# Simulation for a pair of biallelic Loci in LD

The purpose of this note is to give an introduction that how to simulate a pair of loci, which is in gametic disequilibrium. s the commonly used gametic disequilibrium parameter, D, has its upper bound and lower bound upon to the allele frequencies, it seems to be easier to simulate a pair of loci given D', the relative gematic disequilibrium as defined by Lewontin, the value of which is between -1 to 1.

Let the allele frequency of locus A is of $f_A$ for allele A and of $1 - f_A$ for a, and of locus B is of $f_B$ for B and of $1 - f_B$ for b. The frequencies of the four haplotypes of these two loci are $f_{AB}$ (haplotype AB), $f_{Ab}$ (haplotype Ab), $f_{aB}$ (haplotype aB), and $f_{ab}$ (haplotype ab). The Lewontin's measure (Lewontin 1964; Devlin and Risch 1995) of linkage disequilibrium, which is a quantity between -1 and 1, can be specified first for a pair of loci of question

$$D' = \begin{cases} \dfrac{D}{\min\left(f_A(1-f_B), f_B(1-f_A)\right)}, D > 0 \\ \dfrac{D}{\min\left(f_A f_B, (1-f_A)(1-f_B)\right)}, D < 0 \end{cases}$$

in which $D = f_{AB} - f_A f_B = f_{AB} f_{ab} - f_{Ab} f_{aB}$ [Devlin and Risch 1995]. Regardless of the sign of $D$, the denominator in the expression of $D'$ is denoted as $\psi$, and $D = D' \times \psi$.

| | | Locus A | | |
|---|---|---|---|---|
| | | A (1) | a (0) | |
| Locus B | B (1) | AB $f_A f_B + D$ | aB $(1 - f_A)f_B - D$ | $f_B$ |
| | b (0) | Ab $f_A(1 - f_B) - D$ | ab $(1 - f_A)(1 - f_B) + D$ | $1 - f_B$ |
| | | $f_A$ | $1 - f_A$ | |

In simulation, under random mating, the conditional probability of generating the second locus could be expressed as ($t$ is generation)

$$P(L_B = l_B | L_A = l_A) = \frac{f_{l_B} f_{l_A} + (-1)^{|l_B - l_A|} D^t}{f_{l_A}}$$

and $l_X$ indicates the allele. $l_A = 1$, if the allele is A, and 0 for a. Similar for $l_B$.

$$P(L_B = 1 | L_A = 1) = \frac{P(L_B = 1, L_A = 1)}{f_A} = \frac{f_A f_B + D^t}{f_A}$$

$$P(L_B = 0 | L_A = 1) = \frac{P(L_B = 0, L_A = 1)}{f_A} = \frac{f_A(1 - f_B) - D^t}{f_A}$$

$$P(L_B = 1 | L_A = 0) = \frac{P(L_B = 1, L_A = 0)}{1 - f_A} = \frac{(1 - f_A)f_B - D^t}{1 - f_A}$$

$$P(L_B = 0 | L_A = 0) = \frac{P(L_B = 0, L_A = 0)}{1 - f_A} = \frac{(1 - f_A)(1 - f_B) + D^t}{1 - f_A}$$

Devlin B., Risch N., 1995 A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics **29**: 311–22.

Lewontin R. C., 1964 The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. Genetics **49**: 49–67.