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

Ivo M. Chelo and Henrique Teotónio

Instituto Gulbenkian de Ciência, Apartado 14, P-2781-901 Oeiras, Portugal

Correspondence to: Henrique Teotónio, Instituto Gulbenkian de Ciência, Apartado 14, P-2781-901 Oeiras, Portugal; phone: +351 214464684; Email: teotonio@igc.gulbenkian.pt

Running Head: Balancing Selection in C. elegans

The manuscript includes **Supporting Information** with 9 Figures (**S1-S9**) and 3 Tables (**S1-S3**)

Abstract

The role of balancing selection in maintaining genetic diversity during the evolution of populations in novel environments is poorly understood. To address this problem we study the impact of two mating systems, androdioecy and dioecy, on genotype diversity during the experimental evolution of *Caenorhabditis elegans*. We obtained the temporal trajectories of 334 single nucleotide polymorphisms (SNPs), covering 1/3 of the genome, to find extensive allele frequency changes and little loss of heterozygosities after 100 generations. As modeled in forward numerical simulations of experimental evolution, SNP differentiation was consistent with genetic drift and average fitness effects of 2%, if selection acted independently at each locus. Remarkably, inbreeding by self-fertilization was of little consequence to SNP differentiation. Again with simulations, we modeled selection on deleterious recessive alleles to find that it can explain the initial stages of evolution but not the later stages since lower heterozygosities would be maintained. In contrast, models with selection on overdominant loci could explain the heterozygote excess at all periods, in particular if there was negative epistasis or if loci had independent fitness effects. Potential balancing selection was little affected by the degree of self- and cross-fertilization. Overall these findings suggest that selection at single loci, including purging of recessive alleles, underlies most of the genetic differentiation accomplished. They further suggest that maintenance of genetic diversity in large populations can be due to balancing selection at multiple loci.

Introduction

The evolution of sexual populations, when in novel environments, usually depends on pre-existing diversity and on recombination to generate genotypes that can be maintained or lost e.g., (Wang et al. 1999; Meyer and Thomson 2001; Colosimo et al. 2005; Teotonio et al. 2009; Burke et al. 2010; Hancock et al. 2010; Turner et al. 2011). Since genetic drift and directional selection promote loss of diversity, and presumably mutation rates are not high enough to be significant during tens of generations (Hill 1982; Mackay et al. 1992; Christiansen et al. 1998; Denver et al. 2009), maintenance of diversity is typically thought to depend on density- and frequency-dependent dynamics because of population sub-division or because natural environments are temporally and spatially heterogeneous e.g., (Lenormand et al. 1999; Bradshaw and Holzapfel 2001; Grant and Grant 2002; Hanski and Saccheri 2006). Whether balancing selection, specifically selection on overdominant loci, might also underlie the maintenance of diversity when populations evolve in constant and homogeneous environments has long been investigated (Dobzhansky 1937; Lerner 1954; Lewontin 1974; Wright 1978; Barton 1990, 1995; Christiansen 2000; Coyne et al. 2000; Charlesworth 2006; Sellis et al. 2011) but empirical evidence remains equivocal (Hudson and Kaplan 1988; Takahata et al. 1992; Meyer and Thomson 2001; Kroymann and Mitchell-Olds 2005; Charlesworth 2006; Andres et al. 2010).

Detection of overdominant loci in random mating populations is inherently a difficult task. This is because random mating promotes even heterozygosity among individuals and, as a consequence, variation in fitness is reduced (Weir et al. 1980; Barton 1995; Christiansen 2000). With weak balancing selection, diversity will mostly follow the dynamics expected with neutrality (Charlesworth 2006). But even if balancing selection is strong it can be easily confounded with selection on deleterious recessive

alleles because in both there can be correlations between homozygosity states across the genome with reduced fitness (Ohta and Kimura 1970; Ohta 1971; Charlesworth 1991; Bierne et al. 2000). This phenomenon is more severe when recombination rates are low and linkage disequilibrium is strong (Ohta and Kimura 1970; Ohta 1971; Charlesworth et al. 1992; Palsson and Pamilo 1999), and/or when mating occurs between relatives and identity disequilibrium is strong (Charlesworth 1991; Bierne et al. 2000; Nordborg 2000). However, and despite the fact that inbreeding among relatives reduces the effective population sizes (Pollak 1987), purging of deleterious alleles is expected to be more effective as more homozygotes are also produced (Lande and Schemske 1985; Charlesworth et al. 1990; Nordborg et al. 1996). The fitness overdominance created by selection on deleterious recessives should be more transient under inbreeding than random mating. Inbreeding is thus predicted to affect not only the opportunity for balancing selection but, empirically, its manipulation can provide a way to detect overdominant loci maintaining diversity during short-term evolution.

Here we ask if balancing selection maintains genetic diversity when large populations evolve in a novel environment and, if so, how it depends on the population genetic structure imposed by two different mating systems. For this, we performed 100 generations of experimental evolution in *C. elegans* populations under dioecy or androdioecy and described at five time points their genotype frequency distributions at single nucleotide polymorphisms (SNPs) covering 1/3 of the genome.

It has been previously shown that during experimental evolution androdioecious populations had approximately 50% of their hermaphrodites undergo self-fertilization, with remaining ones outcrossing with males (Teotonio et al. 2012). Dioecious populations had obligatory cross-fertilization between males and females. Experimental genotype distributions are compared to those obtained in forward numerical simulations

modeling the demography of the two mating systems and the recombination rates among SNPs. We test for genetic drift alone or together with: selection at single loci, selection at multiple deleterious recessive alleles, or selection at multiple overdominant loci.

Materials and Methods

Experimental evolution design

Construction of the ancestral populations and experimental evolution design has been fully described elsewhere (Teotonio et al. 2012). The ancestral androdioecious population (A_0) resulted from a funnel cross among 16 highly inbred wild isolates. These wild isolates represent most of the known genetic differentiation found in nature (Rockman and Kruglyak 2009; Andersen et al. 2012) and when intercrossed manifest outbreeding depression (Johnson and Hutchinson 1993; Dolgin et al. 2007; Seidel et al. 2008; Teotonio et al. 2012). Each of the wild isolates was crossed in a pair-wise fashion to create two-isolate hybrids, which were subsequently crossed, also in a pair-wise fashion, to obtain the four-isolate hybrids, and so on until the final 16-isolate hybrid population. Equal nuclear and cytoplasm genome contributions between wild isolates were ensured with reciprocal crosses and large sample sizes during the derivation. After obtaining the final 16-isolate hybrid, over 10^5 individuals were frozen at -80ºC following standard protocols (Stiernagle 1999).

The ancestral dioecious population (D_0) was derived by the introgression of the *fog-2(q71)* allele (Schedl and Kimble 1988) into A_0. This allele disrupts spermatogenesis in hermaphrodites without apparent consequences in males. The *fog-2* locus is located at one of telomeres in chromosome V (genetic position: 24.92cM; *Wormbase WS220*). We followed a backcross design of F2 *fog-2(q71)* homozygous females to A_0 males and intercrossing of heterozygous individuals to again obtain homozygous *fog-2(q71)* F2s. This was repeated 9 times at large sample sizes for a total of 22 generations. The *fog-2(q71)* allele is estimated to segregate at a proportion of 10^{-4} in the derived D_0 population.

From each of the two hybrid ancestors, A_0 and D_0, three replicates for experimental evolution were derived (A_{1-3} and D_{1-3}) from revived -80^{0}C stocks of >10^4 individuals each.

For 100 generations, populations were cultured alongside at constant 20^{o}C and 80%RH, under discrete 4-day non-overlapping life-cycles (Teotonio et al. 2012). Each generation started by placing a synchronized cohort of first larval-staged individuals (L1s) at an estimated density of 10^3 in each of ten 9-cm Petri dishes with NGM-lite agar (US Biological) covered with a lawn of *Escherichia coli*, strain *HT114*. At each generation worms grew for 72±2h, feeding *ad libitum*, after which they were washed out of the Petri dishes, mixed per replicate, and subjected to a 1M KOH: 5%NaOCl solution for 5min. This procedure ensures that only eggs survive (Stiernagle 1999). 24±2h later individuals were collected as arrested L1s, after removal of dead larvae and adults, and seeded for the next generation at the appropriate densities (Teotonio et al. 2012). During growth the10 Petri dishes of each population were randomized across racks and shelves within a single incubator. Randomization was also followed with regards manipulation and experimenter across replicates and mating systems. During the experiment, the proportion of males were observed to be stably maintained at 25% in androdioecious populations, which means that 50% of the hermaphrodites self-fertilized at any given generation, as previously shown (Teotonio et al. 2012). Periodical storage of population samples at -80^{o}C was done.

DNA collection and genotyping

Samples were revived from -80^{o}C stocks (each with >10^3 individuals) and cultured for two generations under common conditions. In the third generation, 48 late

L4-staged (reproductively immature) hermaphrodites or females were collected: generation 0 (G0) populations were sampled after four and five years of storage; G10 and G30 populations were sampled after 4.5 years and 4 years of storage, respectively; and G70 and G100 populations after 3 years, and 6 months of storage, respectively. A total of 28 population samples were thus collected. The two G0 samples for each mating system were collapsed into one after verifying that no differences in the several heterozygosity statistics estimated here were found between them (not shown).

Individual genomic DNA was prepared with the ZyGEM prepGEM™ Insect kit following the manufacturer's protocol. Bi-allelic SNPs along chromosomes IV and X were chosen from the genome sequence of the N2, CB4856 and CB4858 wild isolates (*www.wormbase.org: WS220*; **Table S1**). Genotypes were mass determined with allele-specific extension reactions on oligonucleotides generated from PCR-amplified genomic DNA using the *iPlex Sequenom*™ MALDI-TOF platform (Bradic et al. 2011). A total of 468 genotyping runs were done within 9 months, with each run incorporating individuals from at least 2 different population samples.

Data quality control

As previously determined, polymorphic SNPs within the 16 parental wild isolates of the ancestral populations, which are supposed to be fully isogenic, were not considered (Teotonio et al. 2012). Quality control for the present study involved four steps. First, SNPs with more than 80% of missing data were excluded, when considering all individuals irrespective of population sample. This ensured that poor genotyping runs were ignored. Second, all individuals with more than 50% of missing SNPs genotypes were removed since for these individuals the DNA preparation was unsuccessful. Third,

for each population sample separately, individuals in the upper 5% of the frequency distributions of missing data were removed. This ensured removal of outliers if bimodal distributions of missing data were found, while being conservative otherwise. As a last fourth step, all SNPs polymorphic in less than 5 population samples were removed. The distribution of missing data in all individuals analyzed is shown in supporting materials **Figure S1** and within population sample correlations between missing data and individual heterozygosity in **Table S2** (see below for definition of individual heterozygosity). More stringent QC criteria did not change the results presented here (analysis not shown).

The data obtained after quality control encompasses 334 SNPs, for an average number of 42 genotypes per SNP for each of the G10, G30, G70 and G100 population samples, and average number of 88 genotypes per SNP for the G0 population samples. Details on sample sizes can be found in **Table S1**. A total of 187,432 genotypes at chromosome IV and 208,899 at chromosome X were available for analysis. Physical positions among SNPs followed *Wormbase WS220*. For chromosome IV SNPs are at densities of 1.4/100kb, and for chromosome X at densities of 2.2/100kb. Genetic positions among SNPs were obtained by linear interpolation for the two chromosomes defined each with map sizes of 50cM, according to (Rockman and Kruglyak 2009). The function *approx* in R was used for interpolation (**Table S1**).

Genetic diversity statistics

Heterozygosities: To prevent potential problems with sampling low frequency variants, SNPs were removed when their expected heterozygosity under random mating (Hardy-Weinberg) proportions (H_e) was below 0.05 within each population sample. This

corresponds to a probability of >0.98 in detecting alleles segregating at 5% for the average of 42 genotypes sampled in each population (Gibson and Muse 2002), p.271). A fixation index was calculated per SNP as F_{IS}= 1-(H_o/H_e), with H_o being the observed proportion of heterozygotes (Crow and Kimura 1970), p.66). Individual heterozygosity (H_i) was calculated as the proportion of heterozygous SNPs across both chromosomes within each individual. H_i estimates one minus the probability of identity between two SNP alleles within individuals, with high variances in H_i indicating high identity disequilibrium among multiple SNPs (Weir et al. 1980).

Linkage disequilibrium (LD): SNPs with minor allele frequencies <0.05 were removed prior to analysis to prevent bias due to the low power in sampling pair-wise SNP genotypes (Hill 1981). SNPs located in the telomeres were also removed as they were at complete LD (**Table S1**). LD was estimated as the composite genotype disequilibria, Δ, assuming that the genotype probabilities are the products of the gametic probabilities: $r^2=\Delta^2/p_a q_a p_b q_b$; with p and q being the proportions of the most and least common allele, respectively, of SNPs a and b (Weir 1996) p.95). We calculated the genetic distance at which background r^2 was reached for the six chromosome "domains" that have constant recombination rates, as previously defined in (Rockman and Kruglyak 2009) (see also **Table S1**). At each domain we first fitted polynomial regressions of 6th degree (using the *lm* function in R) and then defined the distance at which r^2 decay reached 5% of its initial value, as calculated by taking the first-derivative at points distanced every $5x10^{-3}$cM (see **Figure S2** for an example). Average distances among domains were calculated per replicate population.

Haplotype diversity: Multi-SNP diversity was estimated by phasing SNPs into haplotypes using *fastPHASE 1.2* (Scheet and Stephens 2006). For each population sample 20 random starts of the EM algorithm were employed with 200 haplotypes taken

from posterior distributions. The number of clusters for cross-validation was set to 10 and SNPs with posterior probabilities below 0.9 were considered missing data. Note that the estimation considers within population genetic structure and thus re-constructed individuals are diploids that contain two phased haplotypes. We used this haplotype structured data to seed the simulations of experimental evolution (see below). Effective number of haplotypes was calculated as $h_e=1/\Sigma p_i^2$, with p_i being the proportion of haplotype *i* across all haplotypes (Crow and Kimura 1970); pp.322-327).

Genetic differentiation

Differentiation of derived populations from ancestral populations: Data was characterized for SNP differentiation as the average allele frequency change among replicate populations at G100 from the ancestors at G0, separately by mating system. We employed generalized linear models (GLM) on SNP allele counts with the *logit* link function being used to model quasi-binomial error distributions. Significance was inferred with χ^2 tests on estimated deviances with one degree of freedom (Venables and Ripley 2002), p.187. Bonferroni correction was done with $\alpha=1.5\text{x}10^{-4}$ at each SNP, corresponding to an overall α=5%.

Differentiation of mating systems and replicate populations: We estimated mating system differentiation at each generation, by scanning chromosomes for differences in h_e with windows of 2, 5 and 10 SNPs and step sizes of 1 SNP along genetic position. h_e was right-censored and position was centered. Significant differentiation was inferred when the standard errors between the two mating systems did not overlap for a minimum of five consecutive positions. Variance component analysis was also done on the frequency of shared haplotypes among all populations (Excoffier et al. 1992), using the *amova*

function of the *ade4* package in R (Dray and Dufour 2007). Separately at each generation, random replicate populations were modeled within fixed mating systems.

Effective population sizes

To estimate the expected change of allele frequencies due to the random sampling of individuals across generations we inferred the variance effective population sizes (N_e) (Nei and Tajima 1981; Waples 1989; Barton 1995; Goldringer and Bataillon 2004). This was done at each mating system and for each chromosome separately, thus assuming that replicate populations share a common demography. Under random sampling, the observed allele frequency changes must follow a probability density distribution with mean zero (Goldringer and Bataillon 2004). For this reason we removed all GLM-differentiated SNPs prior to computation. The N_e estimated in this fashion reflects reductions in population sizes due to demography and natural selection at any or all of the replicates. Inclusion of all SNPs regardless of differentiation does not however significantly change the N_e estimated (analysis not shown).

We used the F_c statistic (Waples 1989) to compare the observed allele frequency changes with those obtained from multinomial sampling of alleles with population sizes drawn from a uniform distribution of $2N_e \sim U(50, 3x10^4)$. Each period of experimental evolution was modeled independently: from G0 to G10, G10 to G30, G30 to G70, and G70 to G100. First, "non-recombining" multi-allelic loci were defined, by phasing genotypes within regions where less than 5% of recombinants were expected, as given by the genetic distances in **Table S1**. Second, for each of 10^4 starting values of $2N_e$, drawn from the uniform distribution, allele frequency dynamics were obtained for the three replicate populations at each of the non-recombining loci (only those alleles with

starting frequencies between 5% and 95% were used). The expected F_c per $2N_e$ was calculated as the average among the three replicates. To obtain the probability density distributions of $2N_e$ at each non-recombining locus we identified the 500 runs resulting in the minimum expected F_c difference relative to the average F_c measured among replicates. For fitting we used the non-parametric *density* function in R with Gaussian kernels (Venables and Ripley 2002) p.126; see **Figure S3** for an example). Lastly, probability density distributions were multiplied across loci, separately per chromosome, with the final estimate of N_e taken as the maximum value ($\pm$ 1 $\log_{10}$) of the resulting distributions.

Simulations of experimental evolution

Monte-Carlo forward simulations modelled Wright-Fisher sampling processes in androdioecious and dioecious populations of constant diploid size N_e. Simulations were done from G0 to G100 with genetic drift, or for each of the periods sampled during the experiment when both genetic drift and selection were modelled. Simulated data used to calculate the several heterozygosity statistics was obtained by sampling genotypes in the same numbers as those of the experiment. Simulations were separately done for chromosomes IV and X. Each generation explicitly proceeded with the fertilization of gametes under dioecy or androdioecy, followed by viability selection on offspring and meiotic recombination in surviving adults.

To seed the simulations of each period being modelled, phased genotypes (see above) were sampled with replacement in order to generate populations composed of *Ne* individuals. This data was thus assumed to represent the pool of recombined gametes available for fertilization in the first generation.

Haplotypes are defined by vectors of SNP alleles ordered as in **Table S1**. At the start of each generation, fertilization was modelled with the random sampling of haplotypes irrespective of individual origin for 100% of cross-fertilization under dioecy and 50% of cross-fertilization under androdioecy. Self-fertilization events were therefore a possible outcome of sampling as there is no definition of sex. For androdioecious simulations, self-fertilization of 50% was modelled by the random sampling of individual phased genotypes (pairs of haplotypes). Cross- and self-fertilization rates were previously shown to be stable during experimental evolution at 100% under dioecy and 50% under androdioecy (Teotonio et al. 2012). Selection acted on the (offspring) genotypes as the probability of each being represented at the recombination stage as adults, while keeping total population size constant at *Ne*. Several functions were used to generate the probability weights of the selection process (see below). These weights were attributed to each genotype with the *sample* function in R, and sampling was done with replacement. Recombination in surviving genotypes was done by randomly choosing *Ne*/2 genotypes to undergo one crossover event each. This means that recombination is modelled on meiotic chromosomes of size 50cM and that there is complete crossover interference. Crossover positions were randomly placed in-between any two SNPs by following the probability distribution given by the genetic distances between them (**Table S1**). Crossover events proceeded by switching the downstream ordered set of SNP alleles between haplotype pairs. Recombined haplotypes were then used for fertilization and mating of the N_e adults starting the following generation.

Selection at single loci

We modeled allele selection at each SNP, from G0 and G100, separately by mating system. We used a simple urn model where the probability of a given allele to be passed

on to the following generation was given by its frequency plus a linear coefficient. Fitness was thus defined as: $w_i = 1 + s_i$, with s being the positive or negative selection coefficient of the N2 wild isolate allele i (**Table S1**). 100 simulations were done for each s taken from a uniform grid of 101 points from -0.3 to 0.3. The likelihood of each s was taken as the probability of allele counts observed at G100 in each replicate population, given the mean simulated SNP frequencies. We used the *dbinom* function in R for this purpose. Next, a compound probability was computed by multiplying the probability of each s. The maximum likelihood estimates of s were then tested for significance against zero, using likelihood ratio tests that assumed χ^2 error distributions with 3 degrees of freedom.

Selection at multiple loci

Simulations of experimental evolution were performed with two different selection models, as selection on deleterious recessive alleles or as balancing selection on overdominant loci. The parameter values explored covered similar fitness scales in both models. All SNPs, in both chromosomes and regardless of GLM differentiation, were considered. Models including only subsets of SNPs with regards to GLM-differentiation gave similar results (not shown).

For selection against partially to completely recessive alleles, fitness of a diploid genotype was defined as: $w = (1 - s)^x (1 - hs)^y$; with s and h being positive coefficients, x being the number of homozygous SNPs, and y being the number of heterozygous SNPs. With h=0 the deleterious allele is completely recessive. The deleterious allele was defined in GLM-differentiated SNPs as the allele decreasing in frequency from G0 to G100, and in GLM-undifferentiated SNPs as the alternative allele present in the non-N2

wild isolate, defined in **Table S1**. See (Charlesworth et al. 1990) for further details on the model.

For balancing selection, fitness of a diploid genotype was defined as: $w = 1 + \alpha H_i^k$; with α being the strength of selection and k being a dominance or epistasis coefficient. With two or more loci negative epistasis ($k<1$) implies diminishing-returns of fitness with increasing heterozygosity, while positive epistasis ($k>1$) implies synergism. With multiplicative epistasis ($k=1$) there is additive selection on heterozygosity. See (Navarro and Barton 2002) for further details on the model.

Comparing experimental and simulated data

The distances between experimental and simulated trajectories of H_e, F_{IS} or H_i were calculated as: $dist = \sqrt{\sum_G (h_o - h_s)_G^2}$; with h_o and h_s being the observed and simulated average heterozygosity statistics, respectively, for generations G=10, 30, 70, 100. Non-parametric smoothing of *dist* was used for drawing the figures across the model space parameterized, using the Gaussian kernels and other defaults of the *loess* function in R (Venables and Ripley 2002), pp.230-231).

Data archiving and software

Experimental genotype data is archived at *Dryad* (http://*datadryad.org/*) under **XXXX**. R statistical software was employed for all computations (R Development Core Team 2006). Scripts are available from the authors upon request.

1
2

1

2

3

Results

Genetic differentiation and effective recombination during experimental evolution

GLM-differentiated SNPs between generation 100 (G100) and ancestral populations (G0) were uniformly distributed across both chromosome IV and chromosome X (not shown), involving 38% of the SNPs under androdioecy and 25% of the SNPs under dioecy (**Figure 1**). Differentiation was mostly detected in SNPs with initially intermediate to high heterozygosity (H_o), and evolution did not lead to a great loss of diversity since monomorphic SNPs were few by G100 (4% with androdioecy and 92% with dioecy; **Figure 1**).

The extensive differentiation observed was not correlated with detectable changes in linkage disequilibrium (LD; **Figure 2A**). At G100, and irrespective of mating system, the LD between GLM-differentiated SNPs or between GLM-undifferentiated SNPs exponentially decays with the genetic distance separating them. When LD is calculated between SNPs that differentiated with others that did not we observed lower values until 0.5cM. This result only shows however the very uneven frequency changes between SNPs during experimental evolution. As expected with recombination, the genetic distance at which background LD is reached decreased during evolution at fairly constant rates (**Figure 2B**). Initially, mating systems were distinct with androdioecious populations having significantly higher LD than dioecious populations. This result is mostly due to the extra number of generations involved in the construction of the ancestral dioecious population (Teotonio et al. 2012). From G30 onwards both mating systems had similar LD. At the end of the experiment, background LD was found for genetic distances above 2cM with average r^2 values of ~0.02 (see also **Table S3**), which depending on chromosomal location corresponds to physical sizes above 0.26Mb-1.9Mb (**Table S1**).

Despite similar LD values between mating systems at the later stages of evolution, between G70 and G100, the rate of input of new genetic backgrounds was significantly higher under dioecy than androdioecy. This is shown by measuring the number of haplotypes in windows of 10 SNPs averaged at both chromosomes (**Figure 2C**; for analysis along the chromosomes see **Figure S4**). Initially, about 3% of the haplotypes were novel at each generation (with androdioecy producing more than dioecy), but by G70 this number had decreased to 1.5% in both mating systems. Androdioecious populations stabilized at this value unlike dioecious populations which rebounded and again generated 2% of new haplotypes in the later period of the experiment.

Little differentiation between mating systems was however evident in terms of the relative numbers of haplotypes segregating within each population (**Figure S5**). Although there were initial differences among the two ancestral populations, along both chromosomes and among periods, there was mostly homogeneity of h_e between mating systems. The major exceptions were one ~10cM/2.6Mb region in chromosome X, which showed complex dynamics during evolution but similar h_e values between mating systems by G100, and one ~8cM/3Mb region in chromosome IV that showed much higher h_e values under dioecy than androdioecy by G100. Variance component estimates of differentiation on shared haplotypes further confirmed overall similar responses among mating systems even if significant differentiation was achieved among replicate populations (**Figure S6**).

Expected SNP diversity with genetic drift and selection at single loci

SNP allele frequency changes with experimental evolution can be explained by genetic drift and selection at single loci. To show this we first estimated the expected change in allele frequencies due to the random sampling of individuals across generations as the variance effective population size (N_e). Results from this analysis indicate that on average about 10^3 individuals reproduced at each generation, regardless of mating system (**Figure S7**). There is thus no evidence for population growth bottlenecks or expansions during experimental evolution.

Simulations of experimental evolution from G0 to G100, modeling genetic drift, show that the expected distributions of SNP allele frequency change differ from those observed (**Figure 3**, panels **A** and **B**). In particular, 17% of the SNPs under androdioecy and 28% of the SNPs under dioecy had frequency changes above those expected under neutrality. Of these, 54 SNPs under androdioecy and 75 SNPs under dioecy were also GLM-differentiated from G0 to G100 (not shown).

We next asked how selection at single loci would be consistent with both estimated GLM-differentiation and estimated deviations from expected SNP allele frequency changes under neutrality. Modeling the simulations of experimental evolution with genetic drift and selection in favor or against of a reference SNP allele illustrates that with mean fitness effects of about 2% per generation the observed GLM-differentiation can be achieved with the smaller effect inferred at 0.6% and the largest at 8.4% (**Figure 3C**). Importantly, the distributions of fitness effects do not differ among mating systems, which reveals that genetic drift and potential selection at single loci had similar sampling consequences on SNP differentiation under both dioecy and androdioecy.

Genetic drift cannot however explain the evolution of how mating systems maintained significant population genetic structure. This is revealed by comparing

observed data with the results of neutral simulations done for each evolutionary period: from G0 to G10, G10 to G30, G30 to G70, and G70 to G100. Specifically, the Hardy-Weinberg heterozygosity (H_e) of all SNPs declined with time at slower rates than expected with neutrality under both mating systems (**Figure 4A**). Dioecious populations at G100, but particularly androdioecious populations from G30 onwards, maintained high H_e. Fixation indices (F_{IS}) show a similar pattern (**Figure 4B**). For androdioecy, initially very high F_{IS} rapidly decreased in the first 10 generations, and from then on either maintained or increased to values of about 0.1. This is quite unlike the expected neutral F_{IS} obtained with the simulations, which stabilize at average values of 0.3. For dioecy, F_{IS} was stable throughout evolution as predicted by the neutral simulations. Finally, individual heterozygosities (H_i) were also maintained at higher values than expected with genetic drift (**Figure 4C**). Androdioecious individuals were on average heterozygous at more SNPs than expected, while differences among them (the variance of H_i distributions) were found to be lower than expected with genetic drift throughout the experiment (**Figure S8**). Dioecious populations had also excess heterozygosity at more loci, with H_i being lost at lower rates than expected.

Similar results are obtained for androdioecy, but not dioecy, when comparing observed pair-wise SNP LD with those from simulations of genetic drift. In particular, the background r^2 values were lower than expected at several periods (**Table S3**). Moreover, the background values of r^2 are lower than expected at large distances of >2cM, since predicted self-fertilization rates under androdioecy are higher than those observed during experimental evolution (**Table S3**).

The evolution of population genetic structure and selection on partially-dominant loci

We tested if the evolution population genetic structure was consistent with fitness overdominance emerging indirectly from selection at multiple loci each with deleterious recessive alleles. For this, simulations of experimental evolution were done independently for each period and considering all SNPs. Simulated and experimental data is compared for H_e, F_{IS} and H_i (results for CV(H_i) are not shown but reveal similar dynamics).

For androdioecy, simulations of selection on deleterious alleles show that H_e is maintained at lower values than that observed for experimental H_e, regardless of dominance, for all generations after G10 (**Figure 5A**). Even in models with completely recessive alleles (h=0) and weak selection of s=0.001 less H_e is expected. In contrast to H_e, simulated fixation indices (F_{IS}) are consistent with the observed F_{IS} when $h \leq 0.1$ and s=0.05 (**Figure 5B**). In particular, the rapid decrease in F_{IS} observed at G10, followed by a slow increase in remaining generations, is closely matched by the simulated distributions. However, for weaker selection or higher dominance coefficients, high F_{IS} values are always maintained. Simulation results for average individual heterozygosity (H_i) also indicate that it is maintained at lower values than that observed during the experiment (**Figure 5C**). For example, with h=0.1 and s=0.05, an excellent fit can be found at G10 but in subsequent generations H_i is rapidly lost.

A summary of the differences between observed and simulated values for all periods and all s and h space parameterized is presented in **Figure 6**. For androdioecy, this figure indicates that the expected H_e with selection is to some extent consistent with the observations made at all generations only when s<0.01 (**Figure 6A**). For F_{IS} considerably different parameter ranges are found in order to have a fair match between the simulated and the experimental data (**Figure 6B**). In this case, the best fits are for s>0.05, above which there should be positive correlation between s and h for

experimental F_{IS} to be explained. For H_i, the simulation results have the poorest agreement with the experimental trajectories (**Figure 6C**). When comparing together the three estimates of heterozygosities, there is not a feasible combination of s and h for which simulated distributions closely explain experimental distributions.

With dioecy, modeling selection on deleterious recessives gives qualitatively similar results to those of androdioecy when considering H_e or H_i, but not F_{IS} (**Figure 6DEF**). Under dioecy there is agreement across all statistics for $0.01<s<0.02$ and $h<0.1$. In constrast to androdioecious simulations, F_{IS} results now place expected s and h at lower values while the positive correlation between them is no longer apparent. For H_i, models under dioecy have also better fit than those under androdieocy. Note though that models of genetic drift without selection are usually sufficient to explain maintenance of all heterozygosities during dioecious evolution (as also shown in **Figure 4**).

The evolution of population genetic structure and selection on overdominant loci

Contrary to models of selection on deleterious recessive alleles, models with selection on overdominant loci reveal a range in parameter space that agrees well not only among the three heterozygosity statistics estimated but also among both mating systems. A summary of the results is shown in **Figure 7**, while illustrative examples of androdioecious trajectories are shown in supporting **Figure S9**. In general, and for both mating systems, simulated H_e values closely follow experimental observations with negative epistasis of $k<1$ and selection of any strength (**Figure 7AD**, **Figure S9A**). With additive selection on heterozygosity ($k=1$) or with positive epistasis ($k>1$) simulated H_e usually evolves to considerably higher values than those measured during the experiment, with only weak selection giving the best matches. Differences between

simulated and observed H_e data are however very small and thus this statistic contains little useful information.

Comparisons made with F_{IS} are more informative. In this case, models under androdioecy suggest the existence of negative epistasis (k<1) for selection strenghts of α>3 although the relationship between k and α does not appear to be linear (**Figure 7B**). For k≥1 simulation results predict positive F_{IS} for weak selection or, alternatively, predict negative F_{IS} for strong selection (see **Figure S9B**). Under dioecy, the parameter range for selection is shifted and enlarged relative to androdioecy to also include negative epistasis or additive selection on heterozygosity of any strength and positive epistasis at α<1 (**Figure 7E**). This difference between mating system models is apparent as well when considering only the later stages of evolution, from G70 to G100 (not shown).

Finally, similar results to those of F_{IS} are obtained when considering H_i (**Figure 7CF**). Now though the inference of selection for k≤1 with 3<α<6 is stronger under androdioecy. Further, for positive epistasis, dioecious models are not as good as those obtained when estimating F_{IS}. Regardless of mating system, for simulations with α<3 low H_i values are maintained, in particular with negative epistasis, and for α>6 a higher H_i values are maintained than those which were observed during evolution. Note as well that for most of the parameter space with positive epistasis close to complete linkage disequilibrium would be reached in several regions across the genome (see also **Figure S9C** and results not shown). As with selection on deleterious recessives, dioecious simulations with selection on overdominant loci do not explain much better the observed heterozygosities than neutral models.

1 25
2

1

Discussion

Recent experimental studies have established that the evolution of populations with standing diversity does not lead to a great loss in heterozygosity during tens to hundreds of generations (Nuzhdin et al. 2007; Teotonio et al. 2009; Burke et al. 2010; Turner et al. 2011). Further, signs of selective sweeps (Hill and Robertson 1966; Barton 1998; Kim and Stephan 2002; Innan and Kim 2004; Hermisson and Pennings 2005), whereby large regions have reduced heterozygosity because neutral alleles hitch-hike together with beneficial alleles in linkage disequilibrium (LD), do not appear to be common. Instead, loss of heterozygosity is typically partial and local, suggesting that the loci underlying short-term evolution in novel environments start from intermediate allele frequencies and/or have small fitness effects. Similarly, our results indicate that widespread SNP differentiation was not accompanied by loss of heterozygosity (**Figure 1**), that high pair-wise LD did not correlate with allele frequency changes even for small genetic distances between SNPs (**Figure 2**), and that the expected fitness effects of multiple loci, if evolving independently of each other, were not large (**Figures 3**). Surprisingly, the two mating systems had only minor effects on SNP allele frequency changes, in particular when one considers that under androdioecy, at any given generation, 50% of the individuals reproduced by selfing (Teotonio et al. 2012).

For comparable census sizes and initial standing diversity, inbreeding by selfing was expected to greatly reduce the effective population sizes (N_e) when compared with random mating (Pollak 1987). Reduced N_e should have resulted in less selection efficacy and thus the estimated average fitness effects were expected to have been greater under androdioecy than dioecy if similar SNP differentiation was achieved. The poor resolution of the estimated N_e (**Figure S3**) was not however a reason for not detecting the effects of asssortative mating, since the simulations with genetic drift accurately predicted a

differential loss of Hardy-Weinberg heterozygosity between mating systems (**Figure 4A**). The remarkable result is then that, particularly under androdioecy, observed heterozygosity was clearly in excess given the degree of self- and cross-fertilization realized during the experiment (**Figure 4B**). Furthermore, not only linkage disequilibria (**Table S3**) but also identity disequilibria appeared to be reduced, as individuals were both less homozygous across SNPs and less dissimilar between themselves (**Figures 4C** and **S8**). Together, SNP differentiation at multiple loci was accompanied by excess heterozygosity at multiple loci.

These phenomena of excess heterozygosity during short-term experimental evolution have been measured before (Rumball et al. 1994; Latter 1998; Porcher et al. 2004) and they all suggest that fitness overdominance can be generated from standing diversity. The work of M. Clegg and colleagues in the 1970s, employing *Drosophila melanogaster* populations, is particularly notable since heterozygosity deviations from neutrality were observed under obligatory cross-fertilization, a condition which in our experimental system did not lead to evident heterozygote excess. Specifically, in experiments with populations constructed to bear little ancestral LD at the genome-wide level but complete LD among two or three allozyme markers, M. Clegg and colleagues found that marker LD initial decreased with time considerably faster than predicted with genetic drift (Clegg 1978; Clegg et al. 1980). Similarly, in experiments testing for the consequences of selection against lethal alleles on marker loci, and where the whole genome was constructed to be in complete LD, hitch-hiking with the lethals was observed during the first stages of evolution (Clegg et al. 1976; Clegg 1978; Clegg et al. 1978). Subsequently, markers recovered their heterozygosity at rates proportional to the recombination rate. Also in these experiments, heterozygote defficiency at the marker closest to the lethals were expected under random mating while the marker farthest from the lethals indicated that populations might have been under negative assortative mating,

in a manner closely reminiscent of our dioecious populations (see **Figure 4B**). Overall, heterozygote excess was detectable because there was initial even heterozygosity and also strong LD between markers.

Like us, M. Clegg and colleagues employed numerical simulations of experimental evolution to test for selection on partially-dominant loci and selection on overdominant loci. They were not successful in this in part because the expected distributions of heterozygosities and LD were poorly contrasted to the experimental observations with just a few markers. It has since been recognized however that the heterozygote excess found might have been due to the allozyme markers being overdominant loci themselves. Some of the markers are now known to mediate trade-offs between *D. melanogaster* life-history stages that can generate fitness overdominance by antagonistic pleiotropy, even in experimental populations of constant sizes and little age structure during evolution (Rose 1982; Deckert-Cruz et al. 1997; Teotonio et al. 2009). There is obviously no equivalent information for the SNPs measured here and the assumption of marker neutrality might not be met (**Table S1**). But since a large set of SNPs were followed during experimental evolution, and SNP differentiation was widespead across both chromosomes IV and X, it is already possible to distinguish alternative selection scenarios during evolution.

Between G0 and G10, androdioecious populations rapidly evolved heterozygosities consistent with selection on deleterious recessive alleles (**Figure 5**). These findings were expected as segregation of standing diversity was surely to result in strong selection if the variance in individual heterozygosity and LD were high cf. (Bierne et al. 2000), as they were during initial androdioecious evolution (**Figures 2**, **4C** and **S8**). Interactions between multiple loci might have also generated the necessary fitness overdominance leading to excess heterozygosities (Charlesworth and Barton 1996). This

is because ancestral populations were derived by the hybridization of 16 different wild isolates and it is known that hybrids of *C. elegans* wild isolates show fitness depression (Johnson and Hutchinson 1993; Dolgin et al. 2007), which in part could be due to alleles encompassing multiple functional loci (Gloria-Soria and Azevedo 2008; Seidel et al. 2008; Ghosh et al. 2012). If recombinants of these multiple loci were generated, and were not purged during the derivation of the ancestral populations (Teotonio et al. 2012), they could have contributed to initial fitness overdominance.

Unfortunately, comparison of the two mating systems does not clarify the relative role of segregation and recombination in the initial purging of deleterious alleles at single and multiple loci. The ancestral dioecious populations was expected to have high effective recombination rates and strong selection on multiple loci (Charlesworth and Barton 1996), because of high heterozygosities, but since its derivation involved more generations than the derivation of the ancestral androdioecious population most of the deleterious alleles might have been purged by G0 (Teotonio et al. 2012).

Regardless of mating system consequences, there was marked reduction in LD during the 100 generations of evolution (**Figure 2**), which suggests that many deleterious alleles could continue to be generated and efficiently purged, even if undetectable with our analysis. If so, selection on deleterious alleles is unlikely to have generated fitness overdominance during the remainder of the experiment because alleles that were not purged in the initial generations would have been kept at low frequencies. In an analogous fashion, mutational input of deleterious recessive alleles needed to be very high to generate sufficient fitness overdominance (Charlesworth et al. 1993; Charlesworth et al. 1995; Bierne et al. 2000). Note that deleterious alleles created by mutation needed to have very small selection coefficients and behave neutrality during most evolution ($s<1/N_e$; (Kimura 1983)) to reach sufficiently high frequencies in the later

periods. We have found no such evidence, since neither models with weak selection had good fits with the observations at the later periods (**Figures 5** and **6**), nor there was an increase of excess heterozygosity with time as expected with mutation accumulation (**Figure 4**).

Balancing selection might have caused the observed SNP dynamics between G0 and G10, in particular if there was negative epistasis between overdominant loci (Christiansen 2000; Navarro and Barton 2002). However, selection on deleterious recessives likely overwhelmed balancing selection, if it existed, because during this period an increase in the mean individual heterozygosity was accompanied with an obvious decrease in Hardy-Weinberg heterozygosity. With balancing selection an initially much higher increase in Hardy-Weinberg heterozygosity would have been expected than that observed (compare **Figure 4A** with **Figure S9A**).

From G10 onwards, or soon after, evolution should have involved balancing selection as simulations with overdominant loci provide a good fit with the evolution of population genetic structure in both mating systems (**Figures 7** and **S9**). As for selection on deleterious alleles however, it is difficult to understand if there were consequences of the mating system to balancing selection. This not only because LD was similar between androdioecy and dioecy after G30, but also because the SNP differentiation or haplotype differentiation achieved at G100 was mostly a function of time (**Figures 1**, **S5**, **S6**). Nonetheless, it is interesting to note that from G70 to G100 dioecy appears to have increased the effective recombination rates among multiple loci because the number of new haplotypes generated during this period was higher than those under androdioecy (**Figure 2**). Further, correlated with these late haplotype dynamics, models with negative epistasis or additive selection on heterozygosity also had better fits under dioecy than androdioecy (**Figure 7**), which is consistent with the idea that negative epistastic

balancing selection promotes maintenance of genetic diversity during evolution (see for example Figure 5 in (Navarro and Barton 2002)).

Contingent on the number of fitness loci and the recombination rates between them, diversity at neutral markers can be maintained or lost (Lewontin and Kojima 1960; Christiansen 2000; Kelly and Wade 2000; Navarro and Barton 2002). To illustrate the complex relations between recombination and selection at multiple loci we finish by asking if the mating system differences in effective haplotype numbers found in one outstanding ~8cM/3Mb region in chromosome IV (see **Figure S5**) could have resulted from overdominant selection. In this region, the several heterozygosities estimated above and pair-wise SNP LD were mostly indistinguishable between mating systems, and further, SNPs did not significantly differentiate between G0 and G100 (results not shown). Even so, between G70 and G100, dioecious populations doubled in their numbers of haplotypes relative to androdioecious populations.

We modeled selection as before but controlled the number of overdominant loci (from 2 to 10) and their relative genetic position to the peak of haplotype diversity observed under dioecy at G100. Here we only present a preliminary summary of these analyses: 1) selection maintains observed diversity from G70 to G100 under androdioecy only if there is negative epistasis or additive selection on heterozygosity, irrespective of the numbers of loci; 2) with positive epistasis there is a great loss of haplotype diversity, the extent of which is more severe with increasing numbers of loci, despite of mating system; 3) selection always leads to less haplotypes under dioecy than androdioecy, regardless of epistasis or numbers of loci, and thus to less diversity than that observed at G100; 4) models of genetic drift alone or with selection on multiple deleterious recessives always predicted less diversity than observed at either mating system (and always less diversity than selection on overdominant loci). Together, these results

suggest that increased effective recombination weakens balancing selection even if it is necessary to maintain haplotype diversity.

How then can the evolution of haplotype diversity be reconciled with the evolution of effective recombination rates? Besides the obvious explanation that our proxy for effective recombination is not adequate because it is not independent of how diversity is estimated, a more likely explanation is that recombination modifiers appeared between G70 and G100, for example by gene conversion. Our assumption of fixed recombination rates between SNPs during evolution would then be incorrect. Gene conversion possibly occurs at high rates in *C. elegans* (Semple and Wolfe 1999; Katju et al. 2008) and it could generate genomic re-arrangements which in turn could result in either a suppression or an enhancement of recombination rates. Why gene conversion would be more predominant under one mating system over another, and only at the later stages of evolution, is nonetheless puzzling.

In conclusion, our findings suggest that selection at single loci, including purging of recessive alleles resulting from hybridization, can underlie most genetic differentiation when large sexual populations are faced with novel environments. More significantly, they also suggest that balancing selection is necessary to explain the maintenance of genetic diversity for periods when mutation is expected to have a minor role.

Acknowledgements

We thank S. Carvalho and J. Costa for assistance with genotyping. We thank R. Azevedo, F. Christiansen, A. Cutter, M. Lynch, S. Proulx, M. Rockman, H. Schulenburg and D. Weissman for discussion. N. Barton, J. Hermisson and two anonymous reviewers drew our attention to relevant references and contributed to both the conceptual clarity and the

1presentation of this work. Funding from the *European Research Council*, grant

2stERC/2009-243285 (to H.T).

3

4

Literature Cited

Andersen, E. C., J. P. Gerke, J. A. Shapiro, J. R. Crissman, R. Ghosh, J. S. Bloom, M. A. Felix, and L. Kruglyak. 2012. Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic diversity. Nat Genet 44:285-290.

Andres, A. M., M. Y. Dennis, W. W. Kretzschmar, J. L. Cannons, S.-Q. Lee-Lin, B. Hurle, N. C. S. Program, P. L. Schwartzberg, S. H. Williamson, C. D. Bustamante, R. Nielsen, A. G. Clark, and E. D. Green. 2010. Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. PLoS Genetics 6:e1001157.

Barton, N. H. 1990. Pleiotropic models of quantitative variation. Genetics 124:773-782.

Barton, N. H. 1995. Linkage and the limits to natural selection. Genetics 140:821-841.

Barton, N. H. 1998. The effect of hitch-hiking on neutral geneologies. Genet Res 72:123-133.

Bierne, N., A. Tsitrone, and P. David. 2000. An inbreeding model of associative overdominance during a population bottleneck. Genetics 155:1981-1990.

Bradic, M., J. Costa, and I. M. Chelo. 2011. Genotyping with Sequenom *in* V. Orgogozo, and M. Rockman, eds. Molecular Methods for Evolutionary Genetics. Humana Press, New York.

Bradshaw, W. E., and C. M. Holzapfel. 2001. Genetic shift in photoperiodic response correlated with global warming. Proc Natl Acad Sci U S A 98:14509-14511.

Burke, M. K., J. P. Dunham, P. Shahrestani, K. R. Thornton, M. R. Rose, and A. D. Long. 2010. Genome-wide analysis of a long-term evolution experiment with Drosophila. Nature 467:587-590.

Charlesworth, B., and N. H. Barton. 1996. Recombination load associated with selection for increased recombination. Genet Res Camb 67:27-41.

Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289-1303.

Charlesworth, D. 1991. The apparent selection in neutral loci in partially inbreeding populations. Genet Res 57:159-175.

Charlesworth, D. 2006. Balancing selection and its effects on sequences in nearby genome regions. PLoS Genetics 2:e64.

Charlesworth, D., B. Charlesworth, and M. T. Morgan. 1995. The pattern of neutral molecular variation under the background selection model. Genetics 141:1619-1632.

Charlesworth, D., M. T. Morgan, and B. Charlesworth. 1990. Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. Evolution 44:1469-1489.

Charlesworth, D., M. T. Morgan, and B. Charlesworth. 1992. The effect of linkage and population size on inbreeding depression due to mutational load. Genet Res 59:49-51.

Christiansen, F. B. 2000. Population Genetics of Multiple Loci. John Wiley & Sons, Ltd., New York.

Christiansen, F. B., S. P. Otto, A. Bergman, and M. W. Feldman. 1998. Waiting with and without recombination: the time to production of a double mutant. Theo Pop Biol 53:199-215.

Clegg, M. T. 1978. Dynamics of correlated genetic systems. II. Simulation studies of chromosomal segments under selection. Theor Popul Biol 13:1-23.

Clegg, M. T., J. F. Kidwell, and C. R. Horch. 1980. Dynamics of correlated genetic systems. V. Rates of decay of linkage disequilibria in experimental populations of Drosophila melanogaster. Genetics 94:217-234.

Clegg, M. T., J. F. Kidwell, and M. G. Kidwell. 1978. Dynamics of correlated genetic systems. III. Behaviours of chromosomal segments under lethal segments. Genetica 48:95-106.

Clegg, M. T., J. F. Kidwell, M. G. Kidwell, and N. J. Daniel. 1976. Dynamics of correlated genetic systems. I. Selection in the region of the Glued locus of Drosophila melanogaster. Genetics 83:793-810.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. Villarreal, Jr., M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, and D. M. Kingsley. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science 307:1928-1933.

Coyne, J. A., N. H. Barton, and M. Turelli. 2000. Is Wright's shifting balance process important in evolution? Evolution 54:306-317.

Crow, J. F., and M. Kimura. 1970. An Introduction to Population Genetics Theory. Harper & Row, Publishers, New York.

Deckert-Cruz, D. J., R. H. Tyler, J. E. Landmesser, and M. R. Rose. 1997. Allozymic differentiation in response to laboratory demographic selection of Drosophila melanogaster. Evolution 51:865-872.

Denver, D. R., P. C. Dolan, L. J. Wilhelm, W. Sung, J. I. Lucas-Lledo, D. K. Howe, S. C. Lewis, K. Okamoto, W. K. Thomas, M. Lynch, and C. F. Baer. 2009. A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proceedings of the National Academy of Sciences of the United States of America 106:16310-16314.

Dobzhansky, T. 1937. Genetics and the Origin of Species. Columbia University Press, New York.

Dolgin, E. S., B. Charlesworth, S. E. Baird, and A. D. Cutter. 2007. Inbreeding and outbreeding depression in Caenorhabditis nematodes. Evolution 61:1339-1352.

Dray, S., and A. B. Dufour. 2007. The ade4 package: Implementing the duality program for ecologists. J Stat Software 22:1-20.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Ghosh, R., E. C. Andersen, J. A. Shapiro, J. P. Gerke, and L. Kruglyak. 2012. Natural variation in a chloride channel subunit confers avermectin resistance in C. elegans. Science 335:574-578.

Gibson, G., and S. Muse. 2002. A Primer of Genome Science. Sinauer Associates, Inc. Publishers, Sunderland.

Gloria-Soria, A., and R. B. Azevedo. 2008. npr-1 Regulates foraging and dispersal strategies in Caenorhabditis elegans. Curr Biol 18:1694-1699.

Goldringer, I., and T. Bataillon. 2004. On the distribution of temporal variations in allele frequency: Consequences for the estimation of effective population size and the detection of loci undergoing selection. Genetics 168:563-568.

Grant, P. R., and B. R. Grant. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. Science 296:707-711.

Hancock, A. M., D. B. Witonsky, E. Ehler, G. Alkorta-Aranburu, C. Beall, A. Gebremedhin, R. Sukernik, G. Utermann, J. Pritchard, G. Coop, and A.

Di Rienzo. 2010. Colloquium paper: human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. Proc Natl Acad Sci U S A 107 Suppl 2:8924-8930.

Hanski, I., and I. Saccheri. 2006. Molecular-level variation affects population growth in a butterfly metapopulation. PLoS Biol 4:e129.

Hermisson, J., and P. S. Pennings. 2005. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics 169:2335-2352.

Hill, W. G. 1981. Estimation of effective population size from data on linkage disequilibrium. Genet Res 38:209-216.

Hill, W. G. 1982. Rates of change in quantitative traits from fixation of new mutations. Proc Natl Acad Sci U S A 79:142-145.

Hill, W. G., and A. Robertson. 1966. The effect of linkage on limits to artificial selection. Genet Res 8:269-294.

Hudson, R. R., and N. L. Kaplan. 1988. The coalescent process in models with selection and recombination. Genetics 120:831-840.

Innan, H., and Y. Kim. 2004. Pattern of polymorphism after strong artificial selection in a domestication event. Proc Natl Acad Sci U S A 101:10667-10672.

Johnson, T. E., and E. W. Hutchinson. 1993. Absence of strong heterosis for life span and other life history traits in Caenorhabditis elegans. Genetics 134:465-474.

Katju, V., E. M. LaBeau, K. J. Lipinski, and U. Bergthorsson. 2008. Sex change by gene conversion in a Caenorhabditis elegans fog-2 mutant. Genetics 180:669-672.

Kelly, J. K., and M. J. Wade. 2000. Molecular evolution near a two-locus balanced polymorphism. J Theor Biol 204:83-101.

Kim, Y., and W. Stephan. 2002. Detecting a local signature of genetic hitchhiking along a recombining chromosome. Genetics 160:765-777.

Kimura, M. 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge.

Kroymann, J., and T. Mitchell-Olds. 2005. Epistasis and balanced polymorphism influencing complex trait variation. Nature 435:95-98.

Lande, R., and D. W. Schemske. 1985. The evolution of inbreeding depression and selfing in plants. I. Genetic models. Evolution 39:24-40.

Latter, B. D. 1998. Mutant alleles of small effect are primarily responsible for the loss of fitness with slow inbreeding in Drosophila melanogaster. Genetics 148:1143-1158.

Lenormand, T., D. Bourguet, T. Guillemaud, and M. Raymond. 1999. Tracking the evolution of insecticide resistance in the mosquito Culex pipiens. Nature 400:861-864.

Lerner, I. M. 1954. Genetic Homeostasis. Oliver and Boyd Edinburgh.

Lewontin, R. C. 1974. The Genetic Basis of Evolutionary Change. Columbia University Press, New York.

Lewontin, R. C., and K. Kojima. 1960. The evolutionary dynamics of complex polymorphisms. Evolution 14:458-472.

Mackay, T. F. C., R. F. Lyman, M. S. Jackson, C. Terzian, and W. G. Hill. 1992. Polygenic mutation in Drosophila melanogaster: estimates from divergence among inbred strains. Evolution 46:300-316.

Meyer, D., and G. Thomson. 2001. How selection shapes variation of the human major histocompatibility complex: a review. Ann Hum Genet 65:1-26.

Navarro, A., and N. H. Barton. 2002. The effects of multilocus balancing selection on neutral variability. Genetics 161:849-863.

Nei, M., and F. Tajima. 1981. Genetic drift and the estimation of effective population size. Genetics 98:625-640.

Nordborg, M. 2000. Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. Genetics 154:923-929.

Nordborg, M., B. Charlesworth, and D. Charlesworth. 1996. Increased levels of polymorphim surrounding selectively maintained sites in highly selfing species. Proc Biol Sci 263:1033-1039.

Nuzhdin, S. V., L. G. Harshman, M. Zhou, and K. Harmon. 2007. Genome-enabled hitchhiking mapping identifies QTLs for stress resistance in natural Drosophila. Heredity 99:313-321.

Ohta, T. 1971. Associative overdominance caused by linked detrimental mutations. Genet Res 18:277-286.

Ohta, T., and M. Kimura. 1970. Development of associative overdominance through linkage disequilibrium in finite populations. Genet Res 16:165-177.

Palsson, S., and P. Pamilo. 1999. The effects of deleterious mutations on linked, neutral variation in small populations. Genetics 153:475-483.

Pollak, E. 1987. On the theory of partially inbreeding finite populations. I. Partial selfing. Genetics 117:353-360.

Porcher, E., P. H. Gouyon, and C. Lavigne. 2004. Dynamic management of genetic resources: maintenance of outcrossing in experimental metapopulations of a predominantly inbreeding species. Cons Genetics 5:259-269.

R Development Core Team. 2006. R: A language and environment for statistical computing. http://www.R-project.org, Vienna.

Rockman, M. V., and L. Kruglyak. 2009. Recombinational landscape and population genomics of Caenorhabditis elegans. PLoS Genet 5:e1000419.

Rose, M. R. 1982. Antagonistic pleiotropy, dominance and genetic variation. Heredity 41:63-78.

Rumball, W., I. R. Franklin, R. Franham, and B. L. Sheldon. 1994. Decline in heterozygosity under full-sib and double first-cousin inbreeding Drosophila melanogaster. Genetics 136:1039-1049.

Schedl, T., and J. Kimble. 1988. fog-2, a germ-line-specific sex determination gene required for hermaphrodite spermatogenesis in Caenorhabditis elegans. Genetics 119:43-61.

Scheet, P., and M. Stephens. 2006. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am J Hum Genet 78:629-644.

Seidel, H. S., M. V. Rockman, and L. Kruglyak. 2008. Widespread genetic incompatibility in C. elegans maintained by balancing selection. Science 319:589-594.

Sellis, D., B. J. Callahan, D. A. Petrov, and P. W. Messer. 2011. Heterozygote advantage as a natural consequence of adaptation in diploids. Proc Natl Acad Sci U S A 108:20666-20671.

Semple, C., and K. H. Wolfe. 1999. Gene duplication and gene conversion in the Caenorhabditis elegans genome. J Mol Evol 48:555-564.

Stiernagle, T. 1999. Maintenance of C. elegans. Oxford University Press, Oxford.

Takahata, N., Y. Satta, and J. Klein. 1992. Polymorphism and balancing selection at major histocompatibility complex loci. Genetics 130:925-938.

Teotonio, H., S. Carvalho, D. Manoel, M. Roque, and I. M. Chelo. 2012. Evolution of outcrossing in experimental populations of Caenorhabditis elegans. PLoS ONE:10.1371/journal.pone.0035811.

Teotonio, H., I. M. Chelo, M. Bradic, M. R. Rose, and A. D. Long. 2009. Experimental evolution reveals natural selection on standing genetic variation. Nat Genet 41:251-257.

Turner, T. L., A. D. Stewart, A. T. Fields, W. R. Rice, and A. M. Tarone. 2011. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in Drosophila melanogaster. PLoS Genet 7:e1001336.

Venables, W. N., and B. D. Ripley. 2002. Modern Applied Statistics with S. Springer, New York.

Wang, R. L., A. Stec, J. Hey, L. Lukens, and J. Doebley. 1999. The limits of selection during maize domestication. Nature 398:236-239.

Waples, R. S. 1989. A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics 121:379-391.

Weir, B. S. 1996. Genetic Data Analysis II. Sinauer Associates, Inc., Sunderland.

Weir, B. S., P. J. Avery, and W. G. Hill. 1980. Effect of mating structure on variation in inbreeding. Theor Pop Biol 18:396-429.

Wright, S. 1978. Evolution and the Genetics of Populations: Variability within and among Natural Populations. University of Chicago Press, Chicago.

Figure Legends

Figure 1. SNP differentiation. In panels **A** and **B**, the average H_e among replicate populations at G100 relative to the H_e found in the two ancestrals (A_0 or D_0) for all SNPs measured. Red and blue circles indicate SNPs that differentiated under androdioecy or dioecy, respectively, as estimated with GLM. Filled circles indicate monomorphic SNPs at G100 (as defined by average H_e<0.05 among replicates). Regressions of G100 with G0 are shown as lines for GLM-undifferentiated SNPs (in gray): both models have zero intercepts and slopes close to one, as expected with overall maintenance of diversity during the experiment.

Figure 2. Linkage disequilibrium and effective recombination. In panel **A**, points show the mean r^2 at G100 relative to the genetic distance between SNPs, with the error bars indicating one SEM among replicates. Androdioecious results are shown in red coloring and dioecious results in blue coloring. Results for GLM-undifferentiated SNPs are indicated by gray circles, results between GLM-undifferentiated SNPs with GLM-differentiated SNPs by empty circles, and results for GLM-differentiated SNPs by colored gray circles. In panel **B**, the evolution of the genetic distance at which background r^2 is reached. Red circles indicate the average distance for androdioecious populations and blue for dioecious populations with one SEM (see also **Table S3**). In panel **C**, the rate of input of new haplotypes during evolution, measured in 10 SNP windows along both chromosomes (see also **Figure S4**). Symbols are as in panel **B**.

Figure 3. SNP allele frequency changes expected with neutrality and single-locus selection. In panels **A** and **B**, the SNP allele frequency changes obtained from neutral simulations (empty bars) or observed (gray bars) during evolution from G0 to G100. $3x10^3$ simulations were performed. In panel **C**, the estimated distributions of single-locus

selection coefficients in GLM-differentiated SNPs are shown in red for androdioecious populations and blue for dioecious populations. Fitness effects do not differ among mating systems (Fisher´s exact test p-value=0.222).

Figure 4. Observed and expected evolution of population genetic structure with genetic drift. In panel **A**, the evolution of H_e is shown in red for androdioecious populations and blue for dioecious populations, with associated SEM. Dashed lines limit 95% of the data obtained from $3x10^3$ neutral simulations of evolution for each period. In panels **B** and **C** are shown the evolution of F_{IS} and H_i, respectively [see also **Figure S8** for results on CV (H_i)].

Figure 5. Population genetic structure with selection at deleterious recessive alleles, under androdioecy. In panels **A**, **B** and **C**, illustrative results for H_e, F_{IS} and H_i obtained from simulations with selection on deleterious alleles at each period of evolution. The mean and two SD of 3x100 simulations at each period are presented for each parameter combination of dominance (h) and selection (s) coefficients. Experimental results are presented with one SEM (as in **Figure 4**).

Figure 6. Selection models with deleterious recessive alleles. Distances (*dist*) between observed and simulated average H_e, F_{IS} and H_i trajectories during experimental evolution. Smaller distances indicate better fits. 100 simulations were performed for each parameter combination of h and s (indicated with crosshair symbols in panel **A**). Top panels show the fits for androdioecious evolution and bottom panels for dioecious evolution, as obtained with lowess smoothing. Large symbols indicate the parameter combinations of **Figure 5**.

Figure 7. Selection models with overdominant loci. Distances (*dist*) between observed and simulated average H_e, F_{IS} and H_i trajectories during experimental evolution. 100

simulations were performed for each parameter combination of selection strength α and epistasis k (indicated with crosshair symbols in panel **A**). Top panels are for androdioecious evolution and bottom panels for dioecious evolution. Note that the *dist* scale is the same as in **Figure 6**. See also **Figure S9** for illustrative results at certain parameter combinations under androdioecy.