Purpose:

Apply concepts of Disequilibrium at single and pairs of loci.

Keywords:

Disequilibrium, Gametic Disequilibrium, Linkage Disequilibrium

References:

Captivate: Gametic and Linkage Disequilibrium

Bernardo-Chapter 2

ALA:

Consider a set of 105 Sorghum R lines in a hybrid breeding program. These lines were assayed with DaRT markers, M1 and M2. The DaRT assay is a dominant marker system in which the allele is either present or absent. For the following data set determine if alleles at two loci are in equilibrium. If the alleles are not in equilibrium, estimate the disequilibrium values for the pair of loci.

Assay	Observed
M1M2	66
M1m2	11
m1M2	10
m1m2	18

a. What is the expected number under the assumption that the loci are segregating independently? What are the observed Hint: first you need to estimate the frequencies of each allele

a. First, if these two loci are segregating independently, we need to determine what the expected frequency of genotypes will be. We have to estimate the frequency of each allele at each locus:
- observed $f(m1) = Sqrt(m_1m_1) = .53$, thus observed $f(M_1) = .47$.
- observed $f(m2) = Sqrt(m_2m_2) = .52$, thus observed $f(M_2) = .48$.
- Expected $f(phenotypes|r=0.5$ and $f(alleles))$ are given in Table F.1
- $x_3^2 \sim 27$ *there is sufficient evidence to reject the null hypothesis with respect to the alleles at this locus*. We next need to decide if the disequilibrium is due to disequilibrium in the observed frequencies at each individual locus or in the joint disequilibrium at both loci or due to both sources of disequilibrium.

b. Next decide if the alleles are in coupling or repulsion phase. Hint what is the expected frequency of m1m2, if in coupling phase?

If $f(m_1m_1, m_2m_2) > 1/16$, then coupling.

c. Estimate the recombination frequency between the two loci, conditional on the coupling phase.

Given coupling phase disequilibrium, $f(m_1m_1, m_2m_2) = [(1-r)/2]^2 = 18/105$. Thus, a first estimate of $r \sim 0.17$

d. Estimate the joint disequilibrium Dm1m2. Hint: you must first prove that there is no disequilibrium at individual loci, i.e., Dm1 and Dm2 = 0.

The estimate of joint disequilibrium, Dm_{1,m_2} is $f(m_1, m_2) - f(m_1)f(m_2)$.

$f(m_1, m_2) \sim (1-r)/2 \sim .41$

$f(m_1)^2 = f(m_1m_1) - Dm_1 = .28$
$f(m_2)^2 = f(m_2m_2) - Dm_2 = .27$

$f(M_1\_) = f(M_1M_1) + 2[f(M_1)f(m_1)] - Dm_1 = .72$
$f(M_2\_) = f(M_2M_2) + 2[f(M_2)f(m_2)] - Dm_2 = .73$

So, we have to prove that Dm_1 and $Dm_2 = 0$ before we can calculate Dm_1m_2.

$f(M_1\_) + f(m_1m_1) = 1$

also $f(M_1)^2 + 2 f(M_1) f(m_1) + f(m_1)^2 - 2 Dm_1 = 1$, so $Dm_1 = 0$.

Ditto for Dm_2.

Therefore $Dm_{1,m_2} \sim .41 - (.53)(.52) \sim 0.13$

In summary, there is no disequilibrium at either locus, but there is joint disequilibrium between the alleles at both loci. The joint disequilibrium is due to fewer than 50% recombinants involving both loci. This may be due to linkage, but the data are not sufficient to claim linkage disequilibrium.

Phenotype	Observed Count	Expected freq|r=.5	Expected Count
m1m1, m2m2	18	.06	7
M1__, m2m2	10	.21	20
m1m1, M2__	11	.20	20
M1__, M2__	66	.53	59