

Fst and Population Differentiation

$$F_{ST} = \frac{F_{IT} - F_{IS}}{1 - F_{IS}}$$

$$F_{ST} = \frac{F_S - F_T}{1 - F_T}$$

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

$$F_{ST} = \frac{\text{var}(p)}{\bar{p}(1 - \bar{p})}$$

• F_{ST} is a measure of the diversity of randomly chosen alleles within the same subpopulation relative to that found in the entire population. It is often expressed as the proportion of genetic diversity due to allele frequency differences among populations

Fst

- F_{IT} is the inbreeding coeff of an individual (I) relative to the total (T) pop.
- F_{IS} is the inbreeding coefficient of an individual (I) relative to the subpop (S)
- F_{ST} is the effect of subpopulations (S) compared to the total population (T)

$$F_{ST} = \frac{F_{IT} - F_{IS}}{1 - F_{IS}}$$

- F_S is the probability of IBS of 2 genes drawn at random from the subpopulation (S)
- F_T is the probability of IBS of 2 genes drawn at random from the total population (T)
- F_{ST} is the probability of *identity by state* of 2 genes drawn at random from the total population that is due to subpopulations

$$F_{ST} = \frac{F_S - F_T}{1 - F_T} = \frac{H_T - H_S}{H_T}$$

- F_{ST} as the proportion of variance in allele frequencies that is among populations

$$F_{ST} = \frac{\text{var}(p)}{\bar{p}(1 - \bar{p})}$$

The Average “Heterozygosity” Expected by Randomly Choosing Two Genes From Within a Subpopulation = $H_S = [140(0.36) + 372(0.29)]/512 = 0.309$

<p>H.-W. Het. = $2pq = .36$</p> <p>140 Pueblo Indians</p>		<p>H.-W. Het. = $2pq = .29$</p> <p>372 Australian Aborigines</p>	
M	N	M	N
0.76	0.24	0.176	0.824
<p>↓ Total Population ↓</p>			
M	N		
[140(.76) + 372(.176)]/512 = 0.34	[140(.24) + 372(.824)]/512 = 0.66		

The “Heterozygosity” Expected by Randomly Choosing Two Genes From The Total Population = $H_T = 2\bar{p}\bar{q} = 2(.34)(.66) = 0.449$

$$f_{ST} = (H_T - H_S)/H_T = (.449 - .309)/.449 = 0.312$$

HWD & Inbreeding Coefficients

- H-W Disequilibrium coefficients are sometimes parameterized in terms of an measure of inbreeding (f).

$$p_{AA} = p_A^2 + p_A p_a \cdot f$$

$$p_{Aa} = 2p_A p_a \cdot (1 - f)$$

$$p_{aa} = p_a^2 + p_A p_a \cdot f$$

$$p_{AA} = p_A^2 + \text{var}(p_A)$$

$$p_{Aa} = 2p_A p_a - 2 \cdot \text{var}(p_A)$$

$$p_{aa} = p_a^2 + \text{var}(p_A)$$

$$p_A^2 + f p_A p_a = p_A^2 + \text{var}(p_A)$$

$$f p_A p_a = \text{var}(p_A)$$

$$f = \frac{\text{var}(p_A)}{p_A p_a}$$

$$x_{ijk} = \mu + a_i + b_{ij}$$

variance components: σ_a^2 , σ_b^2

Genotypic data, one population, within- individual level

Source of variation	Degrees of freedom	Sum of squares (SSD)	Variance component
Among Individuals	$N - 1$	$SSD(AI)$	$2\sigma_a^2 + \sigma_b^2$
Within Individuals	N	$SSD(WI)$	σ_b^2
Total:	$2N - 1$	$SSD(T)$	σ_T^2

$$F_{IS} = \frac{\sigma_a^2}{\sigma_{Tot}^2}$$

$$x_{ijk} = \mu + a_i + b_{ij}$$

variance components: σ_a^2 , σ_b^2

Genotypic data, one group of populations, no within- individual level

Source of variation	Degrees of freedom	Sum of squares (SSD)	Variance components
Among Populations	$P - 1$	$SSD(AP)$	$n\sigma_a^2 + \sigma_b^2$
Within Populations	$2N - P$	$SSD(WP)$	σ_b^2
Total:	$2N - 1$	$SSD(T)$	σ_T^2

$$F_{ST} = \frac{\sigma_a^2}{\sigma_{Tot}^2}$$

$$x_{ijk} = \mu + a_i + b_{ij} + c_{ijk}$$

variance components: σ_a^2 , σ_b^2 , σ_c^2

Genotypic data, several groups of populations, no within- individual level

Source of Variation	Degrees of freedom	Sum of squares (SSD)	Variance components
Among Groups	$G - 1$	$SSD(AG)$	$n''\sigma_a^2 + n'\sigma_b^2 + \sigma_c^2$
Among Populations / Within Groups	$P - G$	$SSD(AP/WG)$	$n\sigma_b^2 + \sigma_c^2$
Within Populations	$2N - P$	$SSD(WP)$	σ_c^2
Total:	$2N - 1$	$SSD(T)$	σ_T^2

$$F_{ST} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma_{Tot}^2}$$

$$x_{ijk} = \mu + a_i + b_{ij} + c_{ijk}$$

variance components: σ_a^2 , σ_b^2 , σ_c^2

Genotypic data, one group of populations, within- individual level

Source of Variation	Degrees of freedom	Sum of squares (SSD)	Variance component
Among Populations	$P - 1$	$SSD(AP)$	$n\sigma_a^2 + 2\sigma_b^2 + \sigma_c^2$
Among Individuals / Within Populations	$N - P$	$SSD(AI/WP)$	$2\sigma_b^2 + \sigma_c^2$
Within Individuals	N	$SSD(WI)$	σ_c^2
Total	$2N - 1$	$SSD(T)$	σ_T^2

$$F_{ST} = \frac{\sigma_a^2}{\sigma_{Tot}^2} \quad F_{IS} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_c^2}$$