Lecture 6

Inbreeding and Crossbreeding

Bruce Walsh. July 2005. Asian Institute on Statistical Genetics

Changes in the Mean and Variance Under Inbreeding

Inbreeding

Inbreeding (mating of related individuals) often results in a change in the mean of a trait compared with its value in a random-mating population. Its importance is that inbreeding is generally harmful and reduces fitness. In particular, inbreeding often causes a reduction of the mean value for quantitative traits associated with reproduction and viability.

Inbreeding is intentionally practiced to:

- create genetic uniformity of laboratory stocks
- produce stocks for crossing (animal and plant breeding)

Inbreeding is unintentionally generated:

- by keeping small populations (such as is found at zoos). *Genetic drift* is a special case of inbreeding. The smaller the population, the quicker inbreeding accumulates.
- during selection (which has the effect of reducing the population size relative to the unselected case).

The critical parameter for describing inbreeding is F, *the probability that the two alleles at a locus in an individual are identical by descent*. In an individual inbred to amount F, a randomly-chosen locus has both alleles IBD with probability F and hence is a homozygote.

To compute the genotypic probabilities under inbreeding, suppose we chose a locus at random. Denote the frequency of allele A_1 by p and the freq(A_2) by q. With probability F the two alleles are IBD, and hence this locus is always homozygous, with freq(A_1A_1) = p and freq(A_2A_2) = $q = 1-p$. If the alleles are not IBD, then the genotypic frequencies follow Hardy-Weinberg. Thus, the expected genotypic frequencies under inbreeding become

Genotype	Alleles IBD	Alleles not IBD	Population frequency
A_1A_1	$F \cdot p$	$(1-F)p^2$	$p^2 + Fpq$
A_2A_1	0	$(1-F)2pq$	$(1-F)2pq$
A_2A_2	$F \cdot q$	$(1-F)q^2$	$q^2 + Fpq$

If the genotypes A_1A_1, A_1A_2, A_2A_2 have values of $a, d, -a$, then the mean under inbreeding becomes

$$\begin{aligned}\mu_F &= a \cdot (p^2 + Fpq) + d \cdot (1-F)2pq - a \cdot (q^2 + Fqq) \\ &= a(2p-1) + 2(1-F)pqd\end{aligned}$$

Noting that the mean character value in a random mating population ($F = 0$) is

$$\mu_0 = a(2p-1) + 2pqd,$$

then the mean under inbreeding can be expressed as

$$\mu_F = \mu_0 - 2Fpqd \tag{6.1a}$$

More generally, if there are k loci, the mean is

$$\mu_F = \mu_0 - 2F \sum_{i=1}^{k} p_i\, q_i\, d_i = \mu_0 - B\, F \tag{6.1b}$$

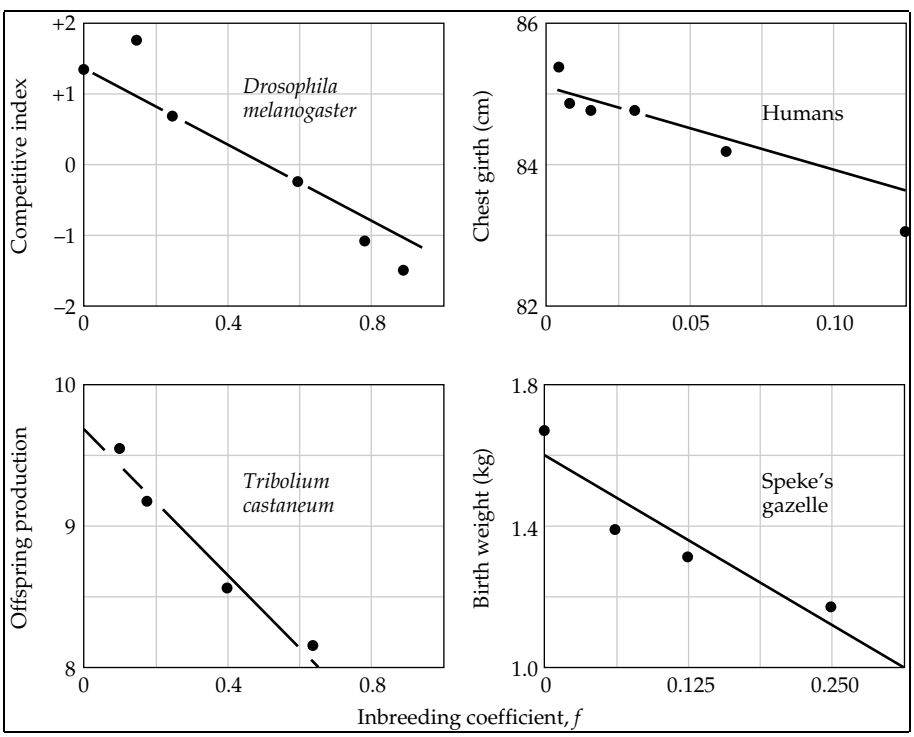
where $B = 2 \sum p_i\, q_i\, d_i$ is the reduction in the mean under complete inbreeding ($F = 1$).

Hence,

- there will be a change of mean value under inbreeding only if $d \neq 0$, i.e., dominance is present.
- for a single locus, if $d > 0$, inbreeding will decrease the mean value of the trait. If $d < 0$, inbreeding will increase the mean.
- with multiple loci, a decrease in the mean under inbreeding (**inbreeding depression**) requires **directional dominance**, with the dominance effects d_i tending to be positive. $d > 0$ occurs when the heterozygote is closer in value to the larger homozygote.
- the magnitude of the change of mean on inbreeding depends on gene frequency, and is greatest when $p = q = 0.5$

Inbreeding Depression in Fitness Traits

Fitness-related traits (such as viability, offspring number and body size) often display inbreeding depression, as the following examples illustrate:



Computing the Inbreeding Depression Coefficient, B

In many cases, lines cannot be completely inbred due to either time constraints and/or because in many species lines near complete inbreeding are nonviable. In such cases, one must estimate the inbreeding depression using a series of lines under partial inbreeding (lines with $F < 1$).

Recalling that

$$\mu_F = \mu_o - BF \tag{6.2}$$

it immediately follows that the slope of the regression of the population mean μ_F on the inbreeding coefficient F estimates the inbreeding depression coefficient B.

The above equation is true if loci combine additively (no epistasis), in which case the change in mean should be directly proportional to F (i.e., a linear function of F), the inbreeding coefficient. If epistasis is present, the change in mean can be a nonlinear (polynomial) function of F. Hence, if epistatis is absent, we expect a linear regression of mean on F to be an adequate fit of the data.

Why do traits associated with fitness show inbreeding depression?

Two competing hypotheses have been proposed:

- **Overdominance Hypothesis:** Genetic variance for fitness is caused by loci at which heterozygotes are more fit than both homozygotes. Inbreeding decreases the frequency of heterozygotes, increases the frequency of homozygotes, so fitness is reduced. Since some inbred lines have means for fitness traits equal to the base population, this explanation cannot be generally true.
- **"Dominance" Hypothesis:** Genetic variance for fitness is caused by rare deleterious alleles that are recessive or partly recessive (heterozygote fitness closer to the fitness of the wildtype). Such alleles persist in populations because of recurrent mutation. Most copies of deleterious alleles in the base population are in heterozygotes. Inbreeding increases the frequency of homozygotes for deleterious alleles, so fitness is reduced. Under the dominance hypothesis, one of the homozygotes has the highest fitness.

While the dominance hypothesis is sufficient to account for inbreeding depression, even a very small fraction of overdominant loci can have a major effect on the level B. Hence, even though most loci that contribute to inbreeding depression may due to uncovering of deleterious recessives, the bulk of the contribution to inbreeding depression could theoretically come from a much smaller fraction of overdominant loci.

Minimizing the Rate of Inbreeding

Given the detrimental effects of inbreeding, breeders and zoo keepers go to great lengths to minimize the rate of increase in F. Inbreeding occurs in any finite population, but the larger the effective population size N_e, the slower the effects of drift. One standard way for maximizing N_e is to ensure that all individuals make an equal contribution of offspring to the next generation, as the effective population decreases as the variance in offspring number increases. Maximal N_e occurs when each male and female in the population leaves exactly the same number of offspring. Alas, in many breeding situations there is a very skewed sex ratio, often due to constraints of reproductive biology (a bull can leave effectively an infinite number of offspring, while a cow typically has only one a year).

Gowe et al (1959) suggest that when the sex ratio of contributing parents is r females to each male, that every male should contribute (exactly) one son and r daughters, while every female should leave one daughter and also with probability $1/r$ contribute a son. Wang (1997) improved on this by imposing that a female contributing a son does not contribute a daughter, but instead that another female from the same male family contributes two daughters.

Change of Variance With Inbreeding

Inbreeding causes a re-distribution of genetic variance within and between lines. For completely additive loci, this can be expressed in terms of the genetic variance ($V_A = V_G$) present in the base population:

	General	$F = 1$	$F = 0$
Between Lines	$2FV_A$	$2V_A$	0
Within Lines	$(1-F)V_A$	0	V_A
Total	$(1+F)V_A$	$2V_A$	V_A

Inbreeding increases genetic variance between lines and decreases genetic variance within lines. With dominance, the expressions are not simple functions of the base population genetic variances, but rather depend on gene frequencies, so there is no simple expression for the re-distribution of variance. Indeed, when non-additive variance is present, the additive variance can actually increase (over some range) with F.

The heritability within any one inbred line (assuming only additive variance) is

$$h_t^2 = \frac{(1-F_t)V_A}{(1-F_t)V_A + V_E} = h_0^2 \frac{1-F_t}{1-h_0^2 F_t}$$

The right-most expression is expressed in terms of heritability in the base population.

Change of Variance with Inbreeding and Mutation

Inbred lines will never completely lose all genetic variance, because new mutational variance is introduced each generation at rate V_M. In the long term the genetic variance within an inbred line will reach an equilibrium level at which the variance gained each generation from mutation is exactly balanced by the variation lost by inbreeding. We consider the particular case where inbreeding is caused by genetic drift in a finite population, in which case the accumulation of inbreeding scales as $1/(2N_e)$, where N_e is the effective population size.

Assume:

- Strictly neutral mutations
- Strictly additive mutations
- Symmetrical distribution of mutational effects

Then at equilibrium

$$V_A = V_G = 2N_e V_M$$

Note that this is the same as that expected within a selection line in mutation-drift equilibrium. For example, with $V_M = 10^{-3}V_E$ and $N_e = 2$ for full sib inbreeding, V_G at mutation-drift equilibrium is $4 \times 10^{-3}V_E$, and the heritability is

$$h^2 = \frac{V_A}{V_G + V_E} = \frac{4 \times 10^{-3}V_E}{4 \times 10^{-3}V_E + V_E} = 0.004$$

which is trivial.

Mutation also contributes to the increase in variance between sublines derived from a common inbred line. The variance among lines from new mutation after t generations is

$$V_B = 2V_M[t - 2N_e(1 - e^{-t/2N_e})]$$

At equilibrium the rate of divergence is $2V_M$ per generation and the total divergence expected is $2tV_M$, which is not negligible.

Changes in the Mean Under Crossbreeding

When inbred lines are crossed, the progeny show an increase in mean for characters that previously suffered a reduction from inbreeding. This increase in the mean over the average value of the parents is called **hybrid vigor** or **heterosis**. Fitness lost on inbreeding is restored by crossing.

To see how heterosis is inbreeding depression in reverse, imagine a large number of inbred lines derived from an outbred base population in which $F = 0$. The mean in each line declines with inbreeding, and the mean of all inbred lines is $\mu_F = \mu_0 - B\,F$. If all these lines are crossed at random, $F = 0$, and the mean of the crossbreeds = μ_0, the mean of the outbred population.

Heterosis can also arise in crosses between outbred (i.e., randomly mating lines), as we detail below.

Heterosis: Change in the Mean Under Crossbreeding

Consider the cross between two particular parental strains (P_1 and P_2). Heterosis depends on the difference in gene frequency between the lines, and the amount of heterosis changes from the F_1 to the F_2. Suppose the crossing scheme is:

$$\begin{array}{ccc} P_1 & \text{x} & P_2 \\ & \downarrow & \\ & F_1 & \\ & \downarrow & \\ & F_2 & \end{array}$$

We will define heterosis as the deviation from the midparental value, so that for the F_1,

$$H_{F_1} = \mu_{F_1} - \frac{\mu_{P_1} + \mu_{P_2}}{2} \tag{6.3}$$

Let the allele frequencies for a diallelic locus in populations 1 and 2, be p and $p + \delta p$, respectively. We assume the genotypes in P_1 and P_2 are in Hardy-Weinberg proportions (which also hold if the lines are completely inbred), giving the means as

$$\begin{aligned} \mu_{P_1} &= (2p-1)a + 2p(1-p)d \\ \mu_{P_2} &= \mu_{P_1} + 2(\delta p)a - 2(\delta p)^2 d \end{aligned}$$

In the F_1, the probability of (say) an A_1A_2 locus is the probability of receiving an A_1 from P_1 and an A_2 from P_2 ($p\,[1-(p+\delta p)\,]$) or an A_2 from P_1 and an A_1 from P_2 ($[\,1-p\,]\,[\,p+\delta p\,]$). Considering the other two genotypes gives the mean of the F_1 (expressed in terms of the means for P_1) as

$$\mu_{F_1} = \mu_{P_1} + (\delta p)a$$

giving the contribution to heterosis from this particular locus as

$$H_{F_1} = \mu_{F_1} - \frac{\mu_{P_1} + \mu_{P_2}}{2} = (\delta p)^2 d \tag{6.4a}$$

Hence, for this locus to show heterosis ($H > 0$), we require both a difference in allele frequencies between the populations ($\delta p \neq 0$) and positive dominance ($d > 0$). Note immediately that over-dominance ($d > a$) is not required for heterosis. Summing over all loci, the heterosis produced by dominance for this cross becomes

$$H_{F_1} = \sum_{i=1}^{n} (\delta p_i)^2\, d_i \tag{6.4b}$$

Hence,

- Heterosis depends on dominance. $d = 0$ = no inbreeding depression and no heterosis.
  As with inbreeding depression, directional dominance is required for heterosis.

- H is proportional to the square of the difference in gene frequency between populations. H is greatest when alleles are fixed in one population and absent in the other (so that $|\delta p_i| = 1$). $H = 0$ if $\delta p = 0$.
- H is specific to each particular cross. H must be determined empirically, since we do not know the relevant loci nor their allele frequencies.

Heterosis in F_2

The F_2 generation is derived by mating the F_1 at random. The gene frequencies in the F_1 are the average of the two parents, so that freq$(A_1) = (p + p + \delta p)/2 = p + \delta p/2$. Since the F_2 is formed by random mating, the genotype frequencies are in HW equilibrium with allele frequency $p + \delta p/2$, and the F_2 mean becomes

$$\mu_{F_2} = a(\,[\,p + \delta p/2\,]^2 - [\,1 - p - \delta p/2\,]^2\,) + d2[\,p + \delta p/2\,][\,1 - p - \delta p/2\,]$$

A little algebra shows that

$$H_{F_2} = \mu_{F_2} - \frac{\mu_{P_1} + \mu_{P_2}}{2} = \frac{(\delta p)^2\, d}{2} = \frac{H_{F_1}}{2} \tag{6.5}$$

so that in the F_2, only half the advantage of the F_1 hybrid is preserved. Since (presumably) random mating also occurs in subsequent generations, the heterosis in future generations is the same as the F_2 heterosis, as the allele frequencies do not change and genotypes remain in Hardy-Weinberg frequencies.

Agricultural importance of heterosis

Heterosis is extremely important in world agriculture. Crosses often show **high-parent heterosis**, wherein the F_1 not only beats the average of the two parents (**mid-parent heterosis**), it exceeds the value of the best parent.

The importance of high-parent heterosis is illustrated by the following estimates of the world-wide contribution of heterosis to both yield and land savings. Here the percent hybrid advantage is the yield increase of the hybrid over the best single variety. (After Duvick 1999).

Crop	% planted as hybrids	% Hybrid yield advantage	Annual added yield Percent	Annual added yield tons	Annual Land savings
Maize	65	15	10	55×10^6	13×10^6 ha
Sorghum	48	40	19	13×10^6	9×10^6 ha
Sunflower	60	50	30	7×10^6	6×10^6 ha
Rice	12	30	4	15×10^6	6×10^6 ha

Crossbreeding in Animals: General Concepts

As the mule illustrates, the importance of between-species hybrids in animal breeding goes back to prehistoric times. The aggressive utilization of crossbreeding to exploit heterosis (as opposed to crosses simply to combine desirable features from two different lines) followed Sewall Wright's (1922) extremely influential publication on crossbreeding (and inbreeding) in guinea pigs.

Animal breeders often distinguish between **individual** and **maternal heterosis**. Individual heterosis is *enhanced performance in a hybrid individual*, while maternal heterosis is *enhanced maternal performance* (such as increased litter size and higher survival rates of offspring in a hybrid mother).

Maternal heterosis is often comparable, and can be greater than, individual heterosis (e.g., Table 6.1). For example, fertility in sheep is improved more by using crossbred ewes (resulting in a 9% maternal heteroic effect) than by using crossbred lambs from purebred ewes (only a 2-3% individual heteroic effect). Maternal effect heterosis is roughly twice as individual heterosis for several traits in European cattle (*Bos taurus*).

Table 6.1. Estimate of individual h^I and maternal heterosis h^M in sheep (estimation of h^I and h^M is discussed later in the lecture notes). Results presented as percentage of parental means. n_I and n_M indicate the number of estimates used for the reported individual and maternal values. **Prolificacy** is the litter size at birth.

Trait	n_I	Mean h^I (in %)	n_M	Mean h^M (in %)
Birth weight	42	3.2	12	5.1
Weaning weight	56	5.0	27	6.3
Preweaning growth rate	19	5.3		
Postweaning growth rate	10	6.6		
Yearling weight	18	5.2		
Ovulation rate			4	-2.0
Fertility	20	2.6	30	8.7
Prolificacy	20	2.8	31	3.2
Birth-weaning survival	29	9.8	25	2.7
Lambs per ewe	20	5.3	25	11.5
Lambs reared per ewe	20	15.2	25	14.7
Total weight lambs/ewe	24	17.8	25	18.0
Carcass traits	7	$\simeq 0$		

Maternal and individual heterosis effects can be combined by using crossbred dams. For example, for total weight of lambs reared per mated ewe has an 18% individual heterotic advantage in a crossbred offspring and an addition 18% advantage (from maternal heterosis) when crossbred ewes are used in place of purebred ewes. This combining of maternal and individual heterotic effects is one reason why three-way crosses are common in animal breeding, generally by crossing a male from line A with a hybrid (i.e., crossbred) female (from a $B \times C$ cross). This strategy exploits maternal heterosis in the female, with the sire line often chosen for its contribution to some production trait.

In theory, one could also consider **paternal heterosis**, increased performance due to paternal effects in hybrid sires, but there has been little evidence to date of this being a significant effect. For example, Bradford et al. (1963) compared purebred versus crossbred sire performance, finding no major differences between them except for a slight (4%) elevation of lamb survival, where the hybrid exceeded both parents.

Synthetics and Rotational Crossbreeding

To maximally exploit heterosis, we ideally would use only F_1 individuals, as the heterotic advantage decreases in the F_2, e.g., Equation 6.5. In much of plant breeding, **terminal crosses** are used, wherein the final hybrid individuals are the endpoints and do not reproduce further. For example, a company sells a farmer F_1 seed, and this seed is generated anew each generation, with the F_1 plants themselves not allowed to reproduce. While such schemes can work in plants with their enormous reproductive potential, they are more difficult in animals. Consider a three-way cross of an A sire to a $B \times C$ dam (with C being the grandam). While artificial fertilization only requires modest numbers of sires from the A and B lines/populations, in large farm animals, the number of offspring is on order of the number of dams. Thus, to produce n triple cross hybrid offspring requires on the order of $2n$ dams

(the grandam to produce the $B \times C$ dam, and this dam herself). One potentially very important application of whole-animal cloning would be the creation of cloned lines from exceptional dams showing outstanding maternal heterosis.

Two approaches have been suggested as potential compromises between the cost of continually providing F_1's and the decrease in performance of F_2. The first is the use of **synthetics**, where n parental lines are chosen and a random-mating population formed by first making all $n(n-1)/2$ pairwise intercrosses between the lines. Sewall Wright gave an approximate expression for the decrease in heterosis in a synthetic as

$$F_2 = F_1 - \frac{F_1 - \overline{P}}{n} \tag{6.6a}$$

where $\overline{P}$ is the average mean of all parents and the F_1 value is the average of all pairwise crosses between lines. Since $F_1 - \overline{P} = H$ is the heterosis, this can be expressed as

$$H_{F_2} = F_2 - \overline{P} = \left(F_1 - \frac{H}{n}\right) - \overline{P} = \left(F_1 - \overline{P}\right) - \frac{H}{n} = H\left(1 - \frac{1}{n}\right) \tag{6.6b}$$

The idea behind synthetics is that, the larger n, the smaller the lost of F_1 heterosis.

A second approach is **rotational crossbreeding**, which continually recycles hybrid individuals. Here, hybrids from the previous generations are crossed (in rotation) to pure lines. For example, a three-bred rotational would use $A \times B$ as the first generation. In generation two, dams from the first generation are crossed to line C. In generation three, dams from generation two are crossed to sires from line A, and the rotation continues over all three lines in subsequent generations. This approach represents a comprise between trying to maintain maximal heterozygosity within a line without having to regenerate the line anew each generation.

Carmon et al. (1956) examined the average (asymptotic) performance of rotational crossbreeding under the assumption of no epistasis. For a two-line (A, B) rotation, the predicted mean is

$$\widehat{R}_2 = \overline{z}_{AB} - \frac{\overline{z}_{AB} - \overline{P}_2}{3}, \quad \text{where} \quad \overline{P}_2 = \frac{\overline{z}_A + \overline{z}_B}{2} \tag{6.7a}$$

Here $\overline{P}$ is the average of the two parental lines and $\overline{z}_{AB}$ the mean value of their cross. Note that the mean performance under rotational crossbreeding is less than the single cross performance. For a three-line (ABC) rotational cross, the predicted mean is

$$\widehat{R}_3 = \overline{SC}_3 - \frac{\overline{z}_{AB} - \overline{P}_3}{7}, \quad \text{where} \quad \overline{SC}_3 = \frac{\overline{z}_{AB} + \overline{z}_{AC} + \overline{z}_{BC}}{3} \tag{6.7b}$$

where $\overline{SC}_3$ is the average of the three single crosses between these three lines and $\overline{P}_3$ the average of the three parental lines.

For a four-line rotation, the order of the rotation matters. Letting the rotation be A, B, C, D, the predicted long-term performances

$$\widehat{R}_4^{(A,B,C,D)} = \overline{SC}_4 - \frac{\overline{SC}_{na} - \overline{P}_4}{15}, \quad \text{where} \quad \overline{SC}_{na} = \frac{\overline{z}_{AC} + \overline{z}_{BD}}{2} \tag{6.7c}$$

As above, $\overline{P}_4$ is the mean of the original lines, $\overline{SC}_4$ is the mean of all six possible single-crosses between the four lines, and $\overline{SC}_{na}$ the average of the two single crosses of non-adjacents lines in the rotations.

As an application of the above expressions, consider the following data for various crosses of Devon and Brahman cattle (from Kidder et al. 1964): The midparent $\overline{P}$, F_1, two-breed rotational crossbred R_2, two-line synthetic S_2, and the backcross (BC) have the following meanss

	Means				
Trait	$\overline{P}$	F_1	R_2	S_2	BC

Weaning weight	154.2	180.5	178.3	170.1	181.4
12-month weight	210.5	246.8	232.2	212.3	233.6
18-month weight	274.9	315.7	296.6	276.6	295.3
12-18 month weight gain	64.4	68.9	64.4	64.4	61.7

Note from these data that the rank of performance is $F_1 > R_2 > S_2 > \overline{P}$. How well does Equation 6.7 predict the rotational crossbred performance? Here the F_1 corresponds to $\overline{z}_{AB}$, so that the predicted equilibrium value is

$$\widehat{R}_2 = F_1 - \frac{F_1 - \overline{P}_2}{3}$$

For example, for weaning weight

$$\widehat{R}_2 = 180.5 - \frac{180.5 - 154.2}{3} = 171.7$$

which is 96% of the observed value. Similarly, the predicted values (and fraction of the actual values) for 12, 18, and gain are, respectively, 234.7 (101%), 302.1 (102%), and 67.5 (104%). Hence, for these data, there is a slight tendency to overestimate the true mean.

Turning to Equation 6.6 to predict the synthetic performance, $n = 2$ and hence for weaning weight

$$\widehat{S}_2 = 180.5 - \frac{180.5 - 154.2}{2} = 167.4$$

which is 98% of the observed value. Similarly, the predicted values (and fraction of the actual values) for 12, 18, and gain are, respectively, 228.65 (107%), 295.3 (107%), and 66.7 (104%).

Estimating the Amount of Heterosis in Maternal Effects

The basic approach for incorporating maternal effects was given in a specific case (three-line crosses) by Magee and Hazel (1959), while a more general treatment was presented by Dickerson (1969). Maternal heterotic effects are judged to be of sufficient importance that significant effort is usually made to use crossbred dams, while sires are very often purebreds. Crossbred dams are very often chosen on the basis of favorable reproductive traits, while sires are often picked for other traits (such as size or carcass traits).

The model used to estimate maternal effects and heterosis is as follows: We can consider the mean value of a line as consisting of an average direct (or individual) effect g^I, a maternal genetic effect g^M expressed through the mother, and even a potential grand-maternal effect expressed through the dam $g^{M'}$. For example, the mean value for line A is

$$\mu_A = \mu + g_A^I + g_A^M + g_A^{M'} \tag{6.8a}$$

as the dam in this cross is from line A and the mother of this dam (the granddam) is also from line A. Crossbred offspring potentially experience an additional heterotic effect h. Hence, the expected mean in a cross with line A as a sire and B as a dam is

$$\mu_{AB} = \mu + \frac{g_A^I + g_B^I}{2} + g_B^M + g_B^{M'} + h_{AB}^I \tag{6.8b}$$

as the individual genetic value of an AB individual is the average of the two lines plus any additional heterotic direct effect h_{AB}^I. In this cross, both the dam and granddam are from line B. Conversely, the expected mean of the reciprocal cross (now with A as the dam) is

$$\mu_{BA} = \mu + \frac{g_A^I + g_B^I}{2} + g_A^M + g_A^{M'} + h_{AB}^I \tag{6.8c}$$

From Equation 6.8a-c, it follows that an estimate of the individual (direct) heterotic effect is the obvious one (Nitter 1978),

$$\frac{\mu_{AB}+\mu_{BA}}{2}-\frac{\mu_{AA}+\mu_{BB}}{2}=h^I_{AB} \tag{6.8d}$$

as the maternal effects cancel. Likewise, the difference in reciprocal crosses

$$\mu_{BA}-\mu_{AB}=\left(g^M_A+g^{M'}_A\right)-\left(g^M_B+g^{M'}_B\right) \tag{6.8e}$$

provides an estimate of the difference in maternal + grandmaternal genetic effects for the two lines.

If the dam is crossbred, then she has the potential of heterotic maternal effects h^M. Likewise, if the granddam is crossbred, there are also potential grandmaternal heterotic effects $h^{M'}$. The final complication is that if h_{AB} is the heterotic (individual or maternal) contribution in the F_1, heterosis can decrease in the F_2 due to recombination breaking up favorable gene combinations. This is incorporated into the general model via a **recombinational loss** term, r_{AB}. As an example of putting all these pieces together, consider the three-way cross using a crossbred dam,

$$\mu_{C\cdot AB}=\frac{2g^I_C+g^I_A+g^I_B}{4}+\frac{h^I_{CA}+h^I_{CB}}{2}+\frac{g^M_A+g^M_B}{2}+h^M_{AB}+g^{M'}_B+\frac{r^I_{ab}}{2} \tag{6.9a}$$

Here since the dam is a crossbred, there is a potential heterotic component (h^M_{AB}), while the granddam of the AB dam is B. The final piece is that the individual heterotic component in the F_1 (the AB) cross may be potential degraded in the F_2, which is accounted for by the r^I_{ab} term. Table 6.2 summarizes these coefficients for a variety of crosses. For example, estimates of the various effects can be obtained by suitably weighted combination of line means. Combining 6.8b and 6.9a gives (Nitter 1978),

$$\mu_{C\cdot AB}-\frac{\mu_{CA}+\mu_{CB}}{2}=h^M_{AB}+\frac{r^I_{ab}}{2} \tag{6.9b}$$

Hence, if recombination effects are small (these would be absence in the absence of epistasis), Equation 6.9b allows for a direct estimate of maternal heterotic effects.

Table 6.2 summaries the weights on the various heteroic components we are trying to estimate for a number of crosses.

Table 6.2. Coefficients for direct (or individual) genetic effect g^I, genetic maternal effects g^M and genetic grandmaternal effects $g^{M'}$, their heterotic counterparts (h) and recombination corrections r for a variety of crosses. After Dickerson (1969).

Cross	Order	g^I	h^I	r^I	g^M	h^M	r^M	$g^{M'}$	$h^{M'}$
P_1	A	A			A			A	
F_1	A·B	$\frac{A+B}{2}$	AB		B			B	
	B·A	$\frac{A+B}{2}$	AB		A			A	
3W	(C·D)·A	$\frac{C+D+2A}{4}$	$\frac{CA+DA}{2}$	$\frac{cd}{2}$	A			A	
	C·(A·B)	$\frac{2C+A+B}{4}$	$\frac{CA+CB}{2}$	$\frac{ab}{2}$	$\frac{A+B}{2}$	AB		B	
	C·(B·A)	$\frac{2C+A+B}{4}$	$\frac{CA+CB}{2}$	$\frac{ab}{2}$	$\frac{A+B}{2}$	BA		A	
B_1	A·(A·B)	$\frac{3A+B}{4}$	$\frac{AB}{2}$	$\frac{ab}{2}$	$\frac{A+B}{2}$	AB		B	
B_1	A·(B·A)	$\frac{3A+B}{4}$	$\frac{AB}{2}$	$\frac{ab}{2}$	$\frac{A+B}{2}$	AB		A	
F_2	$(A\cdot B)^2$	$\frac{A+B}{2}$	$\frac{AB}{2}$	ab	$\frac{A+B}{2}$	AB		B	
F_3	$(A\cdot B)^3$	$\frac{A+B}{2}$	$\frac{AB}{2}$	$ab+$	$\frac{A+B}{2}$	$\frac{AB}{2}$	ab	$\frac{A+B}{2}$	AB

Inbreeding and Crossbreeding Problems

1. Suppose that four genes, each with two alleles affecting body weight have been identified and are segregating in a random-bred strain of mice. The mean weights for the three genotypes (in grams) for each of the four genes are as follows:

Gene A		Gene B		Gene C		Gene D	
A_1A_1	22	B_1B_1	22	C_1C_1	22	D_1D_1	22
A_1A_2	21	B_1B_2	22	C_1C_2	20	D_1D_2	19
A_2A_2	20	B_2B_2	20	C_2C_2	20	D_2D_2	19
freq(A_2)	0.5	freq(B_2)	0.5	freq(C_2)	0.1	freq(D_2)	0.2

Calculate for each gene the inbreeding depression caused by it in the progeny of one full-sib mating, when the allele frequencies in the random-bred population are as shown. Recall that $F = 0.25$ after one generation of full sib mating. What is the total inbreeding depression? What is the value of B?

2. Suppose highly inbred lines of mice are crossed in pairs, and the adult body weights of the two lines crossed and of the F_1 are as shown below. Calculate: (1) the amount of heterosis shown by each cross and (2) the expected body weights in the F_2 generation if random breeding within each crosses was continued. Assume that there are no maternal effects on adult body weight.

	Line A	Line B	F_1
(i)	18	16	20
(ii)	22	16	21
(i)	18	18	24

3. Suppose you have four lines (A through D) and make all pairwise crosses between them. The resulting parental and F_1 means are as follows:

Line	A	B	C	D
A	10	20	14	16
B		14	15	17
C			12	22
D				15

The above table is read as follows: Diagonal elements are the line means, i.e. , 10 for A, 12 for C, while the off-diagonal elements are the means of the crosses, e.g., a $A \times C$ cross has mean 14, etc.

a: What is the expected mean of a four-line synthetic?
b: What is the expected mean of a 4-way rotation with order A:B:C:D?
c: What is the expected mean of a 4-way rotation with order A:C:B:D?
d: What is the expected mean of a 3-way rotation with A, B, and C?
e: What is the expected mean of a rotation with C and D?

Solutions to Chapter Problems 6

1. The contribution to inbreeding depression from each loci is $2Fp_iq_id_i$

	Gene A	Gene B	Gene C	Gene D
d_i	0	1	-1	-1.5
$2p_iq_id_i$	0	0.5	-0.18	-0.48

$$B = \sum 2p_iq_id_i = -0.16, \qquad F \sum 2p_iq_id_i = -0.04$$

Hence, the inbreeding depression is 0.04.

2. Recall that Heterosis $H = F_1 - (\mu_A + \mu_B)/2$, while only half of this is retained in the F_2, with the mean of the F_2 being $H/2 + (\mu_A + \mu_B)/2$

	Line A	Line B	F_1	H	F_2 mean
(i)	18	16	20	3	18.5
(ii)	22	16	21	2	20
(i)	18	18	24	6	21

3. a: Here $n = 4$,

$$\overline{P} = (10 + 14 + 12 + 15)/4 = 12.75$$

$$F_1 = (20 + 14 + 16 + 15 + 17 + 22)/6 = 17.33$$

giving

$$F_2 = F_1 - \frac{F_1 - \overline{P}}{n} = 17.3 - \frac{17.3 - 12.75}{4} = 16.19$$

b: Recall Equation 6.7c ,

$$\widehat{R}_4^{(A,B,C,D)} = \overline{SC}_4 - \frac{\overline{SC}_{na} - \overline{P_4}}{15}, \quad \text{where} \quad \overline{SC}_{na} = \frac{\overline{z}_{AC} + \overline{z}_{BD}}{2}$$

Here, $\overline{P_4} = 12.75$, $\overline{SC}_4 = 17.33$ and

$$\overline{SC}_{na} = \frac{14 + 17}{2} = 15.5$$

giving $\widehat{R}_4^{(A,B,C,D)} = 17.15$.

c: Here

$$\overline{SC}_{na} = \frac{\overline{z}_{AB} + \overline{z}_{CD}}{2} = 21$$

giving $\widehat{R}_4^{(A,C,B,D)} = 16.78$

d: Recalling Equation 6.7b,

$$\widehat{R}_3 = \overline{SC}_3 - \frac{\overline{z}_{AB} - \overline{P}_3}{7}, \quad \text{where} \quad \overline{SC}_3 = \frac{\overline{z}_{AB} + \overline{z}_{AC} + \overline{z}_{BC}}{3}$$

Here $\overline{SC}_3 = 19.33$, $\overline{P}_3 = 13$, giving $\widehat{R}_3 = 18.33$.

e: Recall Equation 6.7a,

$$\widehat{R}_2 = \overline{z}_{AB} - \frac{\overline{z}_{AB} - \overline{P}_2}{3} = 22 - \frac{22 - 13.5}{3} = 19.17$$

Note that this is higher than the predicted means under the above four proposed crosses.