AN ANALYSIS OF THE INSTABILITY OF SEGREGATION-DISTORTER IN *DROSOPHILA MELANOGASTER*

GEORGE L. GABOR MIKLOS* AND SPENCER SMITH-WHITE

School of Biological Sciences, University of Sydney, Australia

Received August 3, 1970

IN 1959 SANDLER, HIRAIZUMI and SANDLER reported the discovery of the Segregation-Distorter locus (*SD*) in *Drosophila melanogaster*. In subsequent papers SANDLER and HIRAIZUMI described a number of puzzling properties of the *SD* system, including translocational modification and instability.

An understanding of the problems of segregation distortion has been hampered by the lack of a suitable method for the analysis of experimental data. The present paper offers such a method; it does not attempt to develop causal explanations.

SANDLER and HIRAIZUMI (1960) reported the phenomenon of instability. When original *SD* chromosomes were examined by crossing *SD/SD*⁺ males with homozygous *SD*⁺ females, the ratio of *SD* to total gametes recovered (which is defined as the *k* value) was in excess of 0.95. However, when recombinant *SD* chromosomes were obtained from *SD/SD*⁺ females, there was an increased male-to-male variation in *k* value, and stocks were defined as being stable, semistable, or unstable, according to the degree of this variation. A stabilizing modifier gene, *St*, was postulated, and located on the right arm of chromosome 2. When *St* was lost from an original *SD* chromosome as a consequence of crossing over, the recombinant chromosome yielded less stable *k* distributions. The hypothesis was examined by introducing *St* into both semistable and unstable lines; it was found that unstable lines were only partially stabilized, whereas semistable ones were rendered fully stable.

The mechanism of instability was envisaged as involving two or more "*SD* states" either at the organismal or cellular levels. It was argued that each male in an unstable line has its own *k* value or *SD* state, and that there is a high mutability from one *SD* state to another, since the sons of males with both high and low *k* values yield similar unstable distributions. Alternatively, it was thought that an *SD* state characterizes each cell, and the *k* value of a male is determined by the proportions of meiocytes in each of several available states.

The SANDLER and HIRAIZUMI hypothesis involves the introduction of new concepts; of *SD* states, *SD*-state mutation, and a stabilizer which controls this mutation. Our analysis dispenses with these assumptions and allows predictions on the relationship of the value of *k* and its variability.

* Present address: Department of Biology, University of California, San Diego, La Jolla, California 92037.

THEORY OF THE MAKE ANALYSIS

The method of analysis involves a concept of "make," and is based on RENDEL's (1967) analysis of scutellar bristle development. *Make* (m) is defined as the total level of all systems leading to the extinction of SD^+ in heterozygous males. An extinction coefficient, e , is defined as the probability of extinction of SD^+ gametes, and this coefficient is directly related to k (Table 1). Extinction or non-extinction of SD^+ involves *alternative* responses, and therefore there must be some level of m which acts as a threshold or switch point. This threshold is designated P , and the extinction coefficient measures the proportion of gametes in which m exceeds P .

Make is assumed to have a normal distribution. It must be variable, otherwise all gametes in an individual would have values above or below P , and k values would be either 1.0 or 0.5. *Make* value is derived by a probit transformation of

TABLE 1

Symbols used in the make analysis

Roman symbols are used for statistics and Greek symbols for parameters

Symbol	Description
k, κ	The ratio of SD to total gametes recovered from an SD/SD^+ male. $k = \frac{SD}{SD + SD^+} \quad [0.5 \leq k \leq 1.0]$
K, K	The mean value of k derived from a number of males. The value of a stock, line, or family.
e	The extinction coefficient. The probability of extinction of SD^+ gametes from SD/SD^+ males. $e = \frac{2k - 1}{k} \quad [0 \leq e \leq 1.0]$
m	The <i>make</i> level of a gamete.
$\bar{m}$	The mean <i>make</i> level of the gametes in a male. It is related to e by the probit transformation.
$\bar{M}$	The mean <i>make</i> level of a number of males.
P	The threshold level of <i>make</i> . <i>Make</i> levels above P cause extinction of SD^+ .
V, σ^2	Experimental and theoretical variances.
S, σ	Standard deviations. Both V and S are used with subscripts.
$V_{k(s)}$	The sampling variance of individual k values. $V_{k(s)} = \frac{k(1 - k)}{n}$ <p>where n is the number of gametes tested.</p>
V_K, S_K	The between-male variance and standard deviation expressed at the K level. $S_K = S_M \cdot dK/dM$
$V_{K(s)}$	The sampling variance of a K value from N males $V_{K(s)} = \frac{K(1 - K)}{N \cdot H_n} \quad \text{approximately,}$ <p>where H_n is the harmonic mean of the n numbers.</p>
V_m	The within-male biological variance of m . It is eliminated by probit scaling.
V_M, S_M	The between-male biological variance and standard deviation of M .
V_G, V_E	The genetic and environmental components of V_M .

FIGURE 1.—The relationship of the extinction coefficient (e) and the *make* value ($\bar{m}$) for an individual male.

e , which in turn is derived from k , and m is scaled in standard deviation or probit units.

This method of analysis is economical, since it involves only the principle of total causation and the assumption of a normal distribution of m . It is illustrated in Figure 1, where the unshaded area under the curve measures the proportion e .

1. The make- k relationship:

The relationship between *make* and k value is fundamental to the method of analysis, and is derived from the relationships given in Table 1. The *make- k* curve is illustrated in Figure 2; it is slightly asymmetrical, and may be divided into five regions with different properties. There are: (a) upper and lower approximately asymptotic regions where dK/dM is very small; (b) an effectively linear mid-region where dK/dM is almost constant and maximal; and (c) upper and lower curvilinear regions where dK/dM changes rapidly with changing M . Values of K within these different regions have properties partially dependent on their positions.

2. *The variances*: Estimates of the k value of individual flies, or the K value of a line, are subject to errors of several kinds.

(a) *Sampling variance*: $V_{k(s)}$ is a binomial variance with zero value at $k = 1.0$ and a maximum of $1/(4n)$ at $k = 0.5$; it is inversely proportional to the number of gametes sampled. Sampling error operates at the given k value and should be estimated at this level, preferably with the use of the angular transformation (SANDLER and HIRAIZUMI 1961). These authors have shown that the total variance significantly exceeds the binomial variance. We consider this extra-binomial variance to be of biological origin.

(b) *Biological variance*: Although $\bar{m}$ and $\bar{M}$ must be estimated from experimental values of k and K , the action is in the opposite direction; k is a consequence of the relationship between $\bar{m}$ and the threshold P in terms of standard deviation units; and similarly K is a consequence of the relationship between $\bar{M}$ and P . When transforming such values to the k scale, it should be noted that

$$V_K = V_M (dK/dM)^2$$

$$S_K = S_M \cdot dK/dM.$$

For convenience, dK/dM values are given in Table 2.

By hypothesis, there is a variance in *make* level V_m within each individual,

FIGURE 2.—The general relationship of K value and *make*.

and probit scaling makes this value constant. Consequently V_m does not contribute to the variance V_M between individuals. V_M , however, is not restricted by probit scaling, and may be small or large for a given $\bar{M}$ value.

$V_{k(s)}$ and $V_{K(s)}$ may be made as small as is experimentally feasible, but this is not true of V_M , which is an estimate of the parameter σ^2_M . An increase in the number of flies used to estimate K only improves the reliability of V_M , which is measured by its own standard deviation.

(c) *The partition of V_M* : The biological variance may be partitioned into genetic and environmental components,

$$V_M = V_G + V_E.$$

TABLE 2

The relationship of dK/dM with K

k	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.5	.0000	.0210	.0400	.0555	.0685	.0805	.0915	.1015	.1115	.1205
0.6	.1300	.1387	.1470	.1550	.1617	.1680	.1740	.1793	.1850	.1895
0.7	.1935	.1960	.1990	.2015	.2030	.2048	.2055	.2060	.2060	.2050
0.8	.2040	.2015	.1988	.1955	.1915	.1867	.1813	.1760	.1695	.1618
0.9	.1540	.1447	.1340	.1230	.1113	.0975	.0820	.0645	.0465	.0260

FIGURE 3.—Theoretical differences of intramale variance on an absolute scale.

Figure 3 illustrates different conditions of intramale variance on an hypothetical absolute scale. Fly *b* is given a standard deviation twice that of a fly *a*, and there is a corresponding difference in their means. On a probit scale, the two normal distributions are identical.

Fly *b* might be more susceptible than fly *a* to environmental change. However, if it is assumed that the susceptibility of a stock conforms to the linear relationship,

$$V_E \text{ (absolute)} = C V_M \text{ (absolute)}$$

then V_E will be constant when measured on the probit scale. Under the range of microenvironments available within culture tubes, V_E will tend to be uniform and should be small: its maximum value must be less than the smallest experimental values of V_M . It follows that where large differences in V_M occur, these must be due to change in the genetic variance. V_G and consequently V_M should be relatively small for inbred stocks, and should be increased by outcrossing.

3. The following *predictions* may be made:

(a) V_M should be substantial, and should greatly exceed the sampling variance except at extreme values of K . The causes of variation between individuals are greater than those within individuals.

(b) If V_M is independent of $\bar{M}$ and constant, then (i) V_K will be small and effectively constant in the upper and lower asymptotic region of the *make-k* curve. The k distributions will be stable. (ii) V_K will be approximately independent of K , constant and maximal, where K lies in the mid-region of the curve ($0.62 \leq K \leq 0.91$). There will be uniform instability. (iii) V_K will change with K value, and will be of moderate size in the two curvilinear regions. There will be semistability.

The stability consequences are illustrated in Table 3 for six values of $\bar{M}$, each having a range of ± 0.3 probits. The predictions apply to the whole range of possible K values.

(c) These predictions are not seriously changed if V_M is variable but still independent of $\bar{M}$. Conformity will be statistical rather than absolute. A mean V_K

TABLE 3

The consequences of uniform variation of make value on the stability of K

<i>SD line</i>	$M \pm S$	<i>K</i> value range	Description
A	3.0 ± 0.3	0.997–0.995	Stable, high <i>K</i>
B	2.0 ± 0.3	0.96 –0.99	Semistable
C	1.0 ± 0.3	0.81 –0.91	Unstable
D	0.0 ± 0.3	0.62 –0.72	Unstable
E	-1.0 ± 0.3	0.53 –0.57	Semistable
F	-2.0 ± 0.3	0.50 –0.51	Stable, low <i>K</i>

derived from a number of similar *K* values should approximate a $V_M (dK/dM)^2$ by *K* curve.

(d) For *K* values lying in the mid-region of instability, distributions of *k* values will be approximately symmetrical. Above and below this region, the distributions will show increasing negative and positive skewness.

(e) Theoretical frequency distributions can be computed for given *K* and V_M values, and may be constructed in the form of histograms. Where published data do not allow the computation of experimental V_M values, a comparison of the shapes of theoretical and experimental histograms provides a comprehensive evaluation of the method.

4. *A comparison:* Some of the predictions of the probit analysis are not unique—they are inevitable. In particular, low variance and marked skewness must occur with very high or very low values of *K*. Very little information about $\bar{M}$ and V_M is provided by extreme values of *K*, and such are not critical for an evaluation of the method.

However, the predictions of the relationship of V_K with *K* over the whole range of *K* values, of the relationship of S_K with the curve of $S_M \cdot dK/dM$ plotted against *K*, and theoretical histogram shapes for given values of $\bar{M}$ and V_M are all precise and unique.

By comparison, the stabilizer-*SD*-state hypothesis proposed by SANDLER and HIRAIZUMI is not predictive. It neither explains the instability of intermediate *K* values, nor does it suggest any relationship of V_K with *K*. The probit method defines the stabilizer gene as one which causes a major upward shift in *K* value. If it moves an initial *K* into the upper asymptotic region of the *make-k* curve, it causes stability; but if it should move an initial value from the lower asymptotic region, it would cause instability. For an understanding of most *SD* data, the probit analysis shows that it is only necessary to assume the existence of major and minor genes which alter the level of *make*. Within lines, genetic heterogeneity for modifiers would cause differences in V_K at given *K* values.

EXPERIMENTAL EVALUATION OF THE METHOD

1. *Examination of published data:* The available data on segregation distortion do not allow precise computations of the *make* variance, V_M . However, comparison of experimental and theoretical histograms can show whether such data conform to normal distributions on the *make* scale, and can determine the approxi-

mate range of variation in V_M values. Such comparisons for histograms published by SANDLER and HIRAIZUMI (1959, 1960) and by SANDLER (1962) are illustrated in Figures 4 to 8. These have been redrawn to a uniform scale for ease of comparison, retaining class intervals but grouping classes below 0.52.

Theoretical histograms have been generated from K medians and variances chosen to approximate the experimental values. From the normal deviates of the class boundaries, $(M - \bar{M})/S$, the corresponding included areas of the normal curve were read from the table of probits, and successive differences were then taken. K medians have been used because, unlike the means, they are not affected by the skewness of the K distributions.

In the following comparisons, distributions covering the whole range of K values are examined, selecting where possible sets based on more than 100 flies. These are grouped in pairs which are related to the major subdivisions of the *make-k* curve of Figure 2.

(a) Stable distributions with K values above 0.97: The experimental histograms of Figure 4 (left) are closely matched by the theoretical ones. Although high K values yield little information about V_M , the distributions are compatible

FIGURE 4.—A comparison of experimental (left) and theoretical (right) histograms, after SANDLER and HIRAIZUMI (1960). Ordinates give the percentage of males in each k value class. The class interval width is 3 k units.

A: *SD-5/cn bw*.

B: *R(SD)-4, St/cn bw*.

FIGURE 5.—A comparison of experimental and theoretical histograms, after SANDLER and HIRAIZUMI (1960).

A: $R(SD)-4/cn\ bw.$

B: $R(SD-5)-24/cn\ bw.$

with variances of the order of 0.3 to 0.4, and the degrees of skewness are conformable with the *make* analysis.

(b) Semistable distributions with K in the range of 0.92–0.95: The correspondence between the experimental and theoretical distributions illustrated in Figure 5 is excellent, especially when allowance is made for sampling fluctuation in those classes which have low theoretical expectations.

(c) Unstable distributions with K in the range of 0.70–0.85: The distributions in this range have been divided into two categories; those derived from established inbred stocks, and those derived from lines newly constructed by outcrossing.

The first category is illustrated in Figure 6. Low variances are expected in inbred stocks, and the V_M values of about 0.3 are in agreement with this expectation. The experimental and theoretical histograms are similar in shape, although the high peak of Figure 6B is rather unsatisfactory.

Distributions of the second category, which are derived from outcrosses, are

highly unstable. The shapes of the experimental and theoretical histograms are again similar. It is surprising that such unusual histogram shapes as those of Figure 7A can be generated from normal *make* distributions merely by the introduction of high variances. The large fluctuations in class frequencies shown by distributions of the kinds illustrated in Figure 7 are expected on the basis of sampling error.

(d) Semistable distributions with K below 0.54: The two comparisons of Figure 8 illustrate somewhat different situations. In A, the distribution was derived from individuals of the fifth generation after the establishment of a new *SD* stock, and heterogeneity has presumably been reduced by inbreeding. In B, the distribution is from the immediate progenies of an outcross program, and a high heterogeneity is to be expected. Both experimental histograms are well matched by their theoretical ones, with the variance in B higher than in A.

(e) The comparisons which have been made establish the following conclusions: (i) All distributions conform with normal variation on the *make* scale. K values near the extremes exhibit degrees of skewness closely conformable with those expected on the basis of a normal distribution in the variation of *make*. (ii) Mean values of V_M are of the order of 0.3 to 0.4 in inbred standard stocks, and show no indication of dependence on K value. (iii) In newly constructed genotypes obtained by outcrossing, especially where the *SD* chromosomes has been

FIGURE 6.—A comparison of experimental and theoretical histograms, after SANDLER and HIRAIZUMI (1960).

A: $R(SD-5)-35/cn\ bw$.

B: $R(SD-36)-1/cn\ bw$.

FIGURE 7.—A comparison of experimental and theoretical histograms, after SANDLER and HIRAIZUMI (1959). The symbol V_M should replace V in each of histograms A and B.

A: $cn\ bw/bw^D \delta \times SD-5 \varphi \rightarrow SD-5/cn\ bw \delta \times cn\ bw \varphi$.

B: $SD-5/al; +/ru$.

passed through females, V_M is increased, and may rise as high as 2.0. It is a reasonable assumption that this increase is due to genetic heterogeneity.

2. *Testing the relationship of dK/dM with V_K :* One prediction of the *make-k* relationship is that at constant V_M and S_M , the directly computed V_K and S_K values should fit curves obtained by plotting $V_M(dK/dM)^2$ and $S_M \cdot dK/dM$, respectively, against K . If a reasonably homogeneous inbred SD stock is crossed as male parent to a number of inbred SD^+ stocks, the different F_1 's should have comparable degrees of homogeneity. Their genetic contributions (V_G) to V_M should therefore be comparable. If the environmental component V_E is generally uniform, then the V_M 's of the different F_1 's should be similar. However, the several F_1 's might well differ in the number and strength of modifiers, so that they would have different K values. It follows that the experimental standard deviations computed directly in terms of K should fit the $S_M \cdot dK/dM$ curve. Data are available for an examination of this prediction from experiments designed for another purpose, i.e., a search for the presence of modifiers in laboratory stocks.

(a) *Materials:* The $SD\ bw$ chromosome was obtained by PEACOCK and ERICK-

FIGURE 8.—A comparison of experimental and theoretical histograms.

A: SANDLER (1962), $R(SD)-4/cn\ bw; Su(SD)/Y$.

B: SANDLER and HIRAIZUMI (1959), $cn\ bw/bw^D \delta \times SD-5 \varphi \rightarrow SD-5/bw^D \delta \times cn\ bw \varphi$.

SON from an $SD-72/cn\ bw$ male. It does not carry the inversion $In(2R)NS$, and has a K value of above 0.99. The stock has been maintained by repeated backcrossing to $cn\ bw/cn\ bw$ for a period of more than three years. Twenty inbred laboratory stocks were used; exact descriptions are unimportant and are not given here.

The F_1 males of each $SD\ bw$ by stock mating were crossed to three $cn\ bw$ virgin females to determine k values. Parents were both less than four days old at the time of mating, and each family consisted of approximately 20 tested males. The females were allowed to deposit eggs over two 3-day periods at $25^\circ C$ and each mating produced about 200 progeny.

(b) *Results*: The results, in so far as they apply to the present topic, are illustrated in Figure 9, where the standard deviations are plotted against K values. The experimental points are derived from tests of between 20 and 120 males, and in some cases different F_1 families with similar K values have been lumped. The points are compared with a theoretical $S_M \cdot dK/dM$ curve with S_M constant at 0.4.

FIGURE 9.—Experimental points of S_K and the curve $\sigma_M \times dK/dM$ for $\sigma_M = 0.4$. Sampling variance $V_{k(s)}$ has not been removed. Its removal gives points which fit the curve of $0.35 dK/dM$.

(c) *Conclusions:* It may be concluded that the prediction of the *make* analysis is substantiated; the experimental points lie reasonably close to the theoretical curve. The data support the view that the *make-k* relationship is valid and biologically meaningful. They also support the inference that the environmental component of the variance is relatively uniform.

The authors wish to thank members of the Genetics Section, Division of Plant Industry, C.S.I.R.O., Canberra, Australia, the Genetics Department of the University of California, Berkeley, and the Department of Biology, University of California, San Diego, for the provision of facilities and helpful discussion.

SUMMARY

A method for the analysis of instability in segregation distortion in *Drosophila* has been presented. The method is based on a concept of an underlying complex system of control, measured as "*make*," and involves a modified probit transformation of K . It is assumed that variances on the *make* scale are normal.

The method allows precise predictions on the relationship of the between-male variance V_K with K , assuming that the variance on the *make* scale is independent of the mean *make* value. It eliminates the need for assumptions of *SD* states and mutation from one state to another. It changes the role of stabilizers from genes which control rates of state mutation to genes which control the level of *make*.

The predictions have been examined in two ways. Firstly by the comparison of experimental and theoretical histograms, and secondly by the fitting of experimental determinations of the standard deviation S_K to the theoretical dK/dM curve. The comparisons fully support the theoretical presentation.

Segregation distortion is controlled by systems which, when measured on the *make* scale, comply with the characteristics of the normal distribution. Thus the variance and the standard deviation on the *make* scale may be used to give a

proper measure of the reliability of estimates of K . Owing to extreme skewness, it is not feasible to estimate fiducial limits directly from K values.

An inspection of data provided by SANDLER and HIRAIZUMI (1959), for which they introduced the concept of translocational modification, shows that a more simple explanation may now be sufficient. The outcrossing involved in these experiments may have added modifiers to reduce the value of K , but it must also have increased genetic heterogeneity.

Finally, the method of *make* analysis specifies the kinds of question which have to be asked and answered in relation to phenomena of segregation distortion such as conditional distortion and deviant sex ratio.

LITERATURE CITED

- FISHER, R. A. and F. YATES, 1963 *Statistical Tables for Biological Agricultural, and Medical Research*. Hafner, New York.
- RENDEL, J. M., 1967 *Canalisation and Gene Control*. Academic Press, New York.
- SANDLER, L., 1962 A directed, permanent, genetic change involving the segregation-distortion system in *Drosophila melanogaster*. *Am. Naturalist* **96**: 161–165.
- SANDLER, L. and Y. HIRAIZUMI, 1959 Meiotic drive in natural populations of *Drosophila melanogaster*. II: Genetic variation at the segregation-distorter locus. *Proc. Natl. Acad. Sci. U.S.* **45**: 1412–1422. —, 1960 Meiotic drive in natural populations of *Drosophila melanogaster*. IV: Instability at the segregation-distorter locus. *Genetics* **45**: 1269–1287. —, 1961 Meiotic drive in natural populations of *Drosophila melanogaster*. VIII: A heritable aging effect on the phenomenon of segregation-distortion. *Canad. J. Genet. Cytol.* **3**: 34–46.
- SANDLER, L., Y. HIRAIZUMI and I. SANDLER, 1959 Meiotic drive in natural populations of *Drosophila melanogaster*. I: The cytogenetic basis of segregation-distortion. *Genetics* **44**: 233–250.