LECTURE NO EIGHT

LINKAGE

Definition

Linkage is the tendency for alleles of different genes to be passed together from one generation to the next. Only genes situated on the same chromosome can show linkage. Genes on non homologous chromosomes are, by definition, unlinked and always show 50% recombination. Parental gametes carry the same set of alleles as were inherited together from one parent. Recombinant gametes carry alleles derived from both parents. The degree of linkage between two genes depends on the frequency of cross-overs that occur between them during meiosis. The closer they are together the less likely a cross-over will occur between them. Groups of genes that are linked are called linkage groups.

Simple measurement of linkage

Two-factor crosses involve crossing a double heterozygote to a double recessive- The frequencies of different phenotypes in the progeny equal the frequencies of different gametes in the double heterozygote. The proportion of recombinant gametes is the recombination frequency. This is used to give a measure of the distance between two genes. The percentage recombination between two genes is taken as the distance they are apart. One percent recombination equals one map unit or centimorgan (cM). In linkage analysis in humans it is more common to utilize molecular techniques to analyze genotypes than to study phenotypes.

Three factor crosses

These are more accurate than two-factor crosses, in that they identify and utilize many of the double cross-overs that are missed by the two-factor cross. They allow ordering of genes and generate additive map distances.

Interference

It is difficult for chiasmata to form close to one another. For this reason the number of double cross-overs observed in three-factor crosses may differ from that predicted from the frequency of single cross-overs. If fewer cross-overs are observed this is termed positive chromatid interference. This is measured as the coefficient of coincidence.

Linkage analysis in fungi

In ascomycete fungi, that have ordered asci, first and second division segregation patterns can be used to determine distance between a gene and the centromere. A similar approach can be used to estimate distances between genes.

Recombination frequency and physical distance

The distance between genes as measured by recombination frequencies is not a precise measure of the physical distance because the frequency of crossing-over varies in different parts of the genome. It is useful for deciding on gene order. Physical maps provide distance between genes in absolute terms.

Meiosis and gametogenesis (C3) The human genome project (F5)

Defination

Linkage is the tendency for alleles of two or more genes to pass from one generation to the next in the same combination. This usually means that the closer together any two genes lie on the same chromosome the more likely they are to show linkage and the stronger that linkage will be. Genes located on different (non-homologous) chromosomes cannot, by definition, show any linkage. As discussed in Topic C3 genes on different chromosomes assort independently at the first division of meiosis.Consider a double heterozygote *An*, *Sb* where the two genes involved are on different chromosomes. Assume that this individual has inherited the dominant alleles *A* and B from one parent, and the recessive alleles a and *b* from the other. In meiosis four genetically different gametes, *AB, Ab, aB* and *ab* will be produced in equal proportions. Those gametes carrying AB or *ab* are referred to as parental and those carrying *aB or Ab* as recombinant. The recombination frequency is obviously 50%. This is the maximum recombination frequency that can be obtained between any two genes, and genes that show this recombination frequency are said to be unlinked.What would have happened had the two genes been located on the same chromosome? Now the parental allele combinations *AB*, or *ab* can only be rearranged (recombined) if a cross-over occurs at a chiasma between the two genes (see Topic C3). The closer the two genes are together the less. likely it is that a cross-over will take place between them. The frequency of recombinants observed will therefore give an indirect measure of how close the two genes lie to each other. In practice the frequency of cross-overs on most chromosomes is high, and this means that genes that lie far apart of the same chromosome show 50% recombination. If two genes show linkage and a third shows linkage to only one of the original two, by definition they must all be on the same chromosome. They are said to constitute a **linkage group.** Thus linkage maps can be built up even though some genes in the group may not show linkage to all other members of the group.

Simple measurement of linkage

Except in certain circumstances where linkage can be measured to the centromere it is necessary to have heterozygosity for at least two genes in any study of linkage. Such a system is called a two-factor cross. An example is set

out below. An F_1 hybrid was made by crossing two pure-breeding strains of tomato, one of which was homozygous for purple fruit and hairy stems, *PP HH,* and the other was homozygous for red fruit and smooth stems, *pp hh.* The hybrid had purple fruit and hairy stems. Thus purple fruit is dominant to red fruit, and hairy stems is dominant to smooth stems. To determine if the genes for fruit color and hairiness of stem are linked, the doubly heterozygous hybrid was crossed to the double recessive (red-fruited and smooth-stemmed) parent. This is a test cross (see Topics Cl and C2). Four classes of progeny were obtained *(Table I):* classes 1 and 2 are the same phenotypes as the original parental strains and are therefore parental: classes 3 and 4 represent new combinations of the alleles at the two genes and are therefore recombinant. If the genes were unlinked a ratio of 1:1:1:1 would be expected between the four possible phenotypes (see Topic C2) Clearly this is not the case. The total number of progeny is 500 of which only 70 are recombinant. The recombination frequency is 70/500, 14%. This is strong evidence that the two genes are linked. They have a map distance of 14 map units between them

Table 1. Phenotypes produced in a test cross between F_1 generation purple-fruited, hairy-stemmed tomatoes (Pp, Hh), and the double recessive red-fruited, smooth-stemmed parent (pp, hh)

Phenotypic class	Frequency
(1) Purple-fruited, hairy-stemmed	220
(2) Red-fruited, smooth-stemmed	210
(3) Purple-fruited, smooth-stemmed	32
(4) **Red-fruited, hairy-stemmed**	38
Total	500

(1% recombination is equivalent to 1 map unit). Map units **are** referred to as centimorgans (cM) in memory of Thomas Hunt Morgan who was the first geneticist to explain linkage.

Is it necessary to use a test cross to determine linkage? It is not essential, and other crosses can be used, for instance, crossing the double heterozygotes together as in a dihybrid cross (see Topic C2) will give data on linkage, as the numbers in recombinant classes of the 9:3:3:1 ratio will be reduced. However, because the double recessive parent contributes nothing to the phenotypes of *the* progeny in the test cross, the ratios obtained represent exactly the ratios of gametes and so this cross gives maximum information on linkage. In current studies of linkage in humans (see Topic F5) recombination frequencies are usually determined by the use of molecular methods to identify differences in DNA sequence (the genotype) rather than the phenotype as described in the examples given here. Working with genotypes allows more information to be obtained from crosses. The two-factor cross has limited use in determination of linkage and gene mapping studies. To map the order of genes along a chromosome and to give more accurate estimates of the distances between these genes at least three genes should be studied in the same cross.

Three factor crosses

A major advantage of a three-factor cross is that genes may be placed in order and that the map distances, at least over relatively short distances, are additive. The worked example given below again involves a test cross. For simplicity, in this example alleles denoted by capital letters are completely dominant over the lower case alleles, and the three genes are simply named after their dominant alleles.

An individual heterozygous at three genes *An*, Nn and Rr is crossed with the homozygous recessive parent *aa*, *nn,* and rr. The frequency of progeny with different phenotypes is given in *Table 1*.

If there were no linkage between these genes each of the eight classes of progeny should have arisen with equal frequency. This is clearly not the case:

classes (1) and (2) represent progeny from gametes where no recombination had taken place between the three genes. These are parental gametes because they retain the parental combination of alleles of the three genes; classes (3) and (4) represent progeny from gametes where recombination has occurred between the genes N and *R;* classes (5) and (6) represent recombination between the genes *A*

and N; classes (7) and (8) are the most important because they represent gametes where recombination has taken place both in the interval between A and *N* and also between N and *R.* These are referred to as double cross-overs; they are less frequent than any of the other classes because they require two independent events

an individual heterozygous far three genes, and one recessive far the same three genes

Phenotypes of progeny	Frequency	Class
Parental		
AN R	347	(1)
a n r	357	(2)
Recombinant		
A N r	52	(3)
a n R	49	(4)
A n r	90	(5)
a N R	92	(6)
A n R	6	(7)
a N r	7	(8)
Total	1000	

to take place. The two least frequent classes of progeny in a test cross of this nature can be used to identify which gene lies between the other two. In this case the central gene is *N.*

To determine the map distance between *the* genes it is necessary to quantify all the recombination events that have occurred.

(i) To determine the map distance between *A* and *N:*

add progeny in classes (5) and (6) (recombination between *A* and N) and progeny in classes (7) and (8) (double cross-over, one of which is between *A* and *N)*

Express the total as a percentage of all progeny.

$$\frac{(90 + 92 + 6 + 7) \times 100}{1000} = 19.5\% \text{ recombination, or } 19.5 \text{ cM}$$

(ii) To determine the map distance between *N* and *R:*

By the same logic as used for the map distance between *A* and N the distance between genes N and R is:

$$\frac{(52 + 49 + 6+7) \times 100}{1000} = 11.4\% \text{ recombination,or } 11.4 \text{ cM}$$

(iii) To determine the map distance between *A* and R:

This requires a summation of all recombinants.

$$(52+49+90+92+6+7+6+7)X100 = 30.9\% \text{ recombination, or } 30.9 \text{ cM}$$

1000

Note that the distances are additive.

A——————————30.9 cM ——————————»R

19.5cM 11.4 cM

A ——————————» N <——————————»**R**

Three-factor crosses are more accurate than two-factor crosses because they detect, and utilize double cross-overs that would go undetected in two-factor crosses. Because more recombination events are detected the calculated recombination frequencies are greater. This is not a problem as recombination frequencies merely reflect the frequency of cross-overs detected rather than act as a physical measurement of actual distance between genes. This point is referred to again in Topic F5.

Interference

Crossing-over takes place at chiasmata. These are physical structures involving two chromatids. Not surprisingly the presence of one chiasmata in a particular chromosome region can reduce the frequency of others forming close to it. This can result in a reduction in the number of double cross-overs observed. In the previous example the observed frequency of double cross-overs is 13 / 1000 (0.013). By using the data for the observed single cross-over between *A* and N and between N and *R we* can predict the expected number of double cross-overs.

Single cross-overs between *A* and *N* = 195

Single cross-overs between N and R = 114

Predicted number of double cross-overs (195/1000) x (114/1000) x 1000 = 22.3

This is greater than the observed number, 13, suggesting that **positive chromatid interference** is observed in this region. The extent of interference is calculated as the **coefficient of coincidence** (S), the observed number of double cross-overs divided by the expected number of double cross-overs. In this case it would be:
S=13/ 22.3= 0.58

Linkage analysis in fungi

Several ascomycte fungi produce asci which hold the haploid ascospores produced after meiosis in a specific linear order, an **ordered tetrad.** This order reflects the organization of the bivalents at meiosis. The distal and proximal pair of ascospores each contain the products of one bivalent after the completion of both meiotic divisions. In some species the ascus contains eight spores. This is simply due to each spore having duplicated by mitosis (see Topic B2) so that the original pattern is displayed by pairs rather than single spores. The example shown in *Fig. 1.* represents the products of meiosis in a fungus heterozygous for pale and dark spores. The asci contain eight spores.

Where the four asci of the same color are found together there have been no cross-overs between the spore color gene and the centromere. This pattern is known as **first division segregation,** because the two phenotypes are physically separated at the first meiotic division. In order to show other patterns a crossover must have taken place between the centromere and the gene for spore color. The color phenotypes are now separated after the second meiotic division, **second division segregation.** Depending on which chromatids are involved in the cross-over different patterns of light and dark spores can be observed in asci showing second division segregation (recombinant asci). These are shown in *Fig 2.* The percentage recombination is determined as:

$$\frac{1/2 \text{ (number of second division segregating asci)} \times 100}{\text{(total number of asci)}}$$

The number of second division segregating asci is halved because in each of these asd only half of the chromatids have had a cross-over between the spore color gene and die centromere. As well as the distance to the centromere the distance between two genes can also be mapped in ordered tetrads using a similar procedure. Map distances can also be determined by counting phenotypes of individual ascospores in fungal species which do not produce ordered tetrads, or when the geneticist is too lazy to micro-dissect each ascospore individually from the ascus. However only analysis of ordered tetrads allows mapping of the distance between a gene and the centromere.

Recombination frequency and physical distance

We have assumed that recombination frequency is a measure of the actual distance between genes. This is only true in a semi-quantitative sense. It is now known that the frequency of cross-overs varies in different regions of the genome. It tends to be lower near centromeres and higher near telomeres. It can also differ between males and females. However gene maps based on linkage data give a reasonably accurate indication of distance between genes and they are most useful in determining the order of genes along the chromosome.