Gene 422

QTL detection and QTL-mapping

Detecting associations between markers and genes (QTL)

- QTL = Quantitative Trait Locus / Loci
- parts of the genome that effect quantitative traits (“genes”)
- We want to know the position of QTL and their effect

SOME IDENTIFIED QTL

- Genestar Marbling
- Genestar Tenderness
- Carwell
- Footrot
- Callipyge Gene
- Booroola Gene
- Inverdale Gene
- Hanna Gene
- Woodland Gene
- Double Muscling Gene
- DGAT Milk Fat%

commercialized

Fecundity

QTL detection with gene markers

Genetic markers

Principle of mapping

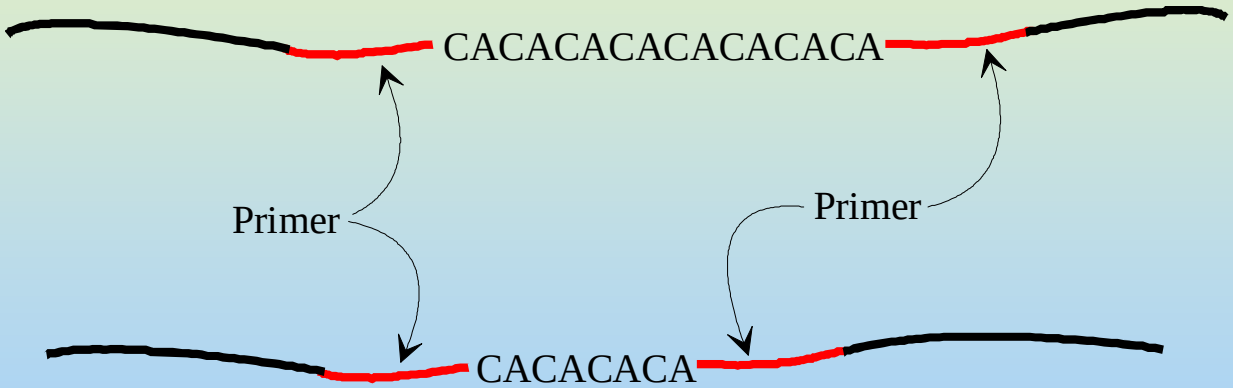
Associations between markers and (quantitative) traits

QTL detection and mapping

Genetic markers

Microsatellites

Microsatellites are DNA regions with variable numbers of short tandem repeats in nucleotide sequence (CA repeats in the diagram here). Each microsatellite locus is recognised or targeted by its primers - the two unique sequences adjacent to the repeating region.



Microsatellites make good genetic markers because they each have many different 'alleles' - ie. there can be many different lengths of the repeat region. With many alleles, most individuals are heterozygous, giving power to note association between marker allele and performance in progeny inheriting a favourable linked QTL allele.

Mapping: Recombination

Parent animal

$$\frac{A \quad B}{a \quad b}$$

Non-recombinant gametes Frequency: (1 – r)	Recombinant gametes Frequency: r
AB	Ab
ab	aB

Classical Mapping: Using a test cross

Parent 1 AABB x aabb Parent 2

F1 AaBb (100%)

F1-gametes	**AB**	**Ab**	**aB**	**ab**
frequencies				
A and B are unlinked:	**25**	**25**	**25**	**25**
A and B linked:	**35**	**15**	**15**	**35**
A and B tightly linked	**48**	**2**	**2**	**48**

Classical Mapping: Using a test cross

Parent 1 AABB x aabb Parent 2

F1 AaBb (100%)

F1-gametes	**AB**	**Ab**	**aB**	**ab**
frequencies				
A and B are unlinked:	**25**	**25**	**25**	**25**
A and B linked:	**35**	**15**	**15**	**35**
A and B tightly linked	**48**	**2**	**2**	**48**
		R	**R**	

R = Recombinants

Classical Mapping: Using a test cross

Parent 1 AABB x aabb Parent 2

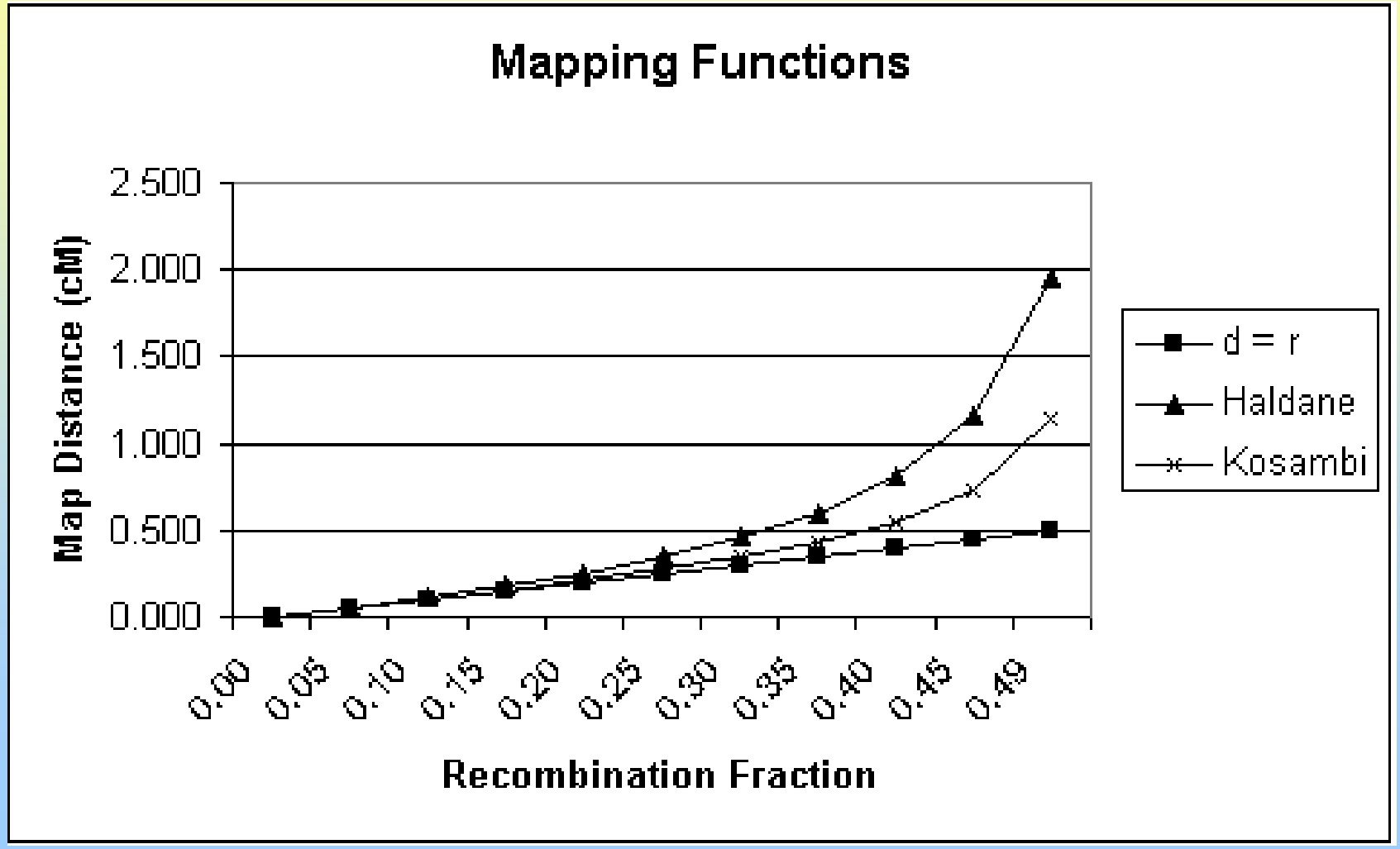
F1 AaBb (100%)

F1-gametes	**AB**	**Ab**	**aB**	**ab**	
frequencies					
A and B are unlinked:	**25**	**25**	**25**	**25**	**50**
A and B linked:	**35**	**15**	**15**	**35**	**30**
A and B tightly linked	**48**	**2**	**2**	**48**	**4**
					% Recomb. Frequency

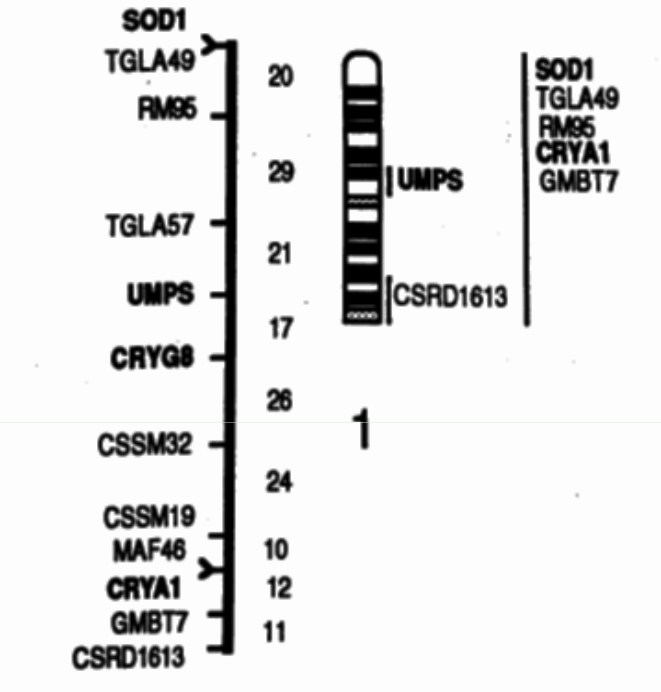
Genetic linkage maps

- Genetic distances determined according to recombination fraction between each pair of loci.
- The unit of measure is Morgans (or centimorgans, cM)
- A mapping function translates a genetic distance (cM) into recombination fractions

Mapping functions



Maps



On the left is the linkage map derived from studies on co-inheritance of alleles. Most loci are microsatelites or RFLPs. Distances between loci are given in units of centimorgans (100 units is equivalent to a 50% recombination fraction).

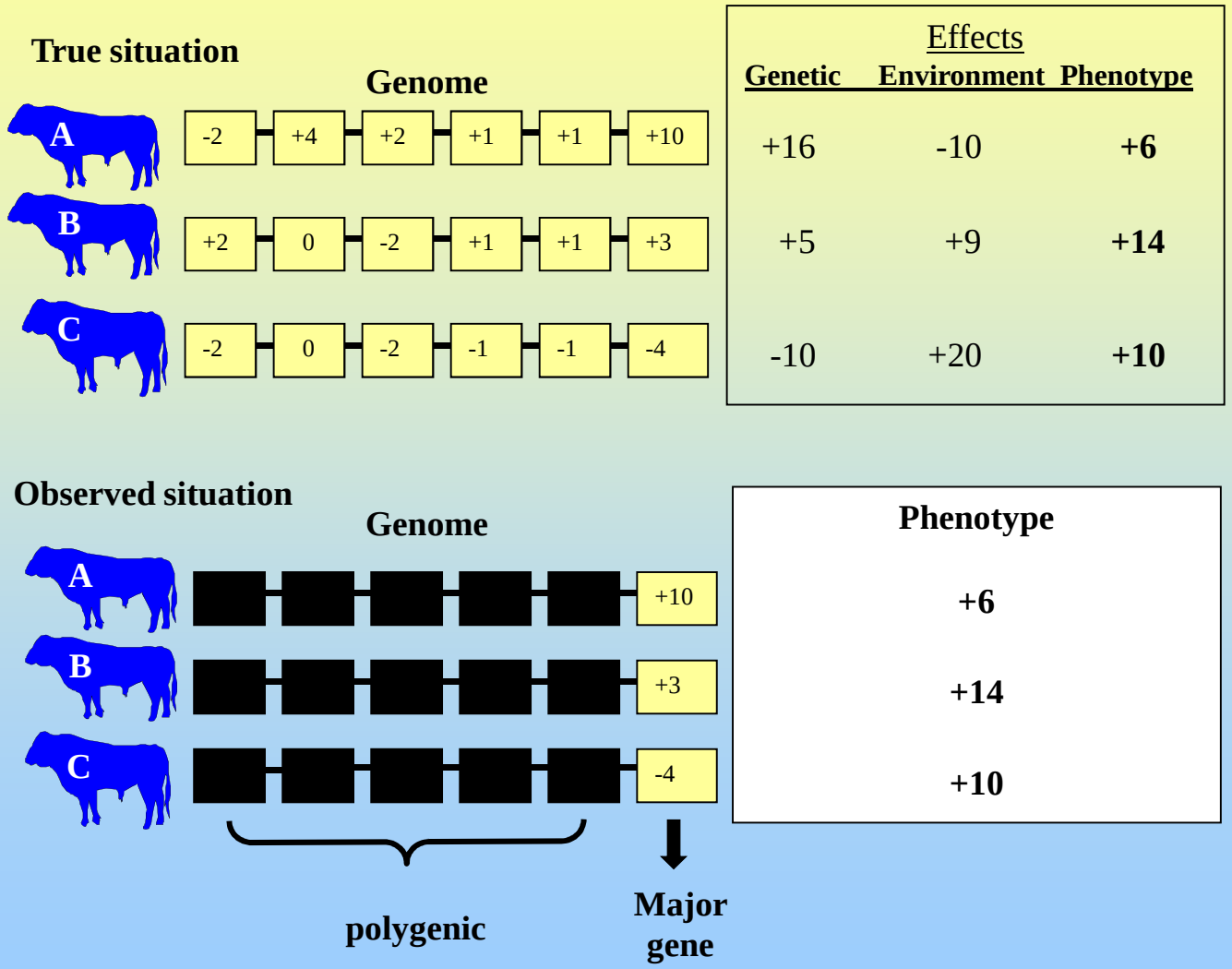
In the center is a representation of the chromosome with its visible banding regions. Some loci have been localised through, for example, observing fluorescing DNA from the gene of interest 'sticking' to particular regions.

On the right is a list of loci associated with chromosome 1 through work with somatic cell hybrid lines. For example, alleles at these loci might all be present in a cell line containing only chromosome 1 from cattle, the rest from, say, hampsters.

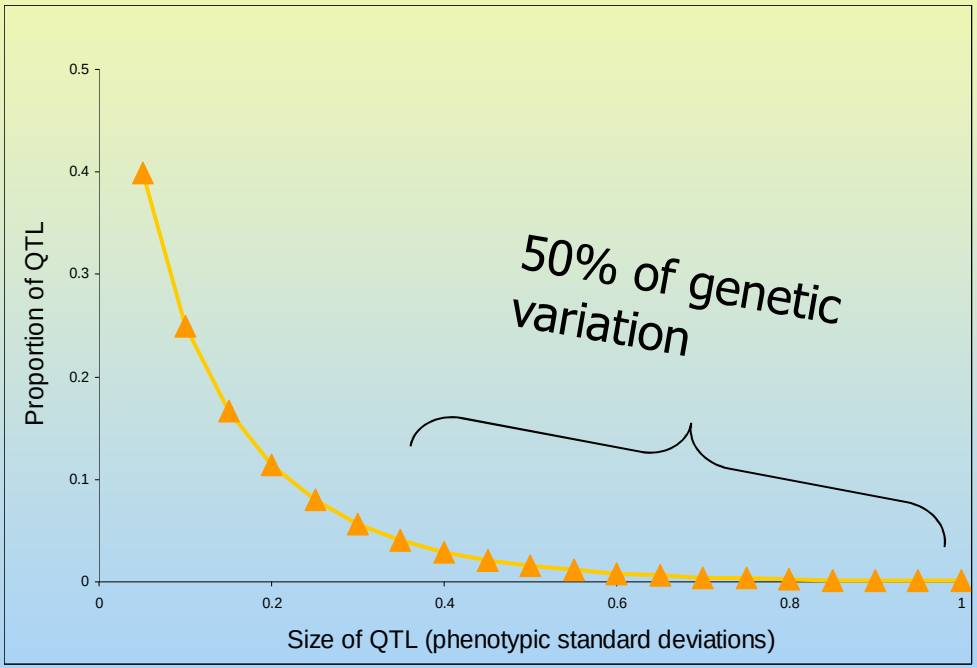
Detecting associations between markers and genes (QTL)

- QTL = Quantitative Trait Loci
- parts of the genome that effect quantitative traits (“genes”)
- We want to know the position of QTL and their effect (of allele substitution)

Polygenes and major genes.



The distribution of QTL effects

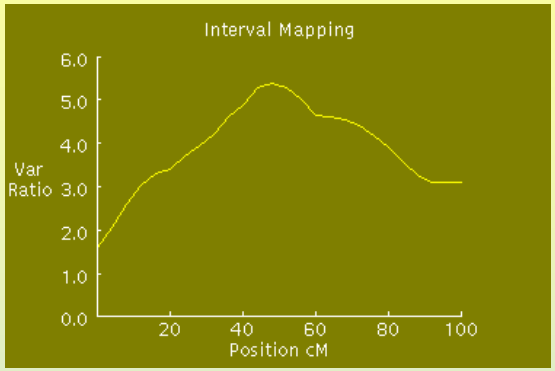


Many small QTL, few of large effect

- Maybe 5-10 large QTL explain the majority of the genetic variance.
- Mapping experiments should be able to detect QTL as small as 0.2σp?

Gene discovery by mapping

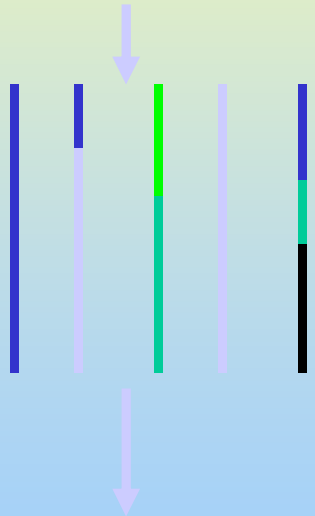
Dairy Cattle: DGAT1 - A success story



1. Linkage mapping **detects a QTL** on bovine chromosome 14 with large effect on fat % (Georges et al1995)

2. Linkage disequilibrium mapping

   refines position of QTL (Riquet et al. 1999)

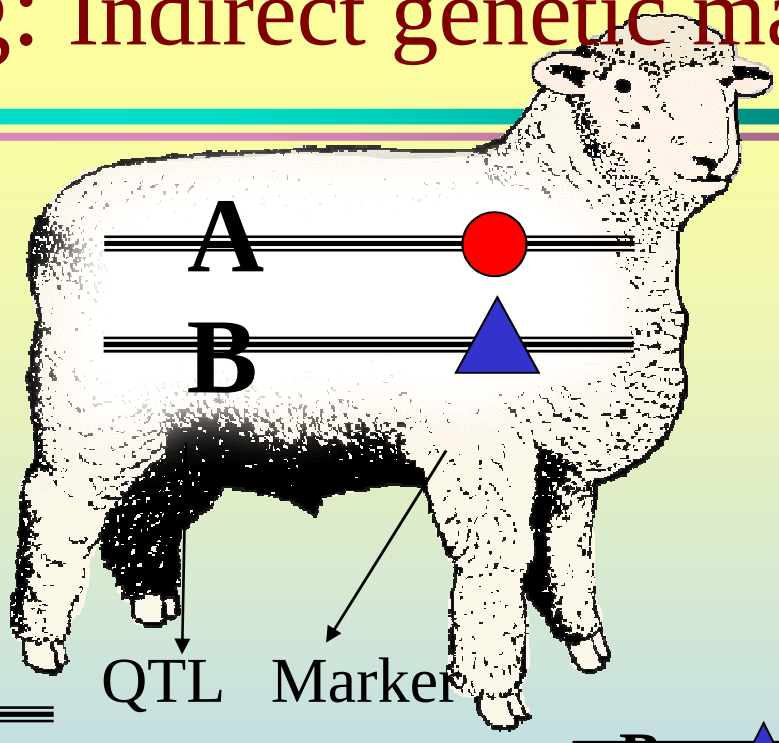


3. Selection of candidate genes.
Sequencing **reveals point mutation** in candidate (DGAT1). This mutation found to be functional - substitution of lysine for analine. Gene patented. (Grisart et al. 2002)

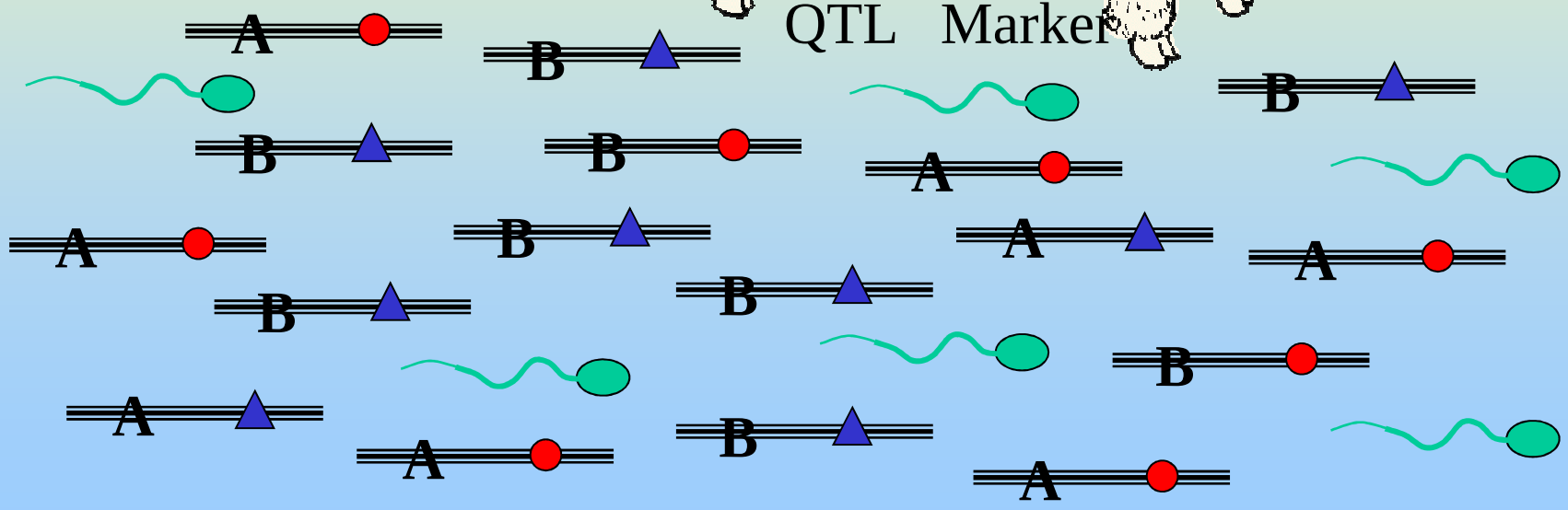
ACCTGGGAG
ACCAGGGAG

Linkage Mapping: Indirect genetic markers

A Ram:

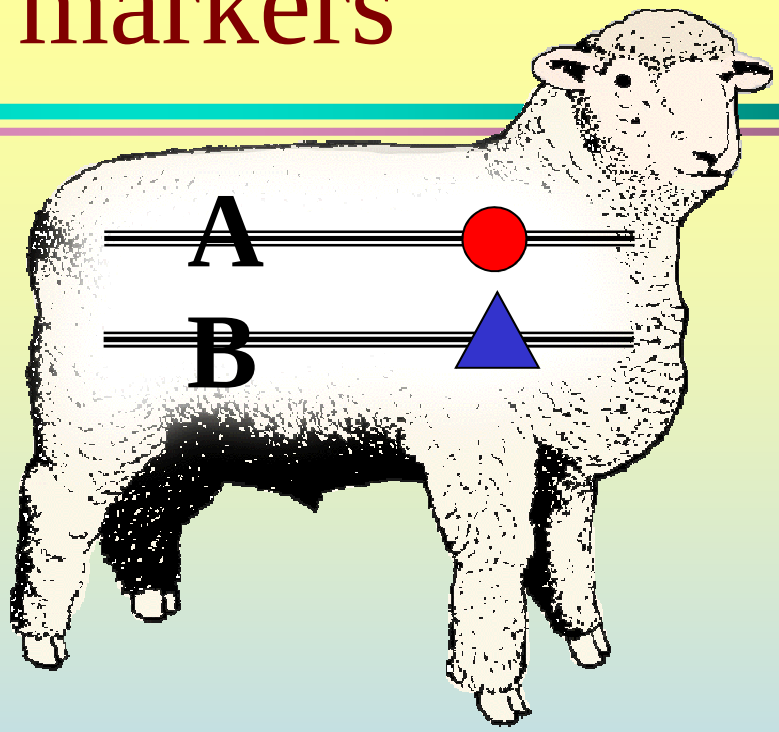


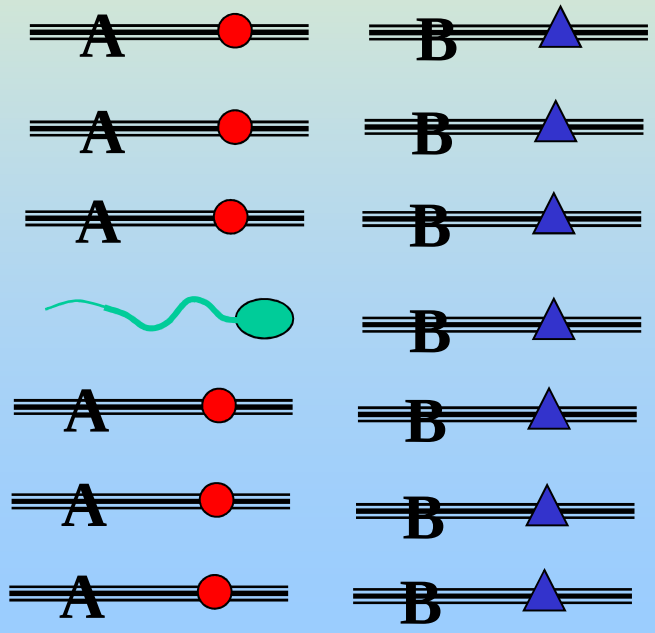
His semen:

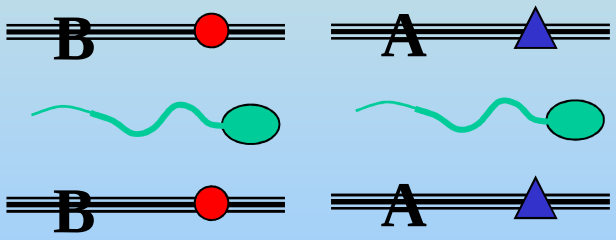


Indirect genetic markers

Order the offspring by marker allele received ...



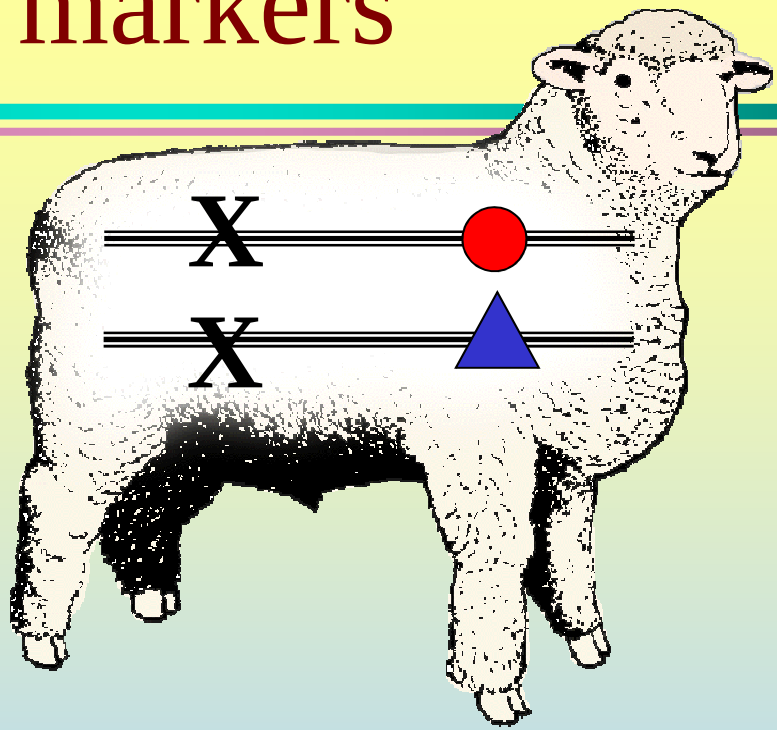


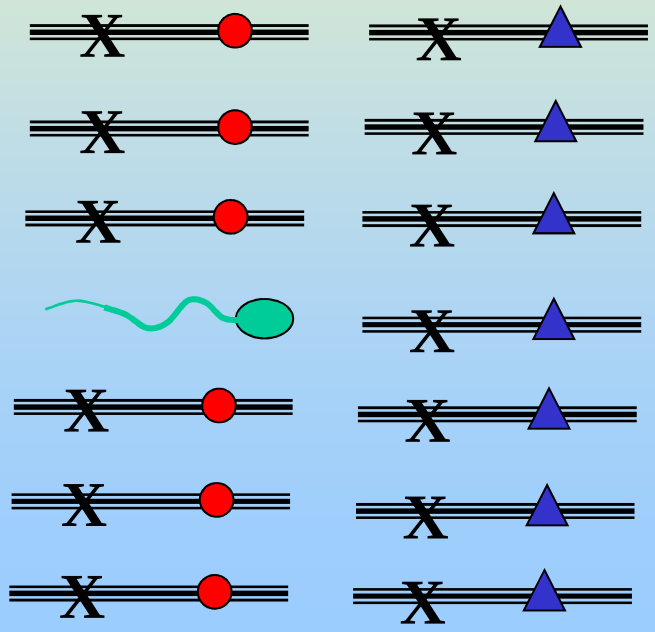


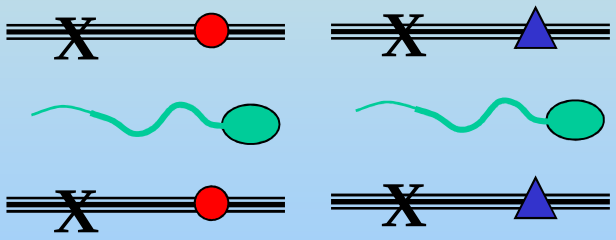
‘recombinants’

Indirect genetic markers

In reality,
We don't see the QTL ...



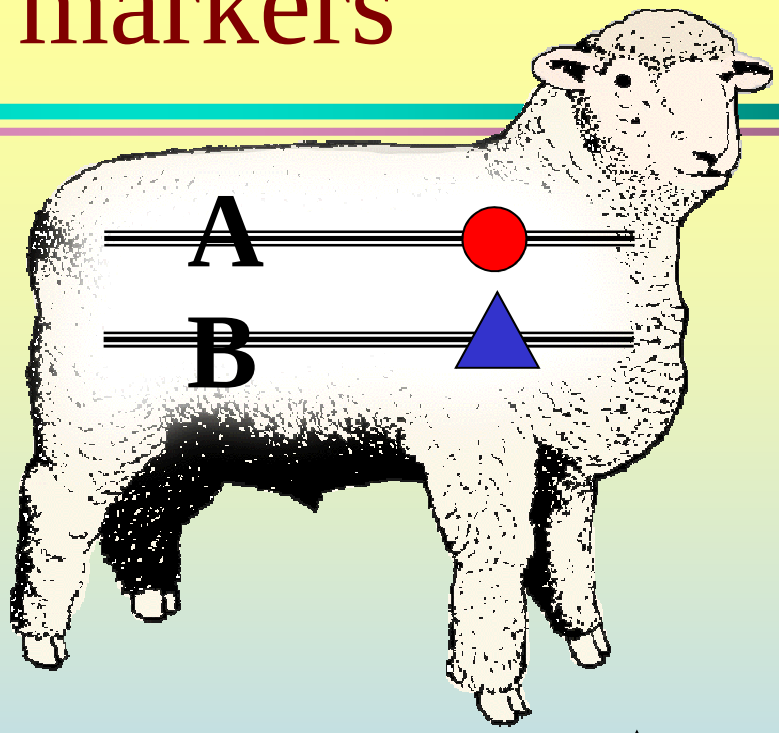


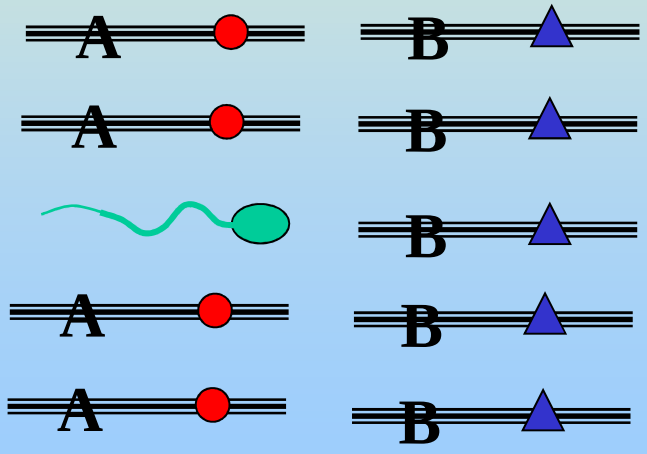


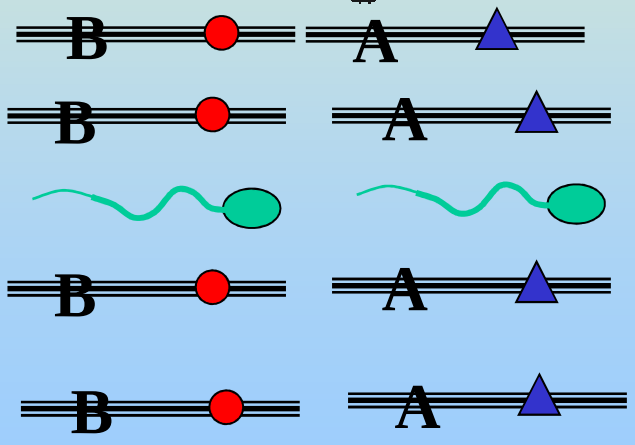
'recombinants'

Indirect genetic markers

If the marker is not linked to the QTL ...



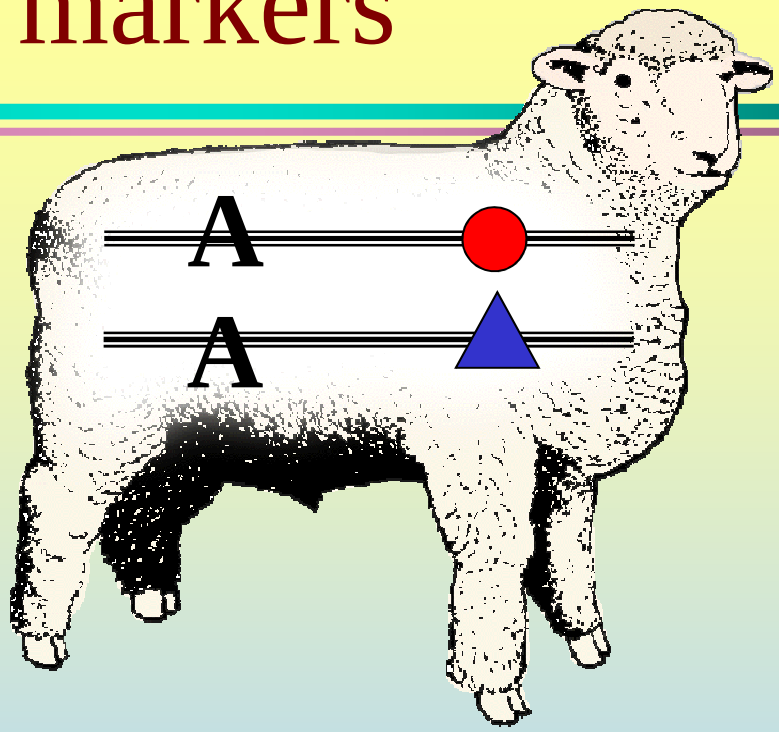


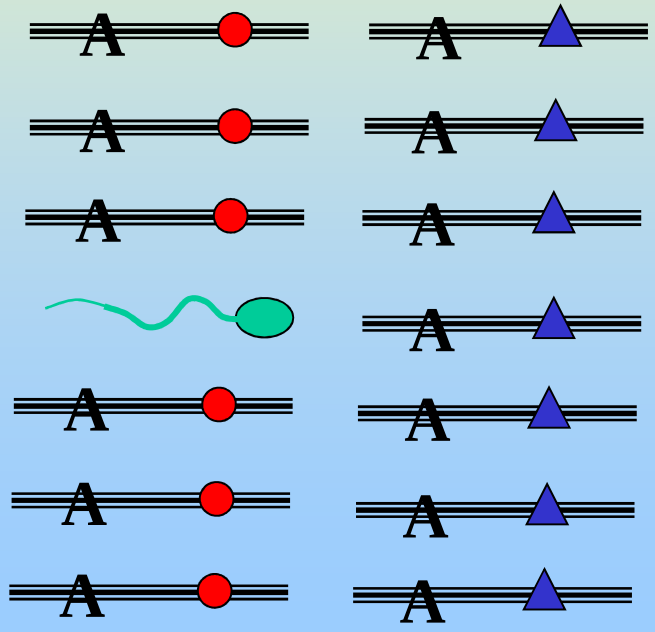


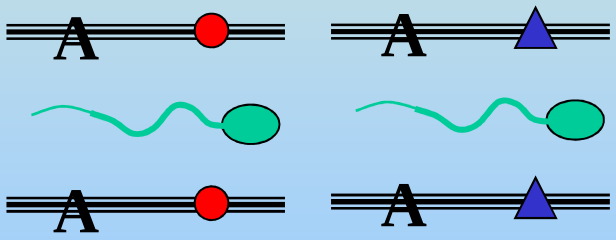
'recombinants'

Indirect genetic markers

If ram is not heterozygote……



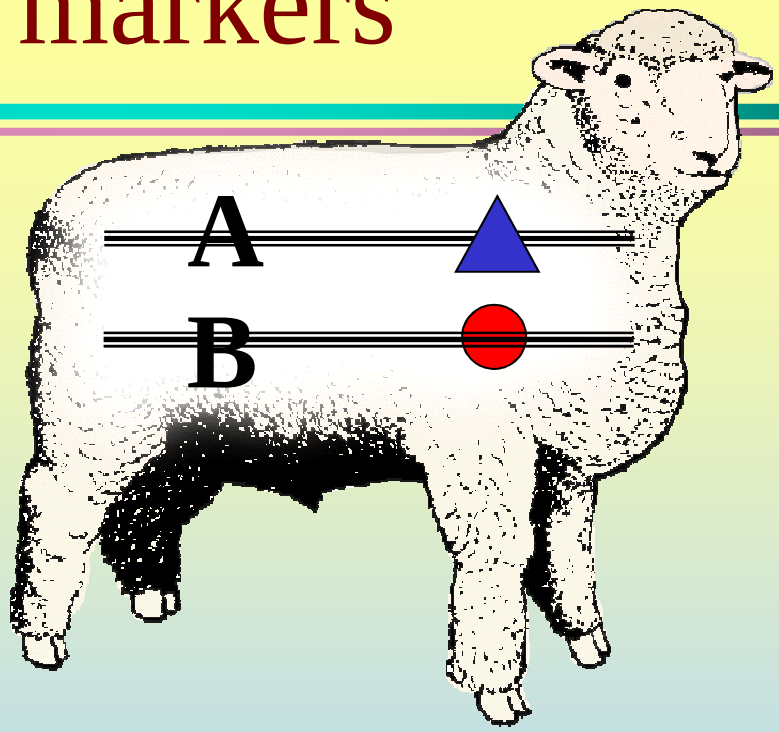


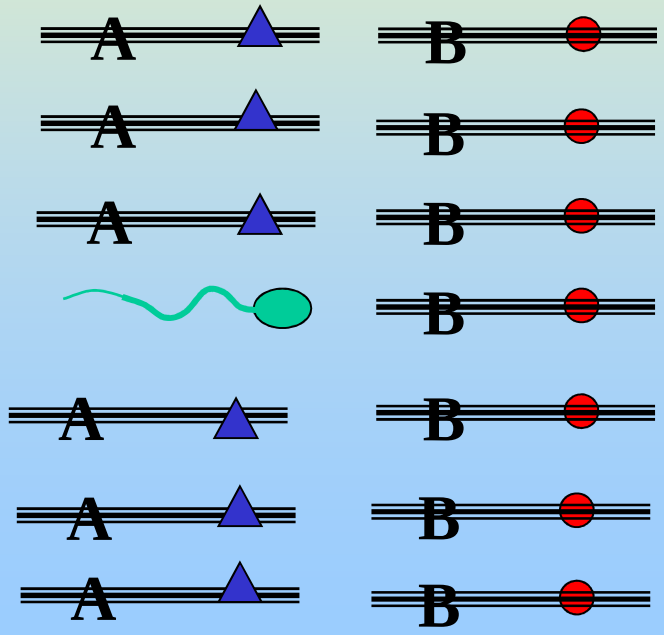


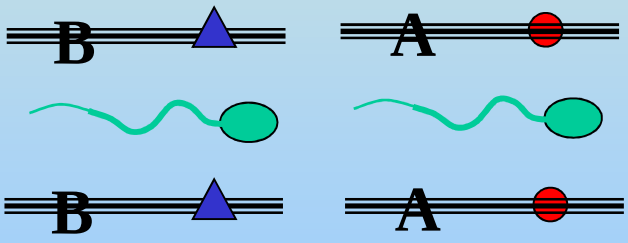
‘recombinants’

Indirect genetic markers

Some rams may have it the other way around…….

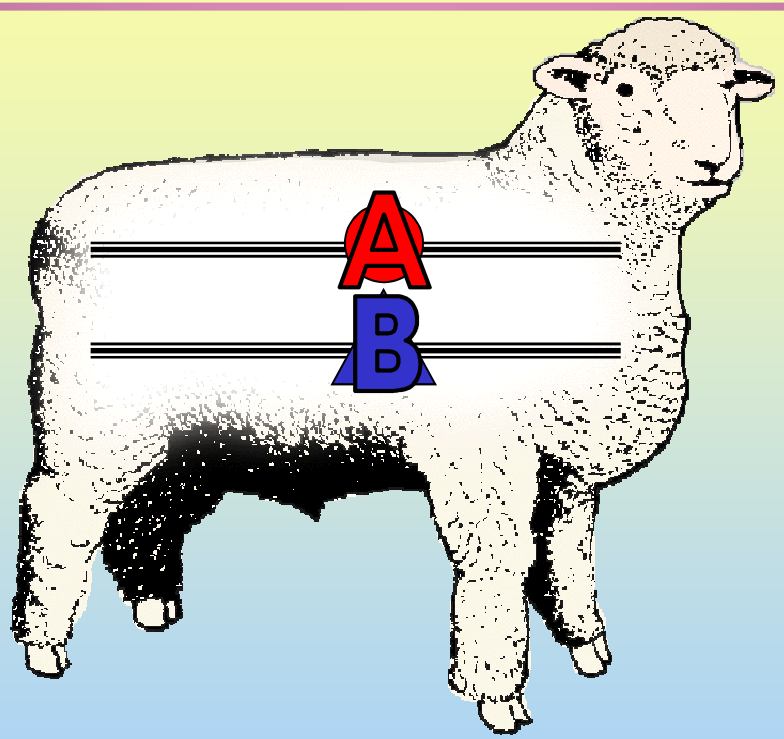






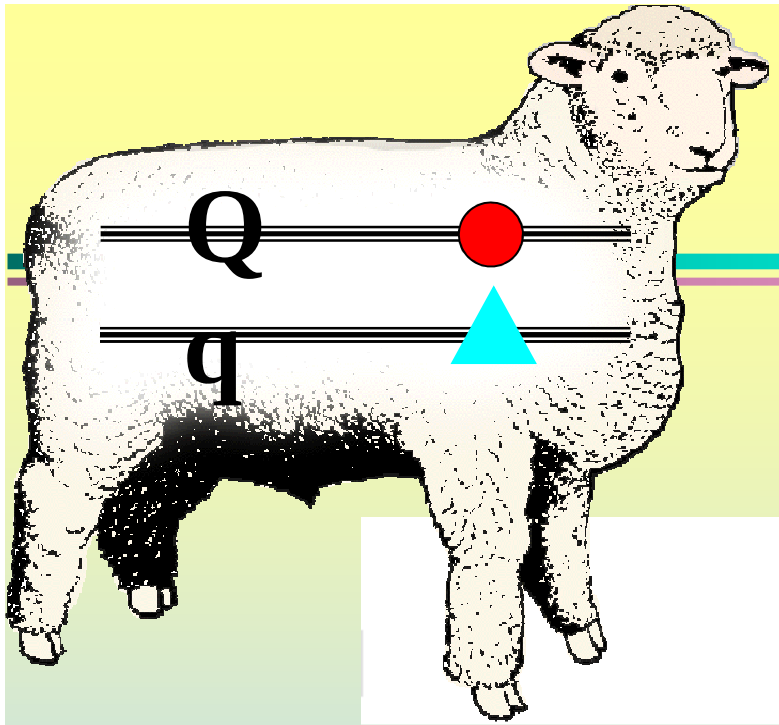
‘recombinants’

Direct genetic markers



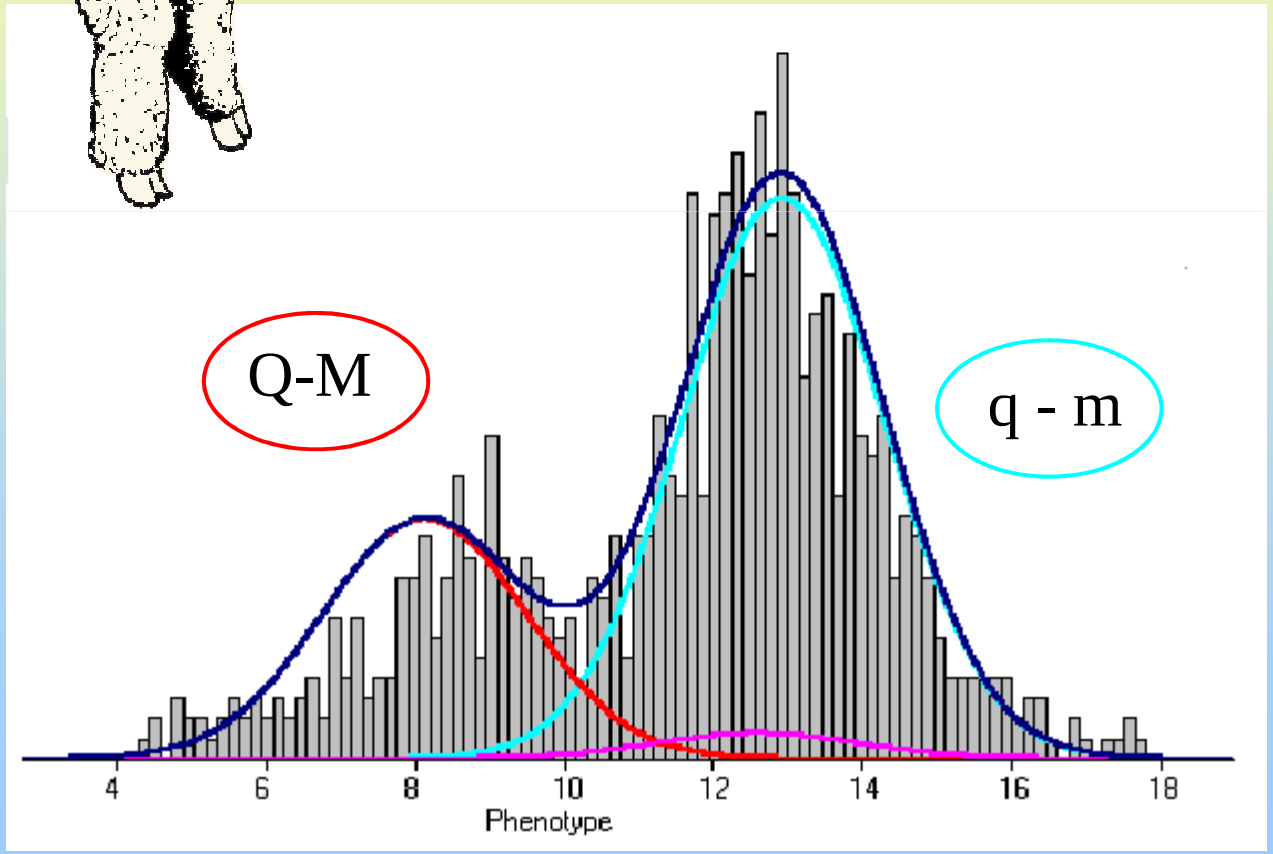
A - always circle, always good

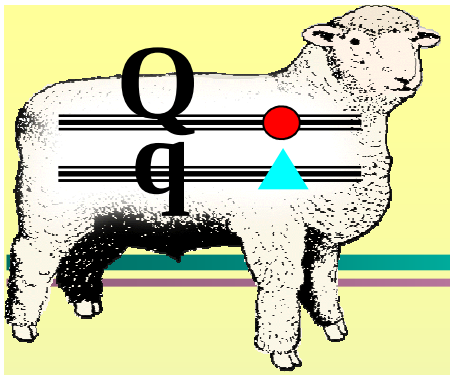
B - always triangle, always bad



Basic principle

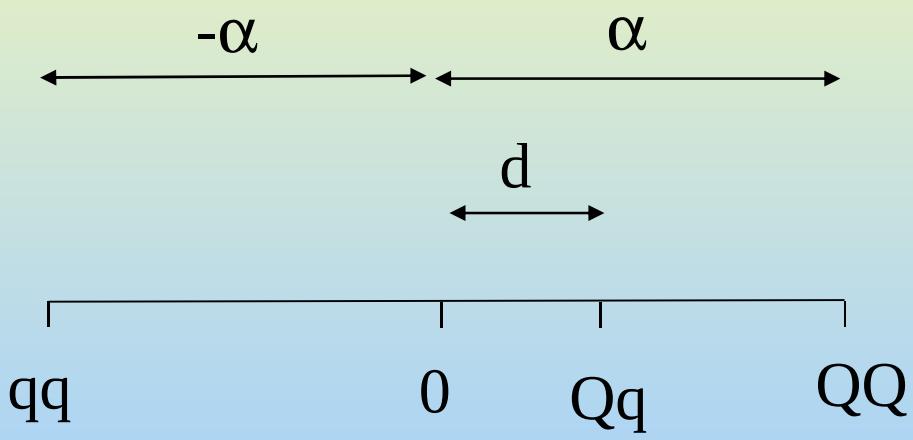
Distribution of the phenotypic values of offspring of sire





Some terminology

QTL effects



α = allele substitution effect

d = dominance effect of the QTL

Determining associations between genetic markers and QTL with single markers

recombination rate / *frequency* between marker allele and QTL allele is r.

Parental genotype

$$\frac{A1 \quad Q}{A2 \quad q}$$

Possible gametes		Recombination?	Gamete probability
A1	Q	no	(1-r)/2
A2	Q	yes	r/2
A1	q	yes	r/2
A2	q	no	(1-r)/2

Note that probabilities hold only for this parental genotype.

Expected progeny means of each marker group

– The difference between progeny receiving Q vs. progeny receiving q from their sire is the **allele-substitution effect:** **α.**

Marker allele from sire	QTL allele from sire	Frequency	Expected mean
A1	Q	(1-r)/2	$\mu + \alpha$
A1	q	r/2	μ
A2	Q	r/2	$\mu + \alpha$
A2	q	(1-r)/2	μ

Marker allele from sire	QTL allele from sire	Frequency	Expected mean
A1	Q	(1-r)/2	$\mu + \alpha$
A1	q	r/2	μ
A2	Q	r/2	$\mu + \alpha$
A2	q	(1-r)/2	μ

Expected difference between A1-group and A2-group:

A1-group:	$((\mu + \alpha)(1-r)/2 + (r/2).\mu) / 0.5 =$	$\mu + (1-r)\,\alpha$
A2-group	$((\mu + \alpha)r/2 + ((1-r)/2)\,\mu) / 0.5 =$	$\mu + r\,\alpha$
Difference (D)		$(1-2r)\alpha$

Note that the difference is affected by both r and α

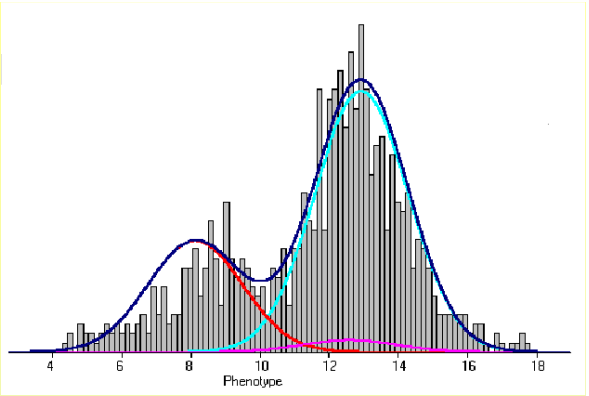
Some possible outcomes

$D = (1 - 2r)\alpha$

Recombination rate M-Q r	QTL-effect Q-q substitution α	Mean of progeny[1] receiving A1-allele	Mean of progeny[1] receiving A2-allele	Marker allele contrast D
0	50	50	0	50
0.1	62.5	56.25	6.25	50
0.2	83	67	17	50
0.3	125	87.5	37.5	50
0.4	250	150	100	50
0.5	1000	500	500	0

[1] Mean is relative to the progeny receiving a q-allele from the sire.

Rehash

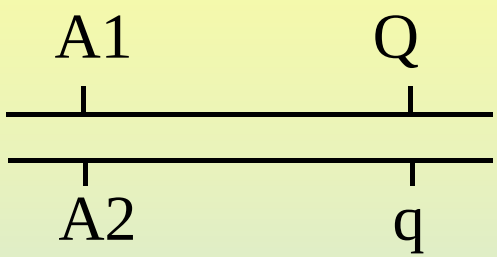


- The presence of a QTL is tested by determining whether the marker classes differ in their trait means
- The difference depends on the recombination fraction and the QTL effect
- Formally, but simply, you would test this with a t-test

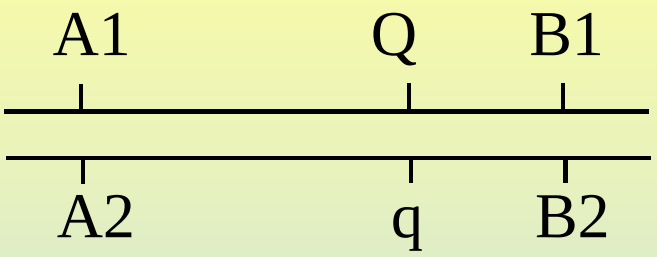
Analysis of trait differences for single markers

- If there is a (significant) marker effect then
  - the marker is linked to a QTL
  - there could be a large QTL far away from marker
    - or
  - can be a smaller QTL but closer to marker

Single vs multiple markers



Single markers: not possible to distinguish between QTL effect and QTL position



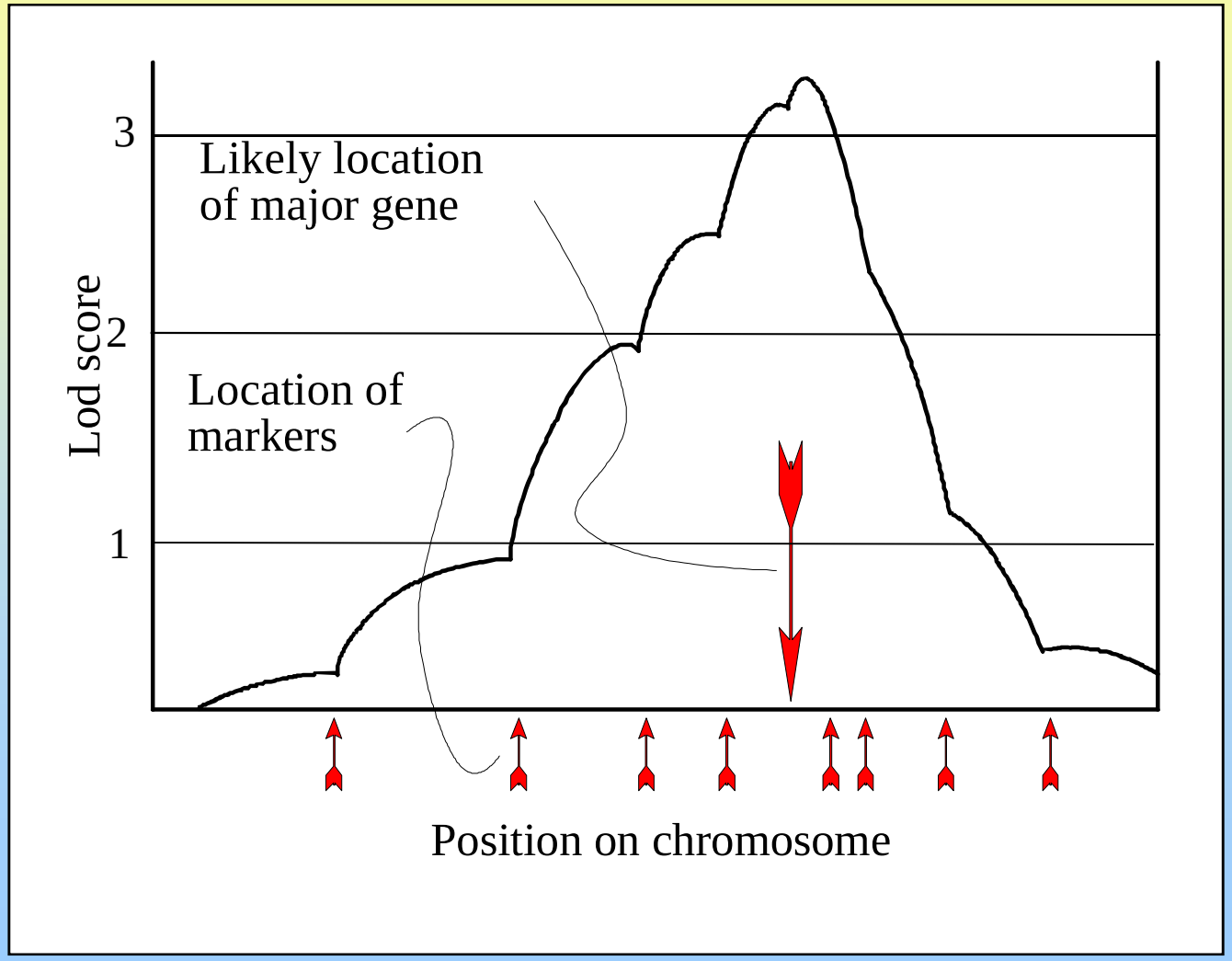
Two (or more) markers: no confounding between QTL effect and QTL position

Proper mapping of a QTL requires the use of multiple marker genotypes

Interval mapping

- Testing pairs of markers
- Many pairs to cover ideally the whole genome
- The position with the highest statistical evidence is the estimated QTL position
- ….or there can be more than one QTL
- Increased power of detection

QTL detection with markers



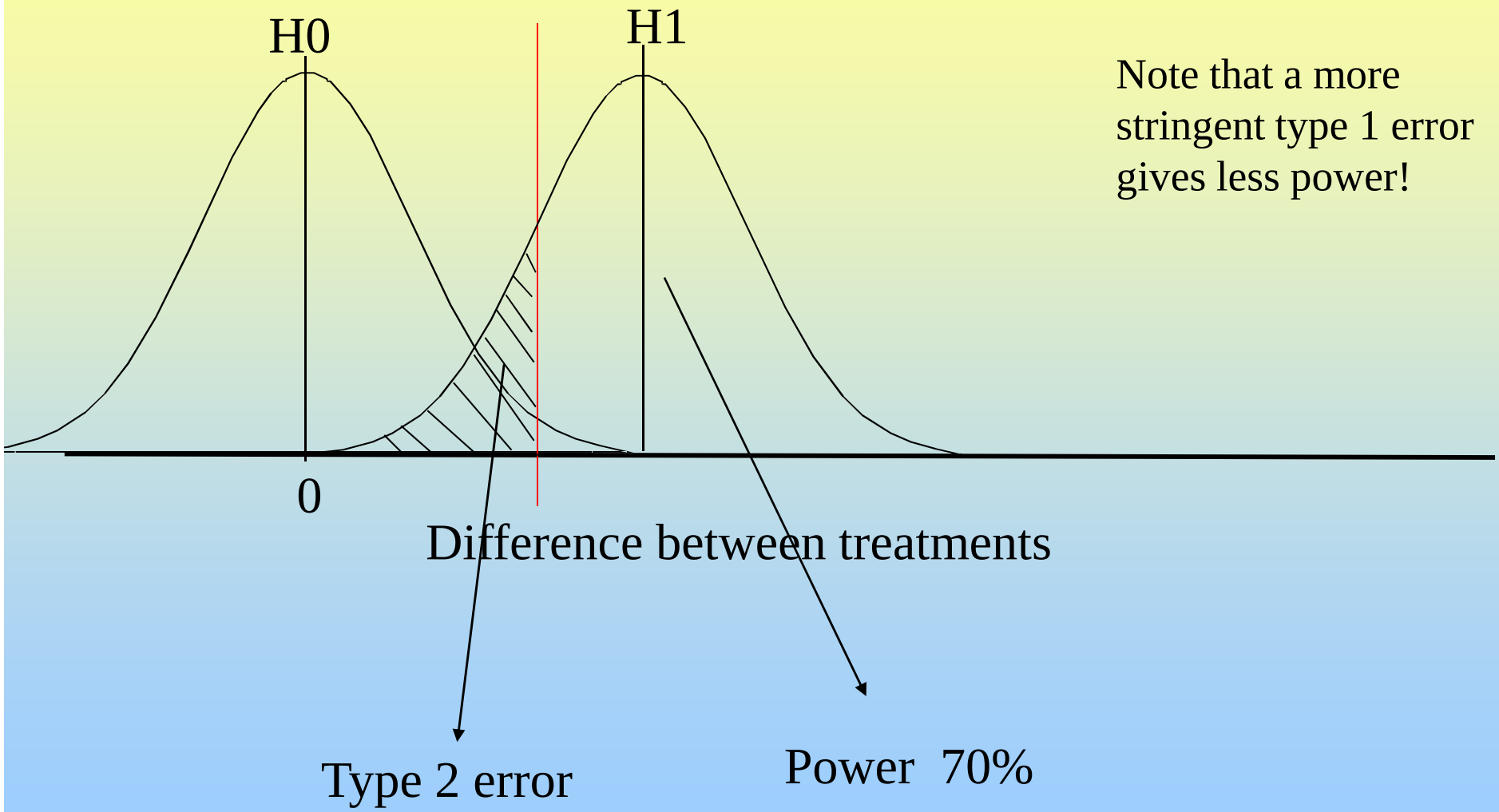
Other mapping methods

- Composite interval mapping
- Joint mapping
- Variance component models

Statistical testing issues

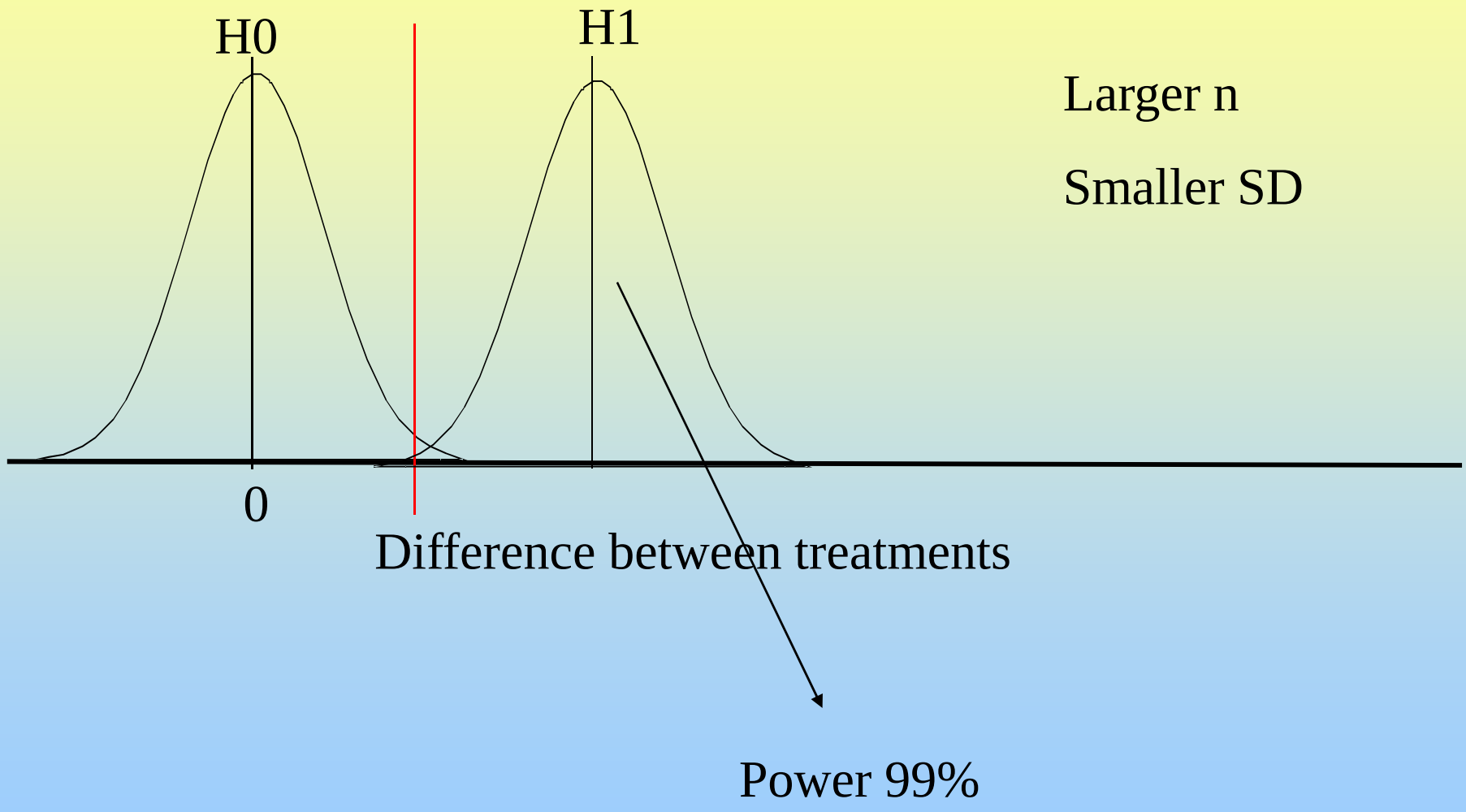
- Power
- Information content
- Significance levels/ Multiple Testing
  - Rule of thumb
  - Bonferroni correction
  - Permutation testing
- Confidence interval
  - 1 LOD drop off
  - Bootstrap

How many animals do we need to measure?



Note that a more stringent type 1 error gives less power!

How many animals do we need to measure?



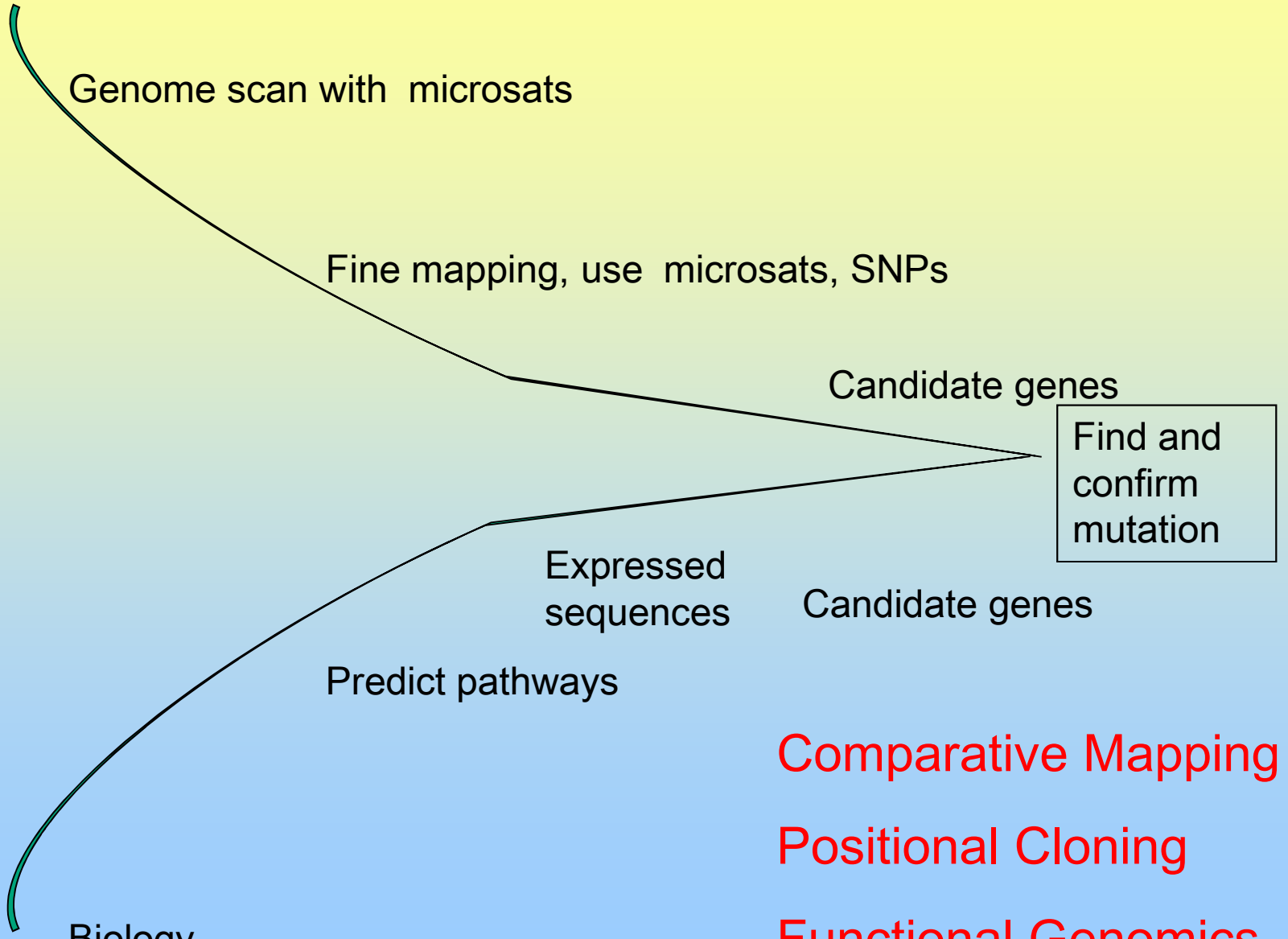
How many offspring needed per ram?

Power of QTL detection?

In phenotypic SD	Power 80%	95%
difference	nr.offspring	
0.4	199	330
0.5	127	211
0.6	88	147
0.7	65	108
0.8	50	82
1	32	53

- Some loss due to recombinations
- Loss due to uninformative progeny
- Need a stricter type 1 error
- Can save genotyping costs by
  - > selective genotyping
    (100%P, 50%G, equal power)
  - > DNA pooling

Pathways for gene discovery



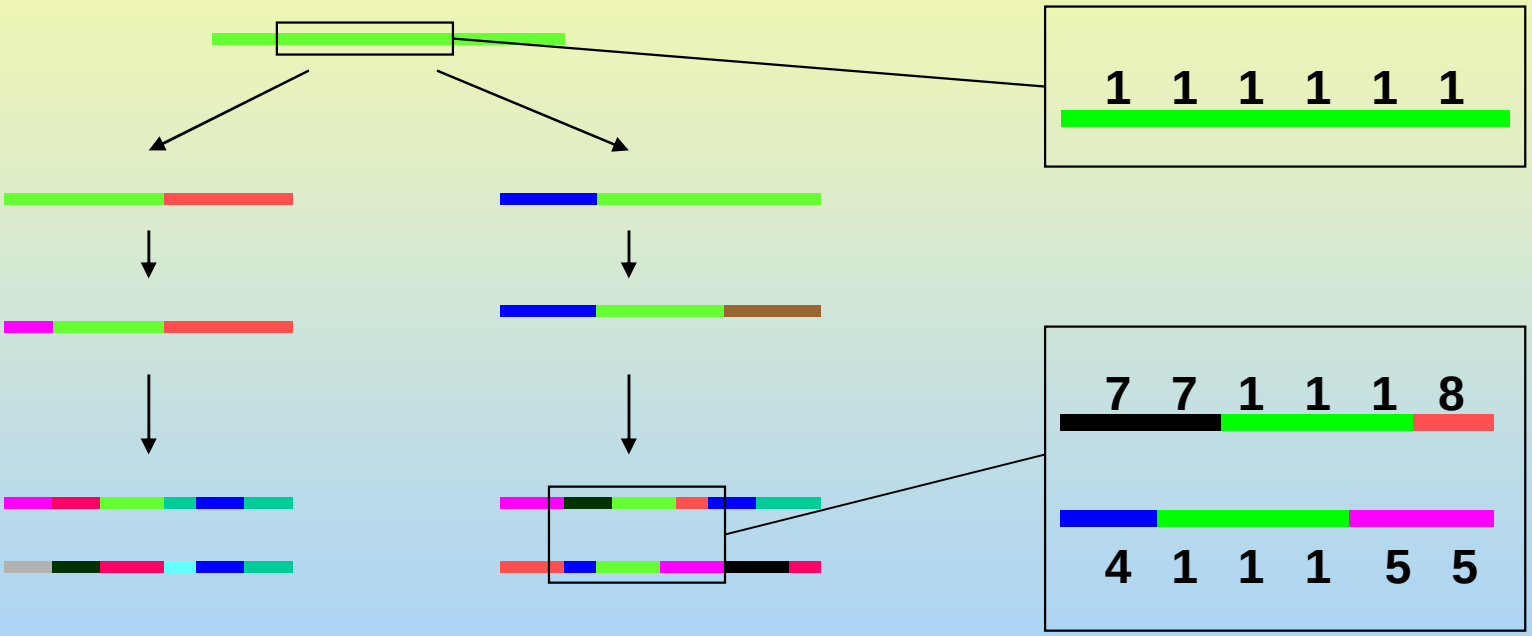
Comparative Mapping

Positional Cloning

Functional Genomics

New generation mapping based on dense SNP markers

LD and IBD probability between haplotypes



IBD depends on # markers in common L+R from putative QTL position

Markers are dense (e.g. 1 cM intervals)

IBD = identity by descent