nature genetics

Experimental evolution reveals natural selection on standing genetic variation

Henrique Teotónio[1], Ivo M Chelo[1], Martina Bradić[2], Michael R Rose[3] & Anthony D Long[3]

Evolution depends on genetic variation generated by mutation or recombination from standing genetic variation. In sexual organisms, little is known about the molecular population genetics of adaptation and reverse evolution[1–11]. We carry out 50 generations of experimental reverse evolution in populations of *Drosophila melanogaster*, previously differentiated by forward evolution, and follow changes in the frequency of SNPs in both arms of the third chromosome. We characterize the effects of sampling finite population sizes and natural selection at the genotype level. We demonstrate that selection has occurred at several loci and further that there is no general loss or gain of allele diversity. We also observe that despite the complete convergence to ancestral levels of adaptation, allele frequencies only show partial return.

Adaptation has been shown to depend on evolutionary history[1–4]. In particular, studies of reverse evolution, defined as the reacquisition of ancestral states[5,6], have demonstrated that adaptation of differentiated populations to common ancestral environmental conditions involves multiple life-history, physiological and developmental phenotypes[3,4,7–11]. However, knowledge of the molecular population genetics of evolution in sexual populations is scarce[12]. If adaptation generally occurs by new mutations followed by their subsequent fixation, then evolution is dependent on events that will not occur among all derived populations, at moderate to small population sizes. On the other hand, if adaptation is generally due to allele frequency changes at loci with standing variation, evolution can proceed in parallel among derived populations experiencing similar environmental conditions[1,6].

A laboratory adaptive radiation has been created over the past 28 years from a single outbred population, by the imposition of sustained environments under defined conditions of high population census sizes ($N > 1{,}000$) and discrete generations[4]. Concomitant with laboratory selection, replicate populations show extensive differentiation for a variety of phenotypes, without any apparent loss of additive genetic variances[4,13]. In 1997, new populations were derived from these populations by reimposition of their common ancestral environment for 50 generations (**Fig. 1**). In the absence of preserved samples from the ancestral population, control populations were defined as those that had been continuously maintained in the ancestral conditions.

We followed 29 populations: 5 replicate control populations ($IB_{1–5}$) at generation 0 and at generation 50 of reverse evolution; 4 replicate populations selected for late-life reproductive success ($O_{1–3,5}$) and their reverse-evolved derivatives ($IO_{1–3,5}$) after 50 generations in the ancestral environment; those selected for increased starvation resistance and their reverse-evolution derivatives ($SO_{1–3,5}$ and $ISO_{1–3,5}$, respectively); and finally, those selected for very early reproductive success and their reverse-evolved derivatives ($ACO_{1–3,5}$ and $IACO_{1–3,5}$).

We genotyped 55 SNPs in all experimental populations (**Supplementary Tables 1** and **2** online). Some SNPs were chosen at loci that have been previously shown to harbor allele variation that correlates with life-history evolution[14–17]. Other SNPs were typed at noncandidate loci, and a third class was typed in regions for which there had been no previous description of polymorphisms. A summary of SNP frequencies among our populations is shown in **Figure 2** and **Supplementary Table 3** online. With the sample sizes used for SNP discovery, we had reasonable power to detect low-frequency variants, and, further, alleles at low frequencies in an ancestor population that increase in frequency in derived populations amplify the possibility of detection of differentiation. About 70% of the SNPs surveyed were at intermediate frequencies (>5%) in the control populations (**Supplementary Table 3**). During the 50 generations of the reverse-evolution experiment, the SNP frequencies in control populations did not seem to have changed (**Fig. 2a**). On the other hand, differentiation was significant for a large fraction of the populations that had undergone divergent (forward) evolution. Furthermore, reverse evolution seemed to have returned many of these divergent frequencies back to the allele frequencies of the controls (**Fig. 2b**).

Correlated changes in allele frequencies at several SNPs can be a manifestation of linkage disequilibrium (LD) between a region undergoing frequency-change and others physically associated with it[18]. Statistics of pairwise associations of allele frequencies among SNPs (r^2 and $|D'|$[19]) show that strong LD rarely extends further than 10 to 100 kb in our study populations (**Fig. 3a,b**), roughly one order of magnitude higher than the LD typically observed in natural

[1]Instituto Gulbenkian de Ciência, Apartado 14, P-2781-901 Oeiras, Portugal. [2]Department of Biology, New York University, New York, New York 10003, USA. [3]Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697, USA. Correspondence should be adddressed to H.T. (teotonio@igc.gulbenkian.pt).

Received 9 July 2008; accepted 27 October 2008; published online 11 January 2009; doi:10.1038/ng.289

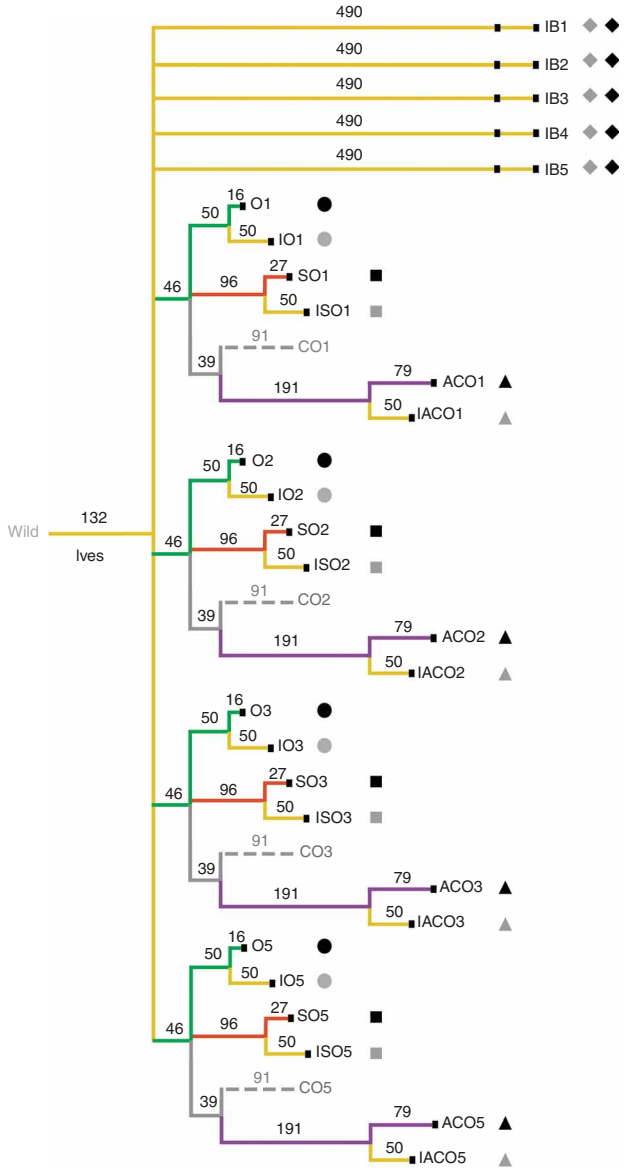

Figure 1 Experimental evolution design. Different laboratory culture regimes were imposed: ancestral two-week life cycle (yellow), late reproduction (green), protracted starvation resistance (red) and very early culture reproduction (magenta). See Methods and ref. 4 for a definition of acronyms and details of laboratory maintenance. Gray dashed lines indicate the demographic controls of the populations selected for increased starvation resistance; these were not followed here. Numbers above each branch indicate the number of generations in each environment at the time of the present study. Dots at the ends of branches indicate the time points where individuals were sampled. Symbols adjacent to these dots indicate the population samples used in the analysis shown in **Figures 2**, **4** and **5**.

populations[20]. Further, it is apparent that LD is unchanged with experimental evolution. An analysis of negative correlations between r^2 and $|D'|$ and physical distance, among other statistics calculated to test for the occurrence of recombination within and among regions, shows that, for physical sizes as small as 50 kb, recombination has occurred in our experiment (**Fig. 3b** and **Supplementary Table 4** online). Thus, we conservatively defined seven regions that seemed to show little evidence of LD among them, and that hence can be argued to be free to independently evolve under our experimental paradigm. To further quantify the role of recombination within these regions, we estimated the genetic distances between the SNP markers that identify the ends of each region (**Supplementary Table 1**). Two regions are genetically large enough (region 2 is ~0.17 cM and region 7 is ~0.05 cM) for recombinant alleles to have arisen and reached appreciable frequency during 50 generations of reverse experimental evolution. For each of these seven genomic regions, referred to henceforth as loci, we reconstructed the multi-SNP haplotypes used in subsequent analysis[21].

We next focused on whether experimental evolution resulted in the gain or loss of genetic variation. First, we determined whether the SNPs used to reconstruct haplotypes were in single-locus Hardy-Weinberg equilibrium. The distribution of observed P values seemed uniform, leading us to conclude that there is little evidence for disequilibria (**Supplementary Fig. 1** online). Second, our estimates of the population genetics parameters $4N_e c$ (ref. 22), where N_e is the effective population size and c the recombination rate per base pair, suggested that the observed patterns of diversity predated the founding of the ancestral population used in this study (data not shown). As the interpretation of these statistics depends on assumptions that our populations are likely not to meet, such as equilibrium among mutation, recombination and selection at each locus, a more direct assessment is to estimate diversity in the form of heterozygosity. In **Figure 3c**, this statistic is shown not to vary substantially between regions and experimental treatments. Overall, diversification seems to result mostly from changes in the frequency of, and recombination between, genotypes already present in the ancestral population from which the entire experimental radiation was derived (**Supplementary Fig. 2** online).

The allele frequency trajectories in the control populations during the 50 generations of reverse experimental evolution should provide a baseline estimate of the sampling effects due to finite breeding size experienced by all the populations in the ancestral environment. For each locus, we calculated effective population size (N_e) from the allele frequency shifts between generation 0 and 50 in the control populations. The analysis was done by forward simulation, assuming a process of gametic sampling preceded by a diploid phase, in order to model recombination at loci 2 and 7. We carried out 10,000 series of 5 simulations each for 50 generations, taking as starting allele frequencies those observed in the control populations at generation 0 for the 5 most common haplotypes in each locus (**Supplementary Fig. 2**). The statistics used to follow allele frequencies in the simulations were the standardized variance in allele frequencies between generations 50 and 0, averaged over the 5 replicate populations[23,24] (Fc), as well as the expected genetic differentiation among replicate populations[25] (Gst). We varied N_e by choosing its values from a uniform distribution, from which we then obtained a probability density function of N_e associated with the observed gene diversity among the control populations. A summary of these simulations is shown in **Supplementary Figure 3** online: Fc-based estimates of N_e are more precise than those based on Gst, and thus we used the former method in our subsequent analysis. The magnitude of sampling finite populations in the ancestral environment during 50 generations was obtained by taking together N_e estimates among loci. Overall, we estimate 232 ± 93 s.e.m. breeding individuals per generation. We note that a N_e smaller than the census size is consistent with both demographic mechanisms as well as potential interactions between these mechanisms and natural selection, distinct to the control populations.

With an estimate of the distribution of N_e in the control populations, we asked whether the observed variation in the forward- and reverse-evolved populations was consistent with those same effective

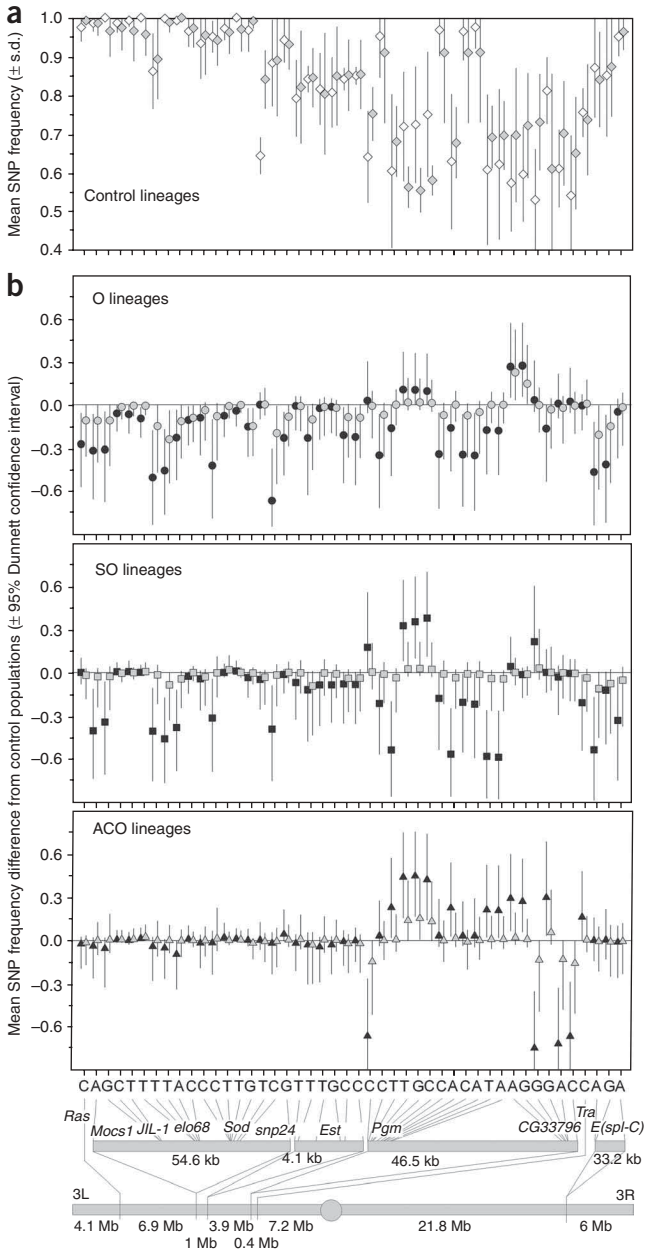

Figure 2 Population dynamics of SNPs. (**a**) Absolute SNP frequencies in the control (IB) populations at generation zero (empty diamonds) and after 50 generations of experimental reverse evolution (solid diamonds). Bars indicate s.d. of the five replicate populations. Changes in control allele frequencies should mostly reflect the effects of genetic drift. (**b**) The allele frequencies of SNPs are shown for the O lineages (circles; O in solid and IO in gray), SO lineages (squares; SO in solid and ISO in gray) and ACO lineages (triangles; ACO in solid and IACO in gray) as the mean difference from the (IB) control populations at generation 50. Associated 95% confidence intervals were obtained after ANOVA modeling each SNP with seven levels (IB, O, IO, SO, ISO, ACO, IACO) and Dunnett post hoc comparisons with the control populations (IB). The bottom diagram illustrates the relative position of SNPs in both arms of the third chromosome, with the circle indicating the centromere, as well as the physical sizes of the regions studied. For each SNP, details on the positions, identities and observed frequencies for all populations can be found in **Supplementary Tables 1–3**. The *x* axis shows the most common (>50%) allele in the control populations. There is extensive divergence in most of the SNPs studied followed by partial or complete convergence to control frequencies after 50 generation of reverse evolution.

population sizes. Simulations were done as before to determine the expected *Fc* of the three experimental contrasts for the duration of divergence among populations from their most recent common ancestor: (i) between the O populations and their reverse-evolved derivatives (the IO), for 66 generations of divergence; (ii) between the starvation-resistance selected (SO) populations and reverse-evolved derivatives (ISO), for 77 generations; and (iii) between the ACO populations and their corresponding reverse-evolved populations (IACO), for 129 generations of divergence. As shown in **Figure 4**, the effects of sampling finite populations are sufficient to explain much of the patterns of experimental differentiation. Only the O lineages at locus 1 ($P = 0.02$) and the SO lineages at locus 7 ($P = 0.03$) seem to have significantly higher differentiation of allele frequencies than those expected.

We have so far considered the effects of mutation, recombination and sampling finite populations in the ancestral environment on allele frequency differentiation between populations. We were also able to assess the impact of natural selection because of the unique feature afforded by experimental design: if natural selection influences patterns of allele frequency change, then we expect divergence from control levels during forward evolution and then convergence toward control levels in the reverse-evolved populations. We first sought to identify the allele, within each independently evolving locus, showing the greatest change in frequency, as assessed by ANOVA and permutation tests[26]. Three experimental contrasts were done: the IB, O and IO contrast; the IB, SO and ISO contrast; and the IB, ACO and IACO contrast. In all, this analysis identified alleles showing significant differentiation at six loci for the O and SO lineages, whereas only three loci had significantly differentiated alleles for the ACO lineages (**Table 1**). The permutation tests and ANOVA were also done among the control populations at generation 50 (**Supplementary Table 5** online). In **Supplementary Figure 2**, we present the dynamics observed in the five most common haplotypes, and the whole dataset can be found in **Supplementary Table 6** online.

We did one final analysis to determine the extent of convergence and, thus, the effects of natural selection. Associations were computed between allele frequency changes in the initially forward-evolved populations, identified using the permutation testing and ANOVA models above, relative to the allele frequency changes observed after the re-imposition of the ancestral environment. It is clear that allele frequency evolution is reversible (**Fig. 5**). There are few examples of stasis and no examples of divergence during the episode of evolution in the ancestral environment. The expected distribution of allele frequencies during divergence and convergence was obtained by 1,000 forward simulations, done as before, but in this case for the whole experimental phylogeny (**Supplementary Fig. 4** online). We estimate that convergence to ancestral allele frequencies during 50 generations of reverse evolution is on the order of 50%, and that this measure is seemingly independent of evolutionary history.

The partial degree of convergence to ancestral allele frequencies might conceivably increase if more generations of reverse evolution are allowed. But because the experimental populations have been previously shown to fully converge to ancestral measures of fitness after 50 generations[7], the present estimate of the extent of convergence at the genotype level suggests that there was a change in allelic effects during the 50 generations of reverse evolution. Other work with experimental reverse evolution in viruses, which depends upon mutation for the creation of diversity, has found that reversibility at

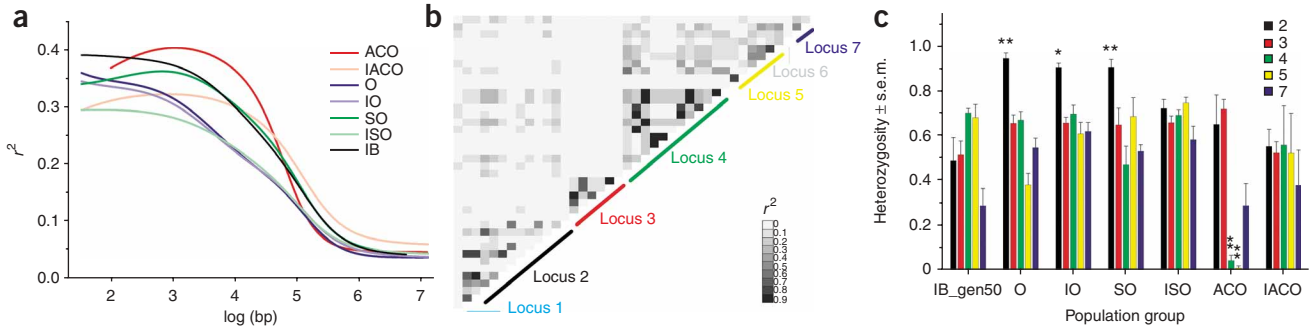

Figure 3 Loci undergoing independent evolution. (**a**) Smoothing splines of r^2 with physical distance for each group of populations. Linkage disequilibrium is high but consistent across the different experimental evolution treatments. (**b**) Pairwise r^2 statistics averaged over all populations. Colored lines indicate the loci we define as independently evolving, as tested by combining over the *P* values of the Lkmax and G4 tests, and negative correlations between linkage disequilibrium r^2 and D' statistics with physical distance, for each single population, among adjacent loci and within loci[22] (**Supplementary Table 4**). Locus 1 is defined by SNP1 (genetic map 3L: 10.13 cM), locus 2 is defined by SNP2,5,9,11–15,19,21,23,24 (3L: 34.07.34.44 cM), locus 3 by SNP25–30 (3L:37.98.37.98 cM), locus 4 by SNP31–36,38–43 (3L:43.59.43.6 cM), locus 5 by SNP44–50 (3L:43.71.43.71 cM), locus 6 by SNP51 (3L:43.99 cM), and locus 7 by SNP52–54 (3R:89.32.89.43 cM). Inversion polymorphisms may increase linkage disequilibrium, but a diagnostic SNP for inversion 3L Payne shows that this inversion is absent in most populations or otherwise only present at low frequencies (O_5, 4%, $n = 70$; IO_5, 2%, $n = 70$; IO_1, 15%, $n = 46$), and results remain unchanged when the inversion marker is taken into account. (**c**) Evolution of haplotype diversity. For each group of populations, expected heterozygosity is shown with s.e.m. for the five loci where haplotype reconstruction was performed and after discarding haplotypes unique to single populations. Heterozygosity at loci 1 and 6 are shown in **Figure 2**. ANCOVA reveal population diversity differences for locus 2 ($F_{6,28} = 6.8$; $P < 0.001$; $R^2 = 0.68$), locus 4 ($F_{6,28} = 9.5$; $P < 0.001$; $R^2 = 0.73$), locus 5 ($F_{6,28} = 7.8$; $P < 0.001$; $R^2 = 0.79$) and locus 7 ($F_{6,28} = 2.7$; $P = 0.04$; $R^2 = 0.45$) but no influence of the sample size covariate. Comparisons with the control populations by Dunnett tests followed significant ANCOVA. $^*P < 0.05$ and $^{**}P < 0.005$. Locus 1 was found to be heterogeneous ($F_{6,28} = 6.5$; $P = 0.001$; $R^2 = 0.67$), with the O populations having significantly higher diversity than controls ($P < 0.001$), and was locus 6 ($F_{6,28} = 4$; $P = 0.008$; $R^2 = 0.54$), with the ACO populations showing less diversity than controls ($P = 0.01$). Most SNP markers defining the haplotypes are at Hardy-Weinberg genotype proportions (**Supplementary Fig. 1**). Fifty generations of reverse evolution usually restore heterozygosity to control levels. Mutation rates would need to be extremely high to explain these results[29], and thus we interpret observed disparities as due to our sampling scheme, which does not allow detection of very low-frequency variants (<1%). We conclude that there is no overall increase or decrease in allele diversity with experimental evolution.

the genotype level is approximately 20% (ref. 3). The difference between that result and ours confirms the prediction that evolution resulting from standing variation is more repeatable than evolution resulting from mutational input.

To our knowledge, this study is the most comprehensive description of the molecular population genetics of adaptation in a sexual species[6,11,12,27]. Adaptive reverse evolution can occur from the sorting and recombination of standing genetic variation at multiple loci.

METHODS

Experimental evolution design. All populations are ultimately derived from a sample of wild *Drosophila melanogaster* collected in Massachusetts in 1975 by P.T. Ives that was maintained in well-defined laboratory conditions until 1980. The ancestral environment features discrete and non-overlapping 2-week generations, with controlled egg and adult density and egg laying taking place during about 2 h after mixing all adults of the same population. Amount of nutrients, temperature, humidity and light schedule were controlled. By 1997, a selective radiation had been derived from the ancestral population, chiefly by imposing varied life-cycle timings ranging from 9-d to 70-d cycles, but, in the case of one group of populations studied here, by imposing starvation sufficient to kill most females and all males before culture reproduction. For experimental reverse evolution, new populations were derived from differentiated populations and the ancestral Ives environment was re-imposed for 50 generations. Control replicate populations ($IB_{1–5}$) were those that had always been maintained in the ancestral regime. The experimental phylogeny used in the present study is shown in **Figure 1**: we studied four replicate populations per group, except for the control populations where five populations were studied, for a total of 29 populations. We refer to groups of populations as sharing evolutionary lineages when populations share a common history of selection, as follows: 'O lineages' ($O_{1–3,5}$ and $IO_{1–3,5}$ populations), populations that share a history for selection for later reproduction; 'SO lineages' ($SO_{1–3,5}$ and $ISO_{1–3,5}$ populations), populations that share a history of selection for increased starvation resistance; and 'ACO lineages' ($ACO_{1–3,5}$ and $IACO_{1–3,5}$ populations), populations that share a history for very early reproduction. At generation 50 of reverse evolution, and after 2 generations of maintenance in

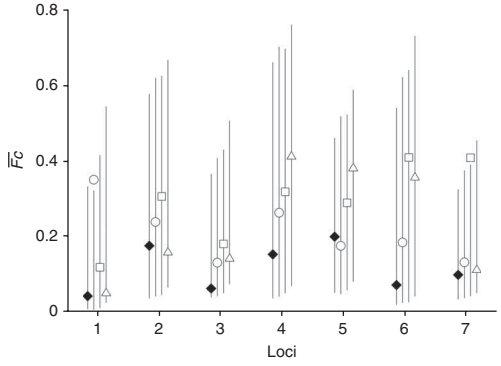

Figure 4 Effects of sampling finite populations. Lines indicate the 95% confidence intervals of the mean of Nei and Tajima's *Fc* statistic[23], calculated over 1,000 forward Wright-Fisher simulations. Confidence intervals were obtained by multinomial sampling of $N_e = 464.2 \pm 184.9$ s.e.m. gametes, with each haplotype assumed to have started at the generation zero control frequencies. Filled diamonds indicate the observed *Fc* statistics in control populations for 50 generations of evolution. Circles indicate *Fc* for the O-IO divergence from their most common recent ancestor (66 generations for paired same numbered replicates), squares for the SO-ISO divergence (77 generations) and triangles for the ACO-IACO divergence (129 generations). The O-IO contrast at locus 1 (*P* value = 0.023) and the SO-ISO contrast at locus 7 (P-value = 0.025) are differentiated. Overall, the effects of sampling finite breeding populations can explain most genotypic changes.

Table 1 Allele differentiation in loci evolving independently

Loci[a]	Test	Experimental contrast: Late-life fitness	Starvation resistance	Early reproduction
1	Haplotype[b]	A**	–	–
	Post hoc[c]	O**, IO*	–	–
2	Haplotype[b]	AGTTACCCGTCG**	AGTTACCCGTCG*	–
	Post hoc[c]	O**, IO*	SO**	–
3	Haplotype[b]	TTTGCC**	TTTGCC**	–
	Post hoc[c]	O**, IO*	SO**, ISO*	–
4	Haplotype[b]	CTCTGCTGTGAC*	CCTGTACACATA**	TCTTGCCACATA**
	Post hoc[c]	O**	SO**	ACO**
5	Haplotype[b]	TAGGAC**	AGCGGT*	AGCGGT**
	Post hoc[c]	O**, IO**	SO*	ACO**
6	Haplotype[b]	–	C*	C*
	Post hoc[c]	–	SO*	ACO*
7	Haplotype[b]	AGA**	AGA**	–
	Post hoc[c]	O**, IO*	SO**, ISO*	–

[a]Inference of independent evolution based upon the occurrence of recombination events among these loci, as presented in **Supplementary Table 4**. [b]Haplotype showing the greatest heterogeneity by one-way ANOVA, when considering the five most common haplotypes across all populations. Per locus and haplotype, three contrasts were performed: late-life fitness (IB versus O versus IO); starvation resistance (IB versus SO versus ISO); and early-life fitness (IB versus ACO versus IACO). [c]Experimental treatment significantly different from controls as assessed by Dunnett's post hoc comparisons. Significance of P values was assessed by permutation tests (see Methods); $^{**}P < 0.005$; $^{*}P < 0.05$; all model residuals follow normality; d.f. = 2 (population group) and d.f. = 10 (error). A dash indicates a nonsignificant result.

a common environment, samples from all populations were frozen at –80 °C, until genomic DNA isolation of individual females was done with the Puregene purification Kit (GENTRA systems). Individuals from the IB control populations at generation zero of reverse experimental evolution were also sampled.

SNP discovery. We initially discovered SNPs in the *Elo68*, *Sod*, *Pgm* and *CG33796* loci using Expand High Fidelity^{plus} PCR amplification (Roche). Nucleotide sequencing was done by dye terminator technology of approximately 4 kb covering these loci, in 20 individuals of the IB_2 population, on an ABI337 automated DNA sequencer (Applied Biosystems). We inferred polymorphisms by observing the chromatograms, and confirmed them by pGEM-Teasy vector cloning (Promega) of relevant PCR products and DNA sequencing. Because of low allele diversity at the *Elo68* and *Sod* loci in the IB_2 population, we further sequenced 20 individuals from the O_2 population. SNPs from all other regions had been previously described and are referenced in **Supplementary Table 1**.

Genotyping. We detected SNPs by mass determination of allele-specific extension oligonucleotide molecules generated from amplified genomic DNA using a MALDI-TOF mass spectrometry platform (Sequenom), which allows multiplexing of up to 30 different SNPs per sample. For each SNP, we designed PCR and extension primers using the MassARRAY design software, with 7.5 ng of template DNA per reaction. Unincorporated nucleotides were deactivated using shrimp alkaline phosphatase. For allele-specific extensions, we used MassExtended primers and a single termination mix. After cleaning, we spotted products onto a Spectrochip, carried out laser scanning using a mass spectrometry workstation (Bruker) and analyzed the resulting spectra with SpectroTYPER-RT software. We assayed a total of 120,733 diploid genotypes, of which only 86,981 were informative (not monomorphic) and could be unambiguously determined. Of these, 10,299 were run in duplicate, with 377 giving conflicting calls; 3.7% is thus the call error of the genotyping technique. Our initial dataset is composed of 54 SNPs, with an average number of 37 individuals per population being genotyped for a total of 71,107 diploid genotypes (**Supplementary Table 2**). We further genotyped a SNP distinguishing the inversion of Payne from the standard *Drosophila melanogaster* sequence.

Observed genotype dynamics. We reduced the initial genotype matrix by removing individuals with less than 90% of their genotype determined and also all SNPs that were called in less than 90% of the samples, either within each population or in the combined group of all populations, depending on the analysis performed (**Supplementary Table 2**). Frequencies were angular transformed before analysis of SNP and haplotype dynamics with experimental evolution. We used general linear models to determine heterogeneity of allele frequencies among groups of populations with different evolutionary histories, followed by a posteriori Dunnett tests contrasting each experimental group with only the control populations (IB), taking into consideration the ANOVA model errors. We defined 'allele' as the haplotype (single SNP or multi-SNP) associated with each of the seven regions showing the greatest differentiation, for each of the three experimental contrasts: late-life fitness (IB versus O versus IO); starvation resistance (IB versus SO versus ISO); and early-life fitness (IB versus ACO versus IACO). To test the significance of the best allele within each region and experimental contrast, we used permutation testing correcting for the bias in the P value associated with ascertaining the most significant allele, using 10,000 permutations[26]. We checked homoscedasticy of group variances with Levene tests and normality of residuals by Kolmogorov-Smirnov tests. Power was estimated assuming a balanced ANOVA design with four observations per group and s.d. as the square root of the mean squares model error. For all models, estimated power was higher than 0.8, whenever differences among groups were inferred (data not shown).

Linkage disequilibrium analysis. The statistics r^2 and $|D'|$ were estimated from the unphased data as composite genotype disequilibria, which assumes random

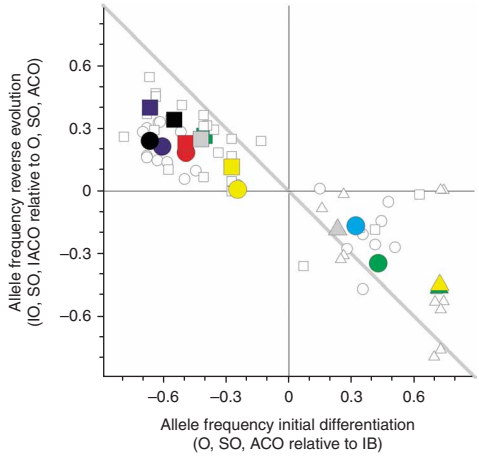

Figure 5 Reverse evolution at the genotype level. Associations between allele frequencies at the regions that have significantly changed from controls in the populations exposed to diversifying environments (O, SO, ACO), and allele frequency changes in the reverse-evolved populations (IO, ISO, IACO) relative to the O, SO and ACO ancestors (**Table 1**). Small open symbols indicate observed values for all replicate populations, and large filled symbols indicate the means for each group of populations: circles for the O and IO populations, squares for SO and ISO populations and triangles for ACO and IACO populations. Colors code the different loci as in **Figure 3**. Observed allele frequency changes are significant, as tested by simulation of the expected distribution of convergence and divergence under the full experimental phylogeny (**Supplementary Fig. 4**). For the O-IO contrast, $\chi^2 = 25.4$ (d.f. = 3, $P < 0.001$); for the SO-ISO contrast, $\chi^2 = 49.2$ (d.f. = 3, $P < 0.001$); and for the ACO-IACO contrast, $\chi^2 = 4.3$ (d.f. = 1, $P = 0.04$). There is no heterogeneity among the three contrasts ($\chi^2 = 2.6$, d.f. = 4). The gray line shows 100% of convergence to ancestral frequencies. The slope of a linear function through these data is –0.43 for the O lineages, –0.47 for the SO lineages and –0.53 for the ACO lineages. Thus, approximately 50% of genotype changes after initial differentiation are reversible within 50 generations. This is likely an underestimate of the true level of convergence, as by examining only those loci showing significant change during forward evolution, we slightly bias the estimate of convergence downward as a result of a sampling error known as the Beavis effect. Nonetheless, given the number of alleles and loci sampled, our data are inconsistent with full convergence to ancestral frequencies during the period studied.

mating so that the genotype probabilities are the products of the gametic probabilities[20]. We obtained splines of r^2 over physical distance with the 'ksmooth' function, a regression-based function using the Nadaraya-Watson kernel estimator, under R[28]. The population genetic recombination parameters ($4N_ec$, where N_e is the effective population size and c the recombination rate per base pair) were estimated with the composite likelihood method of McVean and colleagues and as implemented in LDhat[22]. Input data was phased[21]. Mutation rate was taken at 5.8×10^{-9} bp/generation[29], with effective population size of 1,000.

Haplotype reconstruction was done with fastPHASE 1.2 (ref. 21). We used 20 random starts of the EM algorithm and sampled 200 haplotypes from the posterior distributions. The number of clusters for the cross-validation procedure was set to 10. Genotypes with posterior probabilities lower than 90% were considered as missing data for further analysis. This resulted in phase determination of the alleles present in heterozygote genotypes while maintaining the uncertainty due to original missing data. For subsequent analyses, haplotypes with missing nucleotide information were removed. We estimated haplotype diversity (expected heterozygosity) as the sum of squares of allele frequencies for each population, normalized by the sample size minus one. We obtained expected genetic distances between adjacent SNPs within each locus by locally regressing, using the 'smooth.spline' function in R, the estimated cM positions for each cytogenetic map position on the average sequence coordinate of that cytogenetic band, separately for each chromosome arm. Positions were obtained from Flybase (**Supplementary Table 1**).

In each of our populations, deviations from null models of no recombination were tested with the Lkmax test, where the maximum composite log likelihood $4N_ec$ frequency distribution is estimated with 1,000 permutations of SNP location. G4 tests, comparing the sum of distances between all pairs of SNPs that have all four haplotypes to the distribution obtained from 1,000 permutations of distances, were also done. Finally, we also assessed the significance of negative correlations between either r^2 or $|D'|$ with distance by 1,000 permutations of SNP location[22]. We inferred the overall presence of recombination events by estimating the composite one-tailed P value of all four recombination tests, for a given group of populations, assuming a χ^2 distribution of the P values testing the same hypothesis of no recombination. A region was considered independent from an adjacent one when at least four of these composite P values, among the seven groups tested (IB, O, IO, SO, ISO, ACO, IACO), were below $\alpha = 0.05$.

Simulations of experimental population genetics. Effective population sizes (N_e) in the ancestral culture regime were estimated, for each locus, by the changes in allele frequency among the control populations between generations 50 and 0. We simulated 10,000 series of five multinomial random walks of allele (gametic) frequencies were simulated during 50 generations. Starting allele frequencies were those observed among the five replicate control populations at generation zero, for the five most common alleles plus one allelic class including all other reconstructed haplotypes (**Supplementary Fig. 2**).

We modeled recombination among SNPs within loci for loci 2 and 7 by including a diploid phase before sampling the following generation. The frequency of diploid genotypes was obtained as the product of constitutive allele frequencies. The occurrence of crossing-over events was directly proportional to the expected genetic distance among SNPs (**Supplementary Table 1**). We did not use loci 1 and 6 in these simulations, because they were mostly uninformative, given their biallelic status. For each generation, sampling was produced using the 'rmultz2' function present in Combinat under R, conditional on the allele frequency in each population in the previous generation. Prior probability distributions of $N_e \sim$ U (50, 4000). Posterior distributions of N_e were obtained by following two statistics of population diversity: the mean, among the five replicate simulations of a series, standardized variance for each replicate simulation between generation 0 and 50 (Fc)[23], and among the five replicate simulations differentiation at generation 50 (Gst)[26]. We sampled generation 50 data from the simulated populations, taking into consideration the experimental control sample sizes. The estimated distribution was found by taking 1,000 simulated data around the observed values of Fc or Gst in the control populations (**Supplementary Fig. 3**). As Fc-based estimates were more accurate than those based on Gst, they were used in subsequent analysis.

Four experimental contrasts were done to check the fit between expected and observed population differentiation, under sampling of finite breeding populations assuming starting $N_e \sim$N(464.2, 413.5), as found in the simulations above. We carried out 1,000 simulations, keeping N_e constant in each run. Starting allele frequencies were those observed in the control populations at generation zero. Observed Fc were calculated (i) among the control populations; (ii) among pairwise replicate O and IO populations (that is, as the mean of Fc estimated for O1 and IO1 divergence, O2 and IO2 divergence, and so on; 66 generations of divergence from their most recent common ancestor were simulated); (iii) among pairwise replicate SO and ISO populations (with 77 generations of divergence); and (iv) among pairwise replicate ACO and IACO populations (with 129 generations of divergence).

To determine the expected distribution of convergence versus divergence in allele frequencies, under sampling finite populations, we did 1,000 forward simulations as before, for each experimental contrast (O-IO, SO-ISO and ACO-IACO). However, we considered the whole phylogenetic structure of the experiment. After each simulation, records were kept only of those alleles that were found to be significant in the permutation and ANOVA models during the divergence of O, SO and IACO populations[26]. We used χ^2 statistics to find the significance of convergence.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank T. Aires, D. Brites, J. Costa and I. Marques for support with SNP discovery and genotyping, and G. McVean for support with LDhat. We thank R. Azevedo, S. Carvalho, A. Coutinho, S. Estes, L. Mueller, P. Phillips, S. Proulx, M. Che Soares and É. Sucena for comments on the project and manuscript. Financial support was provided by The National Science Foundation to A.D.L. (DEB-0614429), and Fundação para a Ciência e a Tecnologia (FCT/FEDER POCTI/BSE/48228/2002) and Fundação Calouste Gulbenkian to H.T.

AUTHOR CONTRIBUTIONS

H.T., I.M.C. and M.B. performed the experiments. H.T., I.M.C. and A.D.L. analysed the data. H.T., M.R.R. and A.D.L. conceived the project and wrote the manuscript.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Cohan, F.M. Can uniform selection retard random genetic divergence between isolated conspecific populations. *Evolution Int. J. Org. Evolution* **38**, 495–504 (1984).
2. Travisano, M., Mongold, J.A., Bennett, A.F. & Lenski, R.E. Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* **267**, 87–90 (1995).
3. Crill, W.D., Wichman, H.A. & Bull, J.J. Evolutionary reversals during viral adaptation to alternating hosts. *Genetics* **154**, 27–37 (2000).
4. Teotónio, H. & Rose, M.R. Variation in the reversibility of evolution. *Nature* **408**, 463–466 (2000).
5. Bull, J.J. & Charnov, E.L. On irreversible evolution. *Evolution Int. J. Org. Evolution* **39**, 1149–1155 (1985).
6. Teotónio, H. & Rose, M.R. Perspective: reverse evolution. *Evolution Int. J. Org. Evolution* **55**, 653–660 (2001).
7. Teotónio, H., Matos, M. & Rose, M.R. Reverse evolution of fitness in *Drosophila melanogaster*. *J. Evol. Biol.* **15**, 608–617 (2002).
8. Grant, P.R. & Grant, B.R. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**, 707–711 (2002).
9. Whiting, M.F., Bradler, S. & Maxwell, T. Loss and recovery of wings in stick insects. *Nature* **421**, 264–267 (2003).
10. Whitlock, M.C., Phillips, P.C. & Fowler, K. Persistence of changes in the genetic covariance matrix after a bottleneck. *Evolution Int. J. Org. Evolution* **56**, 1968–1975 (2002).
11. Porter, M.L. & Crandall, K.A. Lost along the way: the significance of evolution in reverse. *Trends Ecol. Evol.* **18**, 541–547 (2003).
12. Colosimo, P.F. *et al.* Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* **307**, 1928–1933 (2005).
13. Teotónio, H., Matos, M. & Rose, M.R. Quantitative genetics of functional characters in *Drosophila melanogaster* populations subjected to laboratory selection. *J. Genet.* **83**, 265–277 (2004).
14. Hudson, R.R., Sáez, A.G. & Ayala, F.J. DNA variation at the *Sod* locus of *Drosophila meanogaster*: an unfolding story of natural selection. *Proc. Natl. Acad. Sci. USA* **94**, 7725–7729 (1997).

15. Deckert-Cruz, D.J., Tyler, R.H., Landmesser, J.E. & Rose, M.R. Allozymic differentiation in response to laboratory demographic selection of *Drosophila melanogaster*. *Evolution Int. J. Org. Evolution* **51**, 865–872 (1997).
16. Verrelli, B.C. & Eanes, W.F. Clinal variation for amino acid polymorphisms at the *Pgm* locus in *Drosophila melanogaster*. *Genetics* **157**, 1649–1663 (2001).
17. Balakirev, E.S., Anisimova, M. & Ayala, F.J. Positive and negative selection in the *β-Esterase* gene cluster of the *Drosophila melanogaster* subgroup. *J. Mol. Evol.* **62**, 496–510 (2006).
18. Smith, J.M. & Haigh, J. The hitch-hiking effect of a favourable gene. *Genet. Res.* **23**, 23–35 (1974).
19. Hartl, D.L. & Clark, A.G. *Principles of Population Genetics* (Sinauer, Sunderland, Massachusetts, 1989).
20. Macdonald, S.J., Pasten, T. & Long, A.D. The effect of polymorphisms in the *Enhancer of split* gene complex on bristle number variation in a large wild-caught cohort of *Drosophila melanogaster*. *Genetics* **171**, 1741–1756 (2005).
21. Scheet, P. & Stephens, M. Fast and flexible model for LD. *Am. J. Hum. Genet.* **78**, 629–644 (2006).
22. McVean, G., Awadalla, P. & Fearnhead, P. A coalescent-based method for detecting and estimating recombination rates from gene sequences. *Genetics* **160**, 1231–1241 (2002).
23. Nei, M. & Tajima, F. Genetic drift and the estimation of effective population size. *Genetics* **98**, 625–640 (1981).
24. Goldringer, I. & Bataillon, T. 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 (2004).
25. Nei, M. Definition and estimation of fixation indices. *Evolution Int. J. Org. Evolution* **40**, 643–645 (1986).
26. Churchill, G.A. & Doerge, R.W. Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963–971 (1994).
27. Hermisson, J. & Pennings, P.S. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**, 2335–2352 (2005).
28. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2006).
29. Haag-Liautard, C. *et al.* Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature* **445**, 82–85 (2007).