Euphytica 10 (1961): 368-378

PREDICTED AND ACTUAL RESPONSES DURING LONG-TERM SELECTION FOR CHEMICAL COMPOSITION IN MAIZE[1])

EARL R. LENG
Professor of Agronomy, University of Illinois, Urbana, Illinois, U.S.A.

With 4 figures

Received 6 June 1961

ABSTRACT

Selection for higher and lower levels of protein and oil content in maize has been conducted for 61 generations. Response to selection has been irregular, and prediction of response by extrapolation of regression trend lines proved seriously inaccurate in at least 9 of 12 predictions. Application of a "selection response" formula also yielded generally unsatisfactory predictions. Analysis of the morphological components of oil content led to very accurate prediction of response in the "High Oil" strain and to a reasonably satisfactory prediction in "Low Oil". It was apparent that in the last 55 generations of selection, selection response for oil content was concentrated in two separate morphological characters – *germ size* in the "high" selection and *oil content of the germ* in selection for lower oil content. A satisfactory analysis of the character "protein content" could not be made from data available.

1. INTRODUCTION

Predictable response to planned selection is considered the basic requirement of a scientific plant breeding program. If response is not predictable to a significant degree, success in a breeding program depends to a disproportionate extent on chance or on application of "breeder's art". While the latter capacity is always valuable in breeding, it cannot be considered an efficient basis for attaining desired objectives in complex selection programs.

Modern literature in breeding and selection theory contains many references to purely theoretical expectations, or to predictions based on short-term actual experience. Rarely are sufficient longterm breeding data made available to provide satisfactory tests of the agreement between predicted and actual responses to selection. Evaluations of incorrect predictions are even more rarely presented, yet the success of future work rests to a great extent on correct appraisal of the validity of prediction methods employed.

The data from 61 consecutive generations of selection for chemical composition in maize, recently presented by LENG (7), provide unique material for evaluating the agreement between predicted and actual response to selection. This experiment is the only case on record in a major economic plant species in which selection for quantitative characters has been conducted at a fixed selection intensity, with the responses accurately recorded, for a large number of consecutive generations.

[1]) Prepared as part of a research project while the author was a Fulbright senior research fellow at the Max-Planck-Institut für Züchtungsforschung, Köln-Vogelsang, Germany, March–July, 1961.

Evaluation of predicted and actual responses in this material is especially interesting because at three stages of the experiment, predictions of future response actually were made. These were by EAST (3) after 10 generations, WINTER (10) after 28 generations, and WOODWORTH *et al.* (11) after 50 generations of selection.

Major discrepancies between predicted and actual response occurred much more frequently than did close agreement between the predictions and the observed results. The problems arising from this situation are of fundamental importance in the evaluation of plant breeding systems.

2. MATERIALS AND METHODS

Selection for both higher and lower levels of oil and protein content in maize was begun by HOPKINS (6) in 1896. Data were recorded as percentage of total protein or oil in the whole grain, on a water-free basis. With minor exceptions in the first 9 generations, 20 percent. of the analyzed ears in each strain were selected to provide seed for the next generation. The size of the analyzed populations was 120 ears per strain from the early generations until the 28th generation; thereafter, 60 ears per strain were analyzed. The "ear-to-row" system of breeding was followed for the first 28 generations; a system of intra-strain reciprocal crossing, with selection based on individual phenotypes, has been followed since the 29th generation. Details of breeding methods, chemical analyses, estimated inbreeding levels, and response trends have recently been presented in another paper (7).

EAST (3) fitted curves to the data presented by SMITH (9) on the results of the first 10 generations of selection. WINTER (10), studying changes in means and variability during the first 28 generations, predicted future responses by extrapolating trend lines fitted to the more recent generations of selection. In a similar manner, WOODWORTH *et al.* (11) made general predictions of response after the 50th generation of selection.

For the present paper, comparisons of predicted and actual responses to selection have been made for three types of prediction, each of which will be discussed separately:

(1). Predictions by previous workers, made by extrapolation of regression lines.
(2). Computation of expected progress by a "selection response" formula.
(3). Prediction after subdivision of characters into components.

3. PREVIOUS PREDICTIONS

The curves fitted by EAST (3) indicated rapid response in all four selected strains during early generations of selection, followed by a decreasing rate of progress. This led EAST to the conclusion that phenotypic levels in all four strains would rapidly become stabilized. WINTER (10) pointed out that the actual results in all strains except "Low Protein" were quite different, since significant further responses were obtained in "High Protein", "High Oil", and "Low Oil" between the 10th and the 28th generations of selection. In fact, response in "High Oil" was greater after the 10th generation than before. The "Low Protein" strain, WINTER pointed out, showed little change in protein content between the 7th and 28th generations of selection.

WINTER suggested that there was no significant difference between the protein content of the "source ear" in the original population and the mean protein content of the strain after 28 generations of selection. Therefore, of EAST's four predictions, three were proven incorrect and only one was substantiated by actual responses during the succeeding 18 generations of selection.

WINTER considered that continued rapid response to selection was likely in the "High Protein" and "High Oil" strains. Pointing out an apparent physiological limit to response in "Low Oil" imposed by a relation between germ size and oil content, he concluded that further response to selection in this strain was approaching a limit. Since "Low Protein" had shown no response to selection for more than 20 generations, further response in this strain appeared unlikely.

Studies of response during 50 generations of selection by WOODWORTH *et al.* (11) indicated that at least two of the four predictions by WINTER were incorrect. In the "High Oil" strain, response followed the expected upward trend. The response of "Low Oil" was evaluated as indicating little progress between the 30th and 50th generations of selection, thus apparently substantiating the prediction of WINTER that a selection limit had been reached in this strain. Examination of the data for these generations indicates that this conclusion may not have been justified, since the mean oil content of "Low Oil" declined from about 1.4 percent. in the 30th generation to about 1.0 percent. in the 50th generation of selection. In the two "protein" strains completely unexpected results were obtained. The "High Protein" strain, after nearly 40 generations of apparently consistent response to selection, suddenly ceased to increase in mean protein content. In fact, response between the 38th and 50th generations appeared to be in the direction *opposite* to that for which selection was exercised. The "Low Protein" strain, previously non-responsive to many generations of selection, showed rapid and relatively consistent response from the 26th to the 50th generation.

WOODWORTH *et al.* therefore suggested that further response could be expected in "High Oil" and "Low Protein", but not in the "High Protein" and "Low Oil" strains. The previous close fit of the "High Oil" response data to a linear regression function indicated that oil content in this strain could be expected to increase at a rate of about .18 percent. per generation of selection. An expected decrease of about .07 percent. protein per generation was indicated in "Low Protein".

Table 1 compares the approximate expected levels of chemical composition, predicted from the WOODWORTH *et al.* evaluation of the 50-generation data, with the actual levels obtained in the 59th, 60th and 61st generations of selection. None of the four strains has responded in the expected manner. Response of the "Low Protein" and "High Oil" strains has been significantly less than expected, while both "High Protein" and "Low Oil" have shown significant responses where no response was anticipated. Thus, each of the four predictions made after 50 generations of selection was found to be incorrect when the next 10 generations of selection actually were carried out.

Thus, a total of 12 predictions were made by extrapolation of response trend lines. Of these, only two or possibly three were substantiated by actual response at the next evaluation. *None of the four selected strains was considered to have shown consistently predicatable response over the entire period of selection.*

TABLE 1. PREDICTED AND ACTUAL RESPONSES TO SELECTION FOR PROTEIN AND OIL CONTENT IN MAIZE, BETWEEN 51ST AND 60TH GENERATIONS

Selected strain	Mean % protein or oil		Selection response, 51st–60th generation		
	Predicted	Actual†	Predicted	Actual†	Difference
High Protein	19.50	22.84	0	+ 3.34	3.34**
Low Protein	4.45	4.96	– .70	– .19	– .51**
High Oil	16.10	14.83	1.80	+ .53	–1.27**
Low Oil	1.00	.77	0	– .23	.23**

† – Means of 59th, 60th and 61st generations of selection
** – Highly significant

4. COMPUTED SELECTION RESPONSE PER DECADE

FALCONER (4) has given formulae for computing expected selection responses from heritability estimates, selection intensity constants, and the phenotypic standard deviations of the character under selection. The most useful formula for this purpose is considered to be:

$$R = i\sigma h^2$$

where

R = expected response per generation
i = selection intensity (1.4 throughout in this study)
σ = standard deviation of the character mean
h^2 = estimated heritability

FALCONER has pointed out that, strictly speaking, response to selection is predictable only from one generation to the next immediately following, since heritability is expected to change under selection, and since the variance may also change. Obviously, such predictions would be of little practical use in a planned breeding program, even if highly accurate. The anticipated usefulness of such a formula stems, according to FALCONER, from the assumption that selection response will be constant over a period of 5 to 10 generations.

Following this reasoning, selection responses were computed by the above formula for each decade (10 consecutive generations) of selection foı chemical composition. The estimates of response for the first decade of selection were made from data obtained in the first two generations of selection, except in the "Low Oil" strain, where the first three generations were used because of an inconsistently extreme response in the second generation. The selection intensity remained constant throughout the entire experiment. Phenotypic standard deviations for each generation of each strain were available from previous computations (10, 11). Difficulty was encountered in obtaining valid estimates of heritability. As a result, "realized heritabilities" were used, as computed from the formula

$$h^2 = \frac{R}{S},$$

where h^2 = realized heritability
R = response per generation of selection
S = selection differential

The resultant values of h^2 were substituted in the selection response formula to obtain predicted average responses per generation for the next decade of selection. The "realized heritability" values thus employed were in general much lower than the estimates of heritability for oil content in maize, given by ALEXANDER (1) and CROWLEY (2) as being in the range 50 to 75 percent.

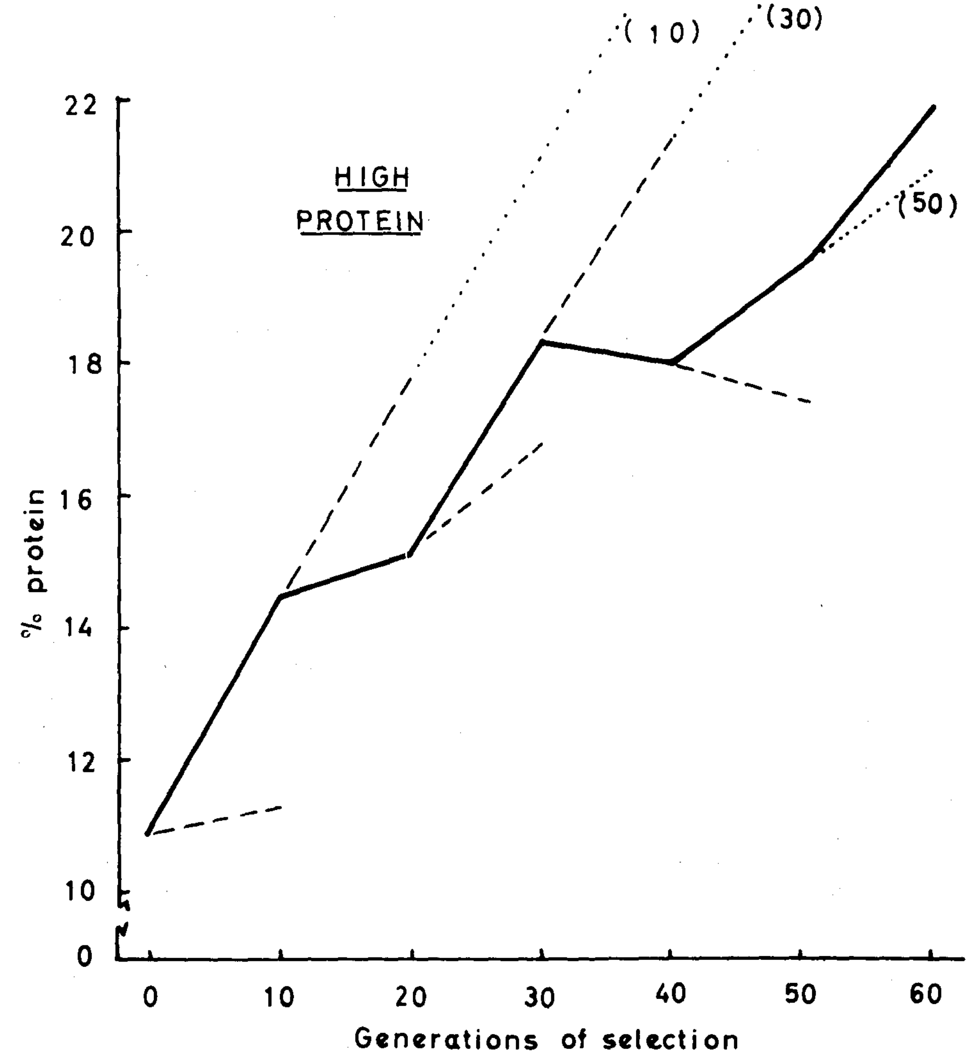

FIG. 1. RESPONSE IN "HIGH PROTEIN"

FIGURES 1–4. ACTUAL RESPONSES TO SELECTION FOR CHEMICAL COMPOSITION COMPARED WITH RESPONSES PREDICTED FROM SELECTION RESPONSE FORMULA. Solid line = actual response; broken line = prediction for 10 generations; dotted line = extrapolated predictions after 10th, 30th and 50th generations.

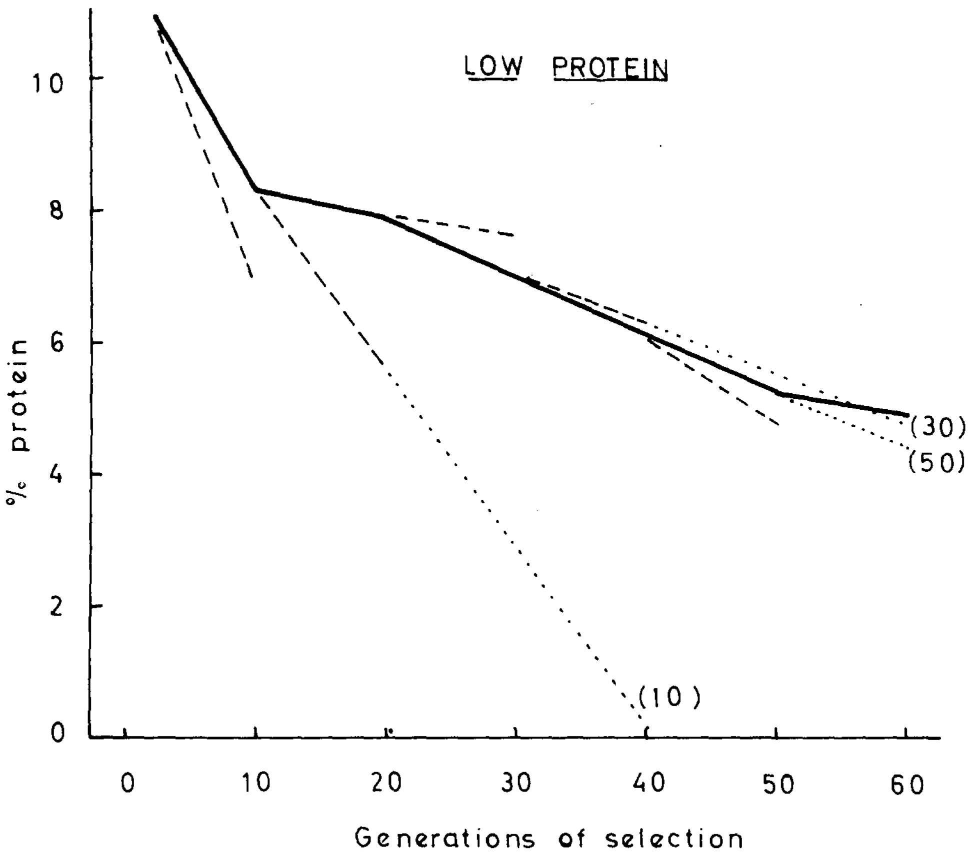

FIG. 2. RESPONSE IN "LOW PROTEIN"

Actual and predicted responses in the four selected strains are compared graphically in Figures 1, 2, 3 and 4. In plotting the actual responses, extreme fluctuations attributable to very unusual seasonal conditions have been eliminated, and the phenotypic levels are shown as means of three generations at the end of each decade of selection. Predictions after the 10th, 30th and 50th generations of selection were extrapolated to facilitate comparisons with the predictions actually made by EAST (3), WINTER (10), and WOODWORTH *et al.* (11).

Predictions in the "High Protein" strain were significantly incorrect in every decade, and showed no consistent direction of error. Extrapolation of the predictions would have led to major under-estimates of future response if made at the beginning of the experiment, or after the 40th or 50th generations, and to very large over-estimates after the 10th or 30th generations of selection. A reasonably close agreement with actual response in the 40th to 60th generations of selection could have been obtained only from extrapolation of predicted response after the 20th generation of selection, and this extrapolation did not fit the actual response between the 20th and 40th generations.

Major over-estimates of response in four of six computations resulted when the selection response formula was applied to the "Low Protein" data. Especially serious

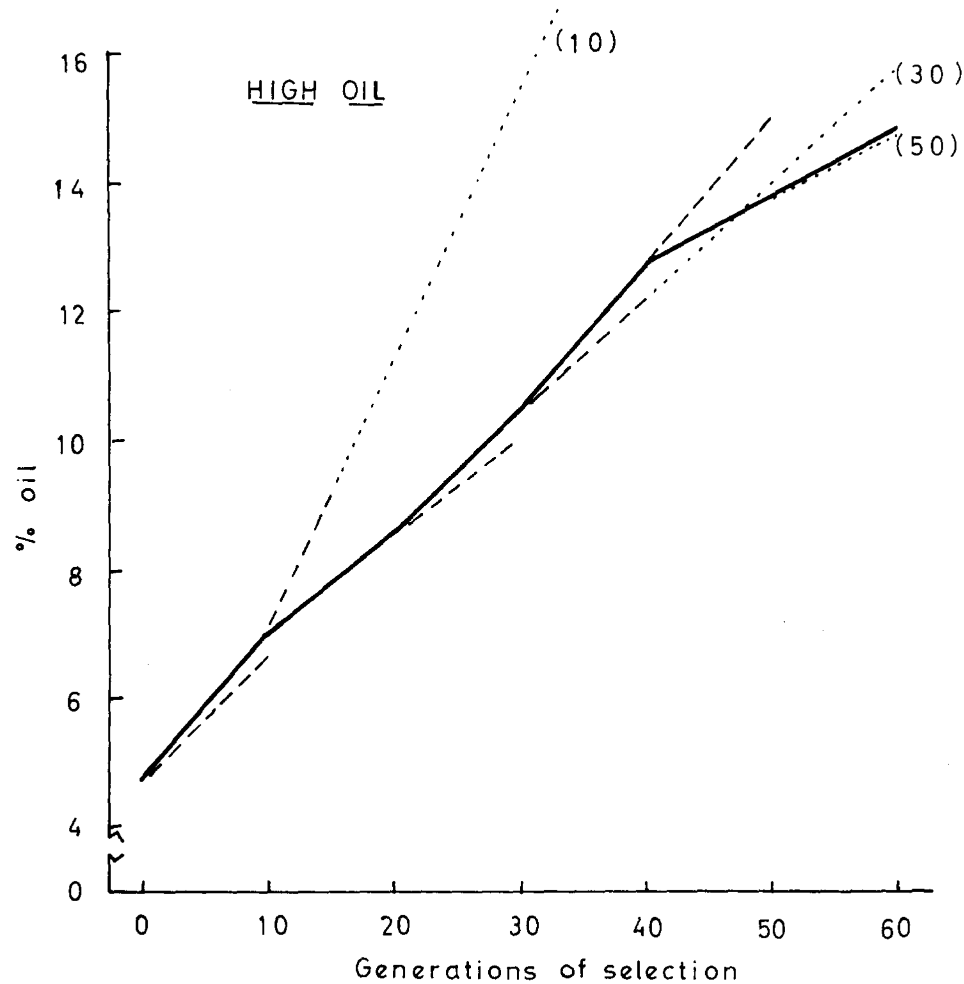

FIG. 3. RESPONSE IN "HIGH OIL"

over-estimates of response were made from the data obtained in the early generations of selection (Figure 2). The predicted response after 20 generations of selection proved to be significantly under-estimated, and only the prediction after 30 generations of selection was in close agreement with the actual response to subsequent selection. It is interesting to note in this connection that the estimate made by WINTER (10) after 28 generations of selection predicted no further progress in this strain, while the "selection response" formula, as applied here, closely reflected the response which was subsequently observed.

Agreement between predicted and actual responses to selection was better in the "High Oil" strain than in the other material. Significant over-estimates of response were made from the data after 10, 30 and 40 generations of selection, while the estimates based on the beginning of the experiment, and the second and fifth decades of selection, were reasonably accurate. The predictions based on the first and fourth

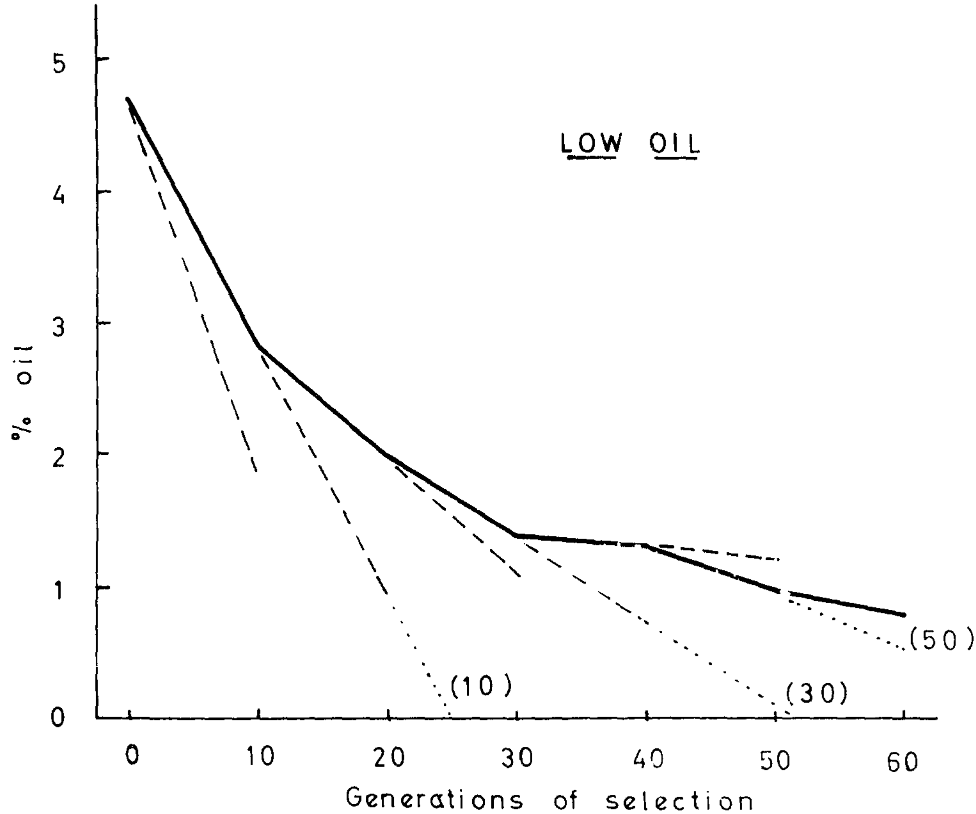

FIG. 4. RESPONSE IN "LOW OIL"

decades of selection were the least accurate, and also agreed the least with predictions made by the other methods discussed in this paper.

All predictions made from the selection response formula in the "Low Oil" strain, except that after 40 generations of selection, were found to be serious over-estimates when compared with the actual responses to selection. All of the predictions based on the first three decades of selection indicated a rate of response which, if extrapolated, would have caused the mean oil content in this strain to reach the level of zero within the subsequent 10 to 20 generations of selection. The response predicted after the 40th generation was significantly underestimated, as compared with actual responses during each of the next two decades of selection. Thus, none of the six predictions made in this strain showed even reasonably close agreement with the actual responses.

Of a total of 24 predictions made by application of the formula, 14 were significant overestimates, 5 were significant underestimates, and 5 were reasonably accurate when compared with actual responses. Only in the "High Oil" strain was more than one of the six possible predictions found to be accurate. Had the higher heritability estimates for oil content, reported by other workers, been used instead of the "realized heritability" values, all of the predictions in "High Oil" would have been gross overestimates. It appears that the retention of variability under selection (discussed more fully by LENG, 7) led in this material to standard deviation values which indicated a higher potential for response than actually was present. Also, in all strains except "High Oil", the estimates of realized heritability apparently gave unsatisfactory bases for the computation of future selection response. As a general conclusion, it appears unlikely that direct application of the selection response formula employed here will lead to useful predictions of response over even as many as 10 generations of selection. It also appears to give no satisfactory basis for predicting selection limits, even when applied after many generations of selection.

5. Predictions based on component analysis

It has previously been shown (7, 8) that subdivision of the character "oil content" into its two major components is helpful in interpreting apparent inconsistencies in response to selection in the two "oil" strains. Although the available data are not extensive, they appear also to be of interest in connection with problems of response prediction discussed in this paper.

Determinations of the two major components of oil content in maize grain – *germ size* and *oil percentage in the germ* – were made at the beginning of the selection experiment by HOPKINS (5) and again after the 6th and 40th generations of selection. For purposes of the present study, changes in these components between the 6th and 40th generations of selection were used to compute an average rate of response per generation. The amount of oil found in parts of the kernel outside the germ were estimated from the 40th generation data. Total oil content in the 60th generation was then predicted from the computed response rates of the two components and the estimated amount of oil outside the germ. The data and method of calculating predicted response are shown in Table 2; Table 3 compares the predictions made by this method with those made from the "selection response" formula after the 40th generation of selection, and with the mean oil percentages actually obtained in the 59th, 60th and 61st generations of selection.

TABLE 2. COMPUTATION OF PREDICTED OIL CONTENT IN 60TH GENERATION FROM RESPONSE TRENDS OF MAJOR PHENOTYPIC COMPONENTS

Character	Observed responses			Predicted, 60th generation
	6th generation	40th generation	Change per generation	
		"High Oil"		
% oil in germ	42.0	47.0	.15	50.0
Germ size, %	14.0	21.5	.22	26.0
Product, % × size	5.9	10.1	–	13.0
Oil outside germ, %	.5	1.9	–	1.9
Total oil % in grain	**6.5**	**12.0**	–	**14.9**
		"Low Oil"		
% oil in germ	30.0	15.0	.44	6.2
Germ size, %	8.0	7.5	.01	7.3
Product, % × size	2.4	1.1	–	.45
Oil outside germ, %	.6	.2	–	.20
Total oil % in grain	**3.0**	**1.3**	–	**.65**

TABLE 3. PREDICTED AND ACTUAL SELECTION RESPONSES IN "HIGH OIL" AND "LOW OIL", 40TH TO 60TH GENERATION

Strain	Predicted % oil		Actual oil %, generations 59–60–61
	"Selection response"	"Component analysis"	
"High Oil"	17.2	14.9	14.8
"Low Oil"	1.20	.65	.77

In the "High Oil" strain, total oil content after 60 generations was predicted almost exactly from the component data available in the 40th generation of selection, while the selection response formula produced a marked overestimate of response. Response to selection in "Low Oil" was slightly overestimated by the component analysis, but seriously underestimated by the "selection response" formula prediction.

From the limited comparisons possible, it is clear that more accurate predictions of selection response in both "oil" strains could be made by analysis of phenotypic components than from treatment of the character "oil content" as a single entity. The component analysis is especially useful in attempts to predict selection limits in these strains, since it reveals that the last 40 to 50 generations of selection for "oil content" actually has been effective primarily only in the *germ size* component in "High Oil" and almost exclusively in the *oil content of germ* component in the "Low Oil" strain. The possibilities of predicting future response and selection limits in this material will be discussed further below.

As yet, no satisfactory basis has been found for a phenotypic analysis of the character "protein content". Thus, no logical basis is apparent for improving predictability of response or selection limits in the "protein" strains.

6. DISCUSSION AND FURTHER PREDICTIONS

Experience over 60 generations of selection for chemical composition in this material makes it clear that meaningful predictions of response rates and selection limits can be made only if a biologically-sound analysis of the character under selection is possible. Simple extrapolation of regression trend lines for the phenotypic mean values, or computations involving heritability estimates and variance data, did not yield satisfactory predictions of response in these complex characters.

Thus, as pointed out above, there appears to be no suitable basis for predicting future response to selection in the "High Protein" and "Low Protein" strains. Past experience suggests in general that response in these strains, especially in "High Protein", is likely to be erratic. To date, there is no indication that a selection limit has been reached in either strain, and no satisfactory evidence for estimating the level of protein content at which a selection limit might be expected.

In the "High Oil" strain, response in the *oil content of germ* component was relatively minor between the 6th and 40th generations of selection. The slow rate of response in this component, after initially rather rapid response, and comparison with oil content found in highly-selected oil-seed species such as the sunflower, soybeans, and others, indicates a probable selection limit in this component may be reached in the vicinity of 50 percent. The *germ size* component in this strain has continued to respond to selection for total oil content, at a rate of about .2 percent. per generation. Germ size may be affected by changes in absolute size of the germ itself, or by changes in the germ-endosperm size relation. Thus, there appears to be considerable latitude for response in this component. At the present rate of response, even assuming that oil content outside the germ does not increase further, and that oil content of the germ becomes stabilized at about 50 percent., a further increase in oil content at a rate of about 1.0 percent. per 10 generations of selection may be anticipated. A selection limit for total oil content is not likely to be reached until the genetic variability for the germ-endosperm ratio is exhausted. Fragmentary evidence (unpublished) suggests that such a limit probably will lie above 25 percent. total oil content.

For all practical purposes, a selection limit in the *germ size* component was reached in the "Low Oil" strain in the first six generations of selection, at a level between 7.0 and 7.5 percent. germ, by weight. Virtually all subsequent response to selection in this strain has been in the component *oil content of germ*, which decreased at a rate of approximately .4 percent. per year for more than 40 generations. Oil outside the germ has been at a very low level for many generations. A selection limit in total oil content of "Low Oil" apparently will be reached when the lowest possible level of oil content in the germ is attained. Present evidence suggests that maize embryos can remain viable when the oil content of the germ is in the range 4 to 5 percent. Thus, the selection limit for total oil content in this strain can be expected to be at least as low as .4 or .5 percent. Sufficient genetic variability appears to remain in the "Low Oil" breeding stock (7) to permit such a response to further selection. Even at the present rate of response, 10 to 20 more years of selection appear necessary before such a limit can be anticipated.

References

1. Alexander, D. E., Some thoughts on breeding high oil maize. Hibridni Kukuruz Jugoslavii **3**, No. 12 (1960): 26–33. (Serbo-Croatian with an English summary).
2. Crowley, J. C., Efficient evaluation of corn inbred lines for combining in oil content. U. of Illinois M.S. thesis, 1959.
3. East, E. M., The role of selection in plant breeding. Pop. Sci. Mo. **77** (1910): 190–203.
4. Falconer, D. S., Introduction to quantitative genetics. Oliver and Boyd, London and Edinburgh (1960). 365 pp.
5. Hopkins, C. G., The chemistry of the corn kernel. Ill. Agr. Exp. Sta. Bul. **53** (1898): 127–180.
6. Hopkins, C. G., Improvement in the chemical composition of the corn kernel. Ill. Agr. Exp. Sta. Bul. **55** (1899): 205–240.
7. Leng, E. R., Long-term selection for chemical composition in maize and its significance for breeding theory. Z. f. Pflanzenzüchtung (in press).
8. Leng, E. R., and C. M. Woodworth, Effects of long-continued selection for oil and protein content in maize. Genetics **38** (1953): 675.
9. Smith, Louie H., Ten generations of corn breeding. Ill. Agr. Exp. Sta. Bul. **128** (1908): 454–575.
10. Winter, F. L., The mean and variability as affected by continuous selection for composition in corn. Jour. Agr. Res. **39** (1929): 451–476.
11. Woodworth, C. M., E. R. Leng and R. W. Jugenheimer, Fifty generations of selection for protein and oil in corn. Agron. Jour. **44** (1952): 60–65.