

THE THEORY OF PLANT BREEDING FOR YIELD

By OTTO H. FRANKEL

Wheat Research Institute, Christchurch, New Zealand

Received 16.viii.46

IN the evolution of cultivated plants, the evolution of productivity has been the predominant factor. It is a generalisation which applies to many crops that, with regard to quality factors, the gap between our cultivated plants and their wild or weed ancestors is smaller in degree and in complexity than it is with regard to yield.

In both directions plant breeders have made spectacular advances in recent years, assisted to a large degree by a fuller understanding of the factors involved. When it is now felt that in various crops the limit of advance has been reached, a close examination of the bases of plant breeding is called for. This paper is chiefly concerned with the problem of yield. The bases of plant breeding for yield are not, however, peculiar to yield alone.

The main increases in yield have been, and are being, achieved by overcoming limiting factors whose effects can be distinguished with a fair degree of certainty, rather than by assembling productivity genes, although the latter process is likely to have accompanied the former. It is well known that the spectacular advances of recent years are all of the former order. The introduction of rust resistant wheats in Western Canada alone are estimated to have raised the wheat production by 41,000,000 bushels annually (Craigie, 1944). Other characters introduced by breeding, though not of equal significance with regard to yield, may be of great agronomic importance. Combine harvesting, for example, would have been impossible in New Zealand without varieties with a high degree of resistance to shattering. Characters of this nature are in the main "observable." The question is raised whether, once selection against these limiting factors has reached its limit, further advance is possible and practicable.

GENERAL

In the first chapter of his book, *Theoretical Bases of Plant Breeding*, Vavilov (1935) outlines the evolution of plant breeding from an art, as old as man himself, through an industry, into a science. It is a "complex science"; for although scientific plant breeding springs from the study of heredity and evolution—Vavilov cites Darwin, Mendel, Johannsen and de Vries as its chief founders—it draws on many other sciences.

The task of plant breeding is the production of improved cultivated

plants. Progress in breeding may result from the application of purely empirical methods, but "to ensure certainty, speed, and pre-eminence in the essential tasks" (Vavilov *loc. cit.*) it is necessary to replace empiricism by scientific method.

To be of value, a theory of plant breeding must be composite, since each phase of the breeding process is to be related to its scientific bases in various fields. According to Vavilov, the following six sections constitute this complex science :—

1. The study of the potentials of original varieties, species and breeds (the phytogeographical basis of breeding).
2. The study of inherited variability (laws governing variation, the study of mutations).
3. The study of the effect of environment on the emergence of varietal characters (variety and environment, the influence of particular factors of environment, the study of the stages of plant development and its application to breeding).
4. The theory of hybridisation of closely and remotely related forms.
5. The theory underlying breeding practice (self- and cross-pollination, vegetative and apogamous reproduction of plants).
6. The study of the chief aims of breeding : immunity to disease, certain physiological characters (winter-hardiness, resistance to drought, photoperiodism), properties useful for industrial purposes and chemical composition.

On these foundations the following are the main phases of scientific plant breeding :—

1. *Collection and classification.*—The collection, observation, description and systematic classification of original material.
2. *Inter-specific.*—The study of the origin, evolution, relationship and crossing behaviour of major systematic units.
3. *Intra-specific.*—The study of the nature, origin, variability (genetic and environmental) and inheritance of varietal differences.
4. *Selection and testing.*—The study of the effects of selection, as defined by system of propagation, nature of variability, and reliability of estimation ; the study of tests of the significance of selection.

These four sections are by no means distinct, with theoretical bases of their own ; on the contrary, they are closely interrelated, each serving to build up the other, and all based on the same foundations :—

Biological.—Plant geography, anatomy, physiology, genetics, cytology, ecology, phytopathology.

Physical.—Biochemistry, biophysics.

Mathematical.—Statistics.

Applied.—Agronomy, technology.

Vavilov's work to a large degree has provided the "modern synthesis," which we may now term the science of plant breeding. In this essay an attempt is made to see how far it goes towards providing the theoretical bases for the four phases of plant breeding, and where and how these bases could be further utilised to raise the productivity of cultivated plants.

FIRST PHASE : COLLECTION AND CLASSIFICATION

The explorations and systematic observations of Vavilov's Institute of Plant Industry have yielded a phytogeographical theory of the evolution of cultivated plants, which was recently summarised and further developed by Darlington (Introduction to Darlington and Janaki Ammal, 1945). The "centres of development," which Vavilov discovered, have yielded essential knowledge and equally essential plant material. The two contributions are closely connected ; for once we know where the greatest range of variation occurs, we can collect the material needed for constructive breeding. Again in setting comprehensive and precise standards of observation and description (*cf.* for example Vavilov, 1922-23) under carefully selected and representative ecological conditions, the Russian school not only discovered new approaches towards a phylogenetic classification, but enabled its plant breeders to select from the almost infinite number and range of variants those needed, and best suited, for their respective breeding tasks.

The theoretical basis for the first phase of scientific plant breeding has therefore been provided. It facilitates two measures : (a) Comprehensive collections of cultivated plants, representing the existing range of variation ; and (b) comprehensive descriptions and classifications of this material, based on morphological, physiological, genetic, cytological, ecological and phytopathological observations and measurements.

Outside the Soviet Union, work as comprehensive as that recently conducted on the Empire Potato Collection (Hawkes, 1941, 1944) is still the exception rather than the rule. Essentially this phase of scientific plant breeding presents new problems of organisation rather than of theory or method. It is obvious that the more widespread the search, the more comprehensive plant collections will be ; that the greater the geographical and ecological range under which observations are made, the greater is their significance ; and that the more searching the description and classification, the more useful it will prove to the plant breeder. Clearly this assembly and survey of world resources in plant variation require a world effort. The magnitude as much as the nature of the task points to the need of international co-operation and co-ordination.

SECOND PHASE : SPECIES FORMATION

If comprehensive description and classification of the full range of variation is the first basis of scientific plant breeding, the knowledge of the modes of origin of discrete systematic units is the second. The cytological and genetic processes of speciation in cultivated plants have now been classified in many instances, and Crane (1940) and recently Darlington (*loc. cit.*) have classified them broadly, according to the nature of the change, into four major groups :—

1. Selection of mutations, old or new.
2. Auto-polyploidy.
3. Crossing of species, without doubling.
4. Crossing of species, with one-sided or two-sided doubling.

The consequences of these changes are manifold. In the first instance, where crosses are involved, that is in groups 3 and 4, chromosome complements are joined whose mutual balance is untested (Mather, 1943) ; yet the combination may provide a novel basis for evolution. Further, as Darlington points out, variability may be released whose amount depends not only on the nature of the cross, that is whether doubling occurs on two sides, one side, or not at all, but also on the internal organisation of the parental chromosomes and on their mutual relationships.

Finally, as a result of such changes, there will be found varying degrees of adaptation to the environment and of approximation to the breeder's specific aim, depending on the nature of the change and of the material, and on the system of propagation. A cross between species which normally are asexually propagated, may lead to a well-adapted, directly usable form. On the other hand, a cross of seed-propagated forms—especially after chromosome doubling—is likely to yield a more or less narrow range of uncertain adaptation, whose principal, though potentially important, advantages are the possession of a new balance, of new gene combinations, and as stated previously, in some instances of newly released variability. Whatever the relationship of the parents or the nature of the change, initially the new form or forms lack that breadth of variation which older species, thanks to mutation, hybridisation and selection, have built up in time.

Two methods are available for broadening the range of variation of such new products : a multiplicity of crosses between varieties of the same parent species, with the further chance of intercrossing their hybrids, and a series of multiple crosses, using the new hybrid in combination with related forms. Both approaches have been used (for example, Müntzing, 1939 ; Zhebrak, 1944 *a* and *b*, 1946) ; both are laborious. They are likely to yield a wider range of variants than would be adapted to, or useful in, any one locality. Hence much that is discarded in one place may be of the utmost value in another. It seems that for their full and rational exploitation such projects call for the widest measure of co-operation.

THIRD PHASE : INTRA-SPECIFIC EVOLUTION

Plant breeding as we know it to-day has its main roots in two scientific discoveries : those of Vilmorin and Johannsen and of Gregor Mendel. The former was exhausted as a source of further progress once the full implications of the distinctness of hereditary and environmental components of development were understood and applied. Mendelian theory, on the other hand, is fundamental to every concept of modern genetic theory. Its developments have affected every sphere of general and applied genetics. With its complementary theories of linkage and gene mutation, mendelism supplies the principal genetic foundation of the neo-Darwinian theory of evolution (*cf.* Huxley, 1942) ; and since, to use Vavilov's phrase, plant breeding is " evolution directed by the will of man," mendelism is the theoretical basis of the directed evolution of cultivated plants. This applies particularly to that phase which has been—and must remain—the main field of plant breeding, that is, the breeding and selection within genetic species.

The value of mendelian theory and its developments, in its application to plant breeding, has been twofold :—

- (a) It has facilitated precise information on the mode of inheritance and on the genetic relationships of a great many characters of economic significance ; in the main such information is confined to characters which are based on a relatively small number of genes with readily observable effects.
- (b) It has provided the genetic background for breeding work on characters whose nature or whose inheritance is complex and to which therefore a simple mendelian type of inheritance cannot apply.

The distinction between complex characters and characters with complex heredity is arbitrary and provisional, yet it may facilitate a useful approach where, by the resolution of a " character " into components, the latter present a clearer genetic picture than does the complex character itself. Yield of grain in wheat, for example, can be resolved into number of ears per plant, number of grains per ear and weight of grain. Some of these may be subject to further resolution, such as the number of grains per ear into number of spikelets and number of grains per spikelet. Such analytical steps would tend to simplify, or even make possible, a genetic analysis.

Many of the characters of interest to the plant breeder, and, in particular, yield characters, are determined both by the genotype of the plant and the environment in which it is grown. These two factors interact so that such characters " exhibit a continuous range of variability in segregating populations " (Smith, 1944). The resolution of this variation into environmental and hereditary components, and the discovery of the nature, organisation and action

of the latter, constitute the main theoretical problems of breeding for yield.

There is general agreement that quantitative characters are based on a number of genes, estimates varying from a few to a few hundred, though in some instances the larger numbers seem unjustified (Boyce, 1946). The majority of authors regard these genes as equivalent in nature and organisation to those which determine qualitative characters. Mather (1941, 1943, 1944), on the other hand, considers the former—the polygenes—can exist as distinct from the other genes in location, constitution, organisation and function. According to his theory the polygenes are partly located in the heterochromatin and are less specific than the euchromatic oligogenes. They are organised in linked groups which, through a system of balance which is adapted to the breeding system and subject to selection, achieves optimal adaptation whilst retaining genetic flexibility. This theory is of great significance in its application to breeding for yield characters. It requires experimental tests on cultivated plants with a variety of breeding systems.

FOURTH PHASE : SELECTION AND TESTING

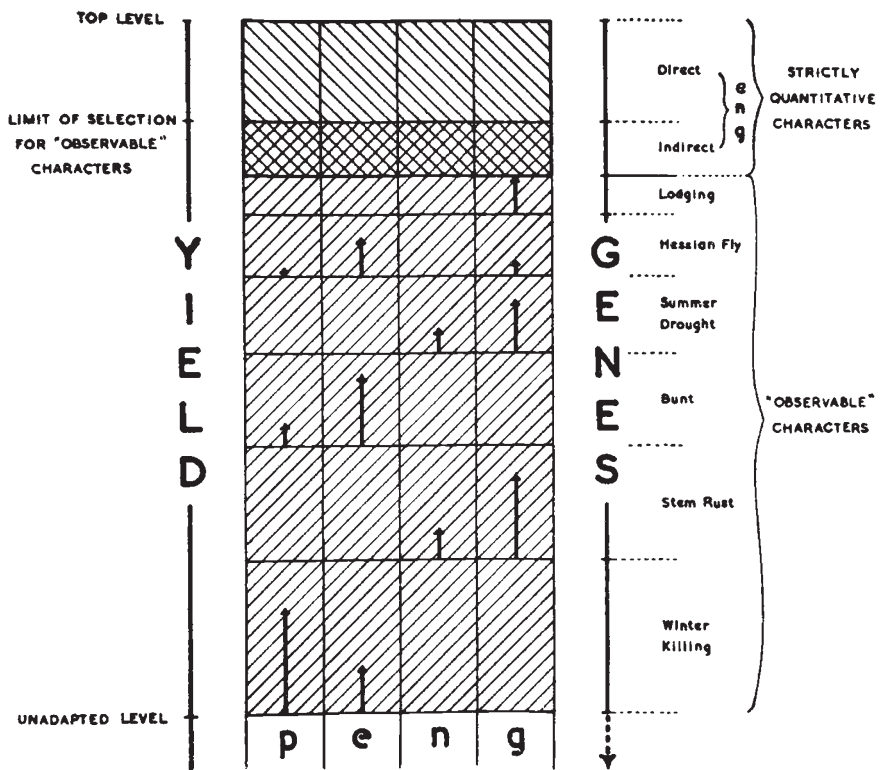
Selection in plant breeding is concerned with the identification of individuals, or groups of individuals, which possess inherited characteristics rendering them superior to the unselected population. Other things being equal, the effectiveness of selection is determined by (i) the number and linkage relations of the genes concerned, and (ii) their penetrance and expressivity under the conditions of selection. These conditions must include the induced environment artificially produced by the breeder to increase penetrance, such as infection with disease agents, freezing, drought, inducement of lodging in cereals, etc. According to this classification, yield characters may be divided into two groups :—

- (a) Qualitative characters based on a relatively small number of genes with a frequency and degree of expression facilitating their recognition, with a reasonable degree of certainty, in single plants and their progenies. Examples are bunt resistance in wheat and crown-rust resistance in oats (Hayes and Immer, 1942).
- (b) Quantitative characters presenting a continuous range of variation, and strongly subject to environmental influence. Examples are weight in tomatoes (Powers, 1942) and in water melons (Poole and Grimball, 1945).

Admittedly this division of yield characters is arbitrary. It is, however, useful in an attempt to analyse the general problem of breeding for yield. This may be shown on a diagrammatic example illustrating the progress of selection. It is assumed that a cultivated plant—in this example a self-fertilising plant, wheat, is used for

simplicity—is introduced into a newly settled area ; or, as an alternative, a newly formed species, with a breadth of variation which permits adaptation, is exposed to selective breeding. Then the yield of that variant within the species which, prior to the commencement of adaptive selection, gives the highest yield, determines the “unadapted level of yield.” It is further assumed that it is possible to assemble, by successive matings and selection, all the + genes of observable characters which are contained in the material ; and hence that the environment, including induced environment, is suited for the recognition of these characters.

Finally, wheat yields under the given conditions are limited, in descending order of magnitude, by winter killing, stem rust, bunt, late-summer drought, Hessian fly and lodging. Wheat yield is shown as determined by its four components, those which are primarily affected by selection being indicated by arrows.



Diagrammatic description of the effect of selection for “observable” characters on yield and its components.

p = number of plants per area *e* = number of ears per plant
n = number of grains per ear *g* = weight of a single grain

As the diagram indicates, resistance to the limiting factors, all of which are more or less observable, is successively built up to the highest degree permitted by the genic complement of the species.

This can be widened by introducing genes from related species—for example, stem-rust resistance from tetraploid into hexaploid species of *Triticum*—at the same time retaining by selection the major characteristics of, and infertility with, the original species. An interspecific cross resulting in a new systematic unit, however, creates a new specific “unadapted level.”

Once the limit of observable selection is attained, further advancement in yield can only be achieved by increases of e , n or g . These yield components are considered here in a purely formal sense. Obviously there may be genes determining the development of any of these individually; on the other hand, genes are likely to control them jointly, by determining factors of the plant metabolism. In the absence of precise knowledge of the physiological processes, we are concerned solely with their results in terms of e , n , g .

The situation as described above is by no means a purely academic problem. It exists in a number of crops, for example in sugar-beet in Europe and in wheat in New Zealand. Wherever it occurs it raises the question whether further breeding is possible or worth while. It therefore demands a close examination of the genetic situation, so far as it is known or deducible, of methods of breeding and selection in use, and of possible improvements in technique.

It must be realised that when this stage is reached, selection for e , n , g is not at its start. In fact it is likely to have proceeded throughout the breeding process, for the following reasons:—

- (a) The aim of improving the yield is ever present, even where observable characters are the principal object of selection. Whether selection for yield itself, that is for e , n , g , is possible in single plants and their immediate progenies will, however, be discussed later.
- (b) In practically every breeding project connected with yield, it is necessary at some stage to conduct tests which are subject to analysis by statistical methods. Even though resistance to one of the limiting factors be the main objective, there is a likelihood of segregation of genes for e , n , g , and scope for selection among the lines which prove resistant to the limiting factor. It must be remembered, however, that at this stage the population in which selection takes place is of necessity limited in size.
- (c) Morphological or physiological characteristics which determine resistance to limiting factors may indirectly act on e , n , g . For example, an enlarged root system, or increased resorption of soil moisture, may not only affect drought resistance, but, through increased resorption of soil nutrients, also e , n , g .

In consequence, the yield components e , n , g may have been raised by indirect selection beyond the level secured for them by the maximum removal of observable limiting factors.

At this point the major question arises : Is it possible to select directly for e , n , g ? The answer depends on the relative values of genetic and environmental variances, that is, on the presence and magnitude of genetic variation, and on the degree to which environmental factors render genetic variation unrecognisable. It is thus necessary to examine whether our provisional classification of yield components proper as "unobservable" in single plants and their immediate progenies,* that is, in the first phase of the selection process, is justified. The implication is of great importance. For should e , n , g prove unrecognisable at this stage, serious restrictions would be imposed on the progress of selection. In this first phase, selection of single plants usually takes place in any generation after the cross, with or without further re-selection, both depending on the system of propagation and on the breeding objectives. The main feature of this phase is the facility it affords for selecting in large populations. The second phase, which commences once numbers have been suitably reduced in the first, is characterised by superior precision, for at this stage replicated tests yield results to which statistical methods can be applied ; but populations, for reasons of space and working capacity, must be limited in size.

It should be borne in mind that, on general grounds, yield components must be expected to depend on a large number of genes. Poole and Grimball (1945) recently found the weight differences between two varieties of water melons based on 12 to 13 genes. Whilst in this case a plant homozygous for all of these could be expected, in the absence of linkage, only once in 17,000,000, even one homozygous for three genes could be expected to be found with a fair degree of reliability only in a population of 200 individuals (Mather, 1938). This consideration emphasises the relevance of large numbers. It also emphasises implicitly the relevance of "observability" in the early stages of the breeding process ; for, if at those stages yield differences due to e , n , g are not detectable, if in fact selection of single plants and their immediate progenies is merely a random selection, we are relegated to selection by means of replicated yield tests open to statistical treatment. There are, however, limitations to the numbers which can be taken into such tests ; as a rule, population size is below that demanded by genetical considerations.

A similar situation occurs in breeding for quality characteristics when methods for determining these require relatively large amounts of material, or where the character is strongly influenced by the environment. This is the case in baking quality of wheat, especially where selection takes place at higher quality levels. Here the application of small-scale tests is unreliable or even misleading (Frankel, 1940). Under New Zealand conditions, selection cannot

* In plants which produce large quantities of seed, replicated trials are possible in the immediate progenies of single plants, for example in cotton (Hutchinson, 1940).

therefore commence until the yield trial stage, when sufficient grain is available for baking and other tests requiring several pounds of wheat or flour. Of course there is always a chance, however small, that the lines first tested in yield trials may contain the rare high-quality genotype. For example, a variety was recently selected with a baking quality and a protein content exceeding those of both parents, Cross 7 and Tainui, and even of the Canadian quality standard, Marquis (Frankel, unpublished).

To both phases of selection apply considerations of the magnitude of genetic variation to be expected, as distinct from frequencies of \pm variants in the population. In many cases, where yield has reached an apparent limit by selection for observable characters, further mean increases in yield of, say, 5 per cent. or less, may still be attainable and of economic value. The standard deviation in extensive lattice yield trials in New Zealand in 1945-46 varied between 6.8 per cent. and 9.2 per cent. of the mean. Increases sufficiently large to be statistically significant can be expected only when the environment is especially favourable for a particular variant. They are therefore of a statistical nature; they occur with a certain frequency and a certain average magnitude, though the latter may be below the limit of statistical significance. Thus, although such increases are both real and desirable, their discernment in replicated trials demands seasonal and regional repetition of large test series. That differences of this nature could not be ascertained in single plants or plots is hardly open to doubt. This view is shared on the whole by Hayes and Immer (1942, p. 34), although these authors believe that "when homozygous lines are available, those that yield most satisfactorily can be isolated through actual comparative-yield trials."

This discussion has been focussed on circumstances which primarily apply to self-fertilised plants, although essentially the problem is universal. Yet in cross-fertilised plants, where the polygenic balance is relational, methods are available for raising yields, beyond the "top level of observable limiting characters," which are not based on the selection of individual plants. These methods are applicable where mating can be controlled either mechanically, as in maize, or biologically, as in self-sterile grasses. The biological and ecological advantages of line mixtures rather than pure lines have been stressed by Harland (1944). It is doubtful whether these methods can be applied in self-fertilised plants. Hybrid mixtures resulting from crosses of related varieties might be considered (Ellerton, personal communication; Frankel and Copp, unpublished). The success of mechanical mixtures of pure lines will depend on their ecological interactions. An extensive experiment on wheat blends did not succeed in raising yields (Frankel, 1939).

Is, then, the answer to the main question, "Is it possible to select directly for yield itself" in the negative? Does the evolution of yield, "directed by the will of man," come to an end when the

factors which limit yield have been removed as far as the genic complement permits?

There are considerations which suggest a slightly more optimistic outlook. So long as genetic variation in yield components exists, there is always the possibility of selection, by chance, of lines with high yield, especially if large numbers are taken into yield trials. There are, in fact, examples of varieties with increased yields produced under the circumstances outlined above. In New Zealand two wheat varieties have been produced (Frankel, 1941, and unpublished) which improve on the previous standards in yield, although not in any observable limiting character. Doubtless there are other examples, especially where breeding has progressed for long periods and where there are no rigours of climate or severe pests; but close definition of the circumstances is required before a yield advance can be assessed as due to residual resistance to limiting observable factors, or to a direct increase of yield itself. However, the aim, as quoted earlier, is "to ensure certainty, speed and pre-eminence in the essential tasks," not to achieve success at random. Ellerton (1944) suggested on the basis of yield tests of F_1 hybrids between English wheats—for which, it is commonly believed, the limit of yield has been reached—that "wheat could be bred for British conditions with yield increases of 20 per cent. at least." If plant breeding becomes an exact science, predictions, such as Ellerton makes, should become generally possible and reliable.

If there are indications that "genetical limits" are merely apparent, "imposed by the shortcomings of current breeding methods" (Ellerton *loc. cit.*), the analysis of the progress of selection for yield places emphasis on investigations which "must aim primarily at increasing the precision with which the tools of the breeder—selection, adjustment of the mating system, and adjustment of population size—can be used to manipulate the heritable variation which constitutes his raw material" (Mather, 1944*b*). Such tools are now available through the application of statistical methods. As Mather (*loc. cit.*) suggests, "further progress in this field will depend on the close co-operation of geneticists, breeders and statisticians." The series introduced by this paper, by examining problems—of an essentially practical nature—involved in the selection for yield itself, is intended as a contribution towards this end.

SUMMARY

The paper discusses the theoretical bases of the science of plant breeding, especially in relation to breeding for yield in its several steps, collection and classification, species formation, intra-specific evolution and selection and testing. It develops the difference between selection for resistance to factors which limit yield and which are more or less "observable" in single plants and their immediate progenies, and typically quantitative factors which

determine yield itself. It raises the question whether selection for the latter is possible in single plants, and if this is not the case, whether selection in replicated yield tests is likely to succeed.

REFERENCES

- BOYCE, S. W. 1946. *Nature*, 157, 657.
- CRAIGIE, J. H. 1944. *Scientific Agric.*, 25, 51-64.
- CRANE, M. B. 1940. *The New Systematics*. London.
- DARLINGTON, C. D., AND AMMAL, E. K. JANAKI 1945. *Chromosome Atlas of Cultivated Plants*. London.
- ELLERTON, S. 1944. *Nature*, 153, 780-783.
- FRANKEL, O. H. 1939. *J. Agric. Sci.*, 29, 249-261.
- " 1940. *Ibid.*, 30, 98-112.
- " 1941. *N.Z. J. Sci. and Tech.*, 22, 303A-308A.
- HARLAND, S. C. 1944. *Institute of Cotton Genetics*, Lima, Peru. Bulletin No. 1.
- HAWKES, J. G. 1941. *Potato Collecting Expeditions in Mexico and South America*. Cambridge. Imp. Agri. Bureaux.
- " 1944. *Ibid.*
- HAYES, H. K., AND IMMER, F. R. 1942. *Methods of Plant Breeding*. New York.
- HUTCHINSON, J. B. 1940. *J. Genetics*, 40, 271-282.
- HUXLEY, J. S. 1942. *Evolution, the Modern Synthesis*. London.
- MATHER, K. 1938. *The Measurement of Linkage in Heredity*. London.
- " 1941. *J. Genetics*, 41, 159-193.
- " 1943. *Biological Reviews*, 18, 32-64.
- " 1944a. *Proc. Roy. Soc.*, B, 132, 308-332.
- " 1944b. *Nature*, 153, 780-783.
- MÜNTZING, A. 1939. *Hereditas*, 25, 387-430.
- POOLE, C. F., AND GRIMBALL, P. C. 1945. *J. Agric. Res.*, 71, 533-552.
- POWERS, LEROY 1942. *Genetics*, 27, 561-575.
- SMITH, H. H. 1944. *Bot. Review*, 10, 349-382.
- VAVILOV, N. I. 1922-23. *Bull. Appl. Bot.*, 13, 149-257.
- " 1935. *Theoretical Bases of Plant Breeding*. Moscow.
- ZHEBRAK, A. R. 1944a. *Acad. Agric. Moscow*, 6, 5-54.
- " 1944b. *Nature*, 153, 549.
- " 1946. *American Naturalist*, 80, 271-279.