

# Definitions

**Heterozygosity:** the probability that two alleles at locus drawn from a population, or an individual, are different

**Inbreeding:** A form of non-random mating between closely related individuals, which leads to reduced heterozygosity within individuals

# Hardy-Weinberg Equilibrium (HWE)

If we assume random mating no selection, no drift, no mutation, and no migration then we can calculate expected genotype frequencies from allele frequencies

Expectations for genotype frequencies are calculated assuming the “random union of gametes”

Example: biallelic locus with A and a alleles

If we denote the frequency of A as p and a as q, then we can derive expected genotype frequencies at HWE

$$E(p_{AA}) = p^2$$

$$E(p_{Aa}) = 2pq$$

$$E(p_{aa}) = q^2$$

where  $E(p_{AA})$ ,  $E(p_{Aa})$ ,  $E(p_{aa})$  are the expected genotype frequencies at HWE

# Hardy-Weinberg Equilibrium (HWE)

If we assume random mating no selection, no drift, no mutation, and no migration then we can calculate expected genotype frequencies from allele frequencies

Expectations for genotype frequencies are calculated assuming the “random union of gametes”

A locus is at HWE when genotype frequencies match expectations under the random union of gametes

Example: biallelic locus with A and a alleles

If we denote the frequency of A as p and a as q, then we can derive expected genotype frequencies at HWE

$$E(p_{AA}) = p^2$$

$$E(p_{Aa}) = 2pq$$

$$E(p_{aa}) = q^2$$

where  $E(p_{AA})$ ,  $E(p_{Aa})$ ,  $E(p_{aa})$  are the expected genotype frequencies at HWE

# The Wahlund effect

Population structure is a form of nonrandom mating that causes deviations from HWE

The Wahlund effect is the deviation from HWE expectations when:

- (1) multiple differentiated populations are sampled
- (2) expectations of genotype frequency under HWE are derived without knowledge of existing population structure between them

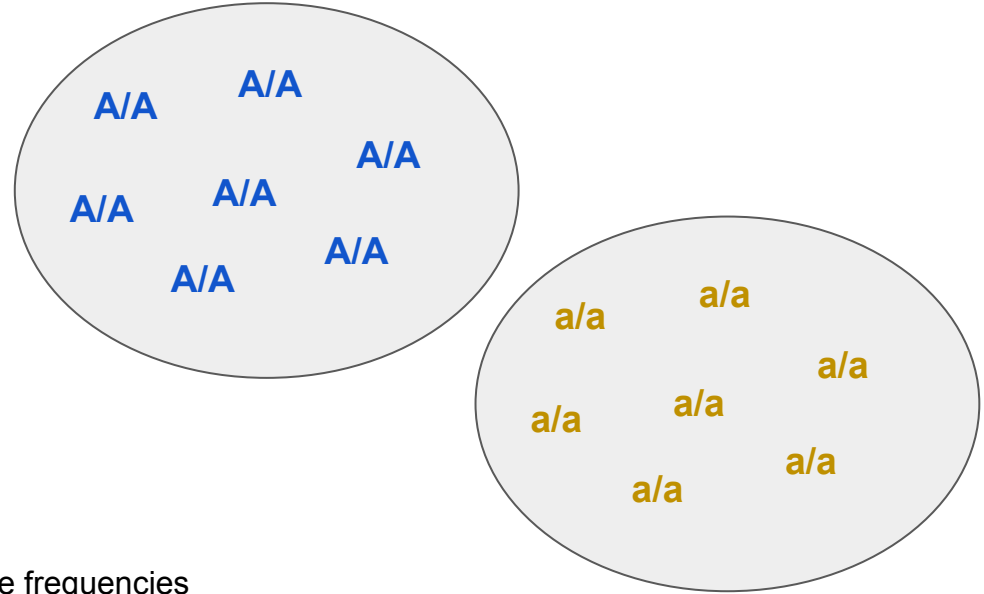
# The Wahlund effect

**Example:** Extreme population differentiation

Consider two populations that are fixed for alternate alleles **A** with frequency  $p$  and **a** with frequency  $q$

The frequency of **A** allele is 1 in population 1 and 0 in population 2, the frequency of **a** allele is 0 in population 1 and 1 in population 2

If we **combine** both populations, global allele frequencies are  $p = 0.5$  and  $q = 0.5$



Expected genotype frequencies  
(under HWE)

$$E(p_{AA}) = p^2 = 0.25$$

$$E(p_{Aa}) = 2pq = 0.5$$

$$E(p_{aa}) = 0.25$$

Observed genotype frequencies

$$O(p_{AA}) = 0.5$$

$$O(p_{Aa}) = 0$$

$$O(p_{aa}) = 0.5$$

Wahlund Effect is observed  
when combining  
differentiated populations

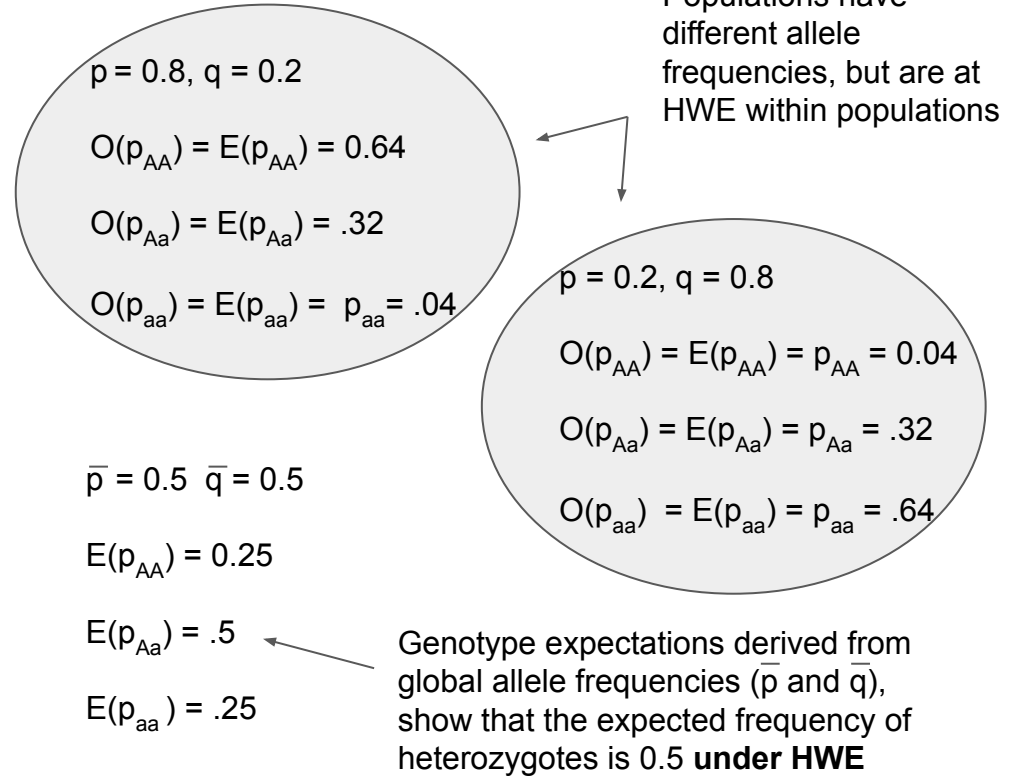
# The Wahlund effect

## Example:

Consider two populations with frequency of A allele denoted by  $p$  and a allele denoted by  $q$

both expected and observed

deficit:  
global expectation -  
group expectation

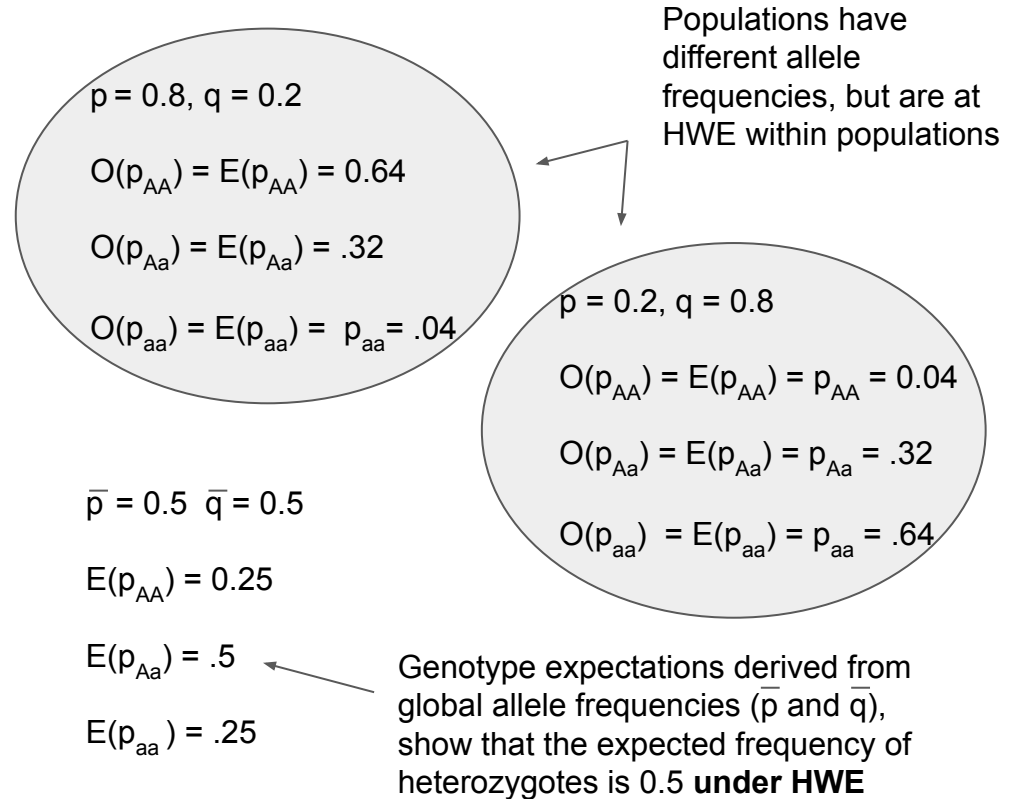


# The Wahlund effect

## Key observations

1. The two populations differ in allele frequencies (i.e.,  $p = 0.8$  in population 1 and  $p = 0.2$  in population 2)
2. Within populations, observed genotype frequencies (e.g.,  $O(p_{AA})$ ) match expectations under HWE (e.g.,  $E(p_{AA})$ )
3. The observed heterozygote frequency of 0.32 in both populations is less than expected under HWE considering both populations in aggregate

Key point: The deficit of observed heterozygotes (observed = .32, expected = 0.5) considering the populations in aggregate is the Wahlund Effect (i.e., the deficit of heterozygotes due to nonrandom mating among populations)



# Inbreeding

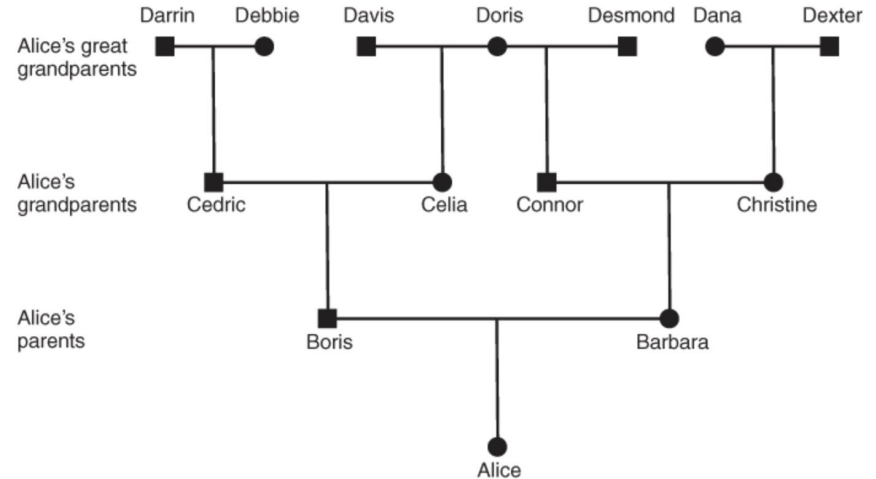
A form of non-random mating between closely related individuals, which leads to reduced heterozygosity within individuals



# Inbreeding

A form of non-random mating between closely related individuals, which leads to reduced heterozygosity within individuals

First cousin inbreeding is the most severe form of inbreeding typically observed in human populations



# The inbreeding coefficient, $F$

$F$  is the probability of inbreeding

Therefore, we can derive expectations for the impact of inbreeding on genotype frequencies expected under HWE

Genotype	Probability of Inbreeding	Probability of random mating	Expected genotype frequencies allowing for inbreeding
AA	$F$	$1 - F$	$p^2 + pqF$
Aa	$F$	$1 - F$	$2pq - 2pqF$
aa	$F$	$1 - F$	$q^2 + pqF$

When  $F = 0$ , expected genotype frequencies match Hardy-Weinberg expectations

When  $F = 1$ , there are zero heterozygotes

# F calculated from the deficit of heterozygotes within individuals

Inbreeding results in reduced observed heterozygosity within individuals

Another way to think of F is as a measure of the reduction of heterozygosity due to inbreeding

This can be quantified by  $F_{IS}$ , the reduction in heterozygosity in individual(s) relative to expectations from random union of gametes in the subpopulation

$$F_{IS} = \frac{H_{exp} - H_{obs}}{H_{exp}} = \frac{1 - H_{obs}}{H_{exp}}$$

Where,

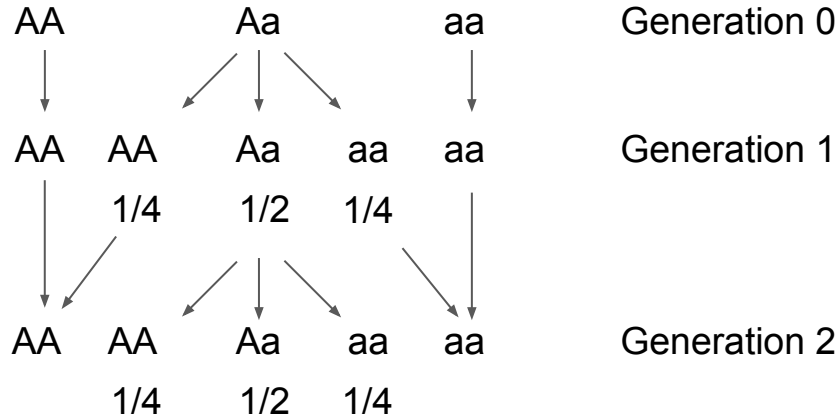
$H_{exp}$  is the expected heterozygosity from subpopulation allele frequencies

$H_{obs}$  is the observed heterozygosity (proportion of genotypes that are heterozygous)

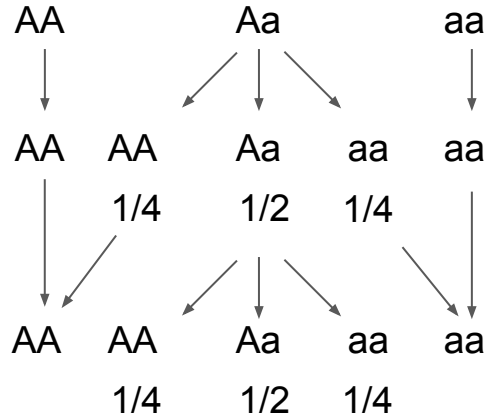
When  $H_{obs} = 0$ ,  $F_{IS} = 1$

When  $H_{obs} = H_{exp}$ ,  $F_{IS} = 0$

# Example: inbreeding under self-fertilization (“selfing”)

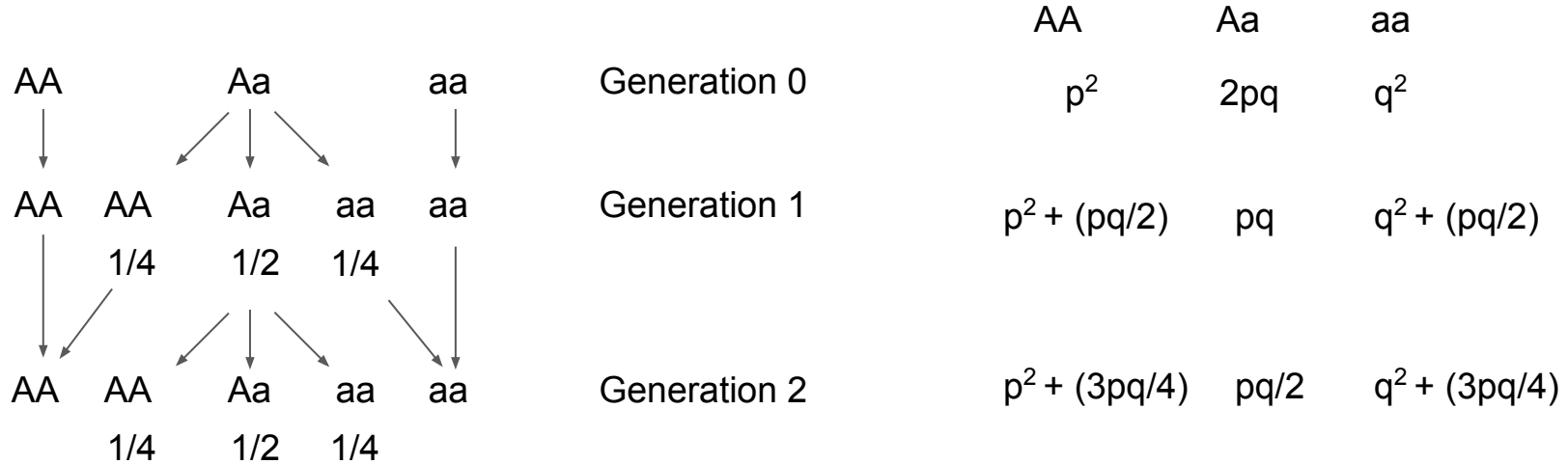


# Example: inbreeding under self-fertilization (“selfing”)



Generation 0 is at HWE			
Generation 0	AA	Aa	aa
	$p^2$	$2pq$	$q^2$
Generation 1	$p^2 + (pq/2)$	$pq$	$q^2 + (pq/2)$
Generation 2	$p^2 + (3pq/4)$	$pq/2$	$q^2 + (3pq/4)$

# Example: inbreeding under self-fertilization (“selfing”)



## Key points:

- (1) under self-fertilization, heterozygosity decreases by  $\frac{1}{2}$  each generation and homozygosity increases by  $\frac{1}{2}$
- (2) Inbreeding does not change the allele frequencies (assuming each genotype has equal fitness)

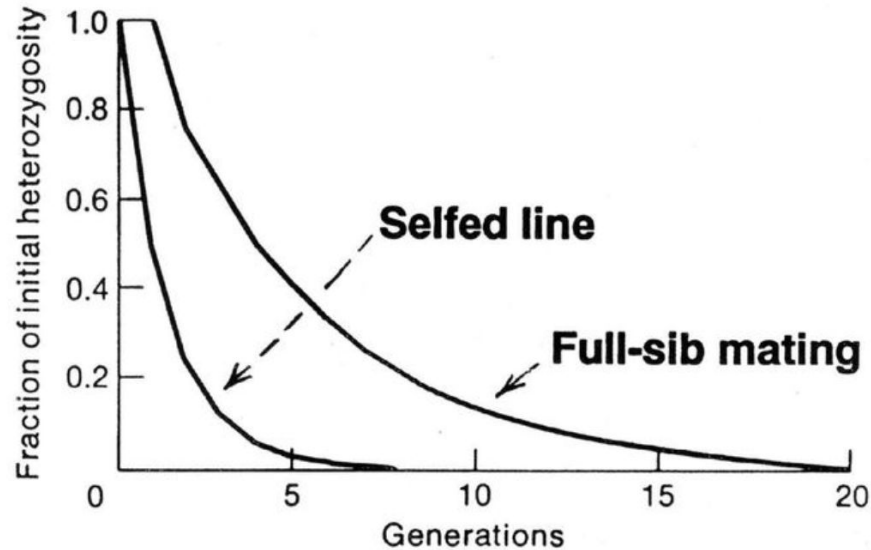
## Example: inbreeding under self-fertilization (“selfing”)

	AA	Aa	aa
Generation 0	$p^2$	$2pq$	$q^2$
Generation 1	$p^2 + (pq/2)$	$pq$	$q^2 + (pq/2)$
Generation 2	$p^2 + (3pq/4)$	$pq/2$	$q^2 + (3pq/4)$

Now, imagine that at Generation,  $p = 0.7$  and  $q = 0.3$ , what are the genotype frequencies in Generation 0, 1, and 2?

	AA	Aa	aa
Generation 0			
Generation 1			
Generation 2			

Milder forms of inbreeding have a similar impact, but take longer to reach homozygosity





# Reductions in heterozygosity due to population structure and inbreeding are both due to non-random mating

The main effect of population structure is to decrease heterozygosity among populations relative to the heterozygosity expected under random mating

Similarly the main effect of inbreeding is to decrease heterozygosity within the same population relative to that expected under random mating

# F calculated from the deficit of heterozygotes within individuals

Inbreeding results in reduced observed heterozygosity within individuals

Another way to think of F is as a measure of the reduction of heterozygosity due to inbreeding

This can be quantified by  $F_{IS}$ , the reduction in heterozygosity in individual(s) relative to expectations from random union of gametes in the subpopulation

$$F_{IS} = \frac{H_{exp} - H_{obs}}{H_{exp}} = \frac{1 - H_{obs}}{H_{exp}}$$

Where,

$H_{exp}$  is the expected heterozygosity from subpopulation allele frequencies

$H_{obs}$  is the observed heterozygosity (proportion of genotypes that are heterozygous)

When  $H_{obs} = 0$ ,  $F_{IS} = 1$

When  $H_{obs} = H_{exp}$ ,  $F_{IS} = 0$

# Hierarchical F-statistics

deficit of heterozygosity of non-random mating

Population show evidence of non-random mating at multiple levels or scales

Sewall Wright developed an index of population structure called the “Fixation Index”, or “F”

Previously, we saw that population structure causes non-random mating among populations and a loss of heterozygosity in subpopulations relative to the global population (quantified by  $F_{ST}$ )

Inbreeding similarly leads to a loss of heterozygosity in individuals relative to the subpopulation (quantified by  $F_{IS}$ )

In fact, we can define losses of heterozygosity at any level in contrast with expectations at any higher level

These are referred to as hierarchical F-statistics

$$F_{IS} = \frac{H_S - H_I}{H_S} = \frac{1 - H_I}{H_S}$$

Reduction in heterozygosity within subpopulations relative to the total

$$F_{ST} = \frac{H_T - \bar{H}_S}{H_T} = \frac{1 - \bar{H}_S}{H_T}$$

Reduction in heterozygosity within subpopulations relative to the total

$$F_{SC} = \frac{H_C - \bar{H}_S}{H_C} = \frac{1 - \bar{H}_S}{H_C}$$

Reduction in heterozygosity within subpopulations relative to regions

$$F_{CT} = \frac{H_T - \bar{H}_C}{H_T} = \frac{1 - \bar{H}_C}{H_T}$$

Reduction in heterozygosity within regions relative to the total

# Inbreeding depression

Inbreeding depression is the reduction in fitness due to increased homozygosity

Fertility is one measure of fitness which can be measured by the number of children

Kinship is a measure of relatedness

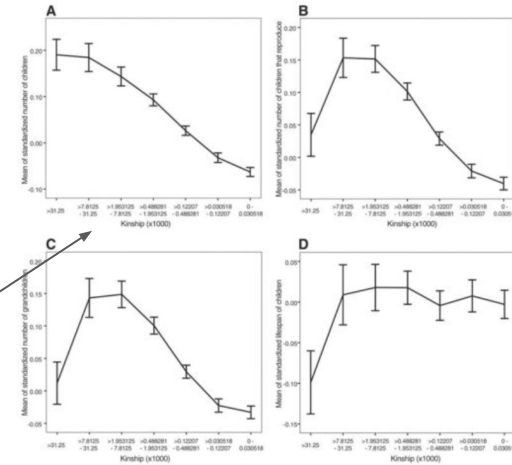
Number of children

Science 8 February 2008:  
Vol. 319, no. 5864, pp. 813 - 816

## An Association Between the Kinship and Fertility of Human Couples

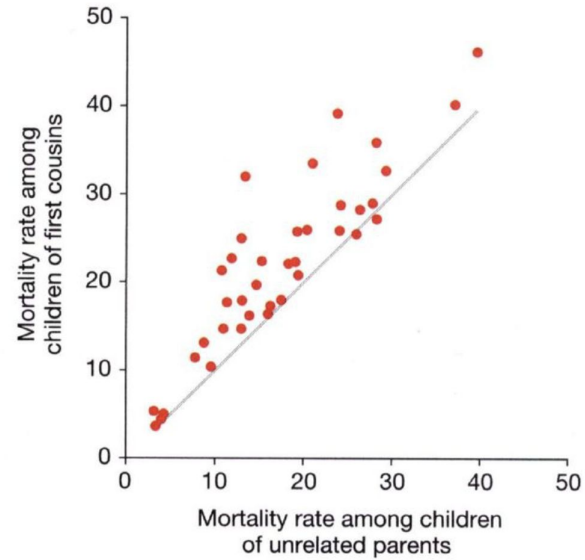
Helgason et al.

Kinship



**Fig. 1.** The relationship between kinship and reproduction among Icelandic couples. The four panels show means and 95% confidence intervals of standardized variables relating to the reproductive outcome of Icelandic couples as a function of seven intervals of kinship. **(A)** shows the total number of children, **(B)** the number of children who reproduced, **(C)** the number of grandchildren, and **(D)** the mean life expectancy of children. **The first interval of kinship represents all couples related at the level of second cousins or closer, the second interval represents couples related at the level of third cousins and up to the level of second cousins, and so on, with each subsequent category representing steps to fourth, fifth, sixth, and seventh cousins and the final category representing couples with no known relationship and those with relationships up to the level of eighth cousins.**

# Inbreeding depression



# Causes of inbreeding depression

Long-standing debate

Recessive deleterious alleles could be masked in the heterozygous state ('dominance hypothesis')

Heterozygotes could be inherently more fit ('overdominance hypothesis')