

### Part III: Linkage disequilibrium.

#### Exercise 7

Whole chromosomes of *Drosophila melanogaster* can be sampled from individuals in natural population, and made completely homozygous by crossing two stocks containing dominantly marked balancer chromosomes that suppress recombination. The genotype for any number of loci on a chromosome is called a haplotype, since the chromosome is essentially one possible haploid gamete type. A total of 47 second chromosomes were sampled from a natural population in Raleigh (North Carolina, USA) and assessed for restriction fragment length variation in the region of the *scabrous* gene (Lai *et al.*, 1994). The following data are numbers of chromosomes of each haplotype for two PstI restriction sites. PstI(-12.0) (locus A) is located 12kb 5' from origin of transcription of the *scabrous* gene, and PstI(5.8) (locus B) is 5.8 kb 3' from the origin of transcription. "+" indicates presence and "-" absence of a site.

Haplotype		Observed	Observed	Expected	Expected
A	B	number	frequency	frequency	number
+	+	4			
+	-	4			
-	+	13			
-	-	26			

We wish (i) to determine whether the two restriction sites are in linkage (dis)equilibrium, and (ii) to calculate  $D$ , the coefficient of linkage disequilibrium and to express the value of  $D$  relative to its theoretical maximum. Consider  $p$ , the frequency of the "+" allele at each locus and  $q$ , the frequency of the "-" allele.

#### Exercise 8

Consider an isolated panmictic population resulting of the mixture of two populations. This population is big enough to neglect the drift; there is no selection and no mutation. For two genes A and B,  $p_1$  and  $p_2$  are the allele frequencies of  $A_1$  and  $A_2$  and  $q_1$  and  $q_2$  those of  $B_1$  and  $B_2$ .  $p_1 = 0.7$  and  $q_1 = 0.3$ .

- What are the frequencies of all possible gametes assuming linkage equilibrium?
- What is the theoretical maximum of the linkage disequilibrium ( $D_{\max}$ ) in this population?
- What are the frequencies of all possible gametes (at generation 0,  $G_0$ ) if there is linkage disequilibrium with  $D$  equal to 50% of its theoretical maximum?
- What will be the linkage disequilibrium at the next generation (at  $G_1$ ) if the 2 loci A and B are distant about i/ 20 cM? ii/ 10 cM? iii/ 3 cM? iv/ what can you conclude?
- What would be the frequencies of all possible gametes (at  $G_1$ ) if the 2 loci A and be are distant about 20cM?
- What would the linkage disequilibrium after 10 generations if the 2 loci A and be are distant about 20cM? Express the value of  $D$  relative to its theoretical maximum.
- What would be the linkage disequilibrium after 20 generations if the 2 loci A and be are distant about 20cM? Express the value of  $D$  relative to its theoretical maximum.
- What would be the frequencies of all possible gametes after 20 generations in this population (the 2 loci A and be are still distant about 20cM)? Does the disequilibrium still detectable?

## Exercise 9

Mukai *et al.* (1974) captured 660 fertilized females of the fruit fly *Drosophila melanogaster* from a natural population in Raleigh (North Carolina, USA). These females were used to found a large experimental population in the laboratory. After about five months (10 generations), 489 third chromosomes in the population were analyzed for allozymes coding for the enzymes esterase-6 (alleles  $E6^F$  and  $E6^S$ ), esterase-C (alleles  $EC^F$  and  $EC^S$ ), and octano dehydrogenase (alleles  $OdH^F$  and  $OdH^S$ ). With 3 loci and two alleles per locus, this makes 8 possible haplotypes. (Please note: in *Drosophila* it is possible to obtain the genotypes of single chromosomes (so called haplotypes). The results were as follows:

Haplotype:	Counts:	Haplotype:	Counts:
$E6^F$ $EC^F$ $OdH^F$	152	$E6^S$ $EC^F$ $OdH^F$	264
$E6^F$ $EC^F$ $OdH^S$	7	$E6^S$ $EC^F$ $OdH^S$	13
$E6^F$ $EC^S$ $OdH^F$	15	$E6^S$ $EC^S$ $OdH^F$	29
$E6^F$ $EC^S$ $OdH^S$	1	$E6^S$ $EC^S$ $OdH^S$	8

The order of the genes in the third chromosome is known to be *E6-EC-Odh*. The recombination fraction between *E6* and *EC* is 0.122, and that between *EC* and *Odh* is 0.002 (that is *E6* and *EC* are rather loosely linked, and *EC* and *Odh* are tightly linked.)

- Carry out an analysis to determine whether there is significant linkage disequilibrium between any of the loci.
- If there is linkage disequilibrium, express the value of *D* relative to its theoretical maximum or minimum.