THE QUANTITATIVE ANALYSIS OF THRESHOLD CHARACTERS USING ASYMMETRY: A STUDY OF THE WITTY CHARACTER IN *DROSOPHILA MELANOGASTER*

M. J. WHITTEN¹

Department of Botany, University of Tasmania, Hobart, Australia

Received March 14, 1966

CERTAIN characters which are discontinuous in their distribution are inherited in a quantitative manner. The occurrence of a number of discrete phenotypic classes for such characters is usually explained in terms of thresholds and an underlying continuous variation (FALCONER 1960). This underlying variation may be regarded as the concentration of one or more substances which influence the phenotypic expression of the character under study. In genetic studies the nature of this underlying variation is usually left unspecified since its specification is not essential to the argument: RENDEL has postulated the existence of a morphogenetic substance (m.s.) or *make* to describe the underlying variation in the case of the scutellar bristles of *Drosophila* (RENDEL 1962). In this paper RENDEL's terminology has been adopted.

The moments of the distribution of the underlying variation are usually determined by a probit analysis (see FALCONER 1960). It is assumed that the number of classes and the size of each class is dependent on the number and location of the thresholds on the underlying scale and the distribution of the underlying variable. When there are only two or three classes present, the problem is generally considered to be one of incomplete penetrance and variable expression (see SANG 1963). However, the problem can be treated in much the same way as cases where there are more than three classes present.

In this paper I will consider the case of three phenotypic classes involving asymmetry where the asymmetry is caused by errors in development. Evidence will be presented which indicates that a simple two-threshold probit analysis may be inadequate in this case. An alternative model is suggested and, although this model involves only one threshold, it permits a more penetrating analysis of the variation on the underlying scale. This underlying variation is partitioned into the following components; additive genetic, nonadditive genetic, environmental and "developmental noise".

The model has been applied to a new eye mutant in *Drosophila melanogaster*. However, an investigation of the literature and preliminary studies have revealed that other characters such as antennaless, various eye mutants, scarp, lobed, eyeless and possibly crossveinless are amenable to the same analysis. Consequently it is felt that the model has general application to characters involving a

¹ Present address: Division of Entomology, CSIRO, Canberra.

breakdown in symmetry. A possible extension of the method of analysis to include cases with more than three phenotypic classes (e.g. certain bristle characters in *Drosophila*) is discussed briefly.

MATERIALS AND METHODS

Three stocks of *Drosophila melanogaster*, homozygous for *wi*: witty (chromosome II—54.6), were derived from a *j*: jaunty stock obtained from the University of Melbourne in 1961. Line SH, highly inbred for 35 generations, has incomplete penetrance of *wi* under normal laboratory conditions. Lines LL and HH show zero and complete penetrance, respectively, under the same conditions. The flies were cultured in vials, on normal semolina medium fortified with yeast, in an unlighted incubator at $25 \pm 0.5^\circ\text{C}$.

Breakdown in symmetry: When the mutant allele witty (*wi*) is present, development may be abnormal. The abnormality is invariably restricted to the lower portion of the eye and is characterized by disoriented and missing ommatidia (see Figure 1). As in scarp (HANSEN and GARDNER 1962), tufts of vibrissae on the anteroventral border of the eye are not uncommon. In fact the two mutants are remarkably similar in their expression but they are not identical. In witty flies the ventral third of the eye is never differentiated from the dorsal two thirds by a horizontal depression as in scarp. Of more significance, and in contrast to

FIGURE 1.—Comparison between wild-type eye and eyes showing the witty phenotype; (a) normal wild-type expression, (b), (c) and (d) varying expression in flies homozygous for the witty gene.

scarp, penetrance is always higher in females. Either or both eyes may be affected so that four phenotypic classes can be observed; $(++)$, $(L+)$, $(+R)$, and (LR) where $(++)$ denotes flies wild type for both eyes, $(L+)$ denotes flies normal for right eye only, etc.

Usually it is assumed that the level of the morphogenetic substance must be above a certain threshold in order that a normal phenotype results. For convenience we are taking an opposite view and propose that the level of m.s. must surpass a threshold to produce an abnormal eye. This premise is used in the two models discussed below. These models are neither exclusive nor exhaustive but they are the simplest and for this reason, discussion is limited to them.

Model I. Each fly produces a certain amount of morphogenetic substance which is shared in some manner, by both eyes. In the $(++)$ phenotype insufficient m.s. is available during the critical period in development to cause either eye to develop abnormally. With (LR) there is sufficient to disrupt development of both eyes while in the case of $(L+)$ or $(+R)$ one eye receives sufficient to cause an abnormality, only to leave the other with an amount insufficient to upset development. Model I implies two thresholds, separating the asymmetrical class from the two symmetrical classes.

Model II. Each eye produces its own m.s. independently. Both sides of the fly have the same genetical make-up, and the external environment in which they develop must be very similar, but because of small accidents in development, i.e. "developmental noise", due to such factors as unequal cell division, the amount of m.s. is likely to be different for each eye. Consider individuals where the genotype and environment are such that their levels of m.s. lie in the vicinity of the threshold. In some individuals developmental noise will cause the amount of m.s. to fall below the threshold for one eye and above it for the other. Whenever this happens an asymmetrical fly is produced. Where the genetic constitution or the environmental circumstances during development produce levels of m.s. well removed from the threshold, developmental errors due to noise will remain cryptic.

In Model I there is competition between sides for m.s. or at least the outcome for one side is dependent on what happens with the other. In Model II each side is independent of the other as regards the final amount of m.s. it produces. In essence, Model I corresponds to the one used by RENDEL (1965) in his analysis of the scutellar bristles while Model II is similar to that proposed by ROBERTSON (1965) and LATTER (1964) for the same character. RENDEL assumes that the total number of bristles is the primary character determined by the interaction of the genotype and the environment. ROBERTSON, on the other hand, considers total bristle number a secondary character, the primary variable determined by the genotype and the general environment being the probability of a bristle developing at a particular site.

It is possible for the witty character to eliminate Model I in favour of Model II. It can be demonstrated that the primary character determined by the genotype and the environment is the *probability* of an eye developing abnormally and not the *number* of abnormal eyes exhibited by an individual. If there is no genetic

or environmental variability present in a particular population, any variability in the amount of m.s. must be due to developmental noise. In such a population if the distribution of the level of m.s. is in the neighbourhood of the threshold so that the proportion of eyes in which the level surpasses the threshold is equal to p , then, assuming Model II, the phenotypes $(++)$, $(L+)$, $(+R)$, (LR) should occur in the binomial frequencies $(1-p)^2$, $p(1-p)$, $(1-p)p$ and p^2 respectively.

If Model I is operating, the proportion of asymmetrical flies can exceed $2p(1-p)$ and this tendency is to be expected if the genetic and environmental components are reduced. It should be noted that both models predict a reduction in the proportion of asymmetrical flies when there is appreciable variability in the amount of m.s. owing to genetic and environmental variability. Determination of the frequencies of the four classes in populations in which genetic and environmental variability is reduced to a minimum should provide a critical test between the two hypotheses.

TABLE 1

Effect of genetic and environmental variability on the frequency of the phenotypic classes $(L+)$, $(+R)$, (LR) , $(++)$ in wi/wi flies

Stock*	Genetic variability	Environ- mental variability	Sex		L+	+R	LR	++	P†	($p_f \cdot p_m$)	$x^\dagger$	$y^\dagger$	($x_f - x_m$)	
SH ₀	low	low	m	obs.	125	106	64	416	0.2525	0.4042	—0.75	0.55	1.20	
				exp.	134	134	45	397						
			f	obs.	163	124	328	103	0.6567		+0.45	0.46		
				exp.	162	162	310	85						
SH ₉	low	low	m	obs.	190	181	160	364	0.3860	0.3737	—0.33	0.49	1.20	
				exp.	212	212	133	337						
			f	obs.	118	115	470	59	0.7697		+0.87	0.52		
				exp.	135	135	451	40						
SH ₂₀	low	low	m	obs.	139	138	62	456	0.2522	0.4316	—0.72	0.4	1.23	
				exp.	150	150	51	445						
			f	obs.	160	143	375	92	0.6838		+0.51	0.43		
				exp.	166	166	360	77						
SH ₂₄	low	high	m	obs.	54	57	219	267	0.4598	0.1336	—0.30	2.2	0.8	
				exp.	148	148	126	174						
			f	obs.	75	54	288	177	0.5934		+0.50	1.8		
				exp.	143	143	209	98						
F ₂ (HH×LL)	high	low	m	obs.	49	51	315	139	0.6588	0.1871	+1.0	2.1	1.4	
				exp.	125	125	240	64						
			f	obs.	35	21	411	52	0.8459		+2.4	2.1		
				exp.	68	68	371	12						

* SH₀, SH₉ and SH₂₀ represent the zero, ninth and 20th generations, respectively, of a line derived from a highly inbred strain of wi/wi flies. F₂ (HH×LL) are the F₂ progeny from a cross between two wi/wi lines, HH and LL having complete and zero penetrance, respectively, under normal laboratory conditions.

† P is the proportion of abnormal eyes; x and y are the mean and standard deviation of the penetrance parameter, ξ . The expected frequencies are calculated using the binomial expansion $(P + (1-P))^2$. An increase in the variance of ξ , whether it be genetic or environmental in origin, causes an increased departure from the frequencies predicted by the binomial.

Highly inbred stocks, homozygous for *wi*, grown under standardised and almost uniform conditions result in populations tending towards the binomial proportions predicted by Model II. For brevity, the results of only three generations of the inbred line, SH, are given in Table 1; similar results were obtained with all other generations in this line. In no cases are the proportion of asymmetrical flies greater than that predicted by the binomial. The positive correlation between departure from the binomial distribution and heterogeneity of background genotype or general environment, indicated in Table 1, suggests that, in the ideal population where this variability is excluded, the distribution would equate to the binomial as predicted in Model II. It will be shown elsewhere that the small but statistically significant deviation from the binomial observed in the inbred line SH (see Table 1) is due largely to changes occurring in the medium after its occupation for some time by larvae. (Penetrance of the witty gene increases with the age of the medium, in contrast to the behaviour of the eyeless gene (MORGAN 1929) and antennaless (GORDON and SANG 1941) where penetrance is lowered as the culture ages.)

The model: Data presented in Table 1 indicate that an increase in the variability of morphogenetic substance through genetic or environmental heterogeneity leads to an increased departure in the frequency of the asymmetrical class from that predicted by the binomial. It should be possible, then, to utilize this deviation from the binomial expectations to measure the variability of the underlying variable in any particular population.

We have argued that developmental noise can affect the amount of m.s. produced by either side of a fly. Consequently, for flies of a given genotype developing in a particular environment the amount of m.s. is not constant but varies from eye to eye.

The situation may be symbolised as follows:

$$\eta = \xi + \delta \quad (1)$$

Where η is a variable representing the amount of m.s. for an eye, δ is a random normal variable whose value is determined by developmental noise, and ξ is a variable whose value is determined by the genotype and the environment.

By definition $E(\delta) = 0$. We now make the assumption that V_δ , the variance of δ , is constant and independent of ξ (the feasibility of this assumption is examined in the discussion and evidence suggesting its validity will be presented in the second paper in this series).

It follows that

$$E(\eta) = E(\xi) \quad (2)$$

$$\text{and } V_\eta = V_\xi + V_\delta \quad (3)$$

The question now arises as to what units we should use to measure η . The threshold would seem to be the logical choice as origin and sd_δ , which we have assumed to be constant, would appear to be the most convenient unit for measuring η .

Thus

$$V_\eta = V_\xi + 1 \quad (4)$$

Consider a population of fixed genotype and environment such that η is distributed about the threshold as in Figure 2, (i.e. $\xi = 0$). A change in genotype or the environment will alter the value of ξ which, in turn, will cause the distribution

FIGURE 2.—Distribution of morphogenetic substance for a population of flies of a given genetic constitution developing in a constant environment such that the distribution is centered around the threshold, i.e., $\xi = 0$. The variability in m.s. is due to developmental noise. A scale is chosen so that there is unit variance in m.s. due to noise. Altering the genotype or the environment has the effect of moving the distribution along the m.s. scale relative to the threshold, the direction depending on the nature of the alteration; e.g. increased temperature increases the penetrance causing a shift in the right-hand direction. It is assumed that such changes alter the value of ξ only and do not affect the variability in m.s. due to noise.

of η to shift along the horizontal axis. Thus in a real population, the mean of the distribution of η is likely to differ with each combination of a genotype and environment. Let π be the probability for any given individual of an eye developing abnormally. Then p , the value that π takes, for a particular individual, is equal to the probability of η surpassing the threshold in that individual. If the value of ξ be u for a particular genotype developing in a given environment, then p and u are related as follows:

$$p = \frac{1}{\sqrt{2\pi}} \int_{-u}^{\infty} e^{-\frac{1}{2}v^2} dv$$

Thus for any fly, π , the probability of an eye being abnormal, is directly related to ξ which is the mean value of η for the corresponding genotype and environment (see Figure 3).

Because ξ gives a direct measure of the probability of an eye developing abnormally, I have called it the penetrance parameter. It should be noted that the value of ξ , unlike the value of η is the same for both eyes in any one fly. It is independent of developmental noise. However, it is determined by an individual's

FIGURE 3.—Relationship between π and ξ .

genotype and the environment in which it develops. It will be shown in subsequent papers in this series that ξ is controlled by background genotype, and by a number of environmental factors. Hence, in any normal population, homozygous for w^i , we can assume that ξ is normally distributed. Under these conditions η , also, must be normally distributed.

For ξ normally distributed, it can be shown that the mean, x , and standard deviation, γ , of ξ are uniquely determined by Z_1, Z_2, Z_3 where:

$$\begin{aligned} Z_1 &= \text{proportion of } (++) \text{ flies} \\ Z_2 &= \text{proportion of } (L+) \text{ and } (+R) \text{ flies} \\ Z_3 &= \text{proportion of } (LR) \text{ flies (see Appendix 1)} \end{aligned}$$

Because of the complicated relationship between x, γ and Z_1, Z_2, Z_3 an Elliott 503 computer was used to construct a table giving the triplet Z_1, Z_2, Z_3 for suitable ranges of x and γ . (The table and computer program are too cumbersome to publish but are available on request.)

Several trials established that x and γ over the ranges $-5 < x < +5$; $0 < \gamma < 2.5$; cover most situations. Thus the measurement of x and γ for any population simply involves calculating Z_1, Z_2, Z_3 by scoring the frequency of the three classes $(++)$, $(L+)/(+R)$ and (LR) and then referring to the computed table, to find the appropriate mean and standard deviation of ξ .

Ignoring genotype/environment interaction, we can expand equation (4) as follows:

$$V_{\eta} = V_G + V_E + 1, \quad (5)$$

where V_G and V_E are the respective genetic and environmental components of V_{ξ} and may be estimated by conventional heritability experiments. Hence it is possible to partition the total variance of morphogenetic substance into its noise, genetic and environmental components.

Perhaps the most commendable attribute of this method of measuring penetrance is the ease with which it can be tested. A series of experiments has been performed in which either the genotype or the environment has been held constant while the other has been altered. The remainder of this paper is concerned with two heritability experiments which demonstrate the application of the model in the analysis of underlying variation. The environmental modification of ξ and the genetical analysis of the genes modifying ξ will be discussed elsewhere.

RESULTS

(a) *Difference in penetrance between sexes*: One point must be discussed before proceeding to the heritability experiments. Penetrance of the witt gene when measured simply as the proportion, p , of abnormal eyes is always higher in females. However, the difference in penetrance between sexes ($p_f - p_m$), is not constant. Maximum difference occurs in the inbred lines (e.g. Table 1, row 1, 2 and 3) while in lines with a history of outbreeding, ($p_f - p_m$) is considerably less (e.g. Table 1, row 5). It was also found that ($p_f - p_m$) decreased in inbred lines when the flies were permitted to develop in a heterogeneous environment (Table

1, row 4). The problem vanishes, or rather changes dimensions, when penetrance, instead of being measured as the proportion of affected eyes, is measured in terms of ξ . The difference in the mean of ξ between sexes ($x_f - x_m$), appears to be constant for all lines homozygous for *wi* irrespective of the value of ξ and the degree of genetic or environmental variability associated with the lines (Tables 1, 2, 3). The value ($x_f - x_m$) fluctuates more in the outcrossed lines but this is to be expected owing to the higher variance of ξ in these lines. It is not difficult to show that, as the variance of ξ increases, ($p_f - p_m$) should decrease. In fact, it is possible to use the value of ($p_f - p_m$) to estimate the variance of ξ . Nevertheless, experience has shown that this estimate is useful only in a confirmatory capacity to the normal method of estimating the variance by calculating Z_1 , Z_2 and Z_3 .

We can conclude that the threshold for females is about one unit lower than for males (see Tables 1-4). Alternatively we may argue that the amount of m.s. is always about one unit higher in females than in the corresponding males.

(b) *Heritability estimate of ξ using the inbred line SH*: The mean, x , and standard deviation, γ , of ξ were calculated for a number of generations of the highly inbred line, SH, homozygous for *wi* (see Tables 1 and 2 for GO, 8, 9 and 20). The values of x and γ were determined for the progeny of a cross in which only (++) flies from SH (G8) were used as parents and also for a cross in which all the parents were (LR). (See Table 2).

Heritability, h^2 , was calculated from the formula $h^2 = R/S$ (FALCONER 1960, p. 189), where R is the response to selection and S is the selection differential. Using the two-way selection experiment described by FALCONER, S and R are given by

$$S = x \text{ (for (LR) parents)} - x \text{ (for (++) parents)}$$

$$R = x \text{ (for progeny of (LR) parents)} - x \text{ (for progeny of (++) parents)}.$$

If x and γ are known for a population it is possible to calculate the mean and standard deviation of ξ for either the (++) or (LR) sub-populations (see Appendix II). The difference in ξ for males and females necessitated each sex being treated separately, the final estimate of heritability for the line being the average heritability for each sex (see Table 4).

TABLE 2

*Results of two-way selection experiment to determine heritability of ξ for SH, an inbred line homozygous for *wi**

Population	Males		Females	
	Mean and sd of ξ	Sample size	Mean and sd of ξ	Sample size
SH ₈	-0.24 $\pm$ 0.46	356	+0.78 $\pm$ 0.43	345
(++) flies from SH*	-0.47 $\pm$ 0.42	...	+0.4 $\pm$ 0.38	...
Progeny of (++) flies	-0.37 $\pm$ 0.3	326	+0.9 $\pm$ 0.49	290
(LR) flies from SH	+0.08 $\pm$ 0.41	...	+0.91 $\pm$ 0.40	...
Progeny of (LR) flies	-0.30 $\pm$ 0.58	569	+0.81 $\pm$ 0.58	472

* To estimate the mean and sd of ξ for (++) and (LR) subpopulations see Appendix II.

TABLE 3

Results of two-way selection experiment to determine heritability of ξ for the outcrossed line F_2 ($LL \times HH$) homozygous for w_i

Population	Males		Females	
	Mean and sd of ξ	Sample size	Mean and sd of ξ	Sample size
LL	-2.6 ± 0.57	...	-1.6 ± 0.57	...
HH	$+2.3 \pm 0.8$	...	$+2.9 \pm 0.8$	...
F_2 ($LL \times HH$)	$+1.0 \pm 2.1$	554	$+2.4 \pm 2.1$	519
($++$) flies from F_2	-1.72 ± 1.2	...	-1.03 ± 1.3	...
Progeny of ($++$) flies	-2.2 ± 1.3	500	-1.1 ± 1.1	476
(LR) flies from F_2	$+2.38 \pm 1.5$	...	$+3.24 \pm 1.7$	...
Progeny of (LR) flies	$+0.8 \pm 2.4$	614	$+1.8 \pm 2.1$	690

FIGURE 4.—The distribution of ξ , before and after selection, in the two-way selection experiment. The unbroken curves represent the probability density function for ξ . For convenience each curve has been plotted to give the same coordinate value at the peak of its distribution. Thus the area under the curve, in each case, is equal to γ , the standard deviation of ξ for that particular distribution. The constant difference in ξ between sexes, necessitates both sexes being treated separately. It is evident that each sex follows a parallel behaviour. The F_2 from a cross between the two extreme lines, LL and HH, both homozygous for w_i , show a greatly increased variance for ξ (cf. Table 2). The distribution of ξ for ($++$) and (LR) subpopulations (broken lines) is calculated theoretically (see Appendix II). The ($++$) and (LR) parents used in the two-way selection experiment were randomly chosen from these two subpopulations. The selection differential is higher for the left hand side of the distribution (b) and this is reflected by the greater response in this direction (c). Note that the variance of ξ in the progeny is proportional to its variance in the selected parents.

(c) *Heritability of ξ for an outcrossed line*: A parallel experiment was conducted using the F_2 generation from a cross between two lines homozygous for witty. One line, HH, had previously been selected for complete penetrance while in the other line, LL, penetrance was effectively zero. A summary of the results for the two-way selection experiment on the F_2 generation are given in Tables 3 and 4 and Figure 4. It is worth noting that the variance of ξ in the progeny is proportional to the variance of ξ in the selected parents. The (++) parents represent a more restricted sample of the parental population than the (LR) parents and this situation is reflected by the fact that their progeny have a smaller variance in ξ .

Measurements of the heritability of ξ give a value of -0.02 for the inbred line and 0.7 for the outcrossed line. The variance in ξ for the inbred line, averaged over a number of generations, is estimated at 0.25 (see Table 1 for sample results). Therefore we can conclude that the variance of ξ , for the inbred line, is essentially environmental, the nonzero value of heritability being due to experimental error.

Assuming that the variance of ξ , attributable to environmental factors, is the same for both the outcrossed and inbred lines (0.5^2 , see Table 2) we can partition the variance of ξ (2.1^2) for the outcrossed line as follows. $V_G = V_\xi - V_E = 2.1^2 - 0.5^2 = 4.2$ (see Table 2).

Since $h^2 = V_A/V_\xi$ where V_A is the additive component of genetic variability, $V_A = h^2 \times V_\xi = 0.7 \times 4.4 = 3.1$. Again ignoring genotype/environment interaction, the nonadditive genetic component, V_D , is given by $V_D = V_G - V_A = 4.2 - 3.1 = 1.1$.

A complete analysis of the variability of ξ and the distribution of m.s. for the inbred and outcrossed lines is summarised in Table 5. The calculation of standard errors for the different sources of variability estimated in this table appears to be very complicated and no attempt has been made to solve this problem.

DISCUSSION

If a character is directly related to the quantity of some metabolite it will vary from individual to individual as the quantity of the metabolite varies; that is, the character itself will be quantitative. In such cases it is sufficient to analyse the phenotypic variability, which in turn, reflects the underlying variation and there

TABLE 4
Heritability of ξ for the outcrossed and inbred lines

	Inbred line, SH		Outcrossed line F_2 (LL $\times$ HH)	
	Males	Females	Males	Females
Selection differential	+0.53	+0.51	+4.1	+4.27
Response	+0.07	-0.09	+3.0	+2.9
Heritability	+0.13	-0.18	+0.73	+0.68
Average heritability	-0.02		+0.70	

TABLE 5

Analysis of variance for the penetrance parameter ξ and the distribution of morphogenetic substance in the inbred and outcrossed lines, homozygous for w^i

Source of variation	Outcrossed line F_2 (LL $\times$ HH)		Inbred line, SH	
	Penetrance parameter (percent)	Morphogenetic substance (percent)	Penetrance parameter (percent)	Morphogenetic substance (percent)
Developmental noise	..	18	..	80
Environmental	6	5	100	20
Additive genetic	70	57	0	0
Nonadditive genetic	24	20	0	0

are no problems in partitioning the variability according to its various sources. In systems of development where the expression at the phenotypic level is not simply related to the quantity of morphogenetic substance (m.s.) but depends on the presence and position of thresholds, the problem of analysis is more complicated. For multiple thresholds the mean and variance of the underlying variation are easily determined by using probits. This technique is adequately described in the literature (FALCONER 1960; RENDEL 1962; and SANG 1963) and needs no further comment here.

When there is only a single threshold operating, the problem of analysis is entirely different and it is generally agreed that little information is obtainable about the distribution of the underlying variable. The present work has shown this obstacle can be circumvented if the character is associated with a breakdown in bilateral symmetry. At first glance one may argue that examples of this sort, e.g. witty, eyeless, scarp, antennaless, crossveinless, etc., are not all-or-none characters because there are three basic classes present; in our notation these would be $(++)$, $(L+)/(+R)$ and (LR) . If this objection were valid, examples of this type would have two thresholds, one separating $(++)$ from $(L+)/(+R)$ and the other between $(L+)/(+R)$ and (LR) and therefore such examples would be amenable to probits.

In this paper an argument has been presented which strongly suggests that in the witty system, only one threshold operates. Each side of the fly enjoys a certain degree of autonomy in determining the phenotype of the eye it produces, since the ultimate outcome is determined by developmental noise which operates independently for each eye. This is indicated by the close approximation of the frequencies of the four phenotypic classes to the binomial distribution where the genetic and environmental variance are experimentally minimised (see Table 1). It would appear unsound in cases similar to the witty system to ignore this situation and use the two-threshold probit analysis which is mathematically simpler but which has little biological basis.

I have shown that the assumption of a model based on a single threshold is not limiting, in the sense that variances on the underlying scale cannot be estimated (see FALCONER 1960, p. 301), but offers considerable advantages. It permits an extension of the analysis beyond the traditional partitioning of variance

into genetic and environmental components and allows us to measure, quantitatively, other components in relation to the effects of developmental noise.

The logical steps in the analysis are as follows:

(1) The proportions of $(++)$, $(L+)/(+R)$ and (LR) phenotypes, Z_1 , Z_2 and Z_3 are measured.

(2) Using these proportions we could determine the distribution of probabilities (π) for the population. However, an answer couched in terms of π would be unsatisfactory both statistically and biologically. The statistical drawbacks are twofold. There would be an accumulation at both limits of the distribution (i.e. at 0 and 1) in populations with large genetic or environmentally determined variability (e.g. Table 1, row 4 and 5; Table 3). Further, in comparing populations or in studying the effects of selection, we would encounter difficulties arising from the interdependence of the mean and variance of probability distributions. Biologically, the probability approach suffers in two ways. Unlike the underlying variation, we have no grounds, *a priori*, for assuming that π has a normal distribution, truncated between 0 and 1. Of more importance, an answer in terms of probability is unsatisfactory because it ignores the presence of the underlying variable. In the present case, we imagine this to be the concentration of some substrate affecting eye development.

To avoid these difficulties we employ the following stratagem:

(3) We assume that each side of the fly has its own supply of m.s. which is largely determined by the individual's genotype and the environment it encounters before the critical period in development. We further assume that there are a number of "random" factors—so-called developmental accidents which cause the amount of m.s. to differ between eyes. Thus in a normal population, the amount of m.s. is a variable (η) which has two components; one component (ξ) is determined by the genotype and the environment while the other (δ) is determined by developmental noise. One basic and central assumption is that the SD of δ is constant, irrespective of the genotype or the environment. This gives a universal constant, so to speak, which permits measurement of changes in the mean value of ξ for a population. Unfortunately it does not seem possible to test this important assumption directly.

(4) The argument may now be taken one step further; we know that the value of ξ is determined by background genotype and the environment. Therefore, in any real population, ξ , also, is a variable whose distribution should be normal. Obviously this variable, which I have called the penetrance parameter, is of more significance than η , at least to the geneticist. Consequently it is important to be able to measure the mean and variance of ξ for any population.

Fortunately, it is a simple operation to relate Z_1 , Z_2 and Z_3 , the respective proportions of the three phenotypic classes $(++)$, $(L+)/(+R)$ and (LR) to the mean, x , and standard deviation, γ , of ξ (Appendix I). In practical application it merely involves reference to a table for the estimation of x and γ . Should we require an answer in terms of η , we use the following transformation:

Mean value of η for the population = x (from equation (2)).

SD of η for the population = $(1+\gamma^2)^{1/2}$ (from equation (4)).

Perhaps the most surprising result from the present work is the high developmental noise component witnessed under normal laboratory conditions. One might expect that the variability in η attributable to developmental errors would not exceed the variability caused by changes in the general environment, where there occur temperature fluctuations of the order of one centigrade degree together with considerable changes in the medium due to aging. It will be shown in the next paper that both these factors affect the value of ξ . Despite this source of variation, developmental noise accounts for 80% of the variance in η in the inbred line and 18% in the outcrossed line (see Table 5). LATTER (1964), using a similar biological model but a different statistical analysis, found that developmental noise accounted for more than 60% of the phenotypic variance on the underlying scale of each of the three characters, sternopleural hair number, abdominal hair number and scutellar bristle number in *Drosophila melanogaster*. These findings give some indication of the difficulty in maintaining symmetry during development.

It was suggested earlier that other examples may exist which are suited to a similar form of analysis of the continuous underlying variation. In each case there is a breakdown in symmetry. However, this one criterion is not sufficient. It must first be established that, in the absence of genetic and environmental variation, the phenotypic classes (++) , (L+), (+R) and (LR) occur in frequencies consistent with the binomial distribution. In reports pertaining to these mutants, raw data are not available except for eyeless (MORGAN 1929) and antennaless (GORDON and SANG 1941). The evidence suggests that both these mutants are amenable to the analysis outlined above. GORDON and SANG have measured the frequency of (++) , (L+)/(+R) and (LR) for an inbred line, homozygous for antennaless. In particular, they studied the effect of the age of medium on penetrance. The point of relevance is that the results for each day

TABLE 6

Comparison between conventional method of measuring penetrance and the method outlined in this paper (using Gordon and Sang's data on antennaless)

Day	(++)	(L+) or (+R)	(LR)	α^*	β^*	Mean and sd of ξ
1	2	19	174	0.99	0.92	2.0 $\pm$ 0.8
2	11	73	69	0.93	0.69	0.5 $\pm$ 0.0
3	52	110	57	0.76	0.52	0.3 $\pm$ 0.0
4	13	24	8	0.71	0.44	...
5	27	42	36	0.74	0.54	0.12 $\pm$ 0.64
6	75	67	80	0.67	0.51	0.1 $\pm$ 1.1
7	46	76	75	0.77	0.57	0.24 $\pm$ 0.7
8	23	77	158	0.91	0.76	0.87 $\pm$ 0.67
9	21	33	87	0.83	0.74	1.0 $\pm$ 1.2
10	1	11	91	0.99	0.93	1.6 $\pm$ 0.2
11	1	0	29	0.97	0.97	2.5 $\pm$ 0.6

* α and β are the measures of penetrance used by GORDON and SANG where $\alpha = Z_2 + Z_3$ (i.e. proportion of flies not (++)); and $\beta = Z_3 + 0.5Z_2$ (i.e. proportion of sites with missing antennae).

represent samples in which both genetic and environmental variabilities are low, and therefore should suffice to test antennalness. The daily totals are recorded in Table 6 together with the mean and standard deviation of ξ and the measures of penetrance used by GORDON and SANG. The standard deviation varies somewhat from sample to sample. However, the higher values can be correlated with pooled heterogeneous samples which give the daily totals and consequently it may be supposed that environmental variability was higher in these cases. On no day is the frequency of the asymmetrical class significantly higher than that predicted by the binomial theory. Hence, the evidence suggests that in the total absence of genetic and environmental variability, a population of flies homozygous for the antennalless mutant would yield phenotypic frequencies in accordance with the binomial theory.

Although I have suggested that measuring the underlying variation of such characters as witty and antennalless by probits is basically unsound, it is nevertheless very interesting to compare results using both methods. A better comparison of the two methods is obtained by constructing a table of corresponding means and standard deviations for the two models rather than by analysing a particular set of data using both methods (see Table 7). For a given mean and standard deviation of the penetrance parameter, Z_1 , Z_2 and Z_3 have been estimated from the computed tables (not included in this paper). The triplet is then used to estimate the mean and standard deviation of the underlying variation using the probit transformation (FISHER and YATES 1938; SANG 1963).

The similarities between the use of probits and the penetrance parameter are striking, but there are differences and these are significant, particularly in lines with little variation of the underlying variable (e.g. SH). Statistically, one must

TABLE 7
Comparison between probits and penetrance parameter ξ

Penetrance parameter		Probits analysis	
Mean	Standard deviation	Mean	Standard deviation
0	0	0.5	0.74
0	0.5	0.5	0.87
0	1	0.5	1.17
0	2	0.5	1.93
0	4.2	0.5	3.79
1	0	1.39	0.71
1	0.5	1.40	0.85
1	1	1.39	1.15
1	2	1.39	1.93
1	4.2	1.39	3.82
2	0	2.0	0.0
2	1	2.24	1.12
2	2	2.28	1.92
2	4.2	2.28	3.80
3	2	3.16	1.92
3	4.2	3.16	3.79

concede that the probit analysis is generally satisfactory for partitioning phenotypic variance into its genetic and environmental components. However, it involves several biological shortcomings and it does not permit as complete an analysis as is allowed by the penetrance parameter suggested in this paper.

The model as it has been outlined so far allows no more than a yes/no classification for each side of the individual. With little modification of the model, cases could be analyzed where there is repetition of the organ, e.g. bristle characters. It would be necessary to assume that each organ is independently determined and that each organ has the same probability of abnormal development. Granted these assumptions we would simply need to replace the binomial distribution with a polynomial of the appropriate order. However, neither condition mentioned above holds with the scutellar system in *Drosophila* (RENDEL 1965) nor is likely to hold in any real situation. Consequently the analysis seems to have no immediate application to repeated organs.

A second possibility showing more promise involves cases similar to cross-veinless. Here both sides are classified as normal or abnormal. The abnormal side may be further classified as to the degree of abnormality. We should be able to use this extra information to investigate the assumption that developmental noise is independent of genotype and the environment.

I wish to thank DR. W. D. JACKSON for his helpful criticisms and suggestions throughout the course of this work. I should also like to thank DR. A. M. HASOFER who supplied invaluable discussion on a prototype of the model proposed in this paper. Help was given by MR. J. DONALDSON in several aspects of the computer programmes. I am indebted to DR. B. D. H. LATTER for his critical reading of the manuscript, and to PROFESSOR E. DEMPSTER for several helpful suggestions.

The work was done while the author held a senior research fellowship granted by the Wheat Industry Research Council of Australia.

SUMMARY

A procedure for measuring penetrance in cases involving asymmetry is presented and applied to a new eye mutant, *wi*: witty. By assuming (1) that the value of some underlying variable must surpass a threshold for abnormal development and (2) that each side of a fly has a particular value which is determined by the genotype, the general environment and developmental noise, it has been possible to determine the contribution to the underlying variation in a particular population due to these three sources. A new parameter, termed the penetrance parameter, whose value is determined entirely by the genotype and the general environment is defined. Reasons are given for preferring the penetrance parameter to probits or the classical methods of measuring penetrance where asymmetry occurs.

LITERATURE CITED

- FALCONER, D. S., 1960 *Introduction of Quantitative Genetics*. Oliver and Boyd, London.
 FISHER, R. A., and F. YATES, 1938 *Statistical Tables*. Oliver and Boyd, London.
 GOODWIN, E. T., 1949 The evaluation of integrals of the form $\int f(x)e^{-x^2}dx$. Proc. Camb. Phil. Soc. Math. Phys. Sci. **45**: 241-245.

- GORDON, C., and J. H. SANG, 1941 The relation between nutrition and exhibition of the gene antennaless. *Proc. Roy. Soc. Lond. B* **130**: 151-184.
- HANSEN, A. M., and E. J. GARDNER, 1962 A new eye phenotype in *Drosophila melanogaster* expressed only at temperatures above 25°C. *Genetics* **47**: 589-598.
- LATTER, B. D. H., 1964 Selection for a threshold character in *Drosophila*. *Genet. Res.* **5**: 198-210.
- MORGAN, T. H., 1929 Variability of eyeless. *Carnegie Inst. Washington Pubn.* **399**.
- RENDEL, J. M., 1962 The relationship between gene and phenotype. *J. Theoret. Biol.* **2**: 296-308. ——— 1965 Bristle pattern in scute stocks of *Drosophila melanogaster*. *Am. Naturalist* **99**: 25-32.
- ROBERTSON, A., 1965 Variation in scutellar bristle number—an alternative hypothesis. *Am. Naturalist* **99**: 19-23.
- SANG, J. H., 1963 Penetrance, expressivity and thresholds. *J. Heredity* **154**: 143-151.

APPENDIX I

Mean and Standard Deviation of the Penetrance Parameter

Our aim is to obtain a relationship between the mean, $\bar{x}$, and the standard deviation, y , of the penetrance parameter, ξ , and Z_1 , Z_2 and Z_3 , the proportions of the phenotypic classes, $(++)$, $(L+)/(+R)$ and (LR) , respectively.

Assume ξ and Π are related as described in Section 3, i.e.

$$p = \frac{1}{\sqrt{2\pi}} \int_{-u}^{\infty} e^{-\frac{1}{2} v^2} dv \quad (1)$$

The probability {of an eye being abnormal given that $\xi = u$ } = p and the probability {that both eyes are abnormal given that $\xi = u$ } = p^2

Thus the proportion of (LR) flies in a population, Z_3 , is equal to the probability of both eyes being abnormal for all the values that ξ assumes in the population. Since ξ is normally distributed with mean, $\bar{x}$, and standard deviation, y , it follows that

$$Z_3 = \frac{1}{\sqrt{2\pi}y} \int_{-\infty}^{\infty} p^2 e^{-\frac{1}{2} \left(\frac{u-\bar{x}}{y}\right)^2} du$$

$$\text{Likewise} \quad Z_2 = \frac{1}{\sqrt{2\pi}y} \int_{-\infty}^{\infty} 2p(1-p) e^{-\frac{1}{2} \left(\frac{u-\bar{x}}{y}\right)^2} du \quad (2)$$

$$\text{and} \quad Z_1 = \frac{1}{\sqrt{2\pi}y} \int_{-\infty}^{\infty} (1-p)^2 e^{-\frac{1}{2} \left(\frac{u-\bar{x}}{y}\right)^2} du$$

Thus if we obtain from an experiment the values of Z_1 , Z_2 and Z_3 we can evaluate the mean, $\bar{x}$, and standard deviation, y , of ξ . However, this would involve solving (2) for every triplet Z_1 , Z_2 and Z_3 which in itself is no mean task. The practical evaluation of $\bar{x}$ and y has been made feasible by computing a table of the triplets Z_1 , Z_2 , Z_3 for a range of values of $\bar{x}$ and y using the Elliott 503 Computer at the University of Tasmania. To do this (2) must be converted to a more suitable form for tabulation.

$$\text{Let} \quad \text{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-t^2} dt$$

After a little algebra we find

$$\begin{aligned} Z_3 &= \frac{1}{4\sqrt{\pi}} \int_{-\infty}^{\infty} (1+k)^2 e^{-v^2} dv \\ Z_2 &= \frac{1}{2} - \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} k^2 e^{-v^2} dv \\ Z_1 &= \frac{1}{4\sqrt{\pi}} \int_{-\infty}^{\infty} (1-k)^2 e^{-v^2} dv \end{aligned} \quad (3)$$

where $k = \operatorname{erf}(vy + x/\sqrt{2})$

Making use of the following relationship (see Goodwin, 1949)

$$\int_{-\infty}^{\infty} f(x) e^{-x^2} dx = \frac{1}{h} \sum_{n=-\infty}^{\infty} f(n/h) e^{-(n/h)^2} - E(1/h)$$

it can be shown that, for $h = 1$, the error term $E(1/h)$ is small and decreases extremely rapidly as h increases. We find that for $h = 4$ and for integral n over the range, $-16 \leq n \leq +16$, the approximation yields values of Z_1 , Z_2 , Z_3 correct to four decimal places for $y < 2.4$. When $y > 2.4$ it is necessary to let $h = 8$ and summate for n over the range -32 to $+32$.

$$\text{Thus we have } Z_3 \simeq \frac{1}{4h\sqrt{\pi}} \sum_{n=-4h}^{4h} (1+b)^2 e^{-(n/h)^2}$$

$$Z_2 \simeq \frac{1}{2} - \frac{1}{2h\sqrt{\pi}} \sum_{n=-4h}^{4h} b^2 e^{-(n/h)^2} \quad (4)$$

$$Z_1 \simeq \frac{1}{4h\sqrt{\pi}} \sum_{n=-4h}^{4h} (1-b)^2 e^{-(n/h)^2}$$

where $b = \operatorname{erf}(ny/h + x/\sqrt{2})$

We are now in a position to tabulate Z_1 , Z_2 and Z_3 for any given values of x and y .

APPENDIX II

The Calculation of the Mean and Standard Deviation of the Penetrance Parameter, ξ , for the (+ +) and (LR) Subpopulations

The r^{th} moment, about the origin, for a continuous distribution function, $f(u)$, may be given by

$$m_r = \frac{\int_{-\infty}^{\infty} u^r f(u) du}{\int_{-\infty}^{\infty} f(u) du}$$

Now the distribution function of ξ for the (+ +) subpopulation is given by

$$\frac{(1-p)^2}{y\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{u-x}{y}\right)^2}$$

Thus for the $(++)$ subpopulation the mean of ξ , $x_{(++)}$, is given by

$$x_{(++)} = \frac{\frac{1}{y\sqrt{2\pi}} \int_{-\infty}^{\infty} u(1-p)^2 e^{-\frac{1}{2}\left(\frac{u-x}{y}\right)^2} du}{\frac{1}{y\sqrt{2\pi}} \int_{-\infty}^{\infty} (1-p)^2 e^{-\frac{1}{2}\left(\frac{u-x}{y}\right)^2} du}$$

After a little algebra we have

$$\begin{aligned} x_{(++)} &= \frac{1}{4\sqrt{\pi}z_1} \int_{-\infty}^{\infty} (\sqrt{2}yv + x) (1-k)^2 e^{-v^2} dv \\ &= x + \frac{\sqrt{2}y}{4\sqrt{\pi}z_1} \int_{-\infty}^{\infty} v(1-k)^2 e^{-v^2} dv \\ &\simeq x + x_1 \end{aligned}$$

where

$$x_1 = \frac{\sqrt{2}y}{4h^2\sqrt{\pi}z_1} \sum_{n=-4h}^{4h} n(1-b)^2 e^{-(n/h)^2}$$

and where k and b are as defined in Appendix I.

It is easily shown that the S.D. of ξ for the $(++)$ subpopulation, $y_{(++)}$ is given by

$$y_{(++)} \simeq \left[\frac{2y^2}{4h^3\sqrt{\pi}z_1} \sum_{n=-4h}^{4h} (n^2(1-b)^2 e^{-(n/h)^2} - x_1^2) \right]^{\frac{1}{2}}$$

Likewise it can be shown that the mean and S.D. of ξ for the (LR) subpopulation is given by

$$\begin{aligned} x_{(LR)} &\simeq x + x_3 \\ y_{(LR)} &\simeq \left[\frac{2y^2}{4h^3\sqrt{\pi}z_3} \sum_{n=-4h}^{4h} n^2(1+b)^2 e^{-(n/h)^2} - x_3^2 \right]^{\frac{1}{2}} \end{aligned}$$

where

$$x_3 = \frac{\sqrt{2}y}{4h^2\sqrt{\pi}z_3} \sum_{n=-4h}^{4h} n(1+b)^2 e^{-(n/h)^2}$$

and k and b are as given in Appendix I.

Thus given Z_1 , Z_2 and Z_3 , it is possible to estimate x and y for the whole population and then to estimate the mean and S.D. of ξ for either the $(++)$ or (LR) subpopulation.