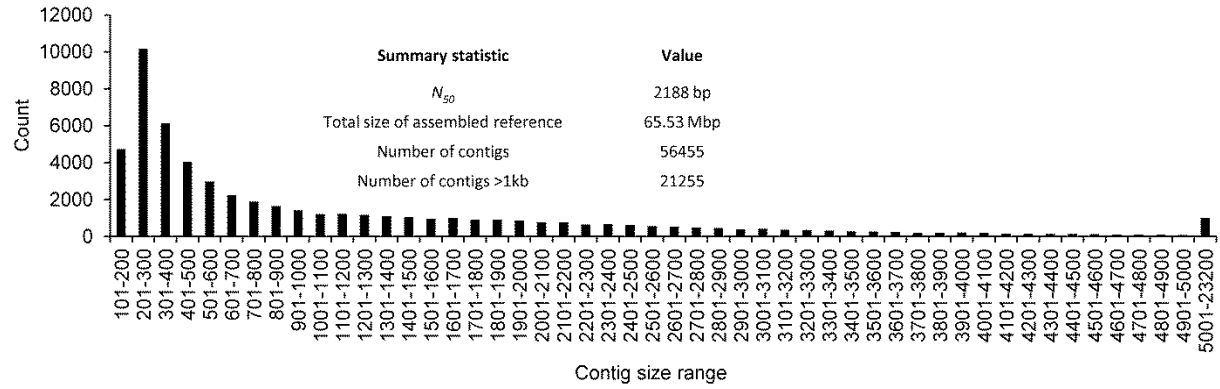**Figure S1** Genome structure implemented in the simulations. Each genome was 100 Mbp comprised of alternating 800 bp noncoding (NC) and 200 bp coding (C) regions. All sites in NC and 25% of sites in C were neutral with h=0.5 and the remaining 75% of C were under varying selection coefficients.

R. Arunkumar *et al.*

Figure S2 Length distribution of all contigs from the *de novo* assembly of six Jamaican genotypes of *Eichhornia paniculata* using Velvet-Oases. The summary statistics for the assembly are shown in the table within the figure. The N_{50} value shows that 50% of all bases in the assembled reference are in contigs of size corresponding to this value or larger.
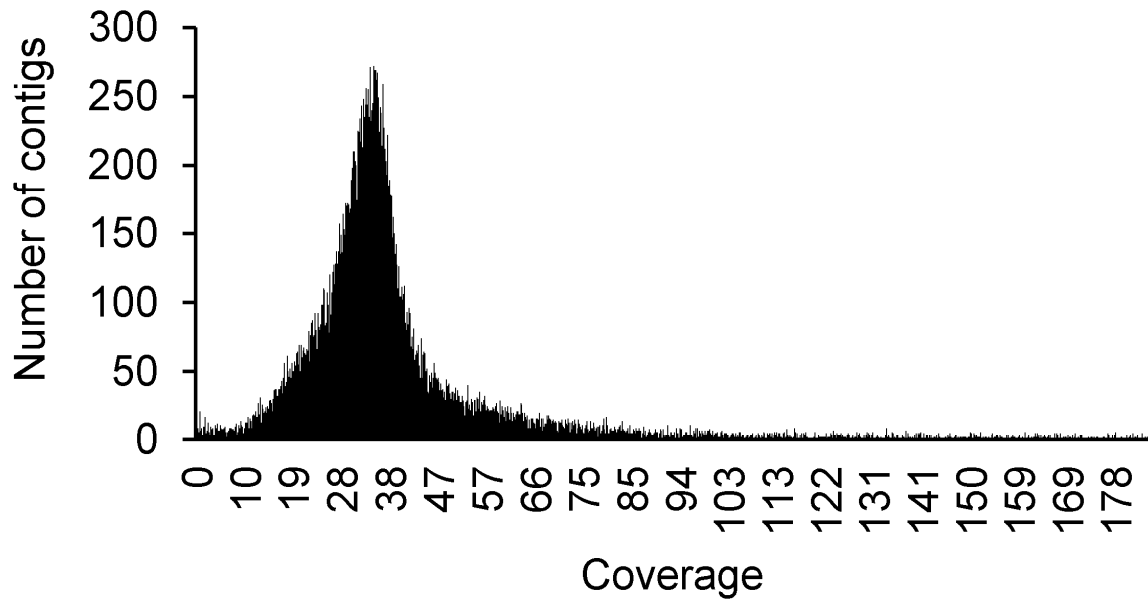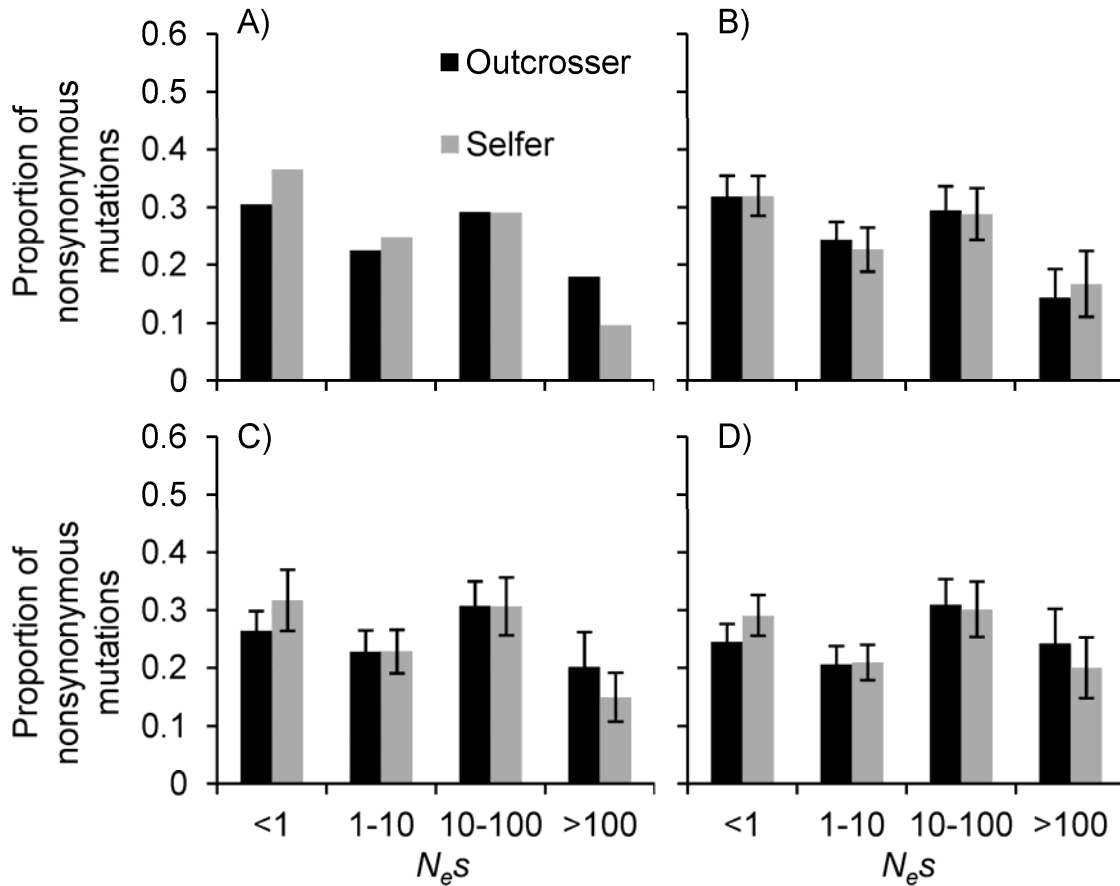
Figure S3 Coverage distribution of genomic reads from one selfing Mexican genotype of *Eichhornia paniculata* mapped to the *de novo* transcriptome reference.

R. Arunkumar *et al.*
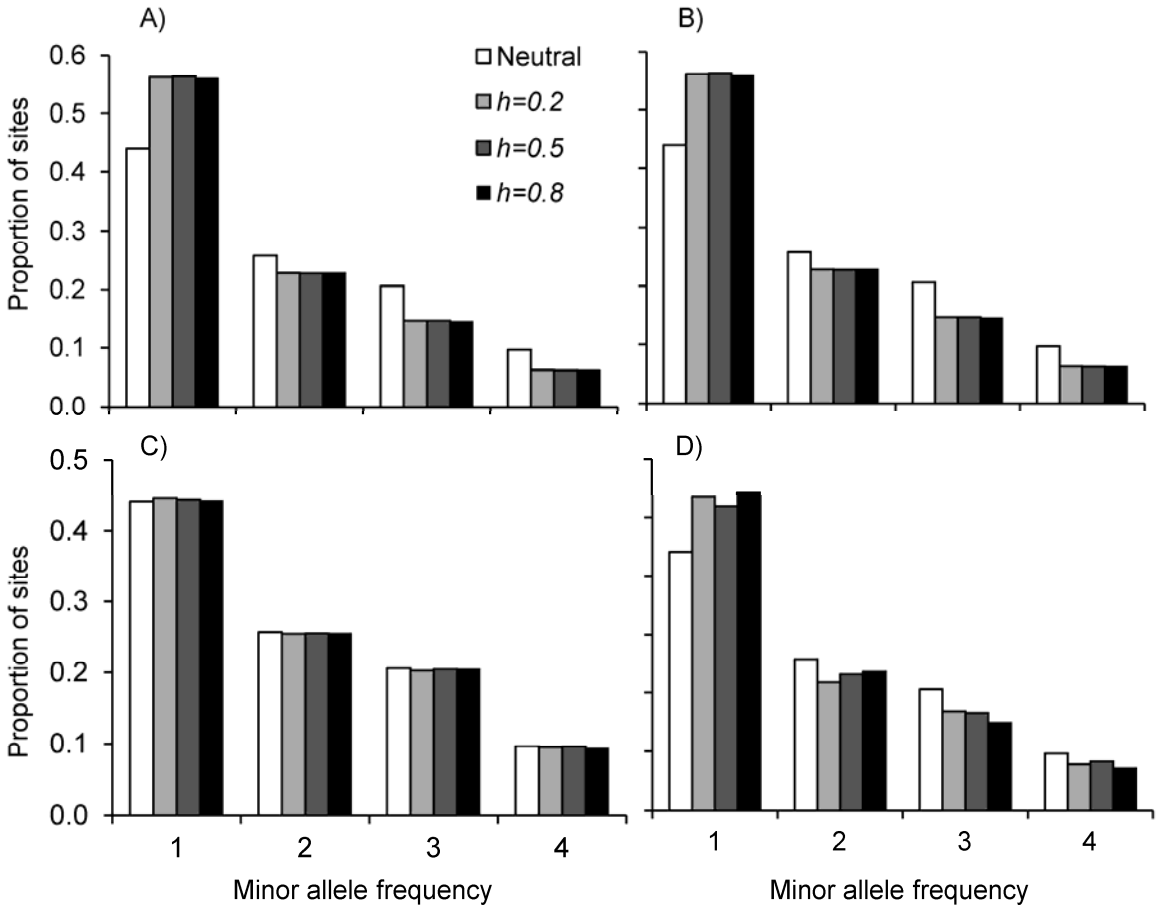
Figure S4 Folded synonymous allele frequency spectra for simulated outcrossing (*t*=1.00) and selfing (*t*=0.02) populations under various fixed dominance coefficients (*h*=0.2, 0.5, or 0.8). We generated the allele frequency spectra by randomly sampling eight genomes from populations of size 1000. To generate the frequency spectra we either used A) mutations unique to outcrossing population, B) mutations unique to selfing population, C) mutations unique to outcrossing populations and shared between outcrossing and selfing populations, or D) mutations unique to selfing populations and shared between outcrossing and selfing populations. The white open bars depict the neutral equilibrium expectation.
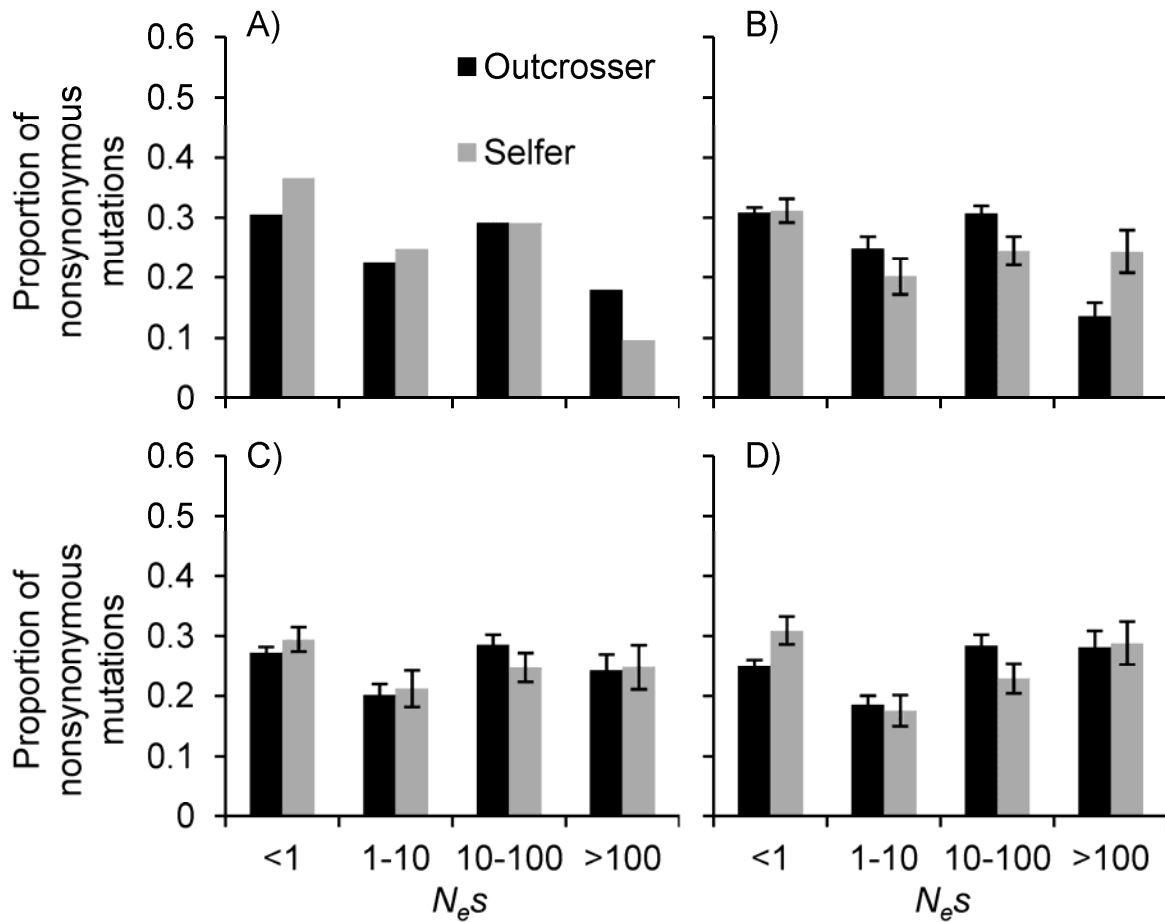
Figure S5 Distribution of fitness effects (DFE) of new nonsynonymous mutations or simulated outcrossing (t=1.00) and selfing (t=0.02) populations under various fixed dominance coefficients (h=0.2, 0.5, or 0.8) when both mutations unique to each population and mutations shared between populations were used to estimate selective and demographic parameters. $N_e s$ is the product of N_e and the selection coefficient (s). Shown are A) the expected DFEs, B) h=0.2, C) h=0.5, and D) h=0.8. Simulations illustrate the change in DFE for both populations after 6N generations after a split from the common outcrossing ancestor. We generated the expected DFE for the outcrossing population represented as a gamma distribution using supplied $β$ and Ns parameters. We generated the expected DFE for the selfing population by scaling supplied Ns parameter by the observed 75% reduction in synonymous diversity in selfers and multiplying it by 1.96 to account for effective dominance levels of mutations in selfing genomes while leaving $β$ unchanged. We generated observed DFEs by randomly sampling and generating allele frequency spectra using eight genomes from populations of size 1000. The coding regions were under various selection coefficients (Ns=0.5-95) all sampled from a gamma distribution with shape parameter ($β$) 0.3. Shown are the mean proportions of sites for each $N_e s$ category and their respective confidence intervals based on 120 simulations.

R. Arunkumar *et al.*

Figure S6 Distribution of fitness effects (DFE) of new nonsynonymous mutations for simulated outcrossing (t=1.00) and selfing (t=0.02) populations under various fixed dominance coefficients (h=0.2, 0.5, or 0.8) when synonymous mutations occurred only within coding regions were used to estimate demographic parameters. $N_e s$ is the product of N_e and the selection coefficient (s). Shown are A) the expected DFEs, B) h=0.2, C) h=0.5, and D) h=0.8. Simulations illustrate the change in DFE for both populations after 6N generations after a split from the common outcrossing ancestor. We generated the expected DFE for the outcrossing population represented as a gamma distribution using supplied β and Ns parameters. We generated the expected DFE for the selfing population by scaling supplied Ns parameter by the observed 75% reduction in synonymous diversity in selfers and multiplying it by 1.96 to account for effective dominance levels of mutations in selfing genomes while leaving β unchanged. We generated observed DFEs by randomly sampling and generating allele frequency spectra using eight genomes from populations of size 1000. The coding regions were under various selection coefficients (Ns=0.5-95) all sampled from a gamma distribution with shape parameter (β) 0.3. Shown are the mean proportions of sites for each $N_e s$ category and their respective confidence intervals based on 120 simulations.
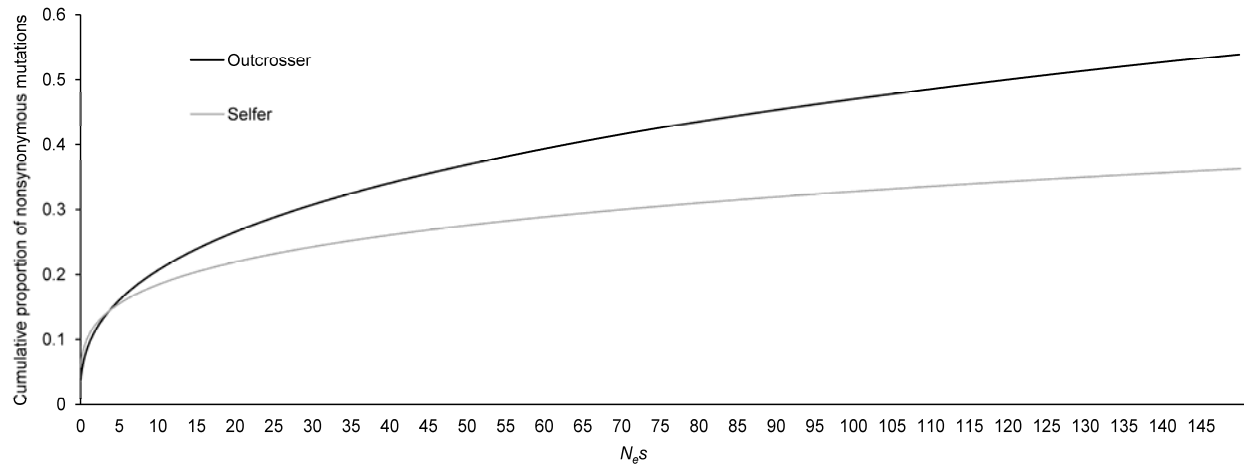
Figure S7 Distribution of fitness effects (DFE) of new nonsynonymous mutations for simulated outcrossing populations with recombination rate (r) of either $5×10^{-8}$ or $1.96×10^{-9}$ per site per generation under various fixed dominance coefficients. Shown are A) the expected DFEs, B) $h=0.2$, C) $h=0.5$, and D) $h=0.8$. Simulations illustrate the DFE for both populations after $16N$ generations. The coding regions were under various selection coefficients ($Ns=0.5-95$) all sampled from a gamma distribution with shape parameter ($β$) 0.3. The population had a constant mutation rate ($μ=7×10^{-9}$ per site per generation) and population size ($N=1000$) during the simulation runs. We generated the expected DFE under $N=1000$ for an outcrossing population represented as a gamma distribution using supplied $β$ and Ns parameters. We expected DFE under $N=250$ by scaling supplied Ns parameters while leaving $β$ unchanged. We generated the observed DFEs by randomly sampling and generating allele frequency spectra using eight genomes. Shown are the mean proportions of sites for each N_es category and their respective confidence intervals based on 120 simulations.

R. Arunkumar *et al.*
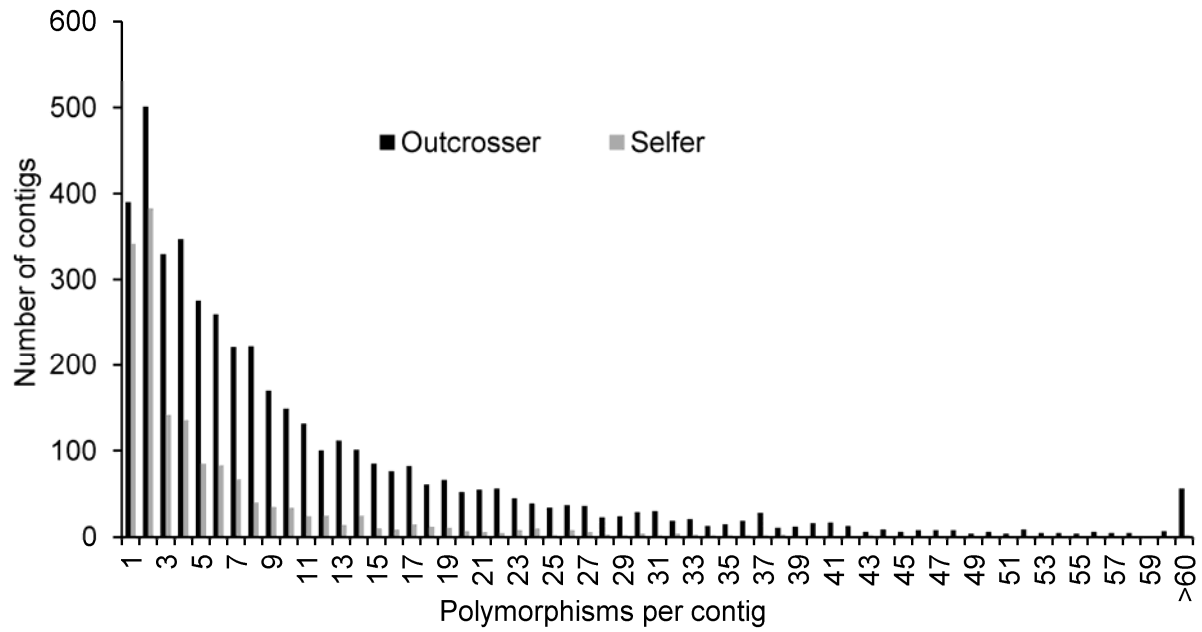
Figure S8 Distribution of polymorphisms per contig for eight outcrossing *Eichhornia paniculata* populations from N.E. Brazil and eight selfing populations from the Caribbean. Of the 16416 loci, 4485 were polymorphic in outcrossers and 1586 were polymorphic in selfers.

Figure S9 Cumulative distribution of fitness effects (DFE) of new nonsynonymous mutations for outcrossing and selfing populations of *Eichhornia paniculata*. $N_e s$ is the product of N_e and the selection coefficient (s). Eight Caribbean selfing samples and eight outcrossing individuals were used to generate the DFEs. Eight outcrossing samples from the 10 that were sequenced were randomly selected to keep the number of chromosomes sampled the same while performing the comparisons. The shape (β) and mean selection against deleterious mutations ($N*E(s)$) parameter estimates from the approach of Keightley and Eyre-Walker (2007) were used to plot the gamma distributions. For outcrossers, β=0.3691, $-N*E(s)$=365.16; for selfers β=0.2517, $-N*E(s)$=3112.84.

R. Arunkumar *et al.*
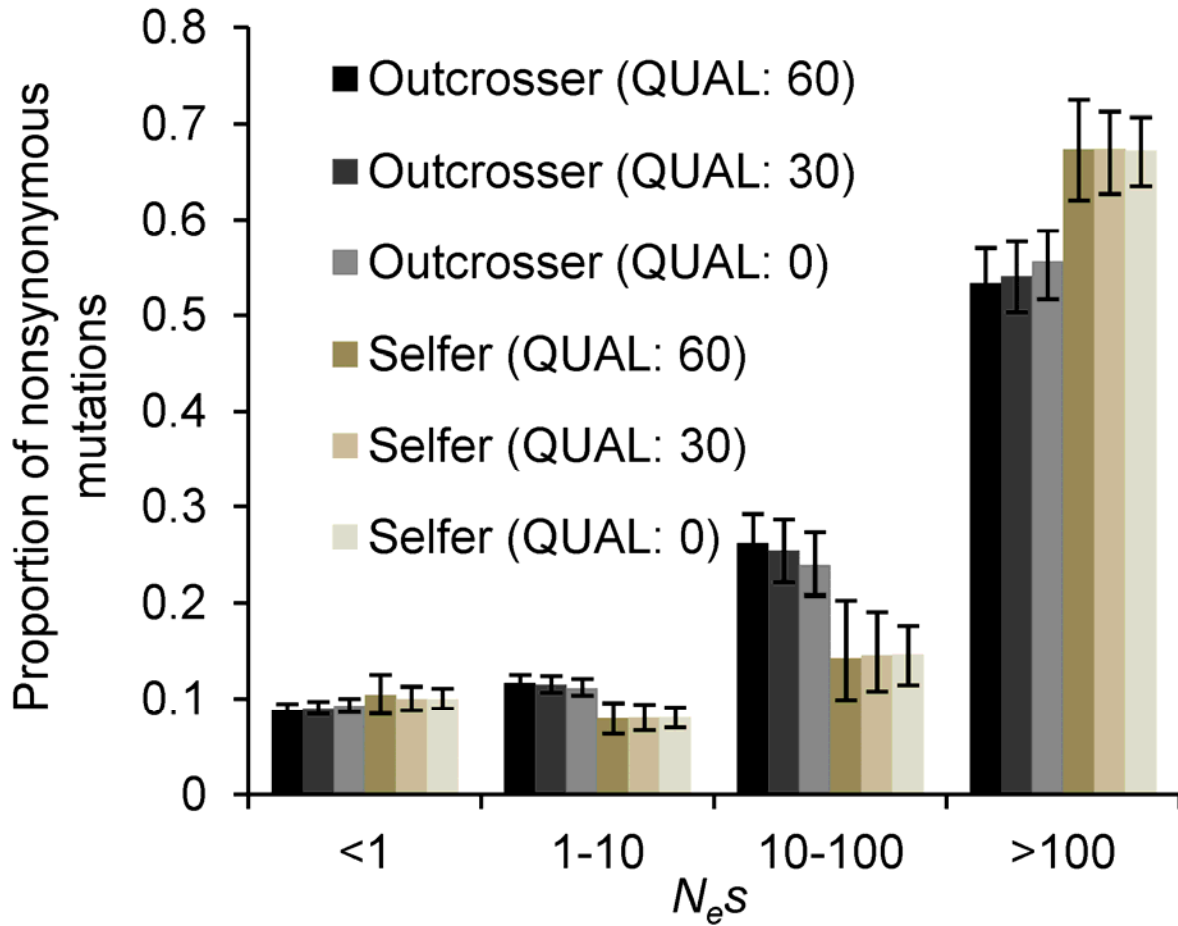
Figure S10 Distribution of fitness effects (DFE) of new nonsynonymous mutations for outcrossing and selfing populations of *Eichhornia paniculata* under varying quality cut-off scores (QUAL) for invariant and variant sites. $N_e s$ is the product of N_e and the selection coefficient (*s*). Eight Caribbean selfing samples and eight outcrossing individuals were used to generate the DFEs. Eight outcrossing samples from the 10 that were sequenced were randomly selected to keep the number of chromosomes sampled the same while performing the comparisons. Error bars for each $N_e s$ category are 95% confidence intervals from 200 bootstrap replicates generated by resampling over loci.
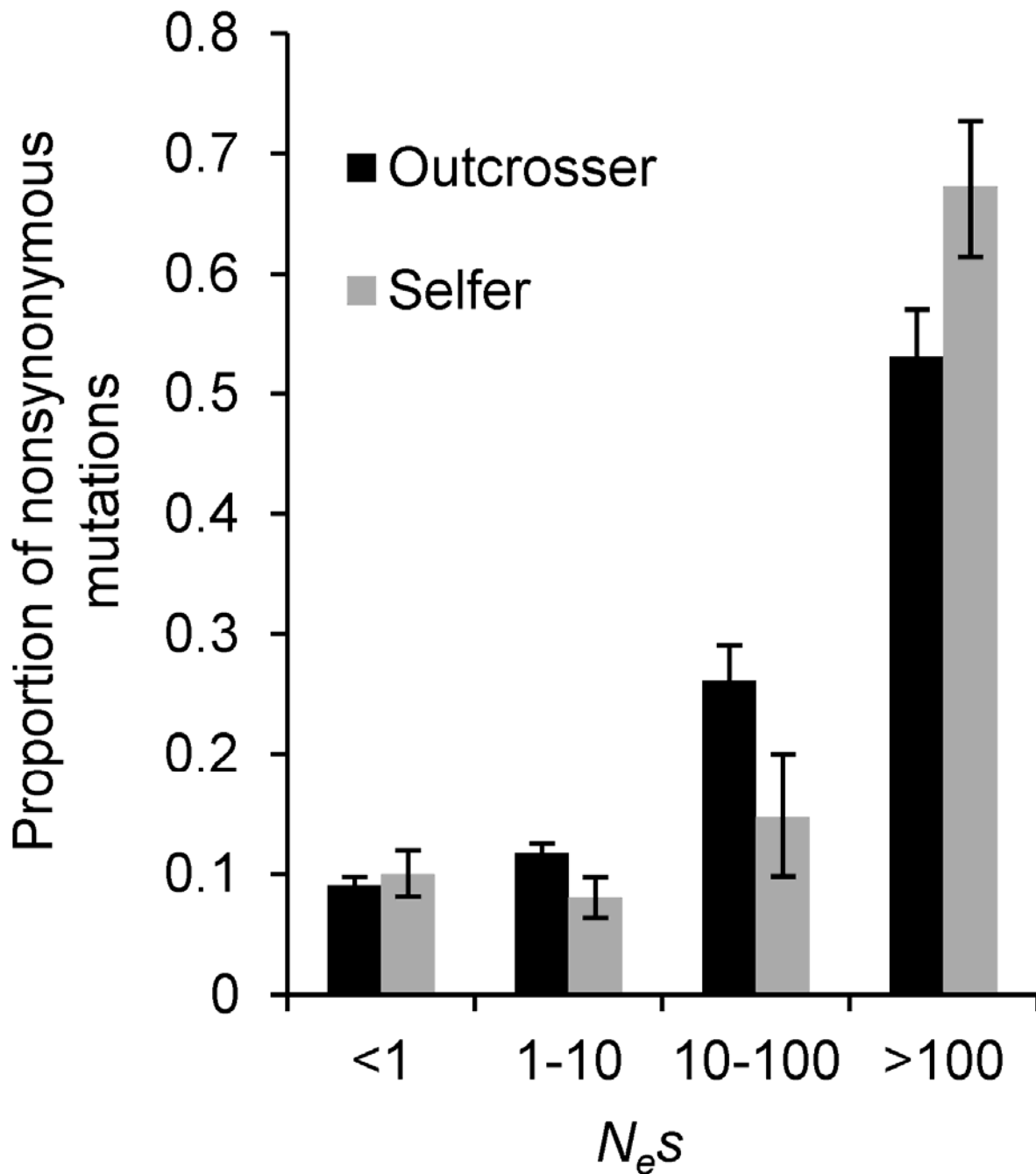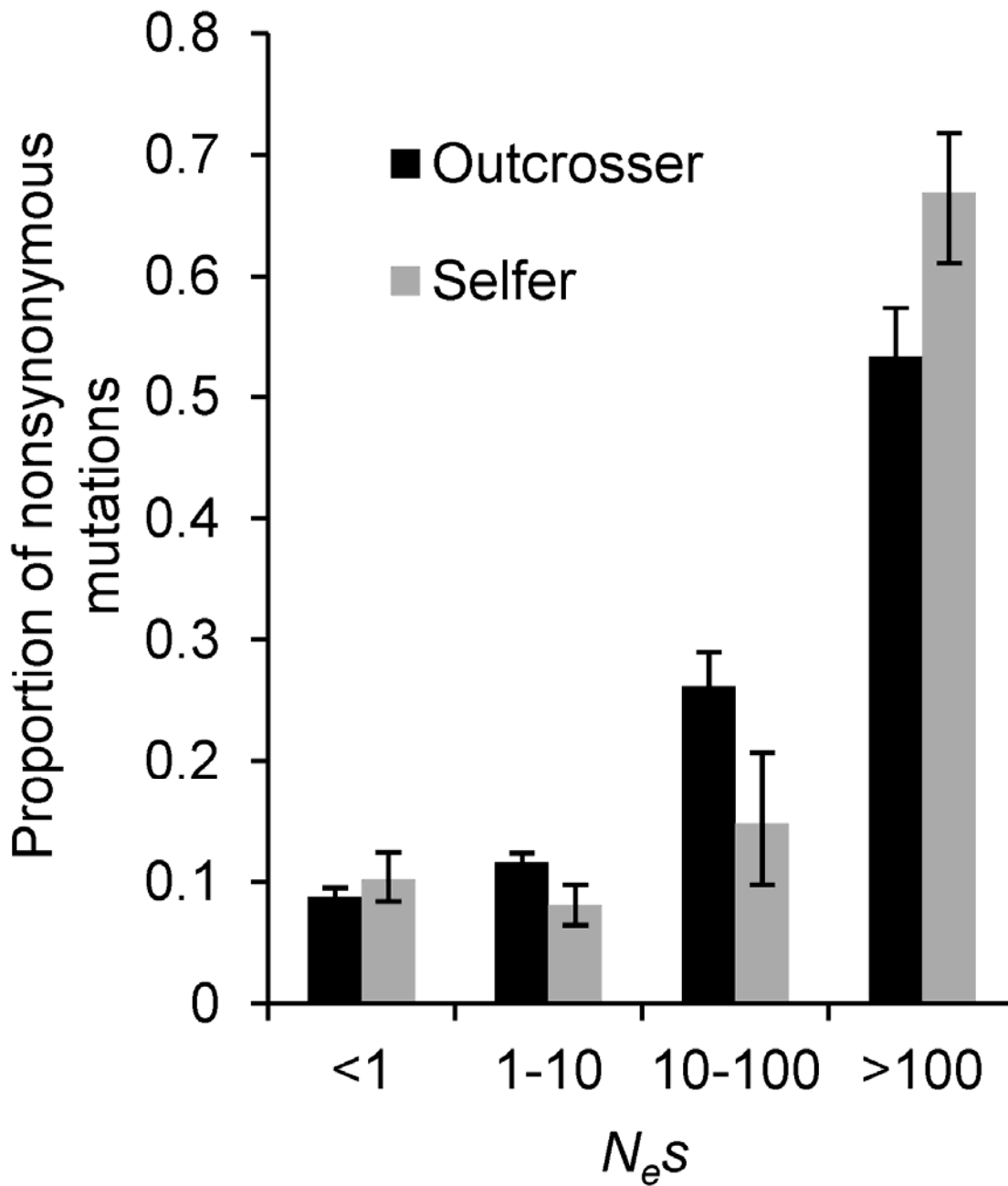
Figure S11 Distribution of fitness effects (DFE) of new nonsynonymous mutations for outcrossing and selfing populations of *Eichhornia paniculata* after 245 loci containing sites heterozygous in two or more selfing populations were excluded from analyses. N_es is the product of N_e and the selection coefficient (*s*). Eight Caribbean selfing samples and eight outcrossing individuals were used to generate the DFEs. Eight outcrossing samples from the 10 that were sequenced were randomly selected to keep the number of chromosomes sampled the same while performing the comparisons. Error bars for each N_es category are 95% confidence intervals from 200 bootstrap replicates generated by resampling over loci.
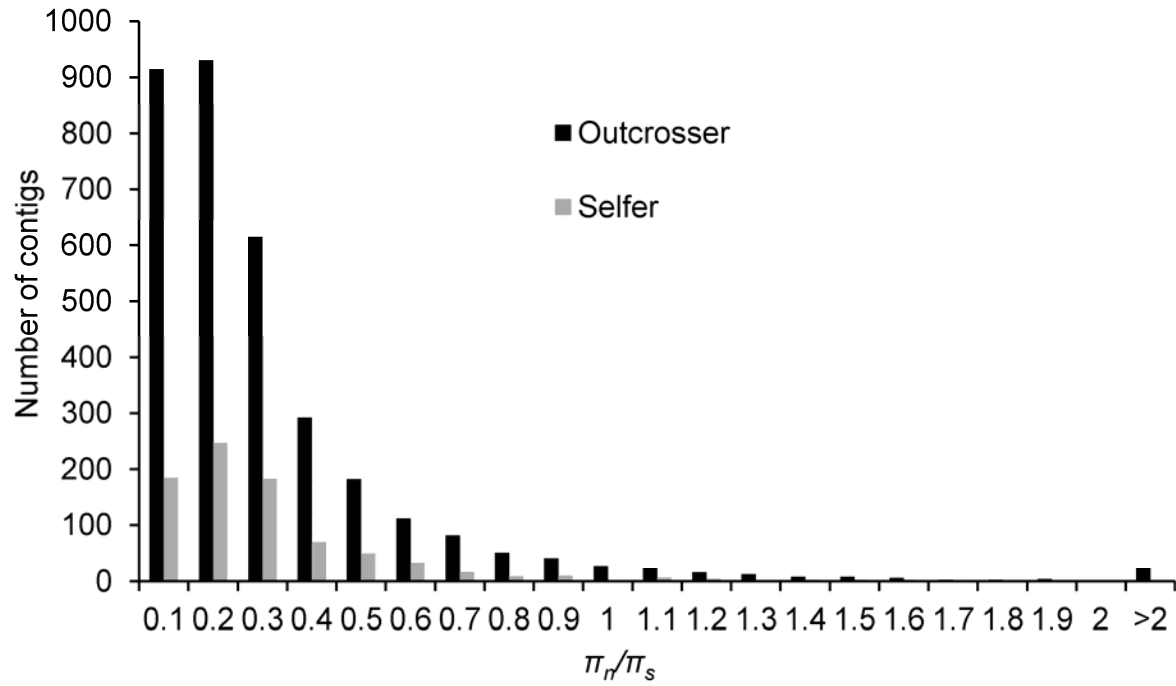
R. Arunkumar *et al.*

Figure S12 Distribution of fitness effects (DFE) of new nonsynonymous mutations for outcrossing and selfing populations of *Eichhornia paniculata* after 73 loci containing sites that violated assumptions of Hardy-Weinberg equilibrium in outcrossing populations were excluded from analyses. $N_e s$ is the product of N_e and the selection coefficient (*s*). Eight Caribbean selfing samples and eight outcrossing individuals were used to generate the DFEs. Eight outcrossing samples from the 10 that were sequenced were randomly selected to keep the number of chromosomes sampled the same while performing the comparisons. Error bars for each $N_e s$ category are 95% confidence intervals from 200 bootstrap replicates generated by resampling over loci.
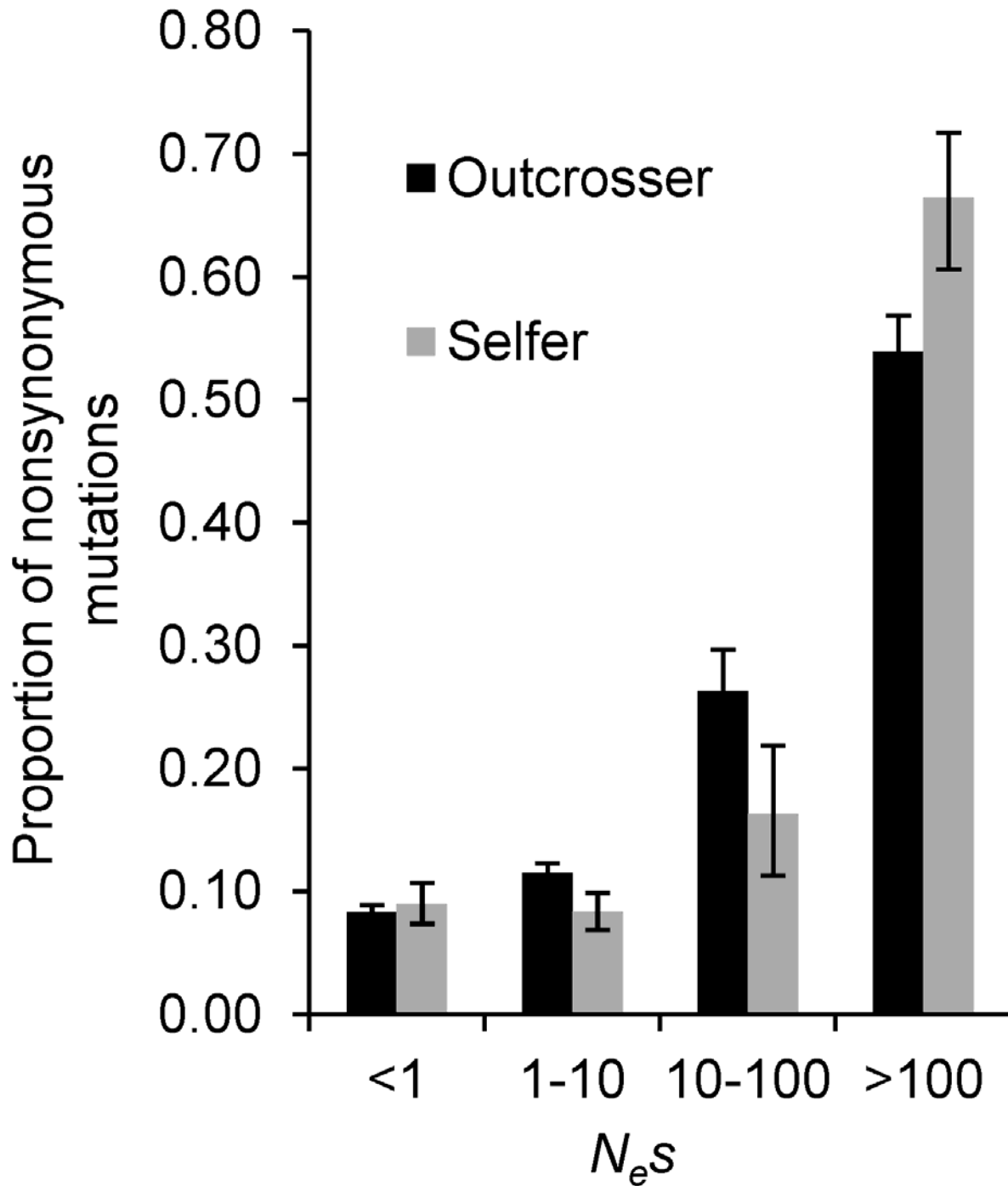
Figure S13 Distribution of π_n/π_s for polymorphic contigs among eight outcrossing *Eichhornia paniculata* populations from N.E. Brazil and eight selfing populations from the Caribbean. Of the 16416 loci, 4485 were polymorphic in outcrossers and 1586 were polymorphic in selfers.

Figure S14 Distribution of fitness effects (DFE) of new nonsynonymous mutations for outcrossing and selfing populations of *Eichhornia paniculata* after 135 loci with $\pi_n/\pi_s > 1$ in either population were excluded from analyses. $N_e s$ is the product of N_e and the selection coefficient (s). Eight Caribbean selfing samples and eight outcrossing individuals were used to generate the DFEs. Eight outcrossing samples from the 10 that were sequenced were randomly selected to keep the number of chromosomes sampled the same while performing the comparisons. Error bars for each $N_e s$ category are 95% confidence intervals from 200 bootstrap replicates generated by resampling over loci.
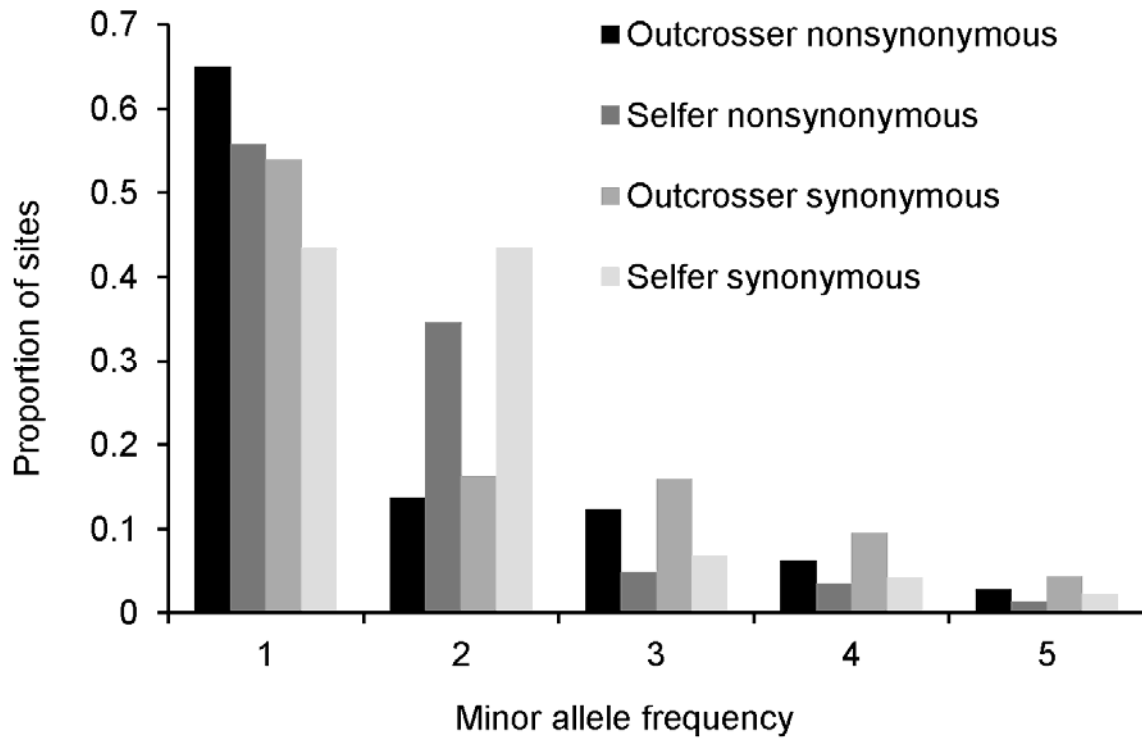
Figure S15 Folded nonsynonymous and synonymous allele frequency spectra for outcrossing and selfing populations of *Eichhornia paniculata*. We used haploid chromosomes from one individual each from 10 outcrossing and 10 selfing populations to generate the frequency spectra. Populations from the entire range of *E. paniculata* were sampled, with outcrossing populations from N.E. Brazil and selfing populations from the Caribbean and Central America.

R. Arunkumar *et al.*

Figure S16 Distribution of fitness effects (DFE) of new nonsynonymous mutations for outcrossing and selfing populations of *Eichhornia paniculata*. $N_e s$ is the product of N_e and the selection coefficient (s). Ten outcrossing samples from N.E. Brazil and 10 selfing samples from the Caribbean and Central America were used to generate the DFE. Error bars for each $N_e s$ category are 95% confidence intervals from 200 bootstrap replicates generated by resampling over loci. We used 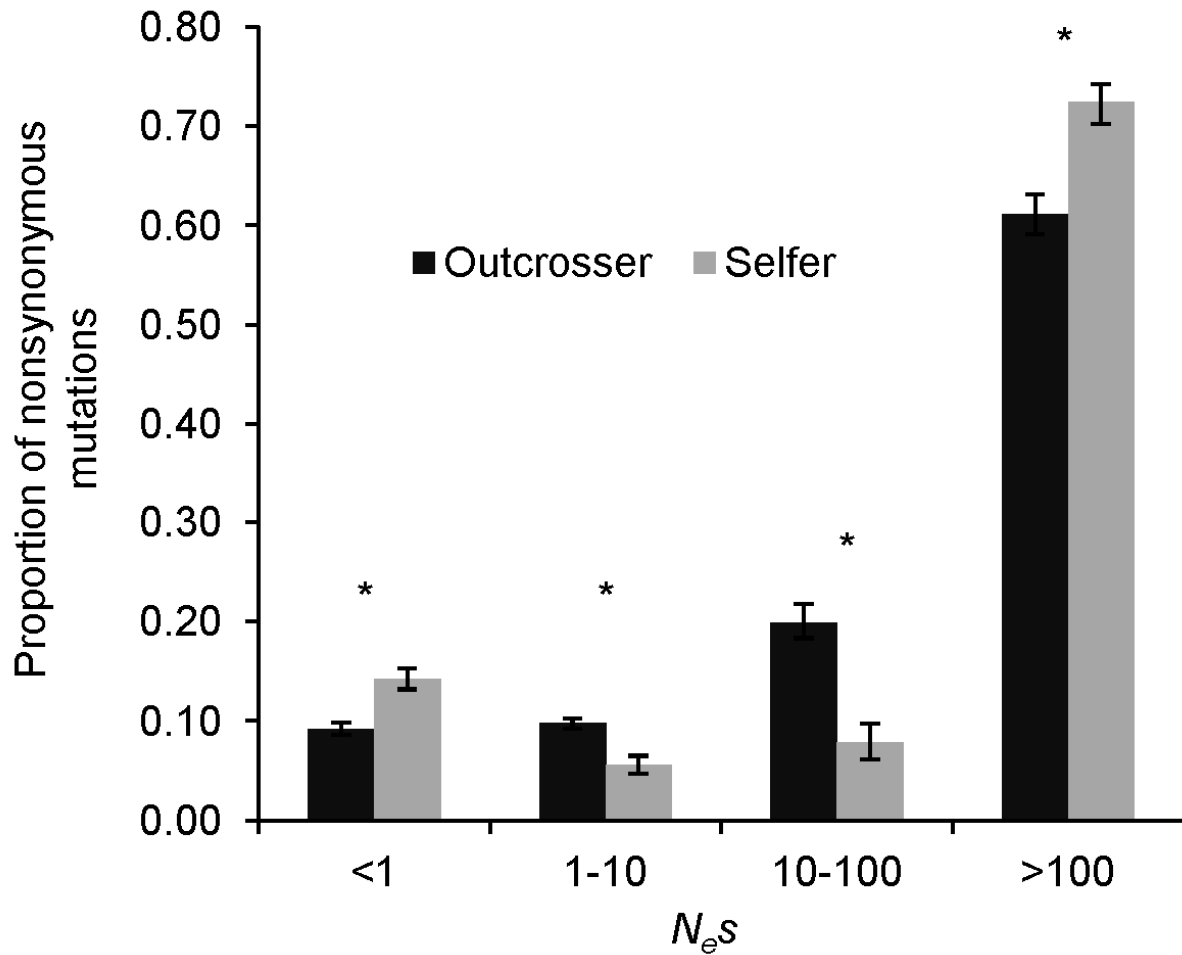a randomisation test (see Keightley and Eyre-Walker 2007) to compare outcrossing and selfing populations and to assess significance at 0.5% level (indicated by *).

Table S1   The sample of *Eichhornia paniculata* individuals used for estimating the efficacy of selection in outcrossing and selfing populations

Code	Country	Province/ Department	Nearest City	Morph structure	Morph[a]
B179	Brazil	Alagoas	Igaci	trimorphic	S
B191	Brazil	Pernambuco	Lajedo	trimorphic	M
B192	Brazil	Pernambuco	Cupira	trimorphic	L
B194	Brazil	Pernambuco	São Caetano	trimorphic	S
B197	Brazil	Pernambuco	Camocim de Sao Felix	trimorphic	S
B199	Brazil	Pernambuco	Caruaru	trimorphic	S
B202	Brazil	Ceará	Quixada	trimorphic	M
B207	Brazil	Ceará	Choro	trimorphic	L
B210	Brazil	Ceará	Canindé	trimorphic	L
B211	Brazil	Ceará	Fortaleza	trimorphic	M
C1	Cuba	Granma	Yara	monomorphic	M´
C3	Cuba	Granma	Cholera	monomorphic	M´
C5	Cuba	Las Tunas/ Camagüey	Camalote	monomorphic	M´
J29	Jamaica	St. Elizabeth	Fullerswood	monomorphic	M´
J30	Jamaica	St. Elizabeth	Cataboo	monomorphic	M´
J31	Jamaica	St. Elizabeth	Slipe	monomorphic	M´
J32	Jamaica	Westmoreland	Little London	monomorphic	M´
J33	Jamaica	Westmoreland	Georges Plain	monomorphic	M´
Mexico	Mexico	San Mateo del Mar, Nr.	Oaxaca	monomorphic	L´

R. Arunkumar *et al.*

		Tehuantepec			
Nicaragua	Nicaragua	Rivas	Rio Las Lajas	monomorphic	L´

[a]L=long-styled morph, L´=selfing variant of the L-morph, M=mid-styled morph, M´=selfing variant of the M-morph, S=short-styled morph. Data from Barrett *et al.* (2009) and Ness *et al.* (2010).

Table S2 Polymorphisms segregating in *Eichhornia paniculata* populations identified in aligned portions of EST-derived nuclear locus EP0314 from Barrett *et al*. (2009) and Ness *et al*. (2010) and the best matching contig from this study

Site	127	148	177	217	228	262	266	298	634	652	655	733
Brazil South selfer (*t*=0)[a]	C	C	G	C	C	G	A	C	C	A/C	C/T	C/?
Brazil South outcrosser (*t*>0.8)[a]	A/C	C/T	G	C/T	C/G	C/G/T	A/G	C/G	C/T/?	A/C/T/?	C/T/?	C/?
Brazil North outcrosser (*t*>0.8)[a]	C	C	G	C	C	G	A	C	C/T	C	C	C/?
Caribbean (*t*<0.15)[a]	A	C	A/G	C	C	T	G	C	T/C	C	C	C
Central America (*t*=0)[a]	C	C	G	C	C	G	A	C	C/?	C/?	C/?	C/?
Brazil_sample1[b]	C	C	G	C	C	G	A	C	C	C	C	C
Brazil_sample2[b]	C	C	G	C	C	G	A	C	C	C	C	C
Brazil_sample3[b]	C	C	G	C	C	G	A	C	T	C	C	C
Brazil_sample4[b]	C	C	G	C	C	G	A	C	C	C	C	C
Brazil_sample5[b]	C	C	G	C	C	G	G	G	C	T	C	C
Brazil_sample6[b]	C	C	G	C	C	G	A	C	C	C	C	C
Brazil_sample7[b]	C	C	G	T	C	C	G	G	C	C	C	C
Brazil_sample8[b]	A	C	A	C	C	G	G	C	T	C	C	C
Brazil_sample9[b]	C	C	G	C	C	G	A	C	C	C	C	C
Brazil_sample10[b]	C	T	G	T	C	C	A	C	C	T	C	C
Caribbean_sample1[b]	A	C	A	C	C	T	G	C	T	C	C	C
Caribbean_sample2[b]	A	C	A	C	C	T	G	C	T	C	C	C

R. Arunkumar *et al.*

Caribbean_sample3[b]	A	C	A	C	C	T	G	C	T	C	C	C
Caribbean_sample4[b]	A	C	A	C	C	T	G	C	T	C	C	C
Caribbean_sample5[b]	A	C	A	C	C	T	G	C	T	C	C	C
Caribbean_sample6[b]	A	C	A	C	C	T	G	C	T	C	C	C
Caribbean_sample7[b]	A	C	A	C	C	T	G	C	T	C	C	C
Caribbean_sample8[b]	A	C	A	C	C	T	G	C	T	C	C	C
Central America_sample1[b]	C	C	G	C	C	G	A	C	C	C	C	C
Central America_sample2[b]	C	C	G	C	C	G	A	C	C	C	C	G

[a]Polymorphisms identified in populations sampled in Barrett *et al.* (2009) and Ness *et al.* (2010)

[b]Polymorphisms identified in the 10 outcrossing and 10 selfing populations from this study

Table S3 **Nonsynonymous and synonymous diversity of simulated outcrossing and selfing populations under fixed dominance levels for all mutations**

Dominance of mutations (h)	Mating system	π_n	π_s	π_n/π_s
0.2	Outcrosser	1.63E-05	3.51E-05	0.46
0.2	Selfer	3.82E-06	8.75E-06	0.44
0.5	Outcrosser	1.36E-05	3.41E-05	0.40
0.5	Selfer	3.67E-06	8.88E-06	0.41
0.8	Outcrosser	1.24E-05	3.34E-05	0.37
0.8	Selfer	3.73E-06	8.73E-06	0.43

R. Arunkumar *et al.*

Table S4   *P*-values from a two-sample *t*-test comparing $N_e s$[a] categories of simulated outcrossing and selfing populations under fixed dominance levels for all mutations

Dominance of mutations (*h*)	<1	1-10	10-100	>100
0.2	0.257	0.479	0.002**	<0.001***
0.5	0.163	0.415	0.609	0.651
0.8	<0.001***	0.181	0.030*	0.275

[a]$N_e s$ is the product of N_e and the selection coefficient (*s*)

*Significant at the 5% level

**Significant at the 1% level

***Significant at the 0.1% level

Table S5 Two-way analysis of variance comparing the $N_e s$[a] categories of simulated outcrossing and selfing populations under varying dominance for mutations

Factor	Degrees of freedom	P-values			
		<1	1-10	10-100	>100
Mating system	1	0.049*	0.012*	<0.001***	<0.001***
Time (N)	5	0.011*	0.089	0.243	0.121
Mating system * Time (N)	5	0.011*	0.056	0.092	0.028*
Error	1428				

[a]$N_e s$ is the product of N_e and the selection coefficient (s)

*Significant at the 5% level

***Significant at the 0.1% level

Table S6 Counts of nonsynonymous and synonymous invariant sites and minor allele frequency spectra bins summed across

loci for eight outcrossing and eight selfing populations of *Eichhornia paniculata*

	Invariant	1	2	3	4
Outcrosser nonsynonymous	5205480	14392	3552	1702	742
Outcrosser synonymous	1438285	19738	6799	3933	1914
Selfer nonsynonymous	5177565	2478	561	402	172
Selfer synonymous	1454240	3562	918	910	340

Table S7   Number of loci containing sites heterozygous across multiple selfing populations from the Caribbean

Sites	Number of loci containing such sites
Heterozygous in two or more populations	245
Heterozygous in three or more populations	140
Heterozygous in four or more populations	113
Heterozygous in five or more populations	97
Heterozygous in six or more populations	87

R. Arunkumar *et al.*

Table S8 *P*-values from randomisation tests as applied in Keightley and Eyre-Walker (2007) comparing the $N_es<1$[a] category between DFEs of eight outcrossing and eight selfing *Eichhornia paniculata* populations under various filtering conditions

Comparisons	*P*-values
Original, Figure 6	0.190
Including sites with QUAL>30, Figure S10	0.190
Including sites with QUAL>0, Figure S10	0.267
Excluding 245 loci with sites heterozygous in >1 selfers, Figure S11	0.370
Excluding 73 loci with sites violating Hardy-Weinberg in outcrossers, Figure S12	0.250
Excluding 135 loci with $\pi_n/\pi_s>1$ in outcrossers or selfers, Figure S14	0.533

[a]N_es is the product of N_e and the selection coefficient (s)