

F-statistics

- Used to evaluate the genetic variation at 3 hierarchical levels:
 - Within subpopulations, among subpopulations, and total variation in the metapopulation
- All measure loss of heterozygosity in relation to HWE:
 - H_i : heterozygosity *observed* within subpopulations
 - H_s : heterozygosity expected with random mating in a subpopulation
 - H_T : the expected heterozygosity if the entire population were undivided undergoing random mating

Measures of heterozygosity

H_i : heterozygosity *observed* within subpopulations

Same as H , the proportion of heterozygotes in a population

H_s : heterozygosity expected with random mating in a subpopulation

= $2p_i q_i$, or the same as H_0 in the inbreeding coefficient, F

H_T : the expected heterozygosity if the entire population were undivided undergoing random mating

= $2p_0 q_0 = 2\bar{p}q$

Note, use proportions for calculations of H , not absolute numbers

F-statistics

- Individuals within subpopulations

$$F_{IS} = \frac{\bar{H}_S - \bar{H}_I}{\bar{H}_S}$$

Measures the proportional reduction in heterozygotes within populations due to inbreeding, same as the inbreeding coefficient, F , but averaged over all subpopulations

F-statistics

- Among populations

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

- Measures the proportional reduction in the heterozygosity of the metapopulation due to differentiation among subpopulations, relative to the expectation with no population subdivision.
- Note, F_{ST} is meaningless for a single population
- F_{ST} is also known as the fixation index because it increases as more subpopulations become fixed for an allele (less gene flow among pops)
- Values from 0.05-0.15 are moderately differentiated, from 0.15-0.25 are greatly differentiated and >0.25 are very greatly differentiated.
- Typically populations are 0.15 or less

F-statistics

- Individuals within populations

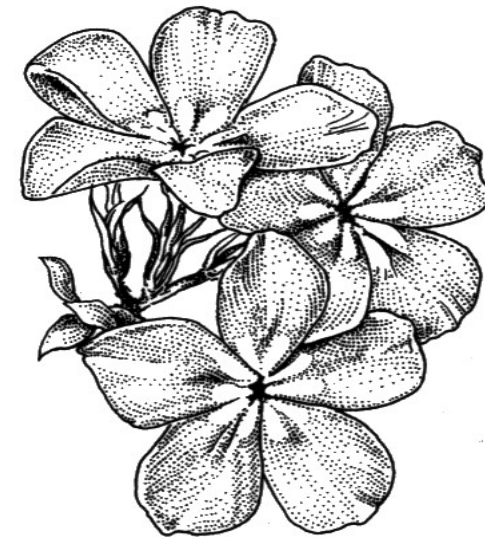
$$F_{IT} = \frac{H_T - \bar{H}_I}{H_T}$$

Measures the reduction in heterozygotes within an individual relative to the metapopulation due to both nonrandom mating and genetic drift among populations. Not as informative as other measures due to the nondifferentiation of inbreeding and drift.

Let's calculate!

Levin (1978) scored allele frequencies at the *Pgm-2* locus in 43 Texas subpopulations of *Phlox cuspidata*. Forty of these subpopulations were fixed for the *b* allele (listed together in the first row of the table below). In the other three subpopulations the frequencies of *b* were 0.49, 0.83, and 0.91, with observed heterozygote frequencies of 0.17, 0.06, and 0.06, respectively:

Subpopulation	p_i	H_i
1–40	1	0
41	0.49	0.17
42	0.83	0.06
43	0.91	0.06



Phlox cuspidata

Magnitude of genetic drift, and thus F_{ST} , depends on population size

- Our earlier examination of drift assumed:
 - Equal numbers of each sex
 - No sexual or natural selection (each individual has an equal probability of successfully contributing gametes to the next generation)
 - Subpopulation size remains the same in each generation
- *Effective population size*, N_e , is the size of an idealized population (all of the above assumptions met) that would experience the same magnitude of genetic drift as the actual population of interest.

F-statistics

$$(1 - F_{IS})(1 - F_{ST}) = (1 - F_{IT})$$

Let's violate some assumptions!

- Nonequal numbers of each sex

$$N_e = \frac{4 N_m N_f}{N_m + N_f}$$

Let's violate some assumptions!

- Individuals have unequal probabilities of contributing offspring to the next generation (i.e., selection)

$$N_e \approx \frac{8 N_a}{V_m + V_f + 4}$$

- What happens if a few dominant males get most of the matings?



Cervus elaphus

What is the effective population size in a population of red deer where there are 33 males and 35 females, with variances of 41.9 and 9.1 respectively?

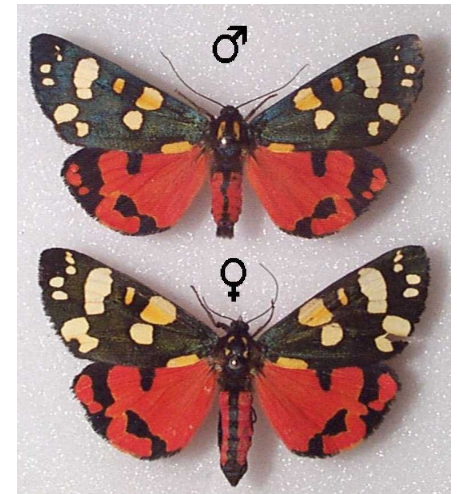
Let's violate some assumptions!

- Population size fluctuates through time

$$\frac{1}{N_e} = \frac{1}{t} \left(\frac{1}{N_1} + \frac{1}{N_2} + \dots + \frac{1}{N_t} \right)$$

What is the effective population size for a population of moths (*Panaxia dominula*) that has the following population sizes over 4 successive years?

11,000 → 2,000 → 11,000 → 16,000



Magnitude of genetic drift depends on N_e

$$\Delta F_{ST} = \frac{1}{2N_e}$$