31

Selection and Crossbreeding I: Basic Concepts

Draft 26 June 2014

Hybridization, the crossing of two (or more) lines or populations is a common feature of applied breeding and is also of considerable interest to evolutionary biologists. There are two distinct reasons to consider crosses between populations. First, breeders cross populations in an attempt to combine the best features of each line/population/species. This can often fail, in quite spectacular fashion. Consider the rabbage, the F_1 between a radish and a cabbage. The idea was to combine a radish root with a cabbage head. Unfortunately, the result was a radish head on a cabbage root (Figure 21. 1). In this case, it would appear that the favorable genes for each species were either recessive and/or the result of strong epistatic interactions.

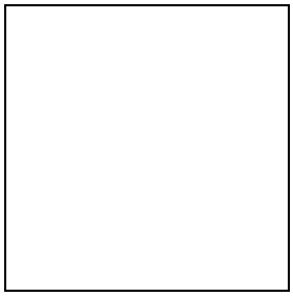


Figure 21. 1. The rabbage.

The second reason to consider cross breeding is heterosis (or hybrid vigor), wherein (unlikely the poor rabbage), the F_1 is superior in some trait to either parent. Heterosis can occur between a cross of two lines that otherwise seem nearly identical. In such cases, the purpose of creating hybrids is not to combine advantageous traits, but rather to uncover favorable genes that are masked by either dominance and/or epistasis. It is this latter aim that is our main focus in this chapter. Heterosis present in an F_1 is usually reduced (often considerably) in an F_2, mirroring the decrease in heterozygosity from the F_1 to the F_2. Given these facts, there has been much effort (especially by plant breeders) to identify and create inbred lines that cross well (such lines are said to show **nicking**) to produce favorable F_1 hybrids. Lines are often classified into **heterotic groups**, wherein individuals from different groups when crossed show significant heterosis, while crosses between individuals within a group show little or no heterosis.

Our treatment of selection and crossbreeding starts by considering some of the types

of crosses that can be used, followed by a short discussion of the role of hybrids in modern agriculture. We then turn to the nuts-and-bolts of crosses: the genetic variances between lines, estimating the effects of lines for combining prediction of lines means, and rotational crossbreeding. Our discussion continues in the next chapter, where we discuss selection for improved crossing lines, creation of synthetic and composite populations, and conclude by discussing methods to detect elite alleles in lines.

TYPES OF CROSSES

The number of different possible crosses one can make is limited only by the energy and imagination of the worker and the reproductive biology of the organism under consideration. The most common is a **single cross** (often denoted **SC**), the F_1 between to lines. With a collection of lines in hand, how does the breeder decide which to cross to produce the best-performing SC? One approach is simply to make all $n(n-1)/2$ single crosses among all n lines, a **diallel design** (LW Chapter 20). The problem is that the full diallel design involves a considerable amount of crosses– 45 possible SCs involving a collection of 10 lines, 190 SCs for 20 lines, 1225 SCs for 100 lines. If one could chose a subset of all lines, then a full diallel is certainly reasonable, but what criteria should be used?

A strategy for selecting parental lines is suggested from the analysis of a full diallel, where one can estimate the **general combining ability** (or GCA) of each line and the **specific combining ability** (SCA) for each cross (Sprague and Tatum, 1942). The GCA of a line is the difference of its average performance in hybrids relative to the grand mean of all line crosses (LW Chapter 20). The expected mean of a cross is the sum of the maternal and paternal general combining abilities, GCA(P) + GCA(M). Since the GCA from a random individual (being a deviation from the grand mean) is zero, if only one parent is known, the expected offspring mean is just the GCA from that parent. The deviation between the predicted and actual values estimates the specific combining ability for that particular cross. The GCA is in effect the breeding value of a line, and (as with breeding values) it is a function of population of lines chosen. Hence, the mean value of the single cross $\ell_i \times \ell_j$ = GCA_i + GCA_j + SCA_{ij}. Thus, if one knew something about the GCA's of the lines in our initial collection, a quite reasonable starting point is to perform a diallel using those with the highest GCAs. While it certainly true that the best line may involve a cross between two lines with very low GCAs, but a very large SCAs, using the GCAs nonetheless offers a good compromise starting point.

The question is how to obtain the GCAs for a series of lines without making all the controlled crosses for a full diallel. Two of the most common approaches are the **topcross design**, in which a particular line (the **testor**) serves as the common male (pollen) parent and the GCAs of female plants are estimated by the performance of their seed. The other approach is a **polycross design** wherein females are allowed to be randomly fertilized from among all the other lines. Again, the elite lines are chosen by those with the best-performing offspring. Under either design, care must be taken to avoid selfing. While manual pollen control can be used if the flowers are sufficiently large and the number of flowers per plant and the number of crosses is small, a much more efficient (and elegant) approach is to use female plants that are genetically male-sterile (e.g., Rogers and Edwardson 1952, Stephens and Holland. 1954).

As we will shortly see, plant and animal breeders often work with more complex hybrids than single crosses. **Triple** or **three-way crosses** (**3W**) crosses involve an F_1 crossed to a third line, e.g., $A \times (B \times C)$, which is often written as $A \cdot BC$. The resulting individuals have 50% of their genes from line A, and 25% from each of lines B and C. **Four-way** (**4W**) or **double crosses** (**DC**) are the crosses of two different F_1's, $AB \cdot CD = (A \times B) \times (C \times D)$ so that the hybrid is the result of equal contributions from four different lines. [While the term double cross is very common in the literature, we will interchangeably use it with four-way cross, as it is easy for the causal reader to assumed that a double-cross involves fewer lines than a

triple cross, while this distinction is clear using three- vs. four-way crosses.] **Modified cross** involve two closely related populations (say A and $A*$). For example, a **modified single cross** is $(A \times A*) \times B$, while a **modified triple cross** is $(A \times A*) \times (B \times C)$. **Multiple cross** are also used (e.g., Sprague and Jenkins 1943), such as the cross using two (different) four-way cross individuals as parents, which involves eight lines.

HETEROSIS

Schull (1952) coined the term **heterosis** to refer in an increase in the vigor or other agriculturally-related traits (which are usually components of fitness such as yield). While the basic concept of heterosis is straight-forward, it can be defined a number of different ways, depending on what reference population is used (Lamkey and Edwards 1999). The most standard usage is **mid-parent heterosis**, defined as the F_1 mean exceeding the average mean of the two parental lines,

$$h = \mu_{F_1} - \frac{\mu_{P_1} + \mu_{P_2}}{2} = \mu_{F_1} - \mu_{\overline{P}} \qquad (21.\ 1)$$

Often the mean of the F_1 exceeds the mean of the best parent, and such **high-parent heterosis** is highly sought after by breeders.

Dominance and Heterosis

While epistasis can generate heterosis, most of the focus in the literature has been on dominance-generated heterosis. No doubt a major reason for this is that dominance is much easier to model, and expressions for predicted line cross means are rather simple when epistasis is negligible. To see under what conditions dominance generates heterosis, consider the cross between two populations, P_1 and P_2. For on a diallelic locus underlying the trait of interest, let the genotypes $\mathbf{AA} : \mathbf{Aa} : \mathbf{aa}$ have genotypic values of $a : d : -a$, and let p and $p + \delta_p$ be the frequency of allele $\mathbf{A}$ in populations 1 and 2 (respectively). A convenient notation will be to denote the means of the P_1 and P_2 parental populations by μ_{11} and μ_{22}, and the F_1 by μ_{12}. Assuming the parental populations are in Hardy-Weinberg, and following Willham (1970, and Willham and Pollak 1985) for their F_1 cross, these means become

$$\begin{aligned} \mu_{11} &= (2p-1)a + 2p(1-p)d \\ \mu_{22} &= \mu_{11} + 2\delta_p\, a - 2\,\delta_p^2 d \\ \mu_{12} &= \mu_{11} + \delta_p\, a \end{aligned} \qquad (21.2a)$$

The resulting (mid-parental) heterosis (for this particular locus) becomes

$$h = \mu_{12} - \frac{\mu_{11} + \mu_{22}}{2} = \delta_p^2 d \qquad (21.\ 2b)$$

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 overdominance ($d > a$) is not required for heterosis. Summing over all loci, the heterosis produced by dominance for this cross becomes

$$h = \sum_{i=1}^{n} \delta_{p_i}^2\, d_i \qquad (21.\ 2c)$$

Dominance thus contributes to heterosis provided it is directional (the d_i's tend to be positive) and their are differences in the allele frequencies between population. Further, heterosis is

expected to increase with the allele frequency differences in the populations being crossed. In the most extreme case, crossing inbred lines, $h = \sum d_i$.

Under what conditions does the mean of the F_1 exceed that of both parents? Summing over all loci, we can rewrite Equation 21. 2a as

$$\mu_{12} - \mu_{11} = \sum_{i=1}^{n} \delta_{p_i} a_i, \quad \text{and} \quad \mu_{12} - \mu_{22} = 2\sum_{i=1}^{n} \delta_{p_i}^2 d_i - \sum_{i=1}^{n} \delta_{p_i} a_i$$

Thus, the hybrid exceeds the best parent (high parent heterosis) when

$$2\sum_{i=1}^{n} \delta_{p_i}^2 d_i \; > \sum_{i=1}^{n} \delta_{p_i} a_i \; > 0 \tag{21. 3a}$$

Note that while overdominance ($d_i > a_i$) facilitates Equation 21. 3, is it not required. For completely inbred lines, Equation 21. 3a reduces to

$$2\sum_{i=1}^{n} d_i \; > \sum_{i=1}^{n} \; a_i \; > 0 \tag{21. 3b}$$

An important point is that the F_1 is *not* in Hardy-Weinberg equilibrium, as the mating is non-random (an individual from population one is always crossed to one from population two). The frequency of heterogyotes in the F_1 is

$$\begin{aligned} \Pr(\mathbf{Aa}) &= \Pr(\mathbf{A} \text{ for Pop 1}) \Pr(\mathbf{a} \text{ for Pop 2}) + \Pr(\mathbf{a} \text{ for Pop 1}) \Pr(\mathbf{A} \text{ for Pop 2}) \\ &= p\,(\,1 - [\,p + \delta_p\,]\,) + (1-p)(\,p + \delta_p\,) \\ &= 2\,\overline{p}\,(\,1 - \overline{p}\,) + \frac{\delta_p^2}{2} \end{aligned} \tag{21.4}$$

where $\overline{p} = p + \delta p/2$ is the frequency of **A** in the hybrid population. Hence, the excess of heterozygotes in an F_1 relative to a random-mating population with the same allele frequency is $(\delta p)^2/2$. If the allele frequencies are close, this excess is small, while its maximal value of $1/2$ occurs with crosses between completely inbred lines (100 percent heterozygotes vs. 50 percent heterozygotes under random mating). After one generation of random mating (the F_2), the amount of heterosis is expected to decrease, reflecting a decrease in the percentage of heterozygotes. In particular, the mean value in the F_2 is $\mu_{F_2} = (2\,\overline{p} - 1)\,a + 2\,\overline{p}\,(1 - \overline{p}\,)d$, and applying Equation 21. 2a plus some simplification yields

$$\mu_{F_2} - \frac{\mu_{11} + \mu_{22}}{2} = \frac{\delta p^2 d}{2} = \frac{h}{2} \tag{21. 5a}$$

Hence

$$\mu_{F_2} - \mu_{F_1} = \frac{h}{2} \tag{21. 5b}$$

so that half the initial heterosis (from dominance) is lost in the F_2. However, since the F_2 is in Hardy-Weinberg equilibrium, future generations are expected to have the same mean and no further reduction in heterosis occurs.

Several authors have extended this basic theory. Cress (1966) examines crosses where more than two alleles are segregating, concluding that negative heterosis is not unlikely, even when overdominance is present. Only in the case of complete dominance will the heterosis be positive (or zero) for all potential vectors of allele frequencies differences between the crossed populations. The effects of epistasis on heterosis have been

examined by Willham and Pollak (1985) and Goodnight (1999). Finally, traits acting in a multiplicative fashion can also generate heterosis, even in the absence of dominance (Richley 1942, Williams 1959, Garifus 1959), and this has been examined by Schnell and Cockerham (1992).

Inbreeding vs. Panmictic Heterosis

Equation 21.2 (showing the effect of dominance on heterosis) assumes that both parental populations are in Hardy-Weinberg. When this is not the case (for example, under inbreeding), the effects of heterosis can be confounded with the effects of inbreeding depression (Lamkey and Edwards 1999). Suppose each parental population is a collection of completely homozygous lines formed by continually selfing members of some initial (random mating) base population. [Note that this situation is different from a cross of two pure lines, where the genotypes are identical within each line and each line is in Hardy-Weinberg, and the above equations hold.] If p is the frequency of allele **A** in the initial base population of P_1, then after complete inbreeding, p of the individuals will be **AA**, while the rest $(1-p)$ are **aa**. Hence the mean of P_1 is just

$$\mu_{11}^{(I)} = ap - a(1-p) = a(2p-1)$$

Likewise,

$$\mu_{22}^{(I)} = a(2[\,p+\delta_p\,]-1)$$

where we have used I to denote the mean under complete inbreeding. Since the F_1 resulting from the cross between these two populations has mean

$$\mu_{F_1} = ap(p+\delta_p) + dp(1-[\,p+\delta_p\,]) + d(1-p)(p+\delta_p) - a(1-p)(1-[\,p+\delta_p\,])$$

the heterosis, based on inbred populations as the parental reference, is

$$h_I = F_1 - \frac{\mu_{11}^{(I)} + \mu_{22}^{(I)}}{2} = d\delta_p + d2p(1-[\,p+\delta_p\,])$$

This can be alternately expressed as,

$$h_I = \frac{d\delta_p^2}{2} + 2d\,\overline{p}(1-\overline{p}\,) \tag{21.6}$$

Lamkey and Edwards (1999) refer to h_I as the **inbred midparent heterosis**, and Equation 21.2b as the **panmictic midparent heterosis**. There are two components to h_I, a contribution from the panmictic midparent heterosis and a contribution from the removal of inbreeding depression. To see this second contribution, note that the difference between the random mating and completely inbred contributions of a locus from population one is

$$\Delta\mu_{11}^{(I)} = \mu_{11} - \mu_{11}^{(I)} = ap^2 + 2dp(1-p) - a(1-p)^2 - a(2p-1) = 2dp(1-p)$$

Lamkey and Edwards use the term **baseline heterosis** to denote the contribution in the F_1 simply from the recovery of the effects of inbreeding depression reducing the parental means. Baseline heterosis, h_B is simply the difference between the panmictic and inbred values,

$$h_B = h - h_I = 2d\,\overline{p}(1-\overline{p}\,) - \frac{d\delta_p^2}{4} \tag{21.7}$$

Since this contribution arises entirely from differences between the random mated and fully inbred values of the parental populations this, we can equivalently express this as

$$h_B = \frac{\Delta\mu_{11}^{(I)} + \Delta\mu_{22}^{(I)}}{2} = \frac{2dp(1-p) + 2d[\,p+\delta_p\,](1-[\,p+\delta_p\,])}{2}$$

which reduces to Equation 21.7.

Dominance vs. Overdominance and Epistasis

While dominance provides a convenient explanation of heterosis, several objections were initially raised to the strict dominance (i.e., no overdominance, $d < a$). First, if no overdominance were involved, one would expect to find F_2 individuals that were as good as the F_1, and this was generally not seen. Example 21.1 shows, one is unlikely to have sufficient sample size to find such individuals if the number of loci contributing to heterosis is moderate to large. Another objection was that the F_2 distribution should be skewed, with the mode greater than the mean, while F_2 populations generally show very little skew. Collins (1921) points out that the later is also a sample effect, see Figure 21.2. Jones (1917) also showed that linkage provides another explanation while these two features (F_2 individuals as good as F_1 individuals, lack of skew in the F_2).

Example 21.1. Just how many F_2 progeny do we have to screen to have a reasonable (say 50%) chance of recovering the largest genotypic value (i.e., one that matches the F_1). We assume complete dominance, so that the presence of a single dominant allele is sufficient to maximize the contribution from a given locus (this is the most favorable]setting). With a pure line cross, then with probability 3/4 at least one favorable allele is present in the F_2. Assuming unlinked loci, the probability of favorable alleles at n loci is $(3/4)^n$. To have a 50% probability that at least one individual has favorable alleles at all n loci requires sampling m individuals, where m satisfies

$$1 - [\,1 - (3/4)^n\,]^m = 0.5$$

This follows as $(1 - (3/4)^n)$ is the probability that a single individual does not contain favorable alleles at all loci and this expression to the mth power is the probability that none of the m individuals do. Solving gives

$$m_{0.5} = \frac{\ln(0.5)}{\ln[\,1 - (3/4)^n\,]}$$

More generally, to have probability 100α requires a sample size of

$$m_\alpha = \frac{\ln(1-\alpha)}{\ln[\,1 - (3/4)^n\,]}$$

For example, for n = 15, 20, 30, and 40 loci, $m_{0.5}$ = 52, 218, 921, 3881, and 68,924. Hence, while there is a reasonable chance of recovering the best genotype with the number of loci is modest (20 or less), this very quickly becomes impossible with more than 30 loci. Further confounding these numbers is that not only must a plant be present in the sample, but we also must be able to assess its genotype with very high confidence, requiring replication across a number of environments (see Chapter 10).

One final remark. If there is (on average) partial dominance (i.e., $a > d$) , then the heterozygote is not as exceptional as the best possible genotype, and hence individuals homozygous for favorable alleles at a large number of loci may actually exceed the F_1.

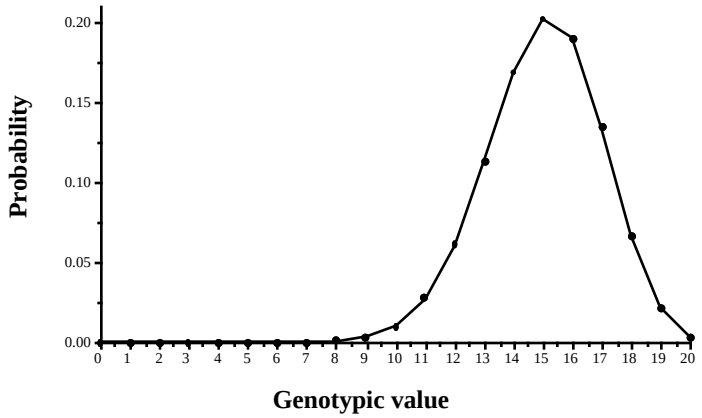


Figure 21.2. The F_2 distribution for 20 completely dominant loci of equal effect ($d = a = 1$). Starting with a cross of pure lines, the probability that a locus contains at least one dominant allele is 3/4, giving a mean value of 20(3/4) = 15, which is also the mode. Note this mean is below the F_1 mean of 20. Although the distribution is indeed skewed, the investigator is unlikely to oberve those individuals with a genotypic value of 8 or less (unless the sample size is enormous). Focusing on the most likely classes of observed individuals (9 to 20), this restrcited distribution shows little skew. After Collins (1921).

AGRICULTURAL EXPLOITATION OF HETEROSIS

The exploitation of hybrid vigor in agriculture traces back at least 5,000 years to the Sumerians (Clutton-Brock 1992), who produced mules by crossing horses (*Equus caballus*) with donkeys (*S. asinus*). Modern agriculture is highly shaped by hybrids, with the heterosis resulting is significantly increased yield in crops. Such gains direct translate into fewer acres that must be farmed to obtain the same total yield, and these land savings are by no means trivial (Table 21.1). Another critical, yet often overlooked, benefit of hybrids to modern agriculture is uniformity, which allows for greater efficiencies in harvesting, such an greatly increased mechanization (Goldman 1999).

Table 21.1. Estimates of the world-wide contribution of heterosis to both yield and land savings. 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	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

Hybrid Corn

Corn (maize) is the third most important crop in the world (behind wheat and rice), and the most important crop in the United States as measured by acres planted and total yield. Almost all of the US crop is **hybrid corn**, the seed from crosses of elite inbred lines (there are small amounts of **open pollinated** (**OP**) corn, but these are largely restricted to use by small independent farmers). The notion of hybrid corn traces back to Shull (1908, 1909), who noted that, because of the significant high parent heterosis seen in many corn crosses, the objective of corn breeders should be to find and maintain the best parental lines for hybrids. By maintaining these elite inbred lines and crossing them as needed, genetically identical hybrid seed can be produced year after year.

The problem with the widespread usage of hybrid seed was that the early inbred lines used by breeders typically had very low seed set. In a (single) cross between two inbred lines, the seed parent is inbred, and although the resulting hybrid seed produce plants with superior yield, each seed parent produces only a small number of seeds. Thus, poor fertility of the initial inbred lines resulted in rather few hybrid seed per plant. Jones (1918, 1922) suggested that instead of using an inbred line as the seed parent, one instead uses a hybrid line, with the hybrid seed being produced by a double cross instead of a single cross. Since the seed parent is a hybrid (a single cross between two inbred lines), it should show heterosis in seed production, resulting superior seed set. This suggestion directly opened up the vast commercial potential of hybrid corn. For example, in 1925 Henry A. Wallace (a farmer who later became Vice-President of the United States) formed the Pioneer Hi-Bred Corn Company of Iowa. From the 1930's to the 1960's, most US hybrid corn was produced by double crosses (Figure 21.3). However, it was soon noted (see below) that the seed from single crosses usually out perform double and three-way crosses. As breeders were able to generate inbred lines with higher yield, hybrid corn based on single crosses became both commercially feasible and also desirable, given the increased yields relative to double crosses. Since the 1970's, most hybrid corn is the US is the result of single crosses. Figure 21.3 shows the yield increase over a roughly 120 year period. The history of hybrid corn is reviewed by Kiesselback (1951) and in particular by Troyer (1999) who presents a detailed history of most of the founding lines that now comprise the whole of US corn production.

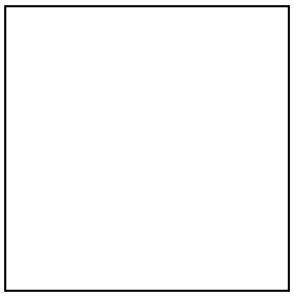


Figure 21.3. Average U. S. Corn yields, 1865-1998. Initially, dominated by open pollinated lines, which were largely replaced by double cross (four-way crosses of inbred lines), and most recent by single crosses. Regressions of yield gain per year b are computed for each of these three periods. Data from USDA and figure after Troyer (1999).

Hybrid corn also offers one of the best cautionary tails for plant breeders – the great **Southern Corn leaf Blight** (SCLB) epidemic from 1970-1971 (Tatum 1971, Ullstrop 1972). This was one of the most damaging plant epidemics in history — in terms of food energy destroyed it was much larger than the potato blight epidemic in the 1840's that produced

widespread famine (Ullstrup 1972). The roots of the epidemic in the United States trace back to what (at the time) appeared to be an elegant genetic solution to a major labor problem in the production of hybrid corn. Corn can self-fertilize, an undiserable feature as it reduce the fraction of hyrbid seeds on a plant. To prevent this, corn breeders manually detassled plants, removing the pollen-shed organ and allowing pollen only to be donated from the parents of interest. Typically one row of pollen plants is planted for every four to six (detassled) rows of seed parents, so that a pollen parent is within two or three rows of any seed parent (any seed set on the pollen plants is ignored). The discovery of a cytoplasmic factor (mitchondrially-encoded **REFERNECE**) that produced sterility in males seemed to be a elegant solution around the hassle of detassling (Jones and Everett 1949, Rogers and Edwardson, 1952). Plants with this T (for Texas, the origin of the line) cytoplasm are denoted by **Tcms**. Seed parents with Tcms cytoplasm produce normal and viable seeds, but no pollen, and hence to not have to the detassled. As a result of the very convenient feature, by 1970, almost 85% of hybrid corn seed in the US contained Tcms. In 1969, a previous unknown strain of the fungus *Helminthosporium maydis* was detected in a few areas in the midwest corn growing region of the central US. Plants contained Tcms were hyper-susceptible to this strain, and the result was over a billion dollar loss (in 1970's dollars) and major angst in both the private and public agricultural sectors. Corn is typically a rather disease-free crop (indeed, Ullstrup 1972 notes that in 1970, the US Department of Agriculture considered corn so healthy that it did not employ any full-time corn pathologists in the US corn belt), so this outbreak was even more of a shock and lead to a considerable focus on increasing the genetic diversity in crops.

Finally, it should be remarked that heterosis in maize is not just restricted to crosses between inbred lines. Richley (1922) reported the results from 244 crosses of open pollinated varieties, finding that 82.4% had yields above the midparental average, and 55.7% of the crosses exceeded the best parent.

Hybridization in Other Crops

A major reason that hybrid corn has been commercially successful is because of its reproductive structures, with pollen being produced on a long tassel, allow for relatively easy manual control of pollination. For many species with perfect flowers (i.e., both male and female parts), manual control of pollination is just not economically feasible because flowers are too numerous, too small, or both. Male-sterile genes are critical to the creation of hybrids in such species, as is chemical sterilization of pollen. In contrast with field crops, horticultural crops have much more favorable economic conditions for the commercial exploitation of heterosis. Here, flowers are typically large and easily handled (unlike the grass-family flower heads of many important field crops). Equally important, the yield from each individual plant has significant commercial value (Dodds 1955), so that individual manipulation of single plants often is economically feasible. Some examples of the successful use of hybridization in horticultural crops in include tomatoes, egg plant, cucurbits, and onions .

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 Wright's (1922) extremely influential publication on crossbreeding (and inbreeding) in guinea pigs. General reviews of heterosis in animals are given by Gowen (1952), Sang (1956), and Sheridan (1981).

Animal breeders often distinguish between **individual** and **maternal heterosis** (e.g., Nitter 1978). 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). Maternal heterosis is often comparable, and can be greater than, individual heterosis (e.g., Table 21.2). For example, Nitter (1978) reports that 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). Cundiff et al. (1974a, 1975b) found that maternal effect heterosis was roughly twice as individual heterosis for several traits in European cattle (*Bos taurus*) .

Maternal and individual heterosis effects can be combined by crossbred dams. For example, for total weight of lambs rear 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 (Nitter 1978). 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 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, 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.

Table 21.2. Estimate of individual h^I and maternal heterosis h^M in sheep (estimation of h^I and h^M is discussed later in this chapter). 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. From Nitter (1978).

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	$\simeq 0$		

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 (i.e., sheep and cows), 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 exception dams showing outstanding maternal heterosis.

Winters (1952) suggested a hybridization scheme, **rotational crossbreeding**, that 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. We examine this scheme in detail at the end of the chapter.

Crossbreeding in Animals: Heterosis in *Bos indicus* $\times$ *Bos taurus* Hybrids

Some interesting results on crossbreeding are found in the literature examining crosses between European (*Bos taurus*) and tropical (*Bos indicus*) cattle (reviewed in McDowell 1985). *Bos taurus* breeds have been selected for significant genetic improvement in a number of traits (such as milk and meat yield), but are adapted for only temperate climates. *Bos indicus* breeds are found in tropical countries, and while they show fairly little genetic improvement, they are well adapted to the local environment. Crosses between *taurus* and *indicus* have been performed in the hope of generating higher performing hybrids that are also adapted to tropical environments.

Bos. indicus $\times$ *B. taurus* hybrids are usually superior to indigenous breeds (i.e., the local breed of *B. indicus*) in milk yield and fitness measures such as calving age and interval. Trail et al. (1985) examined crosses between exotic *B. taurus* breeds (Angus and Red Poll) and indigenous *B. indicus* breeds (Ankole, Boran and Zebu) in Africa. Crossbred (exotic X indigenous) dams showed superior maternal performance over straightbred indigenous dams. However, while maternal effects were apparently superior, the progeny of straightbred (Boran) dams were actually heavier at 24 months than the progeny of exotic $\times$ crossbred dams, suggesting that the 3/4 exotic composition is not as favorable for individual performance as 1/2 exotic, 1/2 indigenous. Interesting, three-breed crosses between an improved breed sire and a cross-bred dam (a second improved breed $\times$ an indigenous breed) generally tend to do poorer than two-breed crosses. This is contrary to the general superiority of three breed crosses of *B. taurus*) in temperate areas (McDowell 1985).

Bos indicus $\times$ *B. taurus* hybrids show higher levels of heterosis than observed in crosses among breeds within each species. This is perhaps not unexpected, as allele frequencies likely have diverged more between species than between the breeds within a species. Everything else being equal (which, of course, is never assured), crosses involving widely allele frequencies differences would show increased heterosis. Interesting, heterosis in crosses of *B. indicus* breeds is higher than is observed in crosses of *B. taurus* breeds. For example, Gregory et al. (1985) found that material heterosis in crosses within *B. indicus* lines was intermediate between levels produced in *B. indicus* $\times$ *B. taurus* hybrids and crosses among *B. taurus* breeds. One explanation is that there is more gene frequency divergence among the various indigenous *B. indicus* breeds than among the *B. taurus* breeds.

While *indicus* $\times$ *taurus* crosses do indeed show significant heterosis, they can also have unanticipated economic disadvantages. McDowell (1985) points out, regarding the hybrids, that "some farmers have reservations that could influence national breeding programs. The two-breed crossbred male is not as temperamental as the favored draft breed, e.g., Hariana in India, and does not move as rapidly for plowing or in performing cartage. The smaller hump of the crossbred is not well suited to handle the traditional wooden yoke. For those

reasons the crossed male has a price discrimination against it in the draft market." Further, crossbreeds typically require supplemental feeding, otherwise they can become nutritionally stressed (McDonnell 1981, 1985).

HETEROSIS IN NATURAL POPULATIONS

Brincic (1954) Genetics 39: 77-88
Wallace (1955) Evolution 9: 302–316

VARIANCES WITHIN AND BETWEEN CROSSES

We now turn our attention in the rest of this chapter to some general machinery for the analysis (and prediction) of crosses and selection among and between crosses. Our first consideration is in the expected covariances between members of a cross and how these are related to genetic variance components.

Hybrid populations are typically outside of both Hardy-Weinberg and gametic-phase equilibrium. Indeed, this is often the purpose for generating a hybrid — to create an excess of certain genotypes (e.g., heterozygotes) over what would be expected in a population under Hardy-Weinberg. Here we explore how the genetic variances are redistributed within and between hybrids. As always, some reference population is required for the genetic variances, in this case it would be the hypothetical population generated by randomly mating the parental lines until equilibrium is reached. For unlinked loci, both Hardy-Weinberg and gametic-phase equilibrium is reached in the F_2. With linked loci, additional generations are required to reach linkage equilibrium.

Covariances for Single Crosses

The basic machinery and approaches for computing the covariances between to relatives (examined in detail in Lynch and Walsh 1998 [henceforth LW] Chapter 7) can be applied to single crosses. Recall (LW Equation 7.12) that the general expression for a genetic covariance between two relatives, z and w, is given by

$$\begin{aligned}\sigma(G_z, G_w) &= 2\theta_{zw}\sigma_A^2 + \Delta_{zw}\sigma_D^2 + (2\theta_{zw})^2\sigma_{AA}^2 + \cdots \\ &= \sum_{u,v}(2\theta_{zw})^u \Delta_{zw}^v \sigma_{A^u D^v}^2 \end{aligned} \tag{21.8}$$

where Θ_{zw} is the coefficient of coancestry (the probability that single alleles drawn at random from z and w are identical by descent) and Δ_{zw} is the coefficient of fraternity (the probability that both alleles at a locus in z and w are identical by descent).

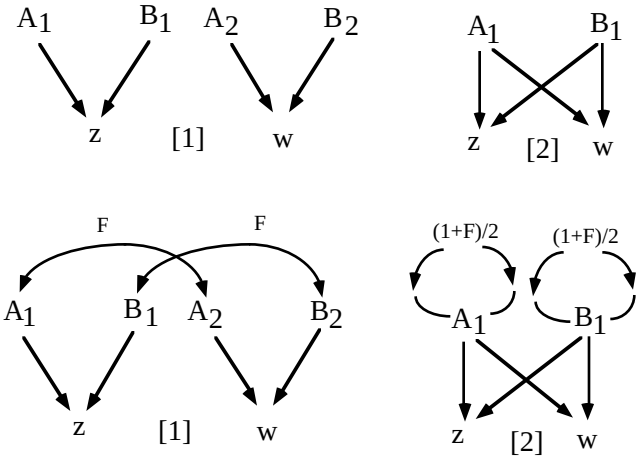


Figure 21.4. **Top:** Two types of relationships between two relatives z and w from a single cross. In situation [1], both relatives have different parents drawn from the two lines, while in situation [2] both relatives have the identical parents. **Bottom:** Any genetic correlations generated are due to correlations between parental alleles. In situation [1], this requires the parental lines be inbred ($F \neq 0$), while relatives have the same parent in situation [2], so that the correlation between parental alleles is $(1+F)/2$ where F is the amount of inbreeding in the line.

Figure 21.4 illustrates the approach for computing Θ_{zw} and Δ_{zw} between two relatives z and w that are the result of a cross between (unrelated) lines A and B. If both sets of parents are different (situation [1] in Figure 21.4), any genetic covariance between the single-cross relatives z and w results only from relationships between their parents (A_1 and A_2, B_1 and B_2), and hence requires one (or both) lines to be inbred for any genetic covariance to be generated. In the absence of any inbreeding ($F = 0$), there is no genetic covariation between the relatives. In the other extreme where both relatives have identical parents (situation [2]), there is a genetic covariance between the relatives even if their parental lines are not inbred. From the path diagrams in Figure 21.4, the coefficients of coancestry and fraternity immediately follow, and these are summarized in Table 21.3.

Table 21.3. Coefficients of coancestry (Θ_{zw}) and fraternity (Δ_{zw}) for two relatives z and w from a single cross when both, one, or no parents are shared. The general values under any level of inbreeding of the crossed lines, as well as the special (but important) cases of a completely inbred and a fully outbred line, are given.

	General		Inbred Lines		Outbred Lines	
Shared Parents	2θ	Δ	2θ	Δ	2θ	Δ
Both	$\frac{1+F}{2}$	$[\frac{1+F}{2}]^2$	1	1	$\frac{1}{2}$	$\frac{1}{4}$
One	$\frac{1+3F}{4}$	$\frac{F(1+F)}{2}$	1	1	$\frac{1}{4}$	0
None	F	F^2	1	1	0	0

Note that if we are crossing completely inbred lines, $F = 1$, it does not matter whether parents are shared, as all have the same genotype. With completely outbred lines ($F = 0$), the

coefficients are the same as for full sibs (when they share both pairs) and half sibs (they only share a single parent). In many plant breeding applications, crosses involve highly inbred lines with many parents from each line being involved. In this case we can generally assume no common parents (i.e., $2\theta = F$, $\Delta = F^2$) without much error (see Cockerham 1961 for an exact treatment). Recalling from ANOVA theory that since the variance between groups equals the within-group covariance, the variance between the mean values of different single crosses equals the covariance within a given single cross. Assuming at least a modest number of parents from each line, the genetic variance between means in a cross becomes

$$\sigma^2_{G(\mathrm{b})} = F\sigma^2_A + F^2\sigma^2_D + F^2\sigma^2_{AA} + \cdots = \sum_{u,v} F^{u+2v}\sigma^2_{A^uD^v} \tag{21.9a}$$

If the lines are not completely inbred ($F \neq 1$), the within-line genetic variance is the difference between the total and between-group genetic variance, or

$$\sigma^2_{G(\mathrm{w})} = \sum_{u,r}(1 - F^{u+2r})\sigma^2_{A^uD^r} \tag{21.9b}$$

The values in Table 21.3 can be used to generate similar expressions when one, or both, parents are shared.

The variance between the means of single crosses can also be expressed in terms of the variances associate with the general (GCA) and specific (SCA) combining abilities. Since the mean of any single cross equals the sum of the general combining abilities of its parents plus the specific combining ability for the cross, and since these (by construction) are uncorrelated, the variance of this sum equals the sum of the variances, or

$$\sigma^2_{\mathrm{b}} = \sigma^2_{GCA}(P) + \sigma^2_{GCA}(M) + \sigma^2_{SCA} \tag{21.10}$$

where GCA(P) denotes the parental (sire/pollen parent) GCA and GCA(M) the maternal (dam/seed parent) GCA. These variances can be written in terms of genetic variance components as

$$\sigma^2_{GCA} = \frac{1}{2}F\sigma^2_A + \frac{1}{2^2}F^2\sigma^2_{AA} + \cdots + \frac{1}{2^u}F^u\sigma^2_{A^u} + \cdots \tag{21.11a}$$

and

$$\sigma^2_{SCA} = \frac{1}{2}F^2\sigma^2_{AA} + \sum_u \left(1 - \frac{1}{2^u}\right)F^u\sigma^2_{A^u} + \sum_{r\geq 1,u} F^{u+2r}\sigma^2_{A^uD^r} \tag{21.11b}$$

As is expected, the variance of specific combining abilities is entirely a function on nonadditive variances. Further, since all the genetic variance components in the specific combining ability variance scale with at least F^2, the effects of SCA are modest at low levels of inbreeding, and become only important as the inbreeding coefficient approaches one.

The careful reader will note that we have still not discussed exactly what population these genetic variances refer to, or how we estimate them. If all individuals involved in the crosses are random mated for several generations, the resulting genetic variances in such a (linkage equilibrium) population are the values we would use above. Estimation of these variance components from three way crosses is examined by Rawlings and Cockerham (1962a) and for double-crosses by Rawlings and Cockerham (1962b).

Covariances for Three- and Four-way Crosses

In a similar fashion, covariances for three- (3W) and four-way (4W) crosses can be obtained, but the bookkeeping becomes a little more involved (Cockerham 1961). Table 21.4 illustrates

this, showing the rather large number of possible comparisons between crosses sharing at least one common line. Another level of complication is that, once again, the individuals from the common line may be the same, or different, individuals. Again, this makes no difference when the lines are highly inbred, but can be an important factor when the lines are only weakly inbred.

Table 21.4. Coefficients of coancestry (Θ_{zw}) and fraternity (Δ_{zw}) for two relatives z and w from three- and four-way crosses. A.BC denotes the three-way cross with line A crossed to a BC single hybrid, while AB.CD denotes the four-way cross where a AB hybrid is crossed to a CD hybrid. For relative comparisons, lines in common are given, while a dash (–) denotes a different line for the two relatives. z and w are assumed to use different parents from any shared lines. After Cockerham (1961).

Three-way (3W) Crosses

Lines in		General F		F =1	
Common	Relative Types	2θ	Δ	2θ	Δ
3	A.BC, A.BC	$3F/4$	$F^2/2$	3/4	1/2
3	A.BC, B.AC	$5F/8$	$F^2/4$	5/8	1/4
2	A.B–, A.B–	$5F/8$	$F^2/4$	5/8	1/4
2	A.B–, B.A–	$F/2$	$F^2/4$	1/2	1/4
2	A.B–, –.AB	$3F/8$	0	3/4	0
2	–.AB, –.AB	$F/4$	0	1/4	0
1	A.– –, A.– –	$F/2$	0	1/2	0
1	A.– –, –.A –	$F/4$	0	1/4	0
1	–.A –, –.A –	$F/8$	0	1/8	0

Four-way (4W) Crosses

Lines in		General F		F =1	
Common	Relative Types	2θ	Δ	2θ	Δ
4	AB.CD, AB.CD	$F/2$	$F^2/4$	1/2	1/4
4	AB.CD, AC.BD	$F/2$	$F^2/8$	1/2	1/8
3	AB.C–, AB.C–	$3F/8$	$F^2/8$	3/8	1/8
3	AB.C–, AC.B–	$3F/8$	$F^2/16$	3/8	1/16
2	A –.B–, A –.B–	$F/4$	$F^2/8$	1/4	1/16
2	AB.– –, AB.– –	$F/4$	0	1/4	0
2	A –.B–, AB.– –	$F/4$	0	1/4	0
2	A –.– –, A –.– –	$F/8$	0	1/8	0

Note that when we consider three- and four-way crosses that the *order* of the crosses is critical. For example, an individual from a $A \times (B \times C)$ triple cross has a higher correlation with another individual from this same cross than with an individual from a $B \times (A \times C)$, even though the same three lines are shared. Note for double-crosses that the two types of DCs sharing the same four parents have the same additive variances (the θ values are the same), but differ by a factor of two in the weighting (Δ) of any dominance variance (and by 2^k for any k-order epistastic term involving dominance).

While Table 21.4 examines the covariances involving all possible pairs, for selection response we are usually more interested in the covariance when all members share the maximal number of common lines (two for a SC, three for a 3W, and four for a 4W). Table 21.5 gives the coefficients for resemblance in this case for both no common parents from any line and for the other extreme case were all parents from each line are shared.

Table 21.5. Coefficients of coancestry (Θ_{zw}) and fraternity (Δ_{zw}) for individuals from the same single, double, or triple crosses. Different parent implies no common parents from any of the shared lines, while same parents is the opposite extreme where both z and w share all the same parents from all common lines..

	Different Parents		Same Parents	
Cross	2θ	Δ	2θ	Δ
Single	F	F^2	$(1+F)/2$	$(1+F)^2/4$
Three-way	$3F/4$	$F^2/2$	$(3/8)(1+F)$	$(1+F)^2/8$
Double	$F/2$	$F^2/4$	$(1+F)/4$	$(1+F)^2/16$

It is instructive to examine the between cross mean differences under the three types of hybrids (SC, 3W, 4W). Assuming parental lines are completely inbred, these are

$$\sigma^2_{G(\text{b, SC})} = \sigma^2_A + \sigma^2_D + \sigma^2_{AA} + \cdots = \sum_{u,v} \sigma^2_{A^u D^v} \tag{21.12a}$$

$$\sigma^2_{G(\text{b, 3W})} = \frac{3}{4}\sigma^2_A + \frac{1}{2}\sigma^2_D + \frac{9}{16}\sigma^2_{AA} + \cdots = \sum_{u,v} \left(\frac{3^u}{2^{2u+v}}\right) \sigma^2_{A^u D^v} \tag{21.12b}$$

$$\sigma^2_{G(\text{b, 4W})} = \frac{1}{2}\sigma^2_A + \frac{1}{4}\sigma^2_D + \frac{1}{4}\sigma^2_{AA} + \cdots = \sum_{u,v} \left(\frac{1}{2^{u+2v}}\right) \sigma^2_{A^u D^v} \tag{21.12c}$$

Likewise, since $\sigma^2_{G(\text{w})} = \sigma^2_{G(\text{total})} - \sigma^2_{G(\text{b})}$, the genetic variation within a particular hybrid line are

$$\sigma^2_{G(\text{w, SC})} = 0 \tag{21.13a}$$

$$\sigma^2_{G(\text{w, 3W})} = \frac{1}{4}\sigma^2_A + \frac{1}{2}\sigma^2_D + \frac{7}{16}\sigma^2_{AA} + \cdots = \sum_{u,v} \left(1 - \frac{3^u}{2^{2u+v}}\right) \sigma^2_{A^u D^v} \tag{21.13b}$$

$$\sigma^2_{G(\text{w, 4W})} = \frac{1}{2}\sigma^2_A + \frac{3}{4}\sigma^2_D + \frac{3}{4}\sigma^2_{AA} + \cdots = \sum_{u,v} \left(1 - \frac{1}{2^{u+2v}}\right) \sigma^2_{A^u D^v} \tag{21.13c}$$

Selection Among Line Crosses

The above variances have direct implications in selection decisions. Suppose each cross is replicated by n individuals in each of r plots, and we base selection on the means (over all replicates) of each cross. Analogous to our development of Equation 8.13 (family-based selection), selecting among crosses (i.e., selecting those crosses with the highest means), the expected response using crosses of type c and a selection intensity of $\overline{i}$ is

$$R_c = \overline{i}\,\frac{\sigma^2_{G(\text{b, c})}}{\sqrt{\sigma^2(\,\overline{z}_c)}} \tag{21.14a}$$

where the variance for the mean of any particular cross follows from Equation 8.38,

$$\begin{aligned} \sigma^2(\,\overline{z}_c\,) &= \sigma^2_c + \frac{\sigma^2_{E_c}}{r} + \frac{\sigma^2_{w_c}}{rn} \\ &= \sigma^2_{G(\text{b, c})} + \frac{\sigma^2_{E_c}}{r} + \frac{\sigma^2_{G(\text{w, c})} + \sigma^2_{E(\text{w, c})}}{rn} \end{aligned} \tag{21.14b}$$

where $\sigma^2_{E_c}$ is the between-plot variance and $\sigma^2_{E(\mathrm{w,\ c})}$ the within-plot variances. Genotype $\times$ environment interactions can be quite important, and the obvious modification of expressions for family-based selection (e.g., Equation 8.39) can be used to incorporate these effects.

The relative order of the expected respondes under the three common hyrbid crosses are $R_{SC} > R_{3W} > R_{DC}$. Relative to the response based on double crosses, if all the genetic variance is additive, the response is 1.5 times greater using three-way crosses, and 2 fold greater with single crosses. These are minimal values, and become even larger as nonadditive variance becomes increasingly important.

Example 21.2. Consider the expected response to selecting among a set of single crosses versus selecting among the set of double crosses from the same initial set of lines. Suppose there is no epistatic variance and no genotype $\times$ environment interactions, so that the between- and within-plot environmental variances are the same for both types of crosses. Ignoring epsitasis, the between-line variances (i.e., the variacne of the means of the crosses) are

$$\sigma^2_{G(\mathrm{b,\ SC})} = \sigma^2_A + \sigma^2_D + E^2, \quad \text{where} \quad E^2 = \frac{\sigma^2_{E_c}}{r} + \frac{\sigma^2_{E(\mathrm{w,\ c})}}{rn}$$

and

$$\sigma^2_{G(\mathrm{b,\ DC})} = \frac{1}{2}\sigma^2_A + \frac{1}{4}\sigma^2_D + \frac{\sigma^2_A/2 + 3\sigma^2_D/4}{rn} + E^2$$

Where the genetic variances are those for the conceptuation population formed by randomly mating all the parents used to construct the various crosses. The response to selection under the two different corssing schemes are

$$R_{SC} = \bar{\imath}\,\frac{\sigma^2_A + \sigma^2_D}{\sqrt{\sigma^2_A + \sigma^2_D + E^2}}$$

and

$$R_{DC} = \bar{\imath}\,\frac{\sigma^2_A/2 + \sigma^2_D/4}{\sqrt{\frac{1}{2}\sigma^2_A + \frac{1}{4}\sigma^2_D + \frac{\sigma^2_A/2+3\sigma^2_D/4}{rn} + E^2}}$$

If the environmental variance E^2 is large relative to the genetic variance, the variances of the means are approximately equal and the ratio of response under these two schemes becomes

$$\frac{R_{SC}}{R_{DC}} \simeq \frac{\sigma^2_A + \sigma^2_D}{\sigma^2_A/2 + \sigma^2_D/4} = \frac{2+\gamma}{1+\gamma/4}, \quad \gamma = 2\sigma^2_D/\sigma^2_A$$

Thus, (under these assumptions) the response using single-cross means produces at least twice the response as using double crosses, with the response approaching a four-fold difference as dominance becomes increasingly more important (γ large). For example, the response is three-fold larger when $\sigma^2_D = 2\sigma^2_A$.

Finally, note that even after we have chosen three particular lines (A, B, C), then (ignoring reciprocals) there are three different **orders** for three-way crosses involving these lines: A$\times$(B $\times$ C), B$\times$(A $\times$ C), and C$\times$(A $\times$ B). Cockerham (1961) showed that the genetic covariance between the means of orders is given by

$$\sigma_{O,3W} = \sigma(\text{A.BC,A.BC}) - \sigma(\text{A.BC,B.AC}) \tag{21.15a}$$

From Table 21.3, the corresponding values of 2θ and Δ for this covariance are

$$2\Theta = 3F/4 - 5F/8 = F/8, \qquad \Delta = F^2/2 - F^2/4 = F^2/4 \tag{21.15b}$$

Using these values, one can compute the within and between group genetic variances, and by applying Equation 21.14a and b compute the response to selection based on orders, that is selecting a set of orders based on the deviation of the mean of that order (i.e., the mean over all crosses in the order) and the grand mean (the mean of all crossed, regardless of order). While the reader might ask about the utility of such a selection (since we would only be selecting among three groups), such covariances allow us to compute a selection index, weighting between order deviations, and cross within an order deviations,

$$I = b_O(\,\overline{z}_{i,\cdot} - \overline{z}_{\cdot,\cdot}\,) + b_c(\,\overline{z}_{i,j} - \overline{z}_{i,\cdot}\,) \tag{21.16}$$

where $\overline{z}_{i,j}$ is the mean of cross j within order i, $\cdot$ denotes averaging over that index, and b_O and b_c are the order and cross within-order weights.

Similarly, (ignoring reciprocals) there are three possible four-way crosses for any set of four lines, AB.CD, AC.BD, and AD.BC. The genetic covariance between the means of the orders is

$$\sigma_{O,DC} = \sigma(\text{AB.CD,AB.CD}) - \sigma(\text{AB.CDAC.BD}) \tag{21.17a}$$

Again referring to Table 21.3, the corresponding values of 2θ and Δ for this covariance become

$$2\Theta = F/2 - F/2 = 0, \qquad \Delta = F^2/4 - F^2/8 = F^2/8 \tag{21.17b}$$

Selection Within Line Crosses

Under between-line selection, one makes a series of crosses and selects those crosses with the highest means. Under within-line selection, we select the best indivdiuals with each cross. The expected response to selection becomes

$$R_w = \overline{\imath}\,\frac{\sigma^2_{G(\text{w, c})}}{\sqrt{\sigma^2(\,\text{within}\,)}} \tag{21.18}$$

Equations 21.13a-13c give the within-cross variances for single, triple and double crosses. There is no genetic variance within single cross lines formed from complete inbreds, as all individuals are heterozygotes. This is not the case with three- and four-way crosses where one (or both) parents are heterozgotes and hence segregate gametes. Ignoring epistasis, the within-line variation for a three-way cross is $\sigma^2_A/4+\sigma^2_D/2$ and $\sigma^2_A/2+(3/4)\sigma^2_D$ for a four-way cross. As with family selection, one could construct a selection index weighting both within- and between-line deviations,

$$I = b_b(\,\overline{z}_{i,\cdot} - \overline{z}_{\cdot,\cdot}\,) + b_w(\,z_{i,j} - \overline{z}_{i,\cdot}\,) \tag{21.19}$$

The optimal weights for the the between b_b and within b_w weighs are given in the index selection chapter.

ESTIMATING HETEROTIC AND CROSSBREEDNG COMPONENTS

GCA, SCA

The Gardner-Eberhart Decomposition of Heterosis

Suppose we have performed a diallel analysis involving all $n(n-1)/2$ pairwise crosses among n lines. Gardner and Eberhart (1966) suggest that we parition the variation in heterosis between crosses as follows:

$$h_{ij} = \mu_{ij} - \frac{\mu_{ii} + \mu_{jj}}{2} = \overline{h} + h_i + h_j + s_{ij} \qquad (21.\text{xx})$$

where the **average heterosis** $\overline{h}$ is the average value of heterosis among all the $n(n-1)/2$ crosses, h_j is the **variety heterosis**, the average heterosis from crosses involving line j expressed as a deviation from the average heterosis (so that $\sum_i h_i = 0$), and $s_{ij} = h_{ij} - (\overline{h} + h_j + h_i)$ is the **specific heterosis** that occurs in the cross $i \times j$. The restriction on the specific heterosis is that $\sum_i s_{ij} = \sum_j s_{ij} = 0$.

Example 21.x: Gardner and Eberhart (1966) examined mean grain yield in six open pollinated mazie varietiess along with all 15 pair-wise (nonreciporical crosses), i.e., using their Design II. The resulting ANOVA table was as follows:

Source	df	MS
Populations	n(n+1)/2 - 1 = 20	23.66
Varieties (v_i)	n - 1 = 5	46.85**
Heterosis (h_{ij})	n(n-1)/2 - 1 = 15	15.93
Average ($\overline{h}$)	1	115.44**
Variety (h_j)	n - 1 = 5	11.94
Specific (s_{ij})	n(n - 3)/2 = 5	7.08

While there were significant varietal effects and a significnat average heterogysoity, there were not signifincatn differences among vareity or specific heterosis. Thus no paricualr combination of lines is expected to have exceptional heterosis relative to the other lines, and hence one can chose the best lines simply from their variteal values.

Connection with GCA, SCA

Gardner and Eberhart discuss three designs. Under their Design I, five types of populations are considered: tvarieties (μ_{ii}), their selfed progenies (μ^s_{ii}), all pairwise crosses (μ_{ij}), and the selfed (μ^s_{ij}) and random-mated (μ^r_{ij}) progeny from these crosses. The advantage of this full design is that it allows for seperate estiamtion of both the contribution α_i from homozgyous loci and the contribution Δ_i from heterozygous loci, in additional to all heterotic components ($\overline{h}, h_j, s_{ij}$). The line means are given by

$$\mu_{ii} = \mu + \alpha_i + \Delta_i \qquad (21.xx\text{a})$$

$$\mu^s_{ii} = \mu + \alpha_i + \frac{1}{2}\Delta_i \qquad (21.xx\text{b})$$

$$\mu_{ij} = \mu + \frac{\alpha_i + \alpha_j}{2} + \frac{\Delta_i + \Delta_j}{2} + h_{ij} \qquad (21.xx\text{c})$$

$$\mu^s_{ij} = \mu + \frac{\alpha_i + \alpha_j}{2} + \frac{\Delta_i + \Delta_j}{4} + \frac{h_{ij}}{2} \qquad (21.xx\text{d})$$

$$\mu^r_{ij} = \mu + \frac{\alpha_i + \alpha_j}{2} + \frac{\Delta_i + \Delta_j}{2} + \frac{h_{ij}}{2} \qquad (21.xx\text{e})$$

Example 21.x: The example covers the defintions of α_i and Δ_i and hence may be skipped by the cahausl reader. Consider the kth locus underlying the character, where the genotypes $\mathbf{A}_k\,\mathbf{A}_k : \mathbf{A}_k\,\mathbf{a}_k : \mathbf{a}_k\,\mathbf{a}_k$ have genotypic values of $\mu^* + a_k \;:\; \mu^* + d_k \;:\; \mu^* - a_k$, and let p_{ik} be the frequency of allele $\mathbf{A}_k$ in line k. For m loci underlying the character, define

$$\alpha = \alpha_1^* - \overline{\alpha}, \quad \text{where} \quad \alpha_i^* = \sum_k^m (\,2p_{ik} - 1)\,a_k, \quad \overline{\alpha} = \frac{1}{M}\,\sum_i^m\,\alpha_i^*$$

Under this construction, the α_i represent the contribution from homozygous loci. In particular, defining $\mu = \mu^* + \overline{\alpha}$, them $\mu + \alpha_i$ is the mean of random inbred lines created from vareity i. The contribution from heterozygous loci is measured by

$$\Delta_i = 2\,\sum_i^m\,(\,p_{ik} - p_{ik}^2\,)\,d_i$$

Finally, the heterotic contribution (from Equation 21.2c) is just given by

$$h_{ij} = \sum_i^m\,(\,p_{ik} - p_{jk}\,)^2\,d_i$$

Design one is rather unworkable, requiring $2n + 3n(n-1)/2 = n(1+3n)/2$ populations. A more useful model is Gardner and Eberhart's Design II, which requires $n(n+1)/2$ populations, all varieties plus all pair-wise crosses. This desing also allows for estiamtion of all heterotic components ($\overline{h}, h_j, s_{ij}$) along with a variety effect $v_i = \alpha_i + \Delta_i - \overline{\Delta}$ where $\overline{\Delta} = \sum\,\Delta_i/n$. Under this parameteization,

$$\mu_{ij} = \mu_v + \frac{v_i + v_j}{2} + \gamma_{ij}\,(\,\overline{h} + h_i + h_j + {}_{ij}\,) \tag{21.xx}$$

where

$$\gamma_{ij} = \begin{cases} 0 & \text{when } i = j \\ 1 & \text{when } i \neq j \end{cases}$$

and the varitey mean $\mu_v = \mu + \overline{\Delta}$.

The final design considered by Gardner and Eberhart is when only the $n(n-1)/2$ pairwise crosses are considered. In this case, the mean μ_c of all crosses is $\mu_c = \mu_v + \overline{h}$, and the vareity effect in crosses is $g_i = v_i/2 + h_i$, and the mean of $i \times j$ is written as

$$\mu_{ij} = \mu_c + g_i + g_j + s_{ij}$$

where g_i is the general combining ability of line i and s_{ij} is the specific combining ability of lines i and j.

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

$$\mu_A = \mu + g_A^I + g_A^M + g_A^{M'} \tag{21.13a}$$

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

$$\mu_{AB} = \mu + \frac{g_A^I + g_B^I}{2} + g_B^M + g_B^{M'} + h_{AB}^I \tag{21.13b}$$

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

$$\mu_{BA} = \mu + \frac{g_A^I + g_B^I}{2} + g_A^M + g_A^{M'} + h_{AB}^I \tag{21.13c}$$

From Equation 21.13a-c, it follows that an estimate of the individual (direct) heterotic effect is the obvious one (Nitter 1978),

$$\frac{\mu_{AB} + \mu_{BA}}{2} - \frac{\mu_{AA} + \mu_{BB}}{2} = h_{AB}^I \tag{21.13d}$$

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

$$\mu_{BA} - \mu_{AB} = \left(g_A^M + g_A^{M'}\right) - \left(g_B^M + g_B^{M'}\right) \tag{21.13e}$$

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, the 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 $_2$ 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,

$$\mu_{C\cdot AB} = \frac{2g_C^I + g_A^I + g_B^I}{4} + \frac{h_{CA}^I + h_{CB}^I}{2} + \frac{g_A^M + g_B^M}{2} + h_{AB}^M + g_B^{M'} + \frac{r_{ab}^I}{2} \tag{21.14a}$$

Here since the dam is a crossbred, there is a potential heterotic component (h_{AB}^M), 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 21.4 summarizes these coefficients for a variety of crosses. As Equation 21.med indicates, estimates of the various effects can be obtained by suitably weighted combination of line means. Combining 21.13b and 21.14a gives (Nitter 1978),

$$\mu_{C\cdot AB} - \frac{\mu_{CA} + \mu_{CB}}{2} = h^M_{AB} + \frac{r^I_{ab}}{2} \tag{21.14b}$$

Hence, if recombination effects are small (these would be absence in the absence of epistasis), Equation 21.14b allows for a direct estimate of maternal heterotic effects. Table 21.5 summarizes estimates of for individual and maternal heterotic effects for a variety of traits in sheep using this approach.

Table 21.5. 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	$g^{M'}$	$h^{M'}$	$r^{M'}$
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	

Using Testors

Top-cross testing, crossing the lines to a common **testor**, is often use to reduce an initial sample of lines to a smaller set of more desirable lines for further testing (i.e., a diallel or three-ways).

Using an inbred line vs. an open-pollinated or synthetic population

broad-based genetic tester (GCA)

inbred line (SCA)

The more negative recessive alleles in the tester, the better (i.e., use a weak line as the testor).

Rawlings and Thompson (1962): Consider a single locus, the genetic variance among the testcross progeny is

$$\sigma^2_t = (1+F)\,\frac{\sigma^2_A}{4}\left(1 + (1-2t)\frac{d}{a}\right)^2 \tag{21.xx}$$

where t is the allele frequency in the testor and σ^2_A the equilibrium additive variance in the population created by randomly mating the tested lines.. If the testor is an inbred line, then $t = 1$ (fixed for the dominant allele, or $t = 0$ (fixed for the recessive). If the testor is an open-pollinated line (or a synthetic variety), then t is the allele frequency in the sample of testor parents used. The larger the among-progeny genetic variance, the greater the discriminatory power of the testor to detect differences among inbred lines. As shown in Figure 21.2, for any degree of dominance (d/a), a recessive testor ($t = 0$) is best.

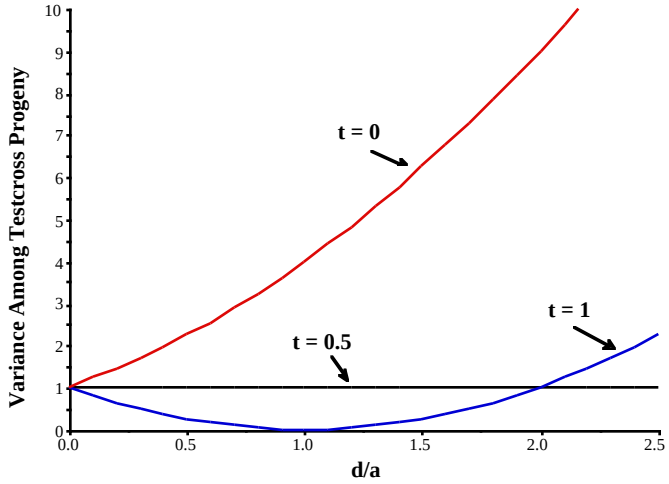


Figure 21.2. The among-testcross progeny genetic variance for a single locus (σ^2_t) as a function of the allele frequency t of the dominant allele in the testor and the degree of dominance (d/a). Modified from XXX and XXX corn book.

Top crosses as predictors of double hybrid performance

Davis (1929) **top cross** – by crossing each line to a common tester, as estimate of their combining ability (and hence predicting hybrid values) can be obtained

Independently suggested by Jenkins and Brunson (1932), who used as a testor a varietal mixture comprised of equal amounts of seeds from 25 open-pollinated corn lines.

Original **recurrent selection**, proposed by Hull (1945, 1952), involves crossing segregating population against a tester strain (serving as one of the parent) and then picking the best parents based on this type of progeny test. Selection for specific combining ability with the tester.

recurrent reciprocal selection (the term originally suggested by Comstock et al. (1949) is now more commonly called **reciprocal recurrent selection**, or **RRS**. Selection for both general and specific combining ability.

pollen parent = **tester parent**, initially open-pollinated line

pre 1970's outbred lines used as testers for estimating GCAs

Inbred lines now used as testers

genetic variance roughly twice as large among inbred testers vs. broad-based (i.e., outbred) testers – **Darrah et al 1972, Russell et al 1973**, Russell and Eberhart (1975)

Zambezi et al (1986): correlations between inbreed testers and GCA were similar to corrections using broad-based testers and GCA.

Jenkins (1934) looked at single crosses, diallel crosses and top crosses, finding correla-

tions of 0.76, 0.73, and 0.61 with actual double-cross values.

ref:

Crop SCi 13: 257-261; Crop Sci 12: 605,

General Remarks on Predicting the Means of Crosses

For various inbreeding and crossing scheme to be superior, requires that nonadditive genetic variance is important, and in particular, that overdomiance is important.

If additive variance is important, a lines performance will also be a good predictor of its crossing ability. Note from Equation 21.cx3 that if $F + 1$, then the variance in GCA equals the additive variance plus increasingly-discounted contributions from higher order additive epistatic terms, while the variance is SCA is all the residual nonadditive variance. Hence, the potential importance of selecting for SCA can be judged from estimates the these variance components. For example, if the additive variance is (say) 60 and the total genetic variance is 80, then 75% (6/8) of the response is from additive variance, and only 25% potentially from SCA, so that 3/4 of the improvements from GCA and only 25% from SCA.

Animal crosses are often three-way: a cross between two heterotic populations to provide superior females (for example, for increased liter size or offspring survivorship due to heterosis) and then a cross to a male from a population carrying some desirable trait.

THREE - AND FOUR-WAY CROSSES

Number of Triple and Double Crosses

n choose 2, or $n(n-1)/2$ potential crosses

For triple crosses ($l_1 \times (l_2 \times l_3)$), with a total of n lines, there are n choose 3, or $n(n-1)(n-2)/3!$ potential sets of parents. However, for each set of parents (l_i, l_k, l_j) there are three different crosses (ignoring reciprocals) : $l_i \times (l_k \times l_j)$, $l_k \times (l_j \times l_j)$, and $l_j \times (l_i \times l_k)$. Hence, there are a total (ignoring reciprocals) of

$$N_{TC}(n) = \frac{n(n-1)(n-2)}{3} \tag{21.xxa}$$

triple crosses involving n lines. Finally, for double crosses involving four lines, $(l_i \times l_j) \times (l_k \times l_\ell)$, with n lines there are n choose 4, or $n(n-1)(n-2)(n-3)/4!$ potentially different sets of four parents. However, as with triple crosses, within each set of four parents, there are several different crosses (three to be exact, with l_i crosses to each of the three parents), e.g., $(l_i \times l_j) \times (l_k \times l_\ell)$, $(l_i \times l_k) \times (l_j \times l_\ell)$, and $(l_i \times l_\ell) \times (l_j \times l_j)$, giving the total number of double crosses as

$$N_{DC}(n) = \frac{n(n-1)(n-2)(n-3)}{8} \tag{21.xxB}$$

For example, with 10 lines, there are 45 single, 360 triple, and 630 double crosses. For 25 lines, these become 300, 6900, and 37950, respectively. Hence, the multiple of multiple line crosses very quickly get out of control.

One suggestion offered by Weber and Wricke (19xx) is that the number of actual crosses may be less than suggested by Equation 21.xx when there is some structure, such as n_1 lines coming from one population and n_2 from another. If the major differences are from between-population crosses, and we ignored within-population crosses as not being of interest, then the number of crosses reduces substantially. For example, the number of single-crosses is now $n_1\, n_2$ (instead of $n(n-1)/2$). Weber and Wricke similar show that the number of double

and triple crosses becomes

$$\begin{aligned} n_{TC} &= \frac{n_1 n_2 (n_1 + n_2 - 2)}{2} \\ &= \frac{n^2(n-2)}{8} \quad \text{when } n_1 = n_2 = n/2 \end{aligned} \tag{21.xxa}$$

$$\begin{aligned} n_{DC} &= \frac{n_1(n_1-1)n_2(n_2-1)}{4} \\ &= \frac{n^2(n-2)^2}{64} \qquad \text{when } n_1 = n_2 = n/2 \end{aligned} \tag{21.xxa}$$

For example, if our ten lines to be crosses consist of two subpopulations of 5 lines each, then the number of triple and double crosses between subpopulations both equal 100 in this case, as opposed to 360 and 630 when considering all triple and double crosses (respectively).

The Problem

A number of breeding strategies use the hybrid of crosses between lines that show heterosis (LW Chapter XX). While a single cross often shows more heterosis that double or triple crosses, the hybrid offspring come from an inbred parent, which may show greatly reduced fertility. Using a single-cross hybrid (i.e. $\ell_1 \times \ell_2$) as the seed parent (for plant breeders) or the mother (for animal breeders) can greatly increase the number of hybrid offspring. Historically, corn breeders have based lines on double cross hybrids resulting from the crosses of four inbred lines, while animal breeders often used three-way crosses, using single-cross females (i.e., a cross between two heterotic populations) which are then crossed to males from another line/population carrying some desirable trait.

The problem breeders face is that if they start with n lines, there are order n^2 single crosses, n^3 triple crosses and order n^4 double crosses (Equations 21.xx-a-c), making it rather unfeasible to test all possible triple or double crosses. Ideally, one could use results from single crosses to draw some inferences on the possible high-performing triples and doubles. Even a strategy that removes just 50% of the lines results in considerable savings.

Further, even if a breeder has chosen four inbred lines from which to form a double cross, there are three possible sets of single cross parents and hence three possible double crosses from each set of four lines ($ij \times k\ell$, $ik \times j\ell$, $i\ell \times jk$). These three possible double crosses often have significantly different trait values (e.g., Doxtator and Johnson 1936). Further, if the choice of seed and pollen parents is critical (i.e., we also need to consider the reciprocal crosses), there are six possible doubles that need to be considered. For example, if maternal effects are important, then so is the genotype of the seed parent, so that a $ij \times kl$ cross may have different values depending on which is the seed and pollen parent.

Predicting Double and Triple Cross Means: Jenkins' Methods

Jenkins (1934) proposed four different estimators (Methods A - D) using single crosses to predict double cross values. Jenkins' **Method A** estimates the double cross value from the means of all six possible single crosses between the four parents comprising the double hybrid, so that if the double hybrid is from by the cross $(\ell_i \times \ell_j) \times (\ell_k \times \ell_l)$, Method A estimates is value as

$$\widehat{D}^A_{ij,kl} = \frac{\overline{z}_{ij} + \overline{z}_{ik} + \overline{z}_{il} + \overline{z}_{jk} + \overline{z}_{jl} + \overline{z}_{kl}}{6} \tag{21.xxa}$$

where $\overline{z}_{ik}$ denotes the mean of the (single cross) $\ell_i \times \ell_j$. Method A involves both GCAs an SCAs. **Method B** is motivated by the observation that genes from each parental line only

interact with alleles from the two parental lines forming the opposite parent (e.g.,, alleles from ℓ_i with alleles from ℓ_i and ℓ_i), suggesting the estimator

$$\widehat{D}^B_{ij,kl} = \frac{\overline{z}_{ik} + \overline{z}_{il} + \overline{z}_{jk} + \overline{z}_{jl}}{4} \tag{21.xxb}$$

Stated another way, Method B considers all single crosses of the four lines *except* the two single crosses that are paired to form the double cross ($\overline{z}_{ij}$ and $\overline{z}_{kl}$ are excluded).

Method C is based entirely on the GCAs of the four parents,

$$\widehat{D}^C_{ij,kl} = \frac{\overline{z}_{i\cdot} + \overline{z}_{j\cdot} + \overline{z}_{k\cdot} + \overline{z}_{l\cdot}}{4}, \qquad \text{where} \qquad \overline{z}_{i\cdot} = \frac{1}{p-1}\sum_{j\neq i}^{p} \overline{z}_{ij} \tag{21.xxc}$$

Finally, **Method D** (a **inbred-variety cross**) is similar to Method C, but each line is crossed to a single **testor** line ℓ_t, with

$$\widehat{D}^D_{ij,kl} = \frac{\overline{z}_{it} + \overline{z}_{jt} + \overline{z}_{kt} + \overline{z}_{lt}}{4} \tag{21.xxd}$$

The advantage of Method D is that only n crosses are required, as opposed to the order n^2 crosses required for Methods A-C. Jenkins (1934) examined the correlation between the predicted and actual double-cross means for six different maize characters using 42 double crosses (all crosses grown in a single location in a single year). Method B had the highest correlation for four of the characters, with Methods A and C had the highest correlation for one character. The average correlations (over the six characters) are 0.648, 0.653, 0.050, and 0.390 for methods A-D, respectively. The lower values for C and D each resulted from one negative correlation, otherwise the methods are rather similar. Based on these results, Jenkins suggests using Method D and choosing the uppermost 50%.

A number of subsequent studies focused on Method B, and found it to be a reasonable predictor (e.g., Doxtator and Johnson 1936; Anderson 1938; Hayes et al. 1943, 1946; Otsuka et al. 1972). Doxtator and Johnson (1936) proposed a modification of Method B for predicting triple crosses;

$$\widehat{T}^B_{ij,k} = \frac{\overline{z}_{ik} + \overline{z}_{ik}}{2} \tag{21.aa}$$

Genotype $\times$ environment interactions are the a very significant complicating factor in that single crosses are typically grown one or two years before the double-crosses are examined. Even if the same locations are used, year and year $\times$ location environmental values are expected to be different. Table 21.c1 presents the data of Hayes et al. (1946), who examine seven maize characters when the single crosses were performed in two years (1940 and 1941) and the double-crosses generated in 1943-4. Note that the correlations between the single cross predictors using crosses in 1940 and 1941 is high, while the correlations of these predicted values with the actual double crosses was lower.

Table 21.c1. Correlations between the predicted double cross values (using Jenkin's Method B) and the actual double cross values when the single and double crosses are performed in different years. Data examines 20 inbred maize lines and 49 double-crosses from a subset of these lines (chosen by Method B). Single crosses were grown in 1940 and 1940, double crosses in 1943 and 1944. r(B40,B41) is the correction between the predicted values using single crosses grown and 1940 and single crosses grown in 1941, r(B40,DC) the correction between the average double cross values (from 1943 and 1944) and their predicted values based on single crosses grown in 1940, r(B41,DC) is similarly defined. After Hayes et al. (1946).

Character	r(B40,B41)	r(B40,DC)	r(B41,DC)
Date silked	0.86	0.71	0.76
Plant height	0.85	0.79	0.71
Ear height	0.60	0.51	0.44
Ear lenght	0.88	0.81	0.72
Kernel rows	0.76	0.71	0.64
% moisture	0.90	0.76	0.64
Yield	0.89	0.72	0.72

The Cockerham and Eberhart Estimators

Several authors have extended Jenkin's basic estimators. Cockerham (1967) notes that a selection index (an appropriately-weighted linear combination, see Chapter XX) can be constructed using the various Jenkins estimators to obtain an optimal estimator, having the highest correlation between the true and predicted values. Otsuka et al. (1972) show that the **Cockerham optimal weight predictor** can be expressed in terms of the general and specific combining abilities of all possible single crosses between the four parental lines,

$$\widehat{\mu}_{ij\cdot k\ell} = \frac{g_i + g_j + g_k + g_\ell}{2} + \gamma \cdot \frac{s_{ik} + s_{i\ell} + s_{jk} + s_{j\ell}}{4} \qquad (21.\text{xxa})$$

where for p lines,

$$\gamma = \left(\frac{H^2 - h^2}{1 - h^2}\right)\left(\frac{1 + (p-4)h^2/2}{H^2 + (p-4)h^2/2}\right) \qquad (21.\text{zzb})$$

where h^2 and H^2 are the narrow- and broad-sense heritabilities, respectively. Recall from Chapter 8 that since we are measuring groups (line means) as opposed to single individuals, that the heritability is a function of the experimental design (e.g.. Equation 8.39). In particular, increasing the number of replicates and/or the number of environments (locations and/or years) also increases the heritability by decreasing the residual environmental variance. Recalling from the definitions of GCA and SCA that for any single cross $\mu_{ij} = g_i + g_j + s_{ij}$, to follows that a value of $\gamma = 1$ corresponds to Jenkin's Method B, while Method C corresponds to $\gamma = 0$. Cockerham shows that his optimal predictor equals Jenkin's method A when all the genetic variation is additive, and equals Jenkin's method B when all the genetic variance is dominant. Note that this later condition requires all the gene action be overdominant, as other types of dominance generally generate at least some additive genetic variance.

Otsuka at el. extend Cockerham's results to provide an optimal predictor of a three-way cross as

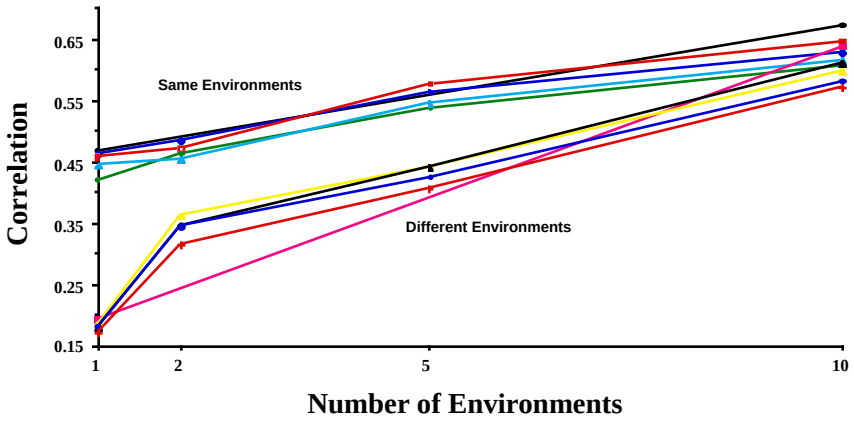
$$\widehat{\mu}_{ij\cdot k} = g_k + \frac{g_i + g_j}{2} + \frac{\gamma}{2}\left(\, s_{ik} + s_{jk} \,\right) \qquad (21.\text{xxa})$$

where γ is again given by Equation 21.zzb.

By far the biggest obstacle to accurate prediction of double-crosses is genotype $\times$ environment interaction (Eberhard and Hallauer, 1968). The usual situation is that a series of single crosses between the potential lines are grown over a set of locations and perhaps over several years, and it is these single cross values for this set of environments that is used to predict the performance of double crosses over a potentially different set of environments (in particular, in future years). Table 21.c1 summarizes the work of Hayes et al. (1946), showing high between year correlations in the predictive values when the single cross lines are grown in different years and their predicted double-cross means compared. The correlation

between the predicted and actual values are less. The situation is even more complicated when trying to use an optimal predictor, which requires estimation of the genetic and G $\times$ E variance components (in order to estimate h^2 and H^2). Example 21.1 details an extensive study by Otsuka at el. (1972), who conclude that the differences between the various predictors is small, especially when lines are examined over multiple environments.

Example 21.1. Otsuka at el. (1972) examined a diallel crossing involving ten elite maize inbred lines. All 45 (non-reciprocal) single crosses and 45 double crosses were examined for yield with two replications of each line over a total of 21 environments during a two year (1964-5) growing periods in the US Corn Belt. Double cross values were predicted from single cross values using Jenkins Methods A - C and two versions of the Cockerham optimal weight predictor. Recall that the optimal predictor requires estimates of both broad- and narrow-sense heritability. Otsuka et al compared the optimal index using variance estimates from the literature (referred to these as "average" optimal weights) and also using variance estimates obtained from the individual crosses from this experiment ("individual" optimal weights). They then examined the correlation between each of these five estimates (A, B, C, average optimal, individual optimal) when the single crosses used for prediction where grown in the same environment as the tested double crosses and the correlation when the SCs were grown in different environments. They name these comparisons over sets of one, two, five, and ten environments. While the reader is referred to their paper for the details, their basic conclusion is summarized in the figure below. The upper bundle of five lines (one for each of the methods) are for correlations within the same environment, the bottom bundle of fines lines for correlations across environments.



The basic conclusions of Otsuka et al. are obvious from this figure. First, the differences between the five methods was small. Second, the correlation between predicted and observed double-cross values is higher when the single cross lines are grown in the same environment as the double crosses. Third, this difference is most apparent when lines are replicated over a few environments. For ten environments, there was little difference between the correlations for the same versus across environments.

Finally, the authors examine the correlation between observed and predicted values as a function of γ, finding that the optimal value was generally close to 0.5 over all number of environments replicated (1,5,10), although the relationship is rather flat. They concluded that the standard approach of using Jenkin's method B is certainly satisfactory, especially when the fewer required lines can be leveraged into more replications across environments. They point out that to examine all 45 dialleles from the 10 line cross, with each replaced twice in five environments requires only 450 field plots. This information is sufficient to predict the values

of all 630 possible DC's. They suggest using the Jenkin's estimate to first choose a subset of the lines and then examine the predicted superior DCs over a number of environments.

Finally, to deal with the complications from epistasis, Eberhart (1964) proposed two other estimators based on using three-way crosses,

$$\widehat{D}^{tij}_{ij,kl} = \frac{\overline{z}_{ij,k} + \overline{z}_{ij,l}}{2} \tag{21.zza}$$

where $\overline{z}_{ij,k}$ is the mean of the three-way cross $(\ell_i \times \ell_j) \times \ell_k$. Such crosses are not uncommon of one has a favored single-cross lines ($\ell_i \times \ell_j$ in this case). Likewise, one can use the average of all three-way crosses,

$$\widehat{D}^{t}_{ij,kl} = \frac{\widehat{D}^{tij}_{ij,kl} + \widehat{D}^{tkl}_{ij,kl}}{2} = \frac{\overline{z}_{ij,k} + \overline{z}_{ij,l} + \overline{z}_{kl,i} + \overline{z}_{kl,j}}{4} \tag{21.zzb}$$

Eberhart showed that Jenkins A and B estimators and both his three-way cross estimates all are unbiased predictors of the double-cross mean in the absence of epistasis and genotype $\times$ environment interaction. He further showed that the estimator

$$\widehat{D}^{t-A}_{ij,kl} = 2\widehat{D}^{t}_{ij,kl} - \widehat{D}^{B}_{ij,kl} \tag{21.zzc}$$

correctly accounts for additive-by-additive epistasis, but it still biased (departs from the true double-cross mean) when dominance epistasis is present. However, Eberhart et al. (1964) found that genotype $\times$ environmental and plot error likely introduce a greater source of error than ignoring epistasis. Provided sufficient replication over environments (years and locations) and sufficient replication within each environment is performed, a correction for epistasis is not unreasonable. However, without sufficient replication it may not be worth the effort (generating the required triple-crosses), unless breeder may already have a desired single cross parental stock on hand, in which case the extra effort is minimal.

Further, while Sprague and Thomas (1967) found significant evidence for epistasis in most crosses of unselected maize lines, they still did not consider these effects to be of major importance.

Stuber et al. (1973) also conclude while epistasis generally results in underestimation of the means three- and four-way hybrids from single cross predictors, the error for most crosses is fairly insignificant. They found that single crosses (when raised in three environments) were as good a predictor of three- and four-way crosses as predictions of the performance of a three or four-way cross in one environment based on the performance of that cross in another environment.

Single- vs. Three- and Four-way Crosses

Single cross offspring different from the offspring of three- and four-way crosses in their genetical uniformity. All SC offspring are expected to be genetically identical (ignoring mutation), while the hybrid parent(s) in three- and four-way crossed generate different gametes via segregation and recombination, and hence produce genetically heterogeneous offspring. This genetic heterogeneity may result in greater environmental buffering (i.e., less extreme G $\times$ E), and potentially greater resistance to insect and plant pathogens. less drop-off in the F_2 for three- and four-way crosses (Figure 21.1). Conversely, an advantage of the right SC is that it presents favor epsistatic combinations much better than three- or four-way crosses in

which recombination in the hybrid parent(s) disrupts any favorable combinations that may have been present in the inbred lines.

Besides requiring one fewer generations to produce seeds, single crosses generally produce offspring that have higher yield than offspring from three- and four-way crosses (e.g., Sprague and Federer 1951, Rojas and Sprague 1952, Eberhart et al. 1964, Eberhart and Russell 1969). Table 21.1 summarizes results of Weatherspoon (1970), who found that for yield in maize, SC had the highest average yields, followed by 3W and DC. The lowest line, highest line, and range also followed this trend, with SC being the most extreme and DC the least extreme. Why might we expect SCs to have a better performance than 3W or DCs? With an SC, there is no recombination and blocks of favorable genes from each of the original lines are preserved in the hybrid offspring. When one of the parents is itself a SC hybrid (as is the case with 3W or DC hybrids), it will undergo recombination that can disrupt favorable combinations. Hence, one might expect a decrease in the heterotic potential as the parents undergo recombination.

Balancing this advantage of SC over 3W and Dc was concern that single cross hybrids were less stable, i.e., showed more extreme genotype-environment interactions. Consistent with this, Eberhart et al. (1964) and Eberhart and Russell (1969) reported that SC less stable (as a group) than 3Ws, DCs, although some SC were just as stable as the best DC.

Of course, there is a major difference between a random collection of SC hybrids and a select group of SC lines chosen for high performance. Lynch et al. (1973) examined both the performance and stability among SC, 3W, and DC lines recommend by the Ontario Corn Committee between 1968 and 1972. This is an elite subset of lines chosen for both yield and stability. Among this set, SC hybrids had a average yield of 3.8 q/ha over 3W hybrids and 6.3 q/ha over DCs. Lynch et al. examined stability from the slope of the Finlay-Wilkinson regression (see LW Chapter xx for details), finding no difference between the groups of hybrids. Thus, among an elite set of lines, SC hybrids are preferable, high higher performance without an apparent loss of stability compare to 3W and DC hybrid lines.

Table 21.2. Comparison of maize yield on single, three- and four-way crosses. After Weatherspoon (1970). All crosses were tested over two years in two locations (with different plant histories) in Iowa.

Cross	Average	Highest	Lowest	Range
Single	65.1	81.5	43.6	37.9
Three-way	62.0	72.9	47.7	25.2
Four-way	60.2	67.7	54.0	13.7

Comparisons for yield and other agronomically-important characters in sorghum have generally found little or no difference three-way and single crosses in either means or ranges (reviewed in Walsh and Atkins 1973).

The whole issue of sample size comparison – greater sample size, larger chance of a more extreme line. From the theory of order statistics, max = $\overline{z} + \overline{i}_{(1,n)}\sigma$ where $\overline{i}_{(1,n)}$ is the largest of the n unit normal order statistics (Ref to selection chapter).

Cockerham (1961)

F_2 vs. F_1 Performance

The recommendation to farmers that they should not plant seed from F_1 hybrid corn was greeted with considerable skepticism, often viewed as a ploy to increase the profit of hybrid seed companies. In reality, there is a sound genetic basis for not using F_2 plants. While the allele frequencies are the same in the F_1 and F_2 populations (in the absence of selection),

there is an excess of heterogosity in the F_1 relative to that expected under Hardy-Weinberg. In a diploid, Hardy-Weinberg is reached in the F_2, so that (in the absence of epistasis) the trait value in the F_3 and subsequent generations remains unchanged. If epistasis is present, then the mean can change until gametes are in linkage equilibrium. Thus some decline is expected in the F_2 as more homozygotes appear (e.g., Richey et al, 1934), while the F_3 and on should remain unchanged (e.g., Kiesselbach 1930). If the organism is a polyploid, Hardy-Weinberg is only reached asymptotically (but typically after four to six generations).

This prediction is demonstrated in Figure 21.h1, which plots the observed F_2 performance (relative to the F_1) in a number of experiment. Performance declined in all types of inbred line crosses (SC, 3W, DC) and also declined in varietial crosses (i.e., between open-pollinated parents) as well. Further note that among inbred lines, the decline was more severe when crosses involved fewer parents, with SC being most severely effected, followed by 3W and then DCs (a point we will return to shortly). Likewise, even though varietial crosses showed a reduction in the F_2's, it was (typically) not as dramatic as for the inbreds.

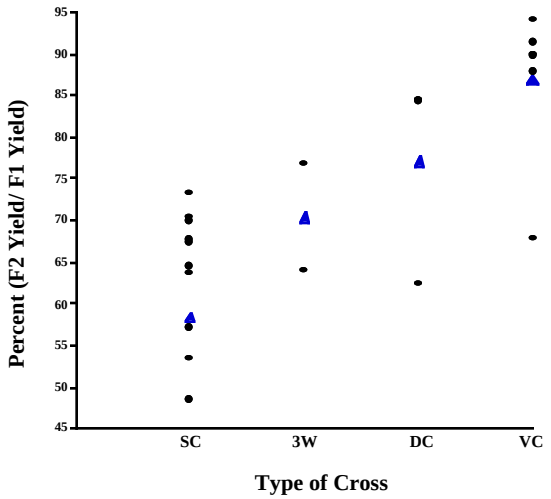


Figure 21.1. The decline from the F_1 to the F_2 in maize yield (data from experiments summarized in XX and YY Table 9.11). Results for three kinds of inbred line crosses are reported: SC = single crosses, 3W = three-way (triple) crosses, DC = double (four-way) crosses. Varietial crosses between open-pollinated parents are denoted by VC. Circles represent values from individual experiments, triangles the mean for each cross type.

As we will discussion in more detail shortly (when examining synthetics), Wright (1922) offered the a simple predictor of F_2 performance (μ_{F_2}) given F_1 performance (μ_{F_1}) and the average value μ_P of the parental lines,

$$\mu_{F_2} = \mu_{F_1} - \frac{\mu_{F_1} - \mu_P}{n} \tag{21.y1}$$

Here n is the number of parental lines: two for a single cross, three for a three-way cross, and four for a four-way (double) cross. Notice that this simple expression predicts that four-way crosses will retain a larger fraction (3/4) of the superiority of the F_1 over the parental line mean ($\mu_{F_1} - \mu_P$), with three-way crosses retaining 2/3 of this superiority, and finally single crosses will only retain 1/2 of the hybrid advantage. Wright's predictor assumes no epistasis and that parental populations are in Hardy-Weinberg proportions. A number of the

early corn breeders found Wright's predictor to be very consistent with their observations for various types of crosses (e.g., Neal 1935). For example, Figure 21.x plots the data of Kinman and Sprague (1945) the observed F_2 values for corn yield with the value predicted from Wright's formula for 45 single crosses. While there is a tendency in these data for the predicted value to overestimate the true values overall the fit is certainly acceptable.

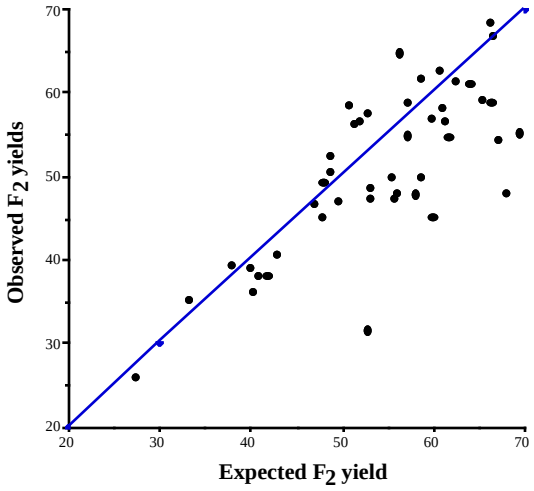


Figure 21.x. Comparison of the observed and predicted F_2 values for yield in single crosses of inbred corn. Data from Kinman and Sprague (1945), who looked at all 45 possible single crosses of 10 inbred lines. The parental lines, F_1's and F_2 were grown in the same environment (year-location).

The careful reader will note that the notion of the F_1, while making perfect sense for single crosses, is more problematic for three- and four-way crosses. In these cases, what we are calling the F_1 (the final hybrid, either $A \cdot BC$ or $AB \cdot CD$) itself involves either one or two F_1's from the parental populations. Hence, Wright's simple expression needs modification to account for three- or four-way crosses. Example 21.x develops these, and shows that under the assumption of no epistasis and parental populations in Hardy-Weinberg, then for a three-way cross,

$$\mu_{F_2 3W} = \mu_{F_1 3W} - \frac{\mu_{F_1 3W} - \mu_P}{2} + \frac{1}{8}\left(\mu_{BC} - \frac{\mu_B + \mu_C}{2}\right) \qquad (21.\text{xxa})$$

Note that parental average here is half the value of line A and one-fourth the values of each of the other two populations, as this is the relative final composition of the line. Note that the final term is positive, provided there is some heterosis in the cross $B \times C$, so that triple crosses are expected to remain more of their F_1 advantage relative to single crosses. For the four-way cross,

$$\mu_{F_2 4W} = \mu_{F_1 4W} - \frac{\mu_{F_1 4W} - \mu_P}{4} + \frac{1}{4}\left(\frac{\mu_{AB} + \mu_{CD}}{2} - \mu_{F_1 4W}\right) \qquad (21.\text{xxb})$$

while the second term has a smaller reduction (1/4 vs. 1/2) than for a single cross, the sign final term is uncertain.

Example 21.xx: Equations 21.xxa and b can be derived using an approach similar to that of Gilmore (1969) and Vencovsky (1973). The idea is that in the absence of epistasis, if P_i denotes a random allele from line i, the expected value of $P_i \times P_i$ is the mean of line i, while the expected value of $P_i \times P_j$ is the value for the single cross between lines i and j. Consider first the three-way cross, $A \cdot BC$. A random allele from the SC ($B \times C$) parent is from line B with probability $1/2$, else it is from line C. Hence, the expected value of the "F_1", the three-way cross, is

$$\mu_{A \cdot BC} = \mu_{F_1 3W} = \frac{\mu_{AB} + \mu_{AC}}{2}$$

For the "F_2", each parent is $A \cdot BC$, so that half of the alleles are from line A, and $1/4$ each from lines B and C. Hence, in the offspring from these parents (the F_2), the expected contributions are

$$\frac{P_A P_A}{4} + \frac{P_A P_B}{8} + \frac{P_A P_C}{8} + \frac{P_B P_A}{8} + \frac{P_B P_B}{16} + \frac{P_B P_C}{16} + \frac{P_C P_A}{8} + \frac{P_C P_B}{16} + \frac{P_C P_C}{16}$$

Collecting terms,

$$\mu_{F_2 3W} = \frac{\mu_A}{4} + \frac{\mu_B}{16} + \frac{\mu_C}{16} + \frac{\mu_{AB}}{4} + \frac{\mu_{AC}}{4} + \frac{\mu_{BC}}{8}$$

Letting $\mu_P = (2\mu_A + \mu_B + \mu_C)/4$ be the weighted average of the parent populations and recalling the expected value of μ_{F_1} from above, this reduces to:

$$\begin{aligned}
\mu_{F_2 3W} &= \frac{\mu_P}{4} + \frac{\mu_{F_1 3W}}{2} + \frac{\mu_A + \mu_{BC}}{8} \\
&= \mu_{F_1 3W} - \frac{\mu_{F_1 3W} - \mu_P}{2} + \left(\frac{\mu_A}{8} + \frac{\mu_{BC}}{8} - \frac{\mu_P}{4} \right) \\
&= \mu_{F_1 3W} - \frac{\mu_{F_1 3W} - \mu_P}{2} + \frac{1}{8} \left(\mu_{BC} - \frac{\mu_B + \mu_C}{2} \right)
\end{aligned}$$

Turning to the four-way cross, $AB \cdot CD$, the expected "F_1" mean (again in the absence of epistasis) is

$$\mu_{F_1 4W} = \frac{P_A P_C}{4} + \frac{P_A P_D}{4} + \frac{P_B P_C}{4} + \frac{P_B P_D}{4} = \frac{\mu_{AC} + \mu_{AD} + \mu_{BC} + \mu_{CD}}{4}$$

With these as the F_1 parents, a random allele has a 25% of being derived from each of the four base lines. Hence, in the F_2, the contributions are

$$\begin{aligned}
\mu_{F_2 4W} &= \frac{\mu_A + \mu_B + \mu_C + \mu_D}{16} + \frac{\mu_{AB} + \mu_{AC} + \mu_{AD} + \mu_{BC} + \mu_{CD} + \mu_{CD}}{8} \\
&= \frac{\mu_P}{4} + \frac{\mu_{F_1 4W}}{2} + \frac{\mu_{AB} + \mu_{CD}}{8} \\
&= \mu_{F_1 4W} - \frac{\mu_{F_1 4W} - \mu_P}{4} + \frac{1}{4} \left(\frac{\mu_{AB} + \mu_{CD}}{2} - \mu_{F_1 4W} \right)
\end{aligned}$$

ROTATIONAL CROSSBREEDING

Example 21.2. Consider a three-bred rotational crossbreding scheme where dams from the previous generation are crossed pure-bred sires in a rotating seqeunce (line A in one genration, b in the next, C in the third and so on). Under this scheme, what fraction of lines from each of the lines are present in any particular generation? The logic is straightforward: in each generation, half of the contribution from the previous dam is passed on, as is half the genes from the sire line. Thus,

		Percentage of lines:		
Generation	Cross	A	B	C
1	$A \times B$	50.0	50.0	0.0
2	$C\times$ gen 1 dam	25.0	25.0	50.0
3	$A\times$ gen 2 dam	62.5	12.5	25.0
4	$B\times$ gen 3 dam	31.3	56.3	12.5
5	$C\times$ gen 4 dam	15.6	28.1	56.3
6	$A\times$ gen 5 dam	57.8	14.1	28.1
7	$B\times$ gen 6 dam	28.9	57.0	14.1
8	$C\times$ gen 7 dam	14.5	28.5	57.0

The asympotical contributions reached are 57.1% for the sire line, 28.6% for the sire line used in the previous generation, and 14.3% for the sire line used two generations previous.

In a similar fashion, for a four breed (A, B, C, D) rotational scheme,

		Percentage of lines:			
Generation	Cross	A	B	C	D
1	$A \times B$	50.0	50.0	0.0	0.0
2	$C\times$ gen 1 dam	25.0	25.0	50.0	0.0
3	$D\times$ gen 2 dam	12.5	12.5	25.0	50.0
4	$A\times$ gen 3 dam	56.3	6.3	12.5	25.0
5	$B\times$ gen 4 dam	28.1	53.1	6.3	12.5
6	$C\times$ gen 5 dam	14.1	26.6	53.1	6.3
7	$D\times$ gen 6 dam	7.0	13.3	26.6	53.1
8	$A\times$ gen 7 dam	53.5	6.6	13.3	26.6

At equilibrium the line contributions cycle among 53.3%, 26.7%, 13.3%, and 6.7%.

Following the logic in Example 21.2, the expected fraction of genetic contributions at equilibrium from each of the k lines where crossbred dams are kept and crossed in rotation to purebreed sires are of the form $(1/2) * I, (1/2)^2 * I, \cdots, (1/2)^k * I$ where

$$I = \sum_{i=0}^{\infty} \left(\frac{1}{2^k}\right)^i = 1 + \frac{1}{2^k - 1} \tag{21.xx}$$

with the largest fraction ($I/2$) from the line last used as a sire and the smallest fraction ($I/2^k$) for the line to be next used as a sire.

Winters et al. (1935)

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

$$\widehat{R}_2 = \overline{z}_{AB} - \frac{\overline{z}_{AB} - \overline{P}_2}{3}, \quad \text{where} \quad \overline{P}_2 = \frac{\overline{z}_A + \overline{z}_B}{2} \tag{21.xxa}$$

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

$$\widehat{R}_3 = \overline{SC}_3 - \frac{\overline{z}_{AB} - \overline{P}_3}{7}, \quad \text{where} \quad \overline{SC}_3 = \frac{\overline{z}_{AB} + \overline{z}_{AC} + \overline{z}_{BC}}{3} \tag{21.xxb}$$

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. At equilibrium, contribution of lines is 57.14, 28.57, and 14.29 percent

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

$$\widehat{R}_4^{(A,B,C,D)} = \overline{SC}_4 - \frac{\overline{SC}_{na} - \overline{P_4}}{15}, \quad \text{where} \quad \overline{SC}_{na} = \frac{\overline{z}_{AC} + \overline{z}_{BD}}{2} \tag{21.xxb}$$

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.

Carmon et al. (1956) conjecture that for an n-line rotation, that an unbiased predictor of performance is of the form

$$\widehat{R}_n = \overline{SC}_n - \frac{\overline{SC}_* - \overline{P}_n}{2^n - 1} \tag{21.xx}$$

where $\overline{SC}_* < \overline{SC}_n$ for $n > 3$. Comparing this to Wright's (1922) expression for S_n, the average performance of a synthetics from n lines (i.e., random mating as opposed to continued rotational crossbreeding among the lines), that $R_n > S_n$, as

$$\left(\overline{SC}_n - \frac{\overline{SC}_* - \overline{P}_n}{2^n - 1}\right) > \left(\overline{SC}_n - \frac{\overline{SC}_n - \overline{P}_n}{n}\right) = \widehat{S}_n \tag{21.xx}$$

Example 21.x. 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, synthetic S, and the backcross (BC)

	Means				
Trait	$\overline{P}$	F_1	R	S	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

How well does Equation 21.xx 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.

Example 21.x. As illustrated in Example 21.xx , it would be nice to have standard errors for the predicted values. Here we illustrate how this is done for the two-bred predictor, but the basic approaches easily extends to most of the other predictors developed in this chapter. Rearranging to collect common terms,

$$\widehat{R}_2 = \overline{z}_{AB} - \frac{\overline{z}_{AB} - \overline{P}_2}{3} = \left(1 - \frac{1}{3}\right)\overline{z}_{AB} + \left(\frac{1}{6}\right)\left(\,\overline{z}_A + \overline{z}_B\,\right)$$

Since the estimates are independent and recalling that $\sigma^2(ax) = a^2\sigma^2(x)$ for a constant a, it immediately follows that

$$\sigma^2\left(\,\widehat{R}_2\,\right) = \left(1 - \frac{1}{3}\right)^2 \sigma^2\left(\,\overline{z}_{AB}\,\right) + \left(\frac{1}{6}\right)^2 \left(\,\sigma^2\left(\,\overline{z}_A\,\right) + \sigma^2\left(\,\overline{z}_B\,\right)\,\right)$$

Table 21.mex. 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 under rotational crossbreeding. The weighting coefficients given here are the weights for the average of that particular component over all the breeds. For example, with a three-bred rotational, the average individual heterosis (averaged over cycles) is $6/7\,\overline{h}^I$, where $\overline{h}^I$ is the average individual heterosis for the three crosses. After Dickerson (1969).

No of Breeds	g^I	h^I	r^I	g^M	h^M	r^M	$g^{M'}$	$h^{M'}$	$r^{M'}$
2 breeds	1	$\frac{2}{3}$	$\frac{1}{3}$	1	$\frac{2}{3}$	$\frac{1}{3}$	1	$\frac{2}{3}$	$\frac{1}{3}$
3 breeds	1	$\frac{6}{7}$	$\frac{3}{7}$	1	$\frac{6}{7}$	$\frac{3}{7}$	1	$\frac{6}{7}$	$\frac{3}{7}$
4 breeds	1	$\frac{14}{15}$	$\frac{7}{15}$	1	$\frac{14}{15}$	$\frac{7}{15}$	1	$\frac{14}{15}$	$\frac{7}{15}$

Literature Cited

Anderson, D. C. 1938. The relation betwen single and double cross yields in corn. *J. American Society of Agronomy* 30: 209–211. [21]

Bailey, T. B. 1980. Selection limits in self-fertilizing populations following the cross of homozygous lines. *In* E. Pollak, O. Kempthorne, and T. B. Bailey, Jr. (eds.), *Proceedings of the international conference on quantitative genetics*, pp. 399–412. Iowa State Univ. Press, Ames. [11, 12]

Bernardo, R. 1990a. An alternative statistic for identifying lines useful for improving parents of an elite single cross. *Theor. Appl. Gene.* 80: 105–109. [21]

Bernardo, R. 1990b. Methods of choosing parents in a peidgree breeding program given an elite inbred tester. *Maydica* 35: 267–271. [21]

Bowman, J. C. 1960. Recurrent selection. I. The detection of overdominance. *Heredity* 14: 197–206. [21]

Bradford, G. E., D. T. Torell, G. M. Spurlock, and W. C. Weir. 1963. Performance and variability of offpsring of crossbred and purebred rams. *J. Animal Science* 22: 617–623. [21]

Busbice, T. H. 1969. Inbreeding in synthetic varieties. *Crop Science* 9: 601–604. [21]

Busbice, T. H. 1970. Predicting yield of synthetic varieties. *Crop Sci.* 10: 265–269.

Busbice, T. H., and C. P. Wilsie. 1966. Inbreeding depression and heterosis in an autotetraploids with application to *Medicago sativa* L. *Euphytica* 15: 52–67. [21]

Carmon, J. L., H. A. Stewart, C. C. Cockerham, and R. E. Comstock. 1956. Prediction equations for rotational crossbreeding. *J. Animal Sciences* 15: 930–936. [21]

Clutton-Brock, J. 1992. *Horse power*. Harvard University Press, Cambirdge, MA.

Cockerham, C. C. 1961. Implications of genetic variances in a hybrid breeding program. *Crop Sci.* **1**: 47–52. [21]

Cockerham, C. C. 1967. Predicion of double crosses from single crosses. *der Zücher* (*Theoretical and Applied Genetics*) 37: 160 –169. [21]

Collins, G. N. 1921. Dominance and the vigor of first generation hybrids. *American Naturalist* 55: 116–133. [21]

Comstock, R. E., H. F. Robinson, and P. H. Harvey. 1949. A breeding program designed to make maximum use of both general and specific combining ability. *Agronomy J.* 41: 360–367. [21]

Coors, J. G., and S. Pandey (Eds). 1999. *The genetics and exploitation of heterosis in crops*. American Society of Agrnomony, Madison, Wisconsin. [21]

Cress, C. E. 1966. Heterosis of the hybrid related to gene frequency differences between two populations. *Genetics* 53: 269–274. [21]

Cundiff, L. V., K. E. Gregory, and R. M. Koch. 1974a. Effects of heterosis on reproduction in Hereford, Angus, and Shorthorn cattle. *J. Animal Sci.* 38: 711–727. [21]

Cundiff, L. V., K. E. Gregory, F. J. Schwulst, and R. M. Koch. 1974b. Effects of heterosis on maternal performance and milk production in Hereford, Angus, and Shorthorn cattle. *J. Animal Sci.* 38: 728–xx. [21]

Davis, R. L. 1929. Report of the Plant Breeder. *Annual Report of Puerto Rice Experimental Station for 1927*, Puerto Rice Experimental Station, Rio Piedras, pp. 14-15. [21]

Demarly, Y. 1963. Genetique des tetaploides et amelioration des plantes. *Ann. Amelio. Plant*. 13: 307–400. [21]

Dessureaux, L., and A. Gallais. 1969. Inbreeding in heterosis in autotetraploid alfalfa. I. Fertility. *Can. J. Genet. Cytol.* 11: 706–715. [21]

Dobbs, K. S. 1955. Hybrid vigor in plant breeding. *Proced. Roy. Soc. London B* 144: 185–192. [21]

Doxtator, C. W., and I. J. Johnson. 1936. Prediction of double cross yields in corn. *J. American Society of Agronomy* 28: 460–462. [21]

Dudley, J. W. 1964. A genetic evaluation of methods of utilizing heterozygosis and dominance in autotetraploids. *Crop Science* 4: 410–413. [21]

Dudley, J. W. 1982. Theory of transfer of alleles. *Crop Sci.* 22: 631–637. [21]

Dudley, J. W. 1984a. A method of identifying lines for use in improving parents of a single cross. *Crop Sci.* 24: 355-357. [21]

Dudley, J. W. 1984b. A method of identifying populations containing favorable alleles not present in elite germplasm. *Crop Sci.* 24: 1053–1054. [21]

Dudley, J. W. 1984c. Theory for identification and use of exotice germplasm in maize breeding programs. *Maydica* 29: 391–407.

Dudley, J. W. 1987. Modification of methods for identifying populations to be used for improving parents of elite single crosses. *Crop Sci.* 27: 940–943. [21]

Dudley, J. W. 1987. Modification of methods for identifying inbred lines useful for improving parents of elite single crosses. *Crop Sci.* 27: 944–947. [21]

Dunbier, M. W., and E. T. Bingham. 1975. Maximum heterozygosity in alfalfa: Results using haploid-derived autotetraploids. *Crop Science* 15: 527 –531. [21]

Duvick, D. N. 1999. Heterosis: feeding people and protecting natural resources. *In* J. G. Coors and S. Pandey (eds). *The genetics and exploitation of heterosis in crops*, pp. 19–29. American Society of Agrnomony, Madison, Wisconsin. [21]

Eberhart, S. A. 1964. Theoretical relations among single, three-way and double cross hybrids. *Biometrics* 20: 522–539. [21]

Eberhart, S. A., and C. O. Gardner. 1966. A general model for genetic effects. *Biometrics* 22: 854–881. [21]

Eberhart, S. A., and A. R. Hallauer. 1968. Genetic effects for yield in single, three-way, and double-cross maize hybrids. *Crop Science* 8: 377 – 379. [21]

Eberhart, S. A., and W. A. Russell. 1969. Yield and stability for a 10-line diallel of single-cross and double-cross maize hybrids. *Crop Science* 9: 357 –361. [21]

Eberhart, S. A., W. A. Russell, and L. H. Penny. 1964. Double cross hybrid prediction in maize when epistasis is present. *Crop Science* 4: 363 –366. [21]

Gallais, A. 1967. Moyennes des populations tétraploids. *Ann. Amélior. Plantes* 17: 215–228. [21]

Gallais, A. 1979. The concept of varietal ability in plant breeding. *Euphytica* 28: 811-823. [21]

Gallais, A. 1990. Applications of the concepts of test value and varietal value to the study of genetic advance in recurrent selection: A synthesis. *Euphytica* 48: 197–209. [21]

Gardner, C. O., and S. A. Eberhart. 1966. Analysis and interpretation of the variety cross diallel and related populations. *Biometrics* 22: 439–452. [21]

Gerloff, J. E., and O. S. Smith. 1988a. Choice of method for identifying germplasm with superior alleles. I. Theoretical results. *Theor. Appl. Gene.* 76: 209–216. [21]

Gerloff, J. E., and O. S. Smith. 1988b. Choice of method for identifying germplasm with superior alleles. I. Computer simulation results. *Theor. Appl. Gene.* 76: 217–227. [21]

Grafius, J. E. 1959. Heterosis in barley. *Agron. J.* 51: 551–554. [21]

Gregory, K. E., J. C. M. Trail, H. J. S. Marples, and J. Kakonge. 1985. Heterosis and breed effects on maternal and individual traits of Bos Indicus breeds of cattle. *J. Animal Sci.* 60: 1175 –1180. [21]

Goodnight, C. J. 1999. Epistatis and heterosis. *In* J. G. Coors and S. Pandey (eds). *The genetics and exploitation of heterosis in crops*, pp. 59–68. American Society of Agrnomony, Madison, Wisconsin. [21]

Gowen, J. W. (ed.), *Heterosis,* Iowa State College Press, Ames.

Hayes, H. K. and R. J. Garber. 1919. Synthetic production of high-protein corn in relation to breeding. *J. American Society of Agronomy* 11: 309–318. [21]

Hayes, H. K., R. P. Murphy, and E. H. Rinke. 1943. A comparison of the actual yield of double crosses of maize with their predicted yield from single crosses. *J. American Society of Agronomy* 35: 60–65. [21]

Hayes, H. K., E. H. Rinke, and Y. S. Tsiang. 1946. The relationship between predicted performance of double crosses of corn in one year with predicted and actual performance of double corsses in later years. *J. American Society of Agronomy* 38: 60–67. [21]

Hill, R. R., Jr. 1971. Effect of the number of parents on the mean and variance of synthetic varieties. *Crop Science* 11: 283– 286. [21]

Ho, Y. T., and R. E. Comstock. 1980. Combining superior alleles form two homozygous populations in a cross-fertilizing species. *Genet. Res.* **XX**: 277–287. [21]

Hohls, T., G. P. Y. Clarke, P. E. Shanahan, and H. O. Gevers. 1995. An unbiased estimator foridentifying lines useful for the improvement of elite single crosses, based on a combining ability model. *Heredity* 75: 155–163. [21]

Hull,F. H. 1945. Recurrent selection for specific combining ability in corn. *J. Amer. Sco. Agronomy* 45: 989–1009. [21]

Jenkins, M. T. 1934. Methods of estimating the performance of double crosses in corn. *J. American Agrn. Society* 26: 199 – 204. [21]

Jenkins, M. T. 1935. The effect of inbreeding and selection within inbred liens of maize upon nybrids made after successive generations of selfing. *Iowa State College J. of Science* 9: 429–450. [21]

Jenkins, M. T., and A. M. Brunson. 1932. Methods of testing inbred lines of maize in crossbred combinations. *J. American Agrn. Society* 24: 523 –530. [21]

Jones, D. F. 1917. Dominance of linked factors as means of accounting for heterosis. *Genetics* 2: 466–479. [21]

Jones, D. F. 1918. The effect of inbreeding and crossbreeding upon development. Connecticut Ag. Station Bull. 207. [21]

Jones, D. F. 1922. The production of single and double first-generation corn hyrbids. *J. American Agrn. Society* 14: 241–252. [21]

Kidder, R. W.,M. Koger, J. H. Meade, and J. R. Crockett. 1964. Systems of crossbreeding for beef production in Florida. *Bull. Agric. Experimental Stations, Florida University* No. 673, 19 pp. [21]

Kinman, M. L., and G. F. Sprague. 1945. Relation between number of parental lines and theoretical performance of synthetic variaties of corn. *J. American Agrn. Society* 37: 341 –351. [21]

Kiesselbach, T. A. 1930. The use of advanced generation hyrbids as parents of double cross seed corn. *J. American Agrn. Society* 22: 614–626. [21]

Kiesselbach, T. A. 1951. (**CHECK SPELLING**) A half-century of corn research. Amer Sci. 39: 629–655.

Koger, M., F. M. Peacock, W. G. kirk, and J. R. Crockett. 1975. Heterosis effects on weaning performance of Brahman-Shorthorn calves. *J. Animal Sci.* 40: 826–833. [21]

Lamkey, K. R., and J. W. Edwards. 1999. Quantitative genetics and heterosis. *In* J. G. Coors and S. Pandey (eds). *The genetics and exploitation of heterosis in crops*, pp. 31–48. American Society of Agrnomony, Madison, Wisconsin. [21]

Lonnquist, J. H. 1949. The development and performance of synthetic varieties of corn. Agronomy Journal 41: 153–156. [21]

Lynch, P. J., R. B. Hunter, and L. W. Kannenberg. 1973. Relative preformance of single cross, three-way cross, and double-cross corn hybrids recommend in Ontario, 1968-72. *Can. J. Plant Sci.* 53: 805-810. [21]

McDowell, R. E. 1981. Limitations for dairy production in developing countries. *J. Dairy Sci.* 64: 2463–XXXX. [21]

McDowell, R. E. 1985. Crossbreeding in tropical areas with emphasis on milk, heath, and fitness. *J. Dairy Sci.* 68: 2418–2435. [21]

Mišević, D. 1989a. Evaluation of three test statistics used to identify maize inbred lines with new favorable alleles no present in elite single cross. *Theor. Appl. Gene.* 77: 402–408. [21]

Mišević, D. 1989b. Identification of inbred lines as a source of new alleles from improvement of elite maize single crosses. *Crop Sci.* 29: 1120–1125. [21]

Neal, N. P. 1935. The decrease in yielding capacity in advanced generations of hybrid corn. *J. American Agrn. Society* 27: 666–670. [21]

Nitter, G. 1978. Breeding utilisation for meat production in sheep. *Animal Breeding Abstracts* 46: 131–143. [21]

Otsuka, Y., S. A. Eberhart, and W. A. Russell. 1972. Comparisons of prediction formulas for maize hybrids. *Crop Science* 12: 325 – 331. [21]

Pfarr, D. G., and K. R. Lamkey. 1992a. Evaluation of theory for identifying populations for genetic improvement of maize hybrids. *Crop Sci.* 32: 663-669. [21]

Pfarr, D. G., and K. R. Lamkey. 1992b. Comparison of methods for identifying populations for genetic improvement of maize hybrids. *Crop Sci.* 32: 670–677. [21]

Powers, L. 1941. Inheritance of quantitative charaacters in crosses of two species of *Lycopersicon*. *J. Agr. Res.* 63: 149–174. [21]

Rawlings, J. O., and C. C. Cockerham. 1962a. Triallel Analysis. *Crop Science* 2: 228 – 231. [21]

Rawlings, J. O., and C. C. Cockerham. 1962b. Analysis of double cross hybrid populations. *Biometrics* 18: 229 – 244. [21]

Reddy, B. V. S., and R. E. Comstock. 1976. Simulation of the backcross breeding method. I. The effects of heritability and gene number of fixation of desired alleles. *Crop Sci.* 16: 825–830. [21]

Richey, F. D. 1966. Moch-dominance and hybrid vigor. *Science* 96: 280-281. [21]

Richey, F. D., G. H. Stringfield, and G. F. Sprague. 1934. The loss in yield that may be expected from planting second generation doulbe-crossed seed corn. *J. American Society of Agronomy* 26: 196–199. [21]

Robsin, O. W., B. T. McDaniel, and E. J. Rincon. 1981. Estimation of direct and maternal additive and heterotic effects from crossbreeding experiments in animals. *J. Animal Science* 52: 44–50. [21]

Rogers, J. S., and J. R. Edwardson. 1952. The utilization of cytoplasmic male-sterile hybrids in the production of corn hybrids. *Agronomy Journal* 44: 8-13. [21]

Rojas, B. A., and G. F. Sprague. 1952. A comparison of variance components in yield trails. III. General and specific combining ability and their interaction with location and years. *Agronomy Journal* 44: 462–466. [21]

Sang, J. H. 1956. Hybrid vigour and animal production. *Animal Breeding Abstracts* 24: 1–6. [21]

Setnz, J. C., H. F. Robinson, and R. E. Comstock. 1954. Relation between heterozygosis and performance in maize. *Agronomy Journal* 46: 514–520. [21]

Schell, F. W., and C. C. Cokcerham. 1992. Multiplicative vs. arbitrary gene acion in heterosis. *Genetics* 131: 461–469. [21]

Sheridan, A. K. 1981. Crossbreeding and heterosis. *Animal Breeding Abstracts* 49: 131–144. [21]

Shull, G. H. 1908. The composition of a field of maize. *American Breeders' Association* 4: 296–301.

Shull, G. H. 1909. A pure-line method in corn breeding. *American Breeders' Association* 5: 51-59.

Sprague, G. F. 1946. Early testing of inbred lines of corn. *J. American Society of Agronomy* 38: 108–117. [21]

Sprague, G. F., and W. Federer. 1951. A comparison of variance components in corn yield traits. II. Error, year × variety, location × variety and variety components. *Agronomy Journal* 43: 535-541. [21]

Sprague, G. F., and M. T. Jenkins. 1943. A comparison of synthetic varieties, multiple crosses, and double crosses in corn. *J. American Society of Agronomy* 35: 137–147. [21]

Sprague, G. F., and L. A. Tatum. 1942. General vs. specific combining-ability in single crosses of corn. *J. American Society of Agronomy* 34: 923–932. [21]

Sprague, G. F., and W. I. Thomas. 1967. Further evidence of epistasis in single and three-way cross yields in maize (*Zea mays* L.). *Crop Science* 7: 355-356. [21]

Stephens, J. C., and R. F. Holland. 1954. Cytoplasmic male-sterility for hybrid sorghum seed production. *Agron. J.* 46: 20-23. [21]

Stingfield, G. H. 1950. Heterozygosis and hybrid vigor in maize. *Agronomy Journal* 42: 143–152. [21]

Stuber, C. W., W. P. Williams, and R. H. Moll. 1973. Epistasis in Maize (*Zea mays* L.): III. Signifiance in predictions of hybrid performances. *Crop Science* 13: 195-200. [21]

Tatum, L. A. 1971. The Southern Corn Leaf Blight epidemic. *Science* 171: 1113-1116. [21]

Trail, J. C. M., K. E. Gregory, H. J. S. Marples, and J. Kakong. 1985. Comparison of Bos Tarus-Bos Indicue breed croses with straightbred Bos Indicus breeds of cattle for maternal and individual traits. *J. Animal Sci.* 60: 1181- 1187. [21]

Troyer, A. F. 1999. Background of U. S. Hybrid Corn. *Crop Science* 39: 601–626. [21]

Ullstrup, A. J. 1972. The impacts of the southern corn leaf blight epidemics of 1970-71. *Annual Review of Phytopathology* 10: 37-50. [21]

Walsh, E. J., and R. E. Atkins. 1973. Performance and within-hybrid variability of three-way and single crosses of grain sorghum. *Crop Science* 13: 267–271. [21]

Weatherspoon, J. H. 1970. Comparative yields of single, three-way, and double crosses of maize. *Crop Science* 10: 157 - 159. [21]

Willham, R. I. 1970. Genetic consequences of crossbreeding. *J. Animal Science* 30: 690 –693. [21]

Willham, R. L., and E. Pollak. 1985. Theory of heterosis. *J. Dairy Science* 68: 2411–2417. [21]

Williams, W. 1959. Heterosis and the genetics of complex characters. *Nature* 184: 527–530. [21]

Winters, L. M., O. M. Kiser, P. S. Jordan, and W. H. Peters. 1935. A six years study of crossbreeding swine. *Minn. Agr. Exp. Stat. Bull.* 320. [21]

Winters, L. M. 1952. Rotational crossbreeding and heterosis. *In* J. W. Gowen (ed.), *Heterosis,* pp. 371–377. Iowa State College Press, Ames. [21]

Wright, S. 1922. The effects of inbreeding and crossbreeding on guinea pigs. U. S. D.A. Bull. 1121. [21]

Zambezi, B. T., E. S. Horner, and F. G. Martin. 1986. Inbred lines as testors for general combining ability in maize. *Crop Science* 26: 908–910. [21]

Zanoni, U. and J. W. Dudley. 1989a. Comparison of different methods of identifying inbreds useful for improving elite maize hybrids. *Crop Sci.* 29: 577–582. [21]

Zanoni, U. and J. W. Dudley. 1989b. Use of F_2 generations to identify inbreds with favortable alleles not present in an elite hybrid. *Crop Sci.* 29: 583–589. [21]

Zanoni, U. and J. W. Dudley. 1989c. Testcross evaluation of F_2 populations form maize inbreds selected for unique favorable alleles. *Crop Sci.* 29: 589–595. [21]

to Incorporate

Patel, R. M., C. C. Cockerham, and J. O. Rawlings. 1969. Selection among diallel classifed variables. *Biometrics* 25: 49–61. [21]

McNew, R. W., and A. E. Bell. 1971. The nature of the purebred-crossbred genetic covariance. *Genet. Res.* 18: 1-7. [21]

McNew, R. W., and A. E. Bell. 1976. Comparison of crossbred nad purebred selection for a heterotic trait in highly selected populations of *Tribolium*. *J. Heredity* 67: 275–283. [21]

Sölkner, J. 1991. The impact of different genetic models on the optimal design of crossbreeding experiments. *Anim. Prod.* 52: 255–262. [21]

Regfe, J. E. O., G. S. Aboagye, and S. Akah. 1993. A method for estimating between-breed differences in crossbreeding parameters and their sampling variances. *Livestock Production Science* 33: 327–340. [21]

Wei, M., and J. H. J. van der Werf. 1994. Maximizing genetic response in crossbreds using both purebred and crossbred information. *Anim. Prod.* 59: 401–413. [21]

==========

Kramer, H. H., and A. J. Ullstrup. 1959. Preliminary evaluations of exotice maize germplasm. *Agron. J.* 51: 687–689. [XX]

Bradford, G. E., D. T. Torell, G. M. Spurlock, adn W. C. Weir. 1963. Performance and variability of offpsring of crossbred and purebred rams. *J. Animal Science* 22: 617–623. [21]

Magee, W. T., and L. N. Hazel. 1959. General and specific combining ability for 154-day weight among three-line crosses in swine. *J. Animal Sciences* 18: 790–795. [21]

Sheridan, A. K. 1981. Crossbreeding and heterosis. *Animal Breeding Abstracts* 49: 131–144. [21]

Sidwell, G. M., D. O. Everson, and C. E. Terrill. 1964. Lamb weights in some pure breeds and crosses. *J. Animal Science* 23: 105–110. [21]

Sorrells, M. E., and W. A. Wlson. 1997. Direct classification and selection of superior alleles for crop improvement. *Crop Sci.* 37: 691–697. [21]

Wei, M., and J. H. J. van der Werf. 1994. Maximizing genetic response in crossbreds using both purebred and crossbred information. *Anim. Prod.* 59: 401–413. [21]

to Get

Baker, L. H., and R. N. Curnow. 1969. Choice of population size and use of variation between replicate populations in plant breeding selection programs. *Crop Sci.* 9: 555-560. [TO GET]

Kenworthy, W. J., and C. A. Brim. 1979. Recurrent selection in soybeans. I. seed yield. *Crop Sci.* 19: 315–318. [TO GET]

Swanson, M. R., J. W. Dudly, and S. G. Carmer. 1974. Simulation studies in autotetraploid populations. II. Effects of double reduction, population size, and selection intensity. *Crop Sci.* 14: 630–636. [TO GET]

NOT USED

Sprague, G. F., W. A. Russell, L. H. Penny, T. W. Horner, and W. D. Hanson. 1962. Effect of epistasis on grain yield in maize . *Crop Science* 2: 205-208. [21]