Repeating pairs of surnames and population genetic structure

Relethford, J.H. (1992) Analysis of Marital Structure in Massachusetts Using Repeating Pairs of Surnames. *Human Biology*, 64(1): 25-33.

References

1. Lasker, G.W. and Kaplan, B.A. (1985) Surnames and Genetic Structure: Repetition of the same Pairs of Names of Married Couples, a Measure of Subdivision of the Population. *Human Biology*, 57(3): 431-440.
2. Relethford, J.H. (1988) Estimation of Kinship and Genetic Distance from Surnames. *Human Biology*, 60(3): 475-492.
3. Crow, J.F. (1980) The Estimation of Inbreeding from Isonymy. *Human Biology*, 52(1): 1-14.
4. Nei, M. (1987) *Molecular Evolutionary Genetics*. pp.130-144, Columbia University Press, New York.
5. Bhatia, K. and Wilson, S.R. (1981) The Application of Gene Diversity Analyses to Surname Diversity Data. *Human Biology*, 88: 121-133.

Hardy-Weinberg principle in population genetics (see 4.)

When X_{11}, X_{12}, X_{22} are genotype frequencies, if the size of the population is sufficiently large, and if mating is completely random and endogamous,

$$X_{11} = {x_1}^2, X_{12} = 2x_1x_2, X_{22} = {x_2}^2 \tag{1}$$

where x_1 and x_2 are gene frequencies of the allele A_1 and the allele A_2, respectively.

In the small population, many factors modify genotype frequencies from this principle. With the fixation index (F),

$$X_{11} = (1-F){x_1}^2 + Fx_1, X_{12} = 2(1-F)x_1x_2, X_{22} = (1-F){x_2}^2 + Fx_2 \tag{2}$$

If inbreeding is only a factor which affect genotype frequencies, the fixation index is same as inbreeding coefficient. In the case that the population is composed of several subdivision,

$$F = \frac{\sigma^2}{\bar{x}(1-\bar{x})}$$

Equations used in Relethford (1992)

$$\mathrm{RP} = \frac{\sum_i \sum_j S_{ij}(S_{ij}-1)}{N(N-1)} \tag{1}$$

where S_{ij} is the number of marriages with groom's name i and bride's name j and N is the total number of marriages ($= \sum_i \sum_j S_{ij}$).

$$\mathrm{RP}_r = \frac{(\sum_i S_i^2 - N)(\sum_j S_j^2 - N)}{N^2(N-1)^2} \tag{2}$$

$$z = \frac{(\mathrm{RP}–\mathrm{RP}_r)}{\mathrm{SE}(\mathrm{RP}_r)} \tag{3}$$