Measures of Genetic Diversity

1. Allele frequencies – Suppose there is a locus B with 3 allels. There are 6 possible genotypes: B_1B_1, B_1B_2, B_2B_2, B_2B_3, B_3B_3. Let N = total # individuals; n_{11} = # B_1B_1 individuals, n_{12} = # B_1B_2 individuals, etc. The frequencies of each genotype are given by

$$B_1B_1 = \frac{n_{11}}{N}; \quad B_1B_2 = \frac{n_{12}}{N} \mathit{etc.}$$

Now, let n_1 = # B_1 alleles = $2n_{11} + n_{12} = n_{13}$; and 2N = total # alleles. Thus, the frequency of

$$B_1 = p = \frac{n_1}{2N} = \frac{2n_{11} + n_{12} + n_{13}}{2N};$$

the frequency of

$$B_2 = q = \frac{n_2}{2N} = \frac{n_{12} + n_{23} + 2n_{22}}{2N};$$

and the frequency of

$$B_3 = r = \frac{n_3}{2N} = \frac{n_{13} + n_{23} + 2n_{33}}{2N}.$$

Furthermore, it can be shown that if the frequency of $B_1B_1 = X_{11}$; $B_1B_2 = X_{12}$; $B_1B_3 = X_{13}$ that $p = X_{11} + ½X_{12} + ½X_{13}$

similarly

$$q = ½X_{12} + ½X_{23} + X_{22}$$
$$r = ½X_{13} + ½X_{23} + X_{33}$$

2. Measures of single gene diversity within populations.
   a. Percentage of polymorphic loci per population (P).
      P = mean proportion of loci polymorphic in population = $\underset{1=1}{\Sigma} P_i / N$; where P_i = proportion of loci polymorphic in a population and N = number of populations.

   b. Mean number of alleles per population.
      A = total number of alleles at all loci divided by the number of loci. This is averaged over all populations.

   c. Effective number of alleles.

$$A_e = \frac{1}{2}$$
$$\Sigma p_i$$

      where p_i is the frequency of the ith allele at a locus. Calculate A_e for each locus and average over all loci and populations.

d. Mean heterozygosity per individual per population

$$H_S = 1 - \Sigma p_i^2 = \Sigma p_i\,(1 - p_i)$$

Where p_i is the frequency of the ith allele. This measure is equivalent to the frequency of heterozygotes under the Hardy-Weinberg expectations.

e. Shannon-Wiener Diversity index.

$$H = \Sigma p_i \log_2 p_i = 3.3219\,(\log_{10} N - \frac{1}{N} \sum_{i=1}^{s} n_1 \log_{10} n_i)$$

where N = total number of alleles

n_i = number of copies of the ith allele.

s = number of alleles.

3. Variation within and among populations.

Single gene genetic variation can be partitioned into variation among populations and variation within populations. Two measures are generally used: Wright's F_{ST} and Nei's G_{ST} statistics.

a. F_{ST}

$$F_{ST} = \frac{\sigma^2_{P_i}}{\overline{P}_i(1-\overline{P}_i)}$$

Where $\sigma^2_{P_i}$ is the variance observed among populations for the frequency of the ith allele and $\overline{P}_I$ = the mean frequency of the ith allele.

Levels of electrophoretic variation in several taxa of plants and animals

Taxa	No. species	H	Source
<u>Plants</u>			
Gymnosperms	28	.207	Hamrick et. al. 1981
Dicots	74	.113	Hamrick et. al. 1979
Monocrots	28	.165	Hamrick et. al. 1979
<u>Animals</u>			
Mammals	184	.041	Nevo et. al. 1984
Birds	46	.051	
Reptiles	75	.083	
Exc parthenogenetic	70	.055	
Amphibia	61	.067	
Fish	183	.051	
Echinoderms	15	.126	
Drosophila	34	.123	
Insects exc. Drosophila	122	.089	
Crustacea	122	.082	
Chelizerta	6	.080	
Mollusca	46	.148	
Brachiopoda	3	.137	
Vermes	6	.072	
Coelenterata	5	.140	
<u>Summary</u>			
Vertebrates	551	.054	Nevo et. al. 1984
Invertebrates	361	.100	Nevo et. al. 1984
Plants	113	.141	Hamrick et. al. 1979

F_{ST} values are calculated for each allele. Typically they are only calculated for the most common allele.

b. Nei's G_{ST}

For each polymorphic locus the total genetic diversity (H_T) for the species at any locus is calculated as:

$$H_T = 1 - \Sigma \overline{p}_i^2$$

where $\overline{p}_i$ is the mean frequency of the ith allele at a locus. The total genetic diversity (H_T) can be partitioned into the genetic diversity within populations (H_S) and the genetic diversity among populations (D_{ST}). Thus,

$$H_T = H_S + D_{ST}$$

H_S is the weighted (by the population sample size) mean of the genetic diversity of each population.

G_{ST} = the proportion of genetic diversity due to the among population component. Thus,

$$G_{ST} = \frac{D_{ST}}{H_T}$$

R_{ST} = the ratio of among population diversity to the within population diversity:

$$R_{ST} = \frac{D_{ST}}{H_S}$$

The influence of plant breeding systems and seed dispersal mechanisms on levels of genetic diversity among populations (From Loveless and Hamrick, 1984).

	Number of Studies	Mean Diversity Among Populations (G_{ST})
<u>Breeding System</u>		
Autogamous	39	.523
Annual	31	.560
Perennial	8	.329
Mixed Mating	48	.243
Outcrossed	76	.118
Animal	32	.187
Wind	44	.068
<u>Seed Dispersal Mechanism</u>		
Gravity	59	.446
Animal-Attached	18	.398
Animal-Ingested	14	.332
Explosive	24	.262
Winged/Plumose	48	.079