Reverse evolution of fitness in *Drosophila melanogaster*

H. TEOTÓNIO,* M. MATOS† & M. R. ROSE*

*Department of Ecology and Evolutionary Biology, University of California Irvine, Irvine, CA, USA
†Centro de Biologia Ambiental, Department Zoologiae Antropologia, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, Portugal

Keywords:

adaptation;
experimental reverse evolution;
fitness;
life-history;
plasticity.

Abstract

The evolution of fitness is central to evolutionary theory, yet few experimental systems allow us to track its evolution in genetically and environmentally relevant contexts. Reverse evolution experiments allow the study of the evolutionary return to ancestral phenotypic states, including fitness. This in turn permits well-defined tests for the dependence of adaptation on evolutionary history and environmental conditions. In the experiments described here, 20 populations of heterogeneous evolutionary histories were returned to their common ancestral environment for 50 generations, and were then compared with both their immediate differentiated ancestors and populations which had remained in the ancestral environment. One measure of fitness returned to ancestral levels to a greater extent than other characters did. The phenotypic effects of reverse evolution were also contingent on previous selective history. Moreover, convergence to the ancestral state was highly sensitive to environmental conditions. The phenotypic plasticity of fecundity, a character directly selected for, evolved during the experimental time frame. Reverse evolution appears to force multiple, diverged populations to converge on a common fitness state through different life-history and genetic changes.

Introduction

One of the most important topics in evolutionary biology is adaptation, particularly the selective and genetic mechanisms that lead to adaptation. Among the many empirical problems that face the study of adaptation is the inherent difficulty of inferring adaptation from the evolution of characters whose relationship with fitness may be questionable. Another important problem facing many empirical studies of adaptation is that they are frequently carried out in environmental conditions which are of unclear evolutionary relevance, if not actually a source of artefacts (vid. Lewontin, 1974; Wright, 1977; Haymer & Hartl, 1982; Hedrick & Murray, 1983; Service & Rose, 1985; Leroi *et al.*, 1994; Reznick & Travis, 1996; Rose *et al.*, 1996; Chippindale *et al.*, 2001). The obvious solution to these difficulties is to study the evolution of fitness itself in interpretable selective contexts, a solution that is also hard to achieve.

Historically, the direct estimation of fitness in sexually reproducing organisms has been difficult, because fitness measures of single genotypes in natural genetic and environmental contexts are usually impossible to obtain, unlike the situation for clonal organisms (e.g, Lenski *et al.*, 1991; Crill *et al.*, 2000). When isogenicity can be contrived, fitness may be estimated, but inbreeding effects can have devastating effects on the fitness of each line, and the whole set of inbred lines will often represent a biased sample of the genetic variation found in the original population. For a long time, intraspecific and interspecific competition experiments have been used in *Drosophila* to infer relative genotypic fitness (Latter & Robertson, 1962; Prout, 1971; Wright, 1977; Jungen & Hartl, 1979; Haymer & Hartl, 1982; Hedrick & Murray, 1983; Mackay, 1986; Wilton *et al.*, 1989; Joshi & Thompson, 1995; Weber, 1996; Fowler *et al.*, 1997; Barton & Partridge, 2000). Differences in relative fitness between wild-caught flies and their laboratory derivatives or between laboratory populations have usually

Correspondence: Henrique Teotónio, Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08540–1003, USA
Tel.: +609 2585587; fax: +609 2581334;
e-mail: teotonio@princeton.edu

been found. But inbreeding and genotype-by-environment interaction make it hard to establish a relationship between these fitness differences and the genetic and environmental circumstances in which the original populations evolved, which renders the scientific value of such findings highly uncertain (Rose *et al.*, 1996).

To address these problems in the study of adaptation one needs, at a minimum, an experimental system in which it is possible to estimate the fitness of members of genetically defined populations relative to the fitness of members of control populations that have already adapted to the experimental environment. If the experimental populations are allowed to evolve in the environment of the previously adapted control populations, their evolution has a well-defined point of reference. One such design is experimental reverse evolution, where the ancestral populations or their equivalent are available for comparison with derived differentiated populations (e.g. Lenski, 1988; Service *et al.*, 1988; Rainey & Travisano, 1998; Burch & Chao, 1999; Crill *et al.*, 2000; Teotónio & Rose, 2000, 2001). This was the experimental paradigm used in the present study. Note that this type of experimental paradigm is very different from both relaxed artificial selection and reverse artificial selection, most importantly with respect to the population sizes involved, the opportunity for natural selection to operate, and the degree of replication. Experimental evolution, forwards or reverse, involves more and larger populations compared with artificial selection, and natural selection is deliberately allowed to occur (see Wright, 1977; Rose *et al.*, 1996; Teotónio & Rose, 2001).

Reverse evolution in sexually outbreeding *Drosophila* populations appears to depend on previous evolutionary history, particularly with respect to the degree of convergence. Yet incomplete convergence to ancestral states could not be explained for most studied characters either by lack of genetic variation or the presence of gene-interaction systems in these *Drosophila* populations (Teotónio & Rose, 2000). However, this earlier study did not consider the reverse evolution of fitness itself, being limited to characters whose relationship to fitness is less obvious. Thus, it was not ascertained whether adaptation is itself dependent on previous evolutionary history. Also, the selective mechanisms that determine reverse evolution were not identified.

In this study we measure a total of 10 characters, including life-history characters and competitive fitness relative to a morphologically marked stock, for populations which have been kept in diverse environments for numerous generations, and their descendent populations selected in the environment of their common ancestor for 50 generations (vid. Teotónio & Rose, 2000). We specifically address: (1) *whether* heterogeneous populations derived from a common ancestor will attain similar degrees of adaptation during reverse evolution; (2) *whether* the evolutionary dynamics of fitness characters are similar to characters whose relation with fitness is less obvious; and (3) the relationship between the ancestral environment and the selective mechanisms behind reverse evolution.

Materials and methods

Experimental populations

All populations were derived from a common ancestor introduced to the laboratory in 1975, the 'Ives' population (Rose, 1984; Teotónio & Rose, 2000). In 1980, five populations were derived and ever-since maintained under Ives-ancestral conditions (B_{1-5}), which comprise a 2-week generation cycle. Another five populations were selected for increased late-life fertility (O_{1-5}), being maintained on a 10-week generation cycle (Rose, 1984). In 1989, populations derived individually from the O populations were selected for increased starvation resistance (SO_{1-5}), with another group maintained as the fed demographic control (CO_{1-5}) (Rose *et al.*, 1992). Both these population-groups were maintained on a 3–4 week generation cycle. Lastly, from the CO populations, a set of five populations was selected for decreased developmental time and increased early fertility (ACO_{1-5}) (Chippindale *et al.*, 1997). These were maintained with generation lengths no longer than 9 days. All populations have been maintained at high population sizes ($n > 1500$), without systematic inbreeding or hybridization, throughout their history. By the time the experiments described here were performed, the O populations had undergone 110 generations of selection for late-life fitness, the SO and CO had undergone 130 generations in their selective environment, and the ACO populations 270 generations.

New populations were obtained from the four groups of selected populations (O, SO, CO and ACO), each new population derived from the same numbered replicate ancestor population (e.g. IO_1 derived from O_1, IO_2 derived from O_2, etc.), and cultured in the common ancestral Ives-environment for 50 generations (Teotónio & Rose 2000). This environment features discrete 2-week generations, with natural selection for early fertility in crowded conditions. Egg-laying occurs within 2 h, density being controlled at 50–100 eggs per vial. Each of the five IB populations was derived from a single B population. These were the control populations in the experimental design. Approximately 440 generations had elapsed since their formation by the time the reverse evolution experiment started (Teotónio & Rose, 2000).

A total of 45 populations were studied: the control IB_{1-5}, the 20 differentially selected populations (O_{1-5}, CO_{1-5}, ACO_{1-5}, and SO_{1-5}), and the 20 returned to the Ives environment (IO_{1-5}, ICO_{1-5}, $IACO_{1-5}$, and ISO_{1-5}). All characters were measured at generation 50 of reverse evolution.

Competitive fitness assays

Three competitive fitness assays were done: male, female, and population fitness. In each of them, morphologically marked populations were competed against the experimental populations in Ives-environment conditions.

For the male fitness assay, five vials were used per population comparison, each containing 25 eggs of the population to be tested and 75 eggs of a brown-eyed *bw* competitor outbred stock (supplied by A.K. Chippindale, UCSB). After 14 days, 20 *bw* females from each assay vial were isolated into separate vials. The proportion of these females with *wild-type* (*wt*) progeny (as a consequence of being fathered by males from the experimental population) was taken as the male fitness estimate. This includes male survivorship, mating success, and viability of offspring (see also Chippindale *et al.*, 2001). Although most vials had all *bw* or *wt* progeny, in approximately 20% of them the progeny were of varied parental origin, because of multiple mating. In such cases, each vial was given a score of 0.75 if the majority of the progeny had *wt* eye colour (65–95%), 0.5 if approximately half was *wt* (35–65%), and 0.25 if a minority of flies were *wt* (5–35%). This scoring system was adopted because the volume of data rendered exact counts impractical.

For the female fitness assay, five pairs of *wt* adults were placed together with 15 pairs of *bw* adults in eight replicate vials per population tested. The females were allowed to lay eggs for 2 h after which all adults were discarded. Egg density was controlled by removing excess eggs. After 14 days, female fitness was estimated as the proportion of *wt* progeny. This estimate includes female survivorship, fertility, and offspring viability.

The population fitness assay employed a two-generation competition contrived so that the marked population could not hybridize with the experimental populations. This was done using a compound autosome stock as the marked competitor because all hybrid progeny with the experimental population will be inviable (Jungen & Hartl, 1979; Novitski *et al.*, 1981). The marked stock no. 1113, with genotype *C(2)EN, b[1] bw[1]; st[1]*, was obtained from the Bloomington *Drosophila* Stock Center. In this assay, seven replicate vials were used per population. In each vial, three pairs of adults from the experimental population were placed together with 17 pairs of the marker stock. These adults were allowed to lay eggs for 24 h, after which egg density was controlled. After 14 days all emerged adult flies were scored for their genotype and allowed to lay for 24 h under crowded conditions for a second generation, egg density again being controlled. The proportion of *wt* adult progeny, after the first and second generation of competition, was used as an estimate of population fitness.

Life-history assays

For developmental time, 10 vials per population were collected with 60 eggs each and emerging adults were counted every 6 h. Population viability, regardless of sex, was taken as the egg to adult survivorship. In the starvation resistance assay, four same-sex flies were placed in each of 10 replicate vials under high humidity conditions but no food. Survivorship was scored every 6 h (Teotónio & Rose 2000).

Early fecundity under low crowding conditions was measured by placing one pair (one male and one female) of 14-day-old-flies in each of the 30 vials assayed per population. These flies were allowed to lay eggs for 24 h. For early fecundity under high crowding conditions, 20 pairs were used for each vial, with 1, 2 and 6 h for the laying period. Eight vials for the 1 h assay and five vials for the 2 and 6 h assays were collected from each population.

Data for developmental time, starvation resistance, and 1 h early fecundity under high crowding conditions for the reverse evolved populations IO_{1-5}, ISO_{1-5}, ICO_{1-5} and $IACO_{1-5}$ were previously reported in Teotónio & Rose (2000).

Data analysis

For each selection history, there are five independent replicate populations. Because of this, our observational units are the mean values of these populations. All group comparisons were made relative to the IB populations.

Two sets of hypotheses were tested. First, we tested whether particular independent groups of populations with different evolutionary histories converged to control IB values. For example, we tested whether O_{1-5} populations were differentiated from the IB populations, and whether the IO_{1-5} converged to the IB control levels. This was done with unpaired two-tailed Student's *t*-tests, with α set to 0.05. Sequential Bonferroni was used for multiple comparison corrections on each character separately, under the composite null hypothesis that all group differences to control are equal to zero (Rice, 1989). The number of tests for all characters were eight, with the exception of fecundity where the number of tests were 32.

Secondly, we tested the general question of convergence to the ancestral values, and the role that previous history has on reverse evolution, using a mixed analysis of variance (ANOVA) with a fixed effects 'selection' factor and a random effects 'history' factor. The historical factor was reflected in the division of populations into four groups according to their specific evolutionary history (the antecedent selection in O, CO, ACO or SO populations). For the selection factor, the populations were divided into two groups: those which had recently undergone Ives-culture and those which had not. A significant interaction would reveal historical effects on

adaptation. For each population the data used were the absolute differences from the mean of the five IB populations. This was done because differences in sign might underestimate average differentiation among 'selection' categories and overestimate the interaction term between 'history' and 'selection', as a result of sign differences in initial differentiation.

For the analysis of differential plasticity in fecundity, analysis of covariance (ANCOVA) was performed on the 1, 2 and 6 h assays. A separate ANCOVA was performed for each of the four evolutionary histories (that is, one for the O ancestry group, including both IO_{1-5} and O_{1-5} populations, one for the CO ancestry group, etc.). Selection (Ives selection vs. non-Ives selection) was the fixed factor with number of hours as the covariate. A significant interaction between 'selection' and 'hours' would indicate a change with reverse evolution in the relationship between observed fecundity and the number of hours allowed for egg laying.

Results

Competitive fitness

The results for the male fitness assay are summarized in Fig. 1a and Table 1. Student *t*-tests revealed that the immediately ancestral O_{1-5}, SO_{1-5}, and ACO_{1-5} populations were differentiated from the IB populations, the CO_{1-5} being marginally differentiated ($P = 0.07$), whereas all reversely evolved populations, IO_{1-5}, ISO_{1-5}, ICO_{1-5}, and $IACO_{1-5}$ converged to control levels. ANOVA indicated that type of selection (populations cultured in Ives vs. non-Ives conditions) to be highly significant, although incidental evolutionary history (O, SO, CO, and ACO ancestry grouping) and the interaction between them were nonsignificant (see Table 1).

Female fitness results are shown in Fig. 1b. As tested with *t*-tests no group of populations was different from the control populations. However, ANOVA on female fitness data showed both a significant 'selection' and 'history' factor (Table 1).

For the population fitness assay, in the first generation, the SO populations were heteroscedastic, as measured by Bartlett's tests. Angular transformation partially corrected this and transformed data were used in subsequent analysis of all populations. Comparisons using *t*-tests revealed that ACO_{1-5} ($P = 0.02$) were significantly less fit in the assay's first generation. All other populations were undifferentiated (see Fig. 1c). However, the CO_{1-5} ($P = 0.07$), IO_{1-5} ($P = 0.07$), and O_{1-5} ($P = 0.08$) were marginally differentiated from control levels. By the second generation of the population fitness assay, most of the populations undergoing assay had cleared the experimental vials of morphologically marked competitor flies (not shown). For this reason, multigenerational statistical analyses could not be pursued. ANOVA did not reveal any significant effects of any factor, when each generation was analysed separately.

Life-history characters

Results from population viability, female developmental time, and female starvation resistance are shown in Fig. 2 and Table 1.

For female developmental time, all but the IO populations were differentiated from the control populations, by *t*-tests. Male developmental time results are similar to female results, with the exception of the ISO populations, which for differences were not significant. ANOVA shows that both selection and interaction between selection and history are significant.

The CO, ICO and O populations were different from the IB populations with respect to viability (Fig. 2). The interaction between selection and history in the ANOVA for population viability was significant (Table 1).

Analysis of both female and male starvation resistance showed that all but the ICO and IACO populations were differentiated from the IB populations. Also for both sexes, only the interaction term in the ANOVA was significant (Table 1). As the SO treatment has disproportionately higher starvation resistance values than those of any other group of populations (Fig. 2), we re-analysed the data using the absolute differences of the assayed populations from the control populations divided by the value of the immediately ancestral differentiated population group. This analysis was undertaken in order to minimize any effects of initial differentiation on the degree of convergence, which could adversely affect the estimation of the interaction term. The ANOVA in this case showed for females that both type of selection ($F_{1,3} = 17.5$; $P = 0.03$) and the interaction term ($F_{3,32} = 3.35$; $P = 0.03$) are significant. In males, selection is marginally significant ($F_{1,3} = 6.41$; $P = 0.09$), whereas the interaction is highly significant ($F_{3,32} = 10.3$; $P < 0.01$).

The early fecundity results are presented in Fig. 3 and also in Table 1. At low adult density with 24 h of egg-laying, ACO and SO are different from the IB values as tested with *t*-tests. The interaction term was the only factor found significant in the ANOVA.

All populations but ISO are differentiated from the control for early fecundity at high density with 1 h egg-laying. In the 2 h assay all but ISO and ICO groups are different from the control. Finally, analysis of the 6-h assay showed that only O, SO, ACO were significantly different from the control. The fecundity data indicate a return to the ancestral fecundity character with increased time of egg-laying in the reverse evolution populations (Fig. 3). ANOVA revealed that type of selection had significant effects in the 1- and 2-h assays, although the interaction was significant for the 1, 2 and 6 h assays (for the 1 h at the 6% level; Table 1).

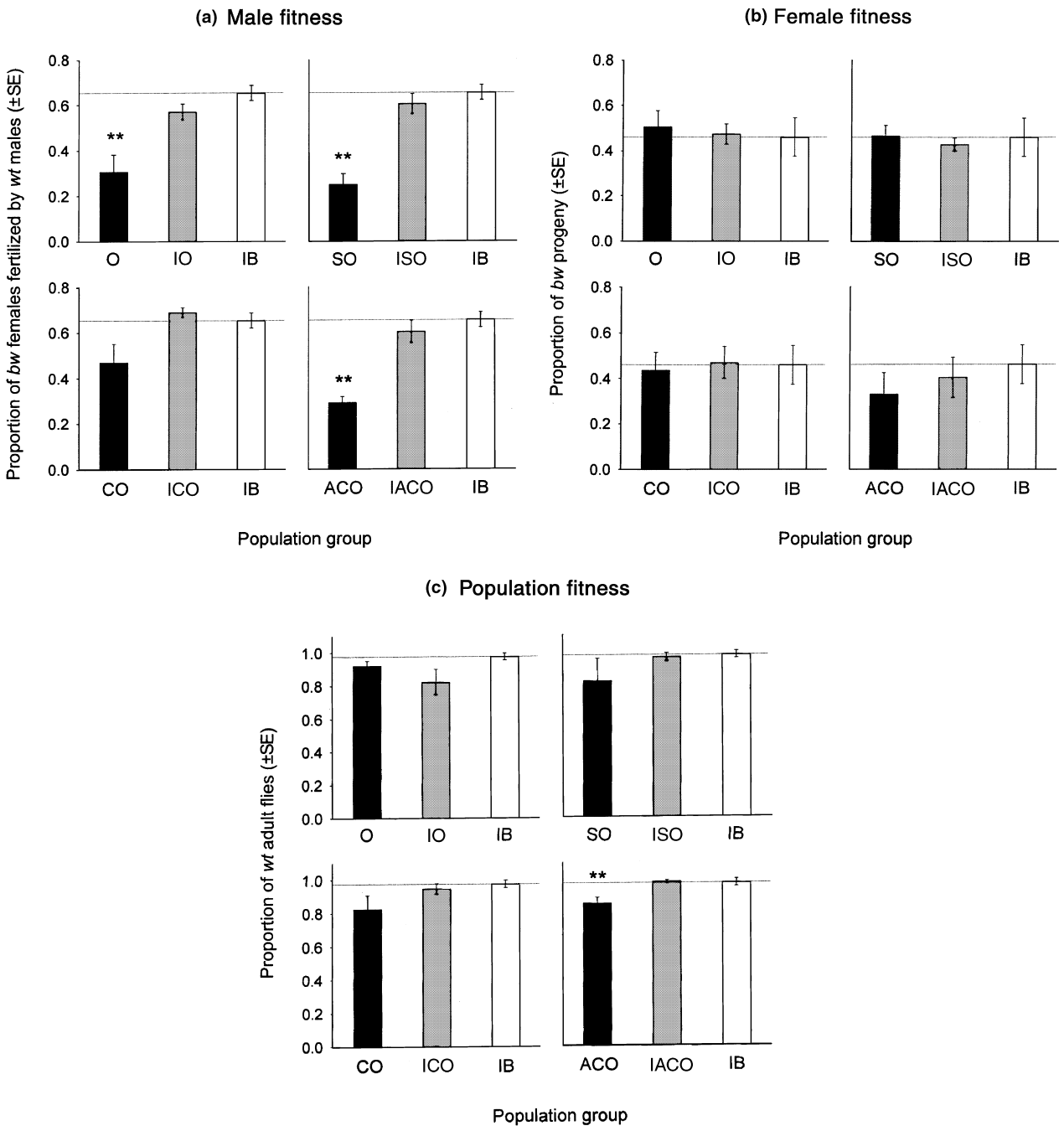

Fig. 1 In (a) Male fitness results are shown. Male fitness was measured as the proportion of tester marked stock females fertilized by males from experimental populations (see Materials and methods). Differentially selected populations are shown by black bars, whereas populations derived from them after 50 generations of reverse evolution are shown by grey bars. Control populations are represented by white bars and their mean value by a horizontal grey line. In (b), female fitness measured as the number of offspring mothered by the tester populations is shown. In (c), generation one of the population fitness assay is shown as the untransformed proportion of experimental flies that survived in competition to marked tester flies; angular transformation was used for the analysis. Error bars represent the standard error of the five replicate populations, for each group. Two-tailed Student's *t*-tests: $*0.01 < P \leq 0.05$; $**P \leq 0.01$. Sequential Bonferroni with number of tests equal to eight: †-tests that lose significance with $\alpha = 0.05$.

Table 1 Mean Squares for a two-factor mixed ANOVA table with selection factor fixed and history factor random.

	Fitness			Developmental Time			Starvation Resistance		Fecundity			
	♀	♂	Population †	♀	♂	Viability	♀	♂	1-hour	2-hour	6-hour	low-density
Selection df=1	0.013*	0.615 **	0.035	1518.8**	1309.5**	2.2E-4	21317	12156	127478*	461081*	816896	793.3
History df=3	0.028**	0.035	0.004	49.7	77.1	2E-3	17105	8773	2814	43050	125887	159.3
Selection × History df=3	9E-4	0.014	0.03	37.6*	36.5*	2.2E-3*	8808**	5778**	6698	46371**	168943**	259.5*
Error df=32	5.3E-3	0.012	0.021	8.64	8.72	6.2E-4	390	122	2412	4238	19921	74.3

significance testing of F statistics: *0.01< p ≤ 0.05; **p ≤ 0.01.
† generation 1 of angular transformed population fitness assay.

The ANCOVA performed on the fecundity data shows that the fecundity characters of all populations responded to 50 generations of Ives-culture, as the selection factor was significant, irrespective of history (not shown). A change of the linear relationship of fecundity with hours was found for the populations with O and SO ancestry, shown by a significant interaction term ($F_{1,26} = 9.2$, $P < 0.01$ for O ancestry; $F_{1,26} = 5.8$, $P = 0.02$ for the SO ancestry), but it was not found for populations with CO and ACO ancestry.

Discussion

Reverse evolution and measures of fitness

Our findings show that the evolutionary dynamics of reverse evolution depend in part on previous selective history, as many of the ANOVA interaction terms reflecting selective ancestry are significant. Our results also show that phenotypic convergence is not universal, particularly for characters whose relationship to fitness

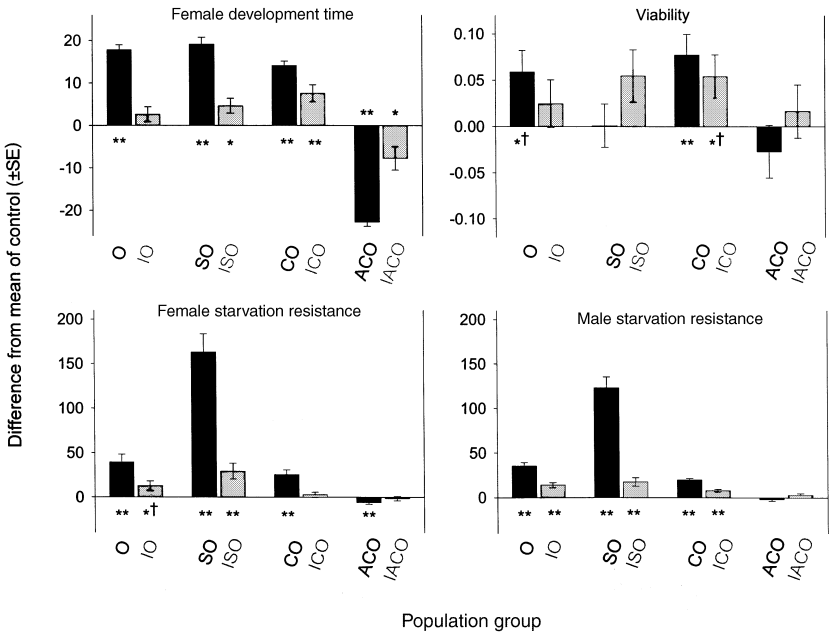

Fig. 2 The upper-left panel shows female developmental time as a difference from control in number of hours from egg to adult. Black bars show the ancestor populations, whereas grey show the reverse evolved populations. Similar results were found for male developmental time data, except that ISO populations were not different from controls. Upper-right panel shows population viability as a difference from control levels in the probability of survival to adulthood. The bottom panels show female and male starvation resistance in number of hours surviving until death. Error bars represent the standard error of the difference between a particular group of populations from the control populations. The development time and starvation resistance data presented for the reverse-evolved populations was previously reported in Teotónio & Rose (2000). Student's *t*-tests: *$0.01 < P \leq 0.05$; **$P \leq 0.01$. Sequential Bonferroni with number of tests equal to eight: †tests that lose significance with $\alpha = 0.05$.

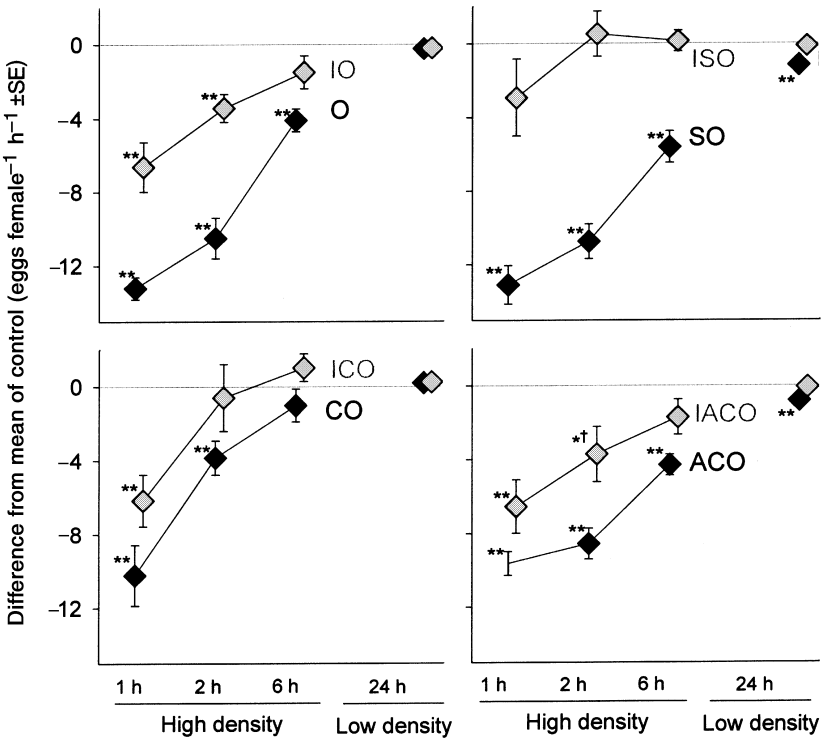

Fig. 3 Early fecundity mean replicate values ± standard error of the difference, with adult density and egg-laying time as environmental variables. All points show the mean difference of each group of populations from the control. The error bars in the low-density environment are too small to be visualized. Student's *t*-tests: *$0.01 < P \leq 0.05$; **$P \leq 0.01$. Sequential Bonferroni with number of tests equal to 32: †tests that lose significance with $\alpha = 0.05$.

is less clear. Similar results have been obtained in experimental studies of adaptation in bacteria (e.g. Travisano *et al.*, 1995). These conclusions are in agreement with several other studies of reverse evolution in both sexual and asexual species (Lenski, 1988; Service *et al.*, 1988; Rainey & Travisano, 1998; Crill *et al.*, 2000; Teotónio & Rose 2000, 2001).

Among the three competitive fitness estimates, the one that gave the clearest results was male fitness. This character was differentiated considerably in the starting populations, converging to a level not significantly different from that of the common ancestral control by generation 50, for all populations selected in the ancestral environment. The particular components of male fitness which are being selected during reverse evolution is unknown and may be different from population to population. For example, rapidity to achieve sexual maturity might be important in populations that take more time to develop and mature than the ancestral populations. In populations selected for late-life fertility, increased starvation resistance, and intermediate age of reproduction, individuals may not attain their peak mating ability by the time that mating usually occurs in the ancestral environment (Rose, 1984; Leroi *et al.*, 1994). In populations selected for accelerated development, on the other hand, selection should be different (Chippindale *et al.*, 1997). Because the generation length of these particular populations is set to 9 days, a male fitness decline after this day may have evolved as a result of a trade-off with very early mating success, producing radically accelerated male senescence. But these possibilities are only speculative at this point.

The total fitness assays did not give much useful information. Differences in total population fitness between the differentiated populations and the controls were negligible, with the exception of the competitive fitness of those selected for accelerated development during the first generation of assay. Taking these results separately from those of the second assay generation they reveal that significant female competitive fitness differences were present, despite the fact that we did not observe those differences in the female fitness assay. However, multigeneration population competition did not reveal any fitness differentiation involving the accelerated development stocks and thus we must conclude that all populations had similar competitive population fitness.

The multigenerational competition did not reveal significant population differences because of the very poor performance of the mutant stock employed. The use

of this type of assay is attractive in strictly sexually reproducing organisms because it can directly assess a full generation of competitive success for a genetically unmanipulated experimental population, by contrast with techniques involving the use of inbreeding to estimate the fitness of single chromosomes (e.g. Fowler *et al.*, 1997; Barton & Partridge, 2000). In our case however, it proved a poor experimental technique.

Although there was a significant selection and historical effect in female competitive fitness estimates, these revealed little differentiation, before or after reverse evolution. This is surprising, given that the ancestral environment imposes selection for reproduction at day 14 after egg collection, and these conditions were closely replicated. The lack of differentiation for females may be due to one of two causes. First, perhaps the test employed is not sensitive enough to population differences because the competitor population has such a good performance in this environment that it masks any differences (but see Chippindale *et al.*, 2001). Secondly, the lack of differentiation may be real. The question is then what component of female fitness is actively being selected, if any? The results for fecundity characters measured directly reveal extensive differentiation of populations. An explanation for the lack of female fitness differentiation might have to wait for experiments which can test the extent to which fecundity plasticity reveals female fitness differences (see below).

It is interesting to note that although male fitness was highly differentiated and responsive to selection, female fitness was neither. As in previous studies with *Drosophila,* the two characters are apparently able to evolve separately from each other (Rice, 1996; Chippindale *et al.*, 2001). The fact that the Pearson correlation coefficient between male and female fitness is equal to -0.002 ± 0.023 reinforces this conclusion.

Alternative routes to adaptation

The specific life-history mechanisms behind reverse evolution were tackled by measuring life-history characters which are closely related to fitness in the ancestral environment: developmental time, starvation resistance and particularly viability and fecundity. Generally the life-history data do not indicate a widespread return to ancestral values after 50 generations of reverse evolution. Our results are not due to a lack of relevant genetic variation or presence of an epistatic genetic architecture (Teotónio & Rose, 2000). In contrast to this, and to the best of our resolution, the significant results from the male competitive fitness assays show a more pronounced convergence to ancestral levels.

These two patterns together suggest that similar degrees of adaptation were achieved through different underlying life-history and genetic mechanisms (cf. Prout, 1971; Wright, 1977; Cohan, 1984; Cohan & Hoffman, 1989; Travisano & Lenski, 1996; Crill *et al.*, 2000). The populations previously selected for increased starvation resistance reverted to ancestral character values in fecundity, remaining differentiated for starvation resistance and developmental time. The populations with accelerated development did not return to ancestral values for fecundity or developmental time. The late-life fertility populations also did not completely return to ancestral fecundity or starvation resistance levels. Nevertheless, all these groups of populations returned to ancestral male fitness levels.

Alternative routes to adaptation have been experimentally observed in *Drosophila* through asymmetry of correlated responses to selection in alternative environments (Shiotsugu *et al.*, 1997) or similar environments (e.g. Cohan & Hoffman, 1989; see also Bohren *et al.*, 1966; Gromko, 1995). Our study however, is one of the first to attempt a direct measure of fitness. Microbial systems have long permitted the simultaneous estimation of fitness and life-history characters. It is well established in the microbial literature that different genetic, physiological and life-history mechanisms are implicated in similar adaptive responses. Generally, this is supported by numerous studies on the evolution of bacterial resistance to phages and antibiotics (e.g. Lenski, 1988; Cohan *et al.*, 1994; review in Lenski, 1998), bacterial adaptation to minimal media (e.g. Travisano *et al.*, 1995; Travisano & Lenski, 1996), viral adaptation to alternative hosts (Crill *et al.*, 2000), and also by the adaptive recovery of deleterious mutations in both viruses and bacteria (Burch & Chao, 1999; Moore *et al.*, 2000).

The role that sexual recombination might have in these evolutionary patterns and their genetic basis has however, not been adequately determined. For example, it has been experimentally determined that bacterial evolution during antibiotic exposure and recovery involve compensatory genetic changes, thus revealing an epistatic mode of action. In sexual populations, because of recombination, such cases might not be so common (Teotónio & Rose, 2001).

Environmental sensitivity of adaptation

Our study illustrates the importance of characterizing adaptation in evolutionarily relevant environments. In particular it shows some of the pitfalls that can occur when not describing the environment properly. For example, if we had measured early fecundity only at low density, no differences between selected populations and controls would have been detected. Furthermore, if early fecundity at high density with 1-h of egg-laying had been the only fecundity character studied, it might have been concluded that reverse evolution did not lead to convergence for fitness characters (Teotónio & Rose, 2000). Nonetheless, with more time for egg-laying, measurable evolutionary convergence occurs. The rapidity of

egg-laying might be what is selected for in the ancestral environment, as well as being the evolutionary constraint involved during convergence. The reverse evolution of fitness might then be explained by different rates of response to an environmental transition at the time of egg-laying, such as the contact with fresh food, CO_2, or the presence of laid eggs. Both amount of nutrients and adult crowding have been previously implicated in the evolution of early fecundity in some of our laboratory populations (Leroi *et al.*, 1994). Here the data indicate that the speed of response to an environmental transition was the focus of selection for early fecundity. This is a striking example of the environmental specificity and dependence of adaptation (Travisano & Lenski, 1996). The findings on fecundity also indicate that the change in plasticity for a fitness character during reverse evolution depended on previous selection history.

Despite the intricacy of these details, it is nonetheless clear that plasticity is eminently evolvable, even in an evolutionary time frame as short as 50 generations.

Complexity in experimental evolution

Overall our findings illustrate the complexity of studying the evolutionary processes of adaptation. Clear expectations of the outcome of reverse evolution were defined *a priori*, but were not fully met. Presumably this reflected the effects of previous selection on the populations analysed, the variable associations between life-history characters and fitness, and the subtle environmental dependence of the expression of these characters. This complexity should not be underestimated, both when designing studies involving experimental evolution and when interpreting their results.

Acknowledgments

For essential technical help we thank Y. Chau, T. Vu, N. Vu, C. Hammerle and E. Gass. A. K. Chippindale and L. D. Mueller have given helpful advice throughout the project. C. L. Burch, G.A.C. Bell, and two anonymous reviewers improved the manuscript with their suggestions. M. Matos received a travel grant from FLAD. H. Teotónio was supported by the Gulbenkian Foundation, FLAD, and PRAXIS / FCT (Portugal).

References

Barton, N.H. & Partridge, L. 2000. Measuring fitness by means of balancer chromosomes. *Genet. Res. Camb.* **75**: 297–313.

Bohren, B.B., Hill, W.G. & Robertson, A. 1966. Some observations on asymmetrical correlated responses to selection. *Genet. Res. Camb.* **7**: 44–57.

Burch, C.L. & Chao, L. 1999. Evolution by small steps and rugged landscapes in the RNA virus ϕ6. *Genetics* **151**: 921–927.

Chippindale, A.K., Alipaz, J.A., Chen, H.-W. & Rose, M.R. 1997. Experimental evolution of accelerated development in *Drosophila*. 1. Development speed and larval survival. *Evolution* **51**: 1536–1551.

Chippindale, A.K., Gibson, J.R. & Rice, W.R. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in Drosophila. *Proc. Natl. Acad. Sci. USA* **98**: 1671–1675.

Cohan, F.M. 1984. Can uniform selection retard random genetic divergence between isolated conspecific populations. *Evolution* **38**: 495–504.

Cohan, F.M. & Hoffman, A.A. 1989. Uniform selection as a diversifying force in evolution: evidence from *Drosophila*. *Am. Nat.* **134**: 613–637.

Cohan, F.M., King, E.C. & Zawadzki, P. 1994. Amelioration of the deleterious pleiotropic effects of an adaptive mutation in *Bacillus subtilis*. *Evolution* **48**: 81–95.

Crill, W.D., Wichman, H.A. & Bull, J.J. 2000. Evolutionary reversals during viral adaptation to alternating hosts. *Genetics* **154**: 27–37.

Fowler, K., Semple, C., Barton, N.H. & Partridge, L. 1997. Genetic variation for total fitness in *Drosophila melanogaster*. *Proc. R. Soc. Lond. B* **264**: 191–199.

Gromko, M.H. 1995. Unpredictability of correlated response to selection: pleiotropy and sampling interact. *Evolution* **49**: 685–693.

Haymer, D.S. & Hartl, D.L. 1982. The experimental assessment of fitness in *Drosophila*. I. Comparative measures of competitive reproductive success. *Genetics* **102**: 455–466.

Hedrick, P.W. & Murray, E. 1983. Selection and measures of fitness. In: *Genetics and Biology of Drosophila* (M. Ashburner, H. L. Carson & J. N. Thompson, Jr, eds), pp. 61–104. Academic Press, Orlando.

Joshi, A. & Thompson, J.N. 1995. Alternative routes to the evolution of competitive ability in two competing species of *Drosophila*. *Evolution* **49**: 616–625.

Jungen, H. & Hartl, D.L. 1979. Average fitness of populations of *Drosophila melanogaster* as estimated using compound-autosome strains. *Evolution* **33**: 371–380.

Latter, B.D.H. & Robertson, A. 1962. The effects of inbreeding and artificial selection on reproductive fitness. *Genet. Res. Camb.* **3**: 110–138.

Lenski, R.E. 1988. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. II. Compensation for maladaptive effects associated with resistance to virus T4. *Evolution* **42**: 433–440.

Lenski, R.E. 1998. Bacterial evolution and the cost of antibiotic resistance. *Intern. Microbiol.* **1**: 265–270.

Lenski, R.M., Rose, R., Simpson, S.C. & Tadler, S.C. 1991. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am. Nat.* **138**: 1315–1341.

Leroi, A.M., Chippindale, A.K. & Rose, M.R. 1994. Long-term laboratory evolution of a genetic life-history trade-off in *Drosophila melanogaster*. 1. The role of genotype-by–environment interaction. *Evolution* **48**: 1244–1257.

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

Mackay, T.F.C. 1986. A quantitative genetic analysis of fitness and its components in *Drosophila melanogaster*. *Genet. Res. Camb.* **47**: 59–70.

Moore, F.B.-G., Rozen, D. & Lenski, R.E. 2000. Pervasive compensatory adaptation in *Escherichia coli*. *Proc. R. Soc. Lond. B.* **267**: 515–522.

Novitski, E., Grace, D. & Strommen, C. 1981. The entire compound autosomes of *Drosophila melanogaster*. *Genetics* **98**: 257–273.

Prout, T. 1971. The relation between fitness components and population prediction in *Drosophila*. I: the estimation of fitness components. *Genetics* **68**: 127–149.

Rainey, P.B. & Travisano, M. 1998. Adaptive radiation in a heterogeneous environment. *Nature* **394**: 69–72.

Reznick, D. & Travis, J. 1996. The empirical study of adaptation in natural populations. In: *Adaptation* (M. R. Rose & G. D. Lauder, eds), pp. 243–289. Academic Press, San Diego.

Rice, W.R. 1989. Analysing tables of statistical tests. *Evolution* **43**: 223–225.

Rice, W.R. 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nature* **381**: 232–234.

Rose, M.R. 1984. Laboratory evolution of postponed senescence in *Drosophila melanogaster*. *Evolution* **38**: 1004–1010.

Rose, M.R., Vu, L.N., Park, S.V. & Graves, J.L. Jr. 1992. Selection of stress resistance increases longevity in *Drosophila melanogaster*. *Exp. Geront.* **27**: 241–250.

Rose, M.R., Nusbaum, T.J. & Chippindale, A.K. 1996. Laboratory evolution: the experimental wonderland and the Cheshire Cat syndrome. In: *Adaptation* (M. R. Rose & G. V. Lauder, eds), pp. 221–241. Academic Press, San Diego.

Service, P.M., Hutchinson, E.W. & Rose, M.R. 1988. Multiple genetic mechanisms for the evolution of senescence. *Evolution* **42**: 708–716.

Service, P.M. & Rose, M.R. 1985. Genetic covariation among life history components: the effect of novel environments. *Evolution* **39**: 943–945.

Shiotsugu, J., Leroi, A.M., Yashiro, H., Rose, M.R. & Mueller, L.D. 1997. The symmetry of correlated selection responses in adaptive evolution: an experimental study using *Drosophila*. *Evolution* **163–172**: 163–172.

Teotónio, H. & Rose, M.R. 2000. Variation in the reversibility of evolution. *Nature* **408**: 463–466.

Teotónio, H. & Rose, M.R. 2001. Perspective: reverse evolution. *Evolution* **55**: 653–660.

Travisano, M. & Lenski, R.E. 1996. Long-term experimental evolution in *Escherichia coli*. IV. Targets of selection and the specificity of adaptation. *Genetics* **143**: 15–26.

Travisano, M., Mongold, J.A., Bennett, A.F. & Lenski, R.E. 1995. Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* **267**: 87–90.

Weber, K.E. 1996. Large genetic change at small fitness cost in large populations of *Drosophila melanogaster* selected for wind tunnel flight: Rethinking fitness surfaces. *Genetics* **144**: 205–213.

Wilton, A.N., Joseph, M.G. & Sved, J.A. 1989. Can chromosomal heterosis in *Drosophila* be explained by deleterious recessive genes? Negative results from a dichromosomal population test. *Genet. Res. Camb.* **53**: 129–140.

Wright, S. 1977. Evolution and the Genetics of Populations, **Vol. 3**. Experimental Results and Evolutionary Deductions. University of Chicago Press, Chicago.

Received: 31 October 2001; revised: 4 February 2002; accepted: 2 March 2002