EVALUATION BY GENETIC SIMULATION OF CHANGES IN A CROSSBRED POPULATION RESULTING FROM SELECTION IN A PUREBRED POPULATION

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

KENNETH EVANS KEMP

1967

This is to certify that the

thesis entitled

Evaluation by Genetic Similation of Changes in a Crossbred Population Resulting from Selection in a Purebred Population

presented by

Kenneth Evans Kemp

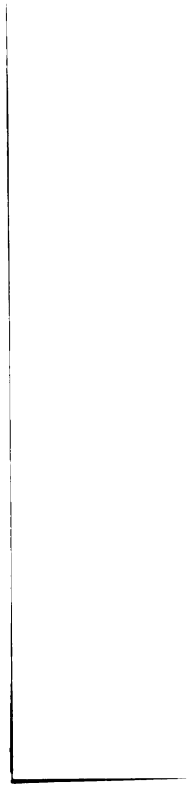
has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Animal Husbandry

R. L. Baugus
Major professor

Date 11 2/67

O-169



EVALUATION BY GENETIC SIMULATION OF CHANGES IN A CROSSBRED POPULATION RESULTING FROM SELECTION IN A PUREBRED POPULATION

By

Kenneth Evans Kemp

AN ABSTRACT OF A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Animal Husbandry

1967

ABSTRACT

EVALUATION BY GENETIC SIMULATION OF CHANGES IN A CROSSBRED POPULATION RESULTING FROM SELECTION IN A PUREBRED POPULATION

by Kenneth Evans Kemp

A proposed breeding plan was evaluated as to its effectiveness in improving the performance of crossbred swine in 5 generations of selection, when selection was entirely within pure breeds. To achieve this goal 2 breeds were simulated, by using a digital computer and a random number generator. A cross of the two breeds was made before the first and after the fifth generations of selection, in order to measure the improvement made in the crossbred pigs. Selection was on 4 traits; 2 within each breed. The selected traits were the 4 major production traits in swine; daily gain, feed efficiency, litter size, and backfat probe. The breeding plan was designed to be one which a purebred swine breeder could follow with the usual facilities and equipment.

In the first breed selection was for feed efficiency and daily gain. Twenty simulated boars were evaluated on the basis of their feed efficiencies, and the 2 most efficient of the 20 were selected to be sires of the succeeding generation. Thirty gilts were randomly generated and the 10 gilts with the largest average daily gains were selected. In standard units, the expected selection differentials were 1.8 and 1.1 for the boars and gilts, respectively. Each gilt bred produced 5 offspring; 3 gilts and 2 boars.

The selection in the second breed was more complicated than that in the first. In this breed, selection was for improved litter size and backfat probe. Each generation 48 gilts and 36 boars were produced. The

20 leanest gilts and the 4 leanest boars were selected and bred together twice. Each mating produced 3 boars and 4 gilts. The 20 leanest of the first litter gilts which were from the 12 dams with the highest average litter size were then selected on backfat probe, and the 4 leanest of the first litter boars from the 8 dams with the highest average litter size were also selected on backfat probe, but with the restriction of only one boar from a litter. This selection was carried forth for 5 generations. The expected selection differentials in standard units were 1.51 and 0.93 for backfat for the boars and gilts, respectively. Selection for improved litter size was on pedigree, and the expected selection differentials were 0.48 and 0.32 standard deviations for the boars and gilts, respectively.

The genetic parameters of the simulated population of swine were taken from estimates found in the literature. The heritabilities of daily gain and feed efficiency were both 0.3, the heritability of litter size was 0.1 and the heritability of backfat probe was 0.5. There was a 0.6 genetic correlation between daily gain and feed efficiency and a 0.4 genetic correlation between daily gain and backfat probe. All other genetic correlations were zero. The phenotypic variances in the initial generation of each run were 0.0289, 0.00072, 5.30, and 0.0196 for daily gain, feed efficiency, litter size, and backfat probe, respectively. The initial means of the traits were set arbitrarily at 1.6 lb./day, 0.31 gain/feed, 8.0 pigs/litter and 1.6 in. for daily gain, feed efficiency, litter size, and backfat probe, respectively.

Daily gain, feed efficiency, and litter size were all under a complete dominance gene model, while backfat probe was a completely additive trait. Each trait was controlled by 20 pairs of independently segregating

loci, i.e., no linkage, and there was no epistasis. All genes had equal effects. The initial frequency of the desired gene for all traits except backfat probe was 0.6, while it was 0.4 for backfat probe since the genetic correlation between daily gain and backfat probe was positive and this forced the genes which increased daily gain to also increase backfat probe, but increasing backfat was undesirable.

The breeding plan was found to be effective in improving the means of all four traits in the crossbred pigs. This was due mainly to an increase in the frequency of the desired gene for all traits, but there was an additional 6% and 4% improvement in daily gain and feed efficiency, respectively, due to heterosis. There was a 2.5% increase in the mean for daily gain from the first to the second cross, and a 3.8% improvement in the mean for feed efficiency. The mean for litter size was least improved, 1%, and the mean for backfat probe was most improved, 14.6%. The increases in the frequencies of the desired alleles were somewhat better. The improvement in the frequency of the desired gene was 3.8% for daily gain, 10.3% for feed efficiency, 9.8% for litter size, and 23.9% for backfat probe. However, the breeding plan was criticized for causing the accumulation of inbreeding at a rapid rate, and the recommendation for larger numbers in the breeding herd was made.

Selection within pure breed I caused a 5.4% and 5.9% improvement in the mean, and a 26.1% and 26.5% improvement in the frequency of the desired gene for daily gain and feed efficiency, respectively. In the second breed, the mean for litter size improved 1% and the frequency of the desired gene improved 9.8%. The mean for backfat probe improved 21%, while the frequency of the desired gene improved 66%. In all cases the means for the unselected traits within each breed deteriorated.

The expected means within the pure breeds were calculated and compared to the observed means. The means for the unselected traits were accurately estimated, but the means for the selected traits were generally underestimated. When a trait was being selected for the discrepancy between the observed and expected means became worse as inbreeding increased.

The expected gene frequencies within the pure breeds were also calculated and compared to the observed gene frequencies. Although the predictions of the gene frequencies were dependent on the means, they were generally better than the predictions of the means, but showed basically similar responses. This was explained, at least partially, by the fact that inbreeding has an effect on the observed mean, but not on the observed gene frequency.

The expected inbreeding was calculated and compared to the observed inbreeding within each of the pure breeds. The agreement between the observed and expected inbreeding was very good in the first breed, but very poor in the second. The discrepancy between the observed and expected inbreeding in the second breed was attributed to the effects of selection and a slower rate of inbreeding which made the effects of selection relatively more powerful in the second breed than it was in the first breed.

The expected genetic variances were computed and compared to the observed genetic variances within each of the pure breeds. The predictions were very close to the observed results, but were generally better for the unselected than the selected traits. The genetic variance in the selected traits tended to be overestimated by a small amount.

The decline from the initial to the final generation of selection in the genetic variance in the selected traits was considerable in both

The expected means within the pure breeds were calculated and compared to the observed means. The means for the unselected traits were accurately estimated, but the means for the selected traits were generally underestimated. When a trait was being selected for the discrepancy between the observed and expected means became worse as inbreeding increased.

The expected gene frequencies within the pure breeds were also calculated and compared to the observed gene frequencies. Although the predictions of the gene frequencies were dependent on the means, they were generally better than the predictions of the means, but showed basically similar responses. This was explained, at least partially, by the fact that inbreeding has an effect on the observed mean, but not on the observed gene frequency.

The expected inbreeding was calculated and compared to the observed inbreeding within each of the pure breeds. The agreement between the observed and expected inbreeding was very good in the first breed, but very poor in the second. The discrepancy between the observed and expected inbreeding in the second breed was attributed to the effects of selection and a slower rate of inbreeding which made the effects of selection relatively more powerful in the second breed than it was in the first breed.

The expected genetic variances were computed and compared to the observed genetic variances within each of the pure breeds. The predictions were very close to the observed results, but were generally better for the unselected than the selected traits. The genetic variance in the selected traits tended to be overestimated by a small amount.

The decline from the initial to the final generation of selection in the genetic variance in the selected traits was considerable in both

breeds. There was a 52% decrease in the genetic variance in daily gain and a 58% decline in the genetic variance for feed efficiency in breed I. In breed II the genetic variance for litter size was reduced by 24% and the genetic variance for backfat probe decreased 46%. The genetic variance for the unselected traits were decreased by 18% or less within each breed. The decreases in the genetic variances were detrimental to the progress made by selection, and again a larger breeding population was recommended, especially for breed I, to reduce the rate of decline due to inbreeding. The increase in population size was considered less crucial for breed II since much of the increase in homozygosity in this breed was apparently due to selection, and was therefore considered unavoidable.

There was very little change in the genotypic correlations, in general. The only exception was the genotypic correlation between daily gain and feed efficiency in the first breed, which declined from an initial 0.59 to a final 0.48.

EVALUATION BY GENETIC SIMULATION OF CHANGES IN A CROSSBRED POPULATION RESULTING FROM SELECTION IN A PUREBRED POPULATION

By

Kenneth Evans Kemp

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Animal Husbandry

1967

EVALUATION BY GENETIC SIMULATION OF CHANGES IN A CROSSBRED POPULATION RESULTING FROM SELECTION IN A PUREBRED POPULATION

By

Kenneth Evans Kemp

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Animal Husbandry

1967

ACKNOWLEDGEMENTS

I am greatly indebted to Dr. William T. Magee, professor in Animal Husbandry, for his assistance in the preparation of this manuscript. I would like to express my appreciation for his counsel and guidance all through my graduate career, and for his unending patience both in and out of the class room.

I appreciate the help and advice extended by Mr. W. L. Ruble in the programming of this work.

I thank Mrs. Kathryn Ide for the typing of this manuscript.

I am grateful to the Department of Animal Husbandry for the financial assistance provided in the form of a C. E. Wilson Research Fellowship.

I am most deeply grateful and indebted to my parents and my wife, Sue, for their continuous support and sacrifice.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The effectiveness of a breeding plan is of paramount importance to the maximization of progress which may be made by selection. Considering the cost and long generation intervals involved in selection experiments with large domestic animals, it would be very difficult to overemphasize the importance of selecting the breeding plan that would maximize the use of existing genetic variability in such populations. Taking these same considerations into account, the need for a method which would accurately predict the response of a domestic population to a specific breeding plan is crucial. Presently there are two general methods for predicting the response of a population to selection under a specific breeding plan. One method is that of mathematical prediction using the existing genetic theory. The other is using biological populations, which have short generation intervals and can be raised economically, as pilot studies of a specific breeding plan.

The limitations of mathematical models in predicting response to selection are many. Presently only equations for relatively simple models have been developed and even these simple models are generally restricted by unrealistic assumptions which must be met if the predictions using these equations are to be accurate. Some of the more frequent restrictions imposed in the use of these equations are infinite population size, hence no random drift or inbreeding, equal numbers of parents of each sex, no epistasis and no linkage. Some equations have been derived which allow for the effects of these sources of variation, but they generally include unestimable quantities which render them useless for predicting the response to selection.

The accuracy of the results of pilot studies using organisms other

than those for which the plan is ultimately intended is always questionable. Although this method has the advantages of being able to use very complex breeding plans and considerably reducing the duration of selection compared to what it would be in the experimental population, the end results of such predictions have very wide fiducial limits. It can never be known whether the two populations, pilot and experimental, have similar population parameters or even, in many cases, similar modes of gene action. To the extent these populations differ from one another, so may the realized response to a specific breeding plan.

In the last decade considerable work has been done using a relatively new method of predicting response to selection. This recently developed method has been called the Monte Carlo method since it is based entirely on random numbers which are generated by digital computers. To date, this method has proved very valuable in studying the effects of various modes of gene action, selection intensity, population size, tightness of linkage, and environmental variation on genetic response to selection. However, very little work has been done in simulating specific populations of biological organisms. Most of the Monte Carlo work has been aimed at answering basic questions of genetic theory rather than simulating responses of actual populations.

This investigation was designed to use the Monte Carlo simulation technique to evaluate the probable effect of a specific selection program and breeding plan on a swine herd. The true validity of the results, of course, cannot be determined until an actual population of swine is subjected to the same breeding plan that is simulated. However, it may be useful in comparing the results of the simulation to those based on genetic theory and mathematical prediction equations to see

where they differ and perhaps to evaluate why they differ. If, at some future date, the results of such a simulation are compared to the results of an actual experimental population and found to be more accurate than either mathematical or pilot population prediction, it could prove to be an invaluable approach to evaluating, accurately, the effectiveness of a specific breeding plan while being considerably faster than pilot studies and yet allowing evaluation of more complex plans than can be handled mathematically. Of course, the accuracy of the simulated results will always be a direct function of the accuracy of the population parameter estimates.

REVIEW OF THE LITERATURE

Since there are numerous aspects of selection and many secondary effects resulting either directly or indirectly therefrom, no attempt will be made to review all literature on all aspects of genetic selection. Any attempt to review all facets of genetic selection programs could be no more than cursory. Therefore, it seems desirable to review the more important factors with some degree of thoroughness in lieu of a superficial treatment of all aspects of this particular investigation. The two aspects of this investigation which seem most pertinent are the Monte Carlo method itself, and the estimates of the genetic parameters of the simulated population. If either of these are not reliable and well founded, then neither will be the resulting predictions of genetic response.

The Monte Carlo Procedure

In 1957, Fraser, an Australian, introduced the application of the Monte Carlo procedure to the simulation of genetic systems with a paper on the techniques he used on Sydney's Silliac computer. In this paper (Fraser 1957a) he introduced the use of logical operations, i.e., bit-by-bit comparisons between words of memory, for determining the state of any given locus when the simulated genes were generated and stored as haploid genotypes in separate words of memory. These logical operations were the logical product or AND operation, the logical non-equivalent or the exclusive OR operation, and the not-sum which is the complement of the inclusive OR operation. For example, if the desired gene is represented by the digit 1 and the undesired gene by the digit 0, given the gametes 001 and 011, each with genes at 3 loci, the above operations

produce the following results. The logical product of the two is 001 and is used to indicate that the third locus is homozygous for the desired allele. The logical non-equivalent of the two gametes is 010 and shows that the second locus is heterozygous. The logical not-sum is 100 and thus indicates that the first locus is homozygous for the unfavorable gene. He also discussed methods for determining phenotypic values, simulating inter-locus interactions, environmental effects, segregation, recombination and selection.

In the second of Fraser's series of papers (Fraser 1957b), he investigated the effects of linkage on rates of advance under selection. The program simulated six loci, all on the same chromosome, and provided for varying recombination frequencies, from independent assortment to complete linkage. Selection was continued over twenty to twenty-five generations for two replications of each of the twenty combinations of parameters. The parameters were population size, selection intensity and linkage. However, population size was confounded with selection intensity since in the large population (100) the number saved for parents was always fifty and in the small population (40) the number saved was always four. Indications from this study were that linkage had no effect until it became tighter than a recombination frequency of 0.025. And the effect of tight linkage was greatly exaggerated in the smaller population under the more intense selection.

The Sydney series of papers was continued by Barker (1958a). In this paper, Barker made the first attempt to simulate actual biological populations. Two populations of Drosophila were simulated. Selection between two autosomal alleles at four stages of the life cycle was

simulated by the methods outlined by Fraser (1957a). The first population was a large population of Drosophila pseudo-abscura in which selection was between the ST and CH chromosomal arrangements in the third chromosome. The experimental and simulated results of this study were in very close agreement. The second population was a small population of Drosophila melanogaster in which the simulation was of the competition between wild type and the glass mutant. The results of the simulation of this population, however, did not compare favorably with the experimental results. This served to point out that simulated results are no better than the estimates of the parameters used in the simulation. It also shows that if the parameters can be accurately estimated, so can the results of selection, using the Monte Carlo procedure.

Barker published a second paper (1957b), the fourth of the Sydney series, which was very similar to his first, except that selection was at sex-linked loci rather than autosomal loci. The results were also similar in that he was successful in one case but not in the other. It showed that sex-linked loci selection may be simulated by a digital computer with the same restriction as to the accuracy of the parameter estimates as in the autosomal case.

Fraser (1960a) resumed his work in the Sydney series on the simulation of genetic systems in a fifth publication. In this paper, he reiterated the general principles of Monte Carlo simulation with particular reference to the Silliac computer. He discussed techniques which could be used to simulate genetic phenomena such as; segregation, identification of genetic structure, formation of gametes, determination of phenotype, simulation of environmental effects, recombination and

selection as he had done in his first paper. He restated the findings of his second paper, that in large populations (100) with low selection intensities (½) there were no periods of slow response followed by periods of sudden response, as proposed by Mather (1943), for any linkage value. He did find, however, that linkage in smaller populations (40) under more intense selection (1/10) seemed to cause results somewhat similar to Mather's, at least in some cases where linkage was very tight. The parameters in the latter population were, in fact, more nearly the same as those used by Mather (1943) in his experiment.

The results of a program which was written to study the effects of dominance and epistasis on selection against the extremes were also presented in this paper to illustrate the use of Monte Carlo in the "variable" parameter method. The degrees of dominance, epistasis and tightness of linkage were under genetic control and, therefore, varied with the genotypes of the individuals. The response to selection was measured as the reduction in variation which occurred, and could be analyzed into components due to homozygosity, dominance, epistasis and linkage. The major effect of selection against the extremes was on dominance since it was, apparently, one of the major causes of deviation from the intermediate values. Selection operated against the causes of deviation from the optimum and tended to cause the development of "canalization" along the path that led to the optimal phenotypic expression. The role ascribed to selection was its discrimination against alleles that increased variability. Such selection strongly favored epistasis when the relation of the genotype to the phenotype could be varied into the form of a sigmoid function. The major weakness of his conclusions on epistasis was that the epistasis he simulated was not

epistasis in the classical sense of being an interation between individual non-allelic genes. Rather the epistasis was a non-linear transformation of the "dominance plus additive" phenotype which simulates interlocus interactions of a sort, but may or may not be a realistic phenomenon.

Fraser's conclusion from the two papers presented was that they served to show the flexibility of Monte Carlo applications to the solution of biometrical genetic problems. He also pointed out that it may serve to bridge the gap between the present status of mathematical genetics and experimental genetics, and that it could serve as a useful tool to maximize the probability of the experimenter obtaining the desired results in actual experiments, if the parameters were first evaluated by Monte Carlo procedures.

Fraser (1960b) discussed the effects of selection against the extremes in the presence of epistasis in a sixth paper of the Sydney series on simulation of genetic systems. In this program he simulated forty loci using logical operations on the maternal and paternal genotypes instead of the "random walk" which he had used in earlier programs to produce offspring. The frequency of the desired gene was 0.5 at each locus in the initial populations. The results of this study showed a greater decrease in variability, for a specified deviation from the initial gene frequency, under selection against the extremes, in the epistatic model than in the additive model. There was a marked affect of population size on genetic fixation, but little, if any, effect due to selection in the epistatic model. Selection in the additive model caused a slow accumulation of genetic fixation. Selection in the complex epistatic systems modified the relation of the genotype to the phenotype into a sigmoid function. Continued selection should, eventually, result in fixa-

tion of the epistasis - determining loci. The fixation of these loci may cause the fixation of the sigmoid relationship between the phenotype and genotype and this will have unexpected effects on the progress of selection towards an extreme phenotype. The initial response to such selection would be, as expected, a straightforward shift of the distribution. However, as selection proceeds the distribution will be moved toward the inflection point of the genotype - phenotype relationship, and this will produce a marked decrease in the phenotypic variance, due to an increase in the proportion of genotypes having the same phenotypes. Further selection for the extreme phenotype will be against genetic extremes, resulting in fixation of the additive loci.

Fraser (1960c) published the seventh paper of the Sydney series as a direct extension of the sixth and as a further investigation of epistasis and selection against the extremes. The primary objective of this study was to examine the effects of varying numbers of offspring per mating and of varying selection intensities. The conclusion of the previous work, that there was no, or little, effect due to selection against extremes in the epistatic model on fixation, was confirmed. It was, therefore, concluded that the trend toward genetic fixation is predominantly controlled by the population parameters of reproductive rate and number of parents. A possible explanation for the "canalization" of a character was presented in the form of a cubic equation of the sigmoid function. This equation, if it holds true in nature, could explain the reduction in phenotypic variance when selection is for an extreme and, therefore, the canalization of the character.

Martin and Cockerham (1960) were the first to publish results from a Monte Carlo genetic simulation in the U.S. The program allowed for varying

numbers of loci, from 1 to 25, recombination fractions from 0.5 to 0.01 and any degree of dominance, but no epistasis. The study was done primarily to investigate the effects of linkage, but the authors also considered the effects of selection intensity, population size and environmental variation. The initial gene frequency in all runs was 0.5, and selection, which was by upper truncation, was based on the individual's phenotype. Selected parents were mated at random with replacement. All loci had equal effects which combined additively, and the genotype and environmental values were additive and uncorrelated.

The results of the first series of simulations were based on five loci. For the additive model with no environmental variation, tight linkage seemed to retard progress and more intense selection caused more rapid progress. However, when environmental variation was present, the linkage effect disappeared. In the case of complete dominance, none of the parameters had any effect on the response rate. The effect of linkage was greatly increased when the initial parental generation was in a state of "loaded repulsion", i.e., all adjacent, linked loci in a repulsion state, than when the initial population was in linkage equilibrium.

The second series of runs was based on twenty loci rather than five. The larger number of loci seemed to allow more progress in the less intensely selected populations because intense selection caused fixation of undesirable genes. Larger population size, under the same recombination frequency and selection intensity, allowed faster progress because of more genetic recombination. Environmental variation had the expected effect of slowing progress.

Baker and Comstock (1961) did a simulation aimed primarily at evaluating the effects of linkage on preserving polymorphism in a population under specified conditions. The model was complete dominance, all loci having equal effects. The initial population was in linkage equilibrium and the initial gene frequency was 0.5 at each locus. The fate of thirty-five loci was studied but only twenty-eight of these affected a single quantitative trait which was under selection. The genes affecting the quantitative trait were subdivided into blocks of four, but all blocks were on the same chromosome. The recombination frequency for adjacent loci was uniform and there was no cross-over interference. There was no epistasis. The recombination frequencies simulated were 0.5 and 0.01. The initial population was produced by 2 parents but the number of parents in succeeding generations varied from 2 to 10. The offspring population size varied from 4 to 100. The variance due to the combined effect of the environment and those genes affecting the quantitative character which were not traced, i.e., simulation of approximately ten other chromosomes, varied from zero to 324. Selection was continued for the number of generations it would take for the expected heterozygosity to become half of the original value, based on the approximation of $\frac{1}{2N}$ as the loss in heterozygosity per generation due to finite population size. Selection was done at random for various parameter combinations and these runs were used as controls for comparing with the free-recombination-selection results and the linkage-selection results.

The resulting data showed that selection caused a decrease in the amount of genetic fixation that took place and that linkage caused an additional reduction in the amount of fixation, although the differences

were not statistically significant in many cases. Contrary to the findings of Martin and Cockerham (1960), there was no retardation of the progress made in changing the genotypic mean due to tight linkage. This discrepancy in results, however, does not seem alarming when one considers the differences in the parameters of the respective populations. For example, the smallest number of parents used by Baker and Comstock was 6, while the smallest number used by Martin and Cockerham was 4. To attain a selection intensity of 0.1, Baker and Comstock selected 6 parents from 60 offspring and 10 parents from 100 offspring, while Martin and Cockerham selected 4 parents from 40 offspring. Since the availability of better genotypes, necessary for progress from selection, is a joint function of selection intensity, population size and recombination fraction, it is apparent that Baker and Comstock's populations would not be expected to show as great effects due to tight linkage as Martin and Cockerham's. This is because when the populations were under the same recombination frequency and selection intensity, Baker and Comstock's populations were larger.

Qureshi (1963) did a Monte Carlo experiment in which he considered the effects of linkage, population size, degree of truncation selection, and environmental variation on genetic response to selection. He simulated forty loci, four linkage groups of ten loci each, which were heterozygous initially and selected for thirty generations, or until all loci were fixed. The design of this experiment was a 3^4 factorial design with two replications of the eighty-one treatment combinations. The levels of each factor used in the experiment were: for population size; 8, 16 and 64 individuals; for linkage 0.5, 0.05, and 0.005; degree of truncation ½, ¼ and 1/8 of the offspring saved as parents for

the next generation; and the levels of environmental variation were 0, $2\sigma_{g_o}^2$ and $8\sigma_{g_o}^2$, where $\sigma_{g_o}^2$ was the expected genetic variance in the base generation. In this experiment Qureshi considered only the case of an additive model. He found all four main effects to be important in changing the genotypic mean of the population. The rate of change in the mean was greater in the larger populations. Linkage was important in all three possible two-way interactions in most generations. There was evidence of a selection intensity by environmental variance interaction in the early generations and a three-way interaction of linkage, selection intensity and environmental variance for all generations beyond the thirteenth.

Noticing the responses of the genotypic means were quite curvilinear up to the point of fixation, he fitted a second degree polynomial of the form

$$E(\bar{g}_t) = a + \beta_1 t + \beta_2 t^2$$

where $\bar{g}_t$ is the genotypic mean in generation t. This equation seemed to fit the response curves satisfactorily up to the point of fixation since R^2 was more than 90% in most cases. Qureshi considered $\hat{\beta}_1$ as an estimate of the change in the genotypic mean and $\hat{\beta}_2$ as an estimate of the change in β_1 over the generations studied. His estimates of β_1 appeared to closely agree with the predicted value of $\Delta\bar{g}$ (where $\Delta\bar{g} = h^2 k/b\sigma_p$) only when there was no linkage and the parental population size was as large as 16.

In this simulation he also studied the changes in the genotypic variance and the fixation of genes. He found that the changes in the genotypic variance were rather erratic except in the cases where population size was large and there was no linkage. Linkage had the largest

effect, by decreasing the genotypic variance, at all levels of the other variables, but this effect was more pronounced in smaller populations. The joint effect of population size and linkage was the major cause of fixation on genes in the first ten generations or so.

In a subsequent report (Qureshi 1964) the study of the effects of population size, linkage, selection intensity, and environmental variation was continued for the case of dominance. In this experiment two specific types of dominance were studied, complete dominance and overominance. In the first case the desired gene was a complete dominant to its allele. The second case was a case of overdominance where the homozygotes were equally preferred with the heterozygote being the most preferred. The design of this experiment was identical to that of his first work with the additive model. In the case of complete dominance, tight linkage caused a negative response to selection in the smaller populations, primarily due to the fixation of undesirable genes. The fixation of the undesirable genes was due to the fixation of desirable genes, caused by selection, with which the undesirable genes were linked. The response was positive only in the largest population. More intense selection generally caused a positive response except when linkage was very tight or population size very small. As was the case for the additive model, the fixation of undesirable genes was due primarily to linkage and population size. Almost no fixation occurred in the larger populations when recombination was moderate to high, while considerable fixation was present in small populations with tight linkage. In the case of overdominance, the total response of the genotypic mean and the fixation of genes over the thirty generations was also apparently due entirely to population size and linkage, although the rate of the

response was evidently affected by the intensity of selection. Low recombination appeared to invariably hinder the effectiveness of selection and the magnitude of this effect was definitely non-linear with respect to parental population size. The expected stable equilibrium which was anticipated by the author was realized when the population was large, selection intensity low and recombination moderate to high. In general, the effect of selection intensity appeared to be more or less additive to the effects of population size and linkage in the case of both models of gene action. A strong interaction between population size and linkage was consistently present with respect to both the response of the genotypic mean and the fixation of loci. The effect of linkage was always more pronounced when the parental population size was reduced below a certain level.

Gill (1965a, b and c) presented a series of three papers in which he investigated the effects of population size, degree of truncation selection, environmental variation, linkage and mode of gene action on the response to selection. Unisexual diploid individuals were simulated and a quantitative characteristic was expressed in both sexes. The trait was determined by forty loci which were equally spaced over eight chromosomes, with two alleles per locus and equal effects at all loci. Equal numbers of parents of each sex were selected by upper truncation and selection was continued over thirty non-overlapping generations, or until all loci were fixed. Nine genetic models were simulated. Three were standard, non-epistatic models, additive, complete dominance and overdominance; while the other six included some sort of epistasis. The overdominance model was the case where the homozygotes were equally preferred. The design of the experiment was a one-sixteenth fractional

replication of a 4^4 factorial design. Because of the confounding in the design, the estimates of the main effects were valid only in the cases where the aliases, interactions in all cases, were considered negligible. The effects of population size and environmental variation were confounded with the two-way interactions of selection intensity with population size and selection intensity with environmental variation, respectively, and these have been inferred to be non-zero. Thus, the main effects of these parameters must be interpreted with caution in this series of investigations.

Gill's first paper (1965a) dealt primarily with the effect of finite population size on the advance in the genotypic mean due to selection. He found that the effect of population size on genetic progress was of major importance for only four of the nine genetic models simulated; complete dominance, overdominance, complementary factors, and dominance-by-dominance conditional epistasis. These were the only models simulated that involved large portions of variation due to dominance. Inbreeding had its greatest effect on the overdominance and dominance x dominance models, causing a negative regression of the genotypic mean on generation number even though, in some cases, there was strong upward selection. Gill concluded that the critical size of a simulated population with respect to the prevention of random extinction of desired alleles was between 16 and 32 individuals for the complete dominance model, while a population of 30 or more was needed when overdominance existed and selection intensity was strong, 1/8 or stronger. The results of this investigation, generally agreed with existing theory.

The second paper (Gill 1965b) dealt with the weakness of existing

theoretical equations for the prediction of expected response to selection in finite populations. Four genetic models were evaluated; additive, complete dominance, optimum number and additive-by-additive conditional epistasis. The response in the additive and complete dominance models was predicted with the usual linear equation, while the predictions of the response in the epistatic models were based on the equation developed by Griffing (1960) for models that include both, or either, linkage and an additive-by-additive variance component. Two methods of prediction were used; the first was based on the parameters in the initial population and the second was based on parameters in an early generation other than the first.

Gill concluded, "The predicted contribution to change in the mean attributed to additive-by-additive genetic variance, in most cases, was far too large over several generations of selection, or for even shorter periods in the smallest population. Random genetic drift and selection appeared to have considerable influence in changing the genetic parameters quickly. The effects of restricted population size and selection on changes in value of genetic parameters and the effects peculiar to a particular mode of gene action combined to obscure the prediction problem so that Griffing's theoretical expression was accurate for more than a very few generations only when fortuitous combination of several factors occurred. However, the magnitude of the discrepancies possibly is larger than it would be in a practical situation because of restrictions in the mechanics of simulations".

The objective of the third paper (Gill 1965c) in the series by Gill was to determine the effects of selection intensity, recombination fraction, level of environmental variation, and mode of gene action on

the response to selection in populations of restricted size. The levels of selection intensity were ½, ¼, 1/6 and 1/8 of the offspring saved for parents. The recombination frequencies were 0.005, 0.05, 0.2 and 0.5 and the levels of environmental variation were 0, $1/3\sigma_G^2$, σ_G^2 and $3\sigma_G^2$, where σ_G^2 is the genotypic variance in the initial population. The effects of these parameters were analyzed for the nine genetic models.

In most populations with complete dominance, intense selection (1/6, 1/8) caused a reduction in the amount of fixation of the recessive allele due to finite population size. This implies that restriction of population size may do little damage to the total potential response to intense selection. In populations with overdominance, selection caused a decrease in fixation in all cases when compared to populations where selection was for homozygous maximums. The results of Gill's work agreed with the findings of Bohidar (1960), who concluded that "dominance makes selection sensitive longer". A negative regression of the genotypic mean on generation number existed when selection was weak, environmental variation was present, and population size small. The regression of genetic merit on generation number was negative for all parameter sets in the overdominance model. The general conclusion was that selection is rather effective in advancing the genetic mean of populations under all models of gene action in which the genotype of the highest merit is homozygous, while selection is weaker than random drift in small populations under mass selection for a character that involves only heterozygous genotypes as optimum. Large amounts of environmental variation caused a decrease in response to selection in the additive model, while smaller amounts affected the rate of improve-

ment in the complete dominance model. These findings are similar to those of Martin and Cockerham (1960). With overdominance, various amounts of environmental variation did not begin to affect the change in the mean until several generations of random drift had caused gene frequencies to deviate from 0.5 which considerably increased the proportion of additive variance. There were no significant differences in the mean due to linkage in populations under the additive or complete dominance models. The mean proportion of loci fixed, and the mean gene frequency, were essentially unaffected by different levels of linkage.

Barker and Butcher (1966) presented the results of a simulation which was designed to test the theory of quasi-fixation as developed by Kimura (1954, 1962). This was the first Monte Carlo simulation ever done which was designed to test a specific hypothesis. Kimura (1954) investigated the process of the change in gene frequency distribution due to random fluctuations of selection intensities. Assuming a pair of alleles lacking dominance, and the selection coefficient (s) fluctuating from generation to generation around a mean value of zero, where s equals the selection coefficient against the heterozygote (i.e., heterozygote fitness = 1-s, disadvantagous homozygote fitness = 1-2s), he showed that the gene frequency distribution will accumulate toward the terminal values of 0 or 1 with increasing time. Given sufficient generations, almost all populations will have the gene either almost fixed in the population or almost lost from it. Kimura described this phenomenon as quasi-fixation and quasi-loss, respectively. The results presented from this simulation were in the interest of confirming Kimura's theoretical analysis, and providing information on the time required for the quasi-fixation process.

Selection in the program was determined by sampling a selection coefficient from the uniform interval $(\bar{s} - (3V_s)^{\frac{1}{2}}, \bar{s} + (3V_s)^{\frac{1}{2}})$, so that the mean and variance of the selection coefficient would agree with a given $\bar{s}$ and V_s. The initial gene frequency was variable in each of 1000 populations, each of which was assumed infinitely large. The number of populations whose gene frequency was in each of 190 specified gene frequency ranges were printed out at specified generation intervals. The gene frequency ranges were wider for intermediate values of gene frequency than for values nearer the ends since primary interest was in what was happening at the ends.

In general, the results confirmed the theoretical expectations of Kimura (1962), but the observed probabilities of quasi-loss of the desirable allele were somewhat higher than the expected in nearly all cases. The probability of quasi-fixation was higher in smaller populations (5000) than in larger populations (50,000) for a given number of generations, all else being equal. The number of generations to final stability of quasi-loss tended to increase with the quantity $k = 2\bar{s}/V_s$, and would be expected to be at least 1000 for $0.5 \leq k < 1.0$. There was an apparent, but unexpected, change in the trend for values where $k < 0.5$ as contrasted to those when $k \geq 0.5$.

Fraser et al. (1966) presented results of a new program (GSD-1) which dealt with inversion polymorphism. The program allowed for varying degrees of linkage and dominance. However, in this series an additive model was used. There were only six loci controlling the trait and two alleles per locus. Thus, there were thirteen possible phenotypes, the phenotype being the digital sum of the six loci and ranging from 0 to 12. Selection was on the basis that all individuals with a phenotype of six

had equal probability of being selected as parents while all others were rejected. Selection was of the symmetrical double truncation type. The population size in all runs was 1024 parents. Pairs of individuals were drawn at random from the set of parents and mated to produce one offspring. This process was repeated until 1024 individuals with phenotypes of six had been produced. Four runs were made to determine the effects of selection on the six locus system. The initial gene frequencies were set equal, i.e., $q_1 = 0.5$, in linkage equilibrium. Recombination was set at 0.25 between adjacent loci.

Selection for the intermediate phenotype clearly favored chromosomes with equal numbers of 0 and 1 alleles since these had the highest probability of combining with another chromosome to produce the required phenotype. All other classes of chromosomes were eventually eliminated. When linkage was tightened to a recombination frequency of 0.00001, the same results were observed as when the recombinations were more frequent. However, the results which took from 20 to 40 generations to attain when recombination was 0.25 were attained by the sixth generation when linkage was essentially complete. Thus, freer recombination greatly decreased the effectiveness of this particular mode of selection. These results point out, however, that whether or not linkage is complete, it is possible to maintain a polymorphic equilibrium in which the heterosis observed is not necessarily due to overdominance. Although intermediate selection imposes overdominance on the primary scale, this is an algebraic artefact due to the mode of definition of fitness. Inversion polymorphism does not necessarily involve overdominance and, therefore, the occurrence of inversion polymorphism is not diagnostic of the occurrence of overdominance.

Parker (1966) did a Monte Carlo simulation to investigate the effects of truncation selection on estimates of genetic correlation. Selection was on one of two quantitative traits and the correlated response in the other trait was measured in each generation. The population was bisexual diploid and the traits were expressed in both sexes. There were forty-eight parents, twenty-four of each sex, each generation, but the number of offspring varied depending on the intensity of the selection for a particular run. Each trait was controlled by forty-eight loci segregating independently, effects were equal at each locus, and gene frequency was arbitrarily set at 0.5 at each locus in the initial generation. The correlation between traits was established by pleiotropic effects of genes which affected both traits. Selection was continued for thirty generations. The design of the experiment was a 3^4 factorial with two replications of each of the eighty-one parameter combinations. The four factors Parker used as sources of variation were selection intensity, genetic correlation, heritability of the primary trait, and heritability of the secondary trait. There was a high, intermediate and low level of each. Two models were investigated; one was additive and the other complete dominance.

In the additive model, the genetic correlation, measured as the correlation of genotypic values in each offspring generation, remained consistently near its initial level at all levels of environment, when the fraction of offspring saved as parents was either the largest or intermediate fraction. At the high level of selection intensity, the genetic correlation decreased. At the low level of selection intensity the genetic covariance was maintained. With greater selection intensity the genetic covariance decreased, but the genetic variances of the

traits decreased proportionately, thus, the genetic correlation was maintained. Truncation selection caused a decline in the genetic correlation in those offspring selected to become parents of the next generation. The amount of reduction depended on the heritability of the selected trait rather than on the intensity of the selection. The correlated response of the unselected trait to selection on the primary trait agreed closely with the expected response.

In the model of complete dominance, the change in the genetic correlation followed essentially the same pattern as in the additive model when selection was by upper truncation. When selection was by lower truncation, however, the decrease in the genetic correlation at high selection intensity was more rapid. The response of the genotypic mean of the unselected trait to selection of the primary trait in opposite directions was quite symmetrical for the first few generations, but became asymmetrical in later generations. At low levels of selection the response was fairly linear, but became distinctly curvilinear at high selection intensities.

Young (1966) presented results of a Monte Carlo study, the first of what is to be a series, on the rate of decay of the additive genetic variance due to selection, and the resulting change in heritability. He also compared the accuracy of predicted response to selection by equations which assumed constant heritability and selection differential with that of equations in which heritability and selection differential were recalculated every generation. His program simulated 1000 offspring each generation, for all practical purposes an infinite population. The trait being selected was controlled by ten loci, with two alleles per locus, and the initial gene frequency at each locus was 0.5. The design

of the experiment was a 3^3 factorial with the experimental variables being selection intensity, heritability value and tightness of linkage. The selection intensities were 80%, 50% and 10% of the offspring population saved for parents. All selection was by upper truncation. High, intermediate and low values of heritability were simulated; 0.9, 0.4, 0.1, respectively. The population was initially in linkage equilibrium and the recombination values were 0.5, 0.2 and 0.05 between adjacent loci, all of which were considered to be part of a single recombinatorial unit. Selection was continued for thirty generations for all parameter sets in each of seven genetic models, making a total of 189 populations.

In this paper only the analyses of the additive and complete dominance models were presented. In the additive model the predicted and realized responses were in closer agreement, in the early generations, when the initial heritability was high than when it was low, for the case where heritability was taken as a constant. However, in either case the expected was far in excess of the simulated response after the tenth generation. The discrepancy between the realized and predicted response began to occur later in the populations with high heritability, but became larger than the discrepancy in the populations with low heritability by the sixth generation. When expected response was calculated on the basis of the preceeding generation heritability, the agreement between the realized and predicted was close in all cases. Prediction was slightly less accurate for the combination of high selection intensity and low heritability. Linkage had no affect on the accuracy of the predicted response or the rate of response. The rate of decay of the additive genetic variance was most rapid under conditions of high selection intensity and high heritability, being similar to a

negative exponential, but for lower values of either parameter the decay curve approached linearity, and the slope became less steep.

Compared to the additive model, the agreement between estimated and realized advances were not as good in the complete dominance model. Under both models, high selection intensity coupled with high heritability tended to overestimate genetic advances, while the combination of low values tended to underestimate it. This trend was especially pronounced in the complete dominance model and particularly large for the overestimate, i.e., the combination of high selection intensity and high heritability. Prediction, in general, was more erratic in the dominance model. Again tightness of linkage had no affect on prediction accuracy or response rate. No undesired alleles were fixed in any of the populations in either genetic model. In the dominance model the overall rate of decay of the additive genetic variance was less rapid than in the additive model. However, a fast initial decay followed by a slow elimination of the remaining portion of the additive variance was generally characteristic of the model. The additive genetic variance decreased more rapidly than the dominance variance in all populations.

Estimates of Genetic Parameters

Heritabilities of Traits Studied

Daily gain has generally been shown to be at least moderately heritable. However, Cockerham (1952), on the basis of 1,980 litters, reported a value of 0.07 when doubling the intra-sire regression of offspring on dam. Also, Reddy *et al.* (1959) estimated the heritability of daily gain to 200 pounds as 0.04, essentially zero. Estimates ob-

tained by Lush (1936), Smith and Donald (1937), Whatley and Nelson (1942), Hazel et al. (1943), Krider et al. (1946), Blunn et al. (1953), Whatley (1956), Craig et al. (1956) and Reddy et al. (1959) range from 0.16 to 0.25. The average of these estimates is very near 0.2. Other workers, however, have indicated that the heritability of daily gain is probably higher than 0.2. Baker et al. (1943), Dickerson and Hazel (1944), Dickerson (1947), Johansson and Korkman (1950), Issawi and Rempel (1961), Park et al. (1963) and Omtvedt et al. (1963) report values ranging from 0.26 to 0.34 which average about 0.3. Still other workers have estimated the heritability of daily gain even higher. The range of the estimates presented by Dickerson and Grimes (1947), Cox (1959b), Brinks (1960), Dillard et al. (1962), Biswas et al. (1963) and Louca and Robison (1965) is from 0.38 to 0.81, indicating the heritability of growth rate in some populations may be quite high. The average of all estimates is about 0.26 when they are weighted according to the numbers of animals on experiment.

Similar to the findings for daily gain, the heritability of feed efficiency (gain/feed) has generally been considered to be at least moderate. The estimates of the heritability of feed efficiency are fewer than those for daily gain. When considering the difficulties involved when estimating the heritability of feed efficiency, this is not surprising. Contrary to the findings for daily gain, no estimates have been presented which would indicate that feed efficiency may be non-heritable in some populations. The estimates of feed efficiency heritability were surprisingly uniform. The estimates presented by Dickerson and Grimes (1947), Whatley (1956), Craft (1958), Park et al. (1963) and Lasley (1964) all fell within the narrow range of 0.26 to

0.31. Only two workers (Dickerson 1947 and Biswas et al. 1963) presented results out of this range, 0.57 and 0.58, respectively. The average of all estimates was 0.34.

Litter size has long been known to be only slightly heritable. The estimates of the heritability of litter size, pigs farrowed alive, range from a negative value, -.15, (Cockerham 1952) to a high of 0.32 (Cummings et al. 1947). Negative estimates can generally be attributed to inbreeding or sampling error. Stewart (1945) reported an estimate of 0.14, Blunn and Baker (1949) estimated it to be 0.22, while Whatley's (1956) estimate was 0.05 and Omtvedt et al. (1963) reported an estimate of 0.10. In general, all estimates were low and the average of all estimates was 0.14.

Considerable work has been done in attempting to reduce carcass fatness in swine by selecting for thinner backfat. As a result several estimates of backfat heritability have been made and most all of these have been consistently high. There are, however, at least three exceptions. Reddy et al. (1959), Hetzer (1963) and Louca and Robison (1965) have all reported estimates lower than 0.4. Most estimates of backfat heritability have been between 0.4 and 0.5. Those estimating backfat thickness within this range include: Lush (1936), Dickerson (1947), Stothard (1947), Johansson and Korkman (1950), Whatley (1956), Hetzer and Zeller (1956), Reddy et al. (1959), Cox (1959b), Dillard et al. (1962), Omtvedt et al. (1963) and Gray et al. (1964). The only two estimates above this range were by Enfield and Whatley (1961) and Zoellner et al. (1963) with estimates of 0.63 and 0.75, respectively. The average of all estimates is 0.49.

Phenotypic Variances of Traits Studied

Several investigators have estimated the phenotypic variation of daily gain in experimental swine populations. There is considerable range in these estimates. The lowest value was reported by Blunn et al. (1953), 0.0108, while the highest was reported by Reimer et al. (1958), 0.0551. Dickerson (1947), England and Winters (1953), Dickerson et al. (1954), Reddy et al. (1959), Brinks (1960), Zoellner et al. (1963), Vogt, et al. (1963) have all estimated the phenotypic variance of daily gain to be intermediate to these values. The average of all estimates reviewed was 0.0289.

Estimates of the phenotypic variance of feed efficiency are seemingly rare. Only three estimates were found in the literature. Dickerson (1947) estimated the phenotypic variance to be 0.1398 $feed^2/gain^2$ while Vogt et al. (1963) estimated it to be 0.0482 $feed^2/gain^2$ and Biswas et al. (1963) reported an estimate of 0.0003 $gain^2/feed^2$. Assuming the range of the normal curve to be five standard deviations, the estimates of the variance which were in terms of feed/gain were converted to units of gain/feed by inverting the limits of the range and then dividing the new range by five. This gives a converted estimate of the standard deviations and the square of that quantity gives the converted estimate of the variance. The average of the estimates when all were converted to the same units was 0.00072 $gain^2/feed^2$.

Estimates of the phenotypic variance in litter size are quite consistent with the exception of the estimate by Lush and Molln (1942) of 1.0. The estimates of Dickerson et al. (1954), Reddy et al. (1958) and Vogt et al. (1963) were 4.58, 5.66 and 5.47, respectively.

The range in estimates of the phenotypic variance of backfat thick-

ness is not too large. The smallest estimate was presented by Zoellner et al. (1963), 0.0113, while the largest estimate was presented by Brinks (1960), 0.0386. The estimates presented by Dickerson (1947), Lerner et al. (1957), Reddy et al. (1959), and Cox (1959a and b) all fell within this range. All estimates were made using hogs that weighed between 200 and 225 pounds. The average of all seven estimates is 0.0196.

Estimates of Genetic Correlations Between Traits

Daily Gain and Backfat Probe. As with other genetic parameter estimates, there is considerable sampling error involved in estimating genetic correlation as is reflected in large ranges in these estimates. The lowest estimate of the correlation between daily gain and backfat was -.55 (Louca and Robison 1965), while the highest was 1.34 (Dickerson 1947). Two other workers, besides Louca and Robison, have estimated the correlation as being negative; Freeden and Jonsson (1957) and Jonsson and King (1962), -.18 and -.26, respectively, both in Danish Landrace swine. The majority of workers, however, have reported moderately high positive estimates. Depape (1954), Cox (1959b), Brinks (1960), Biswas et al. (1963), Zoellner et al. (1963) all report values between 0.53 and 0.70. Only one other estimate was out of this range, 0.12 (Fox 1959). The arithmetic average of these estimates was 0.40. When these values are converted to z (Fisher 1925), averaged, and then converted back to a correlation, the average is 0.29.

Daily Gain and Feed Efficiency. Estimates of the genetic correlation between daily gain and feed efficiency are consistently high, and positive when feed efficiency is expressed as gain/feed. Only three

estimates are below 0.5. Biswas et al. (1963), Vogt et al. (1963) and Zoellner (1963) estimated the genetic correlation to be 0.32, 0.22 and 0.27, respectively. Whereas, Freeden and Jonsson (1957), Reimer et al. (1958), Brinks (1960), Jonsson and King (1962) and Park et al. (1963) have all estimated the correlation to be higher than 0.5. The arithmetic average of all estimates is 0.63. The transformed average is 0.66.

Daily Gain and Litter Size. There has been very little work in which estimates of the genetic correlation between daily gain and litter size have been made. The estimates which have been reported are 0.11 (Reddy et al. 1958), -.30 (Dillard et al. 1962), and 0.06 (Vogt et al. 1963). The average of these estimates is -.04, or essentially zero.

Backfat Probe and Feed Efficiency. The range in estimates of the genetic correlation between backfat and feed efficiency is exceedingly large, -.92 (Biswas et al. 1963) to 0.73 (Brinks 1960). Besides there being a wide range in the estimates, there is no apparent trend toward either a positive or negative value. The average of the estimates presented by Dickerson (1947), Freeden (1953), Anderson (1954), Freeden and Jonsson (1957), Fox (1957), Jonsson and King (1962), Zoellner et al. (1963) is 0.03, while the average of the transformed correlations is -.02.

Backfat Probe and Litter Size. Only three estimates of the genetic correlation between backfat and litter size have been reported. Lerner et al. (1957) reported an estimate of the genetic correlation between backfat thickness and number of embryos recovered as 0.31. However, there was a significant correlation (0.32) between age of dam at breeding

and number of embryos recovered. Since gilts get fatter with age, this apparent relationship of litter size and backfat could probably be explained as an age effect rather than a fatness effect. Reddy et al. (1958), on the other hand, found the genetic correlation to be negative, -.14, between backfat and number of embryos recovered when days of gestation were held constant at 55 days. Dillard et al. (1962) reported an estimate of 0.09 as the genetic correlation between backfat probe at 140 lbs. and litter size. The average of these estimates is 0.08, essentially zero.

Feed Efficiency and Litter Size. No estimates of the genetic correlation between litter size and feed efficiency could be found in the literature.

OBJECTIVES

The primary objective of this study was to evaluate a proposed breeding plan as to its effectiveness in improving the performance of crossbred swine in five generations of selection, when selection is entirely within pure breeds. To achieve this goal two pure breeds were simulated and a cross of the two was made before the first and after the fifth generations of selection, in order to measure the improvement made in the crossbred pigs. Selection was on four traits; two within each breed. The selected traits were the four major production traits in swine; daily gain, feed efficiency, litter size, and backfat probe. The first breed was selected for daily gain and feed efficiency, while the second was selected for litter size and backfat probe.

In addition to evaluating the breeding plan, there were several other objectives in the study. The accuracy of prediction equations in estimating the change in the mean and the change in gene frequency as a response to selection was of particular interest. The rate at which inbreeding accrues, its depressing effects, and the agreement between observed and predicted inbreeding were of interest since selection was strong and the number of parents in both breeds was small. The rate and amount of decline in the genetic variance under such a selection scheme was also considered important since it effects the heritability and selection differential and, as a direct consequence, the rate of improvement in the population. The stability of the genotypic correlations in the presence of inbreeding and selection was also to be evaluated.

METHODS AND PROCEDURE

Initial Genetic Population Parameters

Each trait was affected by genes at twenty independently segregating loci, all with equal effects. No linkage was simulated since it was the objective here to simulate a real swine population and no estimates of the tightness of any existing linkage were found. Hence, although there very probably is some linkage present, in varying degrees, in real swine populations, it was decided to follow the traditional scientific approach of using the simplest hypothesis possible, which may be considered to produce realistic results. By the same reasoning no epistasis was included. There very likely is epistasis present in real swine populations, but since its mode, or modes, of action are unknown, it was felt that it would be as realistic to ignore it as it would be to simulate any particular mode merely for the sake of including some epistasis. From existing Monte Carlo results, it seems that linkage effects in natural populations are probably negligible. Fraser (1957b), Martin and Cockerham (1960), and Qureshi (1964) all found that linkage was not effective unless it was very tight ($r \leq 0.025$), while Baker and Comstock (1961), Gill (1965c) and Young (1966) found no effect due to linkage at any level. If epistasis were present, it would probably have the effect of reducing the response to selection.

The modes of gene action for the traits studied were kept as simple as possible. Daily gain, feed efficiency and litter size were all under a complete dominance model, while backfat probe was under an additive model. The criterion for determining what the mode of gene action for a particular trait should be, was whether or not the trait

generally shows inbreeding depression or heterosis. If the trait generally shows either it was assumed to have a complete dominance mode of gene action, since this is the simplest mode which can account for such behaviour. If the trait generally does not show either, an additive model was assumed.

From the literature review it appears that only two pairs of the four traits showed a consistent genetic correlation. These two pairs were daily gain with feed efficiency and daily gain with backfat probe. The estimates from the literature for these correlations were 0.4 between daily gain and backfat and about 0.6 between daily gain and feed efficiency, when feed efficiency is measured as gain/feed. Although the estimates of the genetic correlation between backfat and feed efficiency were high in some data, they were not consistent, and averaged nearly zero. Therefore, the correlation between these two traits was set to zero in this program. The genetic correlation between daily gain and feed efficiency was set to 0.6. This was achieved by having twelve of the twenty loci affecting each trait in common. The genetic correlation between backfat probe and daily gain was set to 0.4 by having eight of the twenty loci affecting each trait in common. The genetic correlation between backfat probe and daily gain was set at 0.4, since the average of the estimates in the literature was 0.45 when the estimates from Danish Landrace swine were deleted. It was felt these should be deleted since that particular breed of swine has been under intense selection for many generations, and at this time is probably not very similar, genetically, to most of the more popular American breeds of swine. This dissimilarity is expressed by the fact that the estimates which were made involving Danish Landrace swine were negative

while nearly all estimates involving American breeds were moderately to highly positive. All other genetic correlations were set equal to zero.

The initial heritabilities of the traits were established so as to be consistent with the estimates from the literature. When the estimates from the literature review were rounded to the nearest tenth, the estimates of the heritabilities of the traits were 0.3, 0.3, 0.1, and 0.5 for daily gain, feed efficiency, litter size, and backfat probe, respectively. These were the values which were programmed into the computer for the initial generation. No adjustments were made thereafter to attempt to keep the heritabilities constant.

As was the case with the heritabilities, the initial phenotypic variances were established so they would be consistent with the estimates in the literature. The phenotypic variances in the initial generation of each run were programmed to be 0.0289, 0.00072, 5.30 and 0.0196 for daily gain, feed efficiency, litter size and backfat probe, respectively. No adjustments were made in the phenotypic variances after the initial generation.

The initial gene frequency for each trait was 0.6 for the desired gene, except for backfat for which it was 0.4. The initial gene frequency was determined arbitrarily by the author. The rationale for these initial frequencies was that it would seem that most swine populations have been selected to some extent at some time or other, for most, or at least some, of these traits. Consequently, it seems reasonable to assume that the frequency of the desired gene may very well be above 0.5 (the standard initial gene frequency in most Monte Carlo simulations). The reason for the frequency of the desired gene for backfat probe being 0.4 rather than 0.6 is due to a mechanical re-

striction imposed by the genetic correlation. Since the genetic correlation between daily gain and backfat probe was determined to be positive, the genes with pleiotropic effects must affect both traits in the same direction. Thus, the pleiotropic effects which increase daily gain must also increase backfat probe. Since the frequency of the desired gene for daily gain is 0.6, the frequency of the gene with pleiotropic effects which increases backfat probe must also be 0.6. However, increasing backfat is undesirable and, therefore, the frequency of the desired gene for backfat probe is 0.4. An alternative to this method, which would increase the frequency of the desired gene for backfat, would be to reverse the effects of those genes for backfat which have no plieotropic effects on daily gain. If this were done the frequency of the desired gene would be 0.52 rather than 0.4, since in 0.4 of the genes the frequency of the desired gene would be 0.4, but in the other 0.6 of the loci the frequency would be 0.6. This could be taken one step farther and instead of merely reversing the frequencies of the respective genes at the remaining 0.6 of the loci, the frequency of the desired gene would be set to be 0.7. This would cause the overall frequency of the desired gene for backfat probe to be 0.58, very close to 0.6. Neither of these alternatives was chosen, however, since it was felt it would not be worth the extra programming required to accomplish them in view of the fact that the initial frequency of the desired genes in all cases was arbitrary, and there is no substantial reason to assume one initial frequency is more realistic than another.

The initial means of the population for the various traits were established arbitrarily also, to be a value which seemed reasonable. The initial mean for daily gain was 1.6 lb./day, it was 0.31 gain/feed

for feed efficiency, 8.0 pigs per litter for litter size, and 1.6 inches for backfat probe. They were the same in both breeds as were all other initial population parameters. The accuracy of the estimates of the means in the base generation was not of any particular concern, since it has no effect on the net progress made in the cross or in the pure breeds, and thereby has no effect on the accuracy of the evaluation of the breeding plan.

The Breeding Plan

The goal of the breeding plan was to improve the performance of crossbred pigs, such as those produced by commercial hog producers. As stated previously, the breeding plan involved selection in two separate breeds. The traits selected for, as well as the mode of their selection, were different within each breed and, therefore, each breed will be discussed separately. The breeding plan was designed to be a realistic one which most any purebred swine breeder could follow, with the usual facilities and equipment.

In the first breed selection was for feed efficiency and daily gain. In the initial generation an individual feeding trial was simulated for each of 20 randomly generated boars and the 2 most efficient of the 20 were selected for feed efficiency. Thirty gilts were randomly generated and the 10 gilts with the greatest daily gains were selected. The expected selection differentials in terms of standard deviations were 1.8 for the boars and 1.1 for the gilts. Each gilt bred produced 5 offspring; 3 gilts and 2 boars. This is analogous to the situation in a real trial where 3 gilts and 2 boars are randomly selected from each litter produced to be performance tested, since those actually produced from a

given mating are a random sample of those that could have been produced. In the succeeding generations the simulated performance tests and selection of parents were repeated.

The selection in the second breed was more complicated than that in the first. In this breed selection was for litter size and backfat probe. In the base generation 48 gilts and 24 boars were randomly generated according to the conditions presented in the preceeding section. Of the 48 gilts the 20 leanest were selected and bred twice to the 4 leanest boars. However, only offspring of the 12 most prolific gilts were actually produced since in the succeeding generations only offspring from the females with the largest average litter size were used. Litter size was determined solely by the dam. There was a restriction on the boar selection that only one boar from a litter could be selected. Each mating produced 3 boars and 4 gilts. Therefore, each generation 36 boars and 48 gilts were produced. The gilt selection cycle was then repeated by selecting the 20 leanest gilts from the 48 gilts which were first litter offspring of the 12 gilts with the largest average litter size the previous generation. The boar selection cycle was repeated by selecting the 4 leanest boars from the 24 boars which were first litter offspring of the 8 gilts which had the largest average litter size the preceeding generation. The expected selection differentials in terms of standard deviations were 1.51 and 0.93 for backfat for the boars and gilts, respectively. Selection for improved litter size was on pedigree since the offspring of the gilts with the largest average litter size were used for breeding the next generation. The expected selection differentials for litter size were 0.483 and 0.322 standard deviations above the mean for the boars and

gilts, respectively.

Selection was continued for five generations in each breed. As stated previously a cross was made before the first and after the fifth generations of selection, in order to determine the progress selection within the pure breeds made toward improving the performance of the crossbred swine. The parents of these crosses were not selected randomly, but were selected for their respective traits. It was felt this would serve to reduce the sampling error in selecting parents and thereby give a more precise estimate of the improvement made in the cross. In both the first and second cross, 20 gilts from the second breed were selected from 48 on backfat probe, and 4 boars from the first breed were selected from 20 on feed efficiency. Each mating produced 5 offspring and all 100 offspring were evaluated. It was assumed there was no sex effect in any of the traits studied.

Mechanics of the Simulation

In this section the computer, random number generator, mechanics required to accomplish the desired population parameters and the programming logic will be described. No attempt, however, will be made to present the program in detail, since it is lengthy, complex, and quite specific to this particular study.

The computer available for this study was Michigan State University's Control Data 3600. At the time this simulation was made the computer had a single storage module of random-access, magnetic core storage of 32,768 words. Not all of this was available to the user since the system's monitor was contained in the storage module during execution of the user's program. The monitor occupied some 12,000 words of memory

leaving the user approximately 20,000 words. Each word of memory in the CDC 3600 is a 51 bit word. Of these 51 bits, 48 are data bits and 3 are parity bits. The 3600 is considered a very fast computer, being capable of 500,000 additions per second; some ten times faster than the Silliac on which Fraser (1957a) did his original simulation work. Since the computer is exceedingly fast and computer time was available, little effort was made to make the program especially efficient.

The random number generator used was one that was available through MSU's computer systems library, a function called RANF. This function uses the multiplicative congruential method to generate uniformly distributed, pseudo-random numbers. The generator will produce numbers in either fixed or floating point format. If floating point mode is used, the numbers generated range from zero to slightly less than one.

The original multiplicative congruential method was of the form

$$X_{i+1} = X_i \lambda \qquad (\text{mod } m)$$

where X_i is any odd number.

$$\lambda = 5^2 \times 15$$

$$m = 2^{47}$$

This parameter set has been shown to satisfy the conditions for a maximal period by Rotenberg (1960). The period for this generator in the CDC 3600 would be 2^{45}. This procedure was modified by Rotenberg (1960) to the form

$$X_{i+1} = (2^a + 1)\, X_i + c \text{ with } a \geq 2 \text{ and } c \text{ odd.}$$

In RANF $\underline{a}$ was set equal to 10 and c equal to 101, making $\lambda = 2^{10} + 1 = 1025$ and $X_{i+1} = 1025\, X_i + 101$.

The method used in RANF has passed many tests of randomness including a test of the frequency distribution of the random numbers, a

test of the frequency with which a number of certain magnitude was followed by a number of another certain magnitude, a test of the frequency distribution of the length of runs of numbers either above or below the mean, serial correlation tests, and others.

The genic variance was computed by multiplying the phenotypic variance by heritability, the phenotypic variance being that which was obtained from the literature review. The desired genic variance was accomplished by using the equations presented by Lush (1948). Lush presented equations for the computation of the total genetic variance at a single locus, and for partitioning it into components of genic and dominance variance. The equation for computing the total genetic variance at a single locus is of the form

$$\sigma_H^2 = 2q(1-q)(1+F) + 2q(1-q)(1-F)(K-1)\left[2(1-2q) + (K+1)\left[1-2q(1-q)(1-F)\right]\right]X^2$$

where q is the frequency of the desired gene

F is the inbreeding present in the population

K is the degree of dominance; 1 being no dominance and 2 complete dominance

and X is half the difference between the homozygous extremes. This equation simplifies to $\sigma_H^2 = 4q(1-q)^2(2-q)X^2$ when dominance is complete and there is no inbreeding. The equation for partitioning out the genic variance is

$$\sigma_G^2 = \frac{2q(1-q)}{1+F}\left[1 + F + (1-F)(K-1)(1-2q)\right]^2 X^2$$

which simplifies to $\sigma_G^2 = 8q(1-q)(q-1)^2X^2$ when there is no inbreeding and dominance is complete.

The value of X was determined for each trait which was controlled by a complete dominance model by substituting the appropriate gene fre-

quency (0.6 in each case) into the equation for the genic variance, presented above, setting it equal to the genic variance at a single locus and solving for X. By setting the difference between the homozygous extremes at 2X, the desired additive genetic variance for each trait was achieved. In the case of backfat, the equation for genic variance simplifies to $\sigma_G^2 = 2q\,(1-q)\,X^2$ for the case of no dominance and no inbreeding. This form of the equation was used to determine the value of X for backfat probe. In the case of no dominance the genic variance is equal to the total genetic variance.

The difference between the total genetic variance and the phenotypic variance was assumed to be environmental variance. The amount of environmental variance needed to produce the desired phenotypic variance was accomplished by multiplying the square root of the difference between the total genetic variance and the desired phenotypic variance by a Gaussian deviate. Thus producing the desired increase in variance, but not affecting the mean of the population. As a direct result of establishing the desired genic and phenotypic variances, the desired initial heritabilities were also set.

The random normal deviates which were utilized were the 100,000 Gaussian deviates published by the Rand Corporation (1955). The deviates were produced by using the equation $(D + 0.5)\ 10^{-5} = F(\chi)$, where D is a five digit random number and $F(\chi) = {}^{1}/\sqrt{2\pi} \int_{-\infty}^{\chi} e^{-t^2/2}\,dt$, and solving for χ. The five digit random numbers employed in the preceeding equation were produced by an electronic roulette wheel.

The deviates were first punched on cards and then put on tape in card image form (BCD). They formed a matrix which was 10,000 x 10. At the beginning of each run a five digit random number was read into the

program which caused the computer to skip as many rows before the reading of deviates commenced. Also at the start of each run, a five digit random number was input which caused the random number generator to make a random start. A random number from 1 to 10 was then generated and thereby a random column was selected for reading. Once the reading began, the deviates were read sequentially until an end of file was encountered. At this point the randomization procedure was repeated and another random start in the tape was made.

The desired initial gene frequency was achieved by subtracting 0.6 from the number generated by the random number generator. If the difference was positive a mutant allele (0) was entered into the animals genotype. If the difference was zero or negative a 'plus' (1) gene was entered into the animals genotype. Thus the initial gene frequency for all genes which caused a quantitative increase in performance was set at 0.6. For all traits except backfat probe this was also the frequency of the desired gene. However, since in the case of backfat it was desirable to decrease the amount of fat an animal carried, the initial frequency of the desired gene was 0.4.

It was desirable to keep the genes of a given trait for a particular animal in a single word of memory, so that all the genes of an individual animal could be kept in three consecutive words of memory. Thus each group of 20 genes for feed efficiency, litter size and backfat probe occupied separate, single, consecutive words of memory. The genes for daily gain all had pleiotropic effects and therefore were dispersed among the genes for both feed efficiency and backfat probe. Twelve of the twenty loci which controlled daily gain were in the same word of memory as were the feed efficiency genes, while the other eight loci were common with the backfat genes. Consequently, although the animals each had

twenty loci which affected each of four traits, each individual actually had 120 genes rather than 160 as might at first be thought.

Since each animal was diploid for each trait, but each trait, except daily gain, was to occupy only a single word of memory, it was necessary to represent loci with pairs of genes within a single word rather than to represent the diploid individual with consecutive pairs of haploid words as is often the case in Monte Carlo simulation. To achieve this end two bit codes were entered into each word to represent homozygous desired, heterozygous and homozygous mutant loci. The respective codes for each of the three possible conditions were 10, 00 and 01. By shifting the word 2N-1 times (where N is the locus number) circularly to the left and then loading logical with the the mask 1000....01, each pair of loci could be evaluated merely by checking on whether the word was positive, zero or negative. This could be done conveniently with the FORTRAN three-branch IF statement in those subroutines which were written in 3600 FORTRAN.

Although solving for X, half the distance between the homozygous extremes, established the desired genic variance in the population, it did not fix the initial mean for each of the traits to the desired value. Setting the original population mean for those traits which were affected by the complete dominance model was accomplished by substituting the appropriate values into the following equation, which was presented by Lush (1948), and solving for y. The equation was $\mu = n\left[y + (2q(2-q))X\right]$ where μ is the desired initial mean for a particular trait, n is the number of loci, q the initial gene frequency, X half the distance between the homozygous extremes and y is the contribution made by loci homozygous for the mutant allele. In the case of complete dominance $y = \mu/n - 2q(2-q)X$

The preceeding equation was modified slightly to accomodate the calculation of the $\underline{y}$ value for the additive case. In this case the equation was of the form $\mu = n\left[\underline{y} + 2q\underline{X}\right]$ and the solution for y became $y = \mu/n - 2qX$.

In addition to its other fuctions, the program kept a running account of the inbreeding that accumulated each generation. Inbreeding was measured basically as the percentage decrease in heterozygosity. The number of heterozygous loci was counted in the initial or base generation and any inbreeding which accrued thereafter was measured by the quantity $F = (P' - P)/P'$, when P' is the number of heterozygous loci in the initial generation and P is the number in the generation in which the inbreeding accrued. However, since selection was changing the gene frequency in the population at the same time the inbreeding was occurring, the above estimate would be biased, since a certain amount of the increased homozygosity would be due to the change in the gene frequency rather than inbreeding. Therefore an adjustment factor was included into the equation, which would remove the effects of changing gene frequency. The adjustment factor was the quantity $2q'(1-q')/2q(1-q)$, where q' is the frequency of the desired gene in the initial generation and q is the frequency of the desired gene in the generation in which the inbreeding occurred. This adjustment factor was multiplied by the number of heterozygous loci in a particular generation to produce an adjusted number of heterozygous loci (P_A). Mathematically expressed, the adjusted number of heterozygous loci was $P_A = \left[2q'(1-q')/2q(1-q)\right] P$. The equation for inbreeding now becomes $F = (P' - P_A)/P'$, which gives an unbiased estimate of the inbreeding which accrues in the presence of changing gene frequency.

The following proof is offered to show the adjustment factor does

indeed yield unbiased estimates of the inbreeding when gene frequency is or is not changing. The basic assumption for this proof is that the number of loci counted in any generation is a function of the inbreeding and gene frequency in that generation. Mathematically this assumption may be expressed as $P = (1-F)\ 2q\ (1-q)$. With this assumption the equation for the computed inbreeding takes the form

$$F = \frac{2q'(1-q') - (1-F)\ 2q\ (1-q)\left[2q'(1-q')/2q(1-q)\right]}{2q'(1-q')}$$

which may be reduced to

$$\frac{2q'(1-q') - (1-F)\ 2q'\ (1-q')}{2q'(1-q')}$$

which becomes

$$1 - (1-F)$$

and therefore

$$F = F$$

regardless of the gene frequency in the generation for which the inbreeding is being calculated.

The remainder of this section will deal with the programming logic. The sequence of operations and the primary functions of the various subroutines of the program will be discussed briefly.

The entire program was composed of a main program and fourteen subordinate subroutines. The main program and most of the subroutines were written in 3600 FORTRAN, but the subroutines which manipulated individual bits within a word of memory were written in COMPASS, the 3600 assembly language. The function of the main program was basically one of coordination of the subroutines and bookkeeping type operations. Each of the subroutines performed specific functions which were generally repeated many times during the execution of the program.

The first subroutine's only fuctions were to input the random numbers which cycled the random number generator and positioned the random deviate tape to random starting points, and to set indexes which appeared on the punched card output so one run's output could be distinguished from another. The next subroutine generated the animals for the base generation and established the frequency of the desired genes for each trait. The third subroutine read the random deviates from the tape and kept track of the column from which they were to be read. The fourth subroutine evaluated each animals genotype, added the environmental variance, evaluated the phenotypes and calculated the mean, variance, and inbreeding for each generation. The next subroutine computed the genotypic correlation between the traits for each generation. The sixth subroutine ranked the animals in descending order according to their phenotypes. The seventh subroutine selected the animals to be used as parents and generated the required number of gametes for each animal. The next subroutine randomly combined the gametes of the selected parents. The ninth subroutine imposed the restriction that only one boar per litter could be selected as parents on the boars of the second breed. The tenth subroutine simulated the production of two litters of pigs by the gilts in the second breed. The eleventh subroutine was used by other subroutines to decode individual animals' genotypes. The twelfth subroutine loaded individual bits from words in storage for the eighth subroutine. The thirteenth subroutine shifted bits within words of memory for subroutines written in compiler language. The last subroutine was used for output by subroutines written in assembly language.

RESULTS AND DISCUSSION

The results presented in this section are the averages of twenty independent runs through the computer. Prior to each run the random number generator was cycled a random number of times and a random starting point was determined for the random deviate tape. In each run the base populations for each sex within each breed were randomly generated, a cross was made prior to any selection in the pure breeds, each pure breed was selected for its respective traits for five generations, and a cross was made after each breed had been selected for five generations. The gene frequency, genotypic mean, genotypic variance, phenotypic mean, phenotypic variance, inbreeding, and the genotypic correlation between the traits were calculated for each generation within both the pure breeds and in the cross, and were output on punched cards. These data were then summarized.

The Effectiveness of the Breeding Plan

In table 1 the mean and the frequency of the desired gene are presented for each trait for both generations in which crossbred pigs were produced. The breeding plan proved effective in improving all four traits. The most improvement in the mean was made in backfat probe, 14.6%. This was to be expected since the initial heritability of backfat probe was the highest, selection intensity was greatest for this trait, and the initial gene frequency was less than 0.5. The effect of inbreeding on the within line genetic variance was partially counteracted by the effect of the gene frequency approaching 0.5, since the variance is maximum at that point when there are only additive effects, and therefore, the genetic variance and the heritability of the trait tended to be main-

Table 1. PERFORMANCE OF CROSSBRED PIGS PRODUCED BY CROSSING THE PURE BREEDS BEFORE THE FIRST AND AFTER THE FIFTH GENERATIONS OF SELECTION.

	Generation	Mean	Freq.	% Heterosis
Daily gain	Base	1.61	0.607	0
	Fifth	1.65	0.630	6
Feed eff.	Base	0.317	0.633	0
	Fifth	0.329	0.698	4
Litter size	Base	8.06	0.605	---*
	Fifth	8.14	0.664	---
Backfat	Base	1.57	0.423	0
	Fifth	1.44	0.524	0

*Heterosis cannot exist in litter size since it is a completely maternal trait and the dams of the crossbred pigs are purebred dams.

tained for a longer period of time. The second most improved trait was feed efficiency, 3.8%. This too was as expected since the heritability of feed efficiency and daily gain were the same, but selection was more intense for feed efficiency. The improvement in the mean for daily gain was 2.5%. The least affected trait was litter size. Litter size was improved by 1%. This improvement would have been greater if litter size were not a strictly maternal characteristic. However, since this was the case, the mean litter size was strongly affected by inbreeding depression, because the dams of the crossbred pigs were inbred.

When the improvement in the frequency of the desired gene is examined, the results differ from those of the means. The most improvement is still in backfat probe, 23.9%, and the improvement in feed

efficiency is still second, 10.3%. However, contrary to the above results, when conclusions were based on means, litter size is now third most improved, 9.8%, and daily gain is least improved, 3.8%. There are two reasons for this reversal in the rank of these last two traits. First, the mean for litter size was affected by inbreeding depression, while there was 6% heterosis in the mean for daily gain. The second reason for more improvement in litter size than daily gain, despite stronger selection and a higher heritability of daily gain, is due to the downward selection in breed II for daily gain which resulted as a correlated response to the selection for thinner backfat in breed II. Although the mean for daily gain was also affected by the downward correlated response, heterosis cancelled part of this effect on the mean, but, of course, did not affect gene frequency.

Changes in the Pure Breeds

Since the two breeds were selected for different traits, and by different procedures, the results of the changes which took place within each of the pure breeds will be presented separately.

Breed I

Breed I was selected for improved feed efficiency and daily gain. The changes made in breed I are shown in figure 1.1 through 1.4. Figures 1.1 and 1.2 show the improvements in the means and gene frequencies for daily gain and feed efficiency, respectively. The amount of improvement in each of these traits is very similar, 5.4% and 5.9%, respectively, despite stronger selection for feed efficiency. This may be explained by the fact that initially the frequency of the desired gene for feed efficiency increased more rapidly than did the frequency

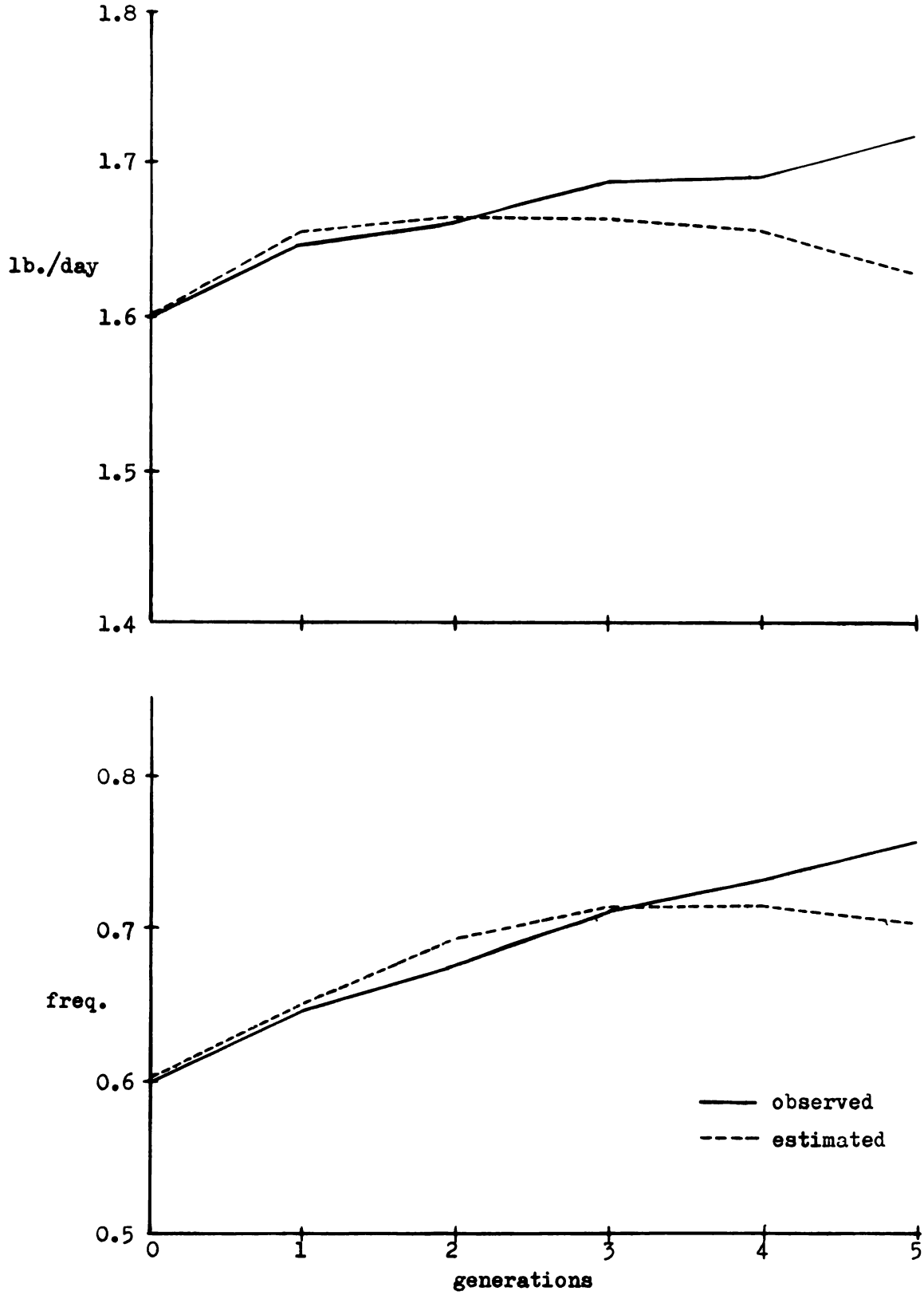


Figure 1.1 Changes in the mean and gene frequency for daily gain within breed I.

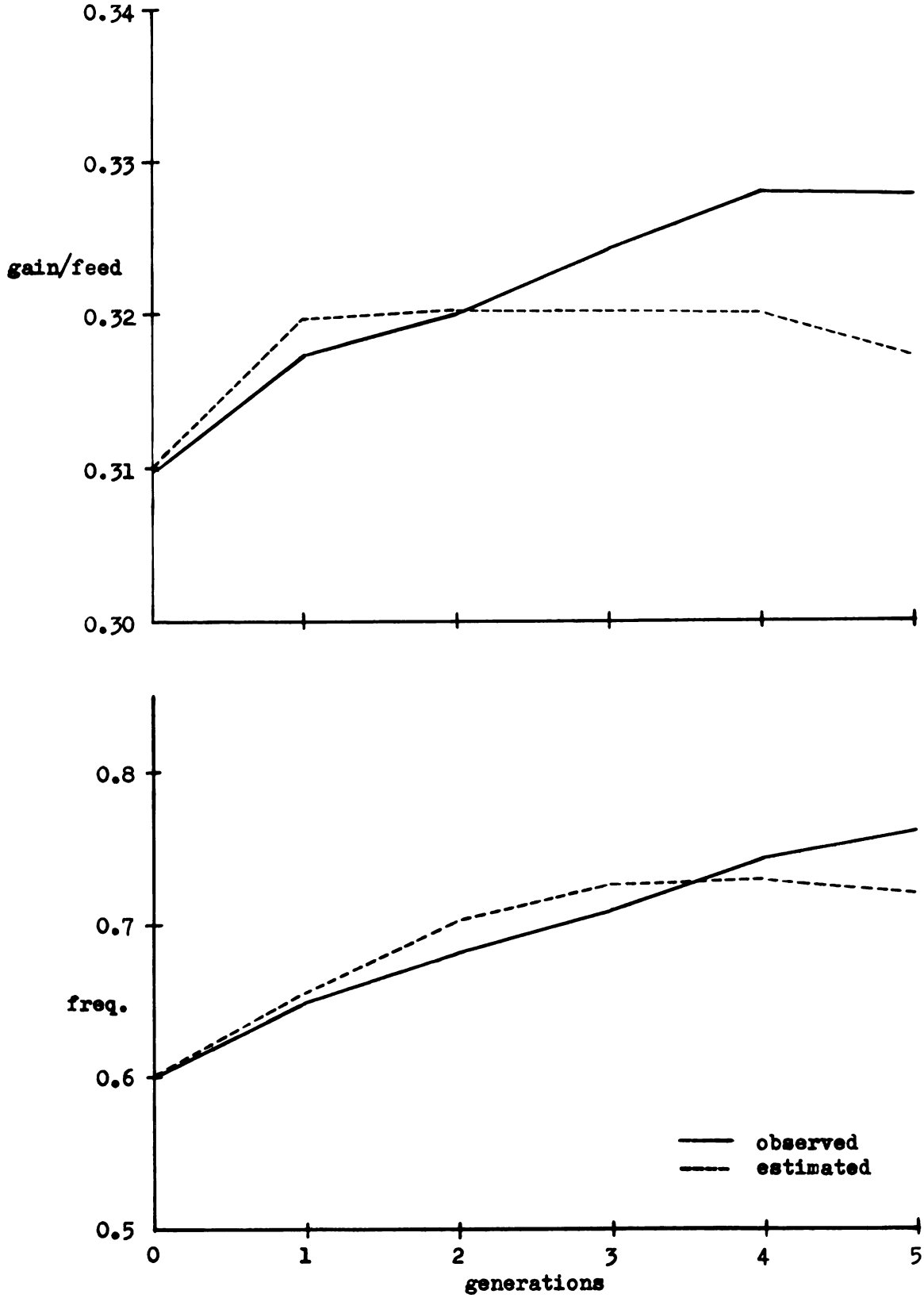


Figure 1.2 Changes in the mean and gene frequency for feed efficiency within breed I.

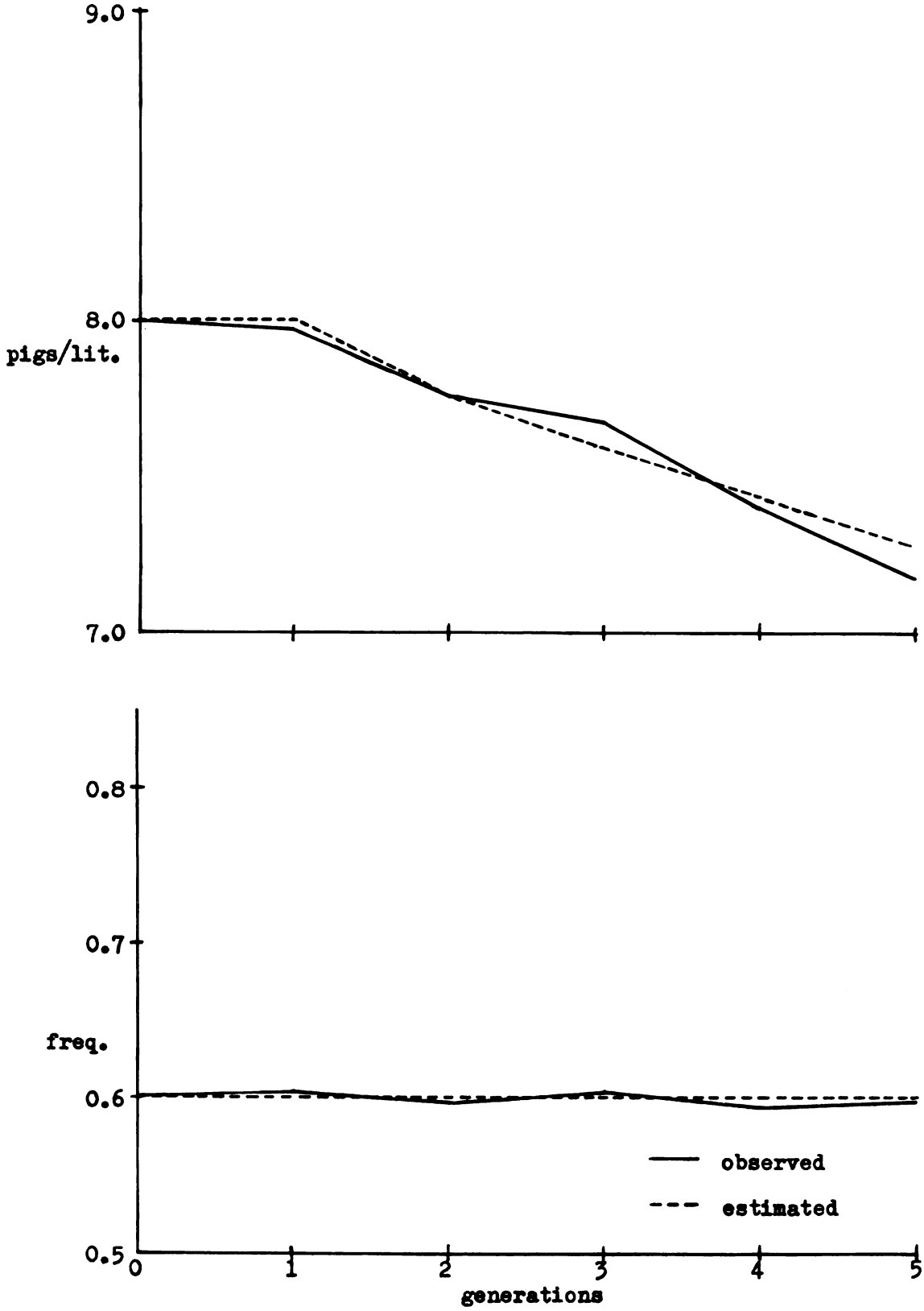


Figure 1.3 Changes in the mean and gene frequency for litter size within breed I.

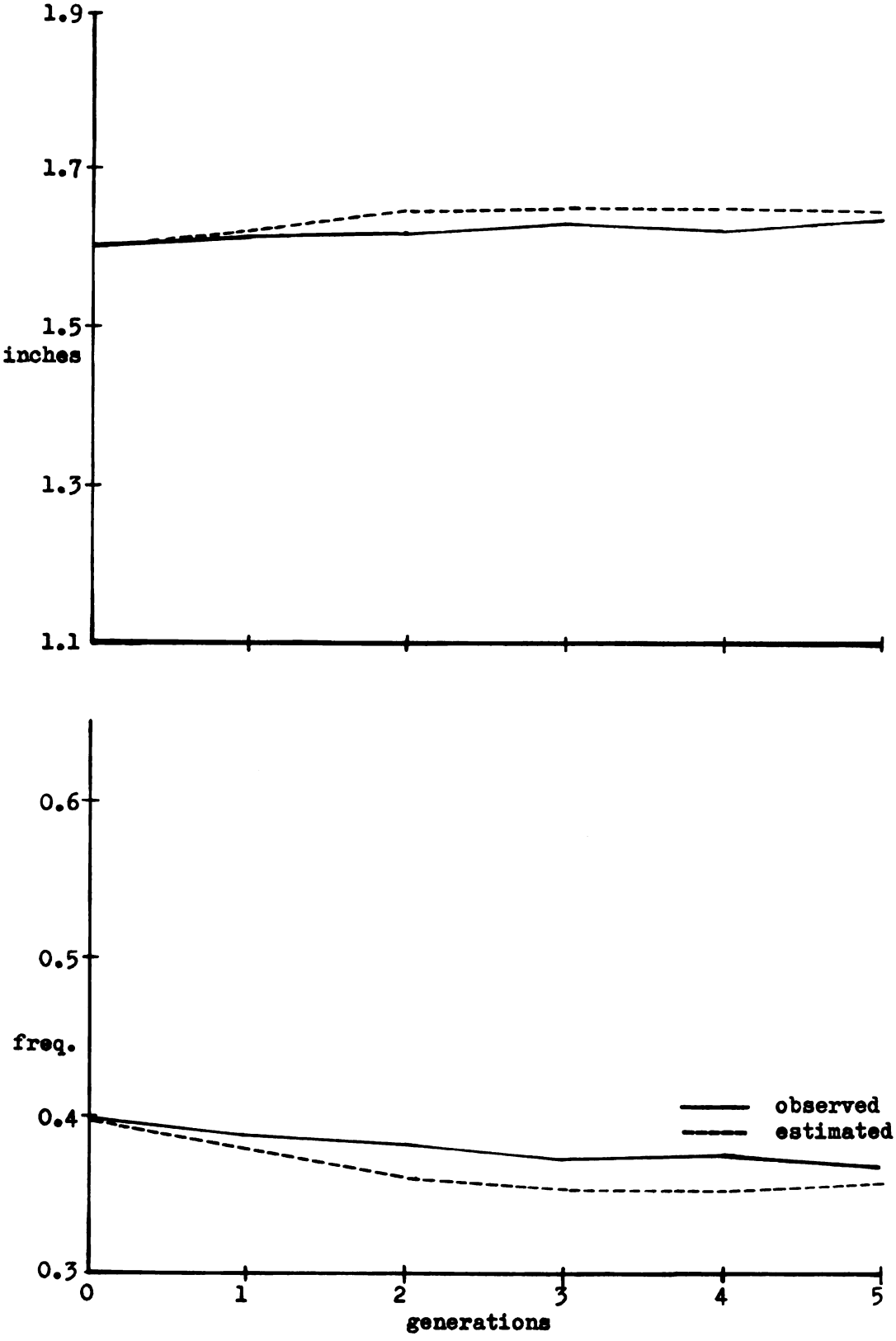


Figure 1.4 Changes in the mean and gene frequency for backfat probe within breed I.

of the desired gene for daily gain. As a result, the genic, genetic and phenotypic variances decreased more rapidly for feed efficiency and, consequently, so did the selection differential and heritability, making further selection less effective, and, thereby, allowing the improvement in daily gain to approach that in feed efficiency. The overall change in the frequencies of the desired genes for daily gain and feed efficiency were also quite similar, 26.1 and 26.5% improvement, respectively.

Both traits which were not selected for in breed I deteriorated over the 5 generations of selection. The mean litter size (figure 1.3) decreased by nearly one pig per litter due to inbreeding depression. The frequency of the desired gene for litter size randomly fluctuated about the initial 0.6 and remained unchanged in generation 5. Since backfat probe (figure 1.4) was positively correlated with daily gain, the pigs became 0.04 inches fatter by generation 5 as a correlated response to selection for daily gain. The frequency of the desired gene for backfat probe declined 7.8% as a correlated response to the change in the frequency of the desired gene for daily gain.

Breed II

Breed II was selected for increased litter size and thinner backfat probe. The changes in breed II are shown in figures 2.1 through 2.4. Figure 2.1 shows there was considerable decline in the mean daily gain, 0.2 lb./day, and nearly 0.1 in the frequency of the desired gene as a correlated response to selection for thinner backfat probe. Expressed as percentages, the declines were 13.0% and 15.6%, respectively. The gene frequency for feed efficiency in breed II (figure 2.2) remained essentially unchanged. However, the mean feed efficiency declined

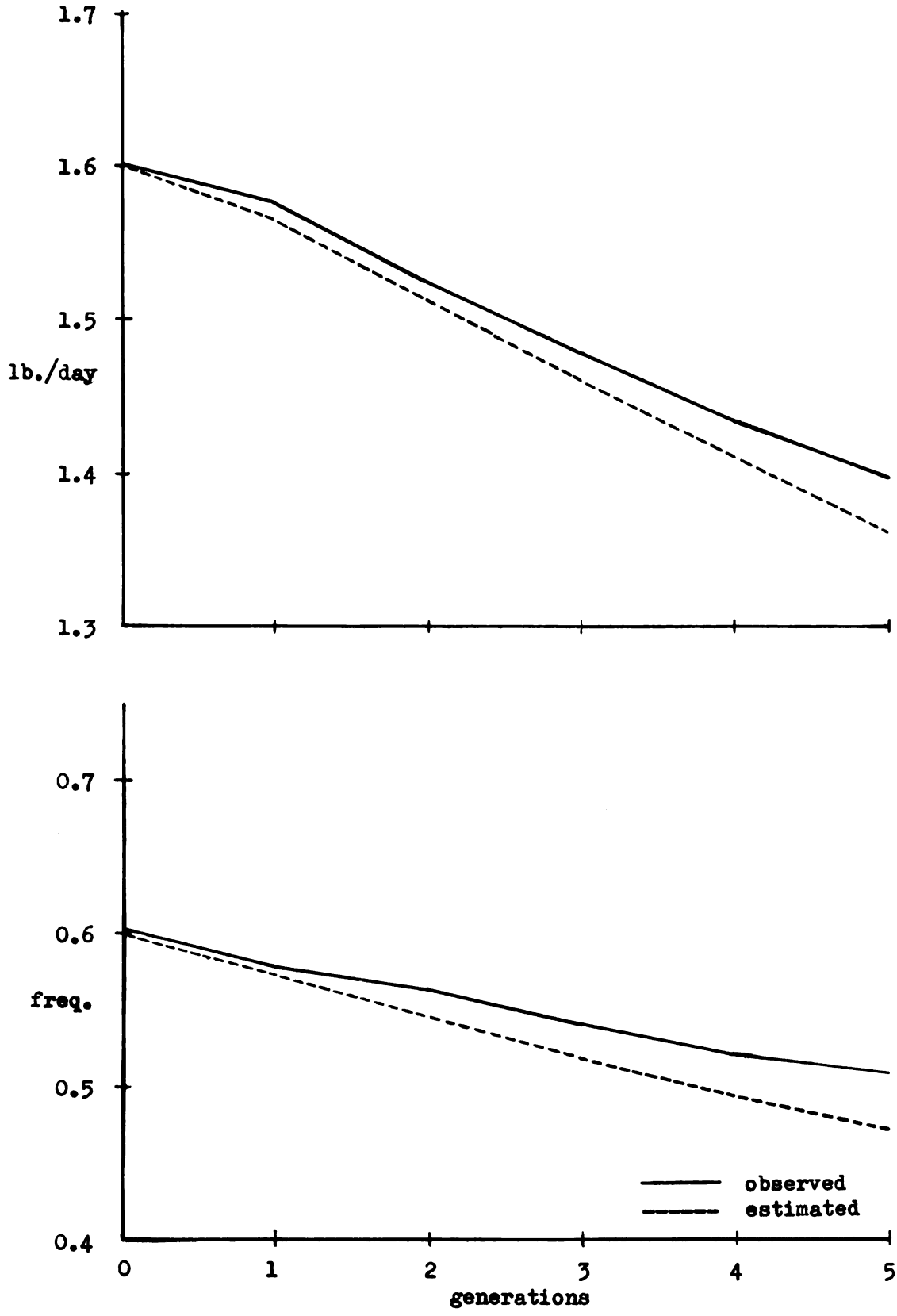


Figure 2.1 Changes in the mean and gene frequency for daily gain within breed II.

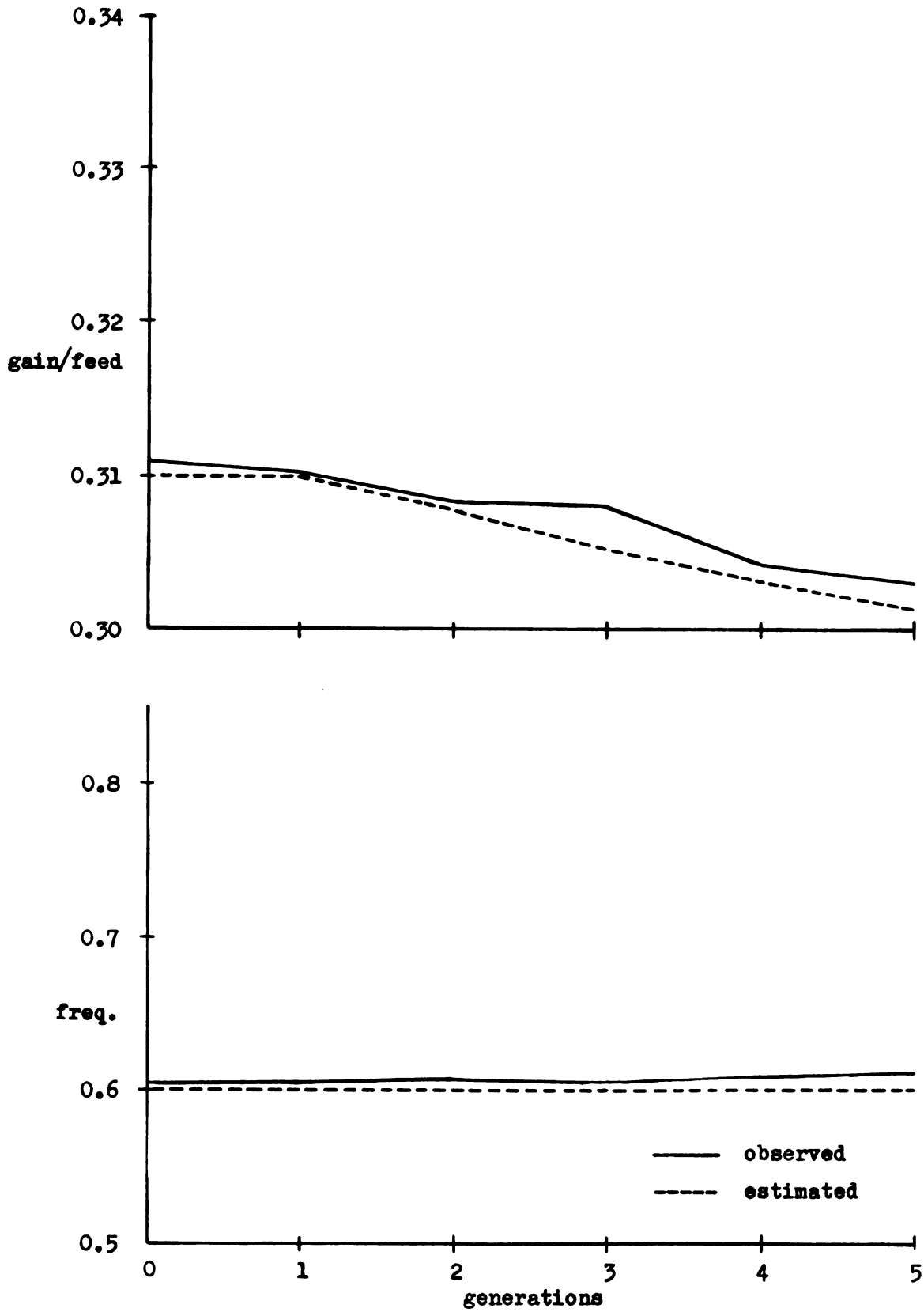


Figure 2.2 Changes in the mean and gene frequency for feed efficiency within breed II.

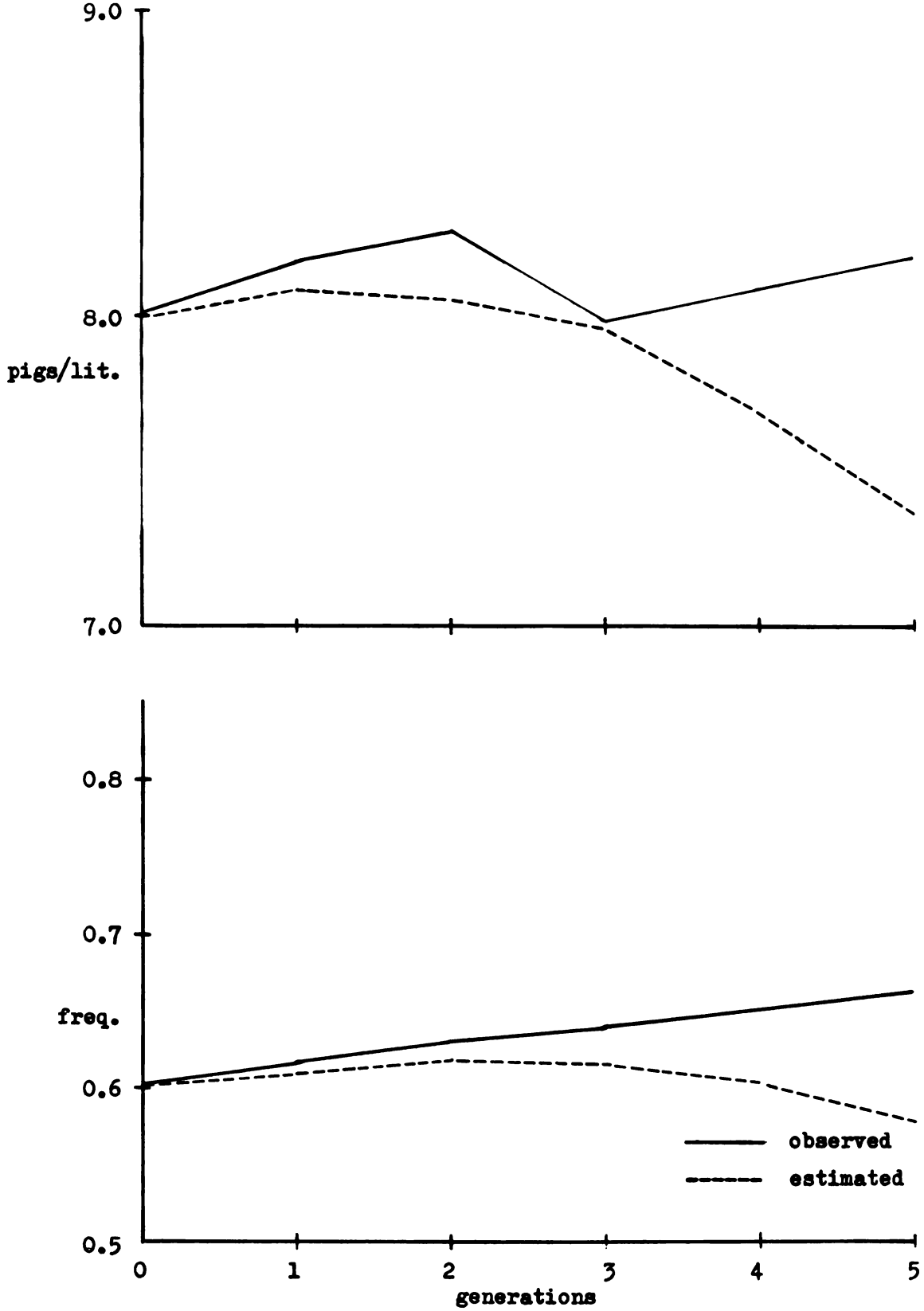


Figure 2.3 Changes in the mean and gene frequency for litter size within breed II.

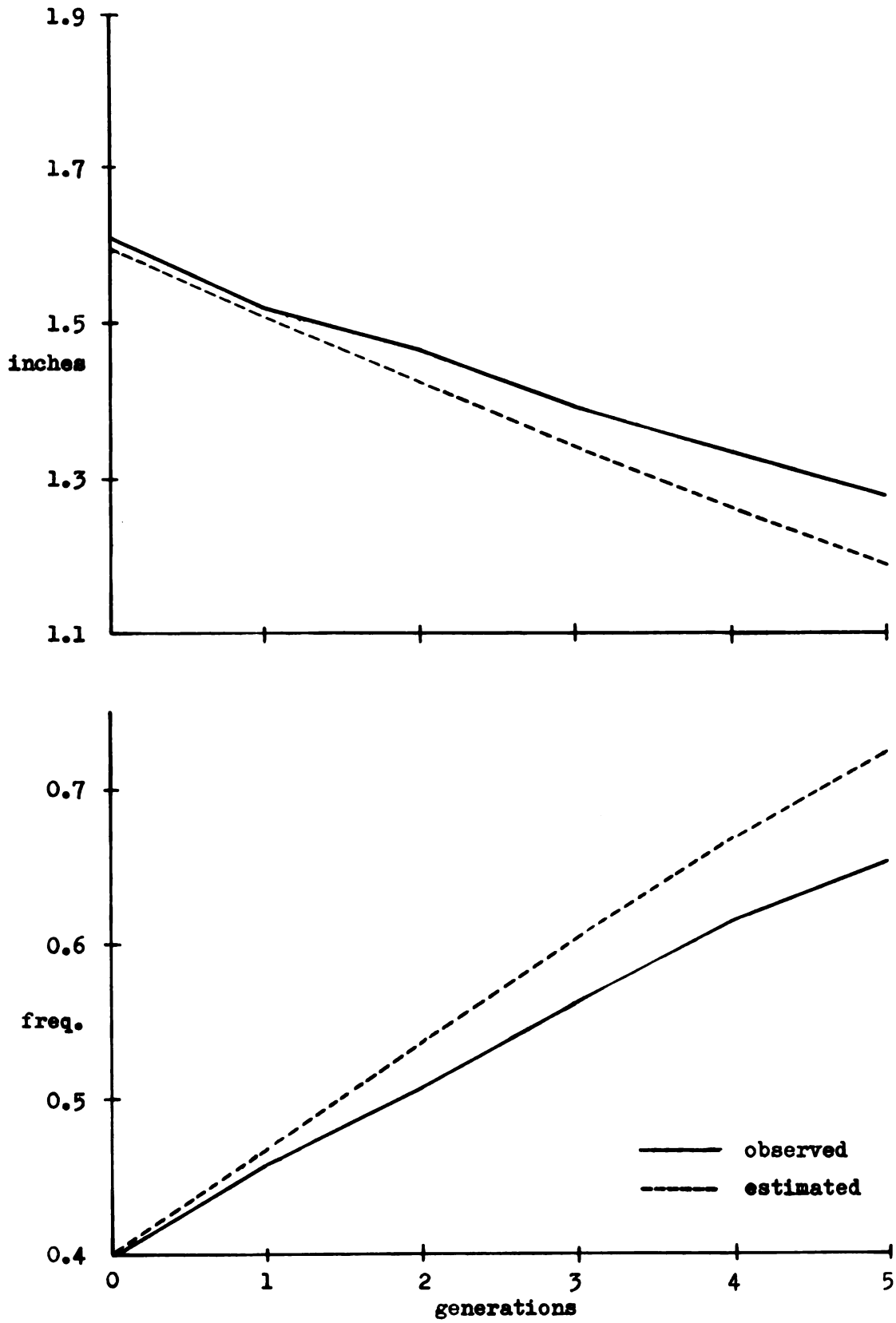


Figure 2.4 Changes in the mean and gene frequency for backfat probe within breed II.

slightly (2.6%) due to inbreeding depression. Figure 2.3 shows the changes in the mean and gene frequency for litter size in breed II. The mean litter size remained essentially unchanged, although the frequency of the desired gene increased 0.06, or 9.8%. The reason for this, of course, was that the effect of increasing the frequency of the desired gene for litter size was cancelled by mounting inbreeding depression. Figure 2.4 indicates that selection for thinner backfat was very effective. The mean backfat probe was decreased by 0.3 inches (20.8%) while the frequency of the desired gene was increased by 0.15 or 65.5%.

The breeding plan was successful in improving all 4 traits in the crossbred pigs, even though selection for thinner backfat and greater average daily gain was antagonistic within the pure breeds. The improvement made in the mean litter size was not realized to its full potential because of the inbreeding depression in the dams, but there was improvement in the frequency of the desired gene for litter size. In addition to the response in the means due to the increased frequency of the desired genes, both daily gain and feed efficiency were affected by heterosis. When measured as the percentage deviation above the average of the parental breeds, the respective amounts of heterosis for daily gain and feed efficiency were 6% and 4%. However, this heterosis included some effects of selection since the breed I boars which sired the cross were selected on feed efficiency.

Accuracy of Predictions of the Means

The estimated and observed means of the traits within each pure breed are shown in figures 1.1 through 2.4. The estimates were arrived at by the following procedure, which is complex but complete. The means

for those traits which were selected for in the first breed were predicted by the equation

$$\bar{X}_{1(i+1)} = \bar{X}_{1i} + \frac{\sigma^2_{A_{1i}}}{\sigma^2_{P_{1i}}} (Z/b)_1 \tfrac{1}{2} \sigma_{P_{1i}} + \left[\frac{\sigma_{A_{1i}}}{\sigma_{P_{1i}}} \frac{\sigma_{A_{2i}}}{\sigma_{P_{2i}}} r_{G_1G_2} (Z/b)_2 \tfrac{1}{2} \sigma_{P_{1i}} \right] -$$

$$2ndp_{i+1} q_{i+1} F_{i+1}$$

where the first subscript on the doubly subscripted variables, i.e., 1 or 2, refers to the trait for which the particular variable is a parameter, 1 for the primary trait and 2 for the secondary, or correlated, trait; i refers to the generation number, 0 through 4; $\bar{X}$ is the mean of the population for the particular trait; $\sigma^2_{A_i}$ is the additive genetic variance for generation i and $\sigma^2_{P_i}$ is the phenotypic variance for generation i (each for the respective traits as is indicated by the first subscript); $(Z/b)_1$ is the selection differential for the primary trait in standard units and $(Z/b)_2$ is the same only for the secondary trait; $r_{G_1G_2}$ is the genetic correlation between the two traits; and the quantity $2ndp_{i+1} q_{i+1} F_{i+1}$ is the inbreeding depression which may be expected in generation i + 1. This quantity was presented by Falconer (1960) as the expected depression in the mean due to inbreeding in a complete dominance model, where n is the number of loci, d is the deviation of the heterozygote from the average of the two homozygotes, p_{i+1} is the expected frequency of the desired allele in generation i + 1, q_{i+1} is the expected frequency of the undesired allele in generation i + 1 and F_{i+1} is the inbreeding expected in generation i + 1. The values of p_{i+1} and q_{i+1} are predicted by iteration of the prediction equation. The value for F_{i+1} is estimated

by the equation presented by Lush (1948), $F_{i+1} = F_i + \frac{N_m + N_f}{8N_mN_f}(1 + F_{i-1} - 2F_i)$ where F_i is the inbreeding in the present generation, F_{i-1} is the inbreeding in the previous generation, N_m is the number of breeding boars and N_f is the number of breeding gilts. The genetic variance, σ_G^2, and the genic variance, σ_A^2, were predicted according to the equations presented by Robertson (1952). Robertson developed equations which predict the within inbred-line genetic and genic variances for a complete dominance model in the case of no selection and low frequency of the recessive allele. The prediction equation for the genetic variance within an inbred line is

$$\sigma_G^2 = \left[0.8q(1-q)(1-F) - q(1-q)(1-2q)(1-F)^3 + 0.2q(1-q) - q^2(1-q)^2(1-F)^6\right] 4nX^2$$

where q is the frequency of the recessive allele, F is the inbreeding coefficient, n is the number of loci and X is half the distance between the homozygous extremes. (The procedure used to estimate q will be described in a succeeding section.) The equation for the genic variance is very similar to that for the genetic variance,

$$\sigma_A^2 = \left[0.6q(1-q)(1-F) - q(1-q)(1-2q)(1-F)^3 + 0.4q\ (1-q) - q^2(1-q)^2(1-F)^6\right] 4nX^2,$$

and as F approaches 1, σ_A^2 approaches three-fourths σ_G^2. Robertson (1952) discussed the affect of selection on the within line genetic and genic variances and concluded that the above equations tend to overestimate the variances when selection is being practiced. However, he also concluded that the bias is small if the rate of inbreeding is rapid and selection against the recessive allele was less than complete. In this study the rate of inbreeding was fairly rapid and the selection against recessive alleles is considerably less than complete. Therefore, these equations should yield estimates with very little bias. The phenotypic

variance is a direct function of the genetic variance being composed of the genetic variance plus a constant amount of environmental variance. The genetic correlation, $r_{G_1G_2}$, in the prediction equation is a constant, 0.6, and is determined by the fraction of genes which the correlated traits have in common. The second quantity in the prediction equation, $\left[\frac{\sigma_{A_{1i}}}{\sigma_{P_{1i}}} \frac{\sigma_{A_{2i}}}{\sigma_{P_{2i}}} r_{G_1G_2} (Z/b)_2 \tfrac{1}{2} \sigma_{P_{1i}}\right]$, is the expected correlated response to selection for a correlated trait as it was developed by Falconer (1960).

The means of the traits which were selected for in breed II were estimated by the standard prediction equation (Lush 1948), since in breed II selection was for the same traits in each sex. This equation was $\bar{X}_{i+1} = \bar{X}_i + \frac{\sigma^2_{A_i}}{\sigma^2_{P_i}} \tfrac{1}{2}\left[(Z/b)_B + (Z/b)_G\right]\sigma_{P_i} - 2ndpqF$, where the subscript i and other symbols are as previously defined, and $(Z/b)_B$ is the selection intensity (in standard units) in the boars and $(Z/b)_G$ is the same, only for the gilts. Of course, in the case of backfat, the quantity 2ndpqF goes to zero, since the heterozygous genotype is equal to the average of the homozygous genotypes and therefore d = 0.

Figure 1.1 shows the predicted and observed results of selection for daily gain in breed I. The estimate of the mean is very close to the observed mean for the first three generations (base, 1 and 2) but begins to diverge from the observed value at generation 3, the divergence becoming significant ($P \leq 0.05$) in generation 4. Figure 1.2 shows the observed and predicted response of the mean for feed efficiency in breed I.

As in daily gain the observed and estimated compare quite well until generation 2. The discrepancy between the observed and estimated thereafter becomes increasingly worse, reaching statistical significance by generation 4 ($P \leq 0.05$). The predicted and realized change in the mean for litter size in breed I is shown in figure 1.3. The estimated and observed changes agree very well throughout all 6 generations for this trait. Figure 1.4 indicates that the estimated and observed correlated response in the mean for backfat probe in breed I agree very well for all 6 generations also.

The observed and predicted changes in the means for breed II are presented in figures 2.1 through 2.4. Figure 2.1 shows the estimated and observed correlated response in the mean for daily gain in breed II. The estimated and observed response is close until generation 5 where they become significantly different ($P \leq 0.05$). Figure 2.2 shows the observed and predicted change in the mean for feed efficiency in breed II. The two are essentially the same except that the estimated response is consistently lower than the observed response. This difference, however, is not statistically significant. Figure 2.3 shows the expected and realized response of the mean for litter size in breed II. The observed response shows considerable fluctuation which is probably due to the interaction of inbreeding and increasing gene frequency. The prediction equation gives a reasonably good estimate until the third generation of selection. After generation 3, the observed and estimated diverge rapidly and the difference is significant ($P \leq 0.05$) by generation 4. The response of backfat to selection within breed II is shown in figure 2.4. Although the observed and estimated appear to be reasonably close, the difference is significant ($P \leq 0.05$) from generation 2 on.

This is, however, the only case in which the prediction overestimates the response to selection. This is partially because backfat is an additive trait and is therefore not affected by the inbreeding depression factor.

The estimated means are consistently lower than the observed means in nearly all cases, except backfat probe in breed II, after the second generation of selection. This difference is especially pronounced when the trait being estimated is being selected for, and generally becomes worse as inbreeding increases. In all cases, except in generation 5 for daily gain in breed II, the agreement between the observed and estimated change in those means for the traits not being directly selected for is very close for all generations.

The deletion of the inbreeding depression factor from the prediction equation causes an even larger bias than does its inclusion, but in the opposite direction. Thus, it appears that modifications must be made in the procedure for estimating the change in the mean in populations exposed to both inbreeding and selection. The accuracy of the estimates of the genetic variances will be discussed in a later section since accurate estimates of genetic variance are necessary for the development of accurate prediction equations.

Accuracy of Predictions of the Gene Frequency

The observed and estimated gene frequencies are shown in figures 1.1 through 2.4. The estimated gene frequencies were established by the solution of a quadratic equation which was presented by Lush (1948)

for the mean of a population with a given gene frequency, a given amount of inbreeding, and a complete dominance gene action model. The original equation is of the form $\bar{X} = \left[(2q + 2q(1-q)(1-F))X + y\right]n$, where $\bar{X}$ is the predicted mean for the generation of interest, q is the frequency of the dominant allele, F is the inbreeding coefficient, X is one-half the deviation of the genotypic value of the homozygous dominant and heterozygous loci above the homozygous recessive loci, y is the genotypic value of the homozygous recessive loci, and n is the number of loci affecting the trait. This equation was expanded and then factored into the components of a quadratic equation. The components are of the form $a = (2F-2)Xn$, $b = (4-2F)nX$, and $c = ny-\bar{X}$, where a, b, and c are the components of the standard equation for the solution of quadratic roots, so that q was solved for by the equation $q = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$. In this case b was negative because if it were positive the solution for q resulted in values larger than one.

Since q is solved for by using the predicted means, it is subject to the same biases that affect the estimates of the means. However, the results are not identical when comparing the estimated and observed gene frequencies with the estimated and observed means for a trait. For example, the mean daily gain in breed I (figure 1.1) is underestimated by generation 3, while the gene frequency is not underestimated until generation 4. The difference in the mean is significant by generation 4 but the difference in the gene frequencies is not significant in any generation. The results for feed efficiency are similar (figure 1.2). In this case the mean in generation 3 is underestimated, but the gene

frequency is overestimated in the same generation. The difference between the means is large by generation 4, but does not become appreciable in the gene frequencies until generation 5. In generation 5, however, the difference in the gene frequencies is still not statistically significant. Although the expected and observed means agree very well for litter size in breed I (figure 1.3), the estimated and observed gene frequencies are even closer. The estimated and observed gene frequency for backfat probe within breed I (figure 1.4) also agree very well over all generations, as do the observed and predicted means.

The breed II results are quite similar to those of breed I. Here again the gene frequencies for the unselected traits agree quite well over all generations. For daily gain (figure 2.1) the results for the gene frequencies are very similar to those for the means, the only difference being that there is no significant difference in generation 5 between the observed and predicted gene frequencies. The predicted gene frequency for feed efficiency in breed II (figure 2.2) is quite similar to the predicted mean to the extent that it consistently underestimates the observed values. However, the agreement is somewhat closer in the gene frequencies because the observed gene frequency shows less fluctuation than does the observed mean. The results for the predicted gene frequency for litter size in breed II (figure 2.3) are similar to those for the predicted mean. The predicted gene frequency is consistently lower than the observed, and the divergence becomes progressively worse in each generation. The divergence of the predicted gene frequency from the observed becomes statistically significant ($P \leq 0.05$) in generation 3. Backfat probe in breed II (figure 2.4) is the only case where the predicted gene frequency is consistently greater

than the observed gene frequency. Contrary to the results based on the means, the differences between the observed and predicted gene frequencies never approach statistical significance.

In general, the predictions of the gene frequencies are better than those of the means, although the two are quite similar in their general response. One reason gene frequency is more predictable than the mean is that it is not affected by inbreeding. The mean is subject to the same causes of variation as is the gene frequency, only in addition it is affected by the distribution of the genes at the various loci. For a given gene frequency, the mean may be quite different depending on whether the dominant genes are at the same or different loci. Thus, there is an effect due to the level of inbreeding on the observed means, but no effect due to inbreeding on the observed gene frequencies. Other than this difference, the results of the predictions for the means and the gene frequencies are very similar.

The Rate of and the Accuracy of the Predictions of Inbreeding

The amount of inbreeding which accrued as a result of the breeding plan is presented in table 2 for each breed. The amount of inbreeding in both breeds is very similar and fairly high by most standards. The accumulation of inbreeding at such a rapid rate is generally undesirable since it depresses the performance of the purebred animals for all traits except backfat and causes the fixation of undesirable alleles due to chance, thus reducing the effectiveness of selection. If this situation were true in a real swine population, it would cost the producer money for the lost performance in his pigs and may cause him to receive lower prices for his breeding stock, especially after the fifth generation of

Table 2. PERCENT OBSERVED AND EXPECTED INBREEDING BY GENERATIONS WITHIN BREEDS.

	BREED I		BREED II	
Generation	% F Observed	% F Estimated*	% F Observed	% F Estimated*
Base	0	0	0	0
1	1	0	1	0
2	8	8	8	4
3	13	14	12	8
4	19	20	18	12
5	23	25	23	15

*Based on $F = F' + \frac{N_m + N_f}{8N_mN_f}(1 + F'' - 2F')$ where ' indicates F in preceeding generation and '' indicates F two generations previous.

selection. At this point the inbreeding seems to be high enough (23%) so that its depressing effects are as great as the effects of increasing the frequency of the desired gene for most of the traits. Thereafter progress would be negative since the effect of the accumulating inbreeding would be greater than the effect of selection. As a result the population mean would actually begin to decrease, although selection for its increase was still being carried on. This would discourage the purebred producer from any further selection since he would be making no apparent progress, although in fact he could still be increasing the gene frequency for the desired allele. For this reason it seems that the breeding plan should be modified. The number of parents in each breed should be increased in order to decrease the rate of inbreeding.

This would allow the producer to realize the benefits of selection, and selection would be more effective since chance would play a less effective role in fixation of loci.

The amount of inbreeding which was expected to have occurred in each breed is also presented in table 2. The equation used to predict the expected inbreeding was one presented by Lush (1948), $F = F' + \frac{N_m + N_f}{8N_mN_f}(1 + F'' - 2F')$, where F is the expected inbreeding in the present generation, F' the inbreeding the previous generation, F'' the inbreeding two generations previous, N_m the number of breeding males and N_f the number of breeding dams. The observed and estimated inbreeding are very similar in the first breed where the number of parents was small and the expected rate of inbreeding was rapid. However, in the second breed the agreement between the observed and expected inbreeding was considerably poorer. In all generations beyond the first, the observed inbreeding was higher than the expected. When these differences were tested, using a t-test, and using pq/n [(Lush 1948), where p is the frequency of the favorable allele, q is the frequency of the unfavorable allele, and n the number of loci (60 in this case)] as the variance of the observed F, all differences were highly significant ($P \leq 0.01$). This, however, is not particularly surprising since there is a correlation between the genic values of the parents for the selected traits due to selection making these animals more similar in their genic values than animals in the population as a whole. This tends to increase the homozygosity in the population above that which is due to finite population size alone. Robertson (1961) theorized that the inbreeding effect is larger than the amount calculated from population size when both selection intensity and heritability are high. This received tentative confirmation from Gill

(1963) when the inbreeding expected for the twentieth generation had already been realized by the fifteenth, in most cases.

The reason the inbreeding in the first breed did not exceed the expected value as it did in breed II is probably due to the rate of inbreeding. The effective number of parents in each breed was calculated according to the formula $\frac{1}{N_E} = \frac{1}{4N_m} + \frac{1}{4N_f}$ (Wright 1931), where N_E is the effective number of parents, N_m the number of sires, and N_f the number of dams. The respective effective numbers of parents for breed I and breed II are 6.7 and 12. Since the effective number in breed II is nearly twice as large as that in breed I, the expected rate of inbreeding in breed I is nearly twice that for breed II. In the presence of such strong inbreeding, selection could have little effect on the correlation of parental genic values in breed I. In addition, the parents in breed I were selected for different, but correlated, traits. This would reduce the correlation between the genic values of the parents which is due to selection in breed I, relative to the correlation in breed II where selection of each sex is on the same traits. This decrease in the correlation would cause an additional reduction in the effect of selection on increasing the homozygosity in breed I.

The Accuracy of the Predictions of and the Stability of Genetic Variance

The observed and predicted genetic standard deviations are presented in tables 3 and 4 for breed I and breed II, respectively. (Standard deviations are presented, instead of variances, to reduce the number of decimal places required.) As stated in a previous section, the prediction equation used to estimate the expected values was one presented by Robertson (1952) for the genetic variance within inbred

Table 3. OBSERVED AND ESTIMATED GENETIC STANDARD DEVIATIONS IN BREED I.

	Daily gain		Feed eff.		Litter size		Backfat probe	
Generation	Observed	Estimated	Observed	Estimated	Observed	Estimated	Observed	Estimated
Base	0.1206	0.1231	0.0202	0.0194	0.9768	0.9631	0.0964	0.0990
1	0.1042	0.1105	0.0164	0.0172	0.9027	0.9631	0.0875	0.0990
2	0.1020	0.1010	0.0160	0.0156	0.9084	0.9624	0.0897	0.0952
3	0.0968	0.0977	0.0161	0.0150	0.9506	0.9497	0.0876	0.0919
4	0.0913	0.0981	0.0143	0.0150	0.9009	0.9343	0.0871	0.0886
5	0.0838	0.1000	0.0131	0.0153	0.8951	0.9178	0.0874	0.0855

Table 4. OBSERVED AND ESTIMATED GENETIC STANDARD DEVIATIONS IN BREED II.

	Daily gain		Feed eff.		Litter size		Backfat probe	
Generation	Observed	Estimated	Observed	Estimated	Observed	Estimated	Observed	Estimated
Base	0.1246	0.1231	0.0199	0.0194	0.9917	0.9631	0.1028	0.0990
1	0.1244	0.1296	0.0191	0.0194	0.8861	0.9434	0.0919	0.1008
2	0.1307	0.1348	0.0207	0.0194	0.9035	0.9275	0.0899	0.0986
3	0.1264	0.1337	0.0187	0.0190	0.9031	0.9327	0.0806	0.0948
4	0.1305	0.1394	0.0195	0.0193	0.8343	0.9513	0.0833	0.0896
5	0.1269	0.1399	0.0184	0.0191	0.8638	0.9755	0.0758	0.0832

lines. The equation was developed for the case of no selection, but in the case of rapid inbreeding and less than complete selection against the recessive allele is only slightly biased upward.

In all cases the estimated and observed genetic standard deviations compare very well. However, as stated by Robertson (1952), they are somewhat biased upward for the traits which are being selected for. However, the difference is small even in generation 5 of the selected traits in both breeds, although it is consistently high. The agreement between the expected and observed variances in the traits not selected for in each breed is generally closer than for the selected traits.

The breeding plan caused considerable decrease in the genetic variance in both breeds for the selected traits. The genetic variance for daily gain in breed I decreased 52%, while the genetic variance for feed efficiency in breed I decreased 58%. The decrease in the unselected traits was considerably less, 16% and 18% for litter size and backfat probe, respectively. The decrease in the genetic variances in breed II are not as large. The decreases in the selected traits in breed II, litter size and backfat probe, are 24% and 46%, respectively. The genetic variance for feed efficiency in breed II declined 14%, but the genetic variance for daily gain increased 4% due to the gene frequency approaching 0.5 in generation 5.

The decreases in the genetic variances for the selected traits have a detrimental effect on the progress which can be made by selection, since it causes a decline in both the selection differential and the heritability of the trait. That amount of the reduction in variation which is due to the increase in the frequency of the desired gene is of course unavoidable since the intention of the selection is to increase the frequency of the desired gene. However, the decrease in the genetic variance which is due to inbreeding could be reduced by increasing the size of the breeding population. In view of the large amount of

inbreeding (23%) which has accumulated by the fifth generation of selection, the population size should certainly be increased if the purebred breeder should happen to have the facilities to do so. The size of the breeding population of course would vary between purebred breeders, depending on the size of their operation. In cases of small scale purebred breeders, where the population size could be no larger than the size simulated in this investigation, it might prove worth while to investigate the efficiency of increasing the number of parents even though this would mean a decrease in the selection intensity. This is especially true in the case of feed efficiency selection in breed I, since the selection intensity for this trait is beyond the limits where the relation between the portion of animals saved for parents and the selection intensity is linear. Lush (1948) states that this relation is nearly linear between 0.8 and 0.2, but drops off when out of this range. When using the size of the breeding population in this simulation as an example, the number of boars saved for parents could be doubled, and this would reduce the selection differential by only 20%. At the same time it would decrease the rate of inbreeding by about 42%. Whether this would have a positive or negative effect on total improvement could very well be the subject of another Monte Carlo investigation. It would certainly not seem advisable for a breeder to use any smaller numbers than have been used in this simulation, since stronger inbreeding would weaken the effects of selection. If a breeder should have facilities large enough to accomodate an increase in the size of the breeding herd in only one of the breeds, it should be in breed I rather than breed II. This is because the effective number of parents in breed I is smaller than that in breed II, and because much of the increase in homozygosity

in breed II is, apparently, due to selection, whereas that in breed I is due to finite population size. Thus, the increase in population size would be ralatively more effective in breed I than in breed II.

The Changes in the Genotypic Correlations

The changes in the genotypic correlations are shown in table 5.

Table 5. OBSERVED GENOTYPIC CORRELATIONS BETWEEN CORRELATED TRAITS WITHIN BREEDS

	Breed I		Breed II	
Generation	Daily gain and feed efficiency	Daily gain and backfat probe	Daily gain and feed efficiency	Daily gain and backfat probe
Base	0.59	0.32	0.62	0.34
1	0.56	0.24	0.56	0.30
2	0.58	0.29	0.61	0.34
3	0.55	0.33	0.61	0.31
4	0.54	0.31	0.58	0.38
5	0.48	0.33	0.58	0.35

Generally the genotypic correlations were maintained and showed little change from the base to the final generations. The only exception to this occurred in breed I, where the observed genotypic correlation between daily gain and feed efficiency declined from 0.59 in the base generation to 0.48 in the final generation. This decline can be accounted for by the fixation of some pleiotropic loci due to selection, and by the fact that the frequency of the desired allele for feed efficiency and daily gain was higher at the pleiotropic loci than at

the non-pleiotropic loci, causing the contribution of each pleiotropic locus to the total variation of each trait to be less than the contribution made by each non-pleiotropic locus, since the initial gene frequency was greater than 0.5. Both the amount of fixation and the frequency of the desired allele are expected to be higher at the pleiotropic loci than the non-pleiotropic loci because the pleiotropic loci are affected by the selection on both traits, but the non-pleiotropic loci are affected by the selection on only one of the traits.

SUGGESTIONS FOR FURTHER RESEARCH

As a result of the experience gained from having done this investigation, a few factors which could improve any further, similar, work are listed below.

1. Perhaps the one greatest weakness of this investigation is the limited number of loci simulated. It would have meant virtually no additional work or computer memory space, to have included at least 4 extra pairs of genes per trait. This is because each computer word of memory in the CDC 3600 is a 48 bit word, and, therefore, could accomodate 24 pairs of genes for each trait in the same number of words required to simulate 20 pairs per trait. Furthermore, since the completion of this work, the 3600 computer facilities at MSU have been greatly expanded by the addition of another module of core memory (32,768 words) and by the implementation of a magnetic drum. This increase in memory capacity would facilitate the use of double precision arrays, instead of single precision arrays, for the storage of genotypes. Therefore, the number of loci affecting each trait could be doubled from 24 to 48 pairs, with relative ease.
2. It is suggested that random normal deviates in further simulation work be generated by the method outlined by Gill (1963), rather than be read off tape, since the arithmetic operations of the 3600 are tremendously faster than the I/O operations.
3. As has already been indicated in the discussion section, it would be interesting to investigate varying degrees of selection intensity in populations of fixed size, in order to determine the optimum

combinations of selection intensity and rate of inbreeding.

4. It would be more interesting to compare two or more proposed breeding plans by the Monte Carlo technique, than to evaluate an individual plan, since in the case of the individual plan there are no criteria with which to measure the efficiency of the utilization of the existing genetic variation.
5. In future work where both pleiotropic and non-pleiotropic genes affect a trait, the changes in the frequency of the desired allele at the two types of loci should be followed separately, in order to determine the effect, quantitatively, pleiotropy has on the change in the gene frequency.
6. Since the completion of this work, it has been pointed out that the use of Z/b as an estimate of the expected selection differential causes an upward bias in the expected selection differential when the parental population size is less than 50. This is because of a discrete distribution of genotypes in the parental population rather than a continuous distribution. To remove such bias in experiments with small parental population size the values from table XX in Fisher and Yates (1953) should be used.

SUMMARY AND CONCLUSIONS

A Monte Carlo simulation of a population of swine was done in order to examine the effectiveness of a proposed breeding plan. The breeding plan was designed to improve the performance in crossbred pigs by selection for improved production traits within pure breeds, a situation analagous to the commercial and purebred swine producers in the U.S. today. In addition to the evaluation of the breeding plan, the predictability of, and the changes in, several population parameters were evaluated.

The proposed breeding plan involved two breeds. Within each of the two breeds, selection was for two different performance traits. In the first breed selection was for feed efficiency and daily gain, while in the second breed selection was for litter size and backfat probe. In the first breed, the two most efficient of 20 tested boars were selected for feed efficiency and the 10 fastest gaining of 30 gilts were selected on daily gain. In the second breed, the 4 leanest boars and the 20 leanest gilts from the 8 and 12 dams with the largest two-litter-average litter size, respectively, were selected for backfat probe. There was a 0.6 genetic correlation between daily gain and feed efficiency and a 0.4 genetic correlation between daily gain and backfat probe. These genetic correlations were caused solely by pleiotropy and were determined by the number of loci which affected both traits in the same direction. All other genetic correlations were zero. The initial frequency of the desired allele was 0.6 for all traits in both breeds, except backfat probe for which it was 0.4. All traits were affected by a complete dominance gene model and twenty pairs of independently segragating genes with equal effects, except backfat probe which was determined by an additive gene model.

Selection was continued for five generations within each pure breed. However, a cross between the pure breeds was made before the first and after the fifth generation of selection in order to measure the effectiveness of the breeding plan in improving the performance of the crossbred pigs. All conclusions are based on the average results of twenty independent runs through the computer.

The breeding plan was found to be effective in improving the mean of all four traits in the crossbred pigs. This was due to both an increase in the frequency of the desired gene for all traits and to the expression of 6% and 4% heterosis in daily gain and feed efficiency, respectively. There was a 2.5% increase in the mean for daily gain from the first to the second cross, and a 3.8% improvement in the mean for feed efficiency. Litter size was least improved, 1%, and backfat probe was most improved, 14.6%. The increases in the frequencies of the desired genes were somewhat better. The improvement in the frequency of the desired gene was 3.8% for daily gain, 10.3% for feed efficiency, 9.8% for litter size, and 23.9% for backfat probe.

Selection within the pure breeds was evaluated. In breed I the means for daily gain and feed efficiency showed 5.4% and 5.9% improvement, respectively, while improvement in the frequency of the respective desired genes were 26.1% and 26.5%. In the second breed, the mean for litter size improved 1% while the gene frequency improved 9.8%. The mean for backfat probe improved 21%, while the frequency of the desired gene for backfat probe increased 66%.

The expected means within the pure breeds were calculated and compared to the observed means. The means for the unselected traits were accurately estimated, but the means for the selected traits were gener-

ally underestimated. In the case of selection, the discrepancy became larger as inbreeding increased.

The expected gene frequencies within the pure breeds were also calculated and compared to the observed gene frequencies. Although the predictions of the gene frequencies were dependent on the means, they were generally better than the predictions of the means, but showed basically similar responses. This was explained, at least partially, by the fact that inbreeding has an effect on the observed mean, but not on the observed gene frequency.

The expected inbreeding was calculated and compared to the observed inbreeding within each of the pure breeds. The agreement between the observed and expected inbreeding was very good in the first breed, but very poor in the second. The discrepancy between the observed and expected inbreeding in the second breed was attributed to the effects of selection and a slower rate of inbreeding which made the effects of selection relatively more powerful in the second breed than it was in the first breed.

The expected genetic variances were computed and compared to the observed genetic variances within each of the pure breeds. The predictions were very close to the observed results, but were generally better for the unselected than the selected traits. The genetic variance in the selected traits tended to be overestimated by a small amount.

The decline from the initial to the final generation of selection in the genetic variance in the selected traits was considerable in both breeds. There was a 52% decrease in the genetic variance in daily gain and a 58% decrease in the genetic variance for feed efficiency in breed I. In breed II the genetic variance for litter size declined 24% and

the genetic variance for backfat probe decreased 46%. The unselected traits within each breed decreased 18% or less. The decreases in the genetic variances were detrimental to the progress made by selection and a larger breeding population was recommended, especially for breed I, to reduce the rate of decline which was due to inbreeding. The increase in population size was considered less crucial for breed II since much of the increase in homozygosity in this breed was apparently due to selection and was therefore considered unavoidable.

There was very little change in the genotypic correlations, in general. The only exception was the genotypic correlation between daily gain and feed efficiency in the first breed, which declined from an initial 0.59 to a final 0.48.

LITERATURE CITED

Anderson, D. E. 1954. Genetic relations between carcass characters, rate, and economy of gain. Ph.D. Thesis, Library, Iowa State College, Ames.

Baker, M. L., L. N. Hazel and C. F. Reinmiller. 1943. The relative importance of heredity and environment in the growth of pigs of different ages. J. An. Sci. 2:3.

Baker, L. H. and R. E. Comstock. 1962. Monte Carlo studies of linkage effects in population genetics. Genetics 47:949. (Abstr.)

Barker, J. S. F. 1958a. Simulation of genetic systems by automatic digital computers. III. Selection between alleles at an autosomal locus. Aust. J. Biol. Sci. 11:603.

Barker, J. S. F. 1958b. Simulation of genetic systems by automatic digital computers. IV. Selection between alleles at a sex-linked locus. Aust. J. Biol. Sci. 11:613.

Barker, J. S. F. and J. C. Butcher. 1966. A simulation study of quasi-fixation of genes due to random fluctuations of selection intensities. Genetics 53:261.

Biswas, D. K., P. V. Hurt, A. B. Chapman, N. L. First and H. L. Self. 1963. Feed efficiency and carcass desirability in swine. J. An. Sci. (abs.) 22:1109.

Blunn, C. L. and M. L. Baker. 1949. Heritability estimates of sow productivity and litter performance. J. An. Sci. 8:89.

Blunn, C. T., G. N. Baker, and L. E. Hanson. 1953. Heritability of gain in different growth periods in swine. J. An. Sci. 12:39.

Bohidar, N. R. 1960. The role of sex-linked genes in quantitative inheritance. Ph.D. Thesis, Library, Iowa State University, Ames.

Boglun, W. J., W. E. Rempel, and R. E. Comstock. 1961. Heritability of litter size in swine. J. An. Sci. 20:566.

Brinks, J. S. 1960. A genetic analysis of carcass and production traits in swine. Ph.D. Thesis, Library, Iowa State University, Ames.

Cockerham, C. C. 1952. Genetic covariation among characteristics in swine. J. An. Sci. 11:738.

Cox, D. F. 1959a. Components of genetic variance in swine as estimated from the regression of phenotypic variances on relationship within pens. J. An. Sci. 18:1464. (Abstr.)

Cox, D. F. 1959b. Genetic and environmental factors influencing gain and fatness in swine. Ph.D. Thesis, Library, Iowa State College, Ames.

Craft, W. A. 1958. Fifty years of progress in swine breeding. J. An. Sci. 17:960.

Craig, J. F., H. W. Norton and S. W. Terill. 1956. A genetic study of weight at five ages in Hampshire Swine. J. An. Sci. 15:242.

Cummings, J. N., L. M. Winters and H. A. Stewart. 1947. The heritability of some factors affecting productivity in brood sows. J. An. Sci. 6:297.

Depape, J. G. 1954. Swine carcass studies. I. Genetic variation and covariation in carcass characteristics. II. The probe as an indicator of carcass merit. Ph.D. Thesis, Library, Oklahoma A.& M. College, Stillwater.

Dickerson, G. E. and L. N. Hazel. 1944. Selection for growth rate of pigs and productivity of sows. J. An. Sci. 3:201.

Dickerson, G. E. 1947. Composition of hog carcasses as influenced by heritable differences in rate and economy of gain. Research Bull. 354, Iowa State College, Ames.

Dickerson, G. E. and J. G. Grimes. 1947. Effectiveness of selection for efficiency of gain in Duroc swine. J. An. Sci. 6:625.

Dickerson, G. E., et al. 1954. Evaluation of selection in developing inbred lines of swine. Research Bull. 551, University of Missouri, Columbia.

Dillard, E. U., O. W. Robison and J. E. Legates. 1962. Selection for low backfat ratio in swine. J. An. Sci. 21:971. (Abstr.)

Enfield, F. D. and J. A. Whatley. 1961. Heritability of carcass length, carcass backfat thickness and loin eye area in swine. J. An. Sci. 20:631.

England, D. C. and L. M. Winters. 1953. The effects of genetic diversity and performance of inbred lines per se on hybrid vigor in swine. J. An. Sci. 12:836.

Falconer, D. S. 1960. Introduction to quantitative genetics. The Ronald Press Co., New York, N. Y.

Fisher, R. A. 1925. Statistical methods for research workers. London, Oliver and Boyd Pub. Co.

Fisher, R. A. and F. Yates. 1953. Statistical Tables for biological, agricultural, and medical research. Edinburgh and London, Oliver and Boyd Pub. Co.

Fox, R. R. 1959. Relations between backfat thickness and other economic traits in swine. Ph.D. Thesis, Library, University of Minnesota. St. Paul.

Fraser, A. S. 1957a. Simulation of genetic systems by automatic digital computers. I. Introduction. Aust. J. Biol. Sci. 10:484.

Fraser, A. S. 1957b. Simulation of genetic systems by automatic digital computers. II. Effects of linkage on rate of advance under selection. Aust. J. Biol. Sci. 10:492.

Fraser, A. S. 1960a. Simulation of genetic systems by automatic digital computers. V. Linkage, dominance and epistasis. In Biometrical Genetics. Kempthorne, O. Ed. Pergammon Press, New Yor, N. Y. p. 70.

Fraser, A. S. 1960b. Simulation of genetic systems by automatic digital computers. VI. Epistasis. Aust. J. Biol. Sci. 13:150.

Fraser, A. S. 1960c. Simulation of genetic systems by automatic digital computers. VII. Effects of reproductive rate and intensity of selection on genetic structure. Aust. J. Biol. Sci. 13:344.

Fraser, A. S., D. Burnell, and D. Miller. 1966. Simulation of genetic systems by automatic digital computers. X. Inversion polymorphism. J. Theoret. Biol. 13:1.

Fredeen, H. T. 1953. Genetic aspects of Canadian bacon production. Canadian Dept. of Agr. Pub. No. 889:38.

Fredeen, H. T. and P. Jonsson. 1957. Genic variance and covariance in Danish Landrace swine as evaluated under a system of individual feeding of progeny test groups. Z. fur Lierzuchtung und Zuchtungbiologic 70:348.

Gill, J. L. 1963. Effect of population size, selection intensity, linkage, and non-additive variability upon genetic change in simulated genetic populations. Ph.D. Thesis, Iowa State University, Ames.

Gill, J. L. 1965a. Effects of finite size on selection advance in simulated genetic populations. Aust. J. Biol. Sci. 18:599.

Gill, J. L. 1965b. A Monte Carlo evaluation of predicted selection response. Aust. J. Biol. Sci. 18:999.

Gill, J. L. 1965c. Selection and linkage in simulated genetic populations. Aust. J. Biol. Sci. 18:1171.

Gray, R. D., L. F. Tribble, B. N. Day, and J. F. Lasley. 1964. Five generations of selection for thinner backfat. Personal communications.

Griffing, B. 1960. Theoretical consequences of truncation selection based on the individual phenotype. Aust. J. Biol. Sci. 13:307.

Hazel, L. N., M. L. Baker and C. F. Reinmiller. 1943. Genetic and environmental correlations between growth rates of pigs at different ages. J. An. Sci. 2:118.

Hetzer, H. O. and J. H. Zeller. 1956. Selection for high and low fatness in Duroc and Yorkshire swine. J. An. Sci. 15:1215.

Hetzer, H. O., W. R. Harvey and W. H. Peters. 1963. Selection for high and low fatness in Duroc and Yorkshire swine. Genetics Today 1:268. (Abstr.)

Issawi, E. L. and W. E. Rempel. 1961. Heritability of growth rate in inbred swine based on a crossbred foundation. J. An. Sci. 20:593.

Johansson, I. and N. Korkman. 1950. A study of the variation in the production traits in the bacon pig. Acta. Agric. Scand. 1:62.

Jonsson, P. and J. W. B. King. 1962. Sources of variation in Danish Landrace pigs at progeny-testing stations. Acta. Agric. Scand. 12(1-2):68.

Kimura, M. 1954. Process leading to quasi-fixation of genes in natural populations due to random fluctuations in selection intensities. Genetics 39:280.

Kimura, M. 1958. On the change of population fitness by natural selection. Heredity 12:145.

Kimura, M. 1962. On the probability of fixation of mutant genes in a population. Genetics 47:713.

Kojima, K. 1961. Effects of dominance and size of population on response to mass selection. Genetical Research 2:177.

Krider, J. L., B. W. Fairbanks, W. E. Carroll and E. Roberts. 1946. Effectiveness of selecting for rapid and slow growth rate in Hampshire swine. J. An. Sci. 5:3.

Lasley, J. F. 1964. Genetics of Livestock Improvement. Prentice-Hall Inc.

Lerner, E. H., D. T. Mayer and J. F. Lasley. 1957. Early embryonic mortability in strain-crossed gilts. Research Bull. 629. University of Missouri, Columbia.

Louca, A. and O. W. Robison. 1965. Heritability and genetic correlations in swine. J. An. Sci. 24:850. (Abstr.)

Lush, J. L. 1936. Genetic aspects of Danish system of progeny-testing swine. Research Bull. 204. Iowa State College, Ames.

Lush, J. L. and A. E. Molln. 1942. Litter size and weight as permanent characteristics of sows. USDA Tech. Bull. 836.

Lush, J. L. 1948. The genetics of populations (Mimeo.) Dept. of Animal Science, Iowa State University, Ames.

Martin, F. C. and C. C. Cockerham. 1960. High speed selection studies. In Biometrical Genetics. Kempthorne, O., Ed. Pergammon Press, New York, N. Y., p. 35.

Mather, K. 1943. Polygenic inheritance and natural selection. Biological Review 18:32.

Omtvedt, I. T., R. L. Artland and R. E. Jacobs. 1963. Improving swine through breeding. University of Minnesota Ext. Bull. 306.

Park, Y. I., J. L. Lush and L. N. Hazel. 1963. Age constant feed efficiency of pigs. J. An. Sci. 22:1109. (Abstr.)

Parker, R. J. 1966. Genetic correlation and response to selection in simulated populations. Ph.D. Thesis, Library, Michigan State University, East Lansing.

Qureshi, A. W. 1963. A Monte Carlo evaluation of the role of finite population size and linkage in response to continuous mass selection. Tech. Report No. MC6. Stat. Lab., Iowa State University, Ames.

Qureshi, A. W. 1964. A Monte Carlo evaluation of the role of finite population size and linkage in response to continuous mass selection. II. Dominance and overdominance. Tech. Report No. MC9. Stat. Lab. Iowa State University, Ames.

Reddy, V. B., J. F. Lasley and D. T. Mayer. 1958. Genetic aspects of reproduction in swine. Research Bull. 666. University of Missouri, Columbia.

Reddy, V. B., J. F. Lasley and L. F. Tribble. 1959. Heritabilities and heterosis of some economic traits in swine. Mo. Ag. Exp. Sta. Bull. No. 689:27.

Riemer, D., R. E. Comstock, W. E. Rempel and A. B. Salema. 1958. Genetic correlations between growth rate and efficiency of feed utilization in swine. J. An. Sci. 17:1137.

Robertson, A. 1952. The effect of inbreeding on the variation due to recessive genes. Genetics 37:189.

Robertson, A. 1960. A theory of limits in artificial selection. Proc. Royal Society, B. 153:234.

Robertson, A. 1961. Inbreeding in artificial selection programmes. Genetical Research 2:189.

Rotenberg, A. 1960. A new pseudo-random number generator. J. Assoc. Comp. Mach. 7:75.

Smith, A. B. D. and H. P. Donald. 1937. Weaning weight in pigs and litter sampling with reference to litter size. J. An. Sci. 28:485

Stewart, H. A. 1945. The inheritance of prolifacy in swine. J. An. Sci. 4:359.

Stothard, J. G. 1947. A study of parental-progeny correlations with Canadian bacon hogs. Sci. Agri. 27:354.

Vogt, D. W., R. E. Comstock, and W. E. Rempel. 1963. Genetic correlations between some economically important traits in swine. J. An. Sci. 22:214.

Whatley, J. A. and R. H. Nelson. 1942. Heritability of differences in 180 day weight and market score in Duroc swine. J. An. Sci. 1:70.

Whatley, J. A. 1956. Heritability in hogs. American Hampshire Herdsman. June, 1956:32.

Wright, S. 1921. Assortative mating based on somatic resemblance. Genetics 6:144.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97.

Young, S. S. Y. 1966. Computer simulation of directional selection in large populations. I. The programme, the additive and the dominance models. Genetics 53:189.

Zoellner, K. O., J. F. Lasley, L. F. Tribble and B. N. Day. 1963. Selection for thinner backfat in swine. Research Bull. 831, University of Missouri.