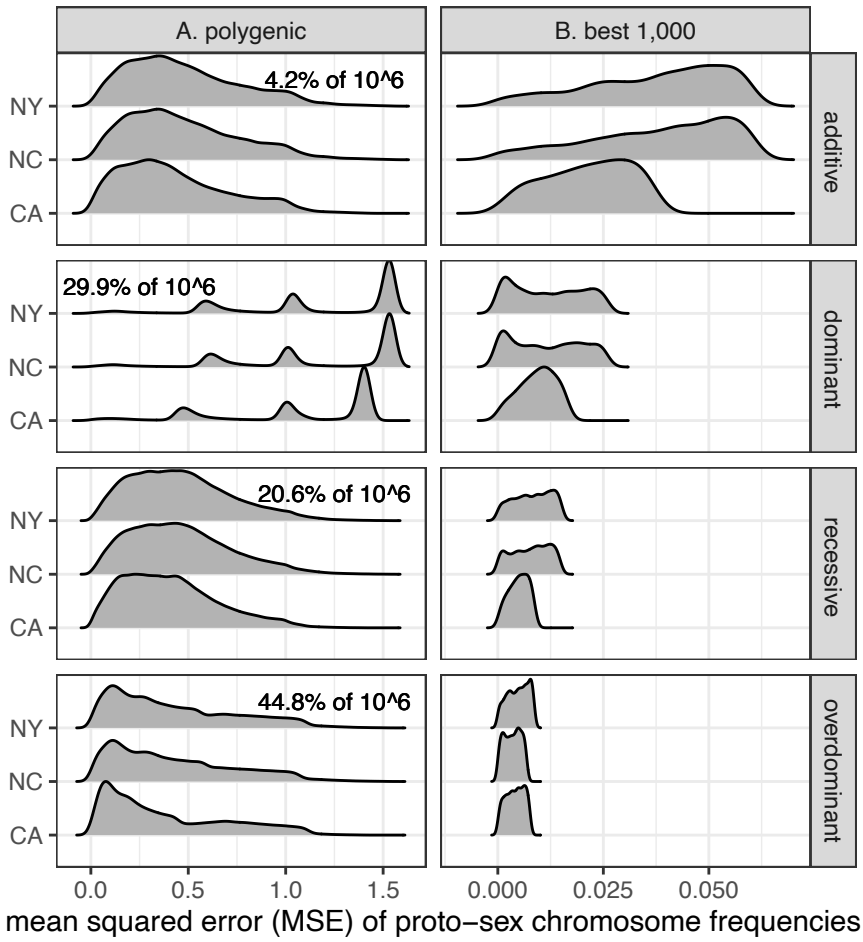
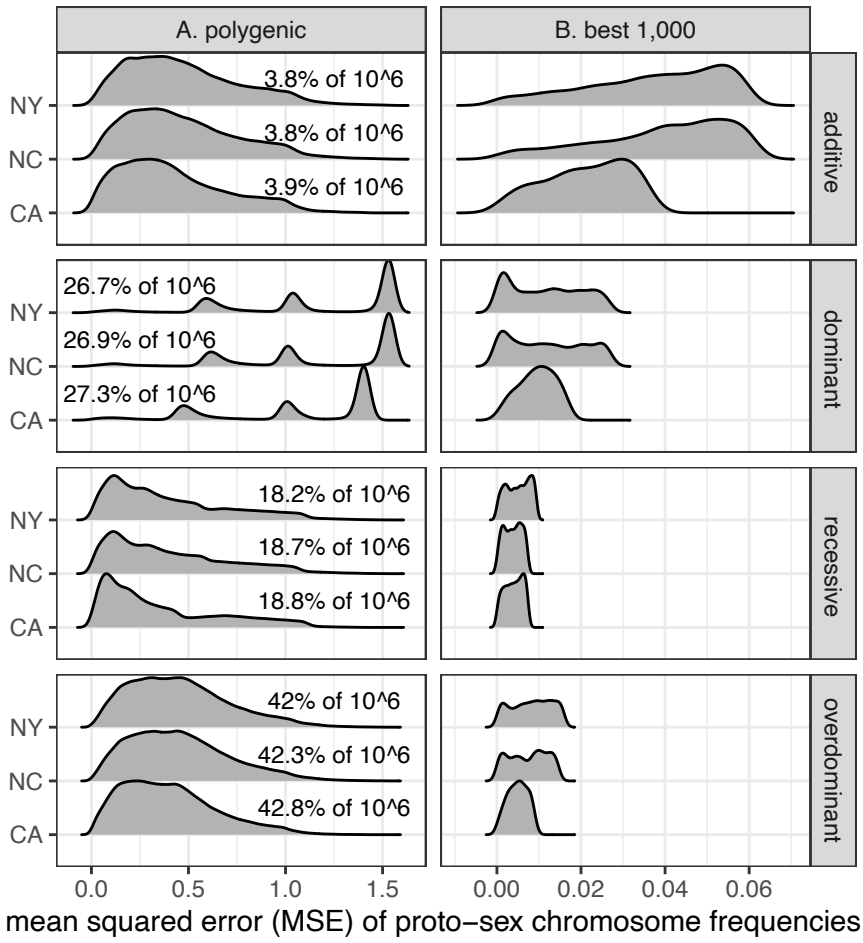
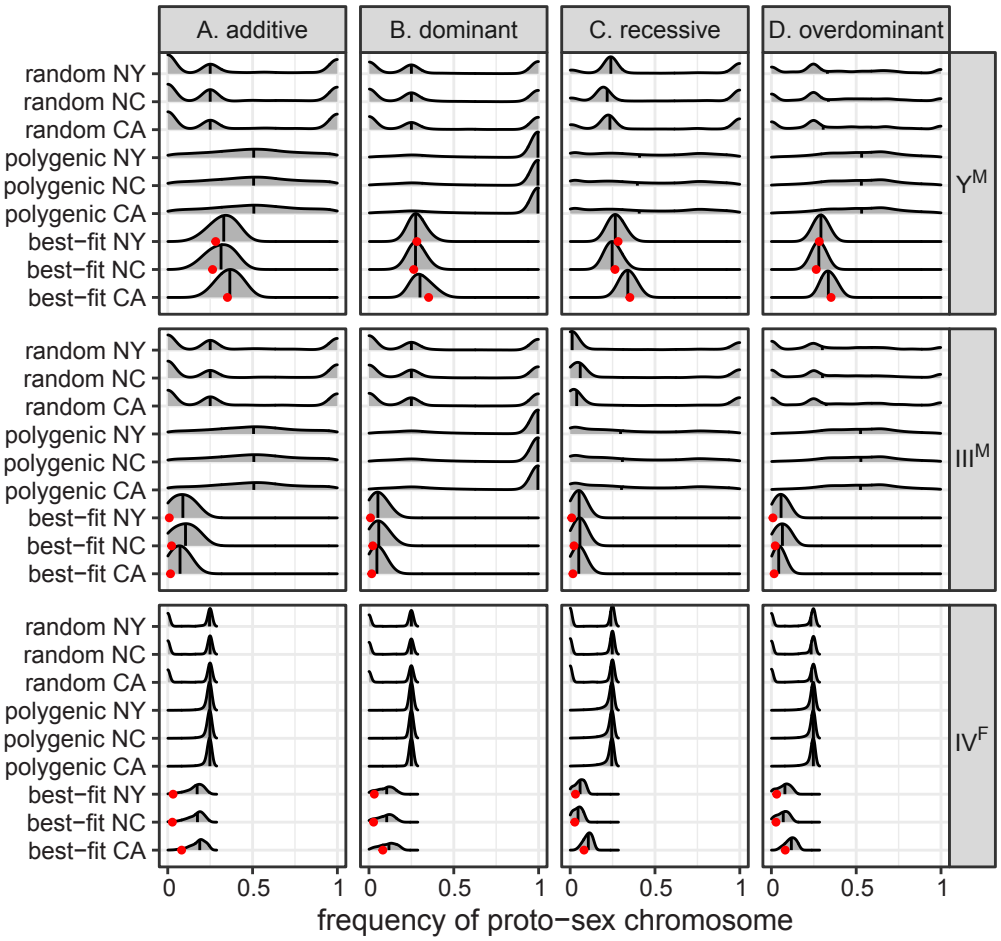


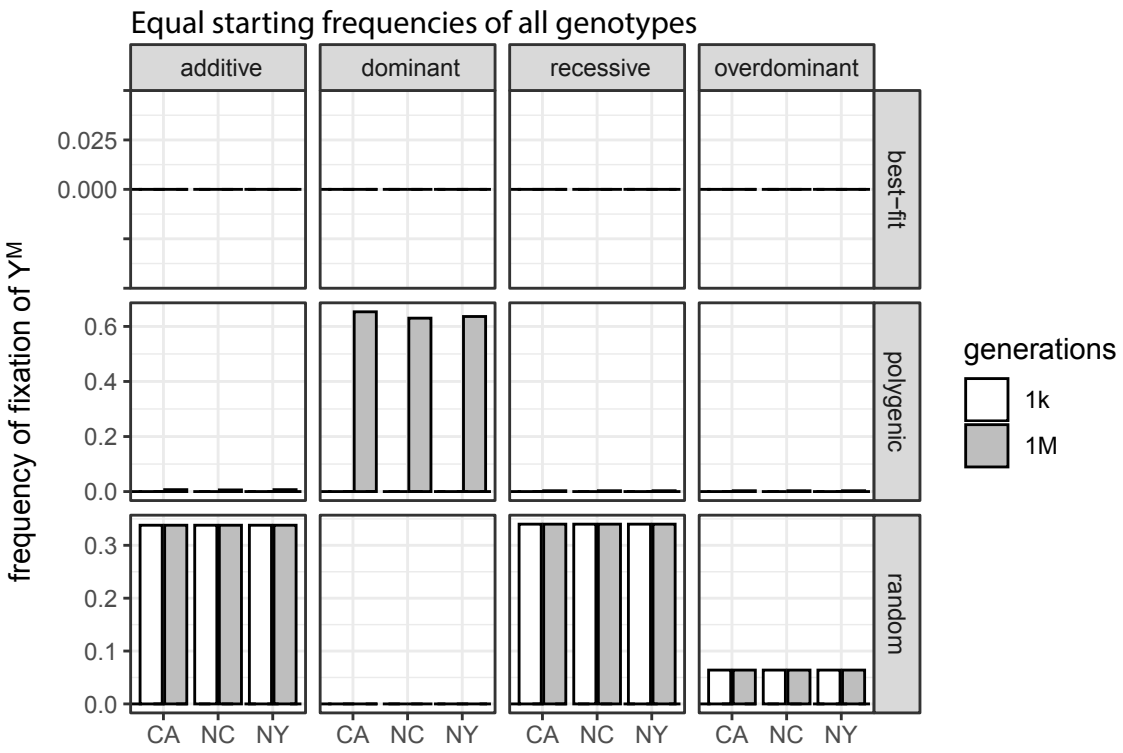
Supplemental Figure S1. Smoothed histograms show the distributions of mean squared error (MSE) values for fitness arrays that maintain PSD (i.e., all proto-sex chromosomes at a frequency >0.1%). MSE is calculated based on the frequencies of Y^M, III^M, and IV^F found in either CA, NC, or NY. Simulations were started with equal frequencies of all genotypes. Fitness arrays are based on additive, dominant, recessive, or overdominant fitness effects of Y^M and III^M. **(A)** Distributions of MSE are shown for all fitness arrays that maintain PSD, with the percent of those arrays that maintain PSD shown in each panel. **(B)** Distributions of MSE are shown for the 1,000 best-fitting arrays with the lowest MSE.



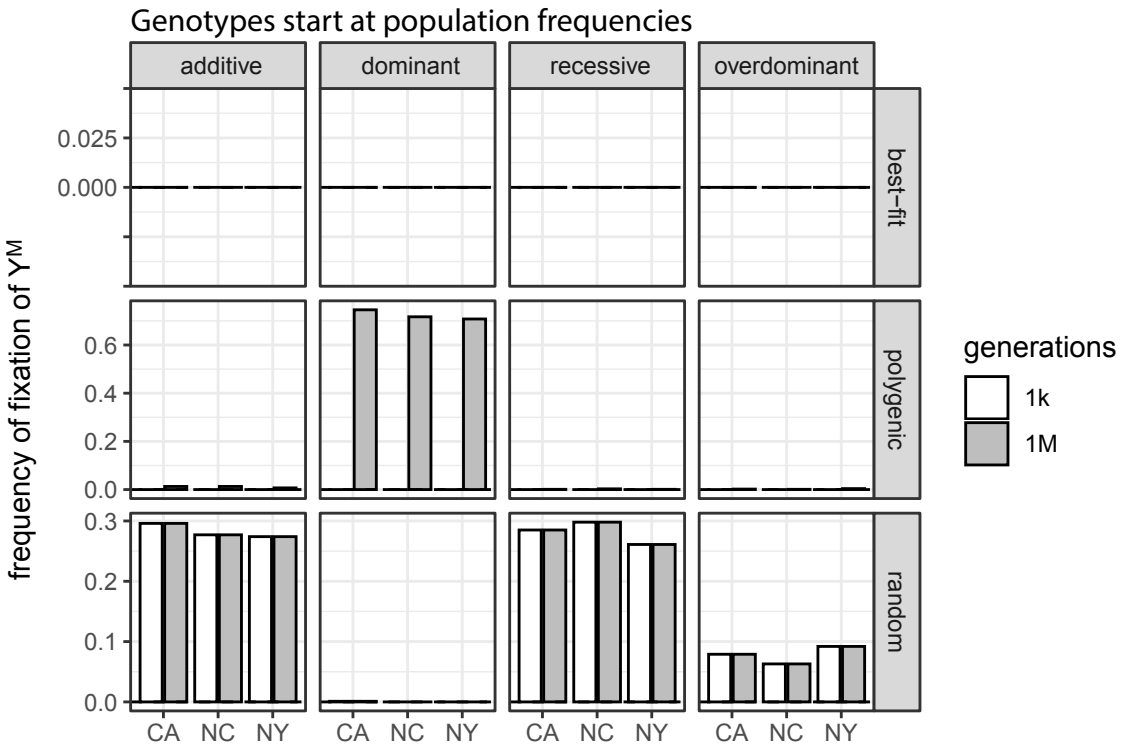
Supplemental Figure S2. Smoothed histograms show the distributions of mean squared error (MSE) values for fitness arrays that maintain PSD (i.e., all proto-sex chromosomes at a frequency >0.1%). MSE is calculated based on the frequencies of Y^M, III^M, and IV^F found in either CA, NC, or NY. Simulations were started with genotype frequencies observed in natural populations. Fitness arrays are based on additive, dominant, recessive, or overdominant fitness effects of Y^M and III^M. **(A)** Distributions of MSE are shown for all fitness arrays that maintain PSD, with the percent of those arrays that maintain PSD shown in each panel. **(B)** Distributions of MSE are shown for the 1,000 best-fitting arrays with the lowest MSE.



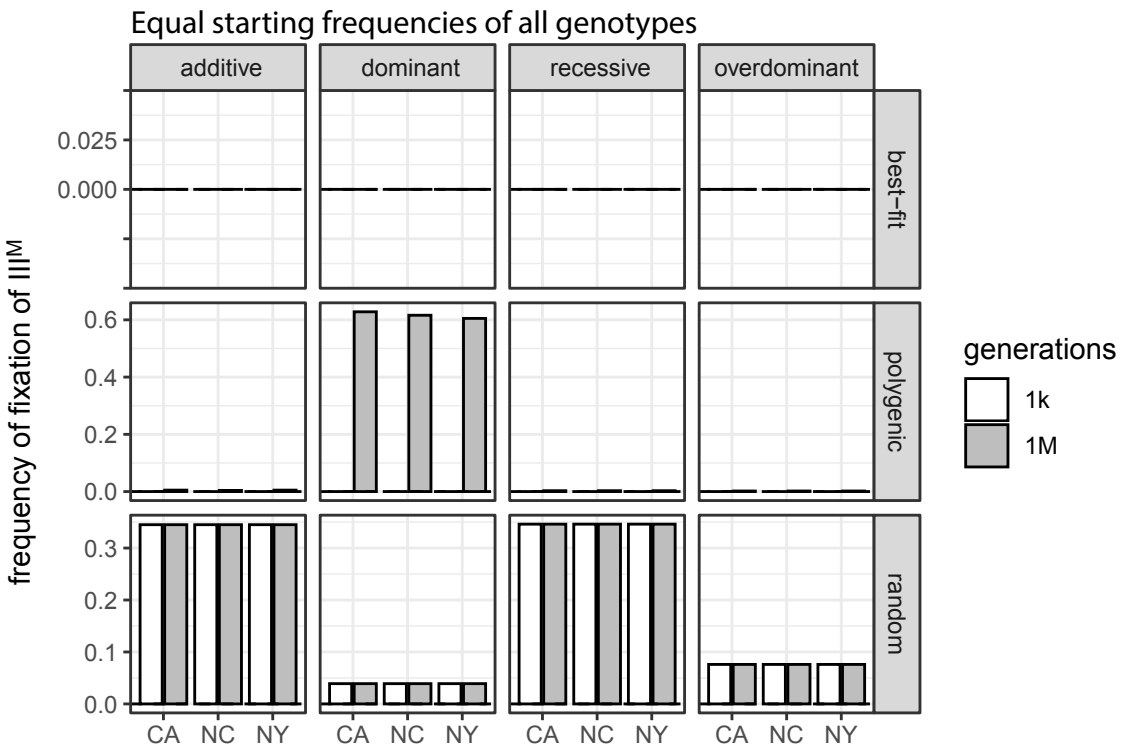
Supplemental Figure S3. Smoothed histograms show the frequency of each proto-sex chromosome (Y^M, III^M, and IV^F) after 1,000 generations in simulations using either 1,000 random fitness arrays (random), fitness arrays that maintain PSD (polygenic), or the 1,000 best-fitting fitness arrays for each population (CA, NC, or NY). Simulations were started with genotype frequencies observed in each population. The vertical line within each histogram shows the median. Red dots show the observed proto-sex chromosome frequencies in each natural population. Fitness arrays were calculated assuming either (**A)** additive, **(B)** dominant, **(C)** recessive, or **(D)** overdominant fitness effects of the proto-Y chromosomes.



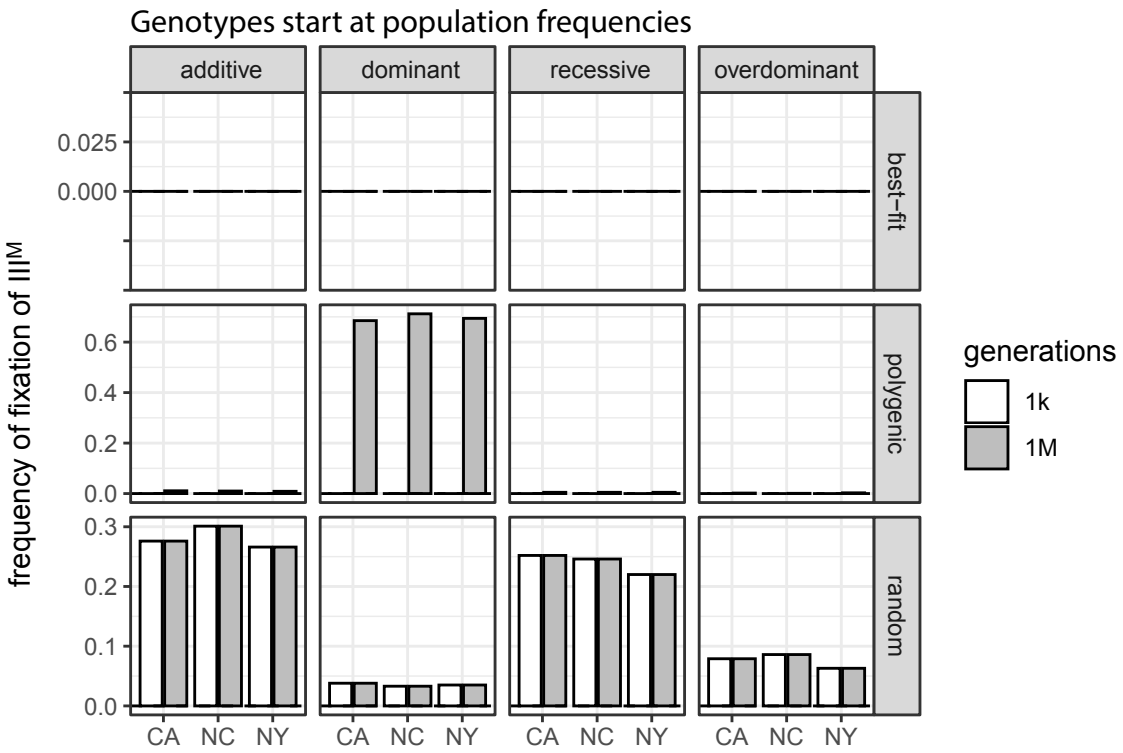
Supplemental Figure S4. The frequency of fixation of the Y^M proto-Y chromosome after 1,000 generations (white) and 1,000,000 generations (gray) is shown for the 1,000 best-fitting genotype fitness arrays (best-fit), 1,000 fitness arrays that maintain PSD (polygenic), and 1,000 randomly chosen fitness arrays (random) in each population (CA, NC, or NY). Best-fitting fitness values and fitness values that maintain PSD were identified using simulation in which all genotypes started at equal frequencies. Fitness effects of each proto-Y chromosome were either additive, dominant, recessive, or overdominant.



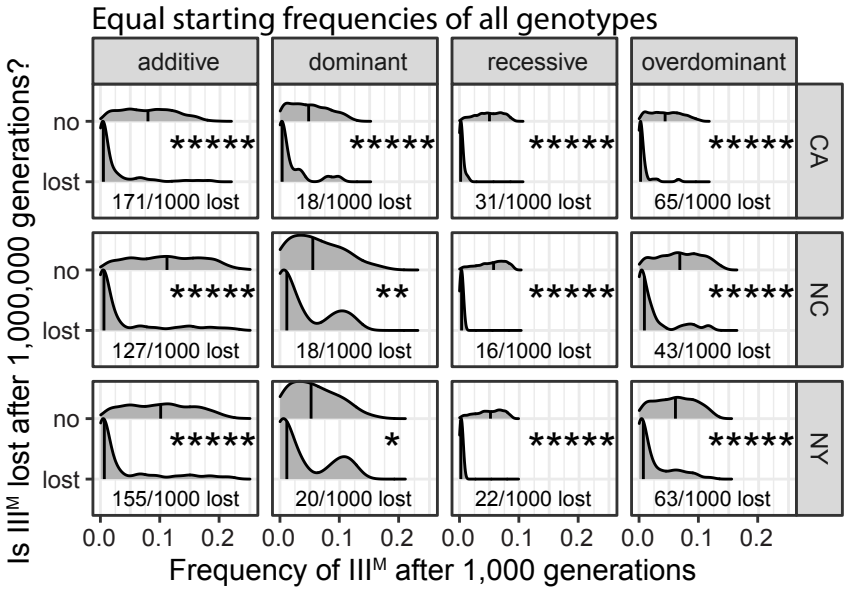
Supplemental Figure S5. The frequency of fixation of the Y^M proto-Y chromosome after 1,000 generations (white) and 1,000,000 generations (gray) is shown for the 1,000 best-fitting genotype fitness arrays (best-fit), 1,000 fitness arrays that maintain PSD (polygenic), and 1,000 randomly chosen fitness arrays (random) in each population (CA, NC, or NY). Best-fitting fitness values and fitness values that maintain PSD were identified using simulation in which all genotypes started at the frequencies observed in each population. Fitness effects of each proto-Y chromosome were either additive, dominant, recessive, or overdominant.



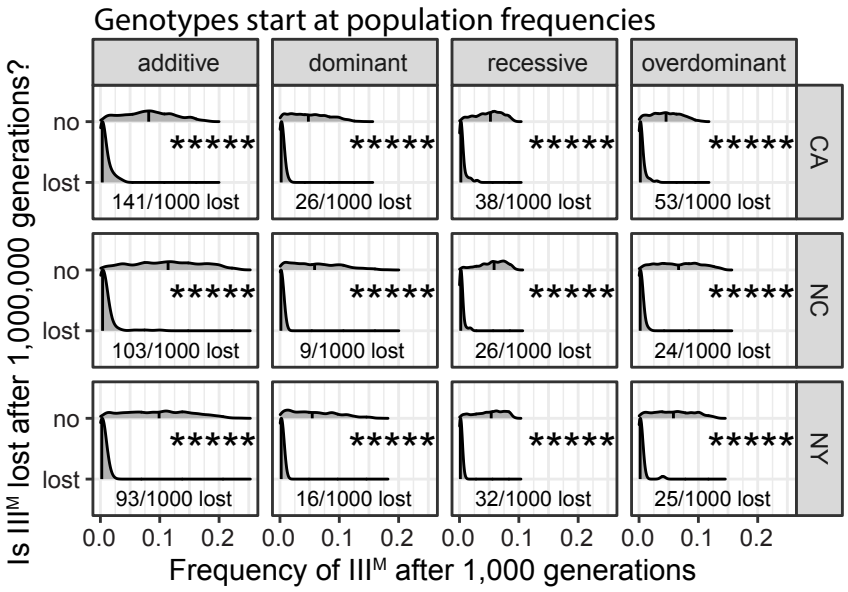
Supplemental Figure S6. The frequency of fixation of the III^M proto-Y chromosome after 1,000 generations (white) and 1,000,000 generations (gray) is shown for the 1,000 best-fitting genotype fitness arrays (best-fit), 1,000 fitness arrays that maintain PSD (polygenic), and 1,000 randomly chosen fitness arrays (random) in each population (CA, NC, or NY). Best-fitting fitness values and fitness values that maintain PSD were identified using simulation in which all genotypes started at equal frequencies. Fitness effects of each proto-Y chromosome were either additive, dominant, recessive, or overdominant.



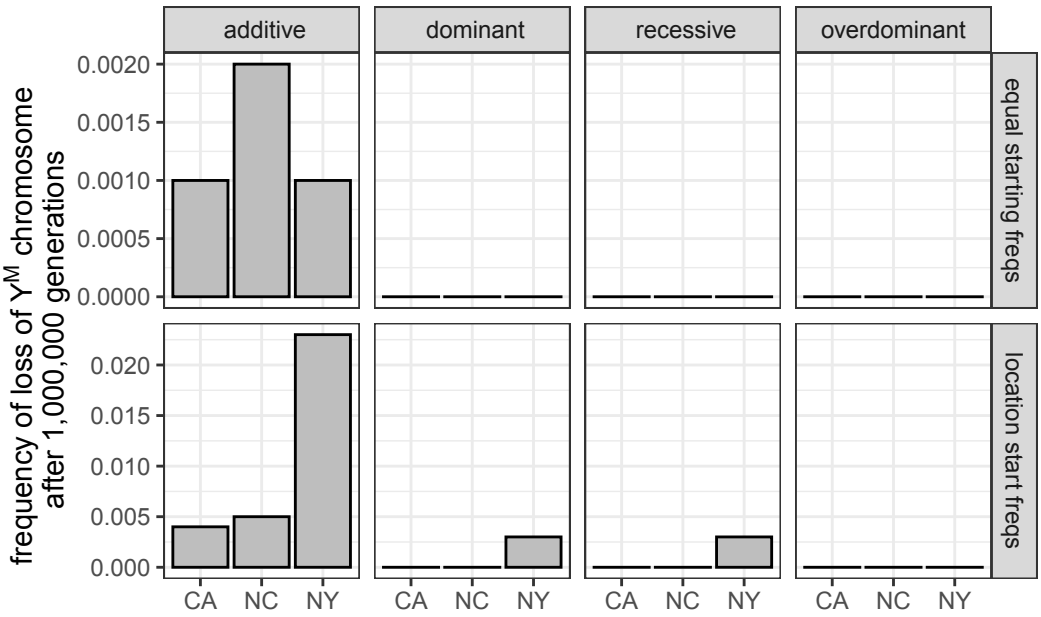
Supplemental Figure S7. The frequency of fixation of the III^M proto-Y chromosome after 1,000 generations (white) and 1,000,000 generations (gray) is shown for the 1,000 best-fitting genotype fitness arrays (best-fit), 1,000 fitness arrays that maintain PSD (polygenic), and 1,000 randomly chosen fitness arrays (random) in each population (CA, NC, or NY). Best-fitting fitness values and fitness values that maintain PSD were identified using simulation in which all genotypes started at the frequencies observed in each population. Fitness effects of each proto-Y chromosome were either additive, dominant, recessive, or overdominant.



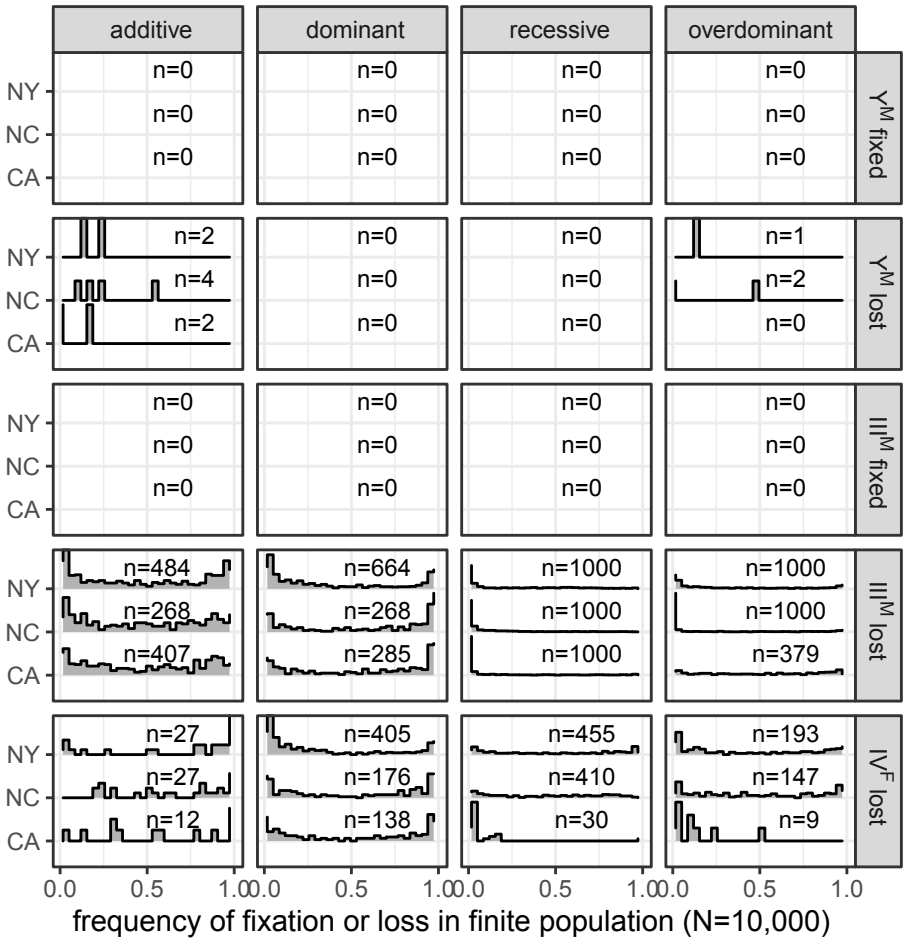
Supplemental Figure S8. Histograms of the frequency of the III^M proto-Y chromosome after 1,000 generations are shown for the 1,000 fitness arrays that produce proto-sex chromosomes most similar to those observed in each natural population (CA, NC, or NY). Fitness arrays are divided into those in which the III^M chromosome eventually reaches a frequency <0.1% after 1,000,000 generations (lost) and those in which the III^M chromosome is maintained at a frequency >0.1% (no). Fractions in each panel indicate how often the III^M chromosome was lost after 1,000,000 generations in the 1,000 best-fitting arrays. Simulations were started with equal frequencies of all genotypes, and fitness effects of proto-Y chromosomes were either additive, dominant, recessive, or overdominant.



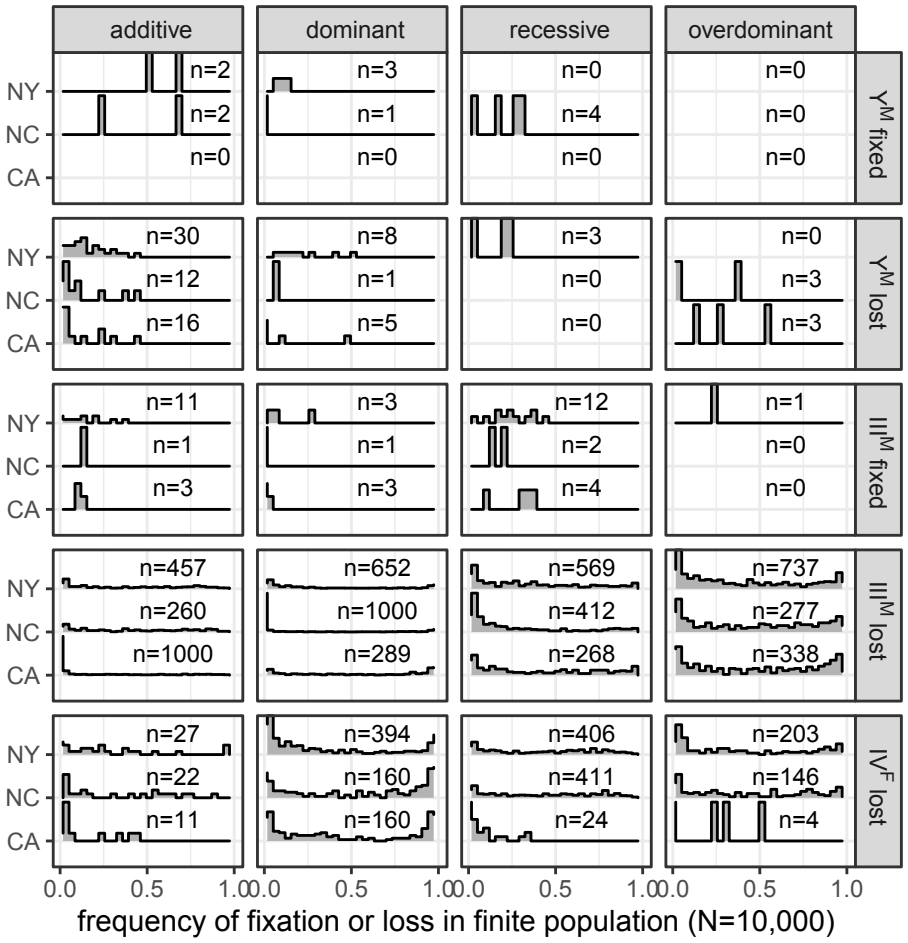
Supplemental Figure S9. Histograms of the frequency of the III^M proto-Y chromosome after 1,000 generations are shown for the 1,000 fitness arrays that produce proto-sex chromosomes most similar to those observed in each natural population (CA, NC, or NY). Fitness arrays are divided into those in which the III^M chromosome eventually reaches a frequency <0.1% after 1,000,000 generations (lost) and those in which the III^M chromosome is maintained at a frequency >0.1% (no). Fractions in each panel indicate how often the III^M chromosome was lost after 1,000,000 generations in the 1,000 best-fitting arrays. Simulations were started with genotypes frequencies observed in each population, and fitness effects of proto-Y chromosomes were either additive, dominant, recessive, or overdominant.



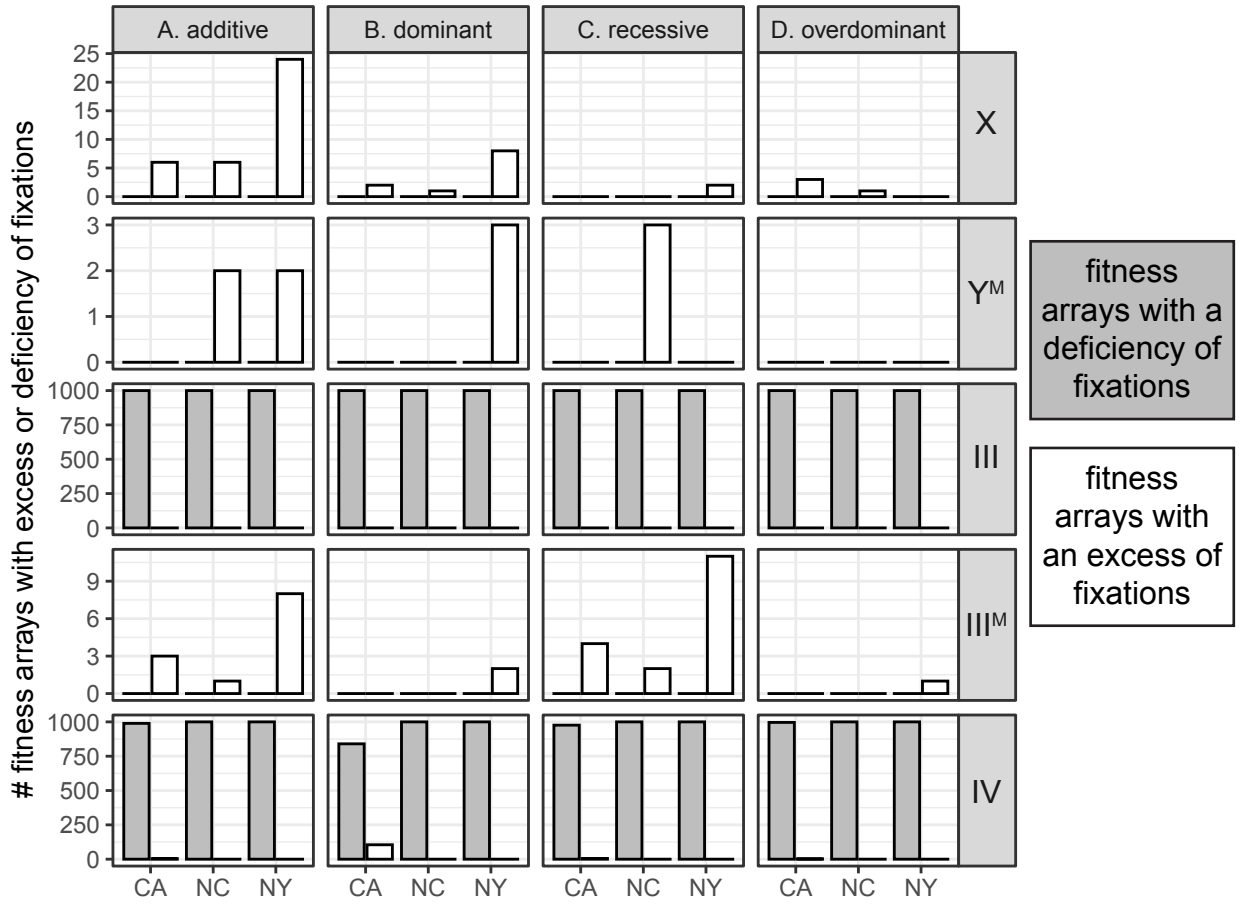
Supplemental Figure S10. Bar graphs show the frequency with which the Y^M chromosome is lost after 1,000,000 generation using fitness effects that maintain proto-sex chromosome frequencies similar to those observed in each natural population (CA, NC, or NY). The best-fitting arrays were identified with simulations starting with equal frequencies of all genotypes (equal starting freqs) or frequencies observed in each population (location start freqs). Genotype fitness values were calculating assuming either additive, dominant, recessive, or overdominant fitness effects.



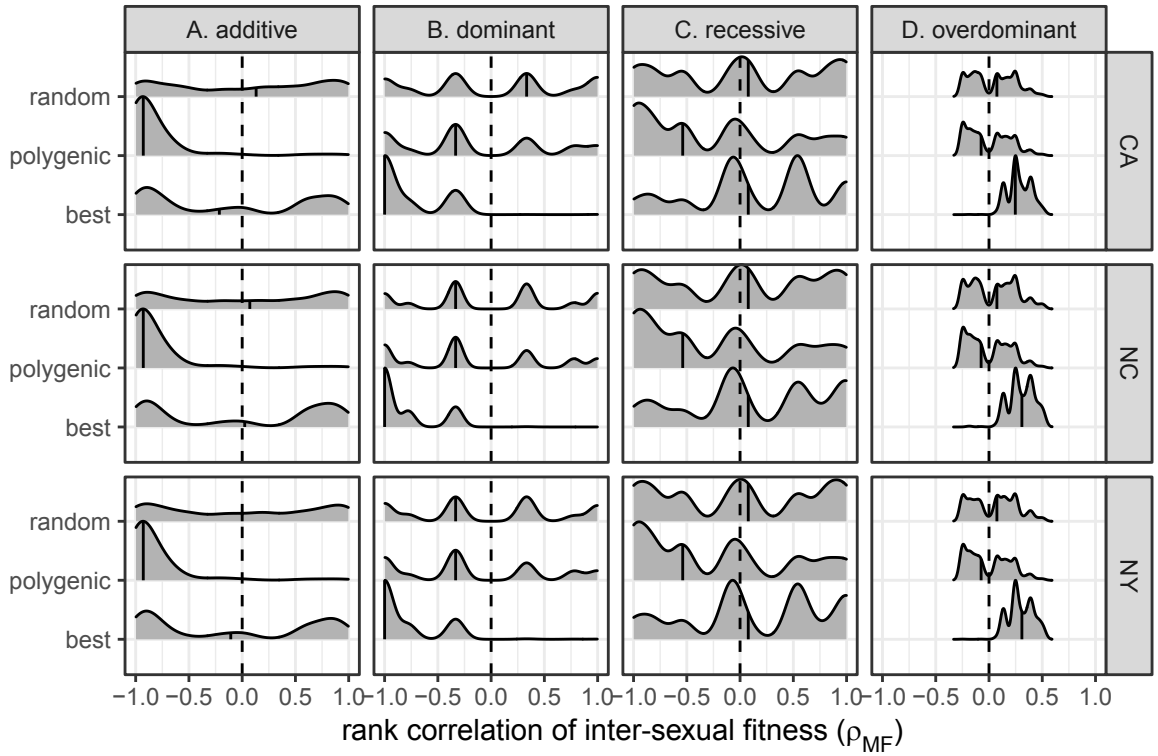
Supplemental Figure S11. Histograms show the distributions of frequency of fixation or loss each proto-Y or proto-W chromosome in each population when $N=10^4$. Numbers in each panel (n) indicate the number of fitness arrays with at least one fixation in 100 simulations. The x-axis is the frequency of fixations in 100 simulations for those arrays with at least one fixation. Simulations were performed with the 1,000 best-fitting selection pressures for each population (CA, NC, or NY) that were identified from simulations started with equal frequencies of all genotypes, using either **(A)** additive, **(B)** dominant, **(C)** recessive, or **(D)** overdominant fitness effects of the proto-Y chromosomes.



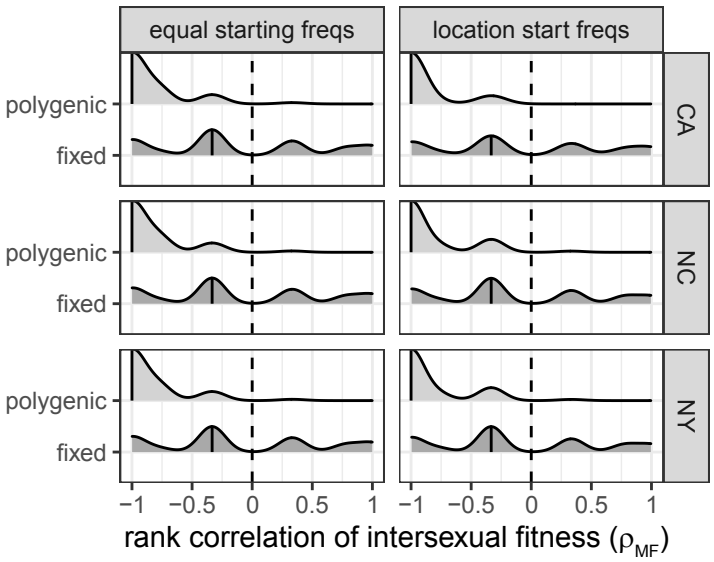
Supplemental Figure S12. Histograms show the distributions of frequency of fixation or loss each proto-Y or proto-W chromosome in each population when $N=10^4$. Numbers in each panel (n) indicate the number of fitness arrays with at least one fixation in 100 simulations. The x-axis is the frequency of fixations in 100 simulations for those arrays with at least one fixation. Simulations were performed with the 1,000 best-fitting selection pressures for each population (CA, NC, or NY) that were identified from simulations started with the observed frequencies of each genotype in each population, using either **(A)** additive, **(B)** dominant, **(C)** recessive, or **(D)** overdominant fitness effects of the proto-Y chromosomes.



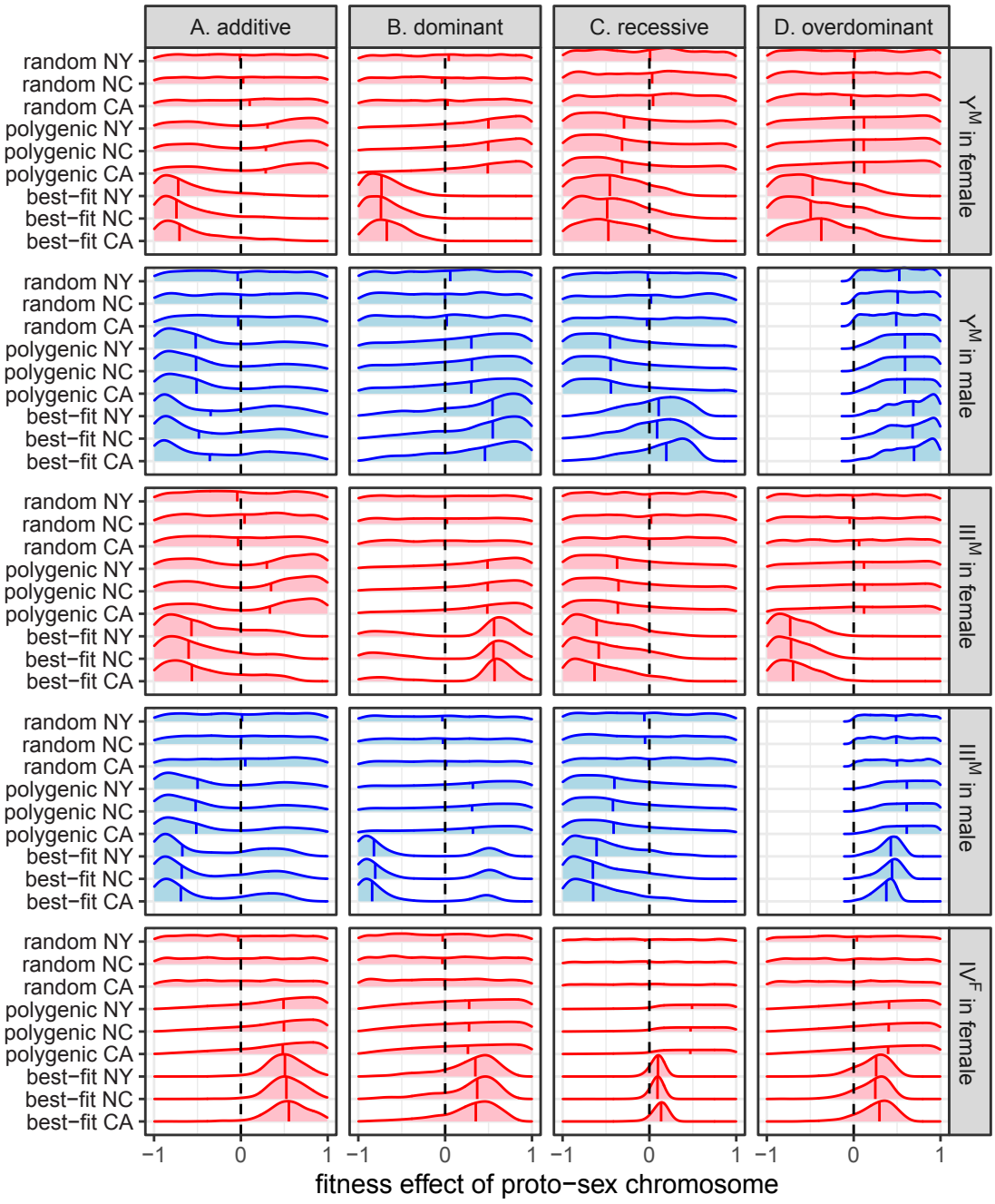
Supplemental Figure S13. Bar plots show the number of fitness arrays that result in a deficiency (gray) or excess (white) of fixations of each proto-sex chromosome (X, Y^M, III, III^M, or IV) relative to simulations in which there are no fitness differences across proto-sex chromosomes (i.e., drift only). The fitness arrays are the 1,000 best-fitting fitness values for each population (CA, NC, or NY) when simulations were started with genotype frequencies observed in natural populations. Fitness effects of the proto-Y chromosomes are either **(A)** additive, **(B)** dominant, **(C)** recessive, or **(D)** overdominant.



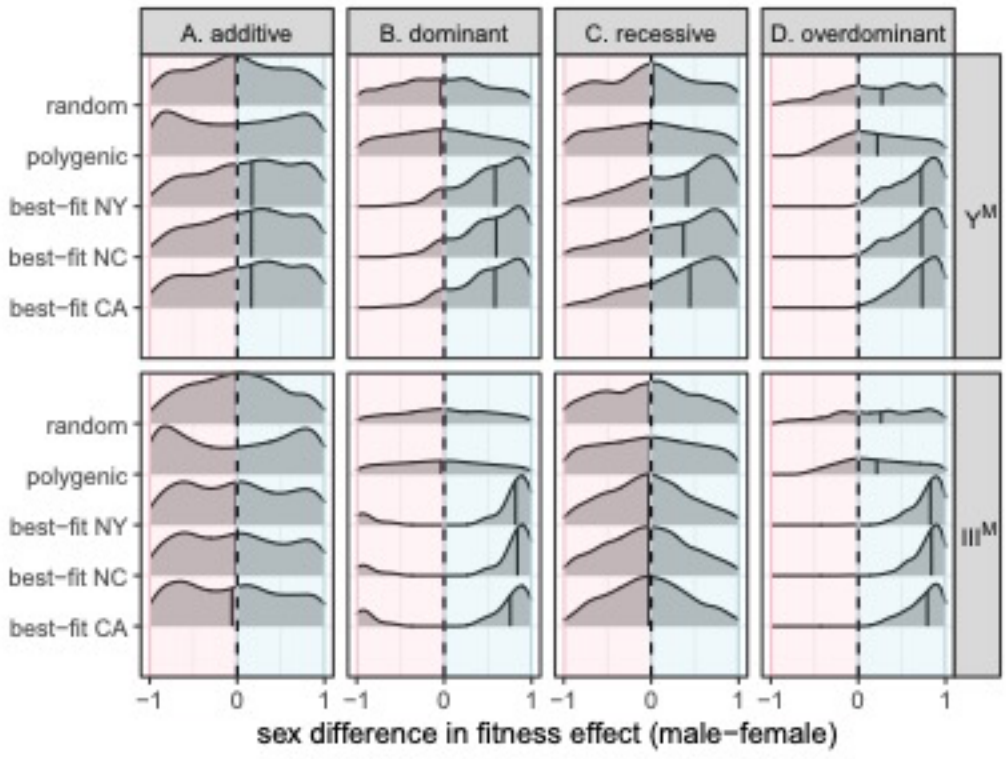
Supplemental Figure S14. Smoothed histograms show intersexual fitness correlations (ρ_{MF}) for the 1,000 best-fitting genotypic fitness arrays for the CA, NC, and NY populations (best), all fitness arrays that maintain PSD (polygenic), and 1,000 random fitness arrays. Simulations were started with genotype frequencies observed in natural populations. Dashed vertical lines show $\rho_{MF} = 0$, and solid vertical lines within histograms show the median. Fitness arrays were calculated assuming either (**A)** additive, **(B)** dominant, **(C)** recessive, or **(D)** overdominant fitness effects of the proto-Y chromosomes.



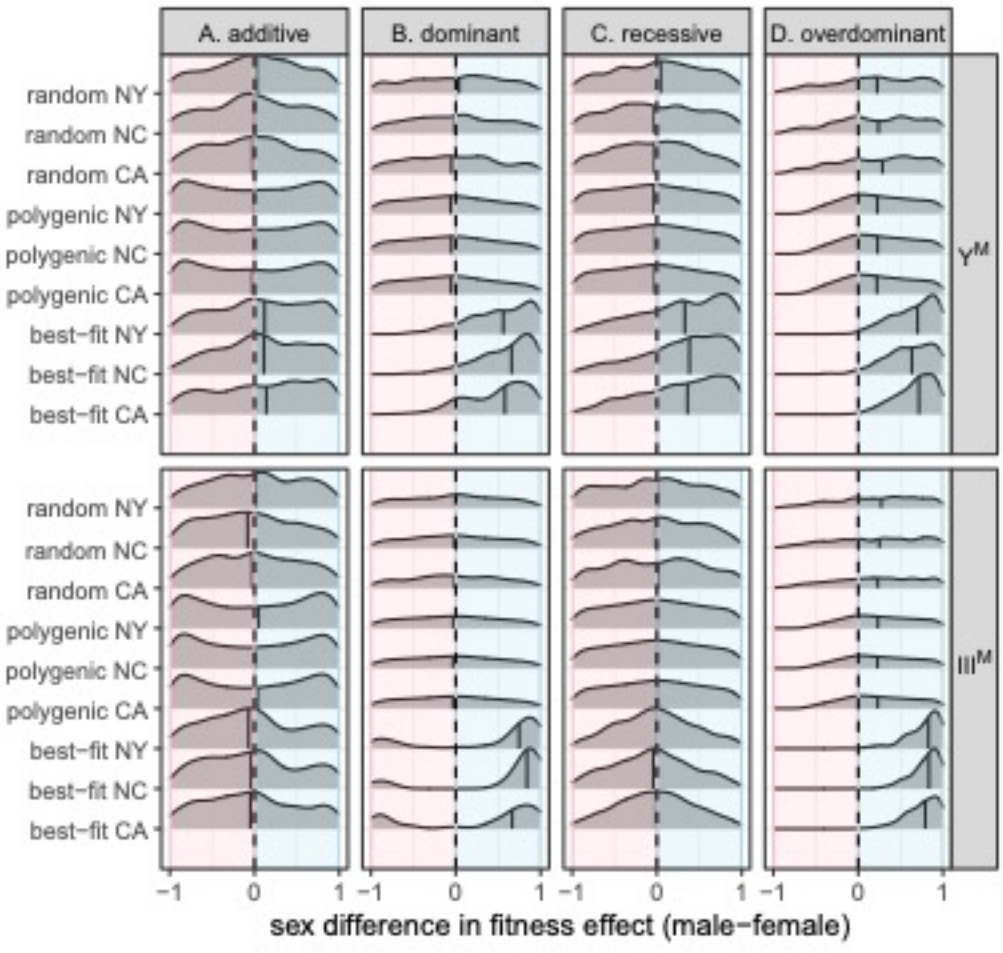
Supplemental Figure S15. Smoothed histograms show intersexual fitness correlations (ρ_{MF}) for the 1,000 best-fitting genotypic fitness arrays based on dominant fitness effects of the proto-Y chromosomes for the CA, NC, and NY populations after 1,000,000 generations. Fitness arrays that maintain PSD for 1,000 generations are divided into those the maintain PSD for 1,00,000 generations (polygenic) and those that allow at least one proto-sex chromosome to reach fixation (fixed). Simulations were started with either equal frequencies of all genotypes (left) or genotype frequencies observed in natural populations (right). Dashed vertical lines show $\rho_{MF} = 0$, and solid vertical lines within histograms show the median.



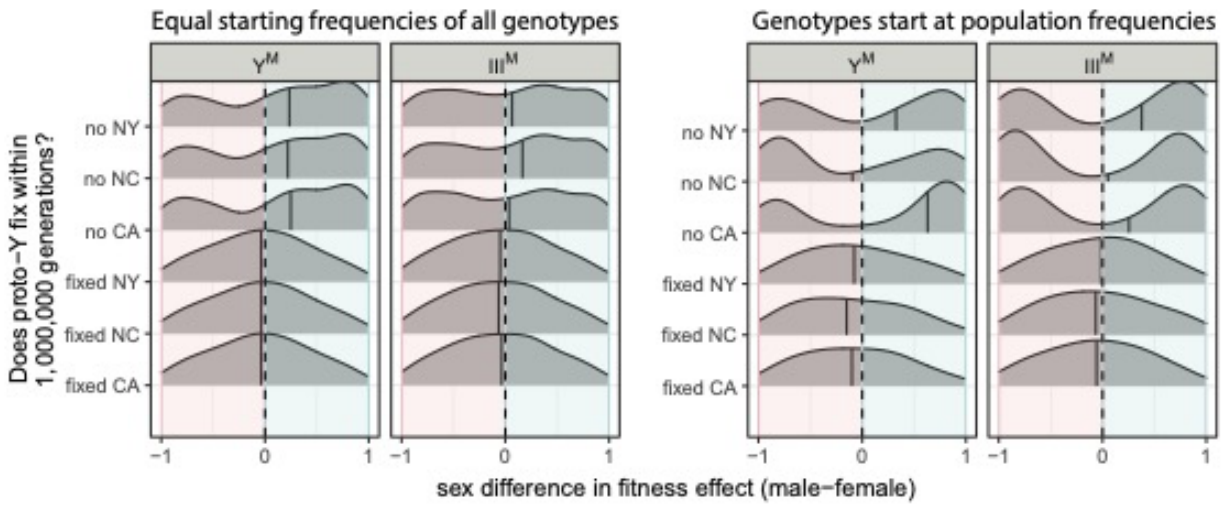
Supplemental Figure S16. Smoothed histograms show the distributions of fitness effects of each proto-sex chromosome in each sex for 1,000 random fitness arrays (random), fitness arrays that maintain PSD (polygenic), or the 1,000 best-fitting fitness arrays for each population (CA, NC, or NY). Simulations were started with genotype frequencies observed in natural populations. The vertical line within each histogram shows the median, and dashed lines in each panel are at fitness value of 0. Fitness arrays were calculated assuming either (**A**) additive, **(B)** dominant, **(C)** recessive, or **(D)** overdominant fitness effects of the proto-Y chromosomes.



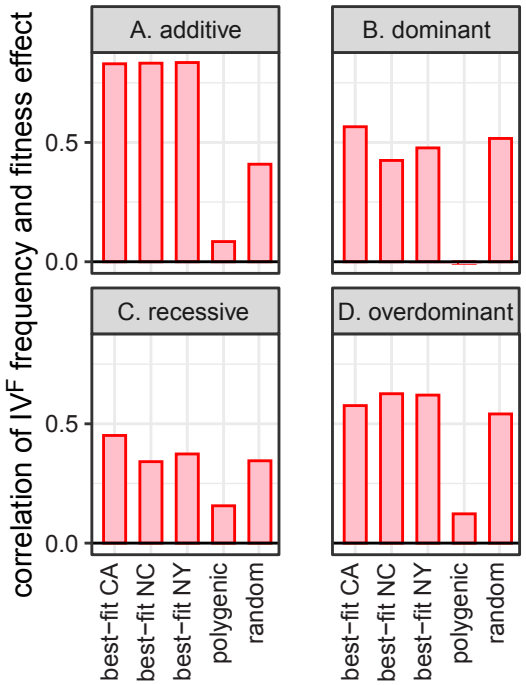
Supplemental Figure S17. Smoothed histograms show the distributions of the difference in the fitness effect of each proto-Y chromosome (Y^M or III^M) between males and females within individual parameterizations of the model. If the fitness effect is greater in males, the values are positive. If the fitness effect is greater in females, the values are negative. Fitness effects of each proto-Y chromosome come from the 1,000 random fitness arrays (random), fitness arrays that maintain PSD (polygenic), or the 1,000 best-fitting fitness arrays for each population (CA, NC, or NY). Simulations were started with equal frequencies of all genotypes. The vertical line within each histogram shows the median, and dashed lines in each panel are a difference of 0 (i.e., equal fitness effects in males and females). Genotype fitness arrays were calculated assuming either (**A)** additive, **(B)** dominant, **(C)** recessive, or **(D)** overdominant fitness effects of the proto-Y chromosomes.



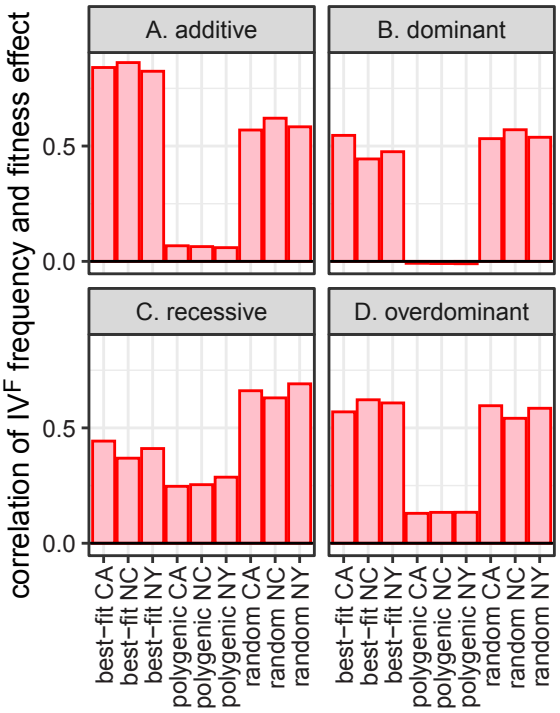
Supplemental Figure S18. Smoothed histograms show the distributions of the difference in the fitness effect of each proto-Y chromosome (Y^M or III^M) between males and females within individual parameterizations of the model. If the fitness effect is greater in males, the values are positive. If the fitness effect is greater in females, the values are negative. Fitness effects of each proto-Y chromosome come from the 1,000 random fitness arrays (random), fitness arrays that maintain PSD (polygenic), or the 1,000 best-fitting fitness arrays for each population (CA, NC, or NY). Simulations were started with genotype frequencies observed in natural populations. The vertical line within each histogram shows the median, and dashed lines in each panel are a difference of 0 (i.e., equal fitness effects in males and females). Genotype fitness arrays were calculated assuming either (**A)** additive, **(B)** dominant, **(C)** recessive, or **(D)** overdominant fitness effects of the proto-Y chromosomes.



Supplemental Figure S19. Smoothed histograms show the distributions of the difference in the fitness effect of each proto-Y chromosome (Y^M or III^M) between males and females within individual parameterizations of the model. If the fitness effect is greater in males, the values are positive. If the fitness effect is greater in females, the values are negative. Simulations were performed using 1,000 fitness arrays based on dominant fitness effects of the proto-Y chromosomes that maintain PSD for 1,000 generations. Fitness arrays that maintain PSD for 1,000 generations are divided into those in which the proto-Y chromosome reaches fixation within 1,00,000 generations (fixed) and those in which it does not reach fixation (no) in each population (CA, NC, or NY). Simulations were started with either equal frequencies of all genotypes (left) or genotype frequencies observed in natural populations (right). The vertical line within each histogram shows the median, and dashed lines in each panel are a difference of 0 (i.e., equal fitness effects in males and females).



Supplemental Figure S20. Bar plots show Spearman's rank order correlatiodn between the frequency the IV^F proto-W chromosome across simulated populations and its fitness effect in the populations. Simulated populations include those that best-fit proto-sex chromosome frequencies in natural populations (CA, NC, or NY), those that maintain PSD (polygenic), and 1,000 random populations (random). Simulations were started with equal frequencies of all genotypes. Genotype fitness in populations was calculated assuming **(A)** additive, **(B)** dominant, **(C)** recessive, or **(D)** overdominant effects of the proto-Y chromosomes.



Supplemental Figure S21. Bar plots show Spearman's rank order correlation between the frequency the IV^F proto-W chromosome across simulated populations and its fitness effect in the populations. Simulated populations include those that best-fit proto-sex chromosome frequencies in natural populations (CA, NC, or NY), those that maintain PSD (polygenic), and 1,000 random populations (random). Simulations were started with the observed frequencies of each genotype in each population. Genotype fitness in populations was calculated assuming **(A)** additive, **(B)** dominant, **(C)** recessive, or **(D)** overdominant effects of the proto-Y chromosomes.