Photograph by Antonia Reeve Photography

Douglas S. Falconer

Annu. Rev. Genetics 1992. 26:1–14

EARLY SELECTION EXPERIMENTS

D. S. Falconer

Institute of Cell, Animal and Population Biology, University of Edinburgh. Zoology Building, West Mains Road, Edinburgh, EH9 3JT, United Kingdom

KEY WORDS: selection, evolution, speciation, mutation, Mendelism

---

CONTENTS

The first experiments on artificial selection, those done before about 1950, are now largely forgotten. Yet the reasons for doing them, and the conclusions drawn from them, are interesting and worth recalling. In this short review I try to show how the basic understanding of selection responses developed.

THE QUESTION AT ISSUE

To understand the motives for the first experiments on artificial selection we have to go back to pre-Mendelian days, when there was disagreement about the kind of variation responsible for evolution by natural selection. Two sorts

0066-4197/92/1215-0001$02.00

of variation were distinguished: continuous variation among essentially normal individuals, and discontinuous variation seen in "sports" or "mutations." The question was whether continuous variation alone could give rise to evolutionary change when acted on by natural selection. Darwin thought that it could. Those supporting his view were headed by Karl Pearson; these were the "Biometricians" or "Selectionists." Those opposed to Darwin thought that evolutionary change was not possible without discontinuous variation. They were headed by Darwin's cousin Francis Galton and by William Bateson, and they were known as the "Mutationists" or, after Mendel's work came to light in 1900, as the "Mendelians." The controversy between the Biometricians and the Mendelians, which in England developed into a vicious war, is not our concern here. The story is well told by Provine in his excellent book *The Origins of Theoretical Population Genetics* (30), on which much of what follows is based. The Mendelians won; they were right, of course, about Mendelism but wrong about the mutation theory of evolution.

The question at issue was whether selection applied to a continuously varying character could produce large and permanent differences comparable to those between species, as Darwin thought it could. The reason for thinking it could not do so originated in Galton's work, published in papers from 1877 (see 30 for references) and in his book *Natural Inheritance* (12) in 1889. He studied human stature and several other continuously varying characters, and found that the offspring of extreme parents were less extreme than their parents; the offspring exhibited a "regression to the mean." (The fact of regression must have been known to observers of human excellence for a very long time, because it is clearly stated in Homer's Odyssey (36). Odysseus' son Telemachus is lamenting his weakness compared with the godlike strength and courage of his father; Athene, in the guise of Mentor the trusted retainer, says to him: "Few are the sons who attain their father's stature: and very few surpass them. Most fall short in merit.")

Galton found that the amount of the regression was roughly two- thirds in all the characters he studied. In modern parlance this means the coefficient of regression of offspring on mid-parent values was two-thirds. Galton saw this as a universal law applying to all continuous variation. The coefficient two-thirds is what we now call the heritability, though of course it is not the same for all characters. With this correction, Galton's Law of Regression is the basic prediction equation for the response to selection, $R = h^2S$, where R is the response, h^2 the heritability, and S the superiority of the parents. Galton, however, thought that the effects of selection must before long be negated by the regression; a balance must soon be reached between the increasingly extreme values of the parents and the regression that brought their offspring back toward the original mean. Thus, the law of regression prevented selection from producing continuous and permanent change. Where he went wrong in extending the effect of regression to more than one generation was that it is not to the original mean that the new generation regresses by a certain proportion, but to that of the previous generation.

Experiments were needed to resolve the differences between the selectionists and mutationists. The first experiments, however, did not settle the matter because the wrong conclusions were drawn from them.

THE MUTATIONISTS

The first people to experiment on the effectiveness of selection were all convinced mutationists. The mutationists' views about selection, which were widely believed at that time, can be summarized in the following propositions:

1. Selection is effective for no more than a few generations;
2. The mean cannot be changed beyond the range of the original variation;
3. The mean after selection is not stable, but reverts to its original level unless selection is continuously applied;
4. Continuous variation is not heritable;
5. For these reasons selection cannot produce a difference large enough to merit the status of a new race or new species.

To test these propositions experimentally, it is obviously necessary, first, to apply artificial selection continuously over more than a few generations and, second, to stop selecting and see if the change made is sustained without further selection. It is curious therefore that none of the experiments done by the mutationists met these requirements, and it is not surprising that their beliefs were at first unaffected by the evidence.

Hugo de Vries

De Vries, Professor of Botany at Amsterdam, was an influential follower of Galton, and one of the three who rediscovered Mendel's papers in 1900. In his book *The Mutation Theory* (7), originally published in its German edition in 1901, he describes several selection experiments, of which I will mention three.

First, he selected maize (*Zea mays*) for the number of seed-rows in the cob. Pollination was not controlled and took place before selection, so only the seed parents were selected. After seven generations of selection for larger numbers of rows, the mean increased from 13 to 20 rows, a change amounting to about three phenotypic standard deviations. The mean equaled but did not exceed the highest individual found before selection. The heritability can be calculated; it was about 40%.

Second, he selected the buttercup *Ranunculus bulbosus* for increased numbers of petals. The normal number is five, but occasional flowers have six or, very rarely, more. In a wild population 91% had five, 6% had six, and 3% had more—up to one flower with 14 petals. In 1887 he grew seeds from flowers with six or more petals and in 1892, after five generations of selection, the mode of 5 had increased to 9, and the upper limit from 14 to 31 among 4425 flowers scored. The proportion of flowers with five petals was reduced from 91 to 9%. This must seem to us now very impressive evidence of the power of selection, yet de Vries discontinued the experiment because he thought no further progress could be made.

The third experiment refers to the improvement made by plant breeders in the sugar content of sugar beet over a period of about 50 years or 25 generations. The original mean was "about 7–8 percent", and at the end it had doubled to 15%. The highest individual value found before selection was 14%, so in this case the mean after selection did exceed the limit of the original variation, though not by much. Enormous numbers were measured. De Vries tells us that one breeder determined the sugar content of 300,000 plants each year. Unfortunately, he does not tell us the numbers selected as parents. In the last generation, grown in 1896, de Vries measured 43,000 plants. This population had a mean of 15.2 ± 0.005% and a standard deviation of 1.1. If the variation in the initial population had been the same as it was at the end, then the change of the mean amounted to seven phenotypic standard deviations.

These experiments did not shake de Vries' belief in the mutation theory and the limited power of selection to produce change.

Wilhelm Johannsen

The work of Johannsen, the Danish botanist, is better known today than that of de Vries because he introduced two important ideas: the distinction between phenotype and genotype (these were his terms), and the idea of pure lines, i.e. lines of descent by continuous inbreeding, whose members are identical genetically. He worked with the cultivated bean, *Phaseolus vulgaris,* a self-fertilizing plant. His choice of a self-fertilizing species, though it partially vitiated his conclusions, was deliberate; he thought that inheritance in a pure line would be simpler and so make the conclusions clearer. He started the work in 1901 and the first results were published in 1903. Details of what he did are not important here (For a full account and references, see ref. 30). He had a population consisting of 19 pure lines, each descended from a single seed. He measured the weights and dimensions of the beans, i. e. seeds. He found that the beans produced by any one pure line varied, and the pure lines differed from each other in the mean weights of their beans. He selected the largest and smallest beans from each pure line and grew the next generation from them. He found that selection within the pure lines was ineffective, showing that the variation within the pure lines was not heritable. The pure lines, however, differed from each other, and selection in the population as a whole was effective. This selection, however, had done no more than pick out the best pure line and multiply it. Consequently, Johannsen thought that selection of continuous variation could not be effective over more that a few generations, and so could not produce evolutionary change; it could only change the mean within the range of the variation originally present. He later crossed the pure lines and noted the additional variation in the F_2 due to Mendelian segregation, but he still maintained that selection could do no better than pick out the best of the F_2 segregants.

H. S. Jennings

Jennings, at the Harvard Museum of Comparative Zoology, repeated Johannsen's work by similar experiments on the size (length) of *Paramecium,* and confirmed Johannsen's conclusions (see ref. 30). Like Johannsen, he failed to realize that recombination among more than a very few genes could create many new genotypes, some of which would be superior to any that existed before.

W. E. Castle

In 1905, Castle, then Assistant Professor at Harvard University, described an experiment with guinea pigs (2). He was then a believer in the mutation theory, though a later experiment made him change his mind. He points out in this paper that no one since Darwin had given much thought to how animal breeders produce new breeds. He asserts that they do this by utilizing mutations or by combining desirable traits from different breeds. In all cases, he says, the variation pre-existed, and is only discovered by the breeder. His experiment with guinea pigs certainly seemed to fit in with this idea.

Among his stock of guinea pigs he found one individual with an extra toe on one hind foot. Normally there are three toes; this one had four. The father, mated to his daughters, later produced four more guinea pigs with an extra toe, and Castle started a four-toed line by breeding exclusively from four-toed individuals. Individuals varied in how well the fourth toe was developed, and he chose those with the best-developed toe. After three generations all individuals had a well-developed extra toe on both hind feet. "This race," he says "was not *created* by selection, though it was *improved* by that means" (his italics). Sewall Wright later obtained the four-toed strain, which was then highly inbred. He crossed it with three normal three-toed strains, also highly inbred. He showed (33) that the four-toed condition was not a simple Mendelian mutation; the difference between the four-toed and one of the three-toed strains must have been due to at least three genes of comparable importance. Wright's classic analysis of the crosses introduced the idea of threshold characters. What Castle had selected on was the underlying continuous variation of liability to develop the extra toe. Castle was right, however, in saying that no "improvement" by selection would have been possible without the spontaneous occurrence of at least one four-toed individual, i. e. a "mutation."

What now seems remarkable about these experiments up to about 1905 is the power of prejudice to influence conclusions. Because of their limitations, the experiments were incapable of disproving either theory; and because their beliefs were not disproved, the mutationists thought they had been proved right.

POST-MUTATIONISTS

W. E. Castle

The experiment that made Castle change his mind about the mutation theory was done with piebald or "hooded" rats. It was started in 1907 and, with the assistance of J. C. Phillips, it was carried on for 17 generations. The hooded pattern is a Mendelian recessive. The head and mid-dorsal line are pigmented, the rest of the body white. There is variation in the width of the mid-dorsal line, and this was the character selected; individuals were scored in grades and quarters of grades above or below 0. The initial range of grades was from roughly +3 to -2 (details of the original population are not given). Selection was made in both directions, for more pigmentation in the "plus" line and for less pigmentation in the "minus" line.

The results of the first eight generations were briefly described by Castle in 1911 (3). By that time he had rejected the mutationist position; the experiment had provided overpowering evidence in favor of the selectionist theory. The mean grades of both lines had progressed up to or beyond the limits of the original variation; the variation within the lines had not been appreciably reduced; and there was no sign that progress was coming to an end. Thus, Johannsen's belief that selection merely sorted out existing genotypes was disproved. Castle here recognized the power of recombination to create new variation.

The experiment up to generation 13 was documented in detail by Castle & Phillips in 1914 (5), and a few more generations were added by Castle & Wright (6) in 1916, bringing the plus line up to 16 generations and the minus line to 17. The numbers measured per generation averaged about 1000 in each line, though they were much smaller at the beginning—150 in the plus line and 55 in the minus line in generation 1. Graphs of the responses are given by Wright (34), with a fuller summary of the experiment than can be given here. The plus and minus lines differed at the end by 12 phenotypic standard deviations as estimated in generations 1 to 3. The plus line was now entirely black on the back and sides, with white only on the belly; the minus line was entirely white except for the head. The means were +4.1 and -2.7, both well outside the range of the original variation. The final distributions were clearly separated at a grade between 3¼ and 3½, with almost no overlap; from generation 14 onwards only 0.4% of the minus line were above this point and only 0.3% of the plus line below it. The final difference between the two lines was clearly as large as the differences between many "races" or species.

The means of the selected parents are given, so the selection differentials and realized heritabilities can be calculated. It is customary now to weight selection differentials by the numbers of offspring measured. Though they did not calculate selection differentials, Castle & Phillips weighted the parental means in this way; they were, as far as I know, the first to do so. There was a very large response in the first generation, thought to be due to one or two factors of large effect (34). Thereafter, there was continuous and steady progress, with a realized heritability of 32% in both lines, until low fecundity brought the experiment to an end.

To find out if the changes made by selection were permanent or required continued selection to sustain them, Castle & Phillips reversed the direction of selection. If there was any tendency for the selected lines to revert to the original level then, they argued, the response to reversed selection would be quicker than the response to forward selection; it was not. Furthermore, they noted that Galton's "regression to the mean" was in the opposite direction when selection was reversed; when selecting away from the original level, the offspring were less extreme than their parents, but when selecting back towards the original level, the offspring were more extreme than their parents. For these reasons, Castle & Philips concluded that the mutationists were wrong in believing that changes made by selection are not permanent. This experiment very convincingly refuted all the propositions believed by the mutationists.

MAIN GENE OR MODIFIERS

Castle had successfully resolved the conflict between the Mutationists and Selectionists. But, in interpreting his results, he introduced a new misconception which was thought by most people to be mistaken. The experiments that followed Castle's had the testing of his idea as one of their main objects. The idea arose from the experiment with hooded rats and was about what had changed during the selection. Castle thought that it was the hooded gene itself that had changed, becoming weaker and thus causing less white in the plus line, and stronger, causing more white, in the minus line. This idea had some plausibility because a mutation, proved to be of the hooded gene itself, arose during the selection—in generation 10 of the plus line; it increased the amount of black by about 2 grades. There was also other evidence, summarized by Sturtevant (31), that Castle thought supported his view. The alternative view was that it is not the main gene that changes, but genes at other loci that modify the expression of the main gene. There was a simple and obvious way of distinguishing between the two hypotheses. This was to cross the selected lines to an unselected wild-type stock and after extracting the hooded homozygotes to backcross these to the same unselected stock. Castle did this after Sewall Wright had suggested it, and he found that after three successive backcrosses the hooded rats derived from the two selected lines were almost indistinguishable, thus proving that the hooded gene had not been changed during the selection. Castle published his recantation in 1919 (4).

Meanwhile, several experiments with *Drosophila* had been started, or completed. They were all very similar in design, execution, and conclusions. Because one of their main aims was to test Castle's idea, the selection in nearly all was on the degree of expression of a major mutant gene; they all used brother × sister matings to propagate the selected lines; and they all led to the conclusion that the main gene did not change, the response being due to modifiers at other loci. This conclusion raised another question: did the response come from the sorting out of the existing variation present in the base population, or from new variation arising during the selection? If the former, then the response would be confined to the first few generations before the inbreeding had eliminated all the pre-existing variation; if the latter, then the response would continue indefinitely at the same rate. All the experiments showed it to be the former; the responses continued for only a few generations—the half-lives being typically about three generations—and in the later generations there was no further response or, at most, only a very slow one. The total progress made was seldom more that one phenotypic standard deviation of the base population, and the final means were well within the original range of variation.

The particular features of the *Drosophila* experiments were as follows.

E. C. MacDowell

This experiment was started in 1913 and published in 1915, 1917, and 1920 (25–27). The character selected was extra dorso-central bristles. Normally there are four; in the original population 5% of flies had one or more extra bristles. MacDowell says that "extra bristles" was inherited as a single Mendelian recessive, but his evidence does not look convincing. "Extra bristles" seems better regarded as a threshold character like the extra petals of buttercups studied by de Vries, or the extra toe of guinea pigs studied by Castle. Selection was continued for 49 generations. The mean number of extra bristles in the base population was 0.08 and after six generations of selection it reached 2.7 in males and 4.1 in females. Thereafter there was no further increase that could be attributed to genetic change. Relaxed selection from generation 6 was not followed by any change, nor was reversed selection from generation 15. The absence of genetic variation in the later generations was confirmed by an absence of parent-offspring correlation (27). MacDowell's conclusions were: that the "main gene" had not changed; that the response was due to selection having "sorted out" genetic differences present in the original flies with extra bristles; and that additional genetic differences were not shown to occur during the course of the selection.

A. H. Sturtevant

The experiments were started in 1916 and published in 1918 (31). The mutant was *Dichaet,* an autosomal dominant with lethal homozygote. It reduces the numbers of dorso-central and of scutellar bristles, both normally four, to between one and eight in total. There were two lines selected for increased and two for decreased bristle number, all with full-sib mating. There was also one line selected in each direction with "crossbreeding" because it was thought that, without inbreeding, progress might be faster and more prolonged. Selection was continued for between 6 and 14 generations, according to the line. The results of the selection are hardly remarkable. Only one of the four inbred lines gave a clear response; its mean increased by one bristle. Both crossbred lines responded; the one selected downwards responded by one bristle and the one selected upwards by less. Of much more interest was a cross made between the upward-selected inbred line and the downward-selected crossbred line, whose means differed by two bristles. Since *Dichaet* is lethal in homozygotes, the *Dichaet* flies in the F_1 were of two sorts according to which of the parent lines the gene came from. If the *Dichaet* genes were now different as a result of the selection, and if no independent modifiers were involved, two consequences would follow: first, the distributions of bristle number in the F_1 and the F_2 would be bimodal, the modes being separated by two bristles, like an equal mixture of the parental lines; and second, the F_2 would not have an increased variance. The result clearly contradicted these expectations. Neither distribution was bimodal, and the F_2 showed an increased variance. Thus, the experiment provided convincing evidence that the response was due to independent modifiers and not to changes in the main gene itself.

F. Payne

In an experiment published in 1918 (see ref. 28), Payne selected for increased number of scutellar bristles, of which there are normally four, and after 38 generations the mean had increased to 9.1. The experiment started from a wild population in which a few individuals had an extra bristle. In this case a major mutant was not involved. Analysis of the selected line identified at least two genes responsible for the response. In the sixth generation a sex-linked recessive mutation called "reduced" appeared, in which the number of scutellar bristles ranged from 0 to 4. The original mutant individual was crossed to wild-type and a new selection experiment was started from the F_3, the results being published in 1920 (28). Selection was made in both directions, with full-sib mating in both lines, and continued for 60 generations in the plus line and for 65 in the minus line. The responses continued for about 20 generations, which was a good deal longer than in the other *Drosophila* experiments. The mean in the plus line increased from 1.2 to 3.3 (mean of sexes) and in the minus line it decreased to 0.02, when 98% of flies had no bristles. The selected lines were kept in mass culture without selection for eight months and their mean bristle numbers did not change, proving that the changes made by selection were permanent.

C. Zeleny

The major gene used by Zeleny was *Bar-eye*, which is sex-linked and is a duplication rather than a point mutation. It reduces the size of the eye, and its effect was measured by counting the number of facets. Selection was made for both increased and decreased facet number, and was continued for 42 generations with full-sib mating. Facet number has a skewed distribution and a standard deviation proportional to the mean. Zeleny saw the need to make a logarithmic scale transformation and was, as far as I know, the first to do so. He described the results in terms of deviations from the original mean in "factorial units", one unit representing a 10% difference of facet number. A preliminary experiment lasting three generations was started in 1914; the main experiment was started in 1917 and published in 1922 (35). The initial responses lasted about five generations, the mean facet number increasing by 60% and decreasing by 35% (sexes averaged over generations 5 to 12). These responses amounted to an increase of 1.7, and a decrease of 1.0, phenotypic standard deviations of the base population. In this experiment, however, there was some progress in the later generations in both lines. In the high line, between generations 25 and 30, the mean increased a further 15%, making a total response of 75% (= 2.1 σ). And in the low line, between generations 12 and 14, the mean decreased a further 7%, making the total response 60% (= 1.7 σ). After these changes the means remained stable till the end of the experiment. Thus, in each line at least one useful mutation, or recombination, occurred after the original variation had been exhausted. Zeleny (35) summarizes his general conclusion thus: "Selection thus merely exercises a sorting effect and further progress after the preliminary sorting is completed is confined to new mutations whose origin is independent of selection." From this remark it seems that Zeleny had not realized that, if more than a very few genes were involved, recombination between them would produce superior genotypes not available for the "preliminary sorting." If he, and the others working with *Drosophila*, had realized this they would surely not have practiced brother × sister inbreeding.

There were many more experiments than those noted here. Goodale, whose work on selection is described below, gives a synopsis of all selection experiments known to him, many on domestic animals, up to 1936 (14). There were 43 experiments, counting plus- and minus-selection separately. Of these, 34 were judged to have been successful and 9 to have failed. Most of these experiments were published before the mid-1920s, so the efficacy of selection had been firmly established by then. What was not established, however, was for how long the selection would be effective and what would be the limit to progress if close inbreeding were not practiced. Before we consider experiments bearing on limits, however, a digression must be made to describe a remarkable and influential experiment on the learning ability of rats.

R. Tryon

All the experiments described so far were done by geneticists. Tryon, in contrast, was a psychologist and he was not primarily concerned with the process of selection itself. He studied the ability of rats to learn a maze. The objectives of applying selection were to find out how the learning ability was inherited, to produce divergent strains, and to identify the behavioral and physiological traits associated with the maze learning. (Tryon's papers are hard to obtain but for accounts of his experiment, with references, see refs. 11, 20, 24, 29).

The experiment was started in 1926 and the first publication was in 1929. Selection was made in both directions for the number of errors made in running the maze, and was carried on for 21 generations. Unfortunately inbreeding was practiced, as in most experiments at that time. The selection in both directions was undeniably successful. The responses continued for about seven generations, after which there was hardly any overlap between the distributions of the two lines. The mean number of errors was reduced by 33% in the "maze-bright" rats and increased by 37% in the "maze-dull" rats. (See refs. 24, 29 for graphs of the responses and ref. 11 for graphs of the distributions.) The experiment provided very convincing evidence that heredity was one of the factors contributing to the differences between individual rats in their ability to learn a maze. Crosses between the selected lines showed that the inheritance was polygenic. Subsequent behavioral studies of the lines showed that the differences were not in general learning ability, but were rather specific; for example, the maze-dull rats were more easily distracted by the noises made by the mechanical maze used for the selection (11).

LIMITS

The Mutationists, particularly Johannsen, thought that selection could do no more than fix the best of the existing genotypes. When the role of recombination was recognized, it was seen that selection could do more, as was shown by Castle's experiment with rats. When this experiment came to an end after

19 generations, steady progress was still being made and there was no sign of it stopping. The *Drosophila* experiments gave no evidence about what recombination could do because of the inbreeding practiced in all of them, which quickly removed all the genetic variation. To answer the question about limits needed experiments carried out for a long time and with the minimum of inbreeding. There was an interval of over ten years before any such experiments were published; four were started before 1940, one of which had been going on since before Mendel's rediscovery, but it took ten years or more before clear answers to the question of limits were obtained.

Illinois Corn Experiment

This was one of the first experiments on selection. It was started in 1896 by C. G. Hopkins at the Illinois Agricultural Experiment Station, and it continues to this day. Maize was selected for increased and decreased oil content and for increased and decreased protein content in the kernels. The original objective was primarily agronomic improvement. There was no deliberate inbreeding. Unlike most of the other experiments so far, the selection was for characters of normal individuals, and not for the the expression of a mutant or an abnormality. The first 23 generations were described by East & Jones in 1920 (9), and the latest account, after 90 generations, was by Dudley & Lambert in 1991 (8). At the time of the 1920 report the experiment had not gone much further than Castle's rat experiment, and it had given very similar results. The high- and low-oil lines had progressed far beyond the range of the original variation, but the protein lines had barely surpassed the original range (34). Subsequently, both low lines came very near to the absolute limits of 0%. In both high lines progress continued unabated, with no indications after 90 generations that limits were being approached. The experiment is unusual in showing such long-continued progress; most experiments reach limits after a much shorter time.

H. D. Goodale

Goodale, who worked at Mount Hope Farm in Williamstown, MA, was mainly a poultry breeder but he did two selection experiments with mice. In 1937 he wrote a paper (13) with the title "Can artificial selection produce unlimited change?" Unlike most people then, he thought it could. In support of his belief he described, though very sketchily, a mouse experiment which he started in 1931 (21). In a self-colored stock he found one male with a few white hairs on its head. This male was mated to relatives and selection was made in subsequent generations for more white on the head. After "some seventeen" generations without deliberate inbreeding, a large part of the face was white in most individuals, and later (17) the whole face from ears to nose was white, with some individuals having white on other parts of the body, usually belly and tail. Later still (probably about 1961), a few mice appeared that were entirely white except for the eyes (21). This experiment in the end certainly supported the belief that there would be no limit short of 100% white.

In his second experiment, Goodale selected mice for increased body weight (at 60 days). He thought it might be possible by selection to make mice the size of rats. The experiment was started in 1930 and there were reports in 1938 (15), 1941 (16), and 1953 (18), after roughly 14, 28, and 50 generations, respectively. There was also a later and much fuller report in 1971 (32) after 84 generations. Progress continued at a nearly constant rate for about 35 generations and then ceased. It was very clear then that a limit to progress by selection had been reached. The mean body weight had been increased by 72% or about seven phenotypic standard deviations. This was a much larger change than had been achieved by earlier experiments, but the mice were still far short of the size of rats.

J. W. MacArthur

MacArthur, at Marlboro College, VT, selected mice, as Goodale had done, for body weight (at 60 days) without inbreeding, but he selected in both directions and he described the experiment much more fully. The experiment was started in 1939 and the main results were described in 1944 (22) and 1949 (23). After 21 generations of selection, there was almost no overlap between the distributions of the selected lines and the original population. The large line had increased by 75% or 6.5 phenotypic standard deviations, and the small line had decreased by 52% or 4.4 standard deviations. The increase of body weight was nearly the same as that obtained by Goodale. Summarizing the responses, MacArthur said (23), "the selection has achieved diminishing returns as the experiment proceeded. Heritability has declined from about 25 to 10% . The size variation is still there, but less of it is genetic." He could not at that stage say that limits had been reached. However, the strains were continued for a further 10 generations by L. Butler at the University of Toronto, and no further progress was made in either line after generation 21 (1). Here again, therefore, was clear evidence of limits to the progress made by selection. The reason for the maize experiment being so different is still a puzzle.

MacArthur's experiment set a fine example for those who followed in the study of experimental selection. [Some aspects of the later work are reviewed in refs. 10 and 19.]

ACKNOWLEDGMENTS

I am very grateful to Professor R. C. Roberts for advice and suggestions.

Literature Cited

1. Butler, L. 1952. A study of size inheritance in the house mouse. II. Analysis of five preliminary crosses. *Can. J. Zool.* 30:154–71
2. Castle, W. E. 1905. The mutation theory of organic evolution, from the standpoint of animal breeding. *Science* 21:521–25
3. Castle, W. E. 1911. *Heredity in Relation to Evolution and Animal Breeding*. New York: Appleton
4. Castle, W. E. 1919. Piebald rats and selection, a correction. *Am. Nat.* 53: 370–76
5. Castle, W. E., Phillips, J. C. 1914. Piebald rats and selection. *Carnegie Inst. Washington Publ.* 195
6. Castle, W. E., Wright, S. 1916. Studies of inheritance in guinea-pigs and rats. *Carnegie Inst. Washington Publ.* 241
7. de Vries, H. 1910. *The Mutation Theory*, transl. J. B. Farmer, A. D. Darbishire, Vols. 1, 2. London: Kegan Paul, Tench, Trubner
8. Dudley, J. W., Lambert, R. J. 1992. Ninety generations of selection for oil and protein in maize. *Maydica* 37:1–7
9. East, E. M., Jones, D. F. 1920. Genetic studies on the protein content of maize. *Genetics* 5:543–610
10. Falconer, D. S. 1989. Selection experiments and the nature of quantitative variation. In *Evolution and Animal Breeding*, ed. W. G. Hill, T. F. C. Mackay, pp. 121–27. Wallingford, UK: CAB Int.
11. Fuller, J. L., Thompson, W. R. 1978. *Foundations of Behaviour Genetics*. Saint Louis: C. V. Mosby
12. Galton, F. 1889. *Natural Inheritance*. London: Macmillan
13. Goodale, H. D. 1937. Can artificial selection produce unlimited change? *Am. Nat.* 71:433–59
14. Goodale, H. D. 1938. Breeding to meet economic needs. Methods applicable to the improvement of live stock. *Am. Nat.* 72:243–67
15. Goodale, H. D. 1938. A study of the inheritance of body weight in the albino mouse by selection. *J. Hered.* 29:101–12
16. Goodale, H. D. 1941. Progress report on possibilities in progeny-test breeding. *Science* 94:442–43
17. Goodale, H. D. 1942. Further progress with artificial selection. *Am. Nat.* 76: 515–19
18. Goodale, H. D. 1953. Third progress report on breeding larger mice. *J. Genet.* 51:580
19. Hill, W. G., Caballero, A. 1992. Artificial selection experiments. *Annu. Rev. Ecol. Systemat.* 23:287–310
20. Hirsch, J. 1967. *Behaviour-Genetic Analysis*. New York: McGraw-Hill
21. Kyle, W. H., Goodale, H. D. 1963. Selection progress toward an absolute limit for amount of white hair on mice. In *Genetics Today; Proc. Int. Congr. Genet., 11th*, ed. S. J. Geerts, 1:154–55. Oxford: Pergamon
22. MacArthur, J. W. 1944. Genetics of body size and related characters. I. Selecting small and large races of the laboratory mouse. *Am. Nat.* 78:142–57
23. MacArthur, J. W. 1949. Selection for small and large body size in the house mouse. *Genetics* 34:194–209
24. McClcarn, G. E., DeFries, J. C. 1973. *Introduction to Behavioural Genetics*. San Francisco: Freeman
25. MacDowell, E. C. 1915. Bristle inheritance in Drosophila. I. Extra bristles. *J. Exp. Zool.* 19:61–98
26. MacDowell, E. C. 1917. Bristle inheritance in Drosophila. II. Selection. *J. Exp. Zool.* 23:109–46
27. MacDowell, E. C. 1920. Bristle inheritance in Drosophila. III. Correlation. *J. Exp. Zool.* 30:419–60
28. Payne, F. 1920. Selection for high and low bristle number in the mutant strain "reduced". *Genetics* 5:501–42
29. Plomin, R., G. E. 1980. *Behavioural Genetics*. San Francisco: Freeman
30. Provine, W. B. 1971. *The Origins of Theoretical Population Genetics*. Chicago: Univ. Chicago Press
31. Sturtevant, A. H. 1918. An analysis of the effects of selection. *Carnegie Inst. Washington Publ.* 264
32. Wilson, S. P., Goodale, H. D., Kyle, W. H., Godfrey, E. F. 1971. Long term selection for body weight in mice. *J. Hered.* 62:228–34
33. Wright, S. 1934. The results of crosses between inbred strains of guinea pigs, differing in number of digits. *Genetics* 19:537–51
34. Wright, S. 1977. *Evolution and Genetics of Populations*. Vol. 3. *Experimental Results and Evolutionary Deductions*. Chicago: Univ. Chicago Press
35. Zeleny, C. 1922. The effect of selection for eye facet number in white bar-eye race of *Drosophila melanogaster*. *Genetics* 7:1–115
36. *The Odyssey of Homer translated from the Greek by T. E. Shaw*, 1940, p. 32. New York: Oxford Univ. Press