

# Population Size and Long-term Selection

*Kenneth Weber*

Biology Department  
University of Southern Maine  
96 Falmouth St.  
Portland, Maine 04103

- I. A BRIEF REVIEW OF EXPERIMENTAL RESULTS
  - A. Response Rates and Population Size
  - B. Fitness Interactions and Population Size
- II. NEW RESULTS
- III. POPULATION SIZE AND SELECTION EFFICIENCY  
LITERATURE CITED

## I. A BRIEF REVIEW OF EXPERIMENTAL RESULTS

### A. Response Rates and Population Size

The effect of population size on selection response is of interest to evolutionary biologists as well as plant and animal breeders. Theory and experiment agree that selection is more efficient in large populations than in small ones, as a mechanism for the sorting and concentration of small additive effects that contribute to narrow-sense heritability. This advantage can be detected even in short-term selection, and becomes substantial in the longer term. Although most of the relevant experiments have been done on animals, the implications for plant breeders are equally clear. This paper will review and evaluate the experimental evidence, before presenting some recent results (Part II) and a short practical discussion of population size effects (Part III).

Quite a few experiments have reported increased response in larger populations. The first of all such experiments was presented in three papers by Frankham et al. (1968a), Jones et al. (1968), and Frankham et al. (1968b). In this study, abdominal bristle number in *Drosophila melanogaster* was selected in replicated combinations of four selection intensities and three population sizes—10, 20, and 40 pairs of parents. By the 12th generation, there was a clear trend of superior response in larger lines (Frankham et al. 1968a, Fig. 1). (At this point the lines of lowest selection intensity were terminated.) By the end of the experiment at generation 50, the differences in response were strikingly in favor of larger populations (Jones et al. 1968). In every comparison of treatment means at the same selection intensity, larger populations responded more than smaller ones. These differences were becoming steadily larger as the experiment progressed, and none of the lines had plateaued at the end. There was no indication that the largest population sizes were nearing a limit where the magnitude of the population size effect would begin to decline. In fact, at each selection intensity the difference between the 40-pair lines and the 20-pair lines was about the same as the difference between the 20-pair lines and the 10-pair lines. Therefore, it seemed likely that even larger populations would have yielded still higher gains within the same amount of time.

In generation 16, new selection lines of the smallest size (10 pairs of parents) were sublined from the lines of the largest size (40 pairs), at the same selection intensities. This part of the experiment addressed the question of whether population size can safely be reduced after selection has had some time to raise the frequency of initially-rare favorable alleles (Robertson 1960; James 1971). The answer to that question was fairly clear in this case: response in all the small sublimes immediately lagged behind response in the large lines, and these differences in response continued to increase to the end of the experiment.

The second published experiment on the effect of population size was reported by Hanrahan et al. (1973), Eisen et al. (1973), and Eisen (1975), who selected on postweaning weight gain in mice using two selection intensities and five population sizes—1, 2, 4, 8, and 16 pairs of parents—replicated in all combinations. This study carefully maintained effective population sizes by selecting only within litters, equalizing the contribution of each family in both sexes. The two smallest population sizes were terminated at generation 14, and the other three at generations 21–27. Within each level of selection intensity, average realized heritability and cumulative selection response increased as functions of population size.

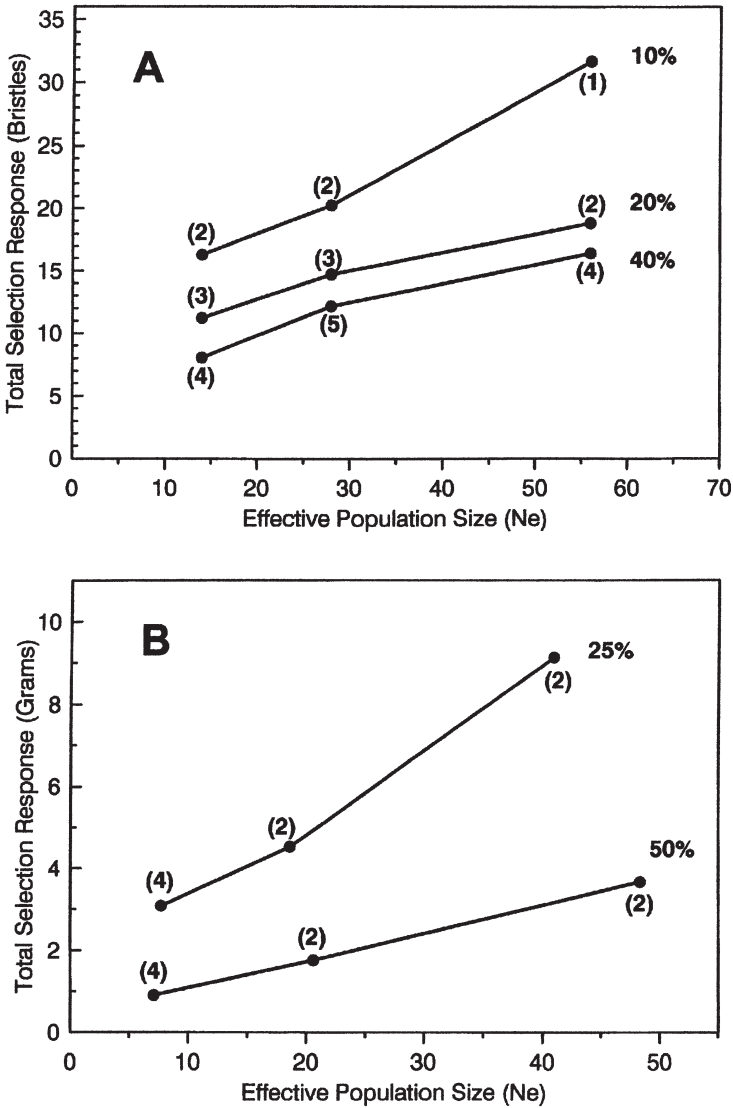


Fig. 11.1. Data extracted from two previous studies and summarized graphically. (A) Mean responses to selection for increased abdominal bristle number in *D. melanogaster*, after 50 generations. Data are for females. From Jones, et al., 1968. Estimated effective sizes following Frankham (1983). (B) Mean responses to selection for increased postweaning weight gain in mice, after 21–27 generations. Responses and calculated effective sizes from Eisen, 1975. In both graphs, each line connects three population sizes selected at the same intensity, given as a percent on right. Numbers of replicates per mean are in parentheses.

These two initial explorations of population size effects, in fruit fly bristles and mouse weight gain, are still in many ways the best. Each included a dual spectrum of population sizes and intensities, and continued long enough to reach firm conclusions about long-term effects. These studies showed that population size has a continuing, cumulative effect: as long as selection continued, the differences in total response between populations of different size continued to increase. The effect of population size was strong even at sizes above those of most selection experiments. However, both these studies ignored one of the most intriguing aspects of their own results, and so I have taken the liberty of distilling the endpoint data from both studies graphically (Fig. 11.1, A and B), to emphasize the following point. The trend of the data in both studies strongly suggests that even larger sizes would have yielded even larger responses, at all selection intensities, within the same number of generations.

In a third, related experiment, Ruano et al. (1975) selected on egg-laying rate in *Tribolium castaneum*, always scoring the same number of females (200) per line, but selecting the top 5%, 10%, 20%, 33%, 50%, and 100% (with two replicates per line) so that intensities were inversely related to population sizes. Higher-intensity lines responded more rapidly for the first 10 generations, but then more slowly. Lines of lower selection intensity (and therefore larger population size) continued to respond at slower rates, but with less of a decline in rate, so that eventually they began to catch up. When the experiment was terminated at generation 32, all lines were still responding, but total response was converging, and only the lines at lowest intensity and largest size were still responding linearly, at an undiminished rate. The design of this experiment doesn't permit the effect of size to be isolated from the effect of intensity, but the overall picture is one of early response dictated by selection intensity, slowly moving toward cumulative response dictated by population size.

Almost all studies have shown an advantage of larger populations that was already apparent in the short term—defined here as shorter than 20 generations (Eisen 1980). For example, in the first two studies mentioned above, the trend in favor of larger size was documented in early reports at generation 12 in Frankham et al. (1968) and at generation 14 in Hanrahan et al. (1973). A number of additional studies have focused only on the short-term effects of population size in selection, and have reinforced these observations.

Barker and Karlsson (1974) studied the effects of intense disruptive selection as a possible prelude to speciation, using sternopleural bristle

number in *D. melanogaster*. This was again an investigation of both selection intensity and population size, in four treatment combinations using selection pressures of 2.5% and 50%, with population sizes of 8 pairs and 40 pairs of selected parents. Selection continued for 17 generations. In the combination at high intensity and large size (which was unreplicated), a significant response was seen in terms of the divergence between the extreme members of the population, and there was also a significant tendency to assortative mating. The other combinations were ineffective.

Silvela et al. (1989) selected for percent oil in maize kernels, at selection pressures of 5% and 17%, with population sizes of 6, 10, and 50 individuals. Small, medium, and large populations had 5, 4, and 2 replicates, respectively, at both selection pressures. Although the experiment only lasted for 10 generations, the results show the advantage of larger size in all comparisons. The final generation trait means are not quite in order of size, but the means of the last 4 (or last 3, or last 2) generations are completely in order of population size across treatments (Silvela et al. 1989, Table 3). Realized heritabilities and regressions of response on generation number were also completely in order of population size. In this experiment the effects of population size were not yet large but were clearly emerging after 10 generations. By contrast, Blow and Briggs (1973) selected for 10 generations for body weight in quail, at a selection intensity of 25% with population sizes of 10, 20, and 30 pairs of parents, and found no effect of population size. There were two replicates per treatment.

Considering all studies of this type, it appears that experiments of short duration will not always detect a clear effect of population size unless there are many replicates and large differences in population size. Nevertheless, several well-replicated selection experiments have detected an advantage in larger populations after surprisingly few generations. Vasquez and Bohren (1982) reported a small, significant effect of population size in an experiment lasting only three generations, selecting for body weight in chickens at three effective sizes ( $N_e$ 's of approximately 4, 7, and 15), at a single selection intensity, with four replicates per treatment. Campo and Turrado (1997) selected on *Tribolium* for an index of adult and pupal weight, at five population sizes (1, 2, 4, 8, and 16 pairs of parents) and four selection intensities, with three replicates per combination, for only four generations. They, too, found a significant effect of population size overall.

These short-term effects of size, among small populations, raise the issue of what endpoints may exist as population size and generation

number both increase. This was the challenge that led to the development of several devices for automated mass selection on *D. melanogaster*. One of these, the “dorsoventrometer” (Weber 1988a), measured the wing-tip height of standing, cold-torpid flies that were homozygous for the allele *raised* (*rsd*), which causes vertical wings. This system permitted rapid selection on large numbers of flies for a trait that was partly morphological, partly behavioral, and perhaps partly physiological. About 20% of the response was due to increases in body size. Populations of three approximate sizes were tested—40, 200, and 1000 selected parents—at 20% selection. Selection was continued for 55 generations. Cumulative response was in order of population size for all replicates at all three sizes (Weber 1990, Figs. 11.2 and 11.3). Treatment means of cumulative response were a nearly linear function of the log of effective population size (Fig. 11.2A).

The second mass-selection system, the “inebriometer” (Weber 1988b), measured resistance to ethanol vapor in flies, in terms of the minutes of exposure to a controlled dose required to make flies fall through a series of baffles into the advancing tubes of a fraction collector. The population sizes were 160 and 1600 selected parents, the selected fraction was approximately 20%, and selection continued for 65 generations (Weber and Diggins 1990). Again, cumulative response was significantly higher in the larger lines (Fig. 11.2B).

In Weber (1990) and in Weber and Diggins (1990), “small” populations that were already large by previous standards were compared with populations that were five or ten times larger. Both of these experiments demonstrated that selection was still more effective in these extremely large populations. However, when two such large sizes are compared, it takes more time for the advantage of larger size to emerge. The larger the populations are, the longer the time required. Yet the differences still emerged well before the end of the experiments. In fact, the faster response of the larger populations was clear by generation 30 for wing-tip height (Weber 1990) and by generation 20 for ethanol resistance (Weber and Diggins 1990). Moreover, the total times involved—55 and 65 generations, respectively—were not extremely long. They were too short for the treatment differences to be attributed solely to the greater accumulation of new mutations in large populations. The lesson seems to be that even very large increments in population size confer an advantage in concentrating small effects that are available in standing genetic variance. Although these experiments enter new size ranges, they raise the same question as before: At what population sizes do larger numbers cease to make a difference in the extraction of genetic variation? Populations of increasing size become deterministic for smaller allelic

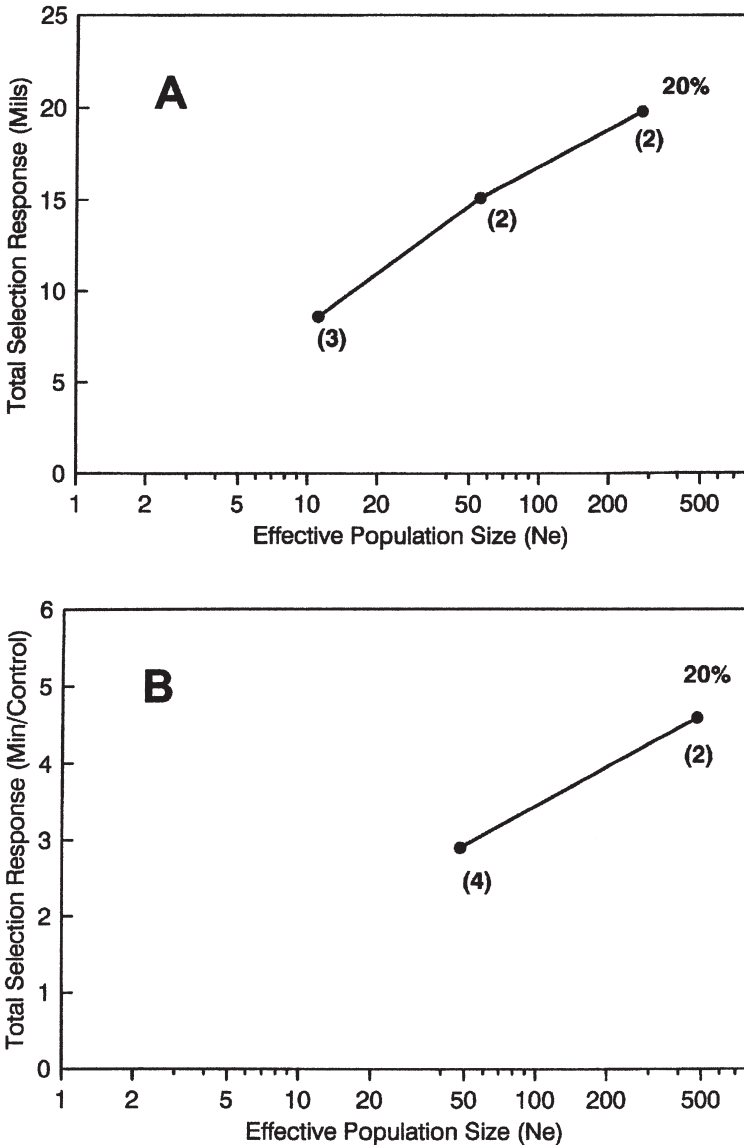
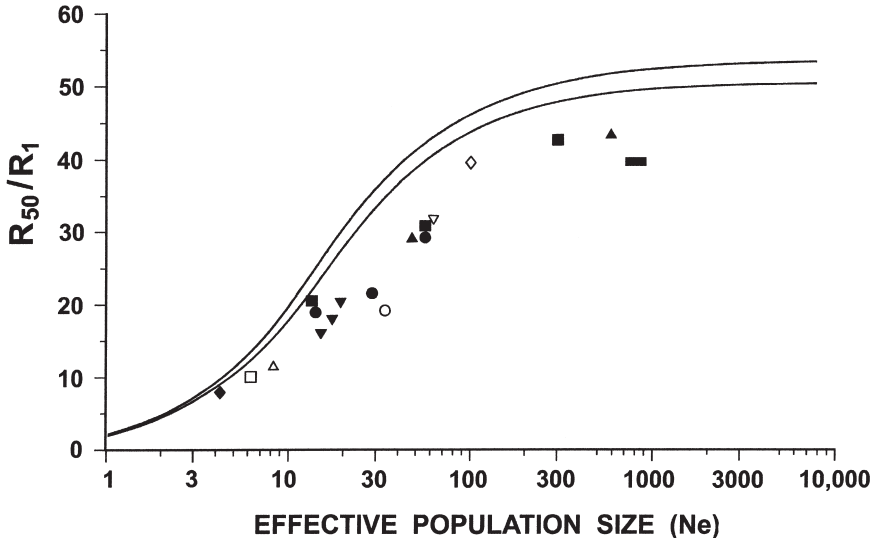


Fig. 11.2. (A) Mean responses to selection for increased wing-tip height in *D. melanogaster* *rsd* homozygotes, after 55 generations. 1 mil = 0.001 in or 0.0254 mm. From Weber (1990). (B) Mean responses to selection for ethanol resistance in *D. melanogaster*, in minutes/control minute, after 65 generations. From Weber and Diggins (1990). Numbers of replicates are in parentheses and percent selected is to the right. Estimated effective sizes in both experiments following Weber and Diggins (1990). Response = final selected phenotype – control phenotype.



**Fig. 11.3.** A summary of 11 selection experiments showing total response in generation 50, normalized by response in generation one ( $R_{50}/R_1$ ). The curves show predicted  $R_{50}/R_1$ , with and without new mutation at a rate of 0.001  $\times$  initial additive genetic variance per generation. Data are from Dudley (1977) on corn protein and oil content ( $\circ$ ), Enfield (1980) on *Tribolium* pupal weight ( $\diamond$ ), Jones, Frankham and Barker (1968) on fruitfly abdominal bristles ( $\bullet$ ), Mather and Harrison (1949) on fruitfly abdominal bristles ( $\Delta$ ), Rasmuson (1955) on fruitfly sternopleural bristles ( $\blacklozenge$ ), Reeve and Robertson (1953) on fruitfly wing length ( $\square$ ), Roberts (1966a,b) on mouse body weight ( $\blacktriangledown$ ), Weber (1990) on fruitfly wing-tip height ( $\blacksquare$ ), Weber and Diggins (1990) on fruitfly ethanol resistance ( $\blacktriangle$ ), Weber (1996) on fruitfly wind tunnel flight ( $\blacksquare$ ), and Yoo (1980) on fruitfly abdominal bristles ( $\nabla$ ). See Weber and Diggins (1990) for methods and for tabulation of all values of response and effective sizes, except for the datapoint from Weber (1996), which was derived later by the same methods.

differences, but these differences must become exhausted in a finite gene pool. It appears that experiments large enough to find this limit would run into timescales where mutation becomes important, and would require parallel selection on initially isogenic lines to quantify that component. The comparison would be quite laborious even with automated systems.

Considering these facts, it seemed that the best way forward in the investigation of large populations was to abandon further comparisons between selected populations and to concentrate all efforts instead on selection at a single size—the largest size possible—for the longest possible time. It has now been 15 years since the beginning of the wind tunnel selection experiment (Weber 1988c). In this experiment two replicate



populations, of approximately 2000 parents each, are selected from the top 2%–6% of flies in a 1.5 m long compartmented wind tunnel, in which they fly upwind towards a light. Mean wind velocity is constant in each chamber but falls evenly along the length of the tunnel via uniform outlets from each chamber. Flies arrive at a final distribution in the wind tunnel reflecting their individual abilities to contend with wind and turbulence. The whole sample is then anesthetized by a sudden transition from air to CO<sub>2</sub>. The phenotype of wind tunnel performance is expressed as the mean apparent flying speed, based on wind tunnel air velocities calculated for each compartment.

Wind tunnel performance responded rapidly to the intense selection. Over the first 100 generations (Weber 1996), the phenotype increased from about 2 cm/sec to 170 cm/sec. By this time, response had slowed, without reaching a plateau. The great differentiation between the phenotypic distributions of the selected lines and the unselected controls (Fig. 11.2) suggested a highly polygenic basis for the trait (Weber 1996). This was supported by the saltational, nearly vertical response to selection on hybrid lines (Fig. 11.5), indicating that selection had created linkages between many contributory alleles.

Although the wind tunnel experiment does not include a comparison of different population sizes, the effect of population size in this experiment and in others can be compared by converting all responses to a common scale. Fig. 11.3 shows results from eleven different selection experiments on ten traits in four organisms, at 17 population sizes. The common scale of response is the ratio of response in generation 50 to response after one generation ( $R_{50}/R_1$ ). Factors such as selection intensity and heritability cancel out. The figure shows the longer continuation of response in larger populations. The two sigmoid lines show  $R_{50}/R_1$ , as calculated from the model of Robertson (1960) and Hill (1982), using the version presented in Weber and Diggins (1990) for the prediction of response at finite generation times. The upper line shows response including the cumulative effect of new mutations. The lower line omits mutation. Under these assumptions, most of the increased response in larger populations at generation 50 comes from better utilization of initial genetic variation.

Traits like bristle number in flies (Jones et al. 1968) and body weight in mice (Eisen et al. 1973) are measured on one individual at a time. By contrast, the traits of wing-tip height, ethanol resistance and wind tunnel performance are mass measurements, subject to increased measurement error. Automation improves, and accuracy declines. Lower accuracy results in lower realized heritability and may also increase the advantage of larger sizes. As Figs. 11.1 and 11.2 show, larger populations

are better with either type of trait, and the advantage appears to extrapolate to population sizes that are still out of experimental range, whether for individually measured traits (Fig. 11.1) or for mass-measured traits (Fig. 11.2). Fig. 11.3 suggests that a wide variety of quantitative traits in many species may be much alike in their dependence on population size.

## **B. Fitness Interactions and Population Size**

To this point, this review has focused solely on heritability and selection response. The second major concern in any selection experiment is the effect of selection on fitness. Fitness is often ignored in both theoretical and experimental studies of selection response, but as a matter of common experience, it is almost as likely that fitness will decline, as it is that the selected trait will respond (Falconer and Mackay 1996). Even in conventionally large selection experiments (Enfield 1977, 1980; Yoo 1980), where inbreeding was not as obvious a problem, severe fitness problems have occurred. Sooner or later in any selection line, there are usually reductions in vitality and reproductive capacity, often associated with unpredictable and peculiar genetic syndromes. Moreover, under selection an opposition usually develops between fitness and the selected trait. This can arise from various causes and often produces a steep, immediate decline in the phenotype when selection is relaxed.

Fitness effects have been studied in several experiments on selection and population size. Frankham et al. (1968b) compared the tendency to regress of selection lines of different sizes, at five-generation intervals throughout their 50-generation experiment on abdominal bristles. In the later generations, several lines showed large declines in the selected trait when relaxed. Upon analysis, these were mainly lines with major recessive lethals affecting the trait. The most extreme case of this was the unreplicated line representing the combination of largest size and highest intensity. Eisen et al. (1973) noted that their two smallest population sizes suffered loss of fertility, while their largest populations suffered general losses of reproductive fitness as measured by size, weight, and survival of litters. Reasoning partly from comparisons with control lines, they attributed the fitness loss in the small lines to inbreeding, and the fitness loss in the largest lines to genetic interactions between fitness and the selected trait. The essential finding of both Frankham et al. (1968b) and Eisen et al. (1973) was that the largest lines showed the most total response to selection, but also the greatest fitness problems.

In the automated selection experiments on wing-tip height (Weber 1990) and ethanol vapor resistance (Weber and Diggins 1990), no measurements of any kind were made of fitness traits or interactions. This

is regrettable, because serious fitness problems did become evident. In the first experiment, the smallest lines became difficult to maintain and perished soon after the experiment terminated. In the second experiment, one of the smaller lines far exceeded the other lines of its size in its response, but this same line also developed extreme viability problems. However, none of the largest lines in either of these experiments showed obvious fitness problems, despite their higher response. These unquantified observations on massive populations contradict the idea that the larger responses of larger populations lead to correspondingly large fitness problems.

In the large wind tunnel selection lines (Weber 1996), fitness has been one focus of the investigation, and tests of two kinds have failed to find evidence of fitness effects. The first type of test was a direct assay of competitive fitness, in which either control or selected females were allowed to compete against white-eyed females from the same tester stock. This is a standard test of fitness that encompasses fecundity, egg-to-adult viability, and larval competitive ability. The test was performed at generation 50 and again at generation 85, and failed both times to show a significantly higher percentage of red-eyed flies emerging in control lines compared to selected lines.

The second type of test looked for a decline in phenotype after relaxation of selection. Lines relaxed in generation 65 for 30 generations, and in generation 85 for 10 generations, did not regress detectably. A lack of regression has been seen occasionally in lines relaxed after extreme selection. For example, in Yoo (1980), out of six replicate populations selected for 86–89 generations for abdominal bristle number, five replicates regressed rapidly after relaxation, but one remained stable, declining only slightly during 35 generations of observation. There are also undoubtedly differences among traits in how easily large changes can be tolerated. For example, in the Illinois long-term selection experiment, as the percent oil nears zero in the line selected for low oil, extreme fitness problems have developed (Dudley and Lambert 1992). Thus, there may be some absolute minimum oil content in the maize kernel. In the wind tunnel lines, the underlying traits that are being increased by selection are probably all fairly compatible with fitness, and unlikely to develop a direct phenotypic antagonism to it. Nevertheless, fitness problems have arisen in other experiments during selection on many seemingly peripheral traits. In the wind tunnel selection lines, the absence of either regression or competitive fitness loss, despite the extreme selection pressure and the long period of response, is rather unusual. It suggests some unusual cause, and the most unusual aspect of these populations is their large size.

## II. NEW RESULTS

Some more-recent results in the ongoing wind tunnel experiment are presented here, as a preview of coming publications. Since generation 100, there has been substantial additional response to selection. Now, as Fig. 11.4 shows, the experiment has finally arrived at a plateau stage where response is very slow. Various aspects of the phenotype appear to show adaptation for wind tunnel performance. One of the most important of these is positive phototaxis. Phototaxis is strongly developed in these flies (Fig. 11.5), to an extent that is obvious even during routine handling. Another aspect of the selected phenotype is larger body size (Table 11.1). These changes add to the evidence for the involvement of many genes in wind tunnel performance.

The earlier results, suggesting a lack of fitness interactions, have also been completely confirmed. New relaxed lines were started in generation 232 and were followed for the next 68 generations. They appear to

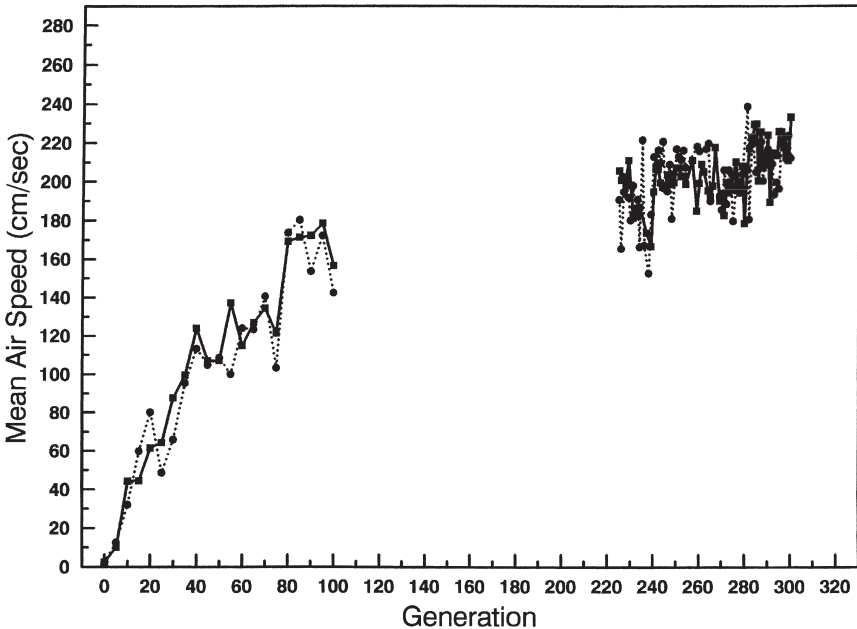


Fig. 11.4. Response to selection for wind tunnel performance in both selected lines. Solid line, population AA1; dotted line, population AA2. First 100 generations as published in Weber (1996). Selection without measurements from generation 100 to 220. Recent generations show that response is leveling off.

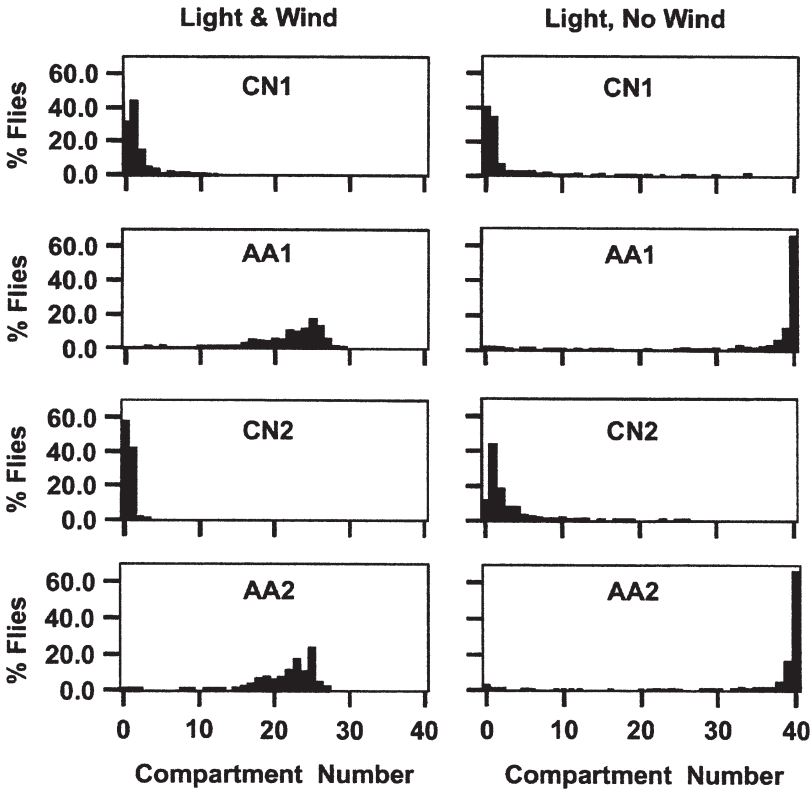


Fig. 11.5. Distributions of flies in the wind tunnel after tests of phototaxis. The combination of light plus wind is the normal condition for a run. Selected populations are AA1 and AA2. Control populations are CN1 and CN2. Tests are from generation 198.

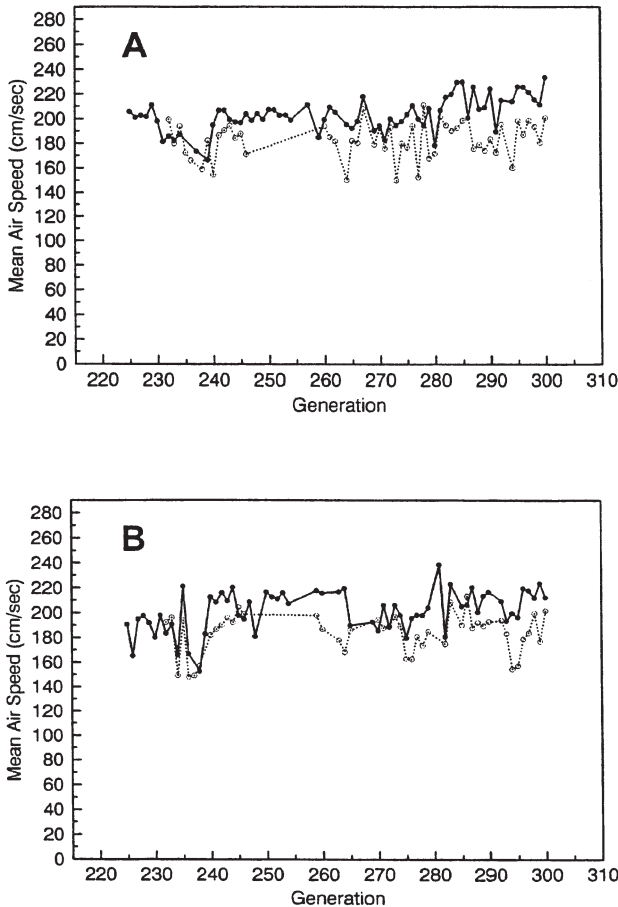
**Table 11.1.** Body weights in control (CN) and selected (AA) wind tunnel lines.<sup>z</sup>

Control (mg ± SD)	Selected (mg ± SD)	Increase (%)
CN1 females 1.24 ± 0.13	AA1 females 1.47 ± 0.34 ***	+19
CN1 males 0.76 ± 0.06	AA1 males 1.01 ± 0.07 ***	+33
CN2 females 1.37 ± 0.11	AA2 females 1.50 ± 0.17 ***	+ 9
CN2 males 0.80 ± 0.10	AA2 males 1.01 ± 0.09 ***	+26

<sup>z</sup>Sample sizes = 100 flies. \*\*\* P < 10<sup>-6</sup>.

have regressed very little, if at all (Fig. 11.6). Moreover, in a combined assay of egg-to-adult survival and larval competitive ability (Table 11.2), at generation 330, the selected lines still show no deficiency compared to controls.

It is interesting to compare these results with several previous experiments, in which measurements of fitness interactions were carried out during direct selection on phototaxis or body size in fruit flies. These experiments reported strong fitness interactions for both these traits, so that the present results would not have been predicted. For example,



**Fig. 11.6.** A comparison of selected populations (solid lines) and relaxed populations (dotted lines) from generation 232 to 300. A. AA1 and its relaxed population. B. AA2 and its relaxed population.

**Table 11.2.** Competitive fitness tests of control (CN) and selected (AA) wind tunnel lines. Means and standard errors of arcsine-transformed proportions of wild-type flies emerging from competition with a white-eyed tester stock. All vials were set up with 150 eggs of white-eyed line and 50 eggs of tested line. N=30 vials per line.

Line	Mean $\pm$ SE
AA1	0.277 $\pm$ 0.014
CN1	0.276 $\pm$ 0.012
AA2	0.317 $\pm$ 0.013
CN2	0.297 $\pm$ 0.011

Dobzhansky and Spassky (1969) selected for both positive and negative phototaxis in *D. pseudoobscura* for 20 generations. When selection was relaxed for 20 subsequent generations, the positive and negative lines both regressed to the point where they became almost indistinguishable. Again, Tantawy et al. (1964) selected for longer and shorter wings in *D. melanogaster* for 20 generations. (Wing length is frequently used as a parameter of body size.) They simultaneously measured egg-to-adult viability and found a decrease in all four selection lines, to approximately 65%–80% of controls (Tantawy et al. 1964, Fig. 11.3). Similarly, when Prevosti (1956) selected for longer and shorter wings on *D. subobscura*, they also measured egg-to-adult viability during the experiment (Figs. 11.3 and 11.4), and found that it was reduced to 40%–80% of controls (Prevosti 1956). These results on phototaxis and body size are typical of the majority of results for other artificially selected traits. Most selection experiments show negative fitness interactions, when these are measured. With prolonged selection, these interactions often become sufficiently deleterious to force the eventual termination of selection.

### III. POPULATION SIZE AND SELECTION EFFICIENCY

The results reviewed in Part I can be summarized by the statement that in panmictic populations founded from the same outbred stock and selected at the same intensity, larger population size accelerates and prolongs response in the selected trait, leading to a higher plateau. The plateau stage, representing the essential exhaustion of the initial, selectable additive genetic variation for the trait, is of course not the ultimate end of response, because mutation continues (Keightley, this volume).

Indeed, in the Illinois long-term selection experiment on *Zea mays*, mutation rates may be high enough to have obscured the loss of initial genetic variation (Walsh, this volume). However, in most experiments the plateau is a recognizable, terminal, perpetual phase of slow response.

In the first one or two generations of selection, the effect of population size is too small to detect. Initial heritability is not detectably affected by the numbers scored, nor by the size of the founding population (short of intentional inbreeding) even if it is very small (Roff and Mousseau 1987). Realized heritability over the somewhat longer short term shows the effect of population size in well-replicated tests. Realized heritability does not predict response at the plateau stage when standing variability is exhausted.

The magnitude of the plateau depends on—and indeed is a kind of assay of—a trait's genetic architecture, especially its polygenicity. At any population size, it is at least theoretically true that in the long term, the most polygenic traits will continue to respond the longest and will evolve the farthest in units of the base-population phenotypic standard deviation. This is the basis for the Castle-Wright formula (Castle 1921; Falconer and Mackay 1996) for effective gene number or apparent polygenicity. Other genetic factors can prolong the period of response, in the sense of slowing it down—factors such as linkage and low initial frequencies of favorable alleles. These and other factors decrease the accuracy of the Castle-Wright formula (Lande 1981; Zeng et al. 1990), without detracting from its logical basis.

Every selection regime involves the operation of selection, tending to fix favorable alleles, and the process of drift, tending to fix alleles at random (Robertson 1960). Drift is reduced in larger populations, hence allelic diversity persists and allows selection more time, permitting greater utilization of genetic variation. But the expected response in larger populations is not only greater at the limit, it is more rapid at all times. Favorable alleles rise in frequency faster in larger populations. They would do so, even if no genes were linked, but the existence of linkage increases the advantage of larger populations. The more numerous recombinants, available in larger populations, dissociate good or bad alleles more effectively from hitchhiking neighbors.

These effects can explain the greater selection response observed in large populations. Aspects of the process were classically modeled by Robertson (1960), and by Hill and Robertson (1966). The incorporation of mutation into this theory (Hill 1982; Hill and Keightley 1988) recognizes the reality that mutations occur during artificial selection, contributing to response (Frankham 1980). This is another reason the



plateau cannot be accurately predicted, but only the response at a given generation (Weber and Diggins 1990).

There is, as yet, no broad experimental basis for the idea that larger populations also have reduced fitness costs during strong (intense or prolonged) selection. It is a reasonable hypothesis for several reasons. First, inbreeding, which has been shown to be one of the main sources of fitness decline in experimental selection (Latter and Robertson 1962), is virtually eliminated in very large populations. Second, natural selection, operating under conditions of somewhat higher population density in mass culture conditions, can cull deficiencies in reproduction and other deleterious effects more effectively. (This is not strictly an effect of larger size, but of conditions attendant to mass culture.) When such culling is done artificially, using family selection on reproductive capacity during a selection experiment, fitness loss can be greatly reduced without reducing trait gains (Frankham et al. 1988). A third reason for the retention of fitness in large populations is that a greater variety of recombinants allows alleles for the trait to rise in frequency with less hitch-hiking of alleles that may depress fitness. Another reason is that, in larger populations, natural selection has a higher probability of ameliorating the deleterious pleiotropic effects of selected alleles. When alleles that are favorable to the selected trait but unfavorable to fitness continue to be elevated in frequency by artificial selection, and cannot be eliminated by natural selection, they must exert a selective force on modifiers that reduce their fitness effects. These weak, second-order effects will be propagated more efficiently through large populations, for the same reasons that large populations respond more efficiently at loci that are the direct targets of artificial selection. Finally, in relation to fitness one must mention the role of mutation in supplying some of the needed modifiers, in numbers proportional to population size.

An interesting question in the background is whether, in the very long term, the efficiency of large populations in sorting out small additive effects, combined with the advantage of large populations in accumulating new mutations, is larger or smaller than the evolutionary potential that may lie hidden within the non-additive component of the gene pool. Following Wright (1931, 1932), many authors believe that the epistatic interactions among genes offer a greater ultimate evolutionary potential than the additive gene action that is directly available to mass selection. This hidden epistatic potential would typically be unavailable to large panmictic populations, according to Wright. In this view, panmictic populations of any size are evolutionary dead ends, and as Wright stressed, mutation supplies no way out. This is a difficult issue that can

only be resolved empirically. Some light may eventually be shed on this by long-term selection in large populations.

## LITERATURE CITED

- Barker, J. S. F., and L. J. E. Karlsson. 1974. Effects of population size and selection intensity on responses to disruptive selection in *Drosophila melanogaster*. *Genetics* 78:715–735.
- Blow, W. L., and D. M. Briggs. 1973. Comparison of selection progress using different population sizes of coturnix quail. *Poultry Sci.* 52:1999.
- Campo, J. L., and H. Turrado. 1997. Population size and selection intensity effects on short-term response for a selection index in *Tribolium*. *J. Anim. Breed. Genet.* 114:107–119.
- Castle, W. E. 1921. An improved method of estimating the number of genetic factors concerned in cases of blending inheritance. *Science* 54:223.
- Dobzhansky, T., and B. Spassky. 1969. Artificial and natural selection for two behavioral traits in *Drosophila pseudoobscura*. *Proc. Natl. Acad. Sci. (USA)* 62:75–80.
- Dudley, J. W., and R. J. Lambert. 1992. Ninety generations of selection for oil and protein in maize. *Maydica* 37:81–87.
- Eisen, E. J. 1975. Population size and selection intensity effects on long-term selection response in mice. *Genetics* 79:305–323.
- Eisen, E. J. 1980. Conclusions from long-term selection experiments with mice. *Z. Tierzuechtg. Zuechtgsbiol.* 97:305–319.
- Eisen, E. J., J. P. Hanrahan, and J. E. Legates. 1973. Effects of population size and selection intensity on correlated responses to selection for postweaning gain in mice. *Genetics* 74:157–170.
- Enfield, F. D. 1977. Selection experiments in *Tribolium* designed to look at gene-action issues. p. 177–190. In: E. Pollak, O. Kempthorne, and T. B. Bailey (eds.), *Proc. Intl. Conf. Quantitative Genetics*. Iowa State Univ. Press, Ames.
- Enfield, F. D. 1980. Long-term effects of selection: the limits to response. p. 69–86. In: A. Robertson (ed.), *Selection experiments in laboratory and domestic animals*. Commonwealth Agr. Bureaux, Slough, UK.
- Frankham, D. S., and T. F. C. Mackay. 1996. *Introduction to quantitative genetics*. 4th Ed. Longman/Essex, UK.
- Frankham, R. 1980. Origin of genetic variation in selection lines. p. 56–68. In: A. Robertson (ed.), *Selection experiments in laboratory and domestic animals*. Commonwealth Agr. Bureaux, Slough, UK.
- Frankham, R. 1983. Origin of genetic variation in selection lines. p. 1–18. In: *Proc. 32nd Annu. Natl. Breeders' Roundtable*, St. Louis, MO.
- Frankham, R., L. P. Jones, and J. S. F. Barker. 1968a. The effects of population size and selection intensity in selection for a quantitative character in *Drosophila*. I. Short-term response to selection. *Genet. Res.* 12:237–248.
- Frankham, R., L. P. Jones, and J. S. F. Barker. 1968b. The effects of population size and selection intensity in selection for a quantitative character in *Drosophila*. III. Analyses of the lines. *Genet. Res.* 12:267–283.
- Frankham, R., B. H. Yoo, and B. L. Sheldon. 1988. Reproductive fitness and artificial selection in animal breeding: culling on fitness prevents a decline in reproductive fitness in lines of *Drosophila melanogaster* selected for increased inebriation time. *Theoret. Appl. Genet.* 76:909–914.

- Hanrahan, J. P., E. J. Eisen, and J. E. Legates. 1973. Effects of population size and selection intensity on short-term response to selection for postweaning gain in mice. *Genetics* 73:513–530.
- Hill, W. G. 1982. Rates of change in quantitative traits from fixation of new mutations. *Proc. Natl. Acad. Sci. (USA)* 79:142–145.
- Hill, W. G., and P. D. Keightley. 1988. Interrelations of mutation, population size, artificial and natural selection, p. 57–70. In: B. S. Weir, J. Eisen, M. M. Goodman, and G. Namkoong (eds.), *Proc. Second Intl. Conf. Quantitative Genetics*. Sinauer Assoc., Sunderland, MA.
- Hill, W. G., and A. Robertson. 1966. The effect of linkage on limits to artificial selection. *Genet. Res.* 8:269–294.
- James, J. W. 1971. The founder effect and response to artificial selection. *Genet. Res.* 16:241–250.
- Jones, L. P., R. Frankham, and J. S. F. Barker. 1968. The effects of population size and selection intensity in selection for a quantitative character in *Drosophila*. II. Long-term response to selection. *Genet. Res.* 12:249–266.
- Lande, R. 1981. The minimum number of genes contributing to quantitative variation between and within populations. *Genetics* 99:541–553.
- Latter, B. D. H., and A. Robertson. 1962. The effects of inbreeding and artificial selection on reproductive fitness. *Genet. Res.* 3:110–138.
- Mather, K., and B. J. Harrison. 1949. The manifold effect of selection. *Heredity* 3:1–52 and 131–162.
- Prevosti, A. 1956. Resultados de la seleccion por alas largas y por alas cortas en dos poblaciones naturales de *Drosophila subobscura*. *Genet. Iber.* 8:73–115.
- Rasmuson, M. 1955. Selection for bristle numbers in some unrelated strains of *Drosophila melanogaster*. *Acta Zool.* 36:1–49.
- Reeve, E. C. R., and F. W. Robertson. 1953. Studies in quantitative inheritance. II. Analysis of a strain of *Drosophila melanogaster* selected for long wings. *J. Genet.* 51:276–316.
- Roberts, R. C. 1966a. The limits to artificial selection for body weight in the mouse. I. The limits attained in earlier experiments. *Genet. Res.* 8:347–360.
- Roberts, R. C. 1966b. The limits to artificial selection for body weight in the mouse. II. The genetic nature of the limits. *Genet. Res.* 8:361–375.
- Robertson, A. 1960. A theory of limits in artificial selection. *Proc. Roy. Soc. Lond. B* 153:234–249.
- Roff, D. A., and T. A. Mousseau. 1987. Quantitative genetics and fitness: lessons from *Drosophila*. *Heredity* 58:103–118.
- Ruano, R. G., F. Orozco, and C. Lopez-Fanjul. 1975. The effect of different selection intensities on selection response in egg-laying of *Tribolium castaneum*. *Genet. Res.* 25:17–27.
- Silvela, L., R. Rodgers, A. Barrera, and D. E. Alexander. 1989. Effect of selection intensity and population size on percent oil in maize, *Zea mays* L. *Theor. Appl. Genet.* 78:298–304.
- Tantawy, A. O., G. S. Mallah, and H. R. Tewfik. 1964. Studies on natural populations of *Drosophila*. II. Heritability and response to selection for wing length in *Drosophila melanogaster* and *D. simulans* at different temperatures. *Genetics* 49:935–948.
- Vasquez, C. G., and B. B. Bohren. 1982. Population size as a factor in response to selection for eight-week body weight in white leghorns. *Poultry Sci.* 61:1273–1278.
- Weber, K. E. 1988a. Systems for measurement and selection of wing-tip height. *Drosophila Information Service* 67:101–103.
- Weber, K. E. 1988b. An apparatus for measurement of resistance to gas-phase agents. *Drosophila Information Service* 67:90–92.

- Weber, K. E. 1988c. An apparatus for selection on flying speed. *Drosophila Information Service* 67:92–93.
- Weber, K. E. 1990. Increased selection response in larger populations. I. Selection for wing-tip height in *Drosophila melanogaster* at three population sizes. *Genetics* 125:579–584.
- Weber, K. E. 1996. Large genetic change at small fitness cost in large populations of *D. melanogaster* selected for wind-tunnel flight: rethinking fitness surfaces. *Genetics* 144:205–213.
- Weber, K. E., and L. T. Diggins. 1990. Increased selection response in larger populations. II. Selection for ethanol vapor resistance in *Drosophila melanogaster* at two population sizes. *Genetics* 125:585–597.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159.
- Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proc. VIth Intl. Congress of Genetics* 1:356–366.
- Yoo, B. H. 1980. Long-term selection for a quantitative character in large replicate populations of *Drosophila melanogaster*. 1. Response to selection. *Genet. Res.* 35:1–17.
- Zeng, Z.-B., D. Houle, and C. C. Cockerham. 1990. How informative is Wright's estimator of the number of genes affecting a quantitative character? *Genetics* 126:235–247.