Evolution, 50(2), 1996, pp. 795–806

EFFECTS ON FITNESS COMPONENTS OF P-ELEMENT INSERTS IN *DROSOPHILA MELANOGASTER*: ANALYSIS OF TRADE-OFFS

STEPHEN C. STEARNS AND MARCEL KAISER
Zoology Institute, Rheinsprung 9, CH-4051 Basel, Switzerland
E-mail: stearns@ubaclu.unibas.ch

Abstract.—We analyzed the trade-offs between fitness components detected in four experiments in which traits were manipulated by inserting small (control) and large (treatment) P-elements into the *Drosophila melanogaster* genome. Treatment effects and the interactions of treatment with temperature, experiment, and line were caused by the greater length and different positions of the treatment insert. In inbred flies, the treatment decreased early and total fecundity. Whether it increased the lifespan of mated females depended upon adult density. Analysis of line-by-treatment-by-temperature interactions revealed hidden trade-offs that would have been missed by other methods. They included a significant trade-off between lifespan and early fecundity. At 25°C high early fecundity was associated with decreased reproductive rates and increased mortality rates 10–15 days later and persisting throughout life, but not at 29.5°C. Correlations with Gompertz coefficients suggested that flies that were heavier at eclosion also aged more slowly and that flies that aged more slowly had higher fecundity late in life at 25°C. The results support the view that lifespan trades off with fecundity and that late fecundity trades off with rate of aging in fruitflies. Genetic engineering is an independent method for the analysis of trade-offs that complements selection experiments.

Key words.—Aging, *Drosophila,* genetic engineering, life-history evolution, lifespan, trade-offs.

Received October 19, 1994. Accepted May 16, 1995.

Two traits trade off when a change in one that increases fitness is linked to a change in the other that decreases fitness. Trade-offs have long played a central role in the analysis of life-history evolution (e.g. Williams 1966; Gadgil and Bossert 1970; Schaffer 1974), for such connections among traits force compromises to be made in the course of evolution. However, basic questions about trade-offs still provoke controversy (Bell 1984a,b; Reznick 1985; Reznick et al. 1986). How should trade-offs be measured? What causes them?

Two natural levels at which to analyze trade-offs—physiology and genetics—correspond to two causes of changes in traits. Physiological trade-offs result from the allocation of limited resources within individuals (Sibly and Calow 1986) and among individuals in families (Stearns 1989). Evolutionary trade-offs are measured among individuals within populations as genetic correlations whose expression is modulated by phenotypic plasticity (Lande 1979; Stearns et al. 1991). Both types of trade-offs play a role in the evolutionary theory of aging (Williams 1957; Rose and Charlesworth 1980; Rose 1991).

The study of trade-offs began, in the late sixties and early seventies, with physiological studies of energy allocation. These led by the midseventies to a sophisticated view of the strengths and limits of the exclusively physiological approach (e.g., Hirshfield and Tinkle 1974). Then, in the late seventies and early eighties, the emphasis shifted to measurements of genetic correlations between traits detected through correlated responses to selection (cf. Dingle and Hegmann 1982). The leading system for such work has been *Drosophila melanogaster.* A substantial tradition of selection experiments (Rose 1984; Luckinbill et al. 1984; Partridge and Fowler 1992, 1993; Hillesheim and Stearns 1992; Zwaan et al. 1995a,b) now suggests that lifespan in *Drosophila* trades off with components of fitness expressed early in life, but which early fitness component is involved has depended on the experiment (Table 1).

Just as work on trade-offs in the 1970s led to a recognition of the limits of the physiological approach, so has the study of correlated responses suggested limits to the exclusively genetical approach. Two lines of evidence indicate that the physiological and genetic approaches should be combined (Stearns et al. 1991). First, the correlated responses to selection found in similar experiments done by different groups have not always been consistent (Table 1), and genotype-by-environment interactions in the larval environment are thought to cause the inconsistency (Zwaan et al. 1995a,b). Second, Leroi et al. (1994) and Leroi et al. (1994) have analyzed Rose's (1984) lines in which a trade-off between early and late fecundity was originally found. After 10 more years of selection, the trade-off seemed to have disappeared. In fact, it was still there but had been obscured by genotype-by-environment interactions and adaptation of each population to unique aspects of its culture regime. Both types of evidence combine physiology and genetics, for genotype-by-environment interactions measure how the expression of genetic variation is modulated by physiological interactions with environmental variation during development.

If genotype-by-environment interactions are the key to understanding the proximate causes of trade-offs, we need new tools to analyze them, for quantitative genetics contains no information on the physiological mechanisms through which the effects of genes can be traced to the phenotype. Genetic engineering may be the tool needed to connect physiology and genetics, for it allows us to manipulate the expression of single genes with partially known physiological effects.

This paper is one in a series on an attempt to genetically engineer the components of fitness by inserting P-elements containing the heat-shock promoter and the gene for elongation factor EF-1α into the *Drosophila* genome (Shepherd et al. 1989; Stearns et al. 1993; Stearns and Kaiser 1993). Until 1993, these experiments, which were done at two temperatures, 25°C and 29.5°C, were interpreted on the assumption that the gene for elongation factor was more strongly expressed at the higher temperature (a design used to control

TABLE 1. Selection experiments on *Drosophila* in which tradeoffs associated with aging have been detected.

Trait selected	Luckinbill[1]	Rose[2] Age at reproduction	Partridge[3]	Hillesheim[4] Weight	Zwaan[5] Development time	Zwaan[6] Lifespan
A. Sample sizes						
Replicates/treatment	2	3	6	3	3	3
Longevity						
Flies/sex/replicate	30	>70[7]	80[7]	30[7]	25[7]	25[7]
Total	240[7]	791[8]	2880[7]	360[7]	300	300
Fecundity						
Females/replicate	30	70[7]	40 or 32[7]	30[7]	—	25[7]
Total	120	416[8]	216	352[8]	—	300
Age and size at eclosion						
Flies/sex/replicate			20 or 32[7]	30[7]	1500[7]	500[7]
Total			312[7]	719[8]	9000	3000
B. Tradeoff detected						
Early fecundity/lifespan	yes	yes	no	yes	[9]	yes
Early fecundity/late fecundity	yes	yes	no	yes	[9]	yes
Size at eclosion/lifespan	[9]	[9]	yes	yes	no	no
Developmental time/lifespan	[9]	[9]	yes	[9]	no	no
Larval competitive ability/lifespan	[9]	[9]	yes	[9]	[9]	[9]

[1] Luckinbill et al. (1984). [2] Rose (1984). [3] Partridge and Fowler (1992). [4] Hillesheim and Stearns (1992). [5] Zwaan et al. (1994a). [6] Zwaan et al. (1994b). [7] Started. [8] Finished. [9] Not investigated.

for position effects). We now know that the additional copy of the gene was not expressed at the temperatures used, neither in Shepherd et al.'s (1989) experiment (Shikama et al. 1994) nor in ours (Kaiser et al., unpubl.). That changed the interpretation of the results from a study of overexpression of a known gene to a study of proper controls in experiments on genetic engineering of phenotypes.

Before describing the experiments, we summarize the evidence that the gene was not expressed. Expression was tested using the S1 nuclease protection assay (Shikama et al. 1994). Line T, which was used by Shepherd et al. (1989) and which expresses the transgenic EF-1α message at 37°C but not at 29°C (Shikama et al. 1994), served as a positive control. In the lines E1 to E6 from the positions experiment the transgenic message could be detected in the lines E1, E2, E3, and E4 heat-shocked to 37°C, but not in the EF lines E5 and E6 heat-shocked to 37°C nor in any of the EF lines grown at 29°C. It was absent also in the control lines. The endogenous signal for the EF-1α mRNA was present in all samples. That there was no detectable transgenic EF-1α message at 29°C indicates that the transgene was not expressed at all or only at a minimal level in our experiments. Thus all treatment effects and the interactions of treatment with other factors (temperature, experiment, line) were caused not by overexpression of the gene but by the greater length of the treatment insert, which was 2 kb longer than the control insert, and by the different positions of the treatment and control inserts (Kaiser et al., unpubl.).

We did four experiments on the impact of P-elements on fitness components in *Drosophila.* In Expt. I, with Shepherd et al.'s (1989) inbred lines, we found strong treatment effects on lifespan and fecundity in mated females that we now interpret as caused by the greater length and different position of the treatment inserts. In Expt. II, we studied the importance of genetic backgrounds using heterozygous lines in which the control and treatment plasmids were inserted at different positions on Chromosome III in six backgrounds derived from isofemale lines. In Expt. III, we studied position effects in heterozygous lines in which the control and treatment plasmids were inserted at six positions on Chromosome III in one genetic background (Stearns and Kaiser 1993). In Expt. IV, we repeated Expt. I with a change in larval temperature to check whether certain effects were due to larval rearing temperature.

In Expts. I, II, and IV the treatment effects confound the size and position of the P-element insert. In Expt. III, the treatment effect measures the impact of the size of the insert by comparing a sample of six insert positions for each size and can be separated statistically from the positions effect.

This paper summarizes the results on trade-offs from all experiments. By "trade-off" we mean genetic correlations among individuals or lines within populations, and we have focused on trade-offs between fitness components expressed early and late in life.

METHODS

Experiments

An additional copy of the gene for elongation factor was inserted into the *Drosophila* genome; it was supposed to enhance protein synthesis at higher temperature. Large-insert lines consisted of *D. melanogaster* transformed with a P-element construct containing the ry+ marker plus the F1 copy of the gene for EF-1α flanked by initiation and termination sequences for heat-shock protein (11.8 kb). Control lines were similarly manipulated but lacked the gene for elongation factor; their insert was 16% shorter (9.9 kb) and at a different position. In Expts. II and III the large-insert and small-insert lines were obtained from a jumpstart-cross in which P-elements already inserted in the genome jumped to new insert

positions (Cooley et al. 1988; Robertson et al. 1988; Bellen et al. 1989). Balancers ensured that the only surviving flies carried the plasmid on Chromosome III. A brother-sister cross between flies with the correct markers yielded flies homozygous for the plasmid construct. For the backgrounds experiment, the test lines were obtained by outcrossing one inbred large-insert (EF 1, insert in centromeric region) and one inbred small-insert line (C2, insert at 64C) with six different inbred isofemale lines to yield six pairs of outbred large-insert and small-insert lines with the genetic backgrounds to be tested. The isofemale lines were derived from flies caught in the region of Basel, Switzerland, and had been maintained by full-sib mating for about 60 generations.

In the positions experiment the test flies were obtained by outcrossing six inbred large-insert and six inbred small-insert lines from the jumpstart-cross with one inbred lab stock (background 5) to yield six pairs of heterozygous large-insert and small-insert lines with the positions to be tested. All inserts were on the third chromosome, but the small inserts (C1—75C, C2—64C, C3—89B, C4—85D, C5—61C, C6 was lost before the position could be measured) were not in the same positions on the third chromosome as the large inserts (EF1—centromeric region, EF2—79F, EF3—96D, EF4—96D, EF5—99B, EF6—99B). The macroposition from banding patterns does not resolve insert positions in the DNA sequence, and analysis of effects of inserts on traits indicated that pairs of treatment lines with the same macropositions often had traits that differed significantly. Therefore we treated every insert as though it were a different position regardless of its macroposition.

Because genotype-by-environment interactions might be the key to understanding trade-offs, we did a detailed analysis of line-by-treatment-by-temperature effects on trade-offs in the backgrounds experiment. This can be detected as the $(abg)_{ijk}$ interaction effect in an ANOVA where i: 1, . . . , 2 indexes treatment a; j: 1, . . . , 2 indexes temperature b; k: 1, . . . , 6 denotes background g; and l: 1, . . . , n denotes either vials or individuals:

$$y = m + a_i + b_j + g_k + (ab)_{ij} + (ag)_{ik} + (bg)_{jk} + (abg)_{ijk} + e_{ijkl} \quad (1)$$

Figure 1 explains the role of line-by-treatment-by-temperature interaction effects in such experiments. This analysis revealed the existence of hidden trade-offs (Figs. 2–4) only detected in the interaction effects.

Table 2 compares the designs of the four experiments. The major differences were these: (1) replication—three controls and three treatments in Expt. I and Expt. IV, six controls and six treatments in Expt. II and Expt. III; (2) genetic backgrounds—different in every line in Expts. I and IV, six pairs of backgrounds matching treatment and control lines in Expt. II, one background in Expt. III; (3) insert positions—different among lines in Expts. I and IV with some on Chromosome II and some on Chromosome III; one treatment position and one control position on Chromosome III in Expt. II, six treatment positions and six control positions on Chromosome III in Expt. III; (4) genetic state of the flies—inbred in Expts. I and IV, heterozygous in Expts. II and III; (5) temperature at which the larvae and pupae were reared—25°C in Expt. I for lifespan and fecundity, 25°C and 29.5°C in Expt. I for age

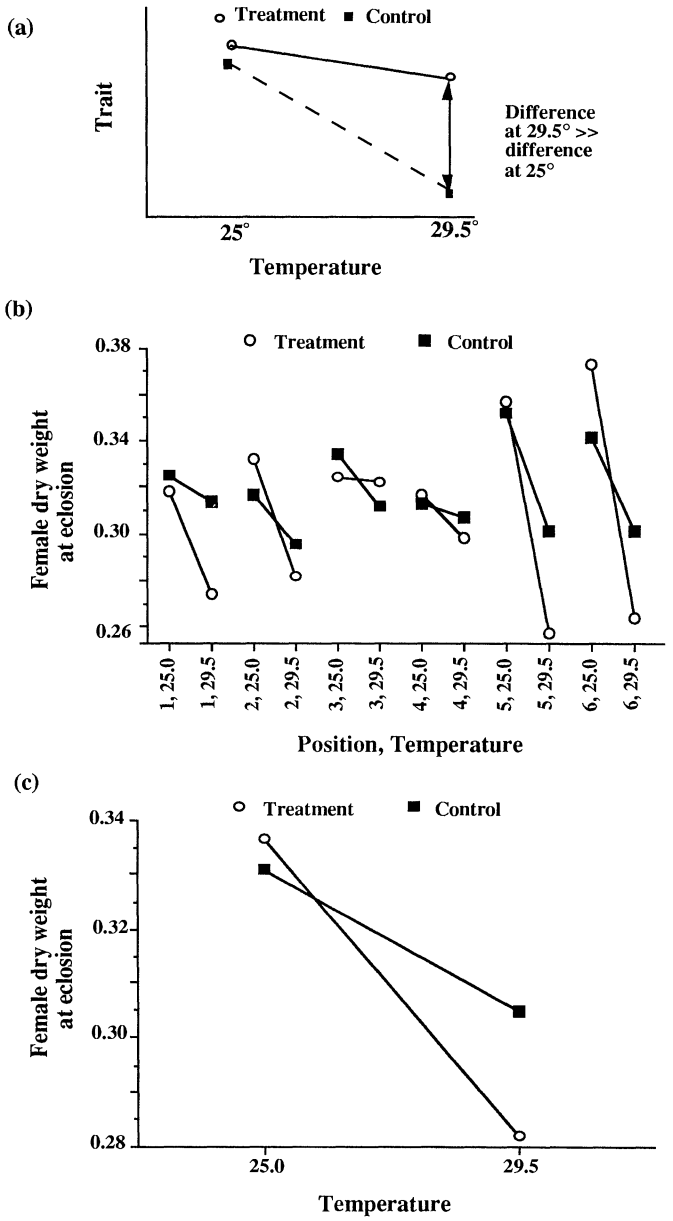

FIG. 1. The role of line-by-treatment-by-temperature interaction effects. (a) The treatment might change the mean value of a trait more than the control at the higher temperature, but most trait means change at higher temperature. One could test for the treatment-by-temperature effects by comparing the difference between treatment and control at 29.5°C with that at 25°C. (b) That simple test is inappropriate, for different lines react differently to the change in temperature. These produce significant three-way interaction effects ($P = 0.0202$, ANOVA, Type III SS), here depicted for position-by-treatment-by-temperature effects on the dry weight at eclosion of females in the positions experiment. In 5 of 6 lines, the reduction in dry weight at higher temperature is stronger in the treatment than in the control lines. (c) In the two-way treatment-by-temperature interaction ($P = 0.0002$), the average effect of treatment at higher temperature is clear: treatment flies were slightly heavier at eclosion at 25°C but much lighter at eclosion at 29.5°C. The larger insert interacts with temperature to reduce dry weight at eclosion in females.

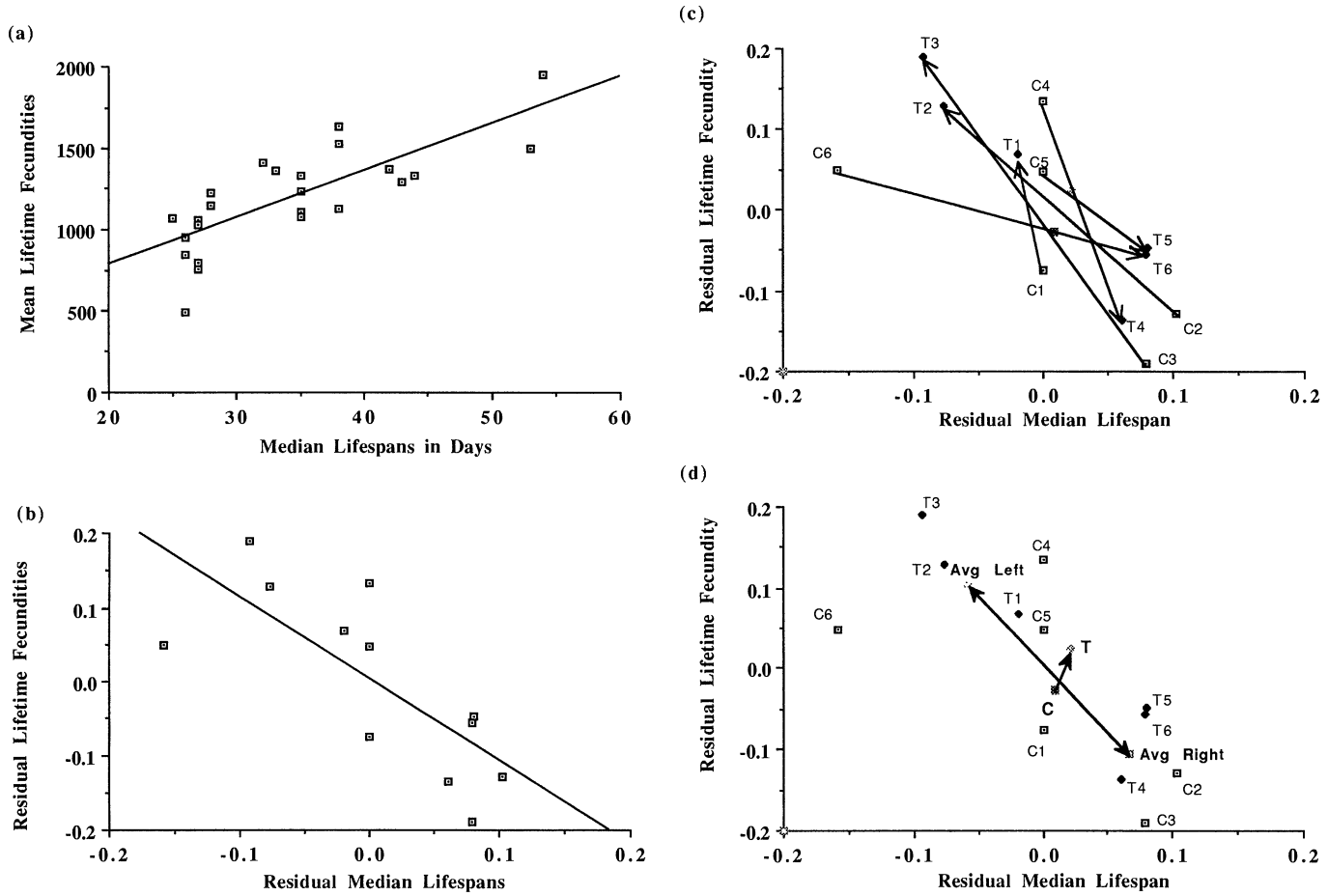

FIG. 2. The dissection of a trade-off. (a) The relationship between lifetime fecundity (numbers of eggs) and lifespan for all lines at both temperatures in Expt. II (backgrounds). (b) The relationship between residual lifetime fecundity and residual lifespan, where the residuals express the line-by-treatment-by-temperature interactions, i.e. the degree to which either fecundity or lifespan was increased by the larger insert at higher temperature in that line. (c) Detailed analysis of changes from control to treatment in the residuals for each of the six genetic backgrounds. (d) There are two ways of calculating the changes. The change from mean control to mean treatment (short line) misleads by obscuring the main effects. The change from left end to right end of the arrows (long line) shows that genetic changes were constrained to move along a trade-off.

at pupation, age at eclosion, and dry weight at eclosion, 25°C and 29.5°C in Expts. II, III, and IV for all traits; (6) sample sizes—small in Expt. I, larger in Expt. IV, quite large in Expts. II and III; (7) number of copies of transfected gene—two (one on each diploid chromosome) in Expts. I and IV, one in Expts. II and III.

To measure lifetime egg production, 30 vials per line were established with one female and two males per vial. A new laying surface with a drop of fresh yeast was presented daily. Dead males were removed and replaced by a young virgin male from the same line. The eggs laid by each female in 24 h were counted daily until the last female died.

Early and late fecundity were defined as follows. In all three experiments, early fecundity was the sum of eggs laid on days 4–14 after eclosion: 14–24 days after birth at 25°C and 12–22 days after birth at 29.5°C in Expts. I and IV; 13–23 days after birth at 25°C and 11–21 days after birth at 29.5°C in Expt. II; 14–24 days after birth at 25°C and 12–22 days after birth at 29.5°C in Expt. III. In Expt. I, late fecundity was the sum of eggs laid on days 21–31 after eclosion; this was 31–41 days after birth at 25°C and 29–39 days after birth at 29.5°C. In Expt. II and III late fecundity was defined as the number of eggs laid from the 15th to the fifth day before the death of the last female. In Expt. II this was 32–42 days after birth at 25°C and 23–33 days after birth at 29.5°C; in Expt. III it was 38–48 days after birth at 25°C and 31–41 days after birth at 29.5°C. In Expt. IV late fecundity was 32–42 days after birth at 25°C and 24–34 days after birth at 29.5°C.

To measure the lifespan of virgins, 10 vials per line and sex were established with 10 two-day-old flies and given one drop of fresh yeast. For mated females, 10 vials per line were established with 10 two-day-old virgin females and 15 two-day-old males. Three times a week the flies were transferred to new vials and the number of dead flies was recorded until the last fly died. In vials with mated females, males were replaced if there were fewer males than females in the vial.

For the analysis of trade-offs with Gompertz coefficients, measured as a property of each line for each treatment and temperature, we used mean fecundity values for all the fe-

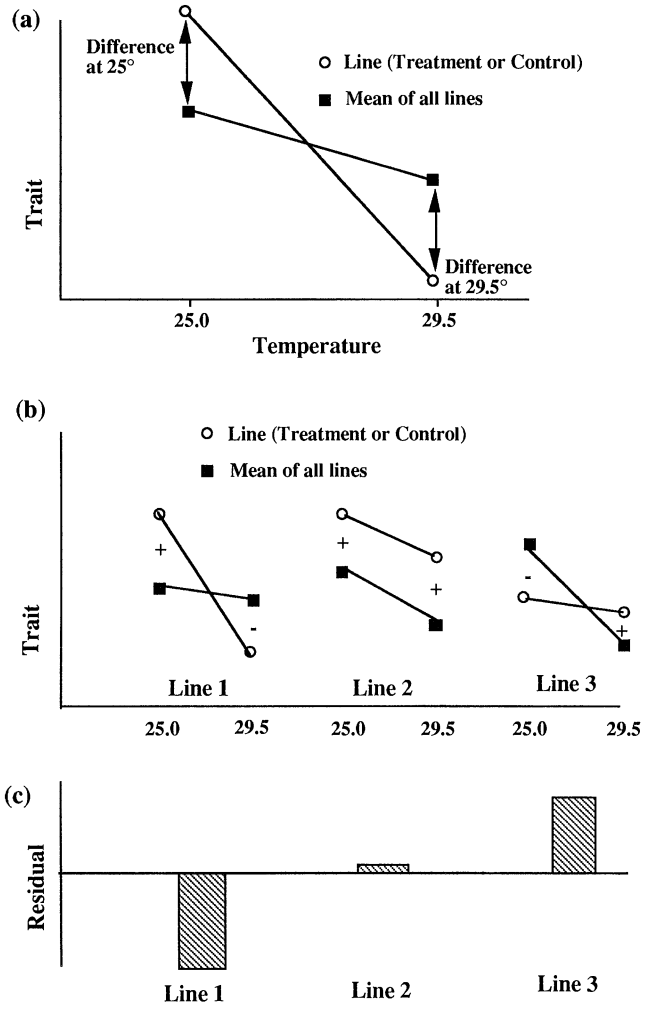

FIG. 3. The calculation of the residuals used in Figures 2 and 4. (a) First the deviation of each line from the overall mean was calculated at each temperature. (b) These deviations express line-by-temperature interaction effects. Line 1 was larger than the mean at 25°C and smaller at 29.5°C; the effect of increased temperature was a larger reduction in the value of the trait in Line 1 than was experienced on average. Line 2 was consistently larger than the mean at both temperatures, just slightly more so at 29.5°C. Line 3 was smaller than the mean at 25°C and larger at 29.5°C. (c) The residuals are calculated by subtracting the difference at 25°C from the difference at 29.5°C. If a residual is negative, increased temperature reduced the trait; if positive, increased temperature increased the trait. The individual residuals depict the line-by-treatment-by-temperature effects. The difference between the average treatment and average control residual depicts the treatment-by-temperature interaction effect that can be tested in the ANOVA.

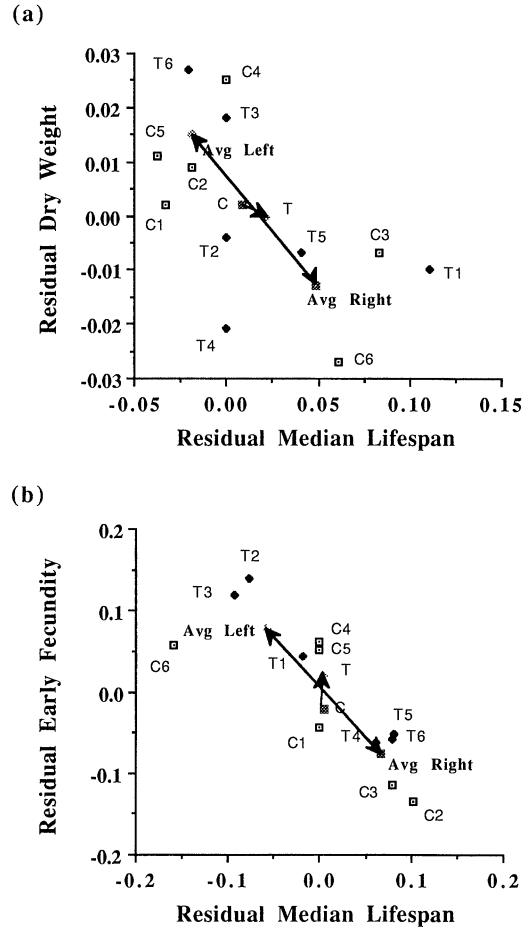

FIG. 4. Plots like Figure 2D. Recall that there are two methods of summarising the changes in residuals caused by the treatment The first—the change from mean control to mean treatment—is misleading, for it produces the short arrow pointing in a direction in which none of the changes occurred. The second—the change from the mean left end of an arrow to the mean right end—produces a long arrow that expresses the pattern in the data much better. (a) Treatment effects on the trade-off dry weight at eclosion versus lifespan in virgin males in Expt. II. (b) Treatment effects on the trade-off early fecundity versus lifespan in mated females in Expt. II. Treatment effects could have been caused either by the larger size of the P-element or by its different insert position or by both.

males in the line. For the analysis of trade-offs between fecundity and other fitness components, we only used data on surviving females that had both early and late fecundity, and there was no effect of mortality on the measurement of late fecundity.

Analysis of Trade-Offs between Group Means

For group means, we had two criteria for deciding whether the treatment, the temperature, or the treatment-by-temperature interaction had an impact on a trade-off: (1) the treatment or treatment-by-temperature interaction significantly increased one trait and significantly decreased the other; and (2) the changes from control to treatment, drawn as arrows on a two-trait plot, were confined to an envelope whose axis suggested a negative relation and whose end-points differed significantly on both the vertical and horizontal axes.

The second criterion requires explanation. Trade-offs detected by the comparison of individuals within populations have both a phenotypic and a genetic component. Expt. II permitted a novel analysis of trade-offs in which the phenotypic and genetic sources of variation could be separated. It is illustrated here for one relationship—mean total lifetime fecundity versus median lifespan. Figure 2 displays a series of plots of the mean total lifetime fecundities and median lifespans. Figure 2a displays all mean fecundities plotted against all median lifespans; the relationship is positive (r^2

TABLE 2. Experiments on the effects of elongation factor on fitness components in *Drosophila melanogaster.*

	Experiment				
Design element	Shepherd	Expt. I Repeat of Shepherd	Expt. II Backgrounds	Expt. III Positions	Expt. IV Control of Expt. I
Control lines	1	3[1]	6	6	3
Treatment lines	1	3	6	6	3
Backgrounds	2	6	6	1	6
Insert positions	2	6	1	6	6
Extra *EF-1α*	2 copies	2 copies	1 copy	1 copy	2 copies
Larvae and pupae[2]	25°	25°	25° and 29.5°C	25° and 29.5°C	25° and 29.5°C
Sample size per line:					
Longevity	300	40	100[3]	100[3]	100[3]
Fecundity	—	40	30	30	30
Eclosion	—	24	120[4]	120[4]	—
Sample size total:					
Longevity	300	1160	7200[4]	7200[4]	1200[4]
Fecundity	—	411	720[4]	720[4]	360[4]
Eclosion	—	1106[5]	2384[6]	2268[6]	—

[1] For lifespan and fecundity assays, there were 2 at 29.5°C. [2] For fecundity and lifespan; for other traits larvae and pupae were raised at both temperatures. [3] For each class: virgin males, virgin females, and mated females. [4] Number starting experiment. Number of larvae surviving to eclosion out of [5] 1200 [6] 2400.

= 0.60). At this level there is no trade-off. Figure 2b plots for each genetic background the treatment-by-temperature interaction effects (defined in Fig. 3); they define the difference in treatment effect at high and low temperature for each of the six genetic backgrounds. The relationship is negative ($r^2 = 0.56$) but not significant. However, each treatment point can be connected with its corresponding control by an arrow (Fig. 2c). In three genetic backgrounds (4, 5, and 6), the treatment-by-temperature interaction increased lifespan and decreased fecundity. In the other three backgrounds (1, 2, and 3) the treatment-by-temperature interaction reduced lifespan and increased fecundity.

The two ways to summarize these changes are depicted in Figure 2d. The short arrow pointing up and to the right describes the change between the average control-by-temperature interaction and the average treatment-by-temperature interaction. It fails to capture the consistency of the changes in all six backgrounds, which all lie within an envelope running from the upper left to the lower right. The long arrow running from the upper left to the lower right describes the change from the average left (T1, T2, T3, C4, C5, C6) to the average right end of an arrow (C1, C2, C3, T4, T5, T6) in Figure 2d, regardless of treatments and controls. It captures the important feature of the plot.

There are also two methods for judging the significance of the pattern. The backgrounds divide into two classes. Lines 1–3 reacted to the treatment by decreasing lifespan and increasing fecundity. Lines 4–6 did the opposite. The first method tests the differences between those two groups of lines for each trait separately by the significance of the background-by-treatment-by-temperature interaction in an ANOVA. It was not significant for either trait. The second method tests the difference between the x and the y coordinates of the two points (upper and lower coordinates on the y-axis, left and right coordinates on the x-axis) with a *t*-test. The left average lifespan is significantly less than the right one ($P = 0.0009$), and the upper average fecundity is significantly larger than the lower one ($P < 0.0001$).

The problem with the analysis of means becomes clear: there are two legitimate senses in which a difference between treatment and control can describe a change corresponding to a trade-off—reduced lifespan and increased fecundity, or increased lifespan and reduced fecundity. If both types of change are present in a data set, they counteract each other, and simply adding all changes together wipes out significant differences and brings one to the misleading conclusion that the treatment had no effect on the trade-off. It would be more accurate to say that all genetic effects of treatment were constrained to occur within the limits of a physiological trade-off, and that the direction of the change caused by the treatment depended on the genetic background. We therefore made plots like Figure 2d for every trade-off in Expt. II.

Analysis of Age-Specific Correlations

Lifespan is the product of survival at all ages up to death and incorporates the effects of trade-offs with effects on all age classes. Because lifespan is composite, it is not a good measure of aging. We therefore calculated birth rates and risks of death (age-specific hazards calculated with SAS procedure Lifetest, SAS Institute, 1985) in Expt. II (backgrounds) for all 12 lines (6 treatments and 6 controls) at both 25°C and 29.5°C for each five-day interval from eclosion to 45 days after eclosion at 25°C and to 35 days after eclosion at 29.5°C. We then calculated the correlation between the birth rate in each five-day interval with both the birth rate and the death rate in each subsequent five-day interval. We used $p<0.01$ as the significance level. It compromises between power and multiple comparisons.

Analysis of Correlations with Gompertz Coefficients

Aging is often defined as an increase in intrinsic mortality rates, and a decrease in intrinsic reproductive rates, with age.

TABLE 3. Tradeoffs detected by combinations of effects on single female traits: (A) magnitude and direction of the direct effects of treatment, the sign indicates the effect of having the larger P element; (B) magnitude and direction of the significant treatment × temperature interaction effects. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$. ( ) = effect not significant.

Trait	Expt. I Biocentre	Expt II. Backgrounds	Expt. III Positions	Expt. IV
A. Direct effects				
25°C				
Development time			−1.2%*	
Lifetime fecundity	+47.0%***		+24.9%*	+11.0%*
Lifespan, mated			+18.0%*	
29.5°C				
Early fecundity			+13.1%**	
Lifetime fecundity	−46.0%***		+14.0%*	−49.0%*
B. Treatment-by-temperature interaction effects				
Dry weight at eclosion	−7.3%***			
Lifespan, mated				
1 female per vial	+30.0%***			(−27.4%)[1]
10 females + 10 males				(+21.5%)[2]
Fecundity				
Early				−48.2%[3]
Lifetime	−77.0%***			(−60.4%)[4]

[1] $P < 0.0002$ tested against individual MS, $P = 0.06$ tested against temperature × line(treatment) MS. [2] $P < 0.0001$ for individual MS, $P = 0.3244$ for interaction MS. [3] $P = 0.0001$ for individual MS, $P = 0.0191$ for interaction MS. [4] $P = 0.0027$ for individual MS, $P = 0.0564$ for interaction MS.

A measure of the mortality component of aging is the Gompertz coefficient, which can be measured as the slope of the linear regression of the natural log of age-specific mortality rates ($ln\ \mu_x$) on age class x (Finch et al. 1990; Tatar et al. 1993; Hughes and Charlesworth 1994). We calculated the Gompertz coefficients for each line for each treatment and temperature in Expt. II, then analyzed the correlations between fitness components and Gompertz coefficients.

RESULTS

Criterion 1: Significant Changes in Both Traits in a Trade-Off

Table 3A lists the change in lifespan and fecundity associated with the direct treatment effect at each temperature where the effects of treatment on fecundity or lifespan were significant. In Expt. I, the treatment increased fecundity at 25°C and increased lifespan at 29.5°C, but there was no clear trade-off within either temperature. In Expt. II, no significant effects were detected. In Expt. III, the treatment increased lifespan and fecundity at 25°C; no trade-off was suggested. In Expt. IV, which differed from Expt. I only in that the larvae had been reared at both temperatures rather than uniformly at 25°C, the treatment increased fecundity at 25°C and decreased it at 29.5°C.

Another criterion for the impact of treatment on a trait is the treatment-by-temperature interaction effect, and Table 3B lists all traits involved in trade-offs with lifespan for which this effect was significant. In Expt. I, the treatment-by-temperature interaction increased the lifespan of mated females by 30%, reduced their weight at eclosion by 7%, and reduced their lifetime fecundity by 77%. In Expts. II and III there were no significant treatment-by-temperature interaction effects on mean lifespan or mortality rates, and therefore no trade-offs with lifespan by this criterion. In Expt. IV, the treatment-by-temperature interaction reduced lifetime fecundity by 60% and early fecundity by 48%. Effects on the lifespan of mated females depended on the environment in which lifespan was measured, and the significance of the effects depended on the statistical criterion used. When females were crowded, the treatment-by-temperature interaction increased lifespan 22%, but the effect was not significant when tested against the temperature-by-line interaction Mean Square. When females were not crowded, the treatment decreased lifespan 27%, and the effect was either significant by one criterion or close to significant by the other.

Thus the treatment-by-temperature interaction consistently reduced the fecundity of inbred females in both Expts. I and IV. It increased the lifespan of inbred mated females in Expt. I when the females were not crowded and in Expt. IV when they were crowded. When they were not crowded in Expt. IV, it decreased their lifespan.

Criterion 2: Interaction Effects Constrained to Lie within a Trade-Off

In Expt. II (backgrounds), we could use the method depicted in Figure 2 to see if the interaction effects of the treatment in six genetic backgrounds were constrained to lie within the envelope of a trade-off. We judged the significance of such effects by two criteria: (1) *t*-tests on the differences between the left and right and upper and lower ends of the lines connecting treatment and controls; and (2) the significance of the treatment-by-line interaction for each trait. Three trade-offs were significant by the first criterion (none were significant by the second): lifespan of mated females ($P = 0.007$) versus total fecundity ($P = 0.003$) (Fig. 2d), lifespan of virgin males ($P = 0.012$) versus dry weight at eclosion ($P = 0.006$) (Fig. 4a), and lifespan of mated females versus early fecundity ($P = 0.002$) (Fig. 4b). Thus in the backgrounds experiment, where no trade-offs associated with lifespan were detected by other methods, such trade-offs existed

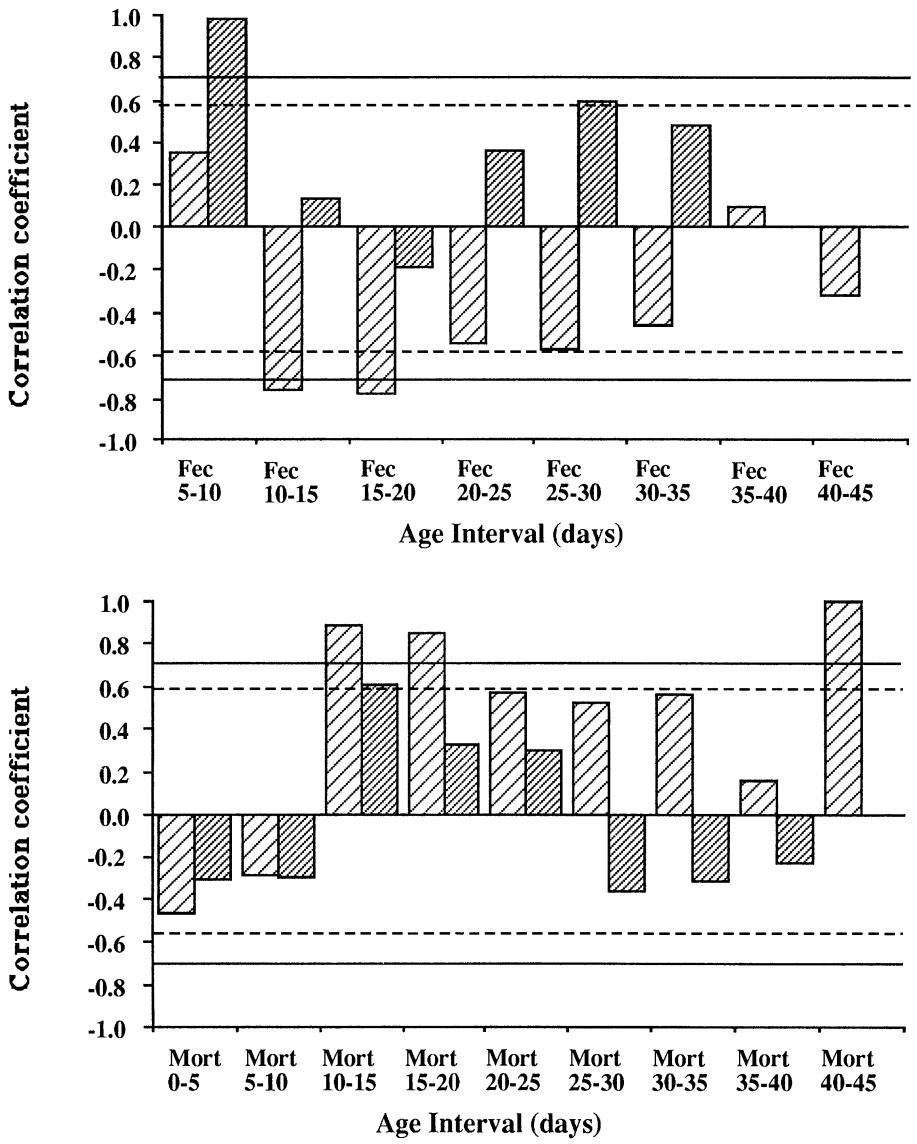

FIG. 5. (upper) Plots of correlation coefficients between birth rates in the first five days after eclosion and later birth rates defined on five-day intervals in Expt. II (backgrounds). (lower) Plots of correlation coefficients between birth rates in the first five days after eclosion and later death rates defined on five-day intervals in Expt. II (backgrounds). Wide hatching: 25°C; narrow hatching: 29.5°C. Solid line: $P = 0.01$; dashed line: $P = 0.05$.

and were significant by one test, but they were hidden in the summary statistics because the effects cancelled each other out.

Criterion 3: Correlations of Age-Specific Birth and Death Rates

The correlations of birth and death rates of five-day age-classes in Expt. II (backgrounds) gave mixed results (Fig. 5) at 29.5°C and clear results at 25°C. At 25°C, 10 to 15 days after birth rates were measured during the five days following eclosion. Flies that laid more eggs had lower birth rates and higher mortality rates for the rest of their lives, with the exception of one age interval for birth rates. Two of seven of the birth-birth correlations were significant, and three of eight of the birth-mortality correlations were significant. The pattern at 29.5°C was mixed and unconvincing.

The correlation analysis also contained information on treatment effects. Combining temperatures, there were 47 correlations between earlier and later birth rates. The treatment decreased 38 (81%) and increased 11 (19%). The correlations are not independent, but a sign test suggests ($P < 0.01$) that the treatment decreased more correlations than it increased, and thus it tended to increase the strength of a trade-off between earlier and later birth rates.

Combining temperatures, there were 64 correlations between earlier birth rates and later death rates. The treatment increased 48 (75%) and decreased 16 (25%). A sign test suggests ($P < 0.01$) that the treatment increased more correlations than it decreased and thus that it tended to increase

the strength of a trade-off between earlier birth rates and later death rates.

Criterion 4: Correlations of Fitness Components and Gompertz Coefficients

The Gompertz coefficients were positive for all lines in Expt. II and varied from 0.014 to 0.028, a twofold range. A two way (treatment-by-temperature) ANOVA using lines as replicates revealed no significant effects of treatment, temperature, or their interaction on the Gompertz coefficients. Correlations of Gompertz coefficients with all female fitness components were done; three were significant—weight at eclosion, early fecundity, and late fecundity. The significant simple correlations of Gompertz coefficients with early and late fecundity were checked with the partial correlation controlling for weight at eclosion.

First, the simple correlation of Gompertz coefficients with dry weight at eclosion was significant ($P = 0.05$, $r^2 = 0.17$) and negative (regression slope = -0.062) for both temperatures taken together. Flies that were heavier at eclosion aged more slowly (Fig. 6a).

Second, the correlation of Gompertz coefficients with late fecundity was significant ($P = 0.002$, $r^2 = 0.63$) and negative (regression slope = -0.00002) at 25°C but was not significant at 29.5°C (Fig. 6b). Flies that aged more slowly had higher fecundity late in life at 25°C but not at 29.5°C. This result was confirmed by the partial correlation analysis ($r = -0.8$ at 25°C, $r = -0.23$ at 29.5°C, critical value for $P < 0.01$ is $r = 0.735$).

Third, the correlation between Gompertz coefficients and early fecundity was significant ($P = 0.05$, $r^2 = 0.34$) and positive (regression slope = $+0.00001$) at 29.5°C but not at 25°C (Fig. 6c). Flies that had higher fecundity early in life also aged more rapidly at 29.5°C but not at 25°C. The significance of this result disappeared in the partial correlation analysis ($r = +0.556$ at 29.5°C, $r = -0.376$ at 25°C, critical value for $P < 0.05$ is $r = 0.602$). The result at 29.5°C was not significant.

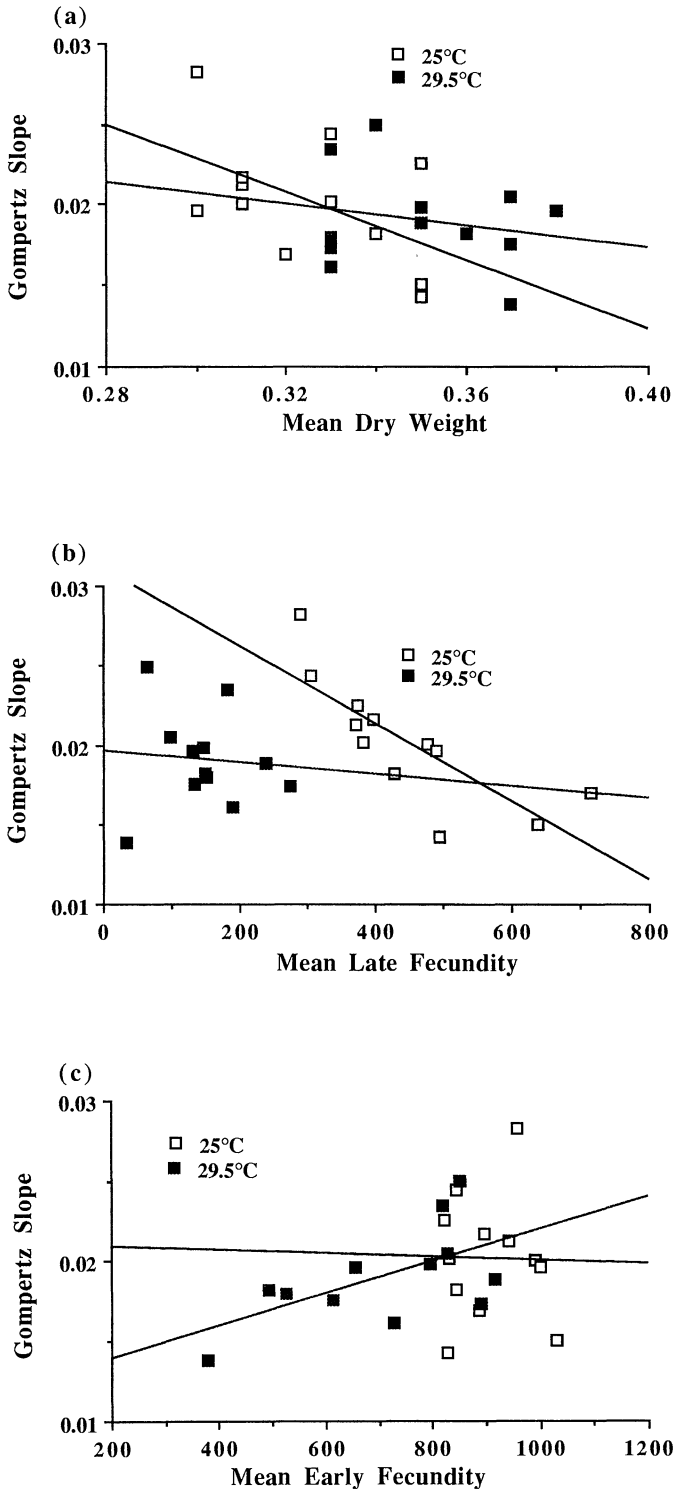

FIG. 6. The relation between Gompertz coefficients and fitness components calculated for each line in Expt. II. (a) Relation with dry weight at eclosion. (b) Relation with late fecundity. (c) Relation with early fecundity.

DISCUSSION

New Tools for Trade-Offs

Genetic engineering is a new tool for the analysis of trade-offs that delivers information independent of prior analyses of correlated responses in selection experiments. To be useful in life-history evolution, such manipulations must have large effects on components of fitness, and it would be helpful if the over- or underexpression of the gene could be traced through a clear causal pathway to the phenotype. In this case, that was not possible, for the gene was not expressed at either temperature. We have interpreted the treatment effects, and the interaction effects involving treatments, as effects caused by the greater disruption to downstream expression of other genes resulting from the greater length and different position of the treatment inserts.

Some evidence from selection experiments with *Drosophila* suggests that flies with low early fecundity live longer (Luckinbill et al. 1984; Rose 1984). Other evidence suggests that larger flies have higher late fecundity and live longer (Partridge and Fowler 1992). Others have found that lifespan trades off with total lifetime fecundity (Zwaan et al. 1995b).

Criterion 1a—direct effects of treatment on both traits involved in a trade-off—does not suggest a trade-off between

lifespan and lifetime fecundity. In Expts. I and IV the treatment increased fecundity at 25°C and decreased it at 29.5°C but had no significant effects on lifespan. No significant direct effects on fecundity or lifespan were detected in Expt. II. No trade-off was detected in Expt. III, where the treatment improved both lifespan and fecundity.

Criterion 1b—significant treatment-by-temperature effects on both traits involved in a trade-off—does suggest that lifespan sometimes trades off with early fecundity. When early and lifetime fecundity were reduced in inbred flies (Expts. I and IV), lifespan was sometimes, but not always, increased. There appears to be a density-by-treatment-by-temperature interaction effect on the lifespan of mated females. The disappearance in Expt. IV of the trade-off at low density noted in Expt. I may be the same phenomenon observed by Leroi et al. (1994a), for Expt. IV was done on the same lines used in Expt. I after three further years of selection for high fecundity on the 14th to 15th day of life. As suggested by Leroi et al. (1994a), the trade-off of fecundity with lifespan probably did exist but was hidden in the interaction effects, a point reinforced by the results of the next method.

Criterion 2—"hidden," counterbalancing effects of treatment on trade-off relations in the line-by-temperature interactions—did suggest clear trade-offs between lifespan and both early and lifetime fecundity and between lifespan and dry weight at eclosion of males. The first two trade-offs resemble those found by Rose (1984), Luckinbill et al. (1984), and Zwaan et al. (1995b). The last, "smaller flies live longer," resembles that found by Hillesheim and Stearns (1992).

Criterion 3—correlations in age-specific birth and death rates—found suggestive evidence for trade-offs at 25° but not at 29.5°C. At 25°C flies that had higher birth rates in the five days following eclosion had lower birth rates and higher death rates starting 10–15 days later and persisting throughout life; these effects were only significant in one third of the age classes. Such effects may be more easily detected in composite traits like lifespan that summarise survival ability over the whole life, as contrasted with 5-day death rates, which vary much more. The treatment effects on the correlation coefficients suggested that the treatment increased the strength of trade-offs between earlier and later fecundity and between earlier fecundity and later survival rates. This supports the existence of "hidden" trade-offs in Expt. II that were also suggested by Criterion 2.

Criterion 4—correlations of Gompertz slopes with fitness components—suggested the existence of a trade-off between weight at eclosion and Gompertz slopes at 29.5°C, where flies that were heavier at eclosion aged more slowly. Flies that aged more slowly also had better fecundity late in life at 25°C but not at 29.5°C. This suggests that larval investment in greater weight at eclosion, perhaps coupled with a decrease in fecundity early in life, results in flies that age more slowly and lay more eggs late in life. In contrast, Tatar et al. (1993) found in *Callosobruchus* that phenotypic manipulation of early reproductive effort resulted in a short-lived increase in age-specific mortality for at least the next 10 days, but not in a difference in Gompertz slopes.

The correlation analysis of age-specific fecundity and mortality rates (Fig. 5) and the analysis of Gompertz slopes (Fig. 6) were apparently inconsistent. The analysis of age-specific rates suggested trade-offs between early fecundity and mortality at 25°C but not at 29.5°C, whereas the analysis of Gompertz slopes suggested a significant trade-off between early fecundity and the rate of increase of mortality rates with age at 29.5°C but not at 25°C. The inconsistency can be explained by noting that different measures were used for age-specific fecundity in the two analysis, and that age-specific mortality rates are not the same thing as the rate of increase of mortality with age. Whether one detects a trade-off or not depends not just on the environment in which the organisms are measured, but also on how one chooses to define the traits to be measured. Gompertz slopes and age specific birth and death rates measure different things; both should be reported and analyzed.

Overall, these results suggest that lifespan trades off conditionally with early and lifetime fecundity in *D. melanogaster* and with age or size at eclosion. Because we did not find that every time one of these traits changed, the others changed as well, we conclude that the connections among them are not simple and are environment- and definition-dependent. We need a model of the organism at a deeper level that contains enough complexity to be empirically sufficient. The lack of agreement between our results and those of Partridge and Fowler (1992) illustrates this problem and has several explanations: their flies may have been different, their selection experiments may have elicited different trade-offs than did these genetic manipulations, or the environments in which they measured the trade-offs may have differed from ours.

The Causes of Trade-Offs: Genetics and Physiology

The discovery of trade-offs with these methods bears on a recent exchange in the literature over the proper measurement of trade-offs. Reznick (1985) and Reznick et al. (1986) claimed that a truly evolutionary trade-off between two traits can only be genetic, and that the best way to measure such a trade-off is through genetic correlations, preferably using the method of correlated response to selection. This stands in contrast to the views of Bell (1984a,b) and Partridge and her coworkers (e.g. Partridge and Farquhar 1981; Partridge et al. 1987), who share the opinion that important information on trade-offs can also be gathered from phenotypic manipulations.

The trade-offs documented here are associated with insertion of P-elements of different sizes and insert positions, their effects being measured in different genetic backgrounds and at different temperatures. That the relations between age-specific birth and death rates and between fitness components and Gompertz slopes were all strongly temperature-dependent (Figs. 5–6) reinforces the conclusion that physiology and genotype-by-environment interactions play a strong role in the expression of trade-offs.

The trade-offs are not associated with overexpression of the gene for elongation factor. The P-element insertions evidently disrupted gene expression in various ways that cannot be precisely measured with these data. That lack of precision actually reinforces, rather than reduces, the surprising nature of the result. Evidently one can disrupt the genome of *D. melanogaster* quasi at random with P-elements with different

lengths and insert positions (the inserts occur wherever P-elements can insert, not with respect to function) and through that manipulation elicit trade-offs otherwise detected as correlated responses to selection. This indicates that trade-offs are created by a physiology built from the effects of many loci. That physiological framework, which we see as establishing the envelope within which genetic changes take place in Figure 2d, can be perturbed by P-element inserts, as was done here, by changing the frequency of alleles segregating at variable loci, as has been done in selection experiments, or through phenotypic manipulations. The results of all three types of manipulations contain useful information.

The Influence of Inbreeding and Insert Dosage on Trade-Offs

In Expts. I and IV, where the flies were inbred and two copies of the insert were present, treatment-by-temperature interaction effects on lifespan and fecundity in mated females were large and significant. In Expts. II and III, where the flies were outbred and one copy of the insert was present, no significant treatment-by-temperature interaction effects on lifespan or fecundity could be detected. This suggests either that the effects on trade-offs of P-element inserts are more easily detected in inbred flies, or that such effects vary strongly with the dosage of the gene. If the differences in the two sets of experiments were primarily due to inbreeding, that would imply that outbreeding decreases the impact of trade-offs. If the differences were primarily due to gene dosage, that would imply that major effects could be elicited with much higher dosages. Both implications deserve further investigation.

Conclusions

Trade-offs are produced by interactions of genetic variation, lineage-specific physiology, and environmental conditions. They were detected in experiments in which analysis of mean response indicated that no trade-off was present. Selection experiments may not give conclusive evidence on the existence of trade-offs, or lack thereof, for if they were repeated under other environmental conditions, they might well elicit other responses. Combining genetic manipulations like these with the products of selection experiments could identify the physiological genetic mechanisms that cause trade-offs.

That the key factors in the evolution of lifespan and aging might only be detectable as two-, three-, or more-way interaction effects may seem to be an uncomfortably complex take-home message. It is, however, a message that is empirically well-supported, it reflects the basic nature of the phenotype as a product of genotype-by-environment interactions, and it is a view that we find foreshadowed in Wright (1969: selection in interaction systems), Mayr (1963: the unity of the genotype), Schmalhausen (1949: increasing complexity of correlation mechanisms), and Lewontin (1974: the problem of analysing into separate elements a number of causes that interact to produce a single result).

Acknowledgments

These experiments were made possible by the patient, precise, and reliable work of A. Mislin, B. Sykes, and H. Zingerli in the laboratory. We are very grateful for their support. Comments by B. Baur, D. Reznick, J. Travis, L. Partridge, and several anonymous reviewers improved the manuscript. Our research was supported by the Swiss Nationalfonds (Nr. 31-28511.90).

Literature Cited

Bell, G. 1984a. Measuring the cost of reproduction. I. The correlation structure of the life table of a planktonic rotifer. Evolution 38:300–313.

———. 1984b. Measuring the cost of reproduction. II. The correlation structure of the life tables of five freshwater invertebrates. Evolution 38:314–326.

Bellen, H. J.,C. J. O'Kane, C. Wilson, U. Grossniklaus, R. K. Pearson, and W. J. Gehring. 1989. P-element-mediated enhancer detection: A versatile method to study development in Drosophila. Genes Devel. 3:1288–1300.

Cooley, L., R. Kelley, and A. Sprading. 1988. Insertional mutagenesis of the *Drosophila* genome with single P-elements. Science 239:1121–1128.

Dingle, H., and J. Hegmann, eds. 1982. Evolution and genetics of life histories. Springer-Verlag, Berlin.

Finch, C. E., M. C. Pike, and M. Witten. 1990. Slow mortality rate accelerations during aging in some animals approximate that of humans. Science 249:902–905.

Gadgil, M., and W. Bossert. 1970. Life history consequences of natural selection. Am. Nat. 104:1–24.

Hillesheim, E., and S. C. Stearns. 1992. Correlated responses in life history traits to artificial selection for body weight in *Drosophila melanogaster.* Evolution 46:745–52.

Hirschfield, M. F., and D. W. Tinkle. 1974. Natural selection and the evolution of reproductive effort. Proc. Nat. Acad. Sci. USA 72:2227–2231.

Hughes, K. A., and B. Charlesworth. 1994. A genetic analysis of senescence in *Drosophila.* Nature 367:64–66.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution applied to brain: Body size allometry. Evolution 34:402–416.

Leroi, A. M., A. K. Chippendale, and M. R. Rose. 1994a. 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.

Leroi, A. M., W. R. Chen, and M. R. Rose. 1994b. Long-term laboratory evolution of a genetic life-history trade-off in *Drosophila melanogaster.* 2. Stability of genetic correlations. Evolution 48:1258–1268.

Lewontin, R. C. 1974. The analysis of variance and the analysis of causes. Am. J. Human Gen. 26:400–411.

Luckinbill, L., R. Arking, M. G. Clare, W. C. Cirocco, and S. A. Buck. 1984. Selection for delayed senescence in *Drosophila melanogaster.* Evolution 38:996–1003.

Mayr, E. 1963. Animal species and evolution. Harvard University Press, Cambridge, MA.

Partridge, L., and M. Farquhar. 1981. Sexual activity reduces lifespan of male fruitflies. Nature 294:580–582.

Partridge, L., and K. Fowler. 1992. Direct and correlated responses to selection for age at reproduction in *Drosophila melanogaster.* Evolution 46:76–91.

———. 1993. Responses and correlated responses to artificial selection on thorax length in *Drosophila melanogaster.* Evolution 47:213–226.

Partridge, L., A. Green, and K. Fowler. 1987. Effects of egg-production and of exposure to males on female survival in *Drosophila melanogaster.* J. Insect Physiol. 33:745–749.

Reznick, D. N. 1985. Cost of reproduction: An evaluation of the empirical evidence. Oikos 44:257–267.

Reznick, D. N., E. Perry, and J. Travis. 1986. Measuring the

cost of reproduction: A comment on papers by Bell. Evolution 40:1338–1344.

ROBERTSON, H. M., C. R. PRESTON, R. W. PHILLIS, D. M. JOHNSON-SCHLITZ, W. K. BENZ, AND W. R. ENGELS. 1988. A stable source of P-element transposase in *Drosophila melanogaster*. Genetics 118:461–470.

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

———. 1991. Evolutionary biology of aging. Oxford University Press, Oxford, U.K.

ROSE, M. R., AND B. CHARLESWORTH. 1980. A test of evolutionary theories of senescence. Nature 287:141–142.

SAS INSTITUTE. 1985. SAS user's guide: Statistics. SAS Institute, Cary, NC.

SCHAFFER, W. M. 1974. Selection for optimal life histories: The effects of age structure. Ecology 5:291–303.

SCHMALHAUSEN, I. I. 1949. Factors of evolution: The theory of stabilizing selection. Blakiston, Philadelphia, PA. Reprint, University of Chicago Press, Chicago.

SHEPHERD, J. C. W., U. WALLDORF, P. HUG, AND W. J. GEHRING. 1989. Fruitflies with additional expression of the elongation factor EF-1α live longer. Proc. Nat. Acad. Sci. USA 86:7520–7521.

SHIKAMA, N., R. ACKERMANN, AND C. BRACK. 1994. Protein synthesis, elongation factor EF-1α expression and longevity in *Drosophila melanogaster*. Proc. Nat. Acad. Sci. USA 91:4199–4203.

SIBLY, R. M., AND P. CALOW. 1986. Physiological ecology of animals. Blackwell Scientific, Oxford, U.K.

STEARNS, S. C. 1989. Trade-offs in life history evolution. Funct. Ecol. 3:259–268.

STEARNS, S. C, AND M. KAISER. 1993. The effects of enhanced expression of elongation factor EF-1α on lifespan in *Drosophila melanogaster* .IV. A summary of three experiments. Genetica 91:167–182.

STEARNS, S. C., G. DE JONG, AND R. NEWMAN. 1991. The effects of phenotypic plasticity on genetic correlations. Trends Ecol. Evol. 6:122–126.

STEARNS, S. C., M. KAISER, AND E. HILLESHEIM. 1993. Effect on fitness components of enhanced expression of elongation factor EF-1α in *Drosophila melanogaster* .I. The contrasting approaches of molecular and population biologists. Am. Nat. 142:961–993.

TATAR, M., J. R. CAREY, AND J. W. VAUPEL. 1993. Long-term cost of reproduction with and without accelerated senescence in *Callosobruchus maculatus*: Analysis of age-specific mortality. Evolution 47:1302–1312.

WILLIAMS, G. C. 1957. Pleiotropy, natural selection, and the evolution of senescence. Evolution 11:398–411.

———. 1966. Adaptation and natural selection. Princeton University Press, Princeton, NJ.

WRIGHT, S. 1969. Evolution and the genetics of populations. Vol. II. The theory of gene frequencies. University of Chicago Press, Chicago.

ZWAAN, B. J., R. BIJLSMA, AND R. F. HOEKSTRA. 1995a. Artificial selection for developmental time in *Drosophila melanogaster* in relation to the evolution of aging: Direct and correlated responses. Evolution 49:635–648.

———. 1995b. Direct selection on lifespan in *Drosophila melanogaster*. Evolution 49:649–659.

Corresponding Editor: D. Reznick