Theor Appl Genet (1990) 80:321–325

THEORETICAL AND APPLIED GENETICS

Prediction of heterosis in crosses between inbred lines of *Drosophila melanogaster*

N. G. Ehiobu, M. E. Goddard and J. F. Taylor

Graduate School of Tropical Veterinary Science, James Cook University of North Queensland, Townsville, Australia

Received August 15, 1989; Accepted March 6, 1990
Communicated by J. S. F. Barker

Summary. The aim of the experiment was to determine if the estimated genetic distance between two populations could be used to predict the amount of heterosis that would occur when they were crossed. Eight lines of known relatedness to each other were produced by eight generations of sib mating and sub-lining. This produced lines that varied in coefficient of coancestry from zero to 0.78. Fourteen reciprocal crosses of these lines were used to measure heterosis for larval viability and adult fecundity. Gene frequencies at six polymorphic enzyme loci were used to estimate the genetic distances between lines, which were then compared with the known degrees of coancestry. The estimated genetic differences were poorly correlated with the known coancestry coefficients ($r=0.4$), possibly due to the small number of loci typed. Also genetic distances were only about 1/3 of what was expected. Selection acting on blocks of genes linked to the enzyme loci probably prevented the expected increase in homozygosity. Coancestry coefficient was correlated with heterosis ($r=0.44-0.71$). This level of correlation implied differences in heterosis among parent lines with the same level of coancestry. This variability is expected if a small number of loci explain most of the heterosis. The average level of heterosis was less than expected after eight generations of sib mating. This is most likely due to selection opposing the increase in homozygosity caused by inbreeding. The combination of these two imperfect correlations resulted in no significant correlation between genetic distance estimated from markers and heterosis.

Key words: Heterosis – Genetic distance – *Drosophila* – Inbreeding

Introduction

It would be useful to be able to predict the amount of heterosis that would occur in crosses between particular breeds or lines of livestock or plants. Even if the prediction was not completely accurate, it could be used in selecting crosses to be compared in field experiments. Glodek (1974) suggested that the greater the genetic distance between breeds, calculated from gene frequencies at marker loci, the more heterosis would occur when they were crossed. Goddard and Ahmed (1982) and Ehiobu and Goddard (1990) developed a more specific but similar theory to predict heterosis.

Heterosis depends on two factors – the increase in heterozygosity in the F_1 compared to the parents, and the change in performance for a given change in heterozygosity. The proportional change in heterozygosity caused by inbreeding is conventionally measured by the inbreeding coefficient F and it is convenient to measure increases in heterozygosity by a similar F statistic (Ehiobu and Goddard 1990). The theory of Goddard and Ahmed (1982) uses gene frequencies at marker loci to estimate the increase in heterozygosity in the cross and experiments on inbreeding depression to estimate change in performance per percent F.

A model of the genetic divergence of population and breeds is one of repeated splitting of populations into subpopulations, which then evolve independently. With finite population size, gene frequencies will drift apart. If drift is the only force causing divergence, then the divergence at marker loci (e.g., enzyme loci) should apply to all loci, including those causing heterosis. The increase in heterozygosity is simply the amount of inbreeding that has occurred since the populations separated.

Ehiobu and Goddard (1990) tested this theory by measuring heterosis among geographically separated

populations of *Drosophila melanogaster* and found only limited success, an experiment was performed in which the degree of relationship between the lines was known independently of the genetic distance estimated from markers.

Eight lines of *D. melanogaster* with known relationships between them were produced by repeated creation of sub-lines during eight generations of sib mating. A partial diallel cross of these lines was performed and heterosis was estimated for larval viability and adult fecundity. Gene frequencies at ten enzyme loci were estimated using electrophoresis and used to estimate the amount of inbreeding that had occurred since each of the lines separated. This allowed testing whether (1) F calculated from marker loci correctly predicts the true relationships between the lines, which was known from the design of the experiment; (2) the relationship between lines predicts which crosses will show the most heterosis; (3) the average heterosis per percent F in the crosses is the same as inbreeding depression per percent F estimated from inbreeding experiments.

Materials and methods

Inbred line development and crossing

The lines originated from two wild, inseminated, assumed unrelated female *Drosophila melanogaster*. Sib mating and sub-lining were carried out for eight generations. Four lines of known relatedness to each other were produced from each of the two wild caught females as shown in Fig. 1. After the eighth generation of sib mating, the lines were maintained at large population sizes ($N=500$) to avoid further inbreeding.

Figure 1 includes inbreeding coefficients expressed as coancestries. The coancestry between any pair of lines (and therefore the inbreeding coefficient in their offspring) is given by the coancestry at the generation at which the lines separated. For instance, the coancestry between lines A and D was 0.25. The inbreeding coefficient within a line was 0.83.

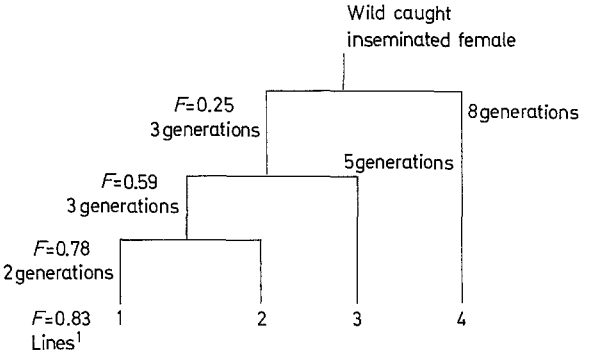

Fig. 1. System of sib mating and sub-lining used in the experiment to produce inbred lines with known relatedness to each other.

[1] Lines 1, 2, 3, and 4 correspond to lines A, B, C, and D from wild-caught female 1 and E, F, G, and H from wild-caught female 2, respectively

Fourteen reciprocal crosses were produced and compared with the eight inbred lines.

Electrophoresis

Starch gel electrophoresis using procedures described by Ehiobu (1985) was used. The enzyme loci assayed were 6-phosphogluconate dehydrogenase (6-*Pgd*), α-glycerophosphate dehydrogenase (α-*Gpdh*), Malic dehydrogenase (*Mdh*), Alcohol dehydrogenase (*Adh*), Esterase-6 (*Est-6*), Phosphoglucomutase (*Pgm*), Acid phosphatase (*Acph*), Octanol dehydrogenase (*Odh*), Xanthine dehydrogenase (Xdh), and larval Alkaline phosphatase (*Aph*). These enzymes were assayed 4 months after the completion of the sib mating.

Genetic distance

The heterosis in a cross is proportional to the difference in heterozygosity between the F_1 and the parent lines. In theory, this in turn is proportional to

$$F-C$$

where

F = inbreeding coefficient of the parent lines,
= 0.83 for our lines,
C = coefficient to coancestry between the parent lines.

Therefore, $(0.83-C)$ will be referred to as the theoretical genetic distance between line.

The proportional increase in heterozygosity for the marker loci was calculated as described by Ehiobu and Goddard (1990) and will be referred to as the genetic distance from markers.

Traits measured

Larval to adult survival (larval viability) and adult fecundity (eggs laid per day) were measured as described by Ehiobu et al. (1989) and Ehiobu and Goddard (1990). A combined trait index (CTI) was calculated by adding larval viability and fecundity together after dividing each by its standard deviation.

Data analysis

Least-squares analysis was carried out based on the model

$$y_{ijkl}=\mu+S_i+S_j+M_j+H_k+e_{ijkl},$$

where

y_{ijkl} = larval viability, fecundity or CTI for the l^{th} replicate from the i^{th} sire line mated to the j^{th} dam line,
μ = population mean,
S_i and S_j = effect of the sire and dam lines,
M_j = maternal effect of the j^{th} line,
H_k = heterosis for the k^{th} cross ($k=1, \ldots 14$),
e_{ijkl} = residual random error.

For further analysis the 14 crosses were grouped into four classes based on level of inbreeding and the heterosis term H_k was replaced in the model by

$$H_k=bF+I_n+C_{nk},$$

where

bF = regression on inbreeding coefficient,
I_n = deviation of mean of inbreeding class from regression line ($n=1, \ldots 4$),
C_{nk} = deviation of individual cross from mean of inbreeding level.

Simulation

To study the effect of selection on a single line during eight generations of full-sib inbreeding, a Monte Carlo simulation was carried out. The effect of large blocks of linked genes was approximated by a single locus at which selection acted equally against all homozygotes. Each generation, a random number generator was used to form parents of the next generation but a proportion s of homozygotes died. After eight generations of inbreeding, the inbreeding coefficient F at this locus was calculated. Two hundred runs at each level of selection ($s=0.25$, 0.5, 0.75, 1.0) were performed.

Table 1. Performance of inbred lines and their crosses

Line/cross[a]	Mean larval viability ±SE (%)	Mean fecundity ±SE (eggs/day)
A	72.0±4.3	25.5±5.0
B	86.0±2.0	29.7±5.6
C	74.0±3.9	31.0±4.7
D	90.0±2.3	30.3±3.1
E	56.5±7.0	58.1±1.5
F	53.0±5.7	38.6±2.6
G	59.0±3.1	55.0±2.2
H	62.0±5.0	41.0±3.3
AB	90.5±3.6	44.6±3.4
AC	56.5±5.1	39.4±5.0
AE	57.0±2.2	50.1±3.9
BD	84.0±2.4	42.7±3.9
BG	81.5±4.3	63.6±4.6
CD	85.5±1.7	56.3±4.3
CG	89.5±2.0	60.9±2.3
DF	91.0±1.5	69.7±2.6
DH	92.0±1.7	66.5±1.9
EF	61.5±3.4	51.1±3.8
EG	67.0±5.0	50.1±4.1
EH	84.0±4.5	54.6±2.7
FG	69.5±4.3	59.7±2.0
FH	84.0±2.5	52.6±4.8

[a] Reciprocal crosses combined

Results

Genetic distance

Despite the large standard errors which apply to the genetic distances, they were lower than theoretical distances (Table 2). Also there were large differences in genetic distance between pairs of lines that were theoretically the same distance apart, although there was a trend for genetic distance to increase with theoretical distance, as shown by the correlation of 0.40 in Table 3.

Performance of inbred lines and heterosis

Table 1 presents the performance of inbred lines and F_1 crosses. Differences between inbred lines in adult fecundity and larval viability were highly significant ($p<0.01$). Inbred lines can be categorized into two groups on the basis of their performance, A–D and E–H, corresponding to the two wild-caught female groups. Lines A–D had very poor fecundity but high larval viability relative to lines E–H. Maternal environment significantly ($P<0.01$) influences fecundity but not larval viability. Strain and maternal effects on CTI approached significance at the 5% level.

Heterosis estimates are presented in Table 2. Heterosis was significant for all traits. The breakdown of these 14 degrees of freedom showed that there was significant

Table 2. Expected inbreeding level, theoretical and genetic distances, and heterosis

Cross	Expected inbreeding	Distance measures		Heterosis		
		Theoretical	Genetic (F from markers) ±SE	Larval viability (%)	Fecundity (eggs/day)	CTI (Standard Deviation units)
AE	0.00	0.83	0.32±0.23	22.8	7.7	2.54
BG	0.00	0.83	0.11±0.08	9.0	21.9	2.13
CG	0.00	0.83	0.23±0.11	23.0	18.1	3.35
DF	0.00	0.83	0.24±0.12	19.5	25.1	3.37
DH	0.00	0.83	0.33±0.16	16.0	30.8	4.16
BD	0.25	0.58	0.24±0.16	−4.0	12.7	0.37
CD	0.25	0.58	0.30±0.21	3.5	25.7	1.92
EH	0.25	0.58	0.19±0.11	24.8	4.5	2.45
FH	0.25	0.58	0.01±0.02	26.5	2.3	2.44
AC	0.59	0.24	0.38±0.27	−16.5	11.3	−0.60
EG	0.59	0.24	0.05±0.02	9.3	−6.3	0.33
FG	0.59	0.24	0.05±0.05	13.5	2.8	1.22
AB	0.78	0.05	0.07±0.05	11.5	17.1	2.28
EF	0.78	0.05	0.15±0.09	6.8	−7.4	0.06
SE**				4.5	4.5	0.40

[a] Inbreeding equals the coancestry of the parent lines
** Standard error of each heterosis estimate

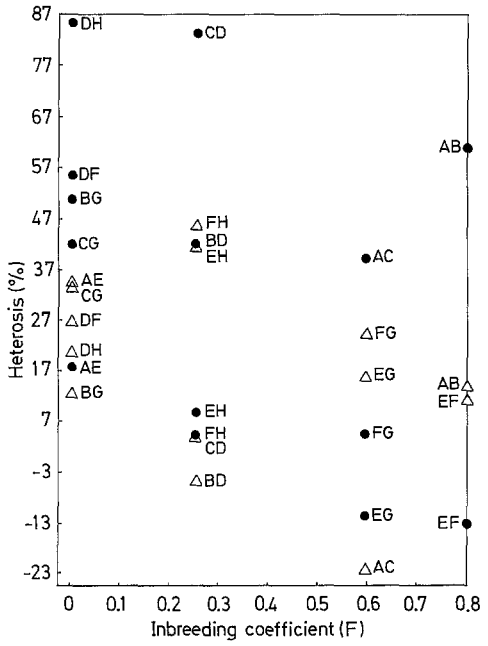

Fig. 2. Distribution of % heterosis estimates by inbreeding coefficient for fecundity (●) and larval viability (△)

Table 3. Correlation estimates between distance and heterosis measures in crosses between inbred lines

	Genetic distances		Heterosis	
	Theoretical	From markers	Viability	Fecundity
Genetic distances (markers)	0.406	–	–	–
Heterosis				
Larval viability	0.440	−0.342	–	–
Fecundity	0.604[b]	0.511[a]	−0.022	–
CTI	0.709[c]	0.094	0.757[c]	0.623[b]

[a] ($P<0.1$)
[b] ($P<0.05$)
[c] ($P<0.01$)

regression on inbreeding coefficient F and significant deviation of individual crosses from the mean of their inbreeding class. The regression of heterosis on F was $-21.4\% \pm 2.9\%$ per 100% inbreeding for larval viability and -22.2 ± 2.5 eggs/day per 100% inbreeding for fecundity. Inbreeding class means did not deviate significantly from the linear regression on F. Figure 2 shows the large variation in heterosis estimates for crosses at the same level of inbreeding. There was a tendency for low heterosis to occur in crosses between lines that had high purebred means.

Table 3 presents correlations between distance measures and heterosis estimates. Theoretical distance had some ability to predict heterosis but genetic distance estimated from markers did not.

Simulation

As the selection coefficient against homozygotes increased ($s=0.25$, 0.50, 0.75, 1.00), the inbreeding coefficient reached after eight generations of full-sib mating fell to 0.73, 0.56, 0.40, and 0.35, respectively. Therefore, under the assumption of our model, selection against homozygotes at a single locus cannot account for F less than 0.35 after eight generations of full-sib mating.

Discussion

Does the genetic distance calculated from markers correctly predict the true relationships between the lines? The accuracy of genetic distance estimates depends on the number of loci typed. Since only six polymorphic loci were used, the standard errors in Table 2 are rather large and quite compatible with the observed correlation of 0.4 between genetic distance and theoretical distance.

However, the estimated genetic distances were considerably less than the theoretical distances. For the unrelated pairs of lines in Table 3 (theoretical distance = 0.83), the average genetic distance from markers was 0.25. Selection acting on blocks of genes linked to the markers would oppose the increase in homozygosity caused by inbreeding (Sved 1975) and so tend to reduce genetic distances. However, the simulation results show that even intense within-line selection cannot reduce the inbreeding coefficient, and hence the genetic distance, below 0.35. Chance effects during the inbreeding process, selection between lines, and selection within lines after the end of inbreeding but before electrophoresis probably explain the low genetic distances observed. Rumball (1974) carried out 18 generations of full-sib mating in 120 independent lines of *D. Melanogaster* and found the homozygosity of marker loci was 80% of that expected.

Does the relationship between lines predict which crosses will show the most heterosis? The theoretical distance was significantly correlated with heterosis but crosses with the same theoretical distance varied greatly in heterosis. If populations diverge by genetic drift the inbreeding coefficient should predict the average difference in gene frequency between them, but individual loci will vary widely. Therefore, if one or a few loci explain most of the heterosis, then the correlation between theoretical distance and heterosis would not be large. The observation that theoretical distance is more closely correlated with the CTI than with larval viability or fecundity is in line with this expectation.

Is the mean heterosis per percent F theoretical distance the same as the inbreeding depression per percent F estimated from inbreeding experiments? Ehiobu et al. (1989) estimated the inbreeding depression from one gen-

eration of full-sib inbreeding ($F=0.25$) to be 24% for larval viability and 28 egg/day for fecundity. Heterosis in this experiment reached a similar level at $F=0.83$. Also Robertson and Reeve (1955) found that crosses between inbred lines gave more heterosis for fecundity than we observed. It seems likely that the selection that had reduced homozygosity at marker at loci had also reduced homozygosity at loci affecting viability and fecundity, and hence reduced heterosis in crosses. Ehiobu et al. (1989) found that slower rates of inbreeding, which allowed time for selection to act, resulted in less inbreeding depression than sib mating.

The results of inbreeding at individual loci are subject to large chance effects. Since the eight inbred lines are descended from two original females, they may not be typical of all inbred lines that could have been developed from the base population. However, this does not affect our major conclusions. The main purpose of the experiment was to determine whether genetic distance could predict which lines showed the most heterosis when crossed. The origin of the lines should not be crucial in this regard. Because they descended from two females, these lines may by chance have shown less inbreeding depression than the base population. However, it is most unlikely that this could account for the large difference between the inbreeding depression observed and that expected from the results of one generation of sib mating by Ehiobu et al. (1989).

Existing breeds of livestock also represent only a small sample of all possible breeds that might have evolped. Their slower rate of inbreeding would not by itself reduce the effects of chance. Thus, breeds of livestock may, by chance, show more or less divergence at some loci than expected, and this will reduce our ability to predict the amount of heterosis that will occur, just as it has in this experiment. It is possible, however, that with slower inbreeding, recessive genes with large deleterious effects would not become common in any line, and so the effective number of loci controlling heterosis would be greater and the results more predictable.

Conclusion

In this experiment, genetic distance estimated from marker loci had no ability to predict heterosis. The theoretical distance based on the coancestry of pairs of lines was significantly correlated with heterosis. If genetic distance was estimated more precisely by using a larger number of markers, the correlation between genetic distance and heterosis would presumably be closer to that between theoretical distance and heterosis. Even then the correlation would be less than unity because of the limited number of loci controlling most of the heterosis in some traits. For this reason heterosis for a combination of traits, such as the CTI, should be more predictable than heterosis for individual traits. Selection tends to reduce the divergence of lines caused by inbreeding both at marker loci and at the loci causing heterosis for quantitative traits.

References

Ehiobu NG (1985) The prediction of heterosis. PhD thesis, James Cook University, Townsville, Australia, pp 61–62

Ehiobu NG, Goddard ME (1990) Heterosis in crosses between geographically separated populations of *Drosophila melanogaster*. Theor Appl Genet (in press)

Ehiobu NG, Goddard ME, Taylor JF (1989) Effect of rate of inbreeding on inbreeding depression of *Drosophila melanogaster*. Theor Appl Genet 77:123–127

Glodek P (1974) Specific problems of breed evaluation and crossing in pigs. In: Proc Working Symp Breed Evaluation and Crossing Experiments with Farm Animals. Schoonerd, Zeist: Research Institute of Animal Husbandry, Zeist, pp 267

Goddard ME, Ahmed AM (1982) The use of genetic distance between cattle breeds to predict the heterosis in crosses. In: Proc 2nd World Congr Genet Appl Livestock Prod, Madrid, Spain, pp 377–382

Robertson FW, Reeve ECE (1955) Studies in quantitative genetics. VIII. Further analyses of heterosis in crosses between inbred lines of *Drosophila melanogaster*. Z. Indukt Abstamm Vererbungsl 86:439–458

Rumball W (1974) The decrease in heterozygosity produced by close inbreeding. Part of PhD thesis, Macquarie University, Sydney

Sved JA (1975) Fitness of third chromosome homozygotes in *Drosophila melanogaster*. Genet Res 25:197–200