

Lecture 7

Inbreeding and Crossbreeding

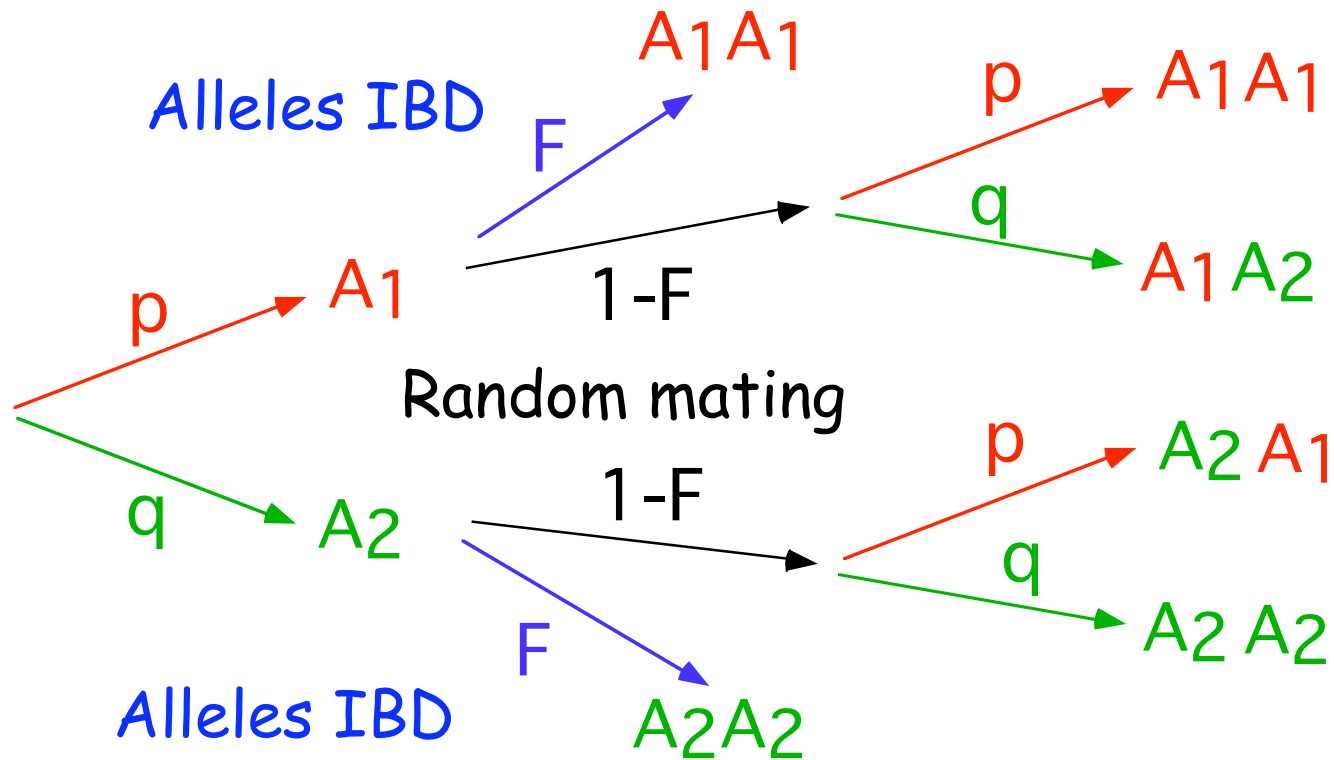
Bruce Walsh lecture notes
Synbreed course
version 3 July 2013

Inbreeding

- Inbreeding = mating of related individuals
- Often results in a change in the mean of a trait
- Inbreeding is intentionally practiced to:
 - create genetic uniformity of laboratory stocks
 - produce stocks for crossing (animal and plant breeding)
- Inbreeding is unintentionally generated:
 - by keeping small populations (such as is found at zoos)
 - during selection

Genotype frequencies under inbreeding

- The inbreeding coefficient, F
 - We (and the literature) interchangeably use f as well
- $F = \text{Prob}(\text{the two alleles within an individual are IBD})$ -- identical by descent
- Hence, with probability F both alleles in an individual are identical, and hence a homozygote
- With probability $1-F$, the alleles are combined at random



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

Changes in the mean under inbreeding

Genotypes	A_1A_1	A_1A_2	A_2A_2
	0	$a + d$	$2a$

$$\text{freq}(A_1) = p, \quad \text{freq}(A_2) = q$$

Using the genotypic frequencies under inbreeding, the population mean μ_F under a level of inbreeding F is related to the mean μ_0 under random mating by

$$\mu_F = \mu_0 - 2Fpqd$$

For k loci, the change in mean is

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

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

- There will be a change of mean value dominance is present (d not zero)
- For a single locus, if $d > 0$, inbreeding will decrease the mean value of the trait. If $d < 0$, inbreeding will increase the mean
- For multiple loci, a decrease (**inbreeding depression**) requires **directional dominance** --- dominance effects d_i tending to be positive.
- The magnitude of the change of mean on inbreeding depends on gene frequency, and is greatest when $p = q = 0.5$

Inbreeding Depression and Fitness traits



Define $ID = 1 - \mu_F / \mu_0 = 1 - (\mu_0 - B) / \mu_0 = B / \mu_0$

Drosophila Trait	Lab-measured ID = B / μ_0
Viability	0.442 (0.66, 0.57, 0.48, 0.44, 0.06)
Female fertility	0.417 (0.81, 0.35, 0.18)
Female reproductive rate	0.603 (0.96, 0.57, 0.56, 0.32)
Male mating ability	0.773 (0.92, 0.76, 0.52)
Competitive ability	0.905 (0.97, 0.84)
Male fertility	0.11 (0.22, 0)
Male longevity	0.18
Male weight	0.085 (0.1, 0.07)
Female weight	-0.10
Abdominal bristles	0.077 (0.06, 0.05, 0)
Sternopleural bristles	-.005 (-0.001, 0)
Wing length	0.02 (0.03, 0.01)
Thorax length	0.02

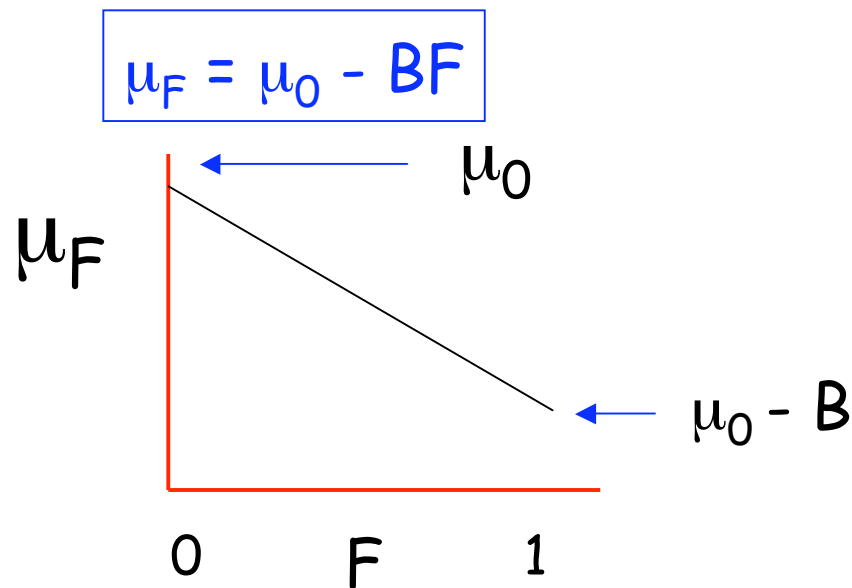
Why do traits associated with fitness show inbreeding depression?

- Two competing hypotheses:
 - **Overdominance:** Genetic variance for fitness is caused by loci at which heterozygotes are more fit than both homozygotes. Inbreeding decreases the frequency of heterozygotes, increases the frequency of homozygotes, so fitness is reduced.
 - **Dominance:** Genetic variance for fitness is caused by rare deleterious alleles that are recessive or partly recessive; such alleles persist in populations because of recurrent mutation. Most copies of deleterious alleles in the base population are in heterozygotes. Inbreeding increases the frequency of deleterious homozygotes, reducing fitness.

Estimating B

In many cases, lines cannot be completely inbred due to either time constraints and/or because in many species lines near complete inbreeding are nonviable

In such cases, estimate B from the regression of μ_F on F,



If epistasis is present, this regression is non-linear, with $C_k F^k$ for k-th order epistasis

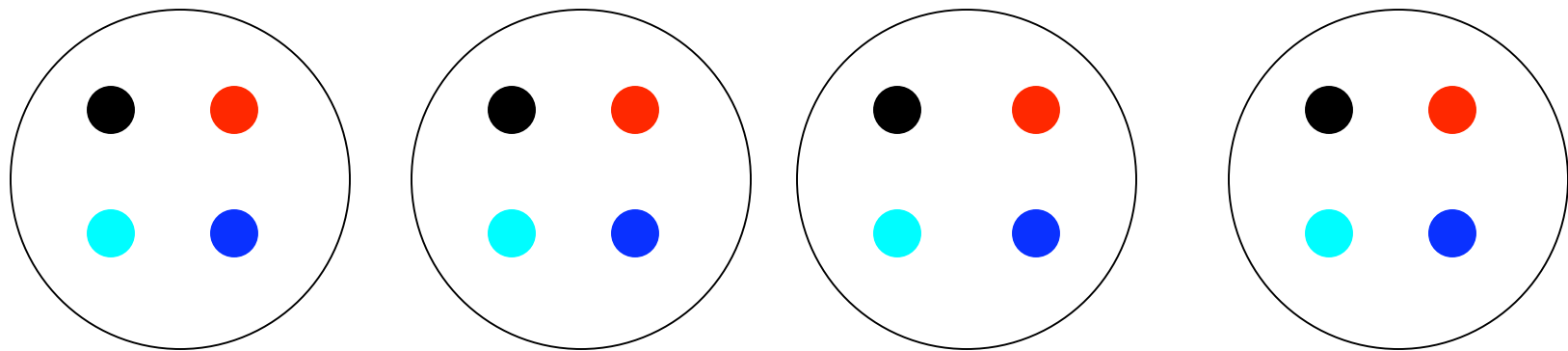
Minimizing the Rate of Inbreeding

- Avoid mating of relatives
- Maximize effective population size N_e
- N_e maximized with **equal representation**
 - N_e decreases as the variance of contributed offspring increases
 - Contribution (number of sibs) from each parent as equal as possible
 - Sex ratio as close to 1:1 as possible
 - When sex ratio skewed (r dams/sires), every male should contribute (exactly) one son and r daughters, while every female should leave one daughter and also with probability $1/r$ contribute a son

Variance Changes Under Inbreeding

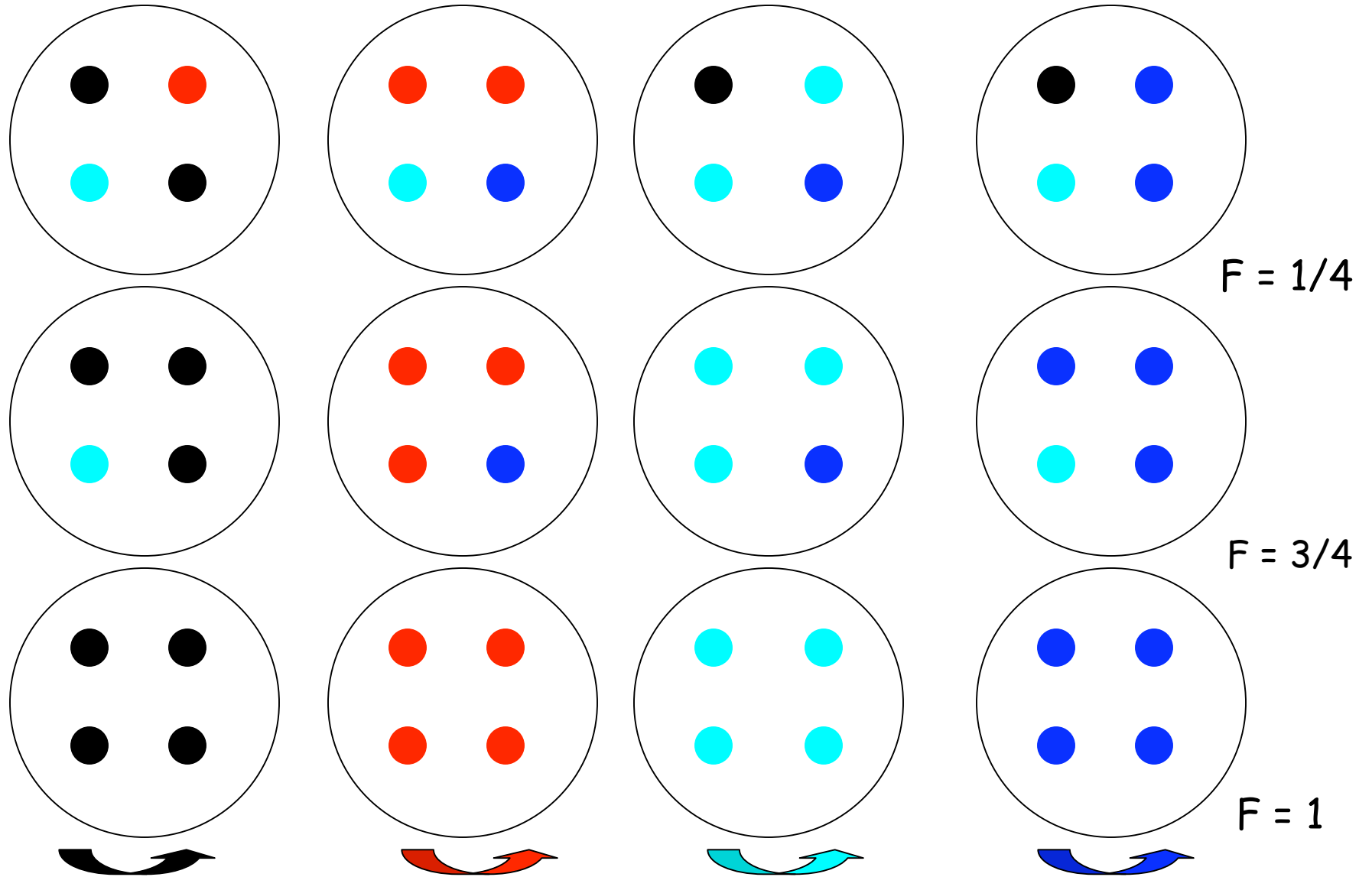
Inbreeding *reduces* variation within each population

Inbreeding *increases* the variation between populations
(i.e., variation in the means of the populations)



$$F = 0$$

Between-group variance increases with F



Within-group variance decreases with F

Variance Changes Under Inbreeding

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

The above results assume ONLY additive variance i.e., no dominance/epistasis. When nonadditive variance present, results very complex (see WL Chpt₁₄10).

Mutation and Inbreeding

- As lines lose genetic variation from drift, mutation introduces new variation
- Eventually these two forces balance, leading to an equilibrium level of genetic variance reflecting the balance between loss from drift, gain from mutation

V_M = new mutation variation each generation, typically $V_M = 10^{-3} V_E$

Assuming:

Strictly neutral mutations

Strictly additive mutations

Symmetrical distribution of mutational effects

$$V_A = V_G = 2N_e V_M$$

Between-line Divergence

The between-line variance in the mean (V_B) in generation t is

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

For large t , the asymptotic rate is $2V_M t$

Implications: Two identical lines will have their difference in means eventually (approximately) following a normal distribution with mean 0 and variance $2V_M t$, e.g., $\mu(1) - \mu(2) \sim N(0, 2V_M t)$

Line Crosses: Heterosis

When inbred lines are crossed, the progeny show an increase in mean for characters that previously suffered a reduction from inbreeding.

This increase in the mean over the average value of the parents is called **hybrid vigor** or **heterosis**

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

A cross is said to show heterosis if $H > 0$, so that the F_1 mean is larger than the average of both parents.

Expected levels of heterosis

If p_i denotes the frequency of Q_i in line 1, let $p_i + \delta p_i$ denote the frequency of Q_i in line 2.

The expected amount of heterosis becomes

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

- **Heterosis depends on dominance:** $d = 0$ = no inbreeding depression and no Heterosis. As with inbreeding depression, directional dominance is required for heterosis.
- **H is proportional to the square of the difference in gene frequency between populations** H is greatest when alleles are fixed in one population and lost in the other (so that $|\delta p_i| = 1$). $H = 0$ if $\delta p = 0$.
- **H is specific to each particular cross.** H must be determined empirically, since we do not know the relevant loci nor their gene frequencies.

Heterosis declines in the F_2

In the F_1 , all offspring are heterozygotes. In the F_2 , random mating has occurred, reducing the frequency of heterozygotes.

As a result, there is a reduction of the amount of heterosis in the F_2 relative to the F_1 ,

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

Since random mating occurs in the F_2 and subsequent generations, the **level of heterosis stays at the F_2 level.**

Agricultural importance of heterosis

Crosses often show **high-parent heterosis**, wherein the F_1 not only beats the average of the two parents (**mid-parent heterosis**), it exceeds the best parent.

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

Hybrid Corn in the US

Shull (1908) suggested objective of corn breeders should be to find and maintain the best parental lines for crosses

Initial problem: early inbred lines had low seed set

Solution (Jones 1918): use a hybrid line as the seed parent, as it should show heterosis for seed set

1930's - 1960's: most corn produced by double crosses

Since 1970's most from single crosses

A Cautionary Tale

1970-1971 the great Southern Corn Leaf Blight almost destroyed the whole US corn crop

Much larger (in terms of food energy) than the great Irish potato blight of the 1840's

Cause: Corn can self-fertilize, so to make hybrids either have to manually detassel the pollen structures or use genetic tricks that cause male sterility.

Almost 85% of US corn in 1970 had Texas cytoplasm Tcms, a mtDNA encoded male sterility gene

Tcms turned out to be hyper-sensitive to the fungus *Helminthosporium maydis*. Result over a billion dollars of crop loss

Crossing Schemes to Reduce the Loss of Heterosis: Synthetics

Take n lines and construct an F_1 population by making **all pairwise crosses**

Allow random mating from the F_2 on to produce a **synthetic population**

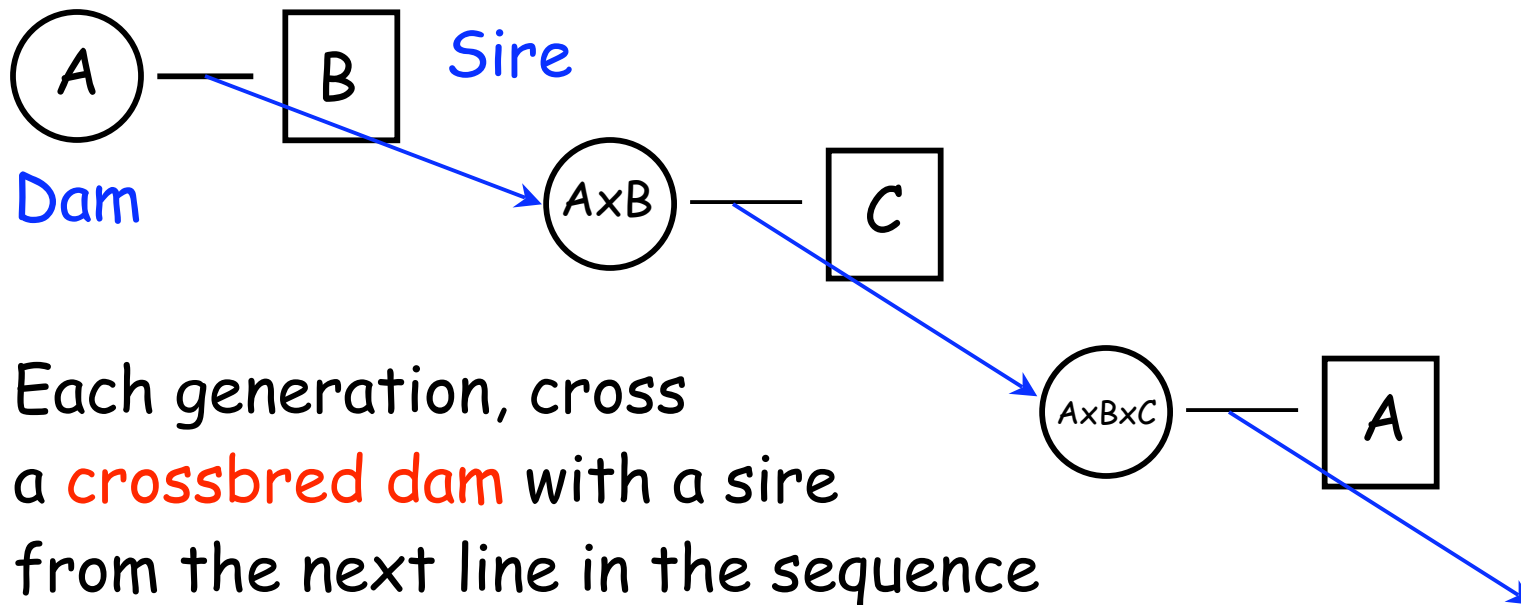
$$F_2 = F_1 - \frac{F_1 - \bar{P}}{n} \quad H/n$$

$$H_{F_2} = H_{F_1} \left(1 - \frac{1}{n} \right) \quad \text{Only } 1/n \text{ of heterosis lost vs. } 1/2$$

Schemes to Reduce the Loss of Heterosis:

Rotational Crossbreeding


Suppose we have three "pure" lines, A, B, C
Originally suggested for pig populations



Each generation, cross
a **crossbred dam** with a sire
from the next line in the sequence

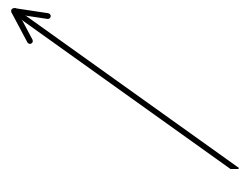
Use of a crossbred dam exploits
maternal heterosis (better moms)

The expected mean value under a two-way rotation:

$$R_2 = \bar{z}_{AB} - \frac{\bar{z}_{AB} - \bar{P}_2}{3}, \quad \text{where} \quad \bar{P}_2 = \frac{\bar{z}_A + \bar{z}_B}{2}$$


Key: Heterosis advantage divided by **3**, not by 2 as in F_2

The expected mean value under a three-way rotation:

$$\hat{R}_3 = SC_3 - \frac{\bar{z}_{AB} - \bar{P}_3}{7}, \quad \text{where} \quad SC_3 = \frac{\bar{z}_{AB} + \bar{z}_{AC} + \bar{z}_{BC}}{3}$$


1/7th of heterosis is lost

Under a 4-way rotation, the order matters:

Mean of all six pairwise crosses

Mean of crosses of nonadjacent lines

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

1/15 of heterosis is lost

Trait	\bar{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 m weight gain	64.4	68.9	64.4	64.6	61.7

Note that $F_1 > R > S > \bar{P}$

For a 2-way rotation:

$$\hat{R}_2 = F_1 - \frac{F_1 - \bar{P}_2}{3}$$

For weaning weight

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

For the 2-breed synthetic, $\hat{S}_2 = 180.5 - \frac{180.5 - 154.2}{2} = 167.4$

Individual vs. Maternal Heterosis

- **Individual heterosis**
 - enhanced performance in a hybrid individual
- **Maternal heterosis**
 - enhanced maternal performance (such as increased litter size and higher survival rates of offspring)
 - Use of **crossbred dams**
 - Maternal heterosis is often comparable, and can be greater than, individual heterosis

Individual vs. Maternal Heterosis in Sheep traits

Trait	Individual H	Maternal H	total
Birth weight	3.2%	5.1%	8.3%
Weaning weight	5.0%	6.3%	11.3%
Birth-weaning survival	9.8%	2.7%	12.5%
Lambs reared per ewe	15.2%	14.7%	29.9%
Total weight lambs/ewe	17.8%	18.0%	35.8%
Prolificacy	2.5%	3.2%	5.7%

Estimating the Amount of Heterosis in Maternal Effects

Contributions to mean value of line A

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

Individual genetic effect (BV)

Maternal genetic effect (BV)

Grandmaternal genetic effect (BV)

Consider the offspring of an A sire and a B dam

Individual genetic value is the average of both parental lines

Contribution from (individual) heterosis

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

Maternal and grandmaternal effects from the B mothers

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

Now consider the offspring of an B sire and a A dam

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

Maternal and grandmaternal
genetic effects for B line

Difference between the two line means estimates
difference in maternal + grandmaternal effects
in A vs. B

Hence, an estimate of individual heterotic effects is

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

Likewise, an estimate of maternal/grandmaternal effects is given by

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

How about estimation of maternal heterotic effects?

The mean of offspring from a sire in line C crossed to a dam from a A X B cross (B = granddam, AB = dam)

Average individual genetic value
(average of the line BV's)

Genetic maternal effect
(average of maternal BV for both lines)

Grandmaternal genetic effect

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

New individual heterosis of C x AB cross

Maternal genetic heterotic effect

"Recombinational loss" --
- decay of the F₁ heterosis in the F₂

One estimate (confounded) of maternal heterosis

$$\mu_{C \cdot AB} - \frac{\mu_{CA} + \mu_{CB}}{2} = h_{AB}^M + \frac{r_{ab}^I}{2}$$