Lecture 9 (Part a) Inbreeding and Crossbreeding

Bruce Walsh lecture notes
Tucson Winter Institute
7 - 9 Jan 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

• $F = \text{Prob(} \text{the two alleles within an individual are IBD) -- identical by descent}$

• 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
<table>
<thead>
<tr>
<th>Genotype</th>
<th>Alleles IBD</th>
<th>Alleles not IBD</th>
<th>frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1A_1$</td>
<td>Fp</td>
<td>$(1-F)p^2$</td>
<td>$p^2 + Fpq$</td>
</tr>
<tr>
<td>$A_2A_1$</td>
<td>0</td>
<td>$(1-F)2pq$</td>
<td>$(1-F)2pq$</td>
</tr>
<tr>
<td>$A_2A_2$</td>
<td>Fq</td>
<td>$(1-F)q^2$</td>
<td>$q^2 + Fpq$</td>
</tr>
</tbody>
</table>
Changes in the mean under inbreeding

Genotypes

<table>
<thead>
<tr>
<th>0</th>
<th>a+d</th>
<th>2a</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1A_1 )</td>
<td>( A_1A_2 )</td>
<td>( A_2A_2 )</td>
</tr>
</tbody>
</table>

freq(\( A_1 \)) = p, freq(\( A_2 \)) = q

Using the genotypic frequencies under inbreeding, the population mean \( \mu_F \) under a level of inbreeding \( F \) is related to the mean \( \mu_0 \) 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 if dominance is present (d not 0)
- 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
Inbreeding depression

Example for maize height

$F_2$, $F_3$, $F_4$, $F_5$, $F_6$
Fitness traits and inbreeding depression

- Often seen that inbreeding depression is strongest on fitness-relative traits such as yield, height, etc.
- Traits less associated with fitness often show less inbreeding depression
- Selection on fitness-related traits may generate directional dominance
Why do traits associated with fitness show inbreeding depression?

• Two competing hypotheses:

  - **Overdominance Hypothesis**: Genetic variance for fitness is caused by loci at which heterozygotes are more fit than both homozygotes. Inbreeding decreases the frequency of heterozygotes, increases the frequency of homozygotes, so fitness is reduced.

  - **Dominance Hypothesis**: 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 homozygotes for deleterious alleles, so fitness is reduced.
Inbred depression in largely selfing lineages

• Inbreeding depression is common in outcrossing species
• However, generally fairly uncommon in species with a high rate of selfing
• One idea is that the constant selfing have purged many of the deleterious alleles thought to cause inbreeding depression
• However, lack of inbreeding depression also means a lack of heterosis (a point returned to shortly)
  - Counterexample is Rice: Losts of heterosis and inbreeding depression
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
Implications for traits

- A series of inbred lines from an F_2 population are expected to show
  - more within-line uniformity (variance about the mean within a line)
    - Less within-family genetic variation for selection
  - more between-line divergence (variation in the mean value between lines)
    - More between-family genetic variation for selection
**Variance Changes Under Inbreeding**

<table>
<thead>
<tr>
<th></th>
<th>General</th>
<th>$F = 1$</th>
<th>$F = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between lines</td>
<td>$2FV_A$</td>
<td>$2V_A$</td>
<td>$0$</td>
</tr>
<tr>
<td>Within Lines</td>
<td>$(1-F) V_A$</td>
<td>$0$</td>
<td>$V_A$</td>
</tr>
<tr>
<td>Total</td>
<td>$(1+F) V_A$</td>
<td>$2V_A$</td>
<td>$V_A$</td>
</tr>
</tbody>
</table>

The above results assume ONLY additive variance i.e., no dominance/epistasis. When nonadditive variance present, results very complex (see WL Chpt 3).
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 allele frequencies 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 F₂

In the F₁, all offspring are heterozygotes. In the F₂, random mating has occurred, reducing the frequency of heterozygotes.

As a result, there is a reduction of the amount of heterosis in the F₂ relative to the F₁,

$$H_{F₂} = \mu_{F₂} - \frac{\mu_{P₁} + \mu_{P₂}}{2} = \frac{(\delta p)^2 d}{2} = \frac{H_{F₁}}{2}$$

Since random mating occurs in the F₂ and subsequent generations, the level of heterosis stays at the F₂ 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.

<table>
<thead>
<tr>
<th>Crop</th>
<th>% planted as hybrids</th>
<th>% yield advantage</th>
<th>Annual added yield: %</th>
<th>Annual added yield: tons</th>
<th>Annual land savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maize</td>
<td>65</td>
<td>15</td>
<td>10</td>
<td>55 x 10$^6$</td>
<td>13 x 10$^6$ ha</td>
</tr>
<tr>
<td>Sorghum</td>
<td>48</td>
<td>40</td>
<td>19</td>
<td>13 x 10$^6$</td>
<td>9 x 10$^6$ ha</td>
</tr>
<tr>
<td>Sunflower</td>
<td>60</td>
<td>50</td>
<td>30</td>
<td>7 x 10$^6$</td>
<td>6 x 10$^6$ ha</td>
</tr>
<tr>
<td>Rice</td>
<td>12</td>
<td>30</td>
<td>4</td>
<td>15 x 10$^6$</td>
<td>6 x 10$^6$ ha</td>
</tr>
</tbody>
</table>
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 potato blight of the 1840’s

Cause: Corn can self-fertilize, so to make hybrids either have to manually detassle 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*. Resulted in 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:

\[
F_2 = F_1 - \frac{F_1 - \bar{P}}{n} H/n
\]

\[
H_{F_2} = H_{F_1} \left(1 - \frac{1}{n}\right)
\]

Only \( 1/n \) of heterosis lost vs. \( 1/2 \).
Synthetics

- **Major trade-off**
  - As more lines are added, the $F_2$ loss of heterosis declines
  - However, as more lines are added, the mean of the $F_1$ also declines, as less elite lines are used
  - Bottom line: For some value of $n$, $F_1 - H/n$ reaches a maximum value and then starts to decline with $n
Types of crosses

• The $F_1$ from a cross of lines $A \times B$ (typically inbreds) is called a single cross.

• A three-way cross (also called a modified single cross) refers to the offspring of an $A$ individual crossed to the $F_1$ offspring of $B \times C$.
  - Denoted $A \times (B \times C)$

• A double (or four-way) cross is $(A \times B) \times (C \times D)$, the offspring from crossing an $A \times B$ $F_1$ with a $C \times D$ $F_1$. 
Predicting cross performance

• While single cross (offspring of A x B) hard to predict, three- and four-way crosses can be predicted if we know the means for single crosses involving these parents

• The three-way cross mean is the average mean of the two single crosses:
  - mean(A x {B x C}) = \left[ \text{mean}(A x B) + \text{mean}(A x C) \right]/2

• The mean of a double (or four-way) cross is the average of all the single crosses,
  - mean({A x B} x {C x D}) = \left[ \text{mean}(AxC) + \text{mean}(AxD) + \text{mean}(BxC) + \text{mean}(BxD) \right]/4