Heredity (1976), **37** (1), 69-81

ESTIMATING THE NUMBER OF GENES IN A POLYGENIC SYSTEM BY GENOTYPE ASSAY

J. L. JINKS and PHILOMENA TOWEY
Department of Genetics, University of Birmingham, Birmingham B15 2TT

Received 5.i.76

Summary

A new method, genotype assay, is described for estimating k the number of genes or more strictly the number of effective factors responsible for variation of a continuous kind. The central feature is the determination of the proportion of individuals in the F_n generation of a cross between two pure breeding lines that are heterozygous at, at least, one locus by an assay of their F_{n+2} grand progeny families. The observed proportion is then equated to a theoretical expectation which is a function of the number of genes involved. Expectations generalised to cover any generation n for experimental designs in which every F_n individual is assayed by comparing two F_{n+2} grand progeny families have been derived for two limiting cases; one in which all genotypic differences are expressed as phenotypic differences and the other where the expression is minimised by imposing the maximum internal and relational balancing out of the contributions of individual gene loci. Equating the observed proportion of heterozygotes to these expectations therefore, leads to an upper and a lower estimate of k corresponding with these two limiting conditions. The reliability and sensitivity of the estimates depends primarily on n the generation chosen for study, the number of individuals (m) assayed from that generation and the number of individuals (l) raised in each F_{n+2} grand progeny family. The two variables m and l being the principal determinants of the variances of the family means set the lower limit to the size of the gene effects that can be detected.

The method is illustrated by assays of the F_3 and F_5 generations of two crosses between conditioned lines of *Nicotiana rustica* for three characters. The estimates are, without exception, as great as or greater than those obtained by alternative procedures. They show large, consistent increases between the F_3 and F_5 that cannot be traced to greater sensitivity of the latter generation and hence are presumably genuine.

1. Introduction

There are two approaches to the estimation of the number of genes, or more correctly effective factors, controlling continuous variation. One approach is based on chromosome assay (Mather and Harrison, 1949; Breese and Mather, 1957; Thoday, 1961; Law, 1967), the other on the statistical properties of distributions (Wright, 1934; Panse, 1940; Mather, 1949; Croft and Simchen, 1965; Jinks, Caten, Simchen and Croft, 1966). Both have reached a high level of sophistication and reliability (see Mather and Jinks, 1971, for review) but chromosome assay, which also locates the genes, can still be carried out on only two or three species and is never likely to be more generally applicable, while the alternative is dependent on a number of assumptions, which if not met, can lead to serious underestimation although in the right circumstances it can give reliable estimates.

An alternative approach that would be more widely applicable than chromosome assay but not so dependent on assumptions as the statistical

method would obviously be welcome. In this paper we describe a new approach, genotype assay, that seems to meet these requirements and we shall compare its use with existing statistical methods by analysing data from one of our *Nicotiana rustica* breeding programmes.

2. Theory

(i) *Frequency of detectable heterozygosity*

The frequency of heterozygotes at any one locus in the nth generation of selfing (F_n) following a cross between two pure breeding lines is $(\frac{1}{2})^{n-1}$. For k loci the probability (P_{Het}) of individuals in this generation being heterozygous at at least one locus is

$$P_{Het} = \left[1 - \frac{(2^{n-1}-1)^k}{2^{n-1}}\right]$$

Heterozygotes can be detected only by the segregation within their progenies and this segregation can be detected only by the differences in the mean and variance of families derived from the segregants that is, by genotype assay. We can proceed, therefore, by selfing each individual in the nth generation. From each individual we can then raise two randomly chosen progeny which may differ in genotype only if the parent was heterozygous at at least one locus. We can test for genetic differences between the two progeny by raising a family from each, by selfing, and comparing their means and variances.

Not all pairs of progeny chosen at random from the progeny of a selfed heterozygote will have different genotypes. At best, therefore, we shall detect only a proportion of the heterozygotes present. The probability of detecting a heterozygote by the proposed procedure will depend on the number of loci at which it is heterozygous. It is necessary, therefore, to determine the probability $P_{Het \cdot r}$ that a heterozygote in the nth generation will be heterozygous at r loci where r can take all values from 1 to k. This has the expectation:

$$P_{Het \cdot r} = \frac{1}{(2^{n-1})^k} \frac{k!}{r!(k-r)!} (2^{n-1}-1)^{k-r}$$

or

$$\frac{1}{(2^{n-1})^k} {}^kC_r (2^{n-1}-1)^{k-r}$$

$P_{Het \cdot r}$ summed over $r = 1$ to k equals P_{Het}.

The probability that the pair of individuals chosen at random from the progeny of a selfed heterozygote will differ is related to the number of loci r at which it is heterozygous as

$$1-(\tfrac{3}{8})^r$$

from which it follows that the frequency of heterozygotes in the nth generation that is detectable by progeny testing two random progeny of each individual in that generation is

$$P_{Max} = \frac{1}{(2^{n-1})^k} \sum_{r=0}^{k} {}^kC_r (2^{n-1}-1)^{k-r} (1-(\tfrac{3}{8})^r)$$

which reduces to

$$P_{Max} = 1 - \left(1 - \frac{5}{2^{n+2}}\right)^k.$$

(ii) *The effect of balance*

Not all genotypic differences lead to phenotypic differences that are detectable as differences in progeny means and variances. In a polygenic system genetic differences may fail to be expressed as phenotypic differences because of internal balance and relational balance (Mather, 1943, 1973). Internal balance will maximise the number of genotypes having identical phenotypes when the additive effects at all gene loci are equal, *i.e.*

$$d_a = d_b = d_c \ldots = d_k.$$

Relational balance will have its maximum effect when there is complete dominance at every locus, *i.e.*

$$h_a = d_a = h_b = d_b \ldots h_k = d_k.$$

If we simultaneously impose these two conditions the frequency of heterozygosity in the nth generation that will be detectable by our procedures becomes:

$$P_{Min} = \frac{1}{(2^{n-1})^k} \sum_{r=0}^{k} {}^kC_r (2^{n-1} - 1)^{k-r} \left(1 - \left(\tfrac{1}{16}\right)^k\right) \sum_{s=0}^{r} (3^s {}^rC_s)^2$$

which reduces to:

$$1 - \left(1 - \frac{1}{2^{n-1}}\right)^k \sum_{r=0}^{k} \frac{{}^kC_r \sum_{s=0}^{r} 9^s ({}^rC_s)^2}{(2^{n+3} - 16)^r}$$

This is clearly a minimal estimate because it allows for the minimum expression of genotypic differences as phenotypic differences which are capable of detection. In contrast the earlier estimate (P_{Max}) is a maximal estimate because it assumes that all genotypic differences are capable of detection as phenotypic differences. In practice, therefore, the true value must lie between these two limits.

(iii) *Estimating* k *and its sensitivity*

In figs. 1, 2 and 3, we have plotted the values of P_{Max} and P_{Min} in the F_2, F_3 and F_5 generations for a range of k values. For any proportion of heterozygotes detectable in any generation there are two corresponding values of k, an upper value obtained from the P_{Min} curve and a lower value obtained from the P_{Max} curve. These are the limits between which the true value falls, its position relative to the limits being determined by the extent to which internal and relational balance are obscuring the genotypic differences.

There are marked differences between the three generations we have chosen for analysis. The F_2 is the most sensitive to changes in k for low values of k but it has the widest upper and lower limits and hence the estimates have the greatest uncertainty. The F_5, on the other hand is the least sensitive

to changes in k but it has the lowest rate of fall off in sensitivity as k becomes large. It also has the narrowest upper and lower limits. It is, therefore, the best generation, from these points of view, for determining the value of k when k is large. The F_3 is intermediate in all respects. There are, however, other factors to be taken into consideration in assessing the relative reliabilities of different generations for estimating k.

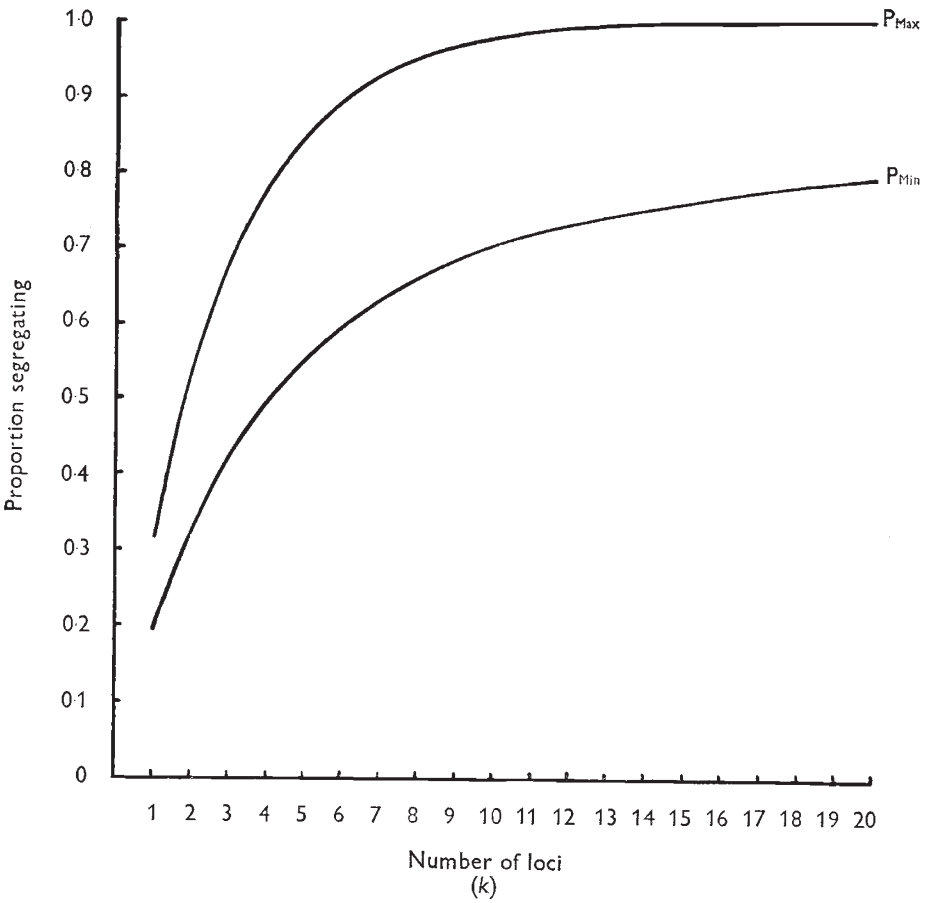

FIG. 1.—$F_n = F_2$. (See Fig. 3)

To estimate k from the F_2, each one of a number (m) of F_2 individuals must be represented by two families in the F_4 each consisting of a number of individuals (l). These two families must differ in mean or variance on the standard statistical criteria with $P \leqq 0{\cdot}05$ before we can conclude that their common grandparent in the F_2 was a heterozygote. The reliability with which we can estimate the frequency of heterozygotes will depend on the magnitude of m. On the other hand, the sensitivity with which we can detect a heterozygote will depend primarily on the magnitude of the error variances of the F_4 family means. This in turn will depend on a number of controllable and uncontrollable factors. Among the controllable factors is l, the family size and the size of the unit of randomisation, single plant or plot consisting

of a number of plants. Among the uncontrollable factors will be the variation within families arising from genetic segregation and the uncontrolled, residual variation in the environment.

If we go to the other extreme and estimate k from the F_5, there is only one predictable change and this is that the variation within the F_7 families due to genetic segregation will be considerably less than that in the F_4. For the

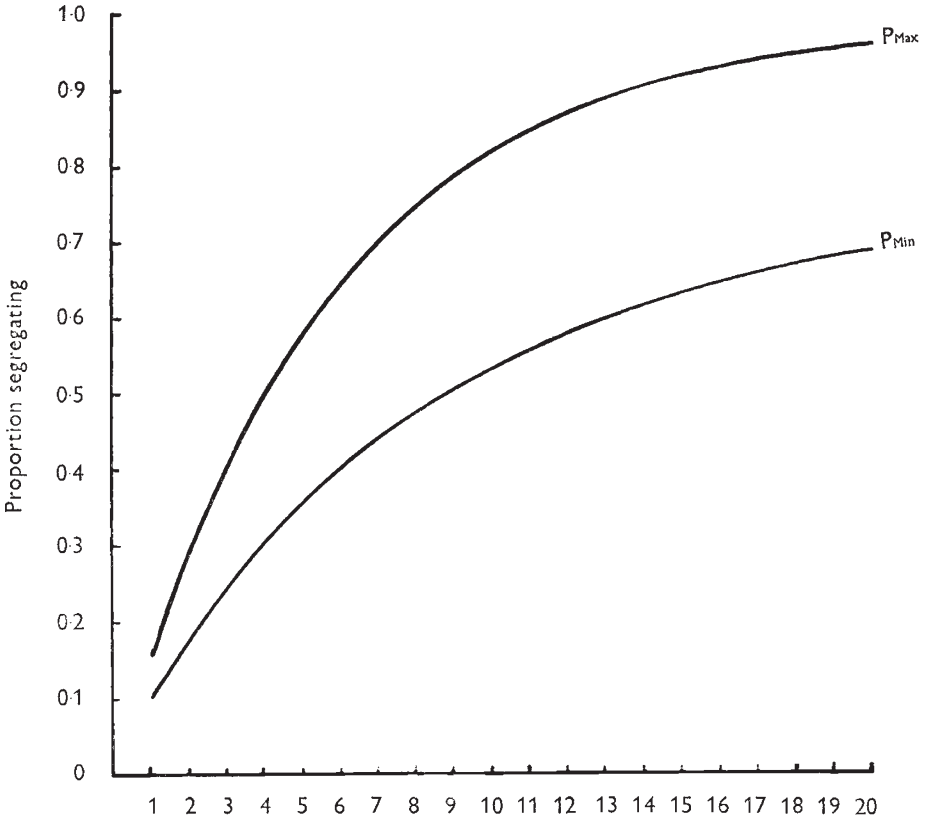

FIG. 2.—$F_n = F_3$. (See Fig. 3)

same family size, l, therefore, we shall have a more sensitive test for differences between F_7 than between F_4 families. In all other respects the sensitivity and reliability of our estimation procedures at the F_2 and F_5 will be limited by the same consideration, namely, that in practice we will only be able to increase m at the expense of l, and vice versa. Within any total size of experiment ($2ml$) there is clearly an optimal strategy for maximising the sensitivity and reliability and this will be pursued in a later paper. There are, however, a few conclusions which can be drawn already.

(*a*) In general the greatest sensitivity for detecting differences between family means will result where each of the l plants in the families of the F_{n+2} generation are individually randomised.

(*b*) Our estimate of the proportion of heterozygotes will in general be an underestimate because we can only claim to have detected a difference in family mean or variance where these are large enough to be significant by the standard criteria. We shall, therefore, be able to detect heterozygosity only where it leads to differences in means and variances that are greater

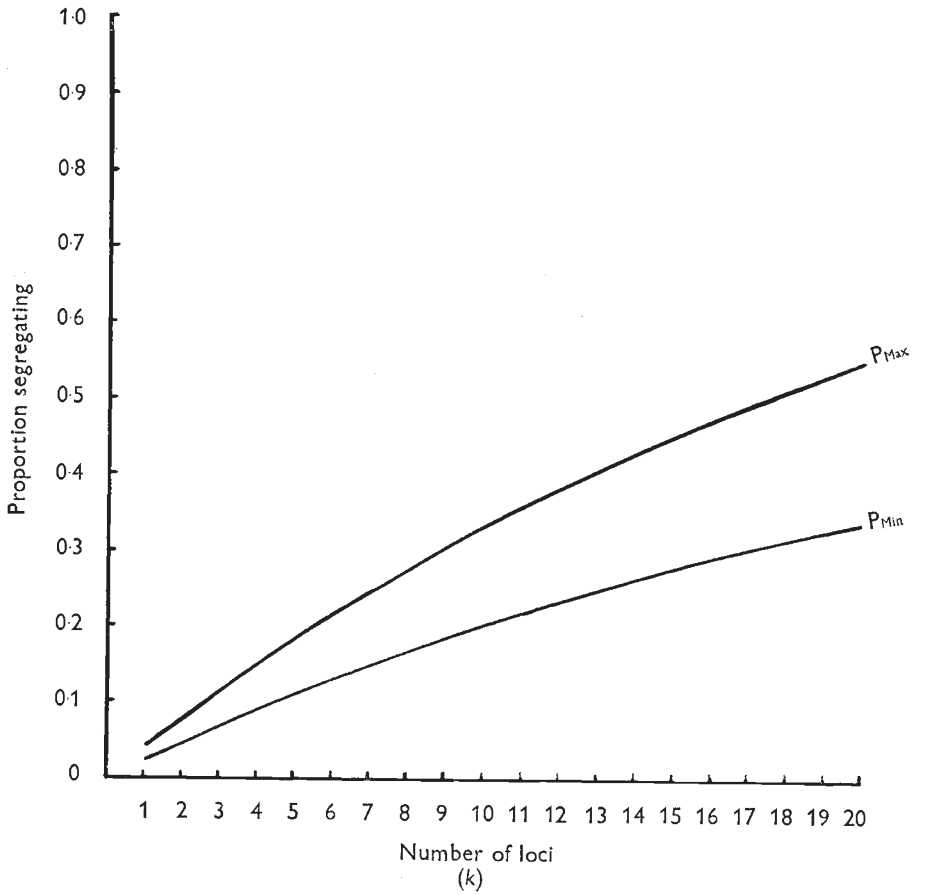

FIG. 3.—$F_n = F_5$.

The relationships between the number of gene loci, *k*, and the proportion of pairs of families in the F_{n+2} generation that are expected to differ because of the heterozygosity of their F_n grandparent. The relationships are given for two limiting sets of assumptions which maximise (P_{Max}) and minimise (P_{Min}) the proportion, respectively.

than a minimum value determined by the errors of their assessment. If we underestimate heterozygosity we will underestimate *k*.

(*c*) It is unlikely that an estimate will be made from an F_5 without first making an estimate from the F_2 or F_3. If the estimate from the F_2 is less than 4 or from the F_3 less than 10, it is not likely to be improved upon in accuracy by any later estimate. If, however, the estimate is greater than these values, greater precision could be expected by re-estimating in a later generation.

(iv) *Assumptions*

In common with all other methods of estimating k, genotype assay estimates the number of effective factors. This is less than the number of genes except in the unlikely circumstance that all the genes segregate independently (Mather and Jinks, 1971). Unless, therefore, all gene loci are unlinked we are underestimating the number of genes and over-estimating their average effect. It would be wrong, however, to infer from this that our method, or any other method of estimation, is dependent upon the assumption of no linkage. Operationally the effective factor is not only the unit which we estimate, it is also the unit of segregation for continuously varying characters and the number of such factors is arguably more relevant to understanding their variation than the number of genes even if the latter could be estimated (Mather and Jinks, 1971).

In common with all biometrical genetical analyses the estimation of k by genotype assay assumes no differential viability between genotypes and hence no selection. Since the method of genotype assay applies only to self fertile plants and all the experimental plants are derived from a pair of homozygous lines any differential viability of different genotypes should be minimal. The usual requirements of a sound experimental design whereby every F_2 plant bred from is equally represented in the progenies of all subsequent generations will further minimise the probability of viability disturbances.

3. Material

The material chosen for illustrating the new method of estimating the number of genes was produced by successive generations of selfing following crosses between contrasting pairs of conditioned lines of *Nicotiana rustica*. The origin and maintenance of these lines has been described by Hill (1967) and Hill and Perkins (1969). All the material that is appropriately structured for the analysis comes from the crosses *p3* × *nil3* and *nk2* × *nil1* chosen for further study from a 3 × 3 diallel set of crosses between conditioned lines (Perkins, Eglington and Jinks, 1971). From both crosses a random sample of F_2 individuals were used to initiate a selfing series which reached the F_7 (Eglington and Moore, 1973; Moore and Eglington, 1973; Moore, 1974). All measurements were made on plants which had been individually randomised and of the many measurements recorded, three only, flowering time, height at flowering time and final height, are available for most of the relevant material and the analyses will, therefore, be confined to these.

The sets of data taken from these two crosses that have the correct structure for our analyses are summarised in table 1. We will take as an illustrative example the first entry. In 1971, 36 F_5 families (F_{n+2} families) of the cross *p3* × *nil3* were raised each consisting of 20 individuals ($l = 20$). These families consisted of 18 pairs ($m = 18$) each member of a pair arising from the same randomly chosen F_3 grandparent (F_n). In this experiment we were, therefore, assaying 18 F_3 plants for their heterozygosity by comparing the means and variances of the pair of F_5 families that arose from each F_3 plant. The corresponding details for the other generations and crosses are summarised in table 1.

TABLE 1

A summary of the dimensions of the data which have the correct structure for estimating k *by genotype assay. Three characters, flowering time, height at flowering time and final height were recorded for both crosses in both generations*

Generation F_n	Crosses	No. of F_n's tested (m)	No. of plants per F_{n+2} family (l)	Year assessed
F_3	*p3*×*nil3*	18	20	1971
	nk2×*nil2*	20	20	1971
F_5	*p3*×*nil3*	72	10	1973
	nk2×*nil2*	80	10	1973

4. RESULTS

Continuing with our illustrative example, the means and variances of the means for flowering time for the 18 pairs of F_5 families of the *p3*×*nil3* cross when grown in 1971 are listed in table 2. The 20 plants planned for each family were divided equally between two replicate, independently randomised blocks. The means and variances in table 2 have been obtained by pooling the estimates from the two blocks. The pooled variances have, therefore, 18 degrees of freedom unless otherwise stated. The probabilities that the means and variances differ between the paired families are given in the last two columns of the table. For the means 8 pairs differ significantly ($P < 0{\cdot}05$) and for the variances 3 pairs differ at this probability after doubling the probability corresponding with the variance ratio. Since, however, the three that differ in their variances also differ in their means only 8 pairs in all differed out of the total of 18. This gives a proportion of detectable heterozygotes in the F_3 grandparents of $0{\cdot}4$. From fig. 2 this proportion can be seen to cut the P_{Min} curve at $k = 7$ and the P_{Max} curve at $k = 4$. By the same procedures we have arrived at the paired estimates listed in table 3 for the other characters, crosses and generations. For comparison we have also listed in the same table some of the alternative estimates that can be obtained from these data. These are:

(i) K_1, the estimate of k obtainable from the parental range and an estimate of the additive genetical component of variation, D.

(ii) K_2, the estimate based on within family variances in the F_3 described by Panse (1940) and extended by Mather (1949), and Cooke and Mather (1962).

(iii) A further estimate of K_1 in which the range within inbred families derived from a cross is substituted for the parental range.

Their derivation, advantages and disadvantages are described in detail by Mather and Jinks (1971). Since the estimate of K_1 from (iii) assumes that the families are highly inbred no estimate has been attempted using this method before the F_5. To avoid the effects of correlated errors we have estimated this range and D from independently randomised experiments as well as by the usual method. We have also used two alternative estimates of D, one by fitting a model to the F_5 data (Eglington and Moore, 1973) the other from an analysis of variance of the F_5 families which assumes that they

TABLE 2

Means and variances of means for flowering time in the m = *18 pairs of* F_5 *families of the* p3 × nil3 *cross grown in 1971. The number of plants in each family* (l) *and the probabilities of the means and variances differing are also listed*

	Character flowering time				
F_5 families	*l*	Mean	Variance of mean	P of *t* test for differences in means	P of *F* test of differences in variances of individuals
1	20	9·3000	0·0706	**	*
2	20	10·9000	0·2072		
3	20	10·1500	0·1681		
4	20	10·6000	0·2606		
5	20	16·0500	0·1736		
6	20	15·3500	0·3703		
7	20	10·8500	0·0703		
8	20	10·8500	0·1292		
9	19	11·1579	0·2595		
10	20	11·7000	0·1833		
11	20	10·6000	0·1583	**	
12	20	12·2000	0·1306		
13	20	9·1500	0·1758		
14	20	9·4500	0·0881		
15	20	7·3500	0·0869	***	
16	20	9·0500	0·0436		
17	18	14·3889	0·1582	+	+
18	20	13·1500	0·3458		
19	16	14·8750	0·2571	**	+
20	15	18·1333	0·7370		
21	20	14·7500	0·0436	***	**
22	20	11·1000	0·1722		
23	20	12·6500	0·3058	*	*
24	19	11·2105	0·1130		
25	19	13·5263	0·2375		
26	12	13·5833	0·4236		
27	20	12·0000	0·1606	*	
28	20	13·2000	0·1289		
29	20	10·2000	0·0250	**	+
30	20	11·0500	0·0581		
31	17	14·5294	0·3289		
32	13	13·3846	1·0328		
33	20	12·8500	1·0458		
34	15	15·0000	0·7051		
35	11	17·6364	0·9764	+	+
36	20	15·6000	0·2022		

[P = 0·05-0·10. + : Not used in subsequent analyses.]
* P = 0·01-0·05.
** P = 0·001-0·01.
*** P < 0·001.

TABLE 3

Estimates of **k**, *the number of effective factors, from the various methods*

Cross	Character	Generation	Parental range K_1	Genotype assay	K_2	Progeny range K_1^*: Analysis of variance: Pooled	Progeny range K_1^*: Analysis of variance: Independent	Progeny range K_1^*: Model fitting: Pooled	Progeny range K_1^*: Model fitting: Independent
p3 × *nil3*	Flowering time	F_3	0	4	3				
				7					
		F_5	1	11		5	3	3	2
				20			4		3
	Height at flowering time	F_3	—	3	—				
				5					
		F_5	8	6		0	2	—	—
				11			1		—
	Final height	F_3	1	2	1				
				4					
		F_5	3	7		4	6	3	1
				12			2		2
nk2 × *nil1*	Flowering time	F_3	0	4	1				
				9					
		F_5	0	9		5	2	4	2
				17			3		3
	Height at flowering time	F_3	—	2	—				
				4					
		F_5	0	5		3	4	—	—
				9			2		—
	Final height	F_3	0	3	1				
				5					
		F_5	0	3		3	1	—	—
				5			4		—

* The sub-headings refer to the alternative sources of the estimates of D and of the progeny range (see text for further details).

are effectively pure-breeding. No estimate has been obtained based on the dominance components because of the relatively poor estimates of the dominance component of variation obtainable from the selfing series.

In a few instances, indicated by a dash in table 3, no estimate could be obtained by a particular method because part of the information required was not available. Wherever the estimate obtained was less than 0·1, it is recorded as zero in the table, otherwise all estimates are rounded up to the next whole number.

5. Conclusions

In every case the estimate from our new method of genotype assay is equal to or larger than those from any of the existing methods (table 3). K_1 estimated from the parental range only gives realistic estimates where the parents have been selected to be opposite extremes for that character. For example, *p3* and *nil3* were chosen as the smallest and tallest of the conditioned

Table 4

The variances of family means averaged over all families in each generation for the F_5 and F_7 generations. The smaller the value the greater the sensitivity in detecting heterozygosity by a difference between family means

Cross	Character	Mean variance of mean F_5	Mean variance of mean F_7
p3 × *nil3*	Flowering time	0·45	0·99
	Height at flowering time	3·03	4·12
	Final height	7·49	13·81
nk2 × *nil1*	Flowering time	0·27	1·10
	Height at flowering time	4·72	6·50
	Final height	11·66	21·81

lines, respectively, and the highest estimates of K_1 come from height at flowering time and final height in the cross between them. The K_1 estimates from the family range in the F_5 are more consistent in value over the two crosses and three characters because they are freed to some extent from the minimising effects of gene dispersion.

For ease of discussion we can take the estimate of k from the new method as that which lies half-way between the upper and lower estimates. For height at flowering time this estimate is 3 and 4 for the two crosses in the F_3 rising to 7 and 8·5 in the F_5 generations. For flowering time it is 5·5 and 6·5 for the two crosses in the F_3 rising to 13 and 15·5 respectively in the F_5. The only inconsistency is for final height in the F_5 where the estimate is 9·5 rising from 3 in the F_3 for the first cross, and 4 in the F_3 and F_5 in the second cross. Only for final height in the second cross, therefore, is there no consistent increase in the estimate between the F_3 and the F_5 generation.

There are a number of reasons why the estimate of k might increase between the F_3 and F_5 some of which we have already discussed. One reason is greater sensitivity of the F_5. This greater sensitivity could arise in two quite distinct ways (section 2 (iii)). It could arise because the variation within families is expected to be smaller in the F_7 than in the F_5 generation due to the greater genetical homogeneity of the members of the same

F_7 family. This, however, has already been more than discounted in the experimental design (table 1) by reducing the family size (l) from 20 in the F_5 to 10 in the F_7 so as to accommodate a greater number of families (m). As a result the variances of the family means, which largely determine sensitivity, are actually larger for the F_7 than for the F_5 (table 4). Not surprisingly therefore, the smallest differences in family means that are significant are smaller in the F_5 than in the F_7. Indeed, if for example, for flowering time in the *p3* × *nil3* we had achieved the same sensitivity in the F_7 as in the F_5 assessments by having families of size 20 and we assume no consequential changes in the family means, we would have a minimum estimate of 29 genes from the F_7 assessments instead of the 11 actually obtained. This is equally the case for the other characters and the other cross. However, to return to the actual situation, the observed increase in the estimate between the F_3 and F_5 generations (assessed in the F_5 and F_7, respectively) clearly cannot be attributed to greater sensitivity from this cause. The other possibility is that it could arise from the greater sensitivity of the P_{Max}, P_{Min} curves (figs. 1, 2 and 3) for the F_5 for the range of k values found in these data. This too does not seem to be the explanation because in the observed range of 2 to 20 (table 3), the F_3 curves are more sensitive than the F_5 at the lower end and equal in sensitivity to the F_5 at the upper end.

Since differences in sensitivity do not appear to offer a satisfactory explanation, the alternative is that a number of effective factors has genuinely increased between the F_3 and the F_5. Because of the nature of effective factors this is in fact expected (Mather and Jinks, 1971) and the additional rounds of recombination between the F_3 and F_5 could be sufficient to increase the number we detect.

It is, of course, quite explicit in the new method that we are only detecting those effective factors whose segregation causes a difference in mean and variance above a certain size, which is the size that achieves significance. For example, for flowering time in the F_3 of both crosses we are detecting between 4 and 9 genes whose segregation cause a difference in family mean greater than 0·8 days. If segregation at some loci are producing differences which are smaller than this in magnitude we will have missed them and hence underestimated the total number of genes involved.

6. Further developments

We have presented a preliminary account of genotype assay appropriate for a simple situation where each individual in the F_n generation is assayed for its heterozygosity through two grand-progeny families in the $F_{(n+2)}$ generation. This has been sufficient to establish its value relative to the alternative methods that are applicable to the same data. It is not, however, necessarily the best strategy for genotype assay, nor are the particular values of l and m of the illustrative examples the best division of resources within the total of 2 ml individuals when two families are used in the $F_{(n+2)}$. If the condition of two grand-progeny families is relaxed we then need to consider the general case of p families within a total of pml individuals. The effects of varying p, with illustrative examples, will be the subject of a later paper.

Acknowledgments.—We are indebted to Professor K. Mather for his invaluable comments on the new procedures and to Mr P. J. Jinks for generalising the probability equations. One of us (P. T.) is supported by an SRC Research Studentship.

7. References

BREESE, E. L., AND MATHER, K. 1957. The organisation of polygenic activity within a chromosome in *Drosophila*. I. Hair characters. *Heredity*, *11*, 373-395.

COOKE, P., AND MATHER, K. 1962. Estimating the components of continuous variation. II. Genetical. *Heredity*, *17*, 211-236.

CROFT, J. H., AND SIMCHEN, G. 1965. Natural variation among monokaryons of *Collybia velutipes*. *Amer. Nat.*, *94*, 451-462.

EGLINGTON, E. G., AND MOORE, C. A. 1973. The nature of the inheritance of permanently induced changes in *Nicotiana rustica*. II. F_4 and F_5 generations of selected crosses. *Heredity*, *30*, 387-395.

HILL, J. 1967. The environmental induction of heritable changes in *Nicotiana rustica*: Parental and selection lines. *Genetics*, *55*, 753-754.

HILL, J., AND PERKINS, J. M. 1969. The environmental induction of heritable changes in *Nicotiana rustica*: Effects of genotype-environment interactions. *Genetics*, *61*, 661-675.

JINKS, J. L., CATEN, C. E., SIMCHEN, G., AND CROFT, J. H. 1966. Heterokaryon incompatibility and variation in wild populations of *Aspergillus nidulans*. *Heredity*, *21*, 227-239.

LAW, C. N. 1967. The location of genetic factors controlling a number of quantitative characters in wheat. *Genetics*, *56*, 445-461.

MATHER, K. 1943. Polygenic inheritance and natural selection. *Biol. Revs.*, *18*, 32-64.

MATHER, K. 1949. *Biometrical Genetics*, 1st edition. Methuen, London.

MATHER, K. 1973. *Genetical Structure of Populations*. Chapman and Hall, London.

MATHER, K., AND HARRISON, B. J. 1949. The manifold effects of selection. *Heredity*, *3*, 1-52 and 131-162.

MATHER, K., AND JINKS, J. L. 1971. *Biometrical Genetics*, 2nd edition. Chapman and Hall, London.

MOORE, C. A. 1974. Conditioned lines of *Nicotiana rustica*. Ph.D. thesis, University of Birmingham.

MOORE, C. A., AND EGLINGTON, E. G. 1973. The nature of the inheritance of permanently induced changes in *Nicotiana rustica*. III. F_5 generation for five characters. *Heredity*, *31*, 112-118.

PANSE, V. G. 1940. The application of genetics to plant breeding. II. The inheritance of quantitative characters and plant breeding. *J. Genet.*, *40*, 283-302.

PERKINS, J. M., EGLINGTON, E. G., AND JINKS, J. L. 1971. The nature of the inheritance of permanently induced changes in *Nicotiana rustica*. *Heredity*, *27*, 441-457.

THODAY, J. M. 1961. Location of polygenes. *Nature*, *191*, 368-370.

WRIGHT, S. 1934. The results of crosses between inbred strains of guinea pigs, differing in number of digits. *Genetics*, *19*, 537-551.