribution of gene counts on chromosome arms and centers. Each point represents one 100 kb genomic window. After normalizing chromosome positions by distance to chromosome midpoint, windows were classified as being in chromosome “centers” (middle half of chromosome) or “arms” (outer half of chromosome). Sina plots (strip charts with points taking the contours of a violin plot) reveal the distribution of gene counts in these windows in chromosome centers and arms for all species. Here, the gene densities of *C. inopinata* when 2,489 transposon-aligning genes are excluded are also plotted (last position on x-axis). Black horizontal lines, means.



Supplemental figure 16. Relationship between gene number and repeat density in chromosome centers only. Plotted are the percent repetitive region by gene count in 100 kb windows across all species. In the case of *C. inopinata*, an additional plot excludes 2,489 transposon-aligning genes (that also do not align to any *C. elegans* proteins) from gene counts.

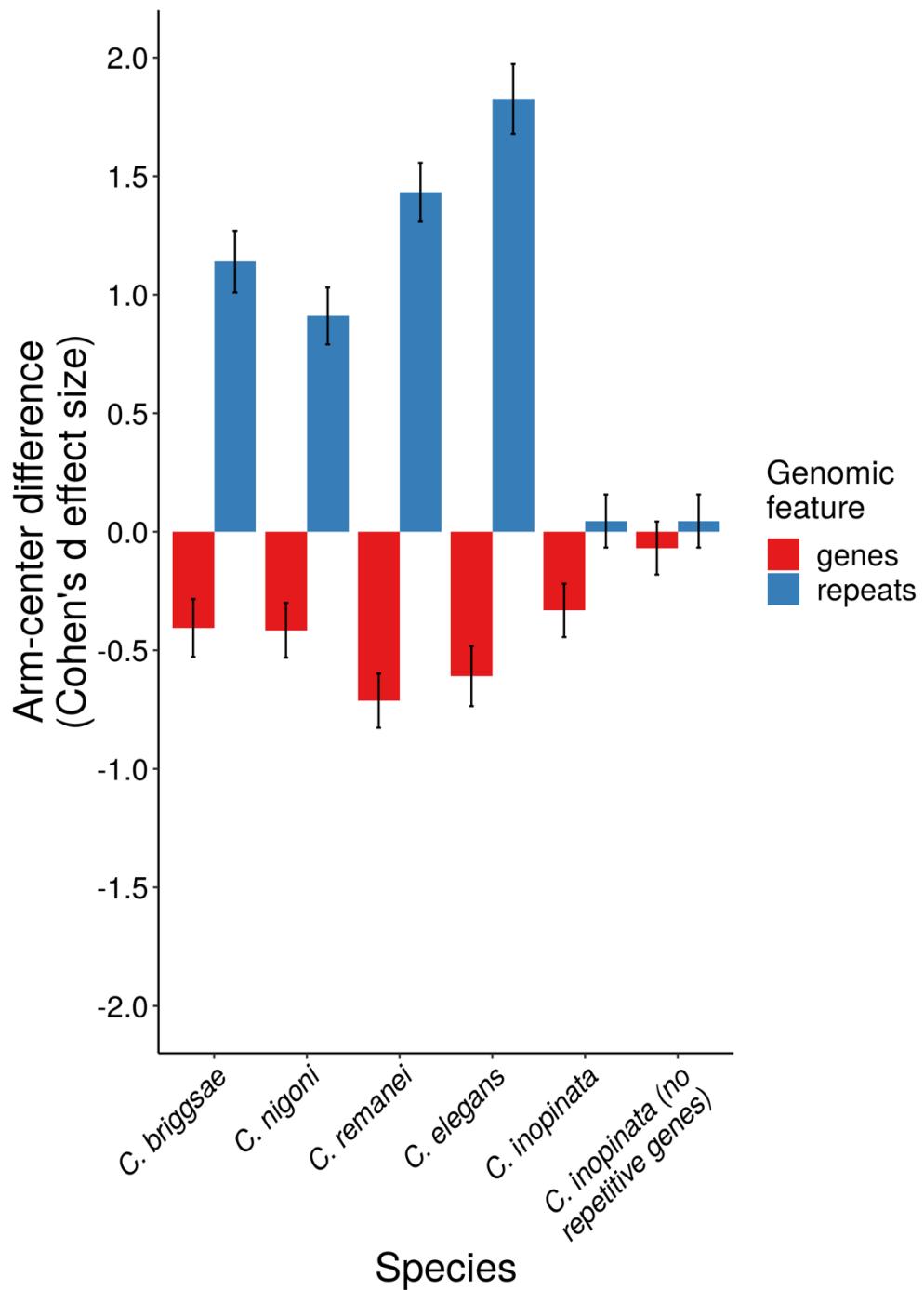


Supplemental figure 17. Relationship between gene number and repeat density in chromosome arms only. Plotted are the percent repetitive region by gene count in 100 kb windows across all species. In the case of *C. inopinata*, an additional plot excludes 2,489 transposon-aligning genes (that also do not align to any *C. elegans* proteins) from gene counts.

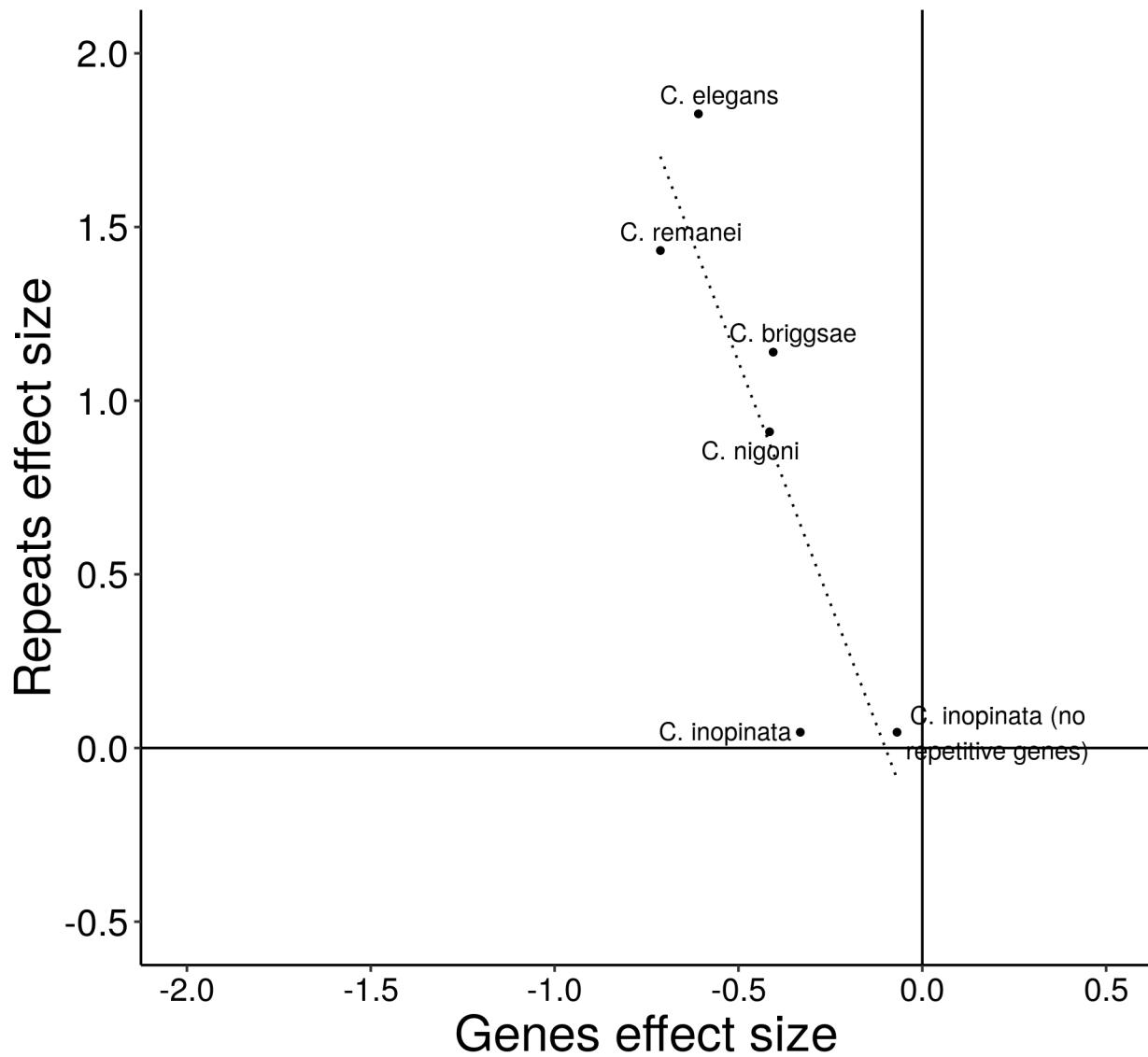


Supplemental Figure 18. Predicted protein-coding genes in all available *Caenorhabditis* genomes align to transposon-related proteins. The parthenogenetic *Diploscapter coronatus* is included as an outgroup.

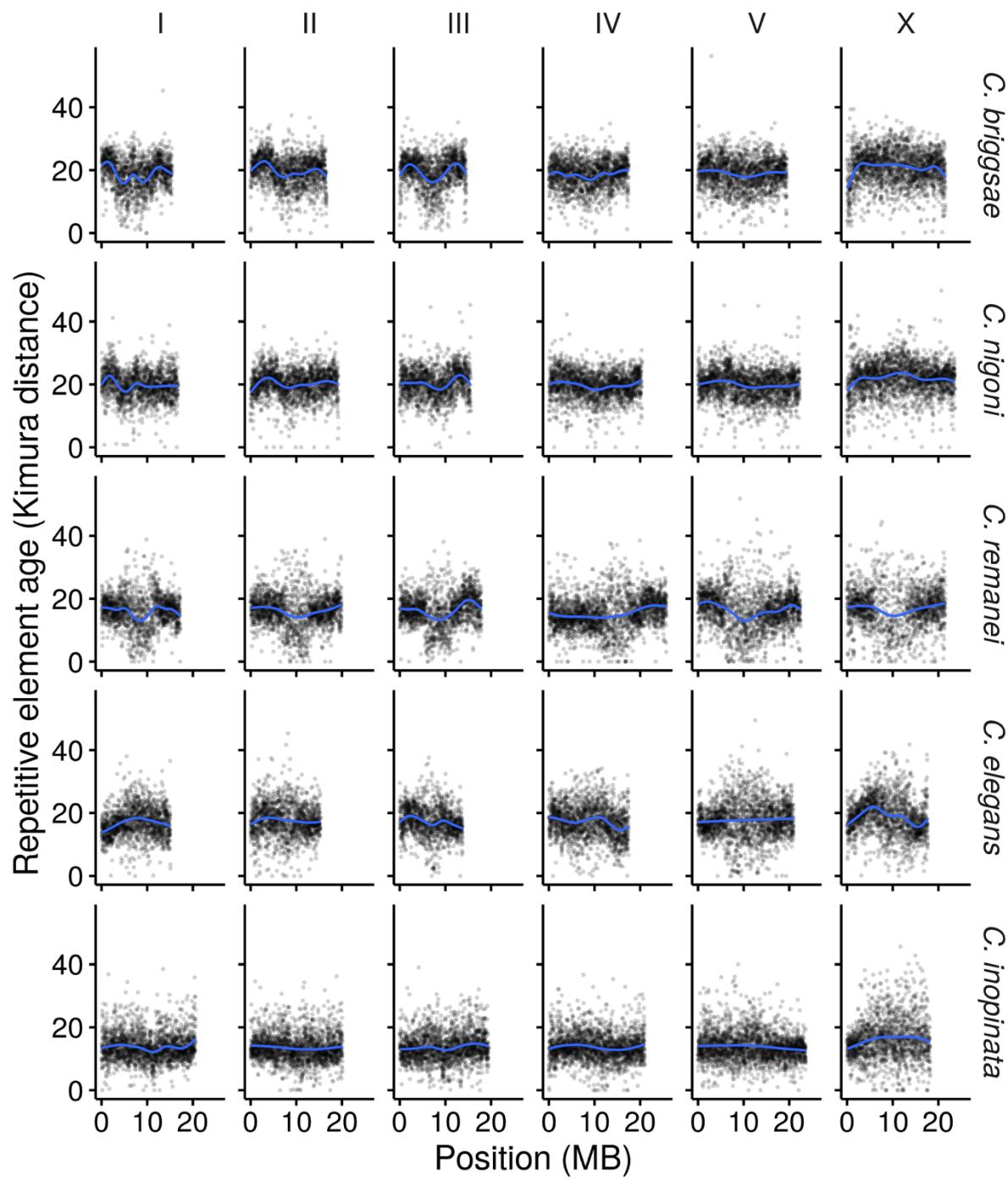
- a. The percentage of protein-coding genes that align to transposons in all available *Caenorhabditis* genomes. Species are roughly ordered by phylogeny and colored by reproductive mode.
- b. The total number of predicted protein-coding genes before (light green) and after (dark green) transposon-aligning genes are removed.



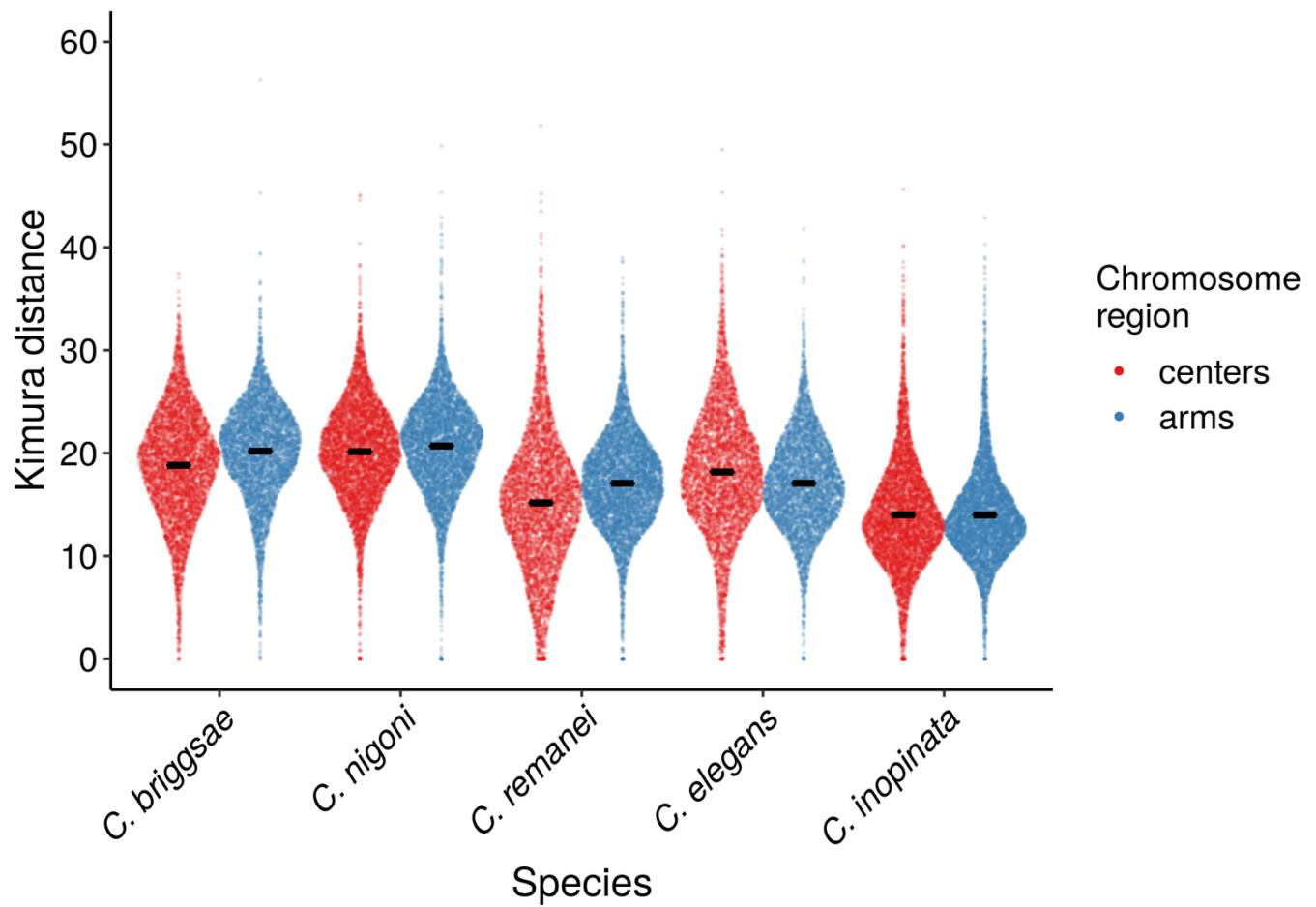
Supplemental figure 19. Effect sizes of chromosome position on repeat and gene density. Plotted are the Cohen's d effect sizes of chromosome arms compared to centers on gene and repeat density. Error bars represent 95% confidence intervals. The gene densities of *C. inopinata* when 2,489 transposon-aligning genes are excluded are also plotted (last position on x-axis). These effect sizes are based on repeat percentages and gene counts in 100 kb genomic windows (same data as Figure 7).



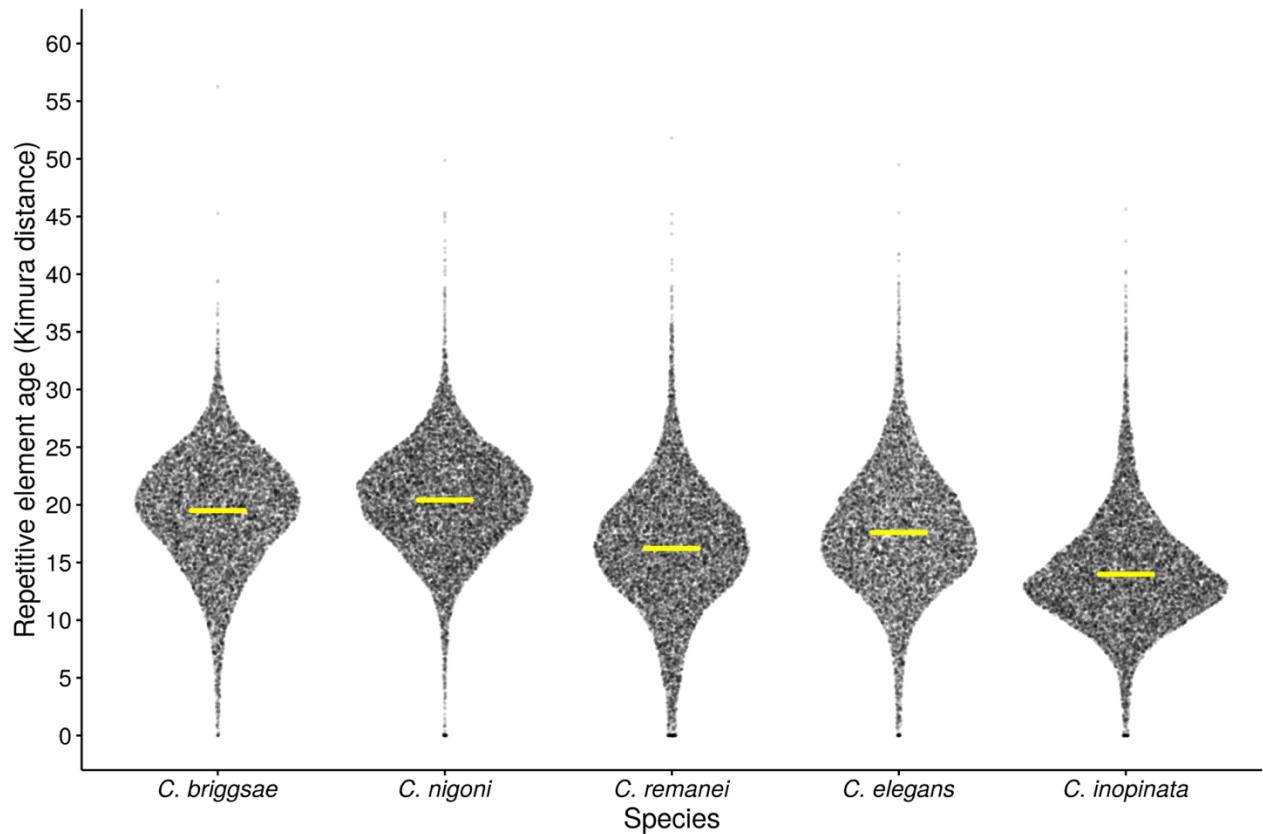
Supplemental figure 20. Effect sizes of chromosome position on repeat density covaries with effect sizes of chromosome position on gene density. Plotted are the Cohen's d effect sizes of chromosome arms compared to centers on gene and repeat density. The gene densities of *C. inopinata* when 2,489 transposon-aligning genes are excluded are also plotted. These effect sizes are based on repeat percentages and gene counts in 100 kb genomic windows (same data as Figure 7 and Supplemental figure 18). The linear fit is significant ($r^2=0.67$; $F=11$; $\beta_1 = -2.8$; $p=0.029$) but dubious as *C. inopinata* is double-counted and the sample is small. When *C. inopinata* (including transposon-aligning genes) is excluded, the fit remains significant ($r^2=0.81$; $F=18$; $\beta_1 = -2.5$; $p=0.024$) but with an even smaller sample.



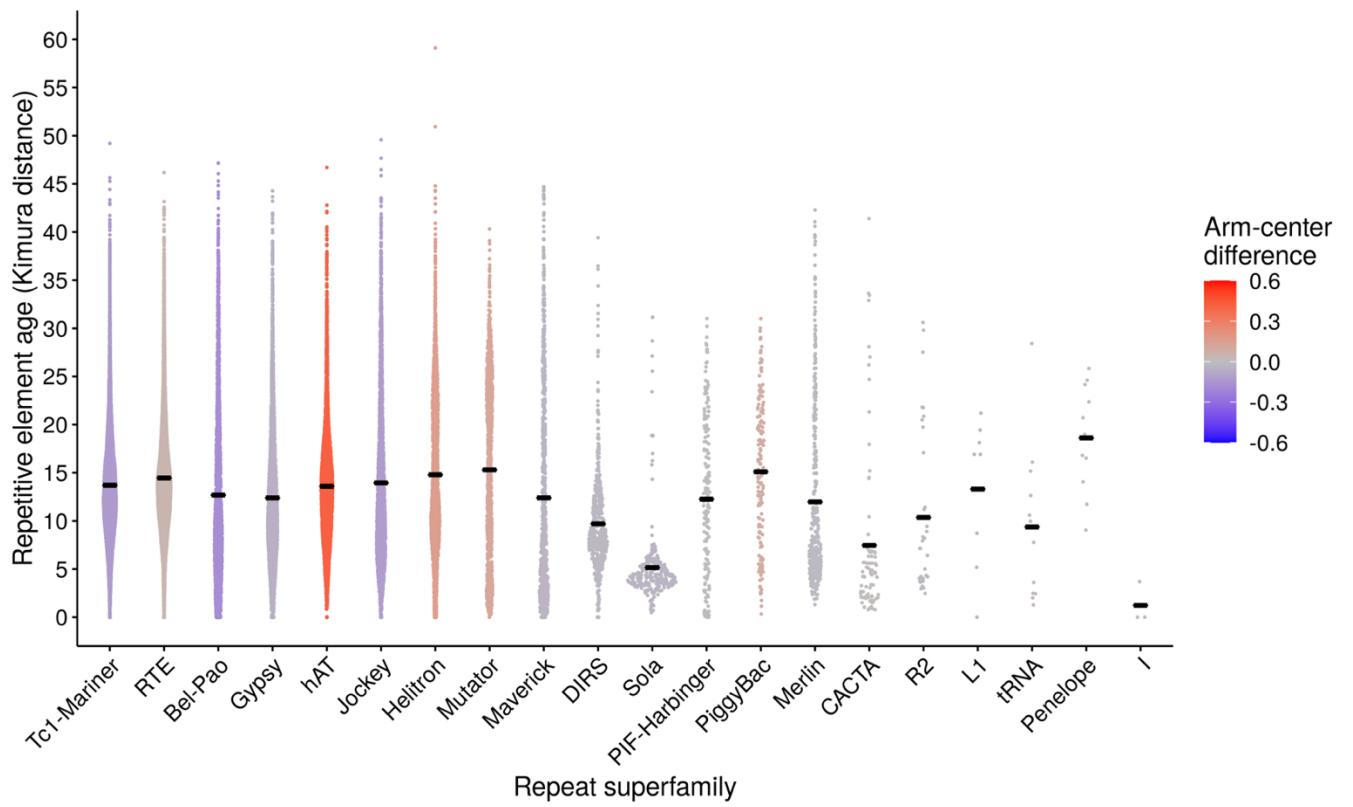
Supplemental figure 21. Repetitive element ages (Kimura distances) among 10kb genomic windows in five *Caenorhabditis* species. Plotted is the average Kimura distance in that window. Blue lines are fit by generalized additive models.



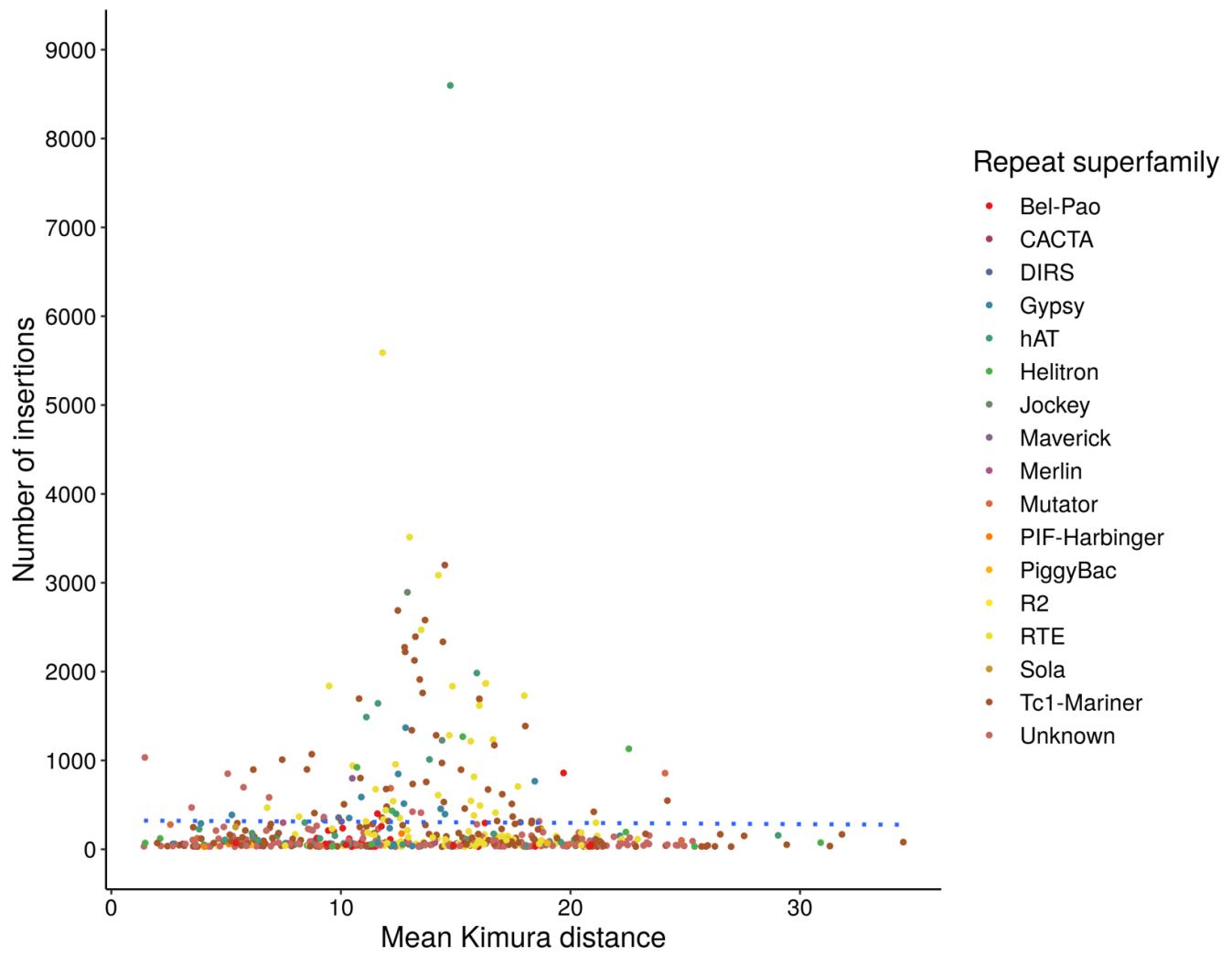
Supplemental figure 22. Distribution of repetitive element ages (Kimura distances) on chromosome arms and centers. Each point represents the mean Kimura distance of one 10 kb genomic window. After normalizing chromosome positions by distance to chromosome midpoint, windows were classified as being in chromosome “centers” (middle half of chromosome) or “arms” (outer half of chromosome). Sina plots (strip charts with points taking the contours of a violin plot) reveal the distribution of repeat ages of these windows in chromosome centers and arms for all species. Black horizontal lines, means.



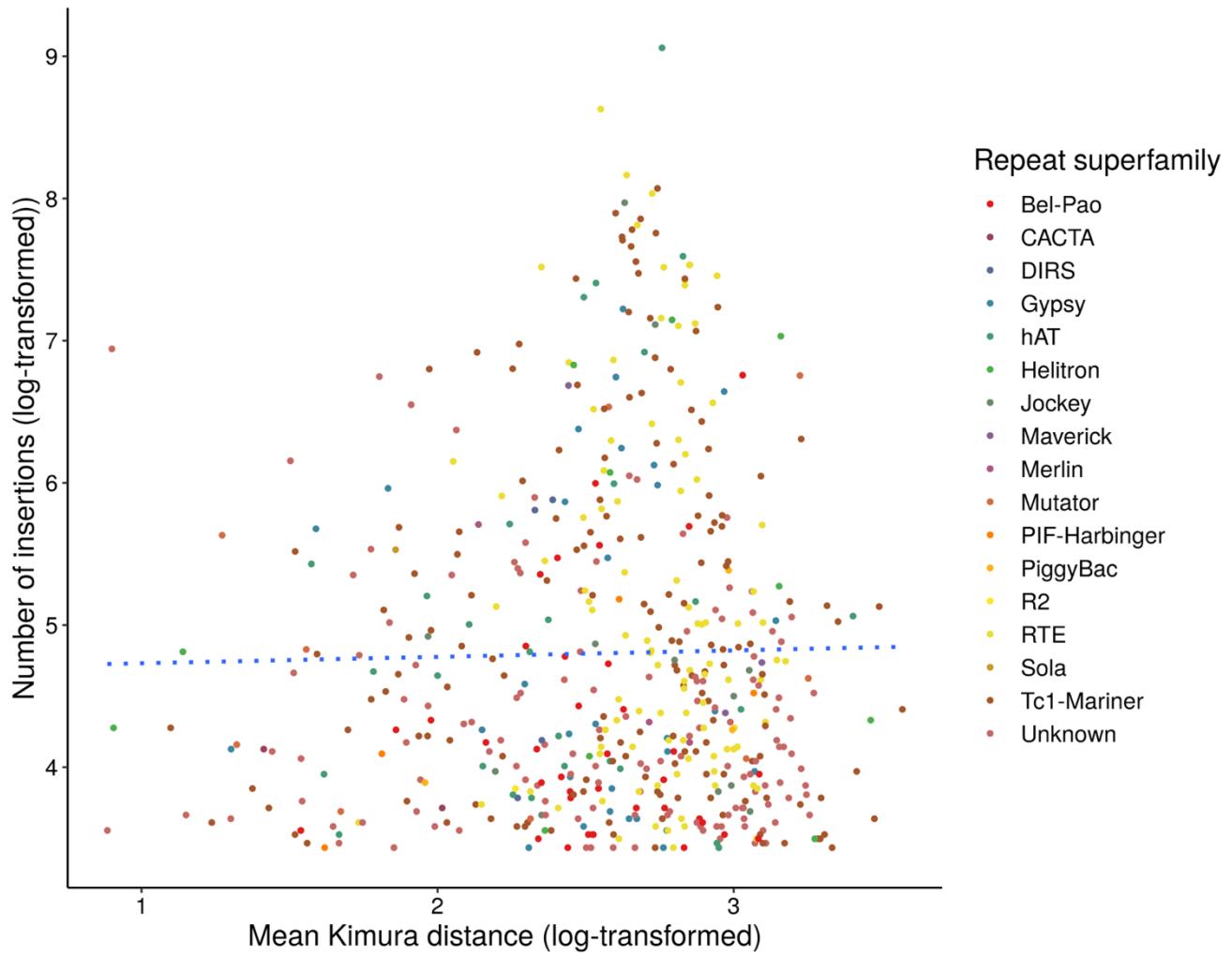
Supplemental figure 23. Distribution of repetitive element ages (Kimura distances) for all species. Each point represents the mean Kimura distance of one 10 kb genomic window. Sina plots (strip charts with points taking the contours of a violin plot) reveal the distribution of repeat ages of these windows. Yellow horizontal lines, means.



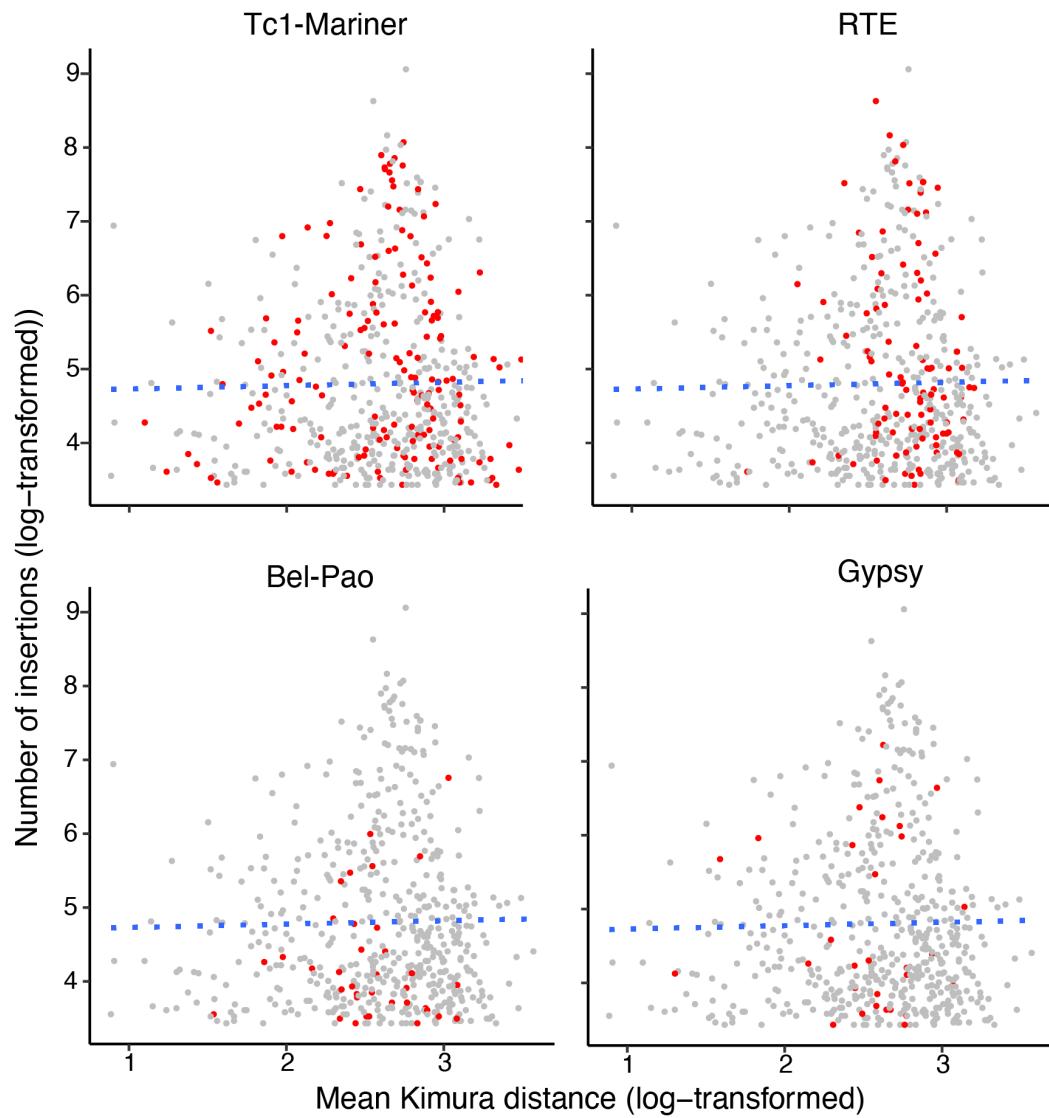
Supplemental figure 24. Abundant and atypically distributed repeat superfamilies are not exceptionally younger nor multimodal in *C. inopinata*. Ordered by decreasing genomic abundance. Each point represents the mean Kimura distance of one 10 kb genomic window that contains an element of that superfamily. Sina plots (strip charts with points taking the contours of a violin plot) reveal the distribution of repeat ages of these windows. The color gradient follows the Cohen's d effect size of chromosome position on repeat density as illustrated in Figure 4b. Black horizontal lines, means.



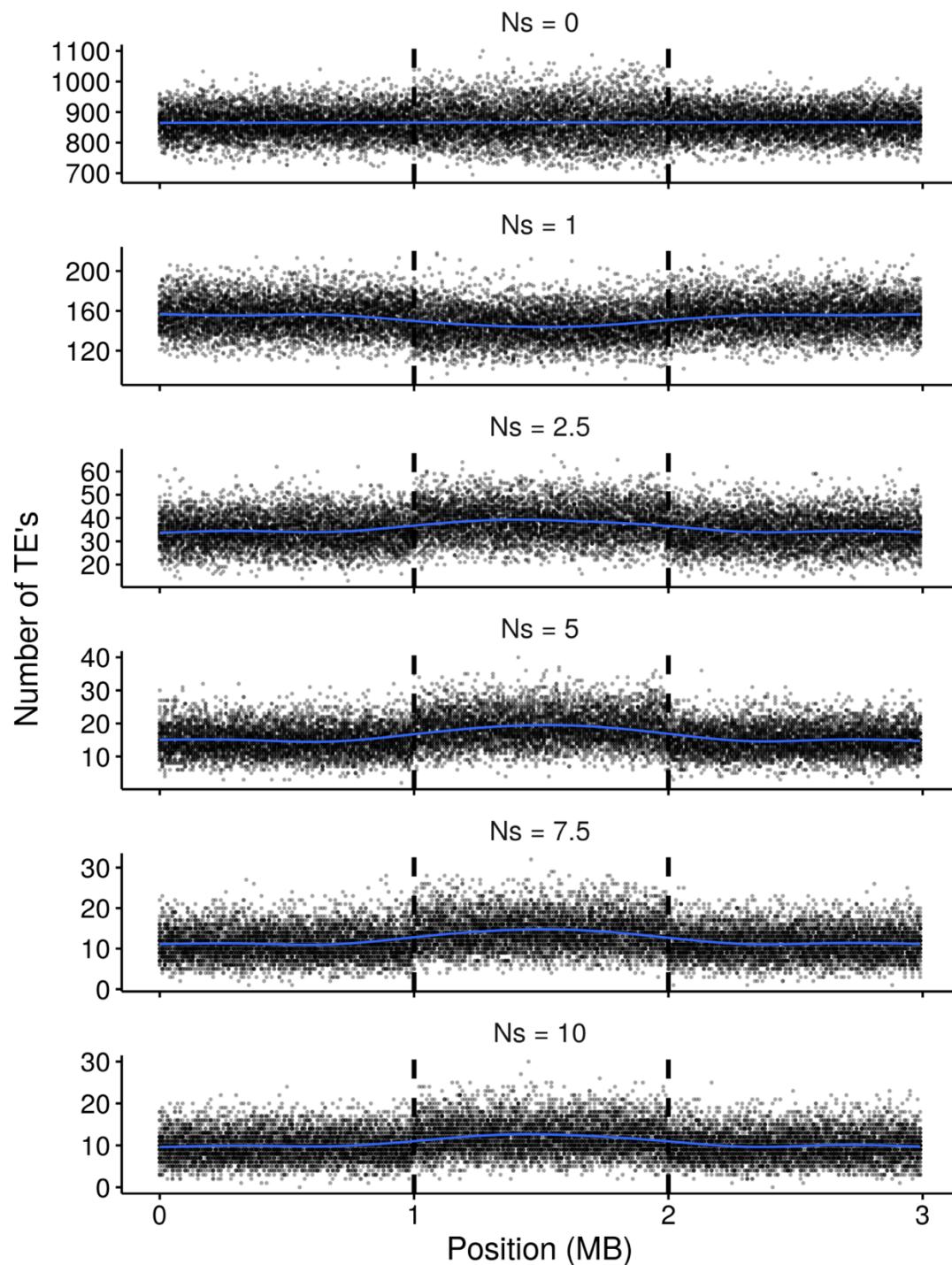
Supplemental figure 25. Relationship between insertion number and average age (Kimura distance) for all repeat clusters with ≥ 30 insertions in *C. inopinata*. Blue line fit by a linear model ($r^2 = -0.0017$; $F = 0.084$; $\beta_1 = -1.4$; $p = 0.77$).



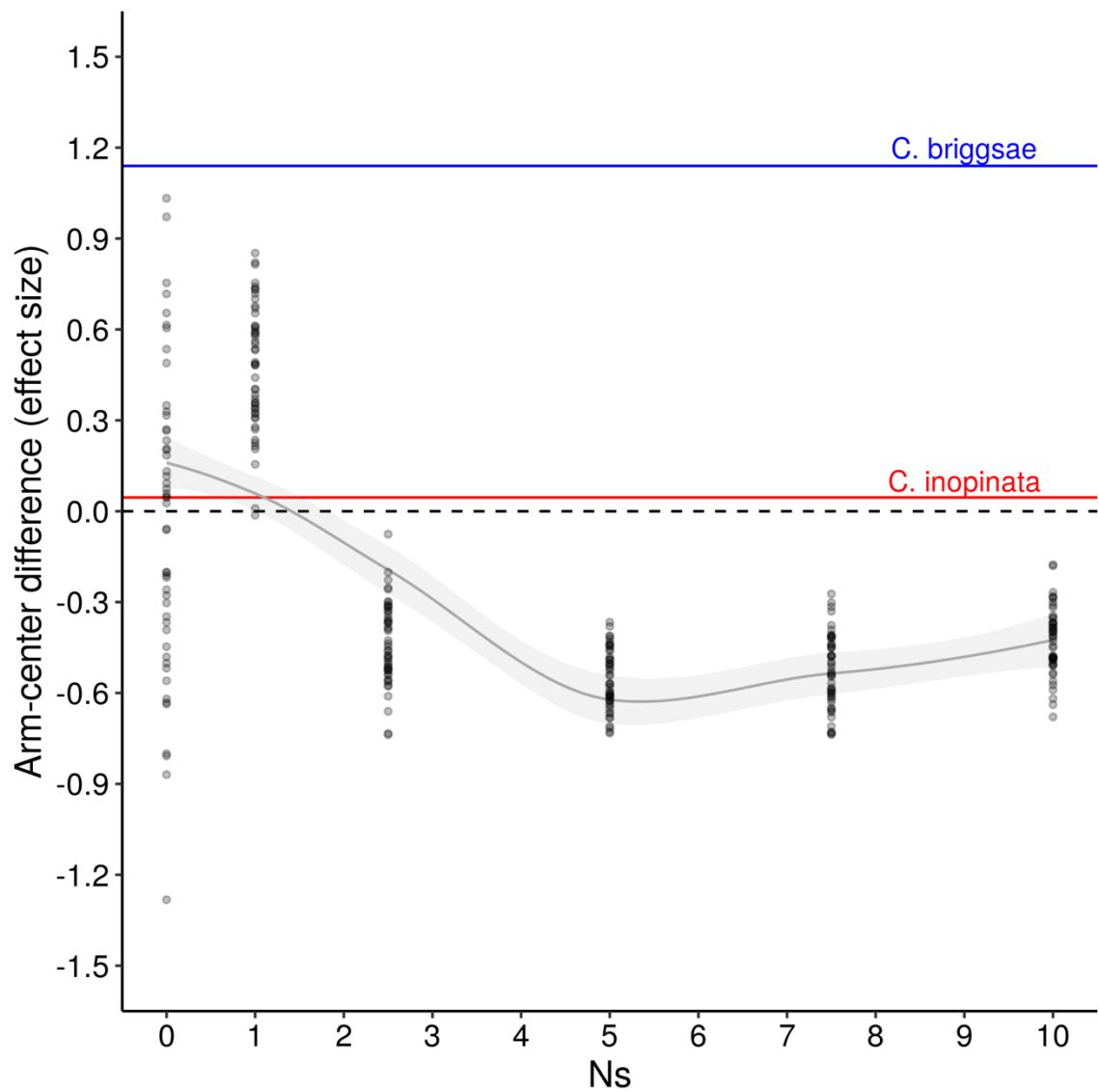
Supplemental figure 26. Relationship between mean insertion number and age for all repeat clusters with ≥ 30 insertions in *C. inopinata*, log-transformed ($\log(\text{variable}+1)$). Blue line fit by a linear model ($r^2 = -0.0015$; $F = 0.18$; $\beta_1 = 0.045$; $p = 0.67$). Same data as Supplemental Figure 25.



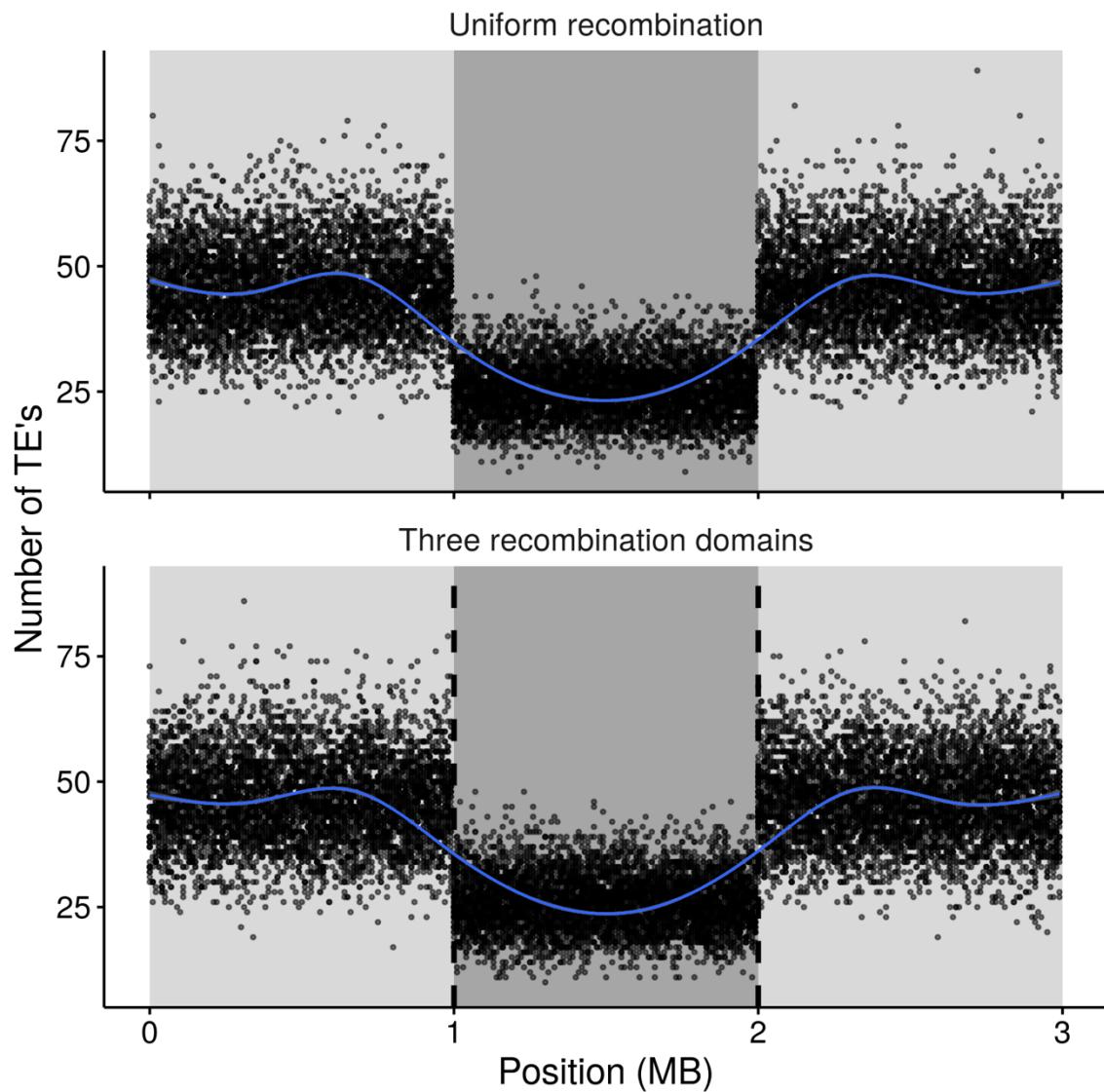
Supplemental figure 27. Relationship between mean insertion number and age for all repeat clusters with ≥ 30 insertions in *C. inopinata*, log-transformed ($\log(\text{variable}+1)$), with points colored by exceptionally abundant and atypically distributed repeat superfamilies. Same data as Supplemental Figures 25-26.



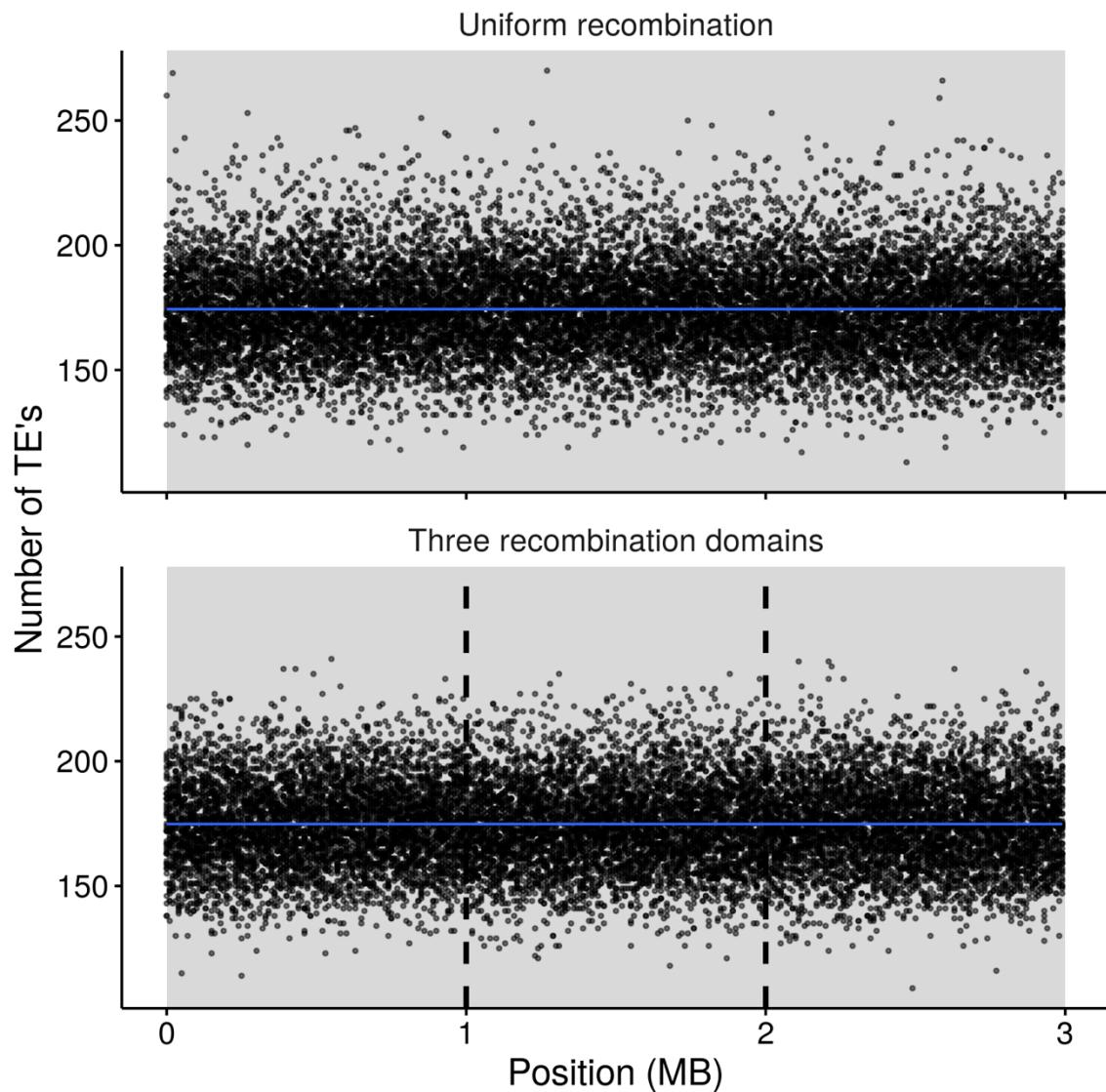
Supplemental figure 28. TE genomic landscapes in simulations with fixed fitness effects of TE insertion and three genomic domains of recombination. Plotted are the total number of transposable element sites in 10kb non-overlapping genomic windows along the chromosome after 50,000 generations. Here, there are three chromosomal domains of recombination (boundaries denoted by the dashed lines; chromosome arms $r=5\times10^{-7}$; chromosome centers $r=1\times10^{-9}$). Blue lines were fit by LOESS regression. Top panel ($N_s=0$) uses the same simulations from Figure 8a, right panel.



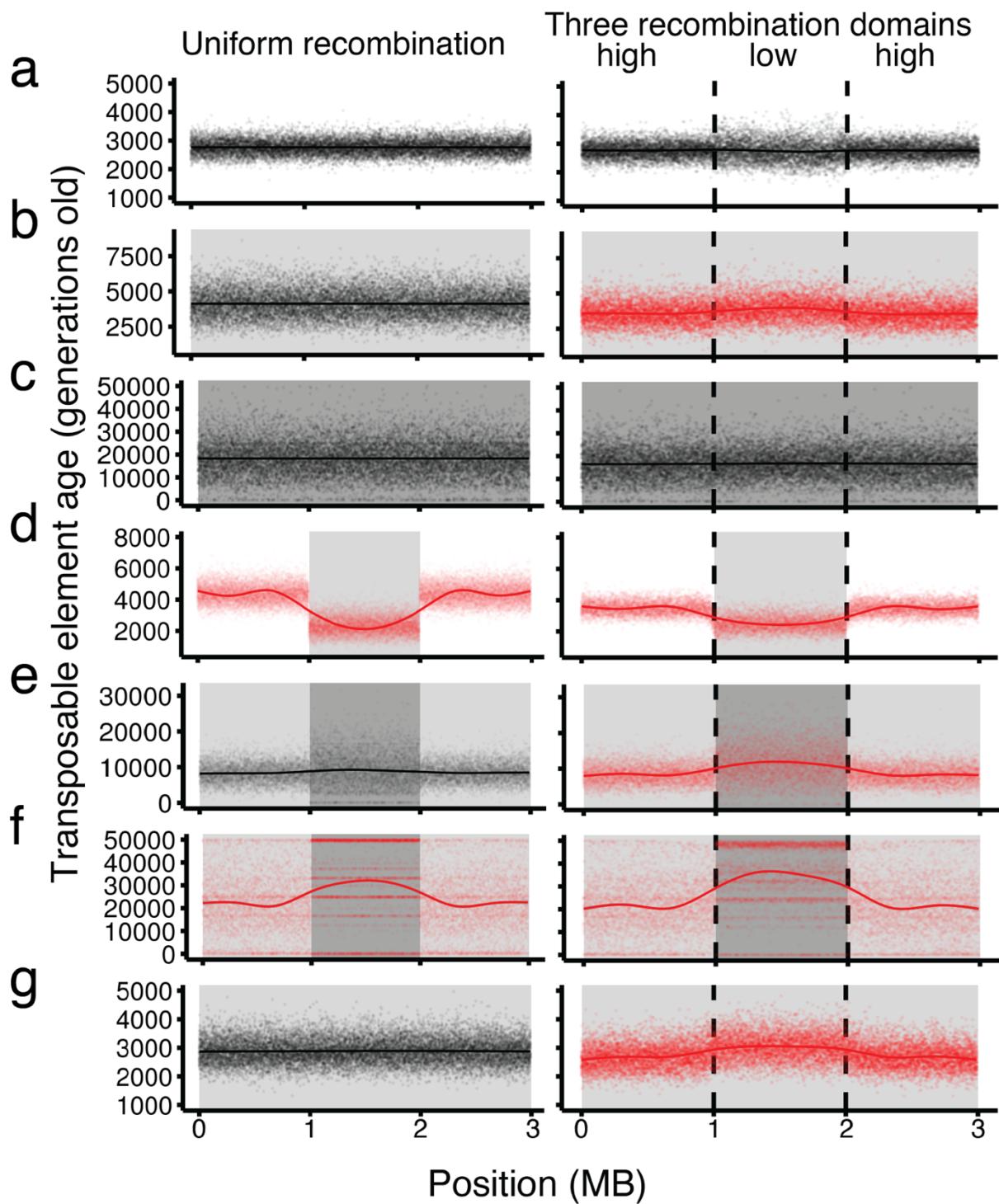
Supplemental figure 29. The impact of fitness effect of TE insertion on the genomic distribution of TE's in evolutionary simulations. Same simulations as Supplemental Figure 28. Each dot represents the Cohen's d effect size and fixed TE insertion effect of one population after 50,000 generations. The solid line was fit by LOESS regression; the gray shaded area represents the 95% confidence interval. Solid horizontal lines denote the global genomic chromosome arm-center Cohen's d effect size in two *Caenorhabditis* species.



Supplemental figure 30. TE genomic landscapes in simulations with selfing. TE insertions are highly deleterious in chromosome centers while weakly deleterious in arms. Plotted are the total number of transposable element sites in 10kb non-overlapping genomic windows along the chromosome after 50,000 generations. Each point represents the total number of transposable element sites in a population in that window. Top, uniformly high rates of recombination along the chromosome ($r= 5 \times 10^{-7}$). Bottom, three chromosomal domains of recombination (boundaries denoted by the dashed lines; chromosome arms $r= 5 \times 10^{-7}$; chromosome centers $r= 1 \times 10^{-9}$). Gray boxes represent chromosomal regions with deleterious fitness effects of transposable element insertion (light gray, mean $s= -0.0006$; dark gray, mean $s= -0.03$). Fitness effects of insertion were drawn from a gamma distribution. Blue lines were fit by LOESS regression.



Supplemental figure 31. TE genomic landscapes in simulations with selfing. TE insertions are weakly deleterious across the entire genome. Plotted are the total number of transposable element sites in 10kb non-overlapping genomic windows along the chromosome after 50,000 generations. Each point represents the total number of transposable element sites in a population in that window. Top, uniformly high rates of recombination along the chromosome ($r= 5 \times 10^{-7}$). Bottom, three chromosomal domains of recombination (boundaries denoted by the dashed lines; chromosome arms $r= 5 \times 10^{-7}$; chromosome centers $r= 1 \times 10^{-9}$). Gray boxes represent chromosomal regions with deleterious fitness effects of transposable element insertion (light gray, mean $s= -0.0006$). Fitness effects of insertion were drawn from a gamma distribution. Blue lines were fit by LOESS regression.

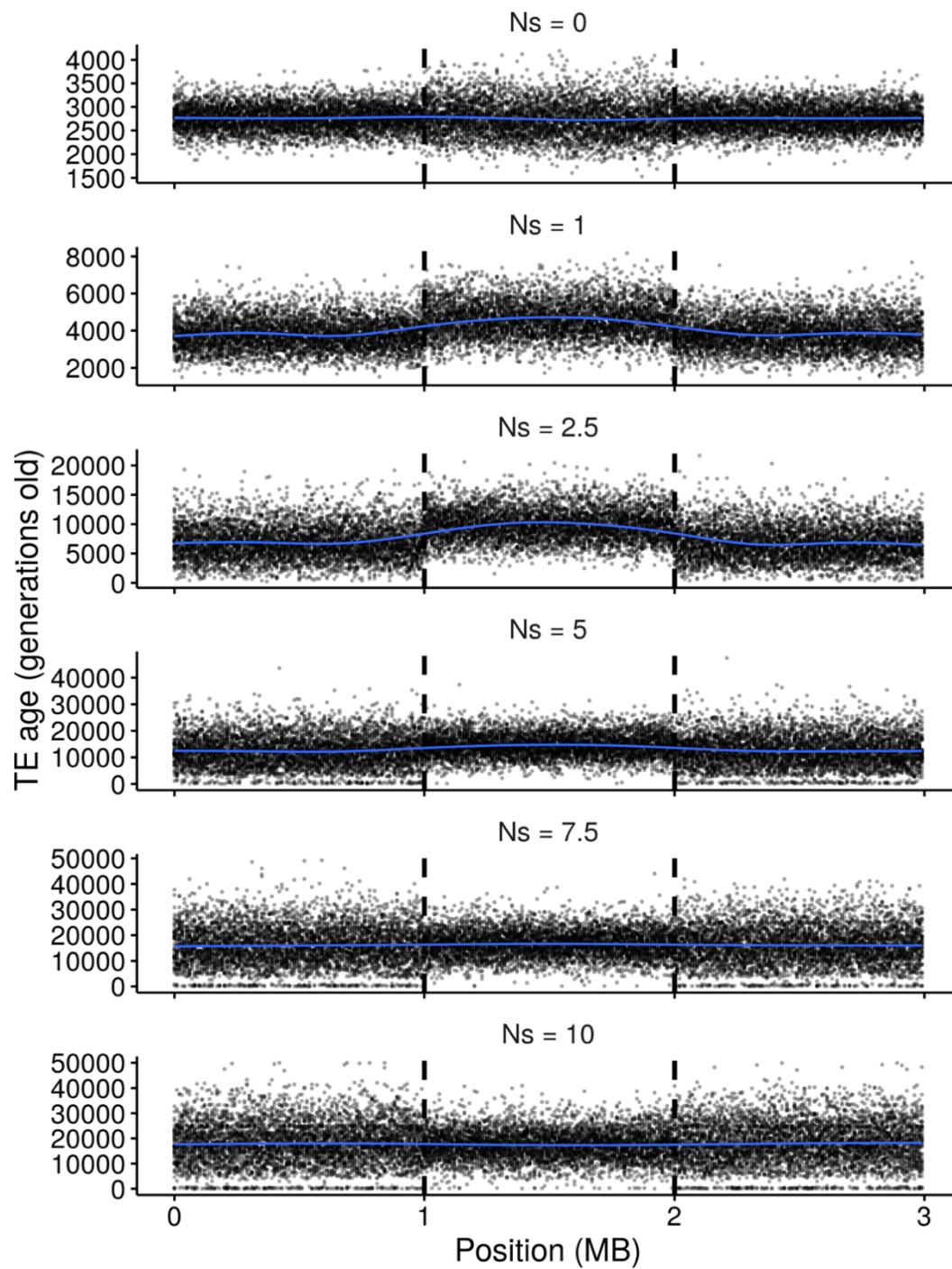


Supplemental figure 32. Ages of TE's in simulations. Plotted are mean age (in generations) of TE insertions in 10kb non-overlapping genomic windows along the chromosome after 50,000 generations among 50 simulations per scenario. Same simulations as in Figure 8. The left column shows scenarios where there were uniformly high rates of recombination along the chromosome ($r = 5 \times 10^{-7}$); the right column shows scenarios where there were three chromosomal domains of recombination (boundaries denoted by the dashed lines; chromosome arms $r = 5 \times 10^{-7}$; chromosome centers $r = 1 \times 10^{-9}$). Gray boxes

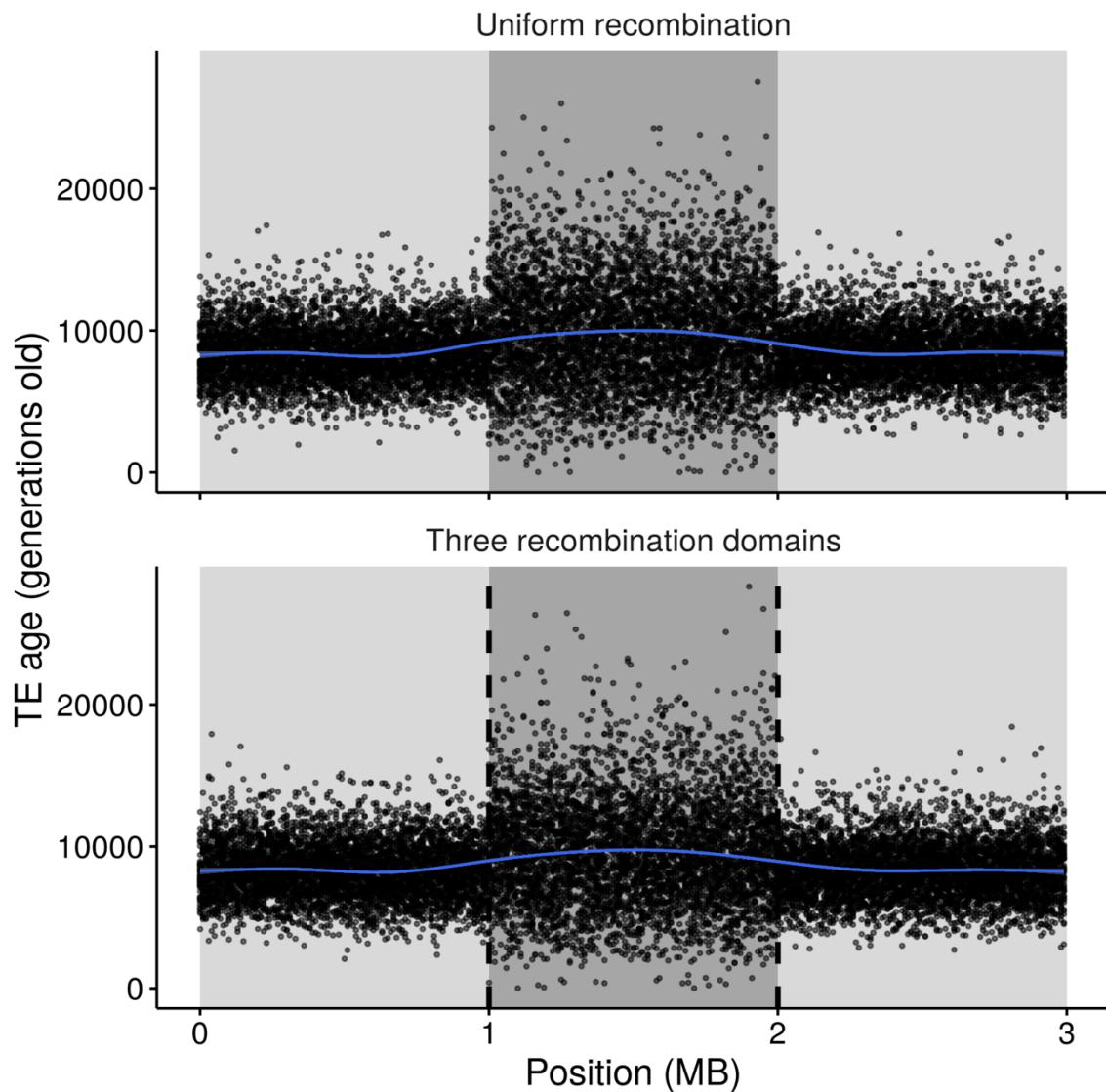
represent chromosomal regions with deleterious fitness effects of transposable element insertion (light gray, mean $s = -0.0006$; dark gray, mean $s = -0.03$; white, $s = 0$ for all transposable element insertions). When not neutral, fitness effects of insertion were drawn from a gamma distribution. Red lines represent scenarios with *TE count* arm-center Cohen's d effect sizes > 0.2 ; black lines show scenarios with no or negligible chromosomal organization of repeats (arm-center Cohen's d effect sizes < 0.2). All scenarios implemented "copy-and-paste" models of transposable element replication with the exception of (g), which used a "cut-and-paste" model.

Brief descriptions of each scenario are described below (see methods for details):

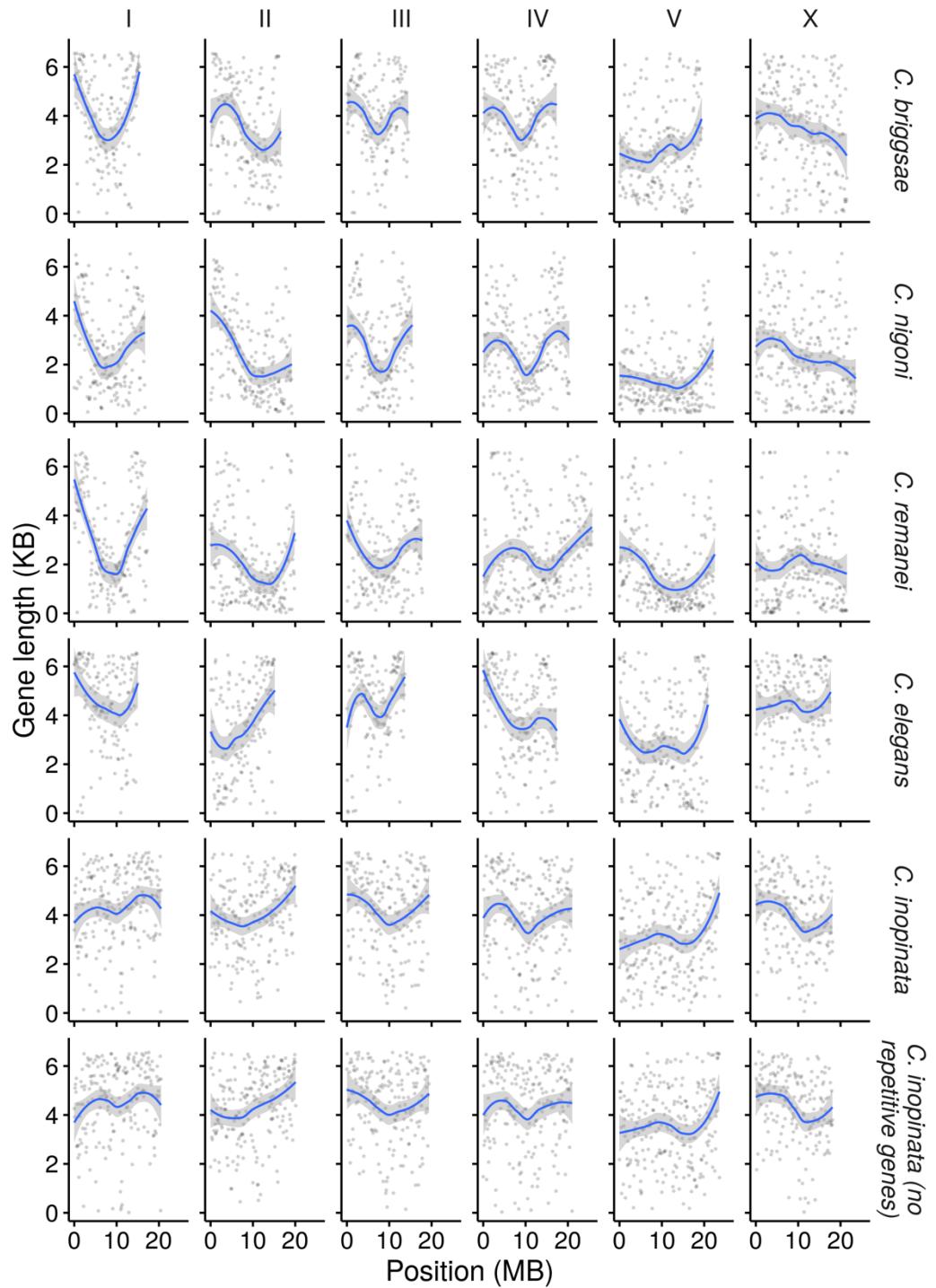
- a. All TE insertions neutral (these populations are not at equilibrium, Supplemental Figure 37).
- b. All TE insertions weakly deleterious.
- c. All TE insertions highly deleterious.
- d. TE insertions weakly deleterious in center; TE insertions neutral in arms.
- e. TE insertions highly deleterious in center; TE insertions weakly deleterious in arms.
- f. TE insertions highly deleterious in center; TE insertions weakly deleterious in arms; "cut-and-paste" mode of TE replication.
- g. All TE insertions weakly deleterious; beneficial mutations introduced ($u = 1 \times 10^{-9}$).



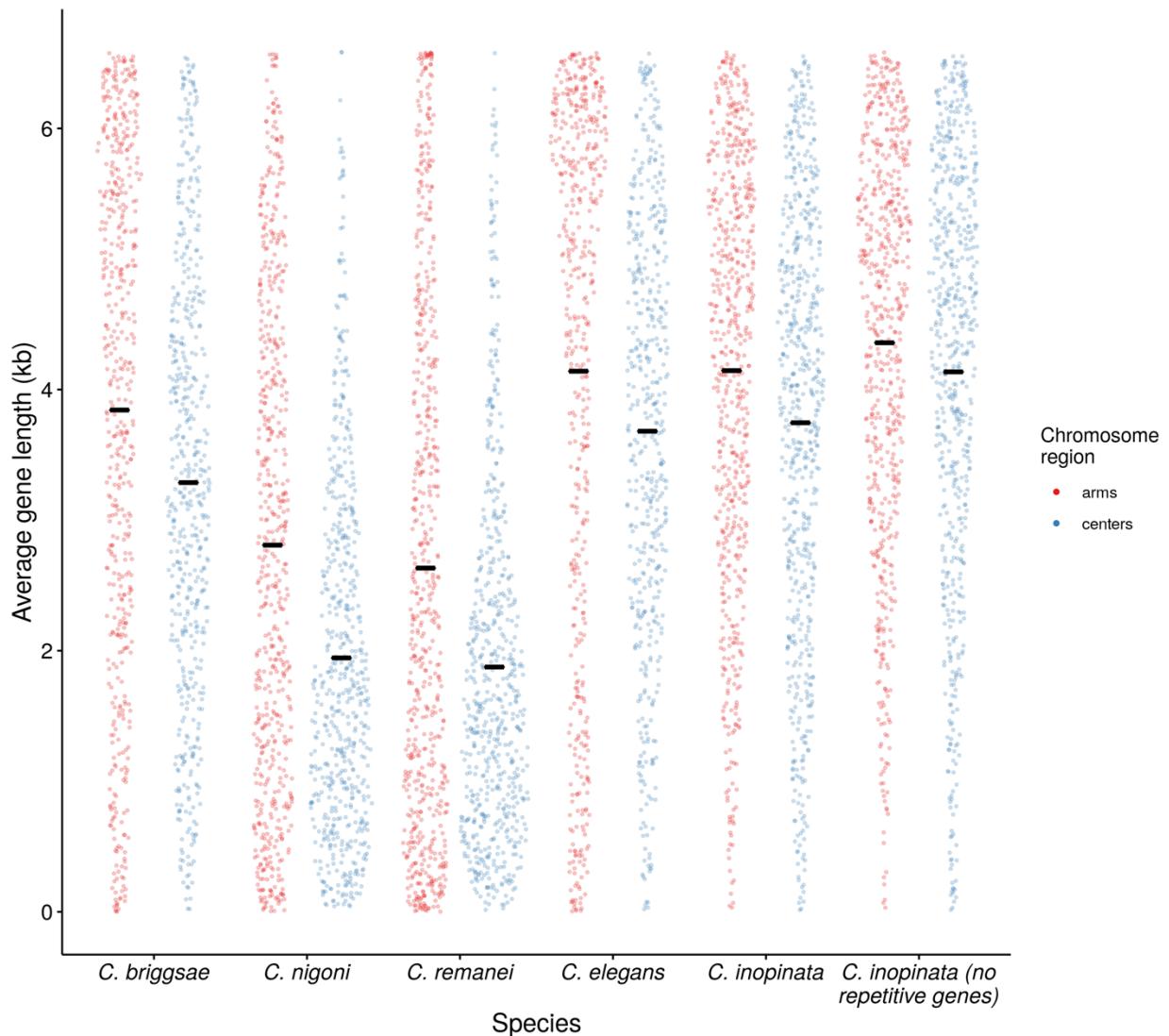
Supplemental figure 33. Genomic landscapes of TE age in simulations with fixed fitness effects of TE insertion and three genomic domains of recombination. Same simulations as in Supplemental Figures 28-29. Plotted are mean age (in generations) of TE insertions in 10kb non-overlapping genomic windows along the chromosome after 50,000 generations among 50 simulations per scenario. Here, there are three chromosomal domains of recombination (boundaries denoted by the dashed lines; chromosome arms $r=5 \times 10^{-7}$; chromosome centers $r=1 \times 10^{-9}$). Top panel ($N_s=0$) uses the same simulations from Figure 8a, right panel. Blue lines were fit by LOESS regression.



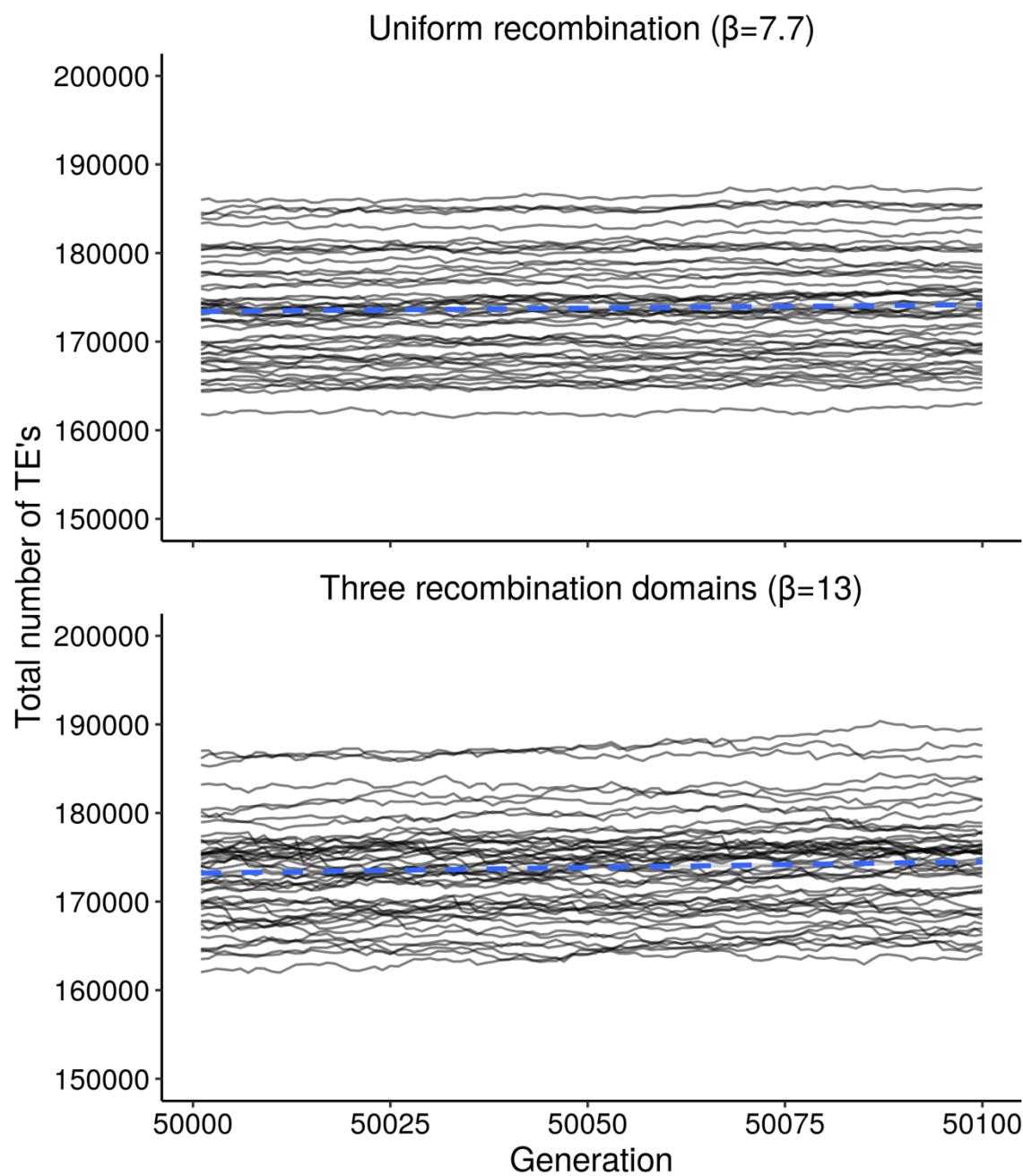
Supplemental figure 34. Genomic landscapes of TE age in simulations with selfing. TE insertions are highly deleterious in chromosome centers while weakly deleterious in arms. Plotted are mean age (in generations) of TE insertions in 10kb non-overlapping genomic windows along the chromosome after 50,000 generations among 50 simulations per scenario. Top, uniformly high rates of recombination along the chromosome ($r= 5 \times 10^{-7}$). Bottom, three chromosomal domains of recombination (boundaries denoted by the dashed lines; chromosome arms $r= 5 \times 10^{-7}$; chromosome centers $r= 1 \times 10^{-9}$). Gray boxes represent chromosomal regions with deleterious fitness effects of transposable element insertion (light gray, mean $s= -0.0006$). Fitness effects of insertion were drawn from a gamma distribution. Blue lines were fit by LOESS regression.



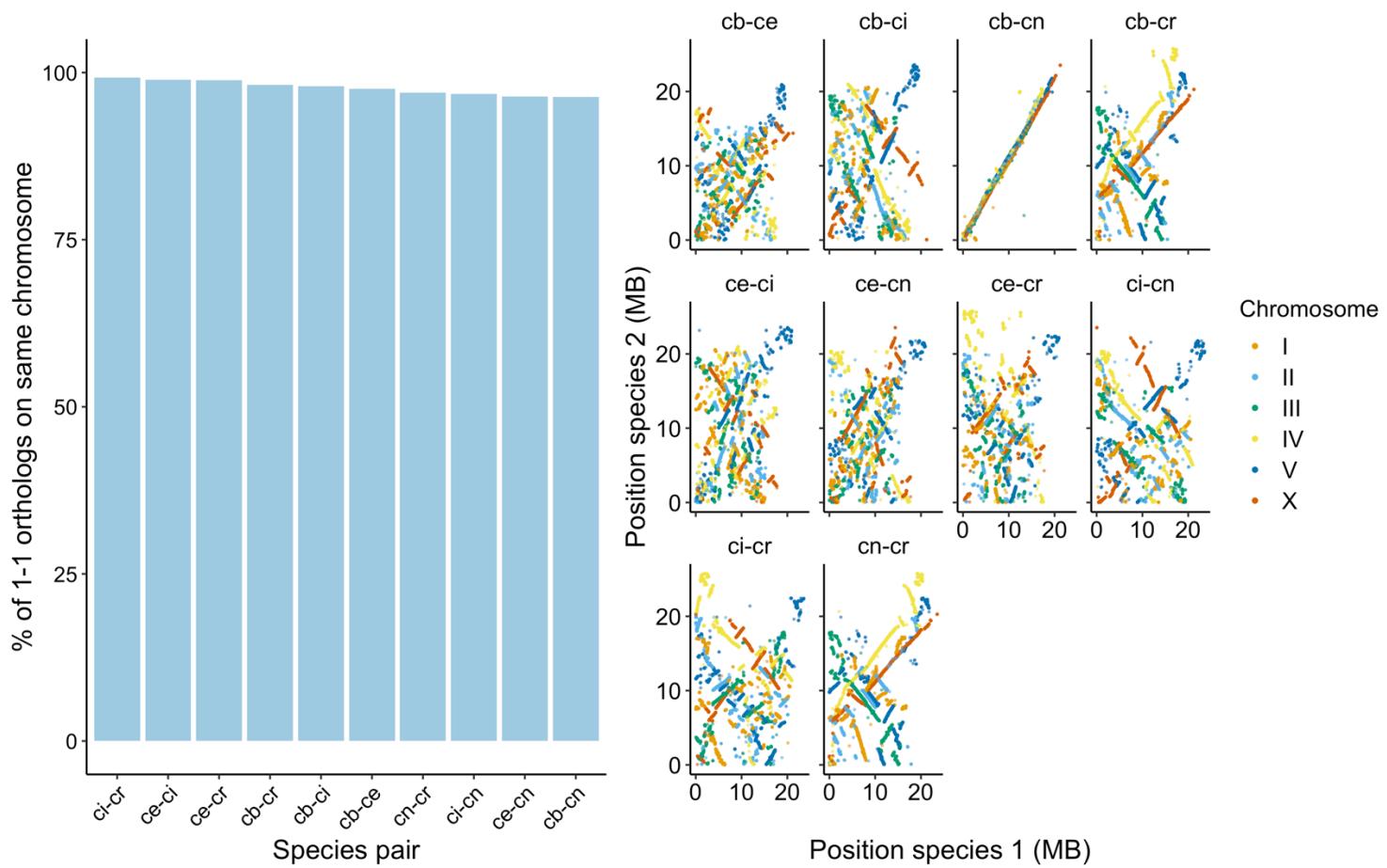
Supplemental figure 35. Gene length among 100kb genomic windows in five *Caenorhabditis* species. Plotted is the average gene length in that window. Here, the genes of *C. inopinata* when 2,489 transposon-aligning genes are excluded are also plotted. The blue line was fit by LOESS local regression; the shaded ribbon is the 95% confidence interval.



Supplemental figure 36. Distribution of gene lengths on chromosome arms and centers. Each point represents one 100 kb genomic window and the average gene length in that window. After normalizing chromosome positions by distance to chromosome midpoint, windows were classified as being in chromosome “centers” (middle half of chromosome) or “arms” (outer half of chromosome). Sina plots (strip charts with points taking the contours of a violin plot) reveal the distribution of repeat densities of these windows in chromosome centers and arms for all species. Here, the gene densities of *C. inopinata* when 2,489 transposon-aligning genes are excluded are also plotted (last position on x-axis). Black horizontal lines, means.



Supplemental figure 37. Simulated populations with neutral TE insertions are not at an equilibrium state and slowly accumulate TE's over time. Each gray line tracks the total TE count of a population in one simulation (50 simulations per scenario). Top, uniformly high rates of recombination along the chromosome ($r = 5 \times 10^{-7}$). Bottom, three chromosomal domains of recombination (chromosome arms $r = 5 \times 10^{-7}$; chromosome centers $r = 1 \times 10^{-9}$). Blue dotted line fit by ordinary least squares regression (uniform recombination: $r^2 = 0.0011$; $F = 6.4$; $\beta_1 = 7.7$; $p = 0.011$; three recombination domains: $r^2 = 0.0040$; $F = 21$; $\beta_1 = 13$; $p = 5.1 \times 10^{-6}$); slope (β) noted above plot.



Supplemental figure 38. *Caenorhabditis* single-copy orthologs reveal high between-chromosome synteny with widespread rearrangements within chromosomes. Left panel, pairwise species comparisons reveal most single-copy orthologous genes ($n=8,553$) to be on the same chromosome in both species. Right panel, dot plots of single-copy ortholog position in all pairwise species comparisons. Only orthologs that have remained on the same chromosome are plotted. With the exception of *C. nigoni*-*C. briggsae*, who are sister species capable of hybridization, all comparisons reveal widespread translocations and inversions within chromosomes. Each point represents the positions of one single-copy ortholog in the species under comparison. cb, *C. briggsae*. ce, *C. elegans*. ci, *C. inopinata*. ch, *C. nigoni*. cr, *C. remanei*.