SUPPORTING INFORMATION for:

The Opportunity for Balancing Selection in Experimental Populations of *Caenorhabditis elegans*

Ivo M. Chelo and Henrique Teotónio

Supporting Information is composed of 9 Figures (**S1-S9**), and three Tables (**S1-S3**) in Microsoft Excel format. This file contains the Figures.

The Tables have the following titles:

Table S1. SNP information and genotype sample sizes.

Table S3. Population sample correlations of missing data with individual heterozygosity.

Table S3. Evolution of linkage disequilibrium and expected self-fertilization rates.

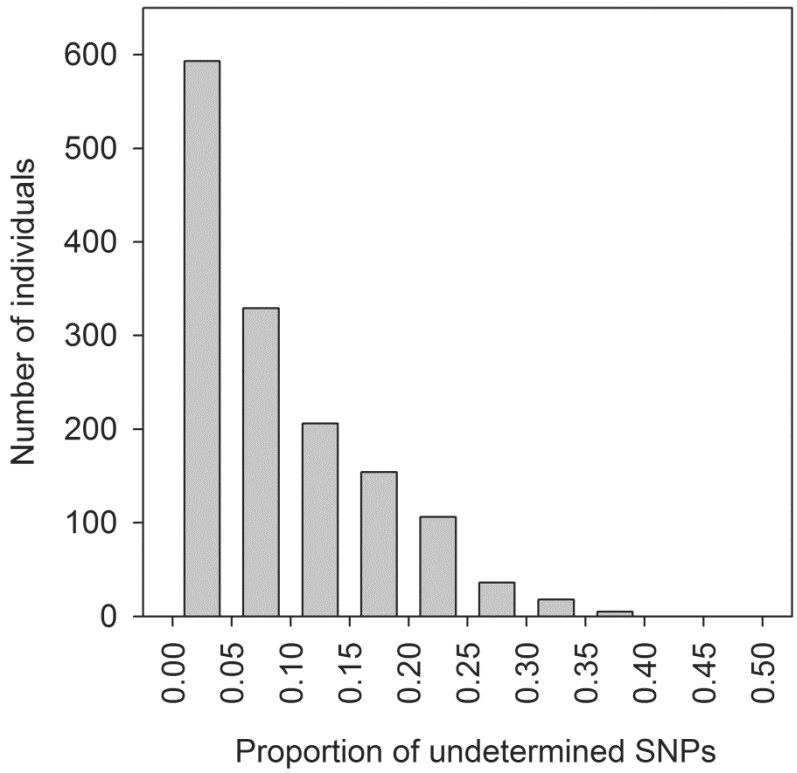

Figure S1. Distribution of missing data after quality control. The number of individuals that constituted the data set for analysis is shown with the proportion of missing SNP data (see also **Table S2** for within population sample correlations between missing data with individual heterozygosity).

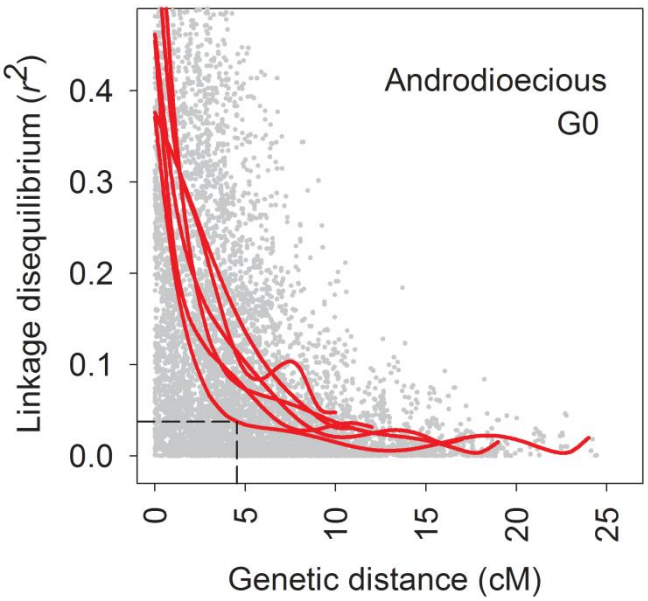

Figure S2. Estimates of linkage disequilibrium by chromosomal recombination rate domains. Pair-wise SNP r^2 is plotted with genetic distance separating SNPs for the ancestral androdioecious population. Each chromosome, excluding tips, can be divided into 3 domains of constant meiotic crossover rates over physical distance; see reference Rockman and Kruglyak (2009). Polynomial functions were found for each one of them (red lines). Dashed lines indicate the background r^2 and the genetic distance at which it is reached at one of the chromosomal recombination domains. See **Table S3** for results under both androdioecy and dioecy during experimental evolution.

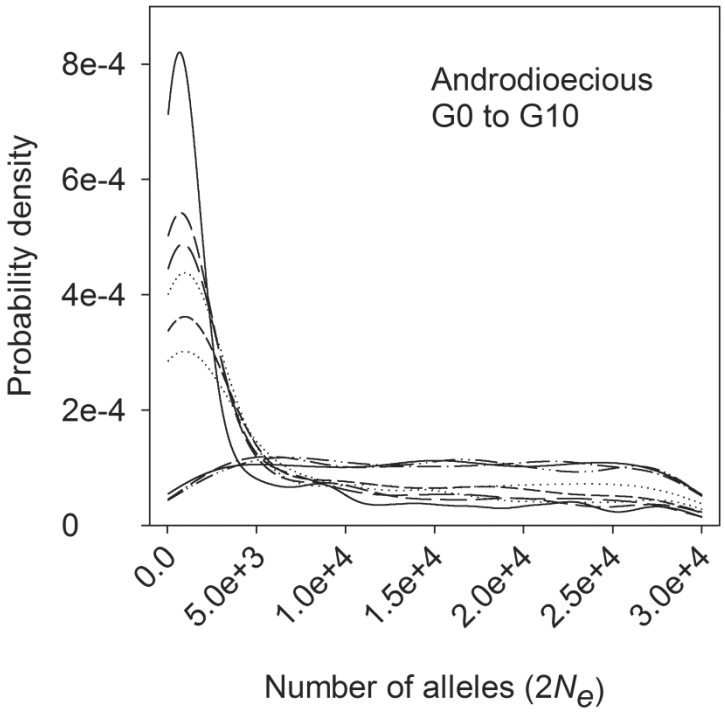

Figure S3. Estimation of effective population sizes. An example of N_e estimation is shown for androdioecious populations between G0 and G10. In chromosome IV, 10 regions were found not to recombine during this period. Lines show the probability density distributions of explaining allele frequency changes from G0 to G10 at each "non-recombining" loci, as obtained with the multinomial sampling of $2N_e$ haplotypes.

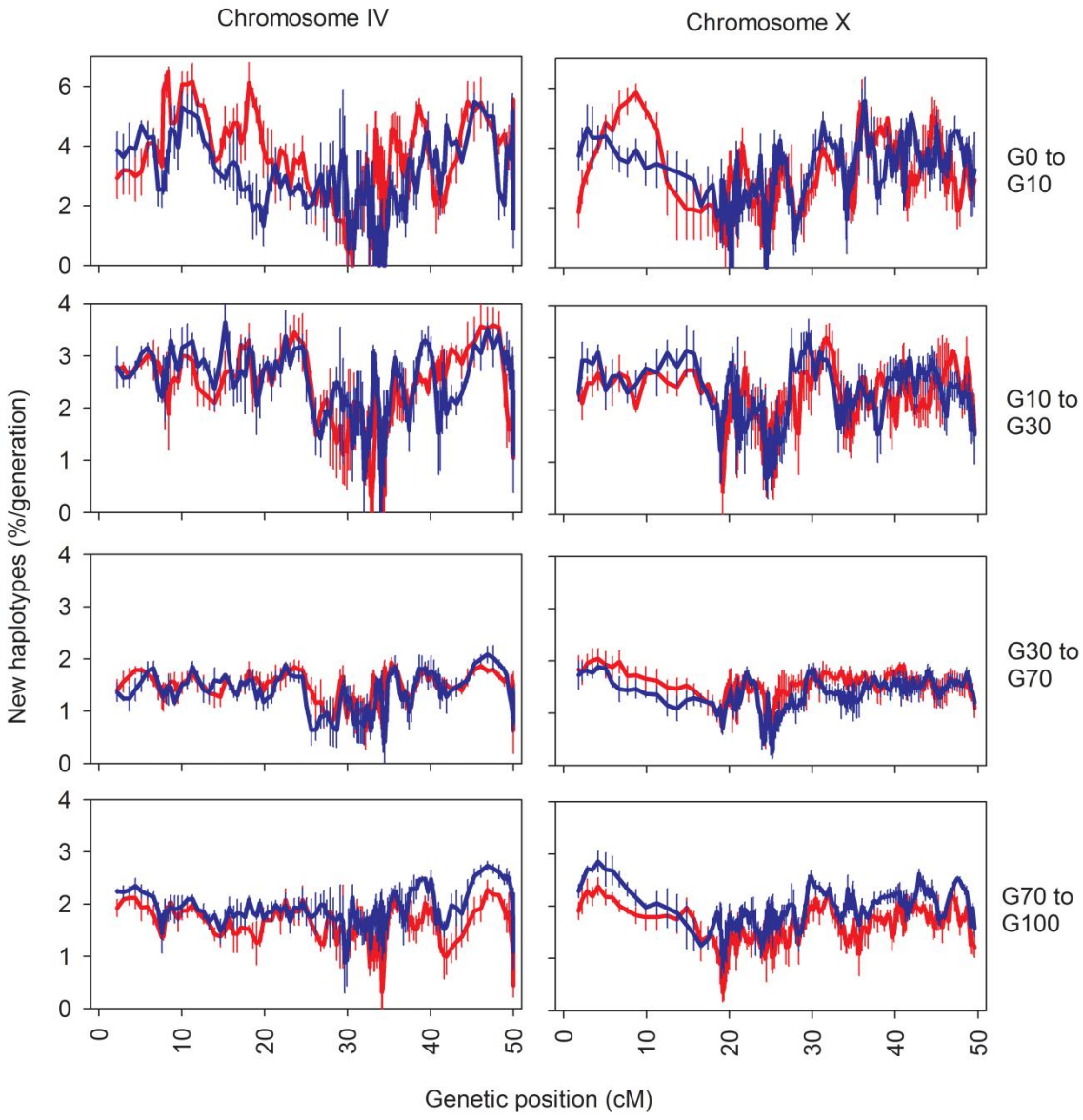

Figure S4. Evolution of effective recombination. Panels show, for each period of evolution, the number of new haplotypes per generation as the percentage of the total number of segregating haplotypes. Average % and one SEM are shown in red for androdioecious populations and blue for dioecious populations. Calculations were performed in window sizes of 10 SNPs and step sizes of 1 SNP along genetic position. Note the different scaling from G0 to G10 relative to other periods.

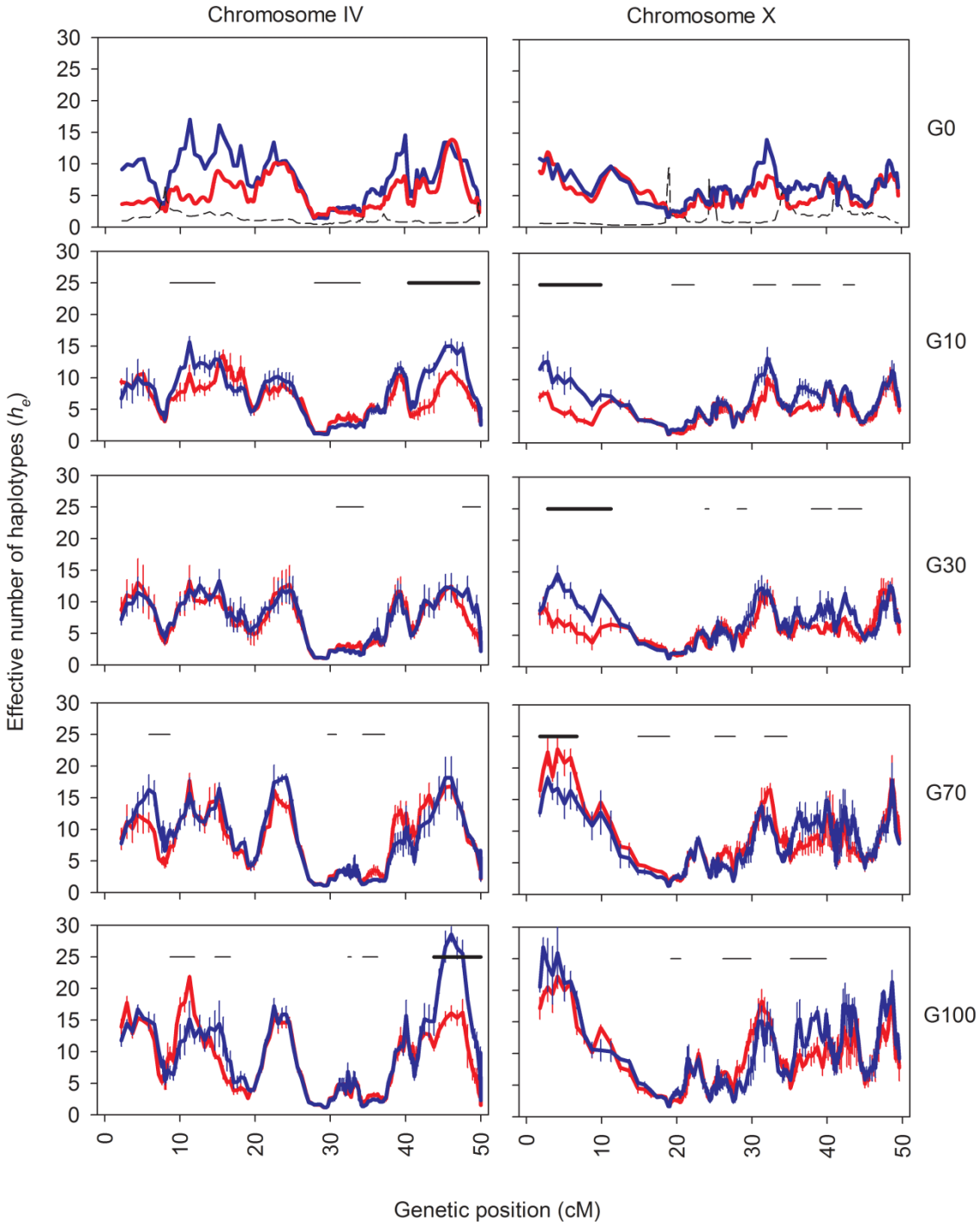

Figure S5. Evolution of effective haplotype numbers. Panels show h_e at each generation as the average and SEM in red for androdioecious populations and blue for dioecious populations. Calculations were performed in window sizes of 10 SNPs and step sizes of 1 SNP along genetic position. Bars indicate regions with significant differentiation in h_e between mating systems. Thick bars indicate the regions which are mentioned in the main text. Dashed lines in top panels indicate the density of SNPs per 100kb.

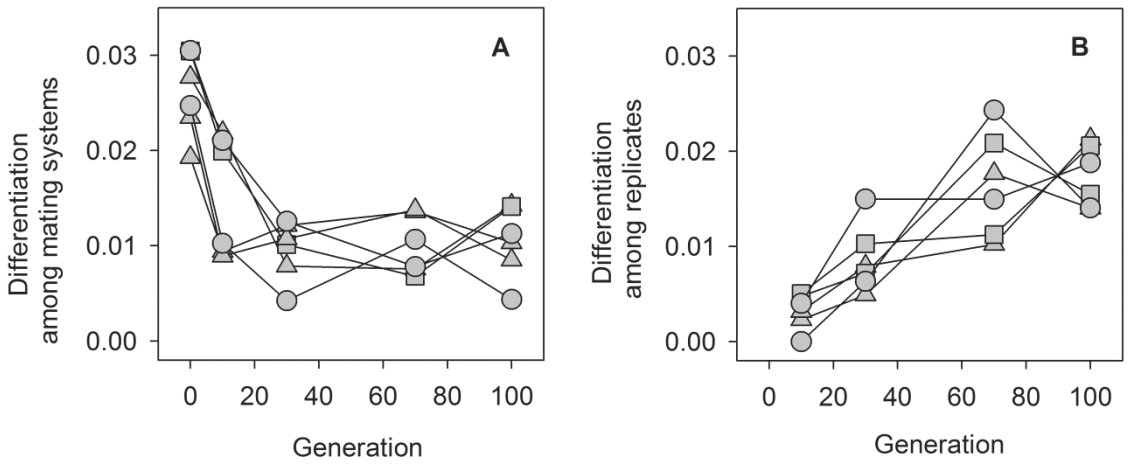

Figure S6. Genetic differentiation among mating systems and replicate populations. In panel **A** is shown the component of variance in allele frequencies due to the mating system or in panel **B** the component due to replicate populations. Alleles were defined by shared haplotypes among all populations independently at each generation, as computed for the average number of haplotypes in both chromosomes at windows of 10 (circles), 5 (squares) or 2 SNPs (triangles).

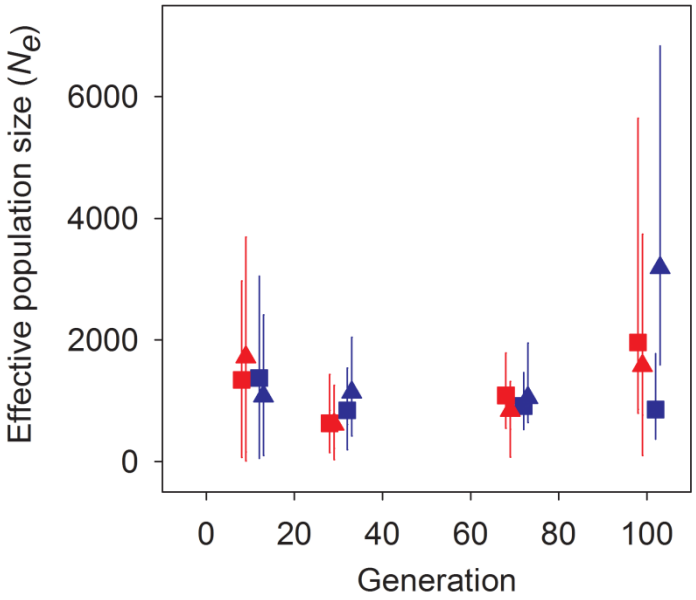

Figure S7. Effective population sizes during evolution. This plot shows the N_e estimated for chromosome IV (squares) and chromosome X (triangles), in red for androdioecy and blue for dioecy. Symbols indicate the maximum of the probability density distribution of sampling $2N_e$ haplotypes within each chromosome (see **Figure S3** above) with error bars indicating ± 1 $\log_{10}$.

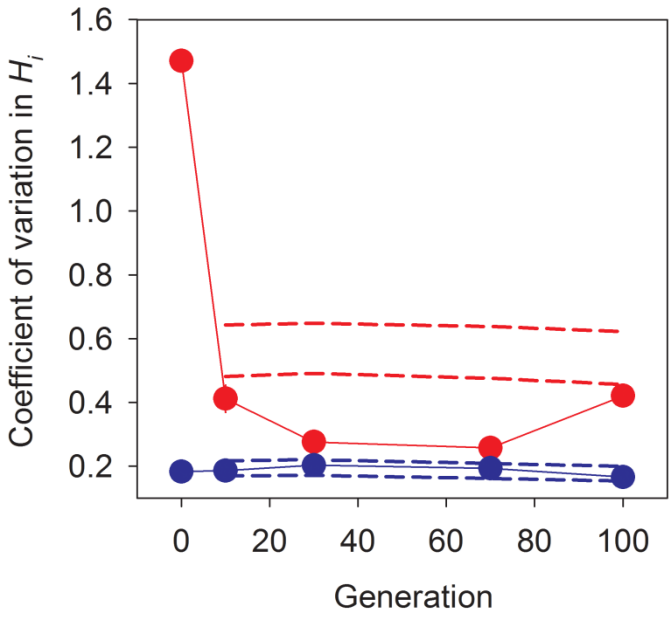

Figure S8. Evolution of the variance in individual heterozygosity. As in the main **Figure 4**, the evolution of CV(H_i) is shown in red for androdioecious populations and blue for dioecious populations with one SEM. Dashed lines limit 95% of the data obtained from 10^3 neutral simulations of evolution at each period.

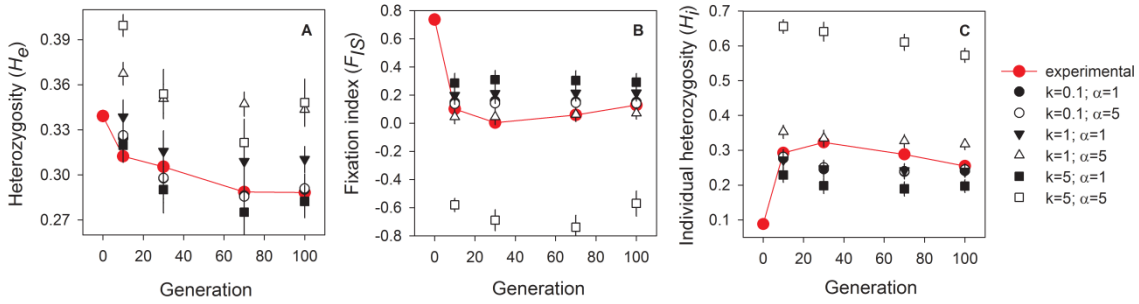

Figure S9. Expected population structure with selection at overdominant loci, under androdioecy. In panels **A**, **B** and **C**, illustrative results for H_e, F_{IS} and H_i obtained from simulations with selection on overdominant loci at each period of evolution. The mean and two SD of 3x100 simulations for each period are presented for each parameter combination of epistasis (k) and selection (α) coefficients. Experimental results are presented with one SEM (as in **Figure 4**).