

# Linkage and Recombination



**T. H. Morgan**  
Nobel Prize 1933



**Calvin B. Bridges**



**Alfred H. Sturtevant**



**Herman Joseph. Muller**  
Nobel Prize 1946

Mendel studied 7 traits and every pair of traits that he reported in his paper segregated independently.

Interpretation: are on different chromosomes.

Peas have  $N = 7$  chromosomes. Somewhat unlikely that each trait is on a different chromosome. In fact we now know they are not.

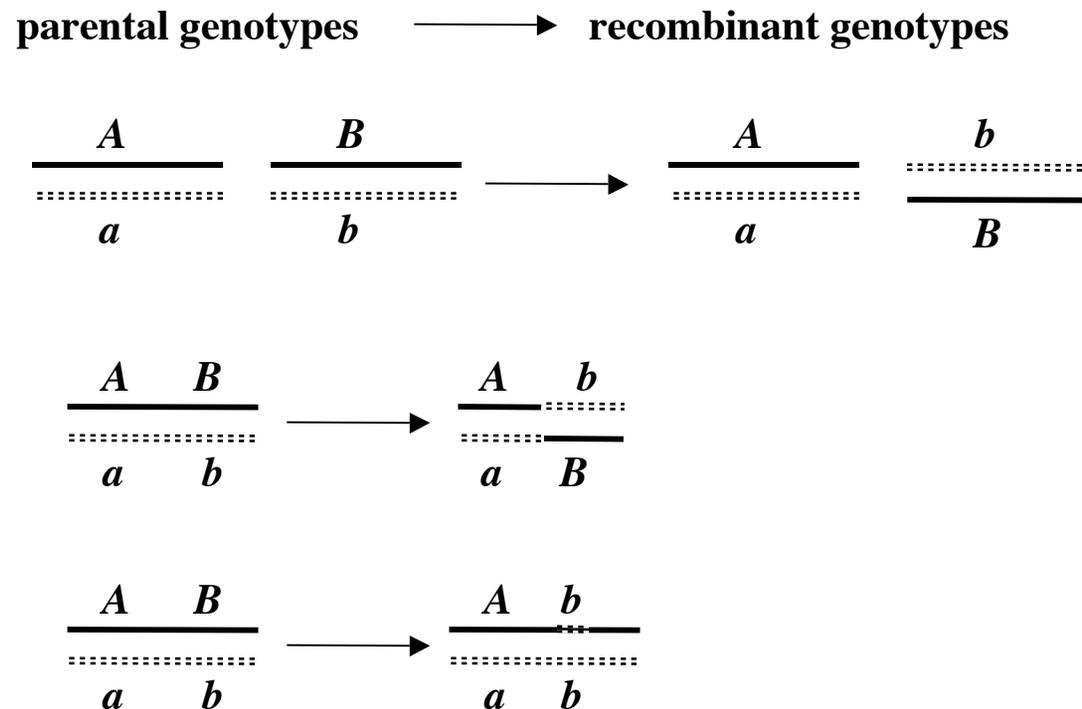
R (round vs. wrinkled) and Gp (green vs. yellow pod) are both on chromosome V (= syntenic) but still segregate independently. This we know is because they are so far apart (ca. 50 cM) that there is on average one crossover between them in every meiosis. This makes them behave as if they are independent = unlinked.

Le (tall, long internode vs. short internode) and V (inflated vs. constricted pod) are both on chromosome III and are so close together that there are very few crossovers between them and they do not segregate independently = linked. If Mendel had done a cross with these two genes, he would have gotten very different results, but he didn't.



Recombination = production of new combinations of alleles at two or more loci. Mechanisms:

1. Independent segregation of genes on different chromosomes.
2. Crossing-over between genes on same chromosome.
3. Gene conversion.



We will study linkage, recombination, and gene mapping as follows:

1. Linkage (as it was first seen and understood in *Drosophila*)
2. Definition and mechanisms of recombination
3. Using recombination frequencies to map genes

Extend timeline:

1866 -----	1900 -----	1902-3 -----	1910-1916
basic rules	rediscovery	chromosome theory	Thomas Hunt Morgan
Mendel	Hugo DeVries	Walter Sutton	Alfred H. Sturtevant
	Carl Correns	Theodore B overi	Calvin Bridges
			H. J. Muller
			<i>Drosophila</i>

Linkage was first seen a few years after Mendel's laws were rediscovered in 1900.

First correctly interpreted ca. 1912 by *Drosophila* research group:

T. H. Morgan      Professor at Cal Tech began working with *Drosophila melanogaster*

Calvin Bridges, Alfred H. Sturtevant began working in lab as undergrads

H. J. Muller graduate student in another department

Morgan got first mutant in 1910,  $w$  = white eyes; next year got two more:  $m$  = miniature wings and  $y$  = yellow body

*Drosophila* gene nomenclature:

	<u>mutant allele</u>	<u>wild type allele</u>
mutant recessive (white eye)	$w$	$w^+$
mutant dominant (Bar eye)	$B$	$B^+$

Linkage of genes in animals and plants is seen as deviations from independent segregation; recombination seen as deviation from complete linkage.

Found red/white trait difference linked to male/female (sex) trait:

$w^+ w^+$  red      **X**       $w -$  white  
 $X X$  female            $XY$  male

↓

$w^+ w$  red      **X**       $w^+ -$  red  
 $X X$  female            $XY$  male

↓

all red females (1 red:0 white)	$1/2 w^+ w^+$ & $1/2 w^+ w$
$1/2$ red males	$1/2 w^+ -$
$1/2$ white males	$1/2 w -$

Frequencies of red and white are different in males and females, therefore eye color and sex don't segregate independently, i.e. sex and eye color are linked.

Sex already shown probably determined by sex chromosomes in Drosophila:

Sex determination in Drosophila:

Karyotype

1 (X) -----o                    sex chromosomes  
1 (Y) -----o---

2        -----o-----           autosomes  
3        -----o-----  
4        -o-

X and Y are heteromorphic pair.

Y has few genes

Females        X X

Males         X Y    X O is sterile male

Sex actually determined by interaction of many genes on X and autosomes, but for the following we can think of it as a trait determined by whether there is one or two copies of the X.

Morgan et al. hypothesized  $m$  and  $w$  are both on sex chromosome.

$m m w^+ w^+ \times m^+ w$  (no alleles on Y)  $\rightarrow F_1$  female

$m w^+$

-----  
-----

$m^+ w$

Did testcross to see what kinds of gametes produced. (Cross to  $m w$  male, or to male of any genotype and look only at male progeny.)

	parentals		recombinants	
Expect female $\rightarrow$ eggs	$m^+ w$	$m w^+$	$m^+ w^+$	$m w$
Expected if independent segreg.	0.25	0.25	0.25	0.25
Expected if complete linkage	0.5	0.5	0	0
Observed partial linkage	0.31	0.31	0.19	0.19

Recombination frequency = # recombinant gametes/total gametes

In this case, recomb. freq. = 0.38 so 38% of chromatids are recombinant for  $m$  and  $w$ .

## Three Kinds Of Recombination Seen In Tetrads

Gene conversion can be distinguished from other kinds of recombination only in tetrads:

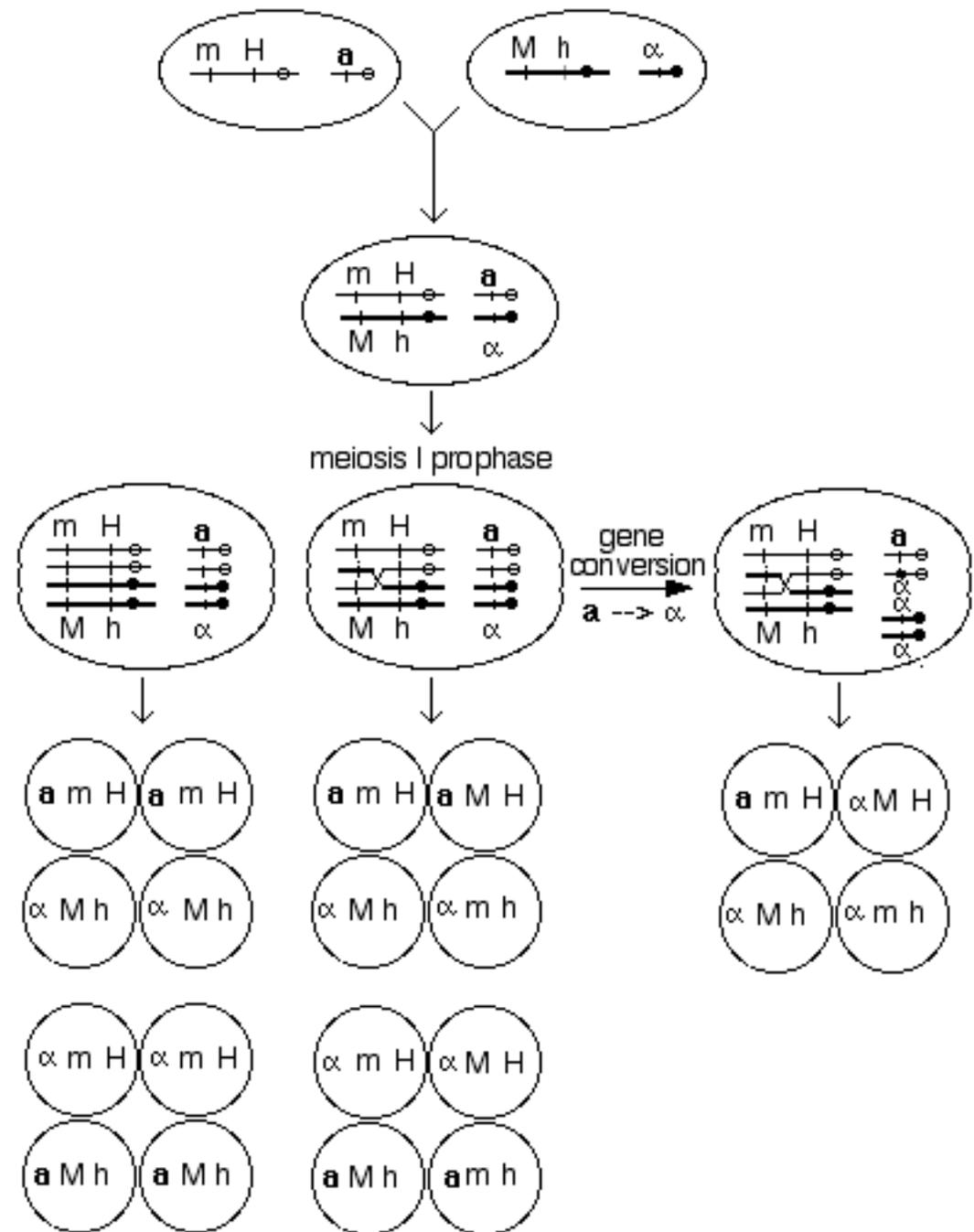
No conversion: all tetrads 2:2

Gene conversion: tetrads 3:1 or 1:3

In random tetrads, conversion genotype  $\alpha M h$  can't be distinguished from same recombinant genotype produced by crossing-over or independent assortment.

How distinguish gene conversion from mutation?

- Much higher frequency
- Often associated with nearby crossover



**When does recombination take place?**

- **Prophase of meiosis I, very high frequency. Chiasmata are physical manifestation. In some organisms, required for normal segregation.**
- **Interphase of mitosis, low frequency**

Genes are linked if show  $< 50\%$  recombination.

Genes are unlinked if show  $50\%$  recombination = independent segregation.

Genes are completely linked if show  $0\%$  recombination; very rare if one looks at very large sample of offspring.

In organisms with prominent haploid stage and tetrads, recombination frequency = frequency of recombinant haploid products = random spores. Also can be calculated from frequency of a certain type of tetrad.

e.g. *ADE8 trp4* X *ade8 TRP4* --> diploid *ADE8 trp4/ade8 TRP4* --meiosis--> many hundreds of asci --> random sample of 100 spores:

<i>ADE8 trp4</i>	38 parental
<i>ade8 TRP4</i>	42 parental
<i>ADE8 TRP4</i>	11 recombinant
<i>ade8 trp4</i>	<u>9</u> recombinant
	100

Recombination frequency =  $20/100 = 0.2$

Recombinants and parentals are in pairs of equally frequent genotypes, because each crossover generates two reciprocal recombinants.

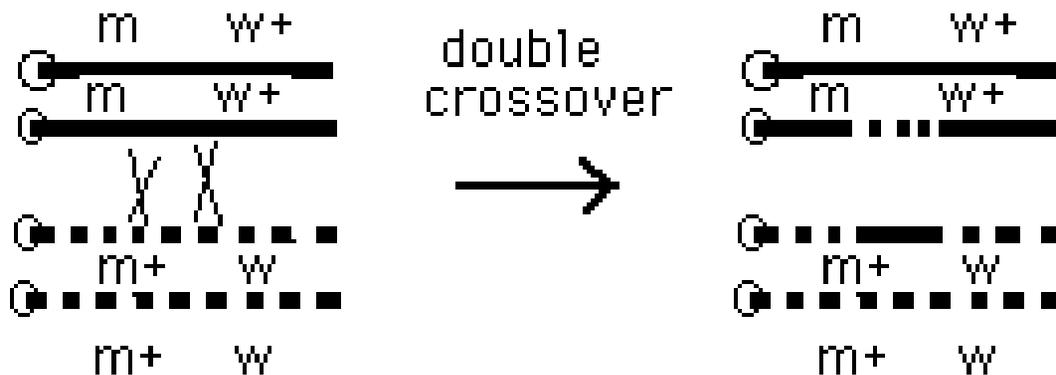


**Double crossovers aren't detected in random gamete samples, so recombination frequencies are always underestimated.**

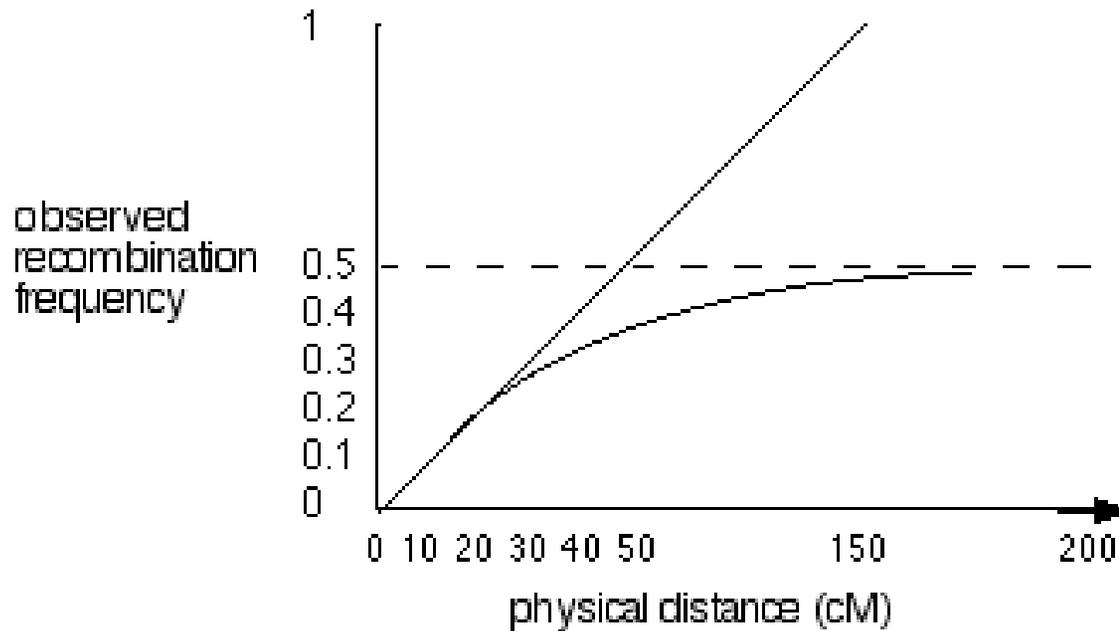
Subsequent studies showed map distances not additive over long distances.

e.g. in Sturtevant's data, he probably observed  $w - m < 33\%$

Reason: double x-overs. Only odd numbers of x-overs between two genes on one pair of chromatids give detectable recombination.



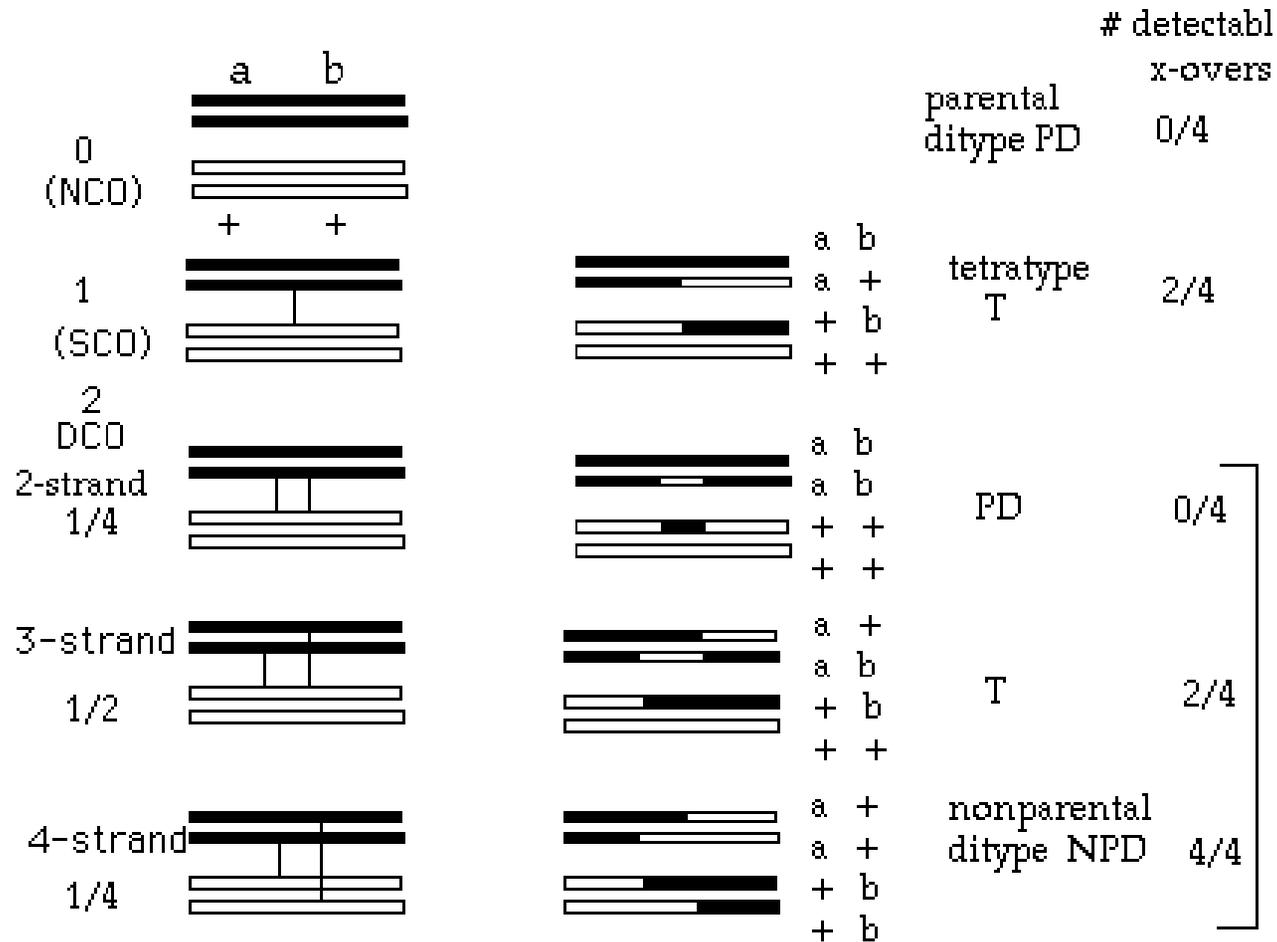
Map function is a mathematical formula that relates the observed number of crossovers to the real number, which is a function of the physical distance. It assumes a simple model in which crossovers are distributed randomly on the chromosome.



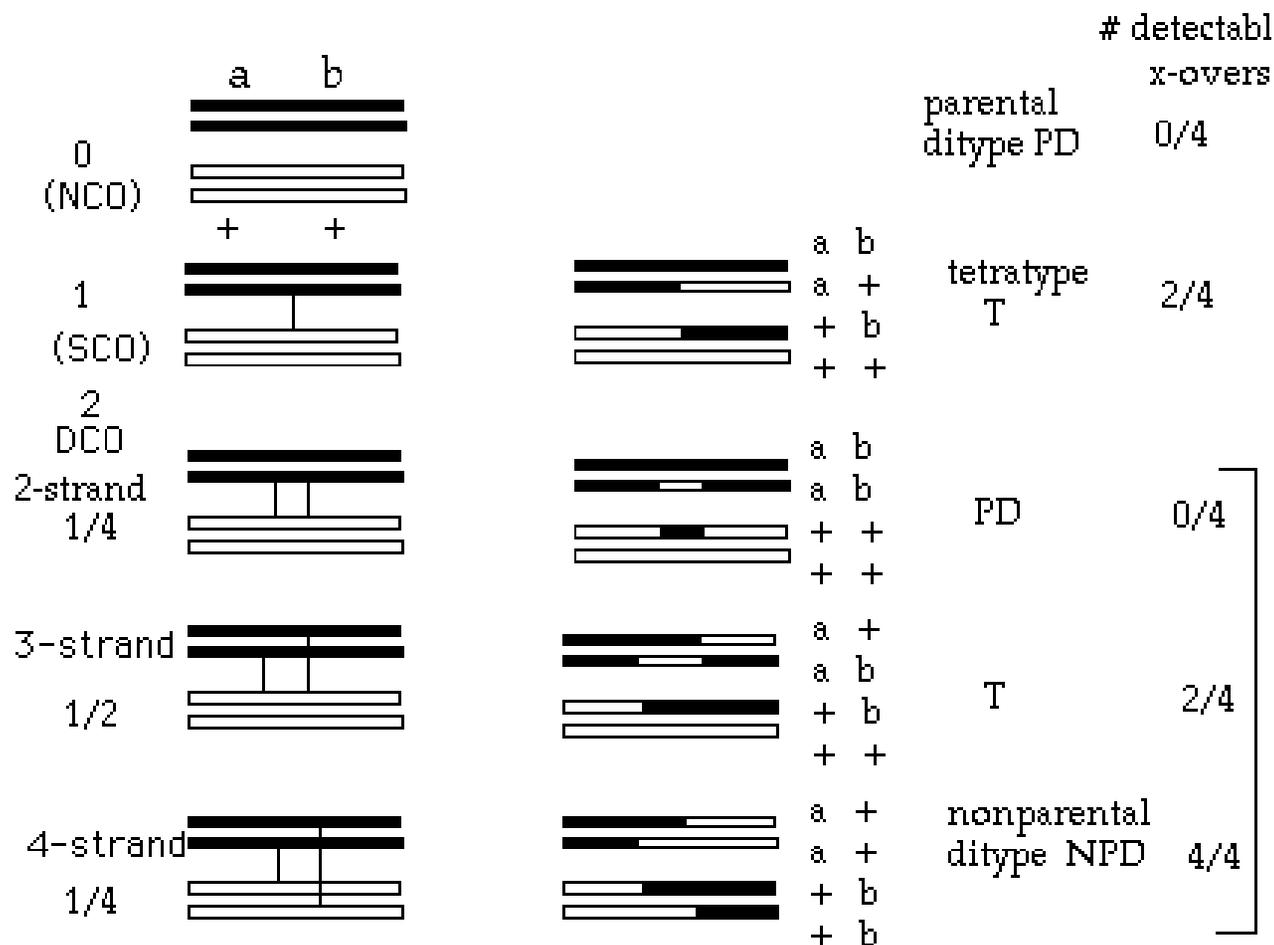
One explanation for why maximum recombination frequency = 50%: as distance increases, # crossovers between any two markers increases, but some of these are double and other even-numbered crossovers. The number of these increases. Number of observed crossovers approaches 0.5 asymptotically.

Tetrad analysis in yeast and *Chlamydomonas* can detect double crossovers.

Cross  $a b \times a^+ b^+$ , diploids undergo meiosis, look at chromosomes in meiosis I:



Double x-overs produce a distinctive type of ascus, the NPD.



Double x-overs produce a distinctive type of ascus, the NPD. Consequently NPDs are a way of estimating the number of DCOs, which will be 4 X the number of NPDs.

Genes unlinked: #PD = #NPD

If every tetrad has a single or double crossover,  $2/4 = 50\%$  of crossovers will be detected; therefore maximum observable frequency of crossing-over is 50%.

## Three-Factor Crosses

In the absence of tetrad analysis, some double crossovers can be detected by using three-factor crosses.

*Drosophila* X-linked genes: yellow body = *y*; cut wings = *ct*; echinus eyes = *ec*

Female heterozygous at all three loci  $y/+ \ ct/+ \ ec/+ \ X \ + \ + \ +$ . This is test cross if look at male progeny because male contributes Y with no genes to male offspring

<u>phenotypes</u>	<u>gametes</u>	<u># males</u>
wild type	+ + +	1080
yellow cut echinus	y ct ec	1071
cut	+ ct +	293
yellow echinus	y + ec	282
yellow	y + +	78
echinus cut	+ ct ec	66
echinus	+ + ec	6
yellow cut	y ct +	<u>4</u>
		2880

Note that the genes are linked; if they weren't, we would have 8 phenotypes and 8 gamete genotypes in approximately equal numbers.

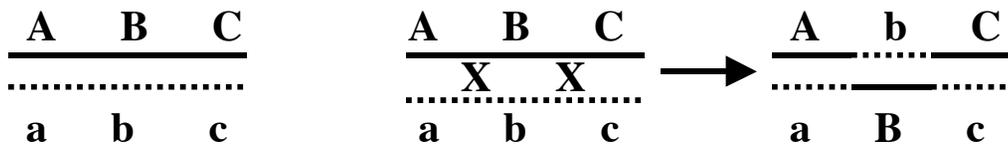
Arranged in pairs of equal numbers, in order of magnitude. Which are parental genotypes? Which are double crossover genotypes?

## Three-Factor Crosses

<u>phenotypes</u>	<u>gametes</u>	<u># males</u>	
wild type	+ + +	1080	parental
yellow cut echinus	y ct ec	1071	parental
cut	+ ct +	293	
yellow echinus	y + ec	282	
yellow	y + +	78	
echinus cut	+ ct ec	66	
echinus	+ + ec	6	double x-over
yellow cut	y ct +	<u>4</u>	double x-over
		2880	

Find parental (most common) and double-crossover (least common) types.

Middle gene is the one that is switched relative to the other two in doubles vs. parentals.



Which gene is in the middle: y, ct, or ec?

## Three-Factor Crosses

<u>phenotypes</u>	<u>gametes</u>	<u># males</u>	
wild type	+ + +	1080	parental
yellow cut echinus	y ct ec	1071	parental
cut	+ ct +	293	
yellow echinus	y + ec	282	
yellow	y + +	78	
echinus cut	+ ct ec	66	
echinus	+ + ec	6	double x-over
yellow cut	y ct +	<u>4</u>	double x-over
		2880	

Analyze data as three two-factor crosses:

<u>y - ct</u>	<u>ct - ec</u>	<u>y - ec</u>
293	293	78
282	282	66
78	6	6
<u>66</u>	<u>4</u>	<u>4</u>
719	585	154
0.250	0.203	0.053

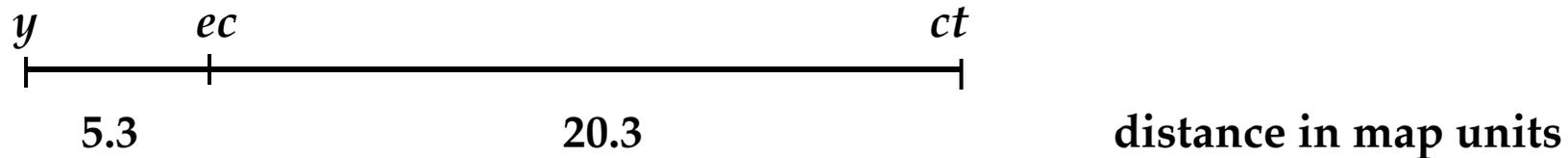
Longest distance is *y - ct*, so these are the outside markers on the map. Agrees with previous conclusion that *ec* is middle marker.

## Three-Factor Crosses

Analyze data as three two-factor crosses:

<u>y - ct</u>	<u>ct - ec</u>	<u>y - ec</u>
293	293	78
282	282	66
78	6	6
<u>66</u>	<u>4</u>	<u>4</u>
719	585	154
0.250	0.203	0.053

Longest distance is  $y - ct$ , so these are the outside markers on the map. Agrees with previous conclusion that  $ec$  is middle marker.



What is best estimate of distance between  $y$  and  $ct$ : 25.0 or  $20.3 + 5.3 = 25.6$ ?



## Interference

**Morgan's group first assumed x-overs occurred independently. Then found out that was wrong:**

**expected doubles if independent:  $P(\text{double}) = P(\text{single y-ec})(P(\text{single ec-ct}) = (0.053)(0.203) = 0.0108$ . Observed doubles =  $10/2880 = 0.0035$ .**

**Or: Expected number of doubles =  $(0.0108)(2880) = 31$ ; observed number of doubles =  $6 + 4 = 10$ . *Can do statistics with numbers, not with frequencies.***

**Observed < expected, therefore one crossover interferes with occurrence of another, but not completely.**

**This works in our favor, means that problem of double crossovers isn't quite as bad as it might be.**

**Also means that we can't predict the number of double crossovers exactly from the number of singles, without correcting for interference.**

**Interference doesn't always happen; mapping very large number of markers in rice showed negative interference over short distances. Maybe because some sites are hotspots for recombination? Recombination frequencies do vary along chromosomes and between chromosomes. *Drosophila*: No crossing-over in males.**

# LINKAGE GROUPS AND CHROMOSOMES

Linkage group = group of genes, each of which is linked ( $r < 0.5$ ) to at least one other.

e.g. new organism

Loci   % recombination

a - b      50

a - c      10

a - d      50

b - c      50

b - d      20

c - d      50

Data allow genes to be put in two linkage groups:

a    10    c

b            20            d

We know that a and c are on the same chromosome, and that b and d are on the same chromosome. Do we know if these two linkage groups are on the same or different chromosomes?

# LINKAGE GROUPS AND CHROMOSOMES

Linkage group = group of genes, each of which is linked ( $r < 0.5$ ) to at least one other.

e.g. new organism

Loci   % recombination

a - b      50

a - c      10

a - d      50

b - c      50

b - d      20

c - d      50

Data allow genes to be put in two linkage groups:

a    10    c

b            20            d

We know that a and c are on the same chromosome, and that b and d are on the same chromosome. Do we know if these two linkage groups are on the same or different chromosomes?  
NO.

Suppose a new mutation e is found. It is linked to both c and d, with recombination frequencies: c-e = 45 and d-e = 35. We now have one linkage group:

a    10    c                    45                    e            35                    d            20            b

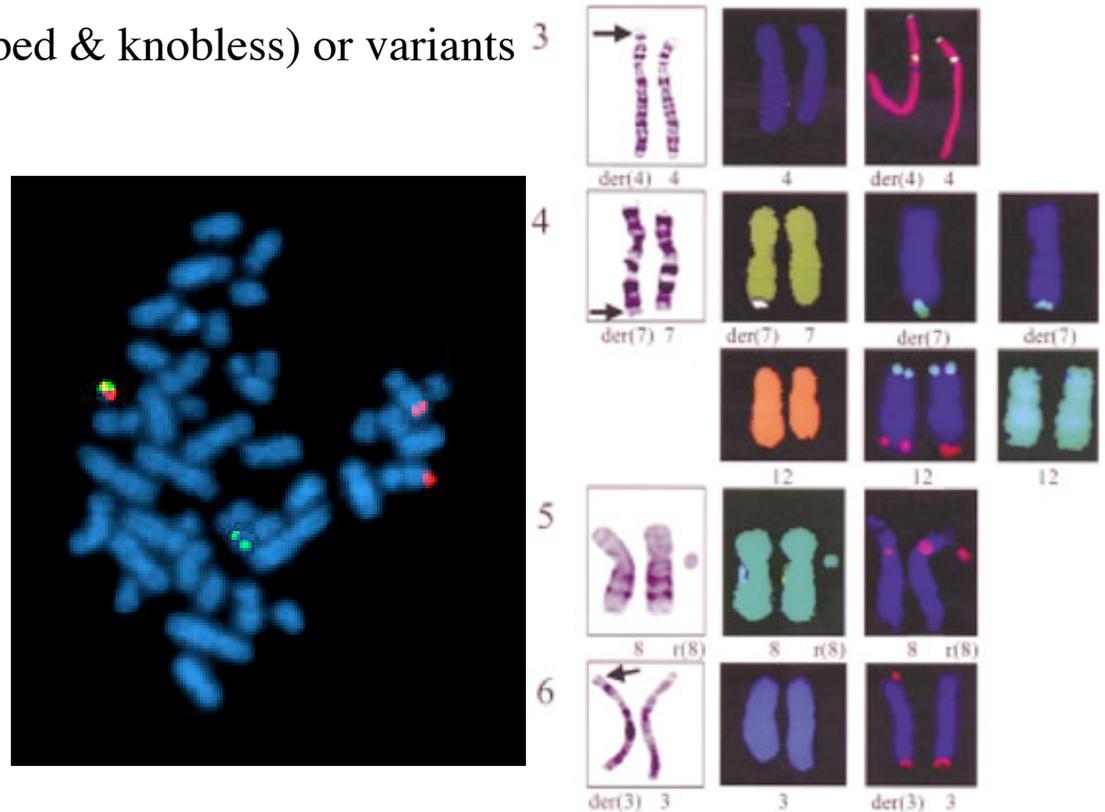
Note that e will show 50% recombination with a and b, even though all of these genes must be on the same chromosome.

If one begins working with a new organism, at first most mutations are unlinked. Eventually some linkage groups appear. The number increases at first, then decreases as mutations are found which combine different linkage groups.

By various genetic tricks, genes can be assigned to specific chromosomes seen in karyotype.

- Use heteromorphic pairs (XY, knobbed & knobless) or variants structure
- Use molecular methods.

e.g. Fluorescent *In Situ* Hybridization (FISH). Cloned DNA segment with Gene of interest is labeled with a fluorescent dye. Squashed metaphase chromosomes are treated to denature the DNA, then the labelled probe is hybridized with the chromosomes.



When one gene is assigned to a specific chromosome, all genes belonging to same linkage group are assigned to that chromosome.

## Predicting the outcome of crosses from linkage maps



$aa BB \times AA bb \rightarrow F_1$

$a B$

-----

Do testcross. What gametes will the  $F_1$  produce?

-----

$A b$

total crossovers  $a-b = 0.15 - 0.10 = 0.05 =$  total freq. recombinants  $a b$  and  $A B$

$a b \quad 0.025$

$A B \quad 0.025$

Parentals =  $1 - \text{recombinants} = 1 - 0.05 = 0.95$

$a B \quad 0.475$

$A b \quad 0.475$

Check:  $0.025 + 0.025 + 0.475 + 0.475 = 1$

## Intragenic Recombination

Gene is long piece of DNA so recombination can occur within as well as between genes. Can have 2 different mutations in same gene --> 2 different mutant alleles.

Can have recombination between the sites marked by different mutations.

Recombination between markers in different genes is called intergenic recombination, even if the crossover event occurs within a third gene.



Recombination between markers in same gene is intragenic recombination.

E.g. Drosophila, chromosome III gene codes for enzyme xanthine dehydrogenase (XDH). Mutants that don't make XDH can't make isoxanthopterin, an eye pigment, so eyes are rosy in color instead of brick red.

Cross  $ry^{23}$  X  $ry^6$  -->  $F_1$  (diagrammed below) --> gametes:

parentals

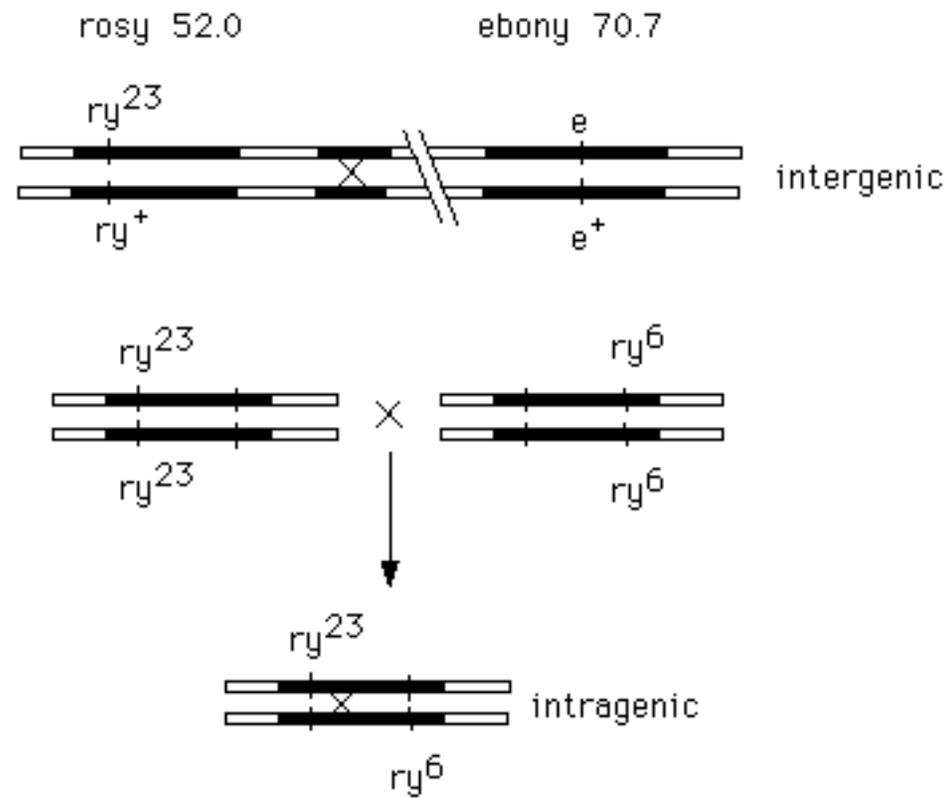
mutant  $ry^{23}$

mutant  $ry^6$

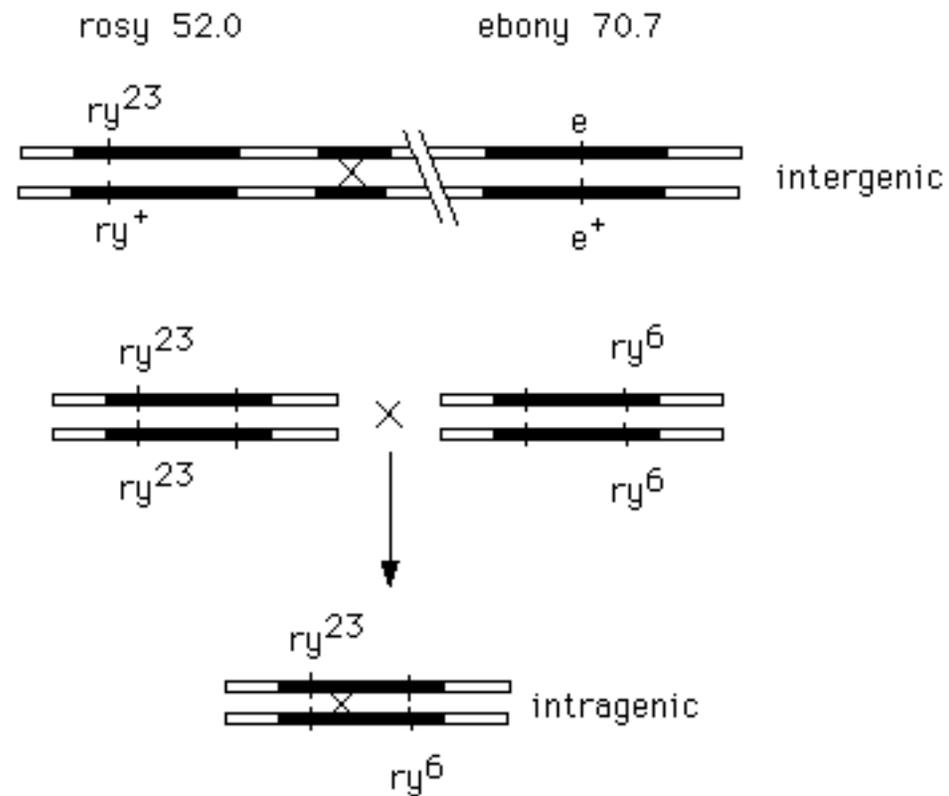
recombinants

double mutants  $ry^{23} ry^6$

wild type  $ry^+$



**What genotypes and phenotypes would one get if there was a crossover between the two *rosy* mutants?**



**What genotypes and phenotypes would one get if there was a crossover between the two *rosy* mutants?**

**Answer: *ry<sup>+</sup>* wild type and *ry<sup>23</sup> ry<sup>6</sup>* double mutant**

## Detecting Intragenic Recombination

The frequency of intragenic recombination is usually low so it can only be detected by selecting recombinants from large numbers of progeny.

*Rosy* gene is about 0.005 cM long, i.e. recombination between markers at the ends of the gene occurs in about 0.005% or  $0.00005 = 5 \times 10^{-5}$  of all gametes from a heterozygous female. Would have to look at  $10^5$  progeny to be sure of finding one or a few recombinants.

Only practical because can select for wild type recombinants: rear progeny on medium with added purine. XDH required to detoxify purine, so purine kills *rosy* mutant larvae but allows wild type larvae to live. In above cross, these are recombinants.

Easier in microorganisms (yeast, bacteria, bacterial viruses).

**Section 15 (this one) is on the web in complete version.**

**Homework deadline extended to Wednesday.**

## **WHAT IS A GENE?**

**Initially a gene was a hereditary factor that had two or more alleles which determined the difference between two or more alternative phenotypes.**

- **Different genes controlled different aspects of phenotype. Gene = unit of function.**
- **Mutation changed one allele to another. Gene = unit of mutation**
- **Different genes could be separated by recombination. Gene = unit of recombination**

**These units usually agreed with each other, until genetic analysis was extended to bacteria and viruses in which rare genotypes can be selected and detected.**

**Then complications arose:**

- **Unit of mutation is a single base pair, not a whole gene.**
- **Mutations within a gene, even in adjacent base pairs, can be separated by recombination.**
- **Two different genes identified by recombination or mutation may control the same phenotype.**

**How can we define a gene?**

## How can we define a gene?

### 1. Sequencing. Problems:

- **Not foolproof.**
- **Impractical in many cases; complete genomes available for very few eukaryotes.**
- **Most genes in most organisms found initially by finding mutants with different phenotypes.**

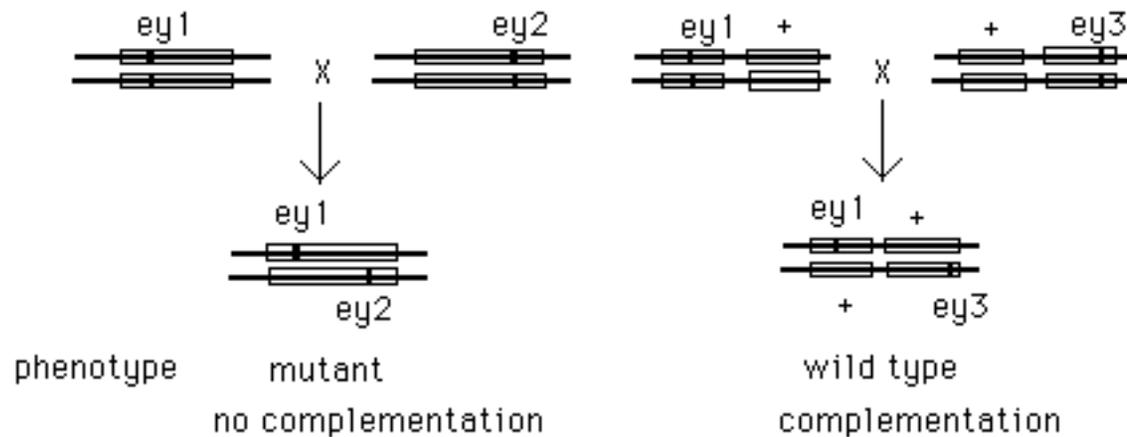
### 2. Complementation Test

**E.g. Drosophila: want to identify all genes required to make eyes. All expected to have eyeless phenotype or some kind of abnormal eye shape. Select many recessive *ey* mutants; don't know whether they represent 1, 2, or more genes. Do complementation test to see which ones are allelic and how many genes they represent.**

Suppose *ey1* and *ey2* are in same gene, but *ey3* is in a different gene.

*ey1 ey1* X *ey2 ey2* --> *ey1 ey2* heterozygote, look at phenotype.

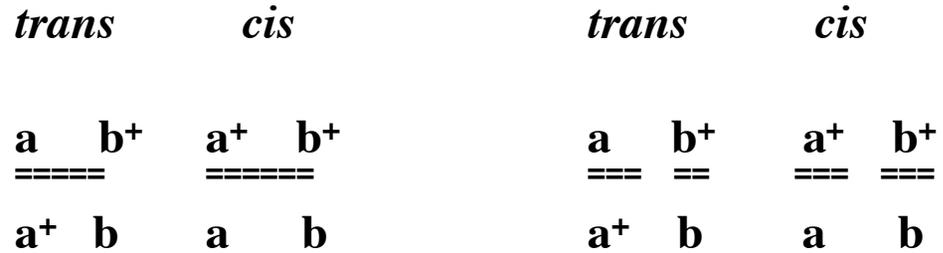
*ey1 ey1* X *ey3 ey3* --> *ey1 ey3* heterozygote, look at phenotype.



**For recessive genes, if make double mutant:**

mutations allelic (in same gene)	mutant	no complementation
mutations nonallelic (in different genes)	wild type	complementation

***Trans* means mutations are on different homologues; *cis* means are on same homologue:**



**What will be the phenotypes of the heterozygotes in the *cis* configuration?**

***Trans* means mutations are on different homologues; *cis* means are on same homologue:**

<b>trans</b>	<b>cis</b>	<b>trans</b>	<b>cis</b>
<u>a</u> <u>b<sup>+</sup></u>	<u>a<sup>+</sup></u> <u>b<sup>+</sup></u>	<u>a</u> <u>b<sup>+</sup></u>	<u>a<sup>+</sup></u> <u>b<sup>+</sup></u>
a <sup>+</sup> b	a   b	a <sup>+</sup> b	a   b



**Complementation test also called *cis-trans* test because must be done in *trans* to distinguish mutations in the same or different genes. In the *cis* configuration, will get wild type either way.**

**Genes defined by complementation are sometimes called cistrons.**

**By defining alleles, we are also defining genes:**

**Mutations that don't complement each other are said to be in same complementation group (or rarely, cistron). Each complementation group = one gene.**



**Complementation actually means that two genomes complement each other because each one has the wild type allele of one gene and the mutant allele for another.**

## **Mitotic Recombination**

**Recombination occurs during mitosis as well as meiosis. But although there are more than ten mitotic divisions and only one first meiotic divisions in a germ line, recombination is more frequent in meiosis when synapsis helps bring homologous sequences together.**

**Mitotic recombination occurs between any homologous sequences, which may be on the same chromosome, sister chromatids, or nonhomologous chromosomes as well as on homologues.**

**Will look at this during Discussion.**

