

Repetitive DNA

- Satellite DNA
- Minisatellite DNA
- Microsatellite DNA
- Transposable elements
- LINES, SINES and other retrosequences
- High copy number genes (e.g. ribosomal genes, histone genes)
- Multifamily member genes (e.g. hemoglobin, immunoglobulin)

Satellite DNA

Unit - 5-300 bp depending on species.

Repeat - 10^5 - 10^6 times.

Location - Generally heterochromatic.

Examples - Centromeric DNA, telomeric DNA. There are at least 10 distinct human types of satellite DNA. A single type may be more than 1% of the genome (equivalent to 3 entire *E. coli* genomes).

Human satellite DNA is prone to be multimeric or hierarchical in structure. Human α satellite DNA (centromeric) is typically 171 bp long present as dimers (342 bp) or up to 16'mers (2736 bp) as the repeating units. Generally less length variation than minisatellites or microsatellites.

Human β satellite DNA is present as 30,000 - 60,000 copies of a 68 bp monomer (2,040,000 - 4,080,000 bp) on the metacentric chromosome 9 and the acrocentric chromosomes 13, 14, 15, 21, and 22. It is a pericentromeric repeat in humans.

Examples of Satellites from *Drosophila virilus*.

Satellite	Primary Sequence	Copies per genome	Percent of genome
I	ACAAACT	1.1×10^7	25%
II	ATAAACT	3.6×10^6	8%
III	ACAAATT	3.6×10^6	8%
<hr/>			41%

Minisatellite DNA

Unit - 15-400 bp (average about 20).

Repeat - Generally 20-50 times (1000-5000 bp long).

Location - Generally euchromatic.

Examples - DNA fingerprints. Tandemly repeated but often in dispersed clusters. Also called VNTR's (variable number tandem repeats).

Human λ 33.1 minisatellite (62 bp)

AAGGGTGGGCAGGAAGTGGAGTGTGTGCCTG
CTTCCCTTCCCTGTCTTGTCCCTGGAAACTCA

Human λ 33.5 minisatellite (17 bp)

YGGCAGGAGGGGAGG

Microsatellite DNA

Unit - 2-4 bp (most 2).

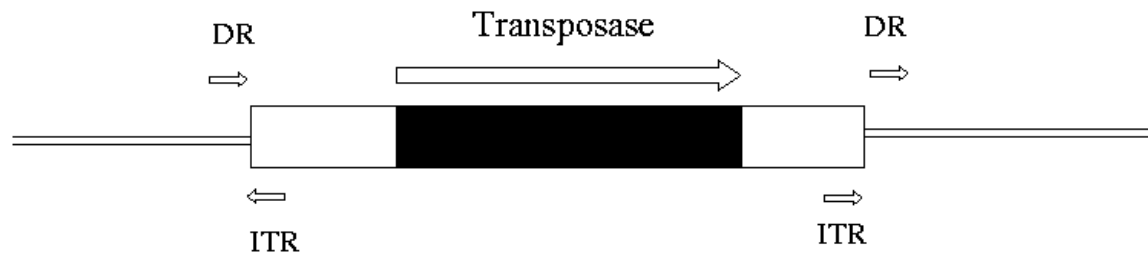
Repeat - on the order of 10-100 times.

Location - Generally euchromatic.

Examples - Most useful marker for population level studies. This example is from a water snake ...

```
...TCCAGACAAGGTGGTGTGTGTGTGTGTGTG  
TGTGTGTGTGTGTTTCTCCAGTGAGATTTA...
```

Minimal structure of a transposable element



Transposable elements in eukaryotes: a few examples

Maize - Ac-Ds = Activator (encodes a transposase), Dissociation (encodes an enzyme that promotes chromosome breakage).

Drosophila melanogaster - P-element = most famous because of its use as a vector to insert foreign DNA into Drosophila. Causes hybrid dysgenesis when crossed between strains.

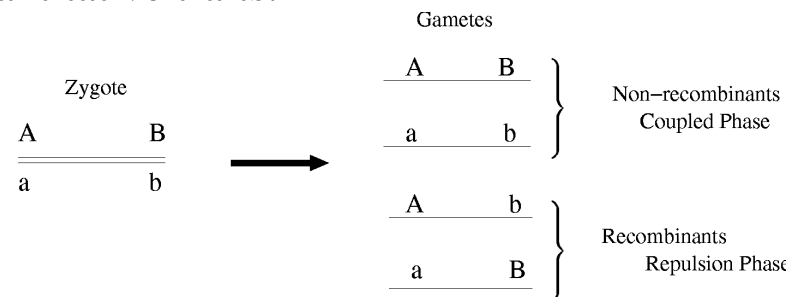
Many organisms - Tc1/Mariner elements. The Mariner element family is exceptionally widespread in animals (from nematodes to mice) and are particularly common in insects (humans have a mariner related element but it is not active).

Repetitive DNA that makes use of an RNA intermediate

	Reverse Transcriptase	Transposition	Presence of LTR's	Virion Particles	Example
Retron	yes	no	no	no	rev.trans.gene
Retroposon	yes	yes	no	no	LINES
Retrotransposon	yes	yes	yes	no	copia
Retrovirus	yes	yes	yes	yes	HIV
Pararetrovirus	yes	no	yes	yes	hepadnaviruses
Retrosequences	no	no	no	no	SINES

Linkage

Linkage Disequilibrium: alleles at two (or more) loci are present together more (or less) often than predicted by their frequencies. A bad example, but one that is familiar to everyone is blond hair and blue eyes. These two traits appear together more often than predicted by a chance combination of their frequencies. This is a bad example since these traits are multilocus, quantitative traits.



Physical distance and probability of recombination are positively correlated.

Probability of recombination is however not a simple function of distance. For example, the probability of a recombination event occurring over 50 kb of heterochromatin is much less than the equivalent probability over 50 kb of euchromatin.

Chiasmata are the physical manifestation of recombination.

If more than one chiasmata form between two markers, there will be recombination events between the markers but they will appear as non-recombinant.

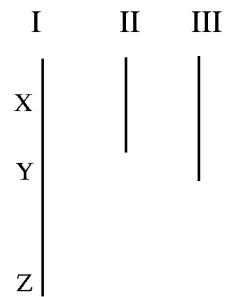
As the distance between markers becomes large, the probability of recombination will approach $1/2$.

If the markers are on different chromosomes, the probability of apparent recombination is also $1/2$. Thus, an apparent lack of linkage does not mean the markers are or are not on the same chromosome.

3I03 - Eukaryotic Genetics

Linkage groups are a collection of genetic markers that each segregates non-randomly with one or more of the other markers of the group.

Ultimately a linkage group is equivalent to a chromosome.



X, Y and Z are in linkage group I, if X & Y are linked and Y & Z are linked even though X & Z may be apparently unlinked.

Linkage Maps:

A recombinant frequency (RF) of 1 percent is defined as 1 centimorgan (1 cM).

If you observe A — B recombine with an apparent distance of 5 cM and that A — C recombine with an apparent distance of 3 cM then ...

$$\begin{array}{l}
 C \xrightarrow{3} A \xrightarrow{5} B \\
 A \xrightarrow{3} C \xrightarrow{2} B \\
 B \xrightarrow{5} A \xrightarrow{3} C \\
 B \xrightarrow{2} C \xrightarrow{3} A
 \end{array}
 \left. \vphantom{\begin{array}{l} C \\ A \\ B \\ B \end{array}} \right\} \text{hence only a} \\
 \text{relative order}$$

A test of the recombination rate between B and C will tell the order but still not the direction.

Multiple recombinants will reduce the apparent linkage between distant markers. Hence the best estimate is the sum of the distances between the closest markers.

A ——— B ——— C

but

A ——— C

Even when using the sum of the distances between the closest markers it should be corrected for the possibility of multiple recombinants.

A ————— B is measured as 0.3 cM

B ————— C is measured as 0.4 cM

but

A ————— C is measured as 0.68 cM

As markers become closer the distances might deviate significantly from additivity. This phenomena is called “interference”.

This is due to the fact that a single chiasmata will generally reduce the probability of another chiasmata forming nearby. (It is also possible (though rare) to get positive interference).

A genetic map \neq a physical map.

Physical mapping can be done via

- cytologically
- FISH
- Deletion mapping
- Interspecies somatic cell hybrids
- Pulsed Field Gel Electrophoresis

Genetic Markers Used to Map Genomes

Traditional Markers

- Genes that cause a visible morphological difference
- Genes that cause a detectable biochemical difference

Molecular Markers

- Restriction Fragment Length Polymorphism (RFLPs)
- Simple Sequence Length Polymorphisms
 - Minisatellites (VNTRs)
 - Microsatellites (micros)
- Single nucleotide polymorphisms (SNPs)
- Expressed Sequence Tags (ESTs)
- Sequence Tagged Sites (STSs)

Genome Mapping

- YACs
- BACs
- Automated Sequencing
 - directed
 - shotgun
- Contigs

Using the genome map to find genes:

- Reverse the process. Use the genome sequence to generate the markers to follow in test crosses.
- Bioinformatics
- QTL / 'complex trait' mapping
 - cross inbred lines that differ in some trait
 - isolate lines with small introgressed regions as identified by the genome markers