The association between a marker and a QTL

sire

Gene location ⟶ +10 | | -4

Marker location ⟶ M | | m

sire   dam

+10   -4   q   q

Recombination rate = 10%

M   m   m   m

X

progeny

+10   q   -4   q   -4   q   +10   q

M   m   m   m   M   m   m   m

90%   10%
1-r   **r**
parental   recombinants
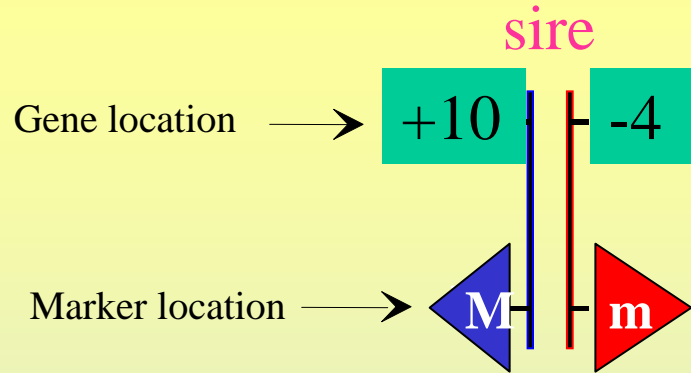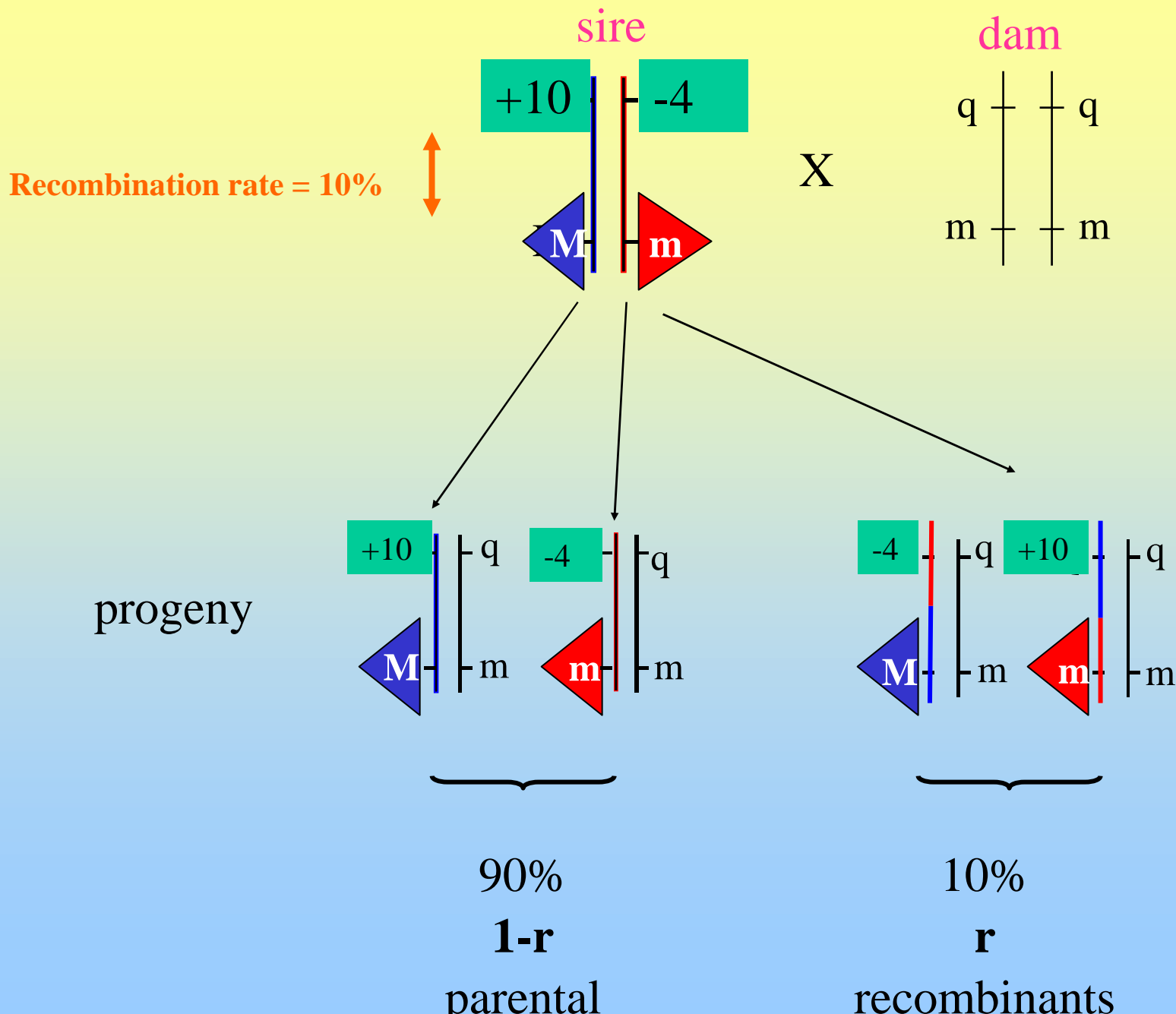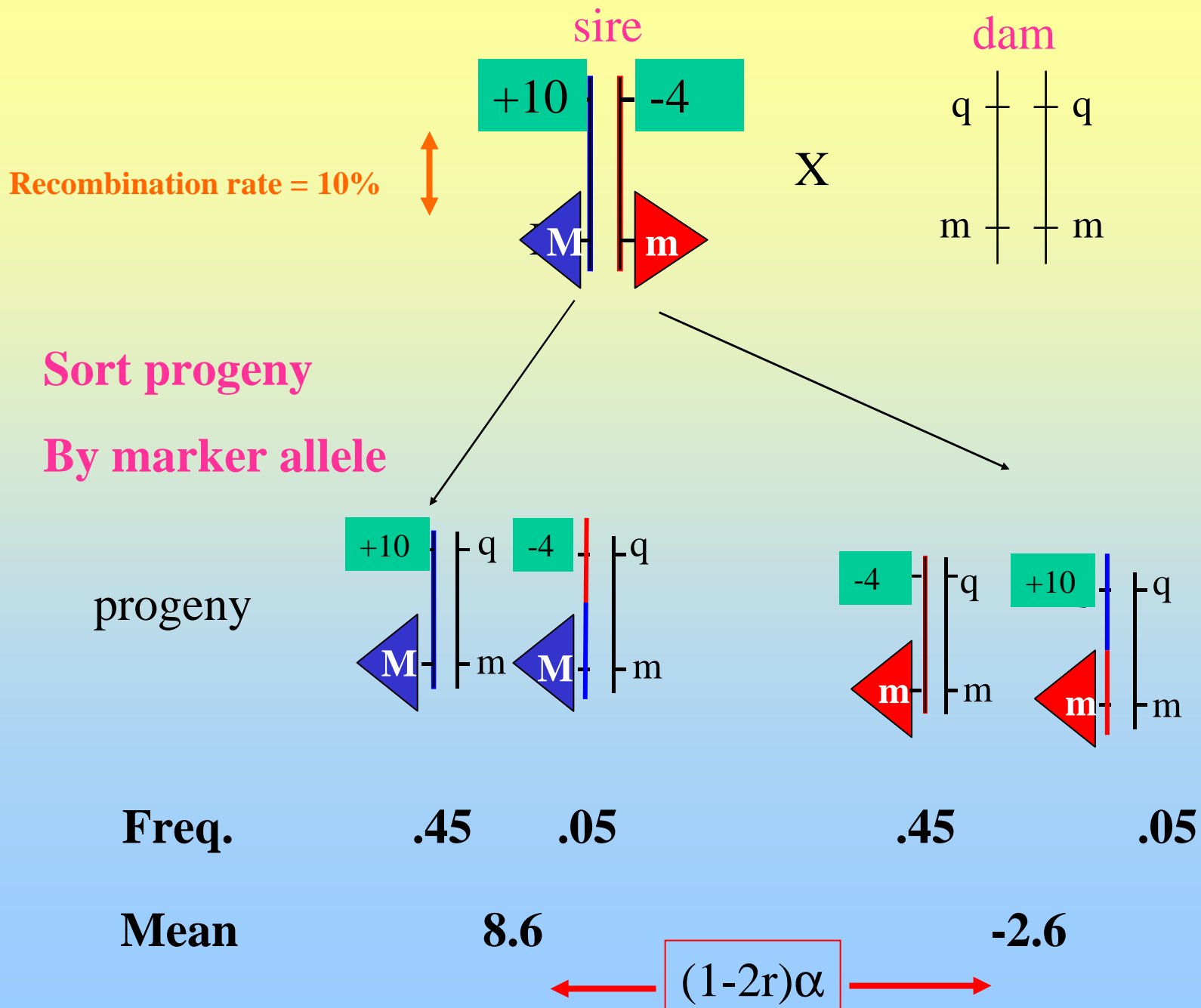
Simple QTL detection with single markers

Sire genotype: $\dfrac{M \qquad Q}{m \qquad q}$

Possible gametes	Recombination?	Gamete Probability
M Q	no	(1-r)/2
m Q	yes	r/2
M q	yes	r/2
m q	no	(1-r)/2

$\Sigma = r$ $\Sigma = 1 - r$

sire
dam

+10   -4

Recombination rate = 10%

M   m

q   q

m   m

X

Sort progeny

By marker allele

progeny

+10 q -4 q

M m M m

-4 q +10 q

m m m m

Freq.   .45   .05   .45   .05

Mean   8.6   -2.6

$(1-2r)\alpha$

ram   ewe

-4   +10

Recombination rate = 50%

X

q | | q

m | | m

M   m

progeny

+10 — q   -4 — q

M — m   M — m

-4 — q   +10 — q

m — m   m — m

Freq.   .05   .45   .05   .45

Mean   -2.6   8.6
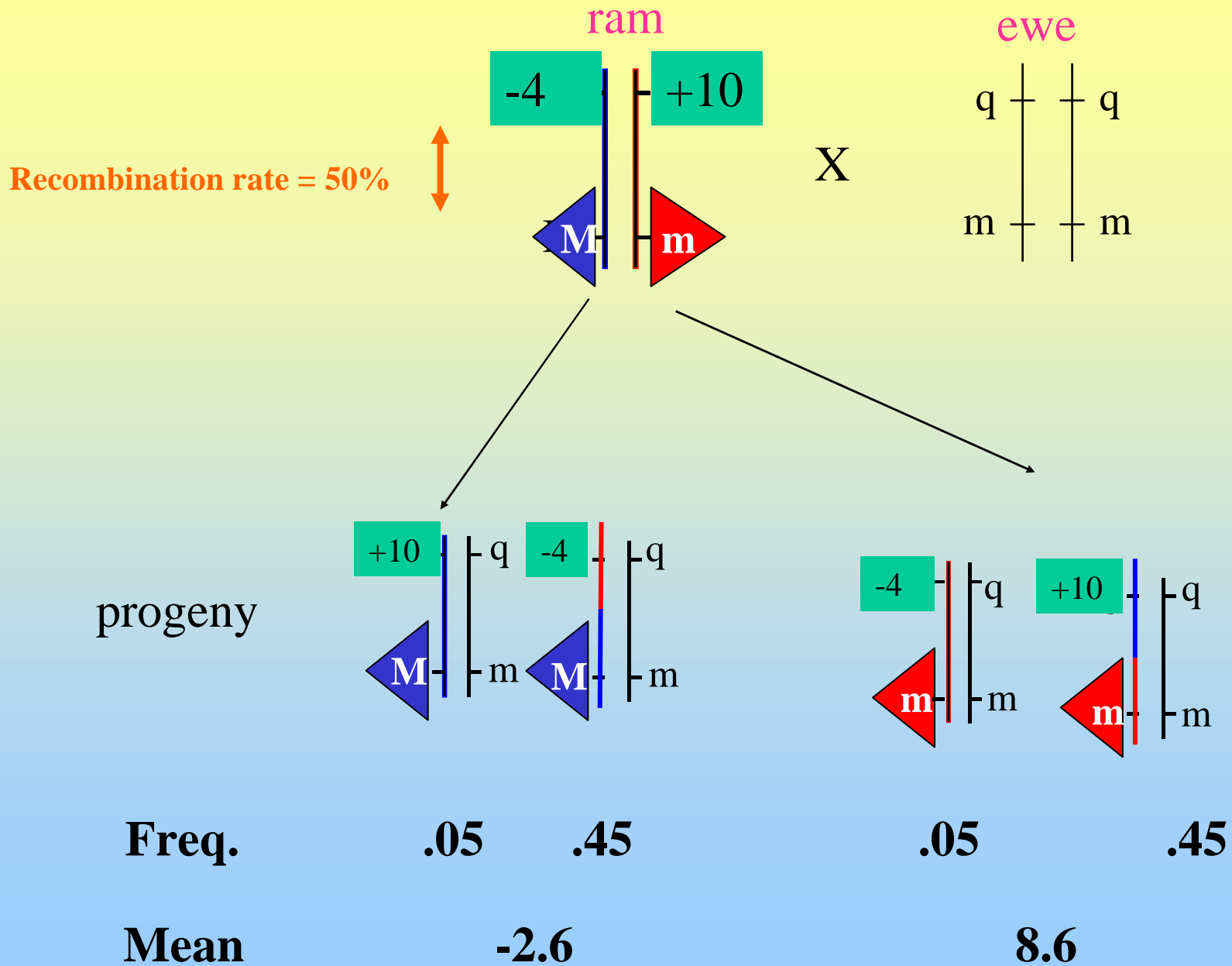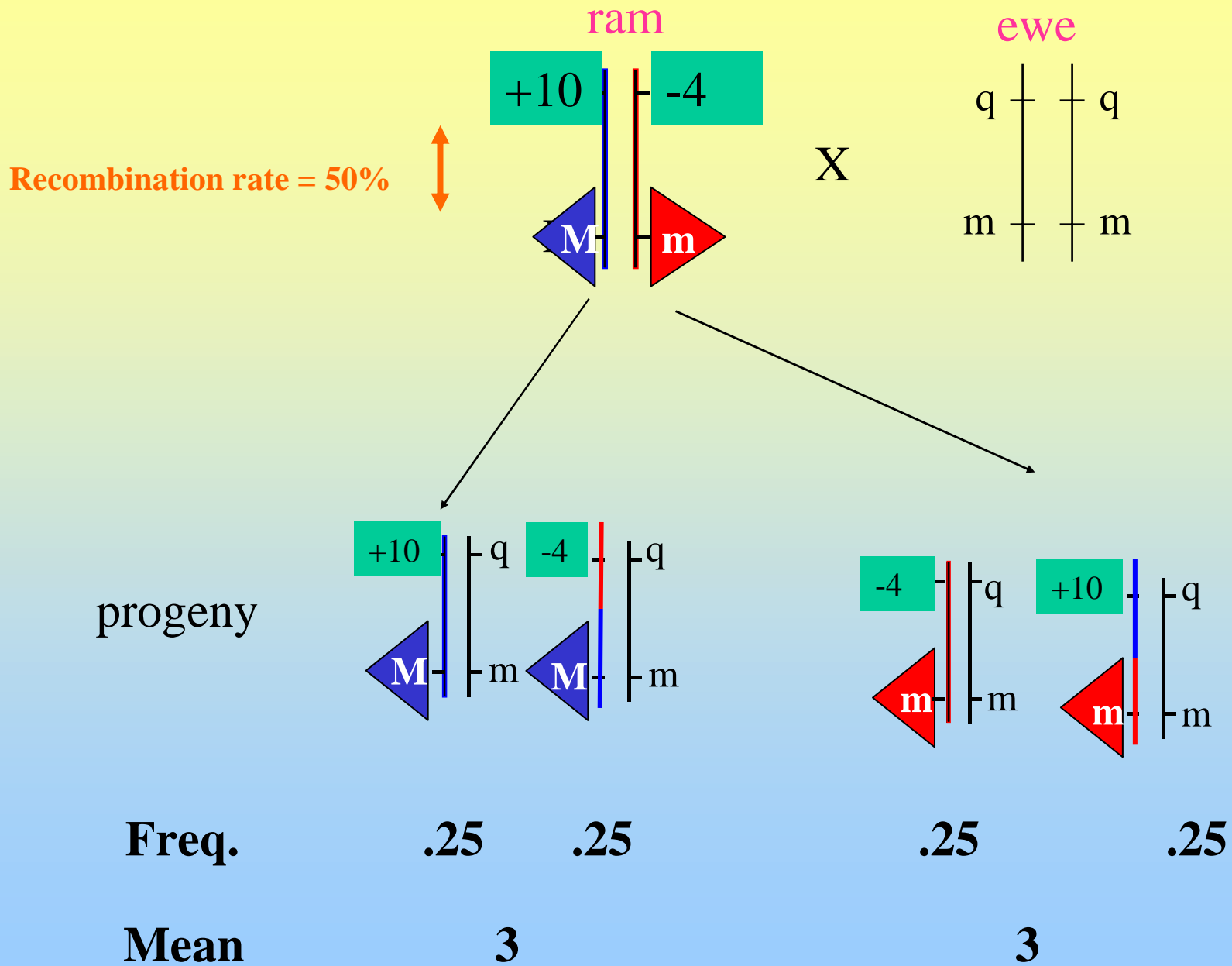
Expected means of progeny groups

Marker allele obtained from sire	QTL allele obtained from sire	frequency	Expected mean of progeny
M	Q	$(1-r)/2$	$\mu + \alpha$
M	q	$r/2$	μ
m	Q	$r/2$	$\mu + \alpha$
m	q	$(1-r)/2$	μ

Mean of M-group: $((1-r)(\mu + \alpha)/2 \ + \ r.\mu/2) \ / \ 0.5 =$ $\mu + (1-r)\, \alpha$

Mean of m-group: $(r.(\mu + \alpha)/2 \ + \ (1-r).\mu/2)/0.5 \ =$ $\mu + r\, \alpha$

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

Combinations of recombination rate (r) and QTL effect (α) giving rise to the same difference (D)

Difference between marker groups $(M - m) = (1-2r)\alpha$

Recombination rate M-Q (=r)	QTL-effect (Q-q substitution =α)	Mean of progeny[1] receiving M-allele	Mean of progeny[1] receiving m-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

To get independent estimates of r and α: > 1 marker locus &/or deeper pedigree

The effect of QTL status in dams

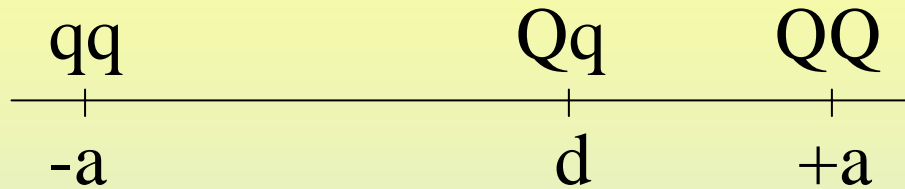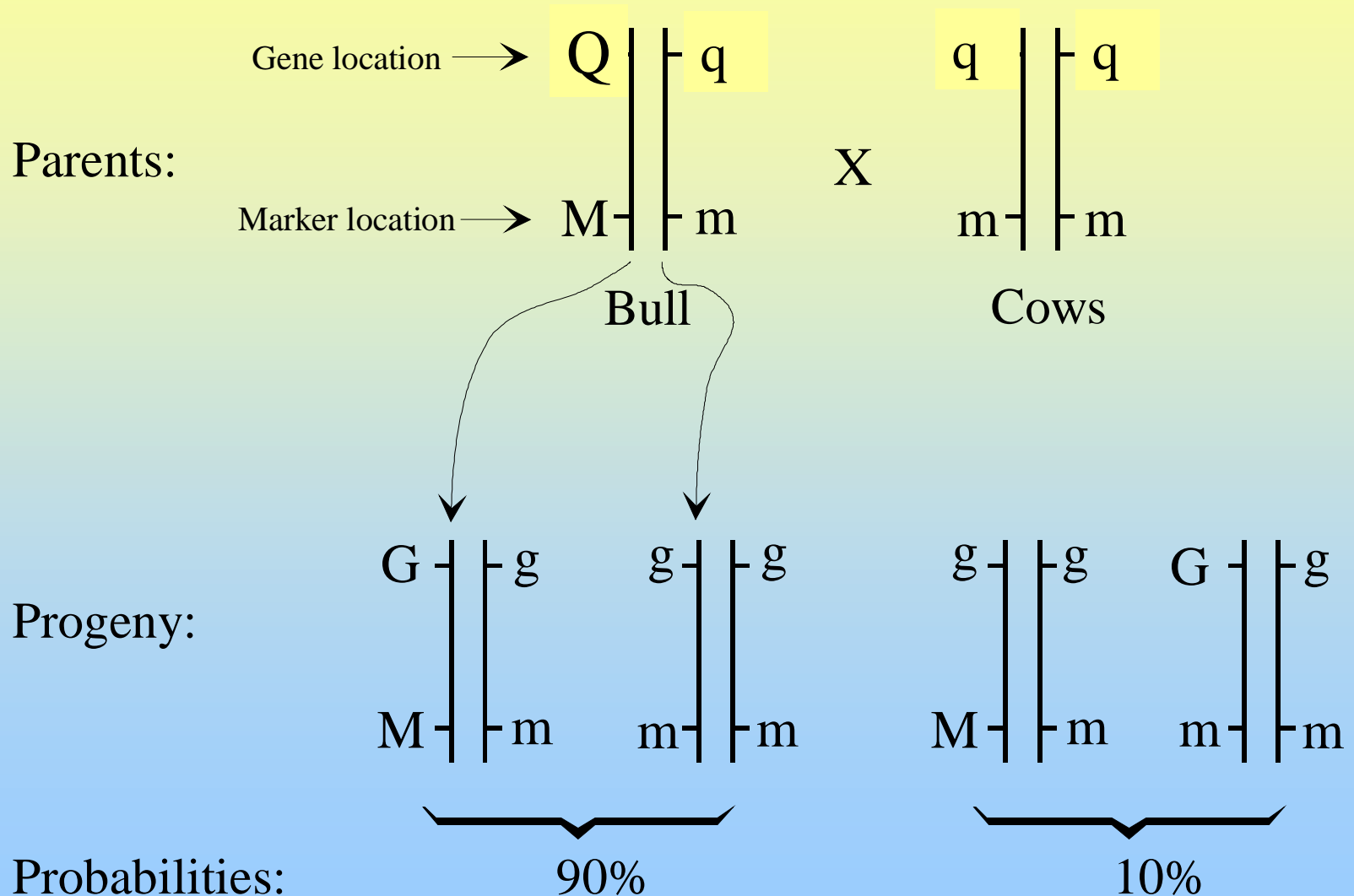| Dam population contributes	allele substitution effect (α) represents
q-alleles	$\mu_{Qq} - \mu_{qq}$
Q-alleles	$\mu_{QQ} - \mu_{Qq}$
freq(Q) = p	$p(\mu_{QQ} - \mu_{qQ}) + (1-p)(\mu_{Qq} - \mu_{qq})$

The effect of QTL status in dams

$$qq \quad\quad\quad\quad\quad Qq \quad\quad QQ$$

$$-a \quad\quad\quad\quad\quad\quad d \quad\quad +a$$

The average effect of an allele substitution …

$$\alpha = a + (q - p)d$$

$p = 1:$	$\mu_{QQ} - \mu_{Qq}$	$a - d$
$p = 0:$	$\mu_{Qq} - \mu_{qq}$	$a + d$

A backcross using inbred lines

Parents:

Gene location → Q | q   X   q | q

Marker location → M | m   m | m

Bull   Cows

Progeny:

G | g   g | g   g | g   G | g

M | m   m | m   M | m   m | m

Probabilities: 90%   10%

Backcross using inbred lines

- Difference between "M progeny' and "m progeny" is $(1-2r)\alpha$

- However, with inbred lines we have extra information:

  - QTL allele frequencies are 1 and 0 (or 0 and 1) in the inbred lines *if* the QTL is segregating.

Thus $\alpha = a + (q - p)d$ is: (Note error in notes!! p28)

$a + (1-0)d = a + d$ if the dam population is qq

$a + (0-1)d = a - d$ if the dam population is QQ

Backcross using inbred lines

- $\alpha = a + d$ if the dams are qq

- $\alpha = a - d$ if the dams are QQ

- Difference between M and m progeny $(1-2r)\alpha = (1-2r)(a+d)$ or $(1-2r)(a-d)$.

- If we make both backcrosses, we can get independent estimates of a and d from estimates of a+d and a-d.

F2 Cross – inbred parental lines

Sire   Dams

Gene location $\longrightarrow$ Q $\dashv$ $\vdash$ q   Q $\dashv$ $\vdash$ q

Parents: X

Marker location $\longrightarrow$ M $\dashv$ $\vdash$ m   M $\dashv$ $\vdash$ m

Table 1.

Sperm ↓	Eggs → Frequency	QM $\frac{1}{2}(1-r)$	qm $\frac{1}{2}(1-r)$	Qm $\frac{1}{2}r$	qM $\frac{1}{2}r$
QM	$\frac{1}{2}(1-r)$	a MM	d Mm	a Mm	d MM
qm	$\frac{1}{2}(1-r)$	d mM	-a mm	d mm	-a mM
Qm	$\frac{1}{2}r$	a mM	d mm	a mm	d mM
qM	$\frac{1}{2}r$	d MM	-a Mm	d Mm	-a MM

F2 Cross – inbred parental lines

Table 1.

Sperm ↓	Eggs → Frequency	QM ½(1-r)	qm ½(1-r)	Qm ½r	qM ½r
QM	½(1-r)	a MM	d Mm	a Mm	d MM
qm	½(1-r)	d mM	-a mm	d mm	-a mM
Qm	½r	a mM	d mm	a mm	d mM
qM	½r	d MM	-a Mm	d Mm	-a MM

Marker genotype	Predicted frequency	Equals
MM	$(½(1-r))^2 + 2½r½(1-r) + (½r)^2$	¼
Mm	$2[(½(1-r))^2 + 2½r½(1-r) + (½r)^2]$	½
mm	$(½(1-r))^2 + 2½r½(1-r) + (½r)^2$	¼

F2 Cross – inbred parental lines

Table 1.

Sperm ↓	Eggs → Frequency	QM ½(1-r)	qm ½(1-r)	Qm ½r	QM ½r
QM	½(1-r)	a MM	d Mm	a Mm	d MM
Qm	½(1-r)	d mM	-a mm	d mm	-a mM
Qm	½r	a mM	d mm	a mm	d mM
QM	½r	d MM	-a Mm	d Mm	-a MM

Marker genotype	Predicted merit	Equals
MM	$$\frac{(½(1-r))^2a + 2½r½(1-r)d + (½r)^2(-a)}{¼}$$	$(1-r)^2a + 2r(1-r)d + r^2(-a)$
Mm	$$\frac{2[(½(1-r))^2d + 2½r½(1-r)(a-a) + (½r)^2d]}{½}$$	$[(1-r)^2 + r^2]d$
mm	$$\frac{(½(1-r))^2(-a) + 2½r½(1-r)d + (½r)^2a}{¼}$$	$(1-r)^2(-a) + 2r(1-r)d + r^2a$

F2 Cross – inbred parental lines

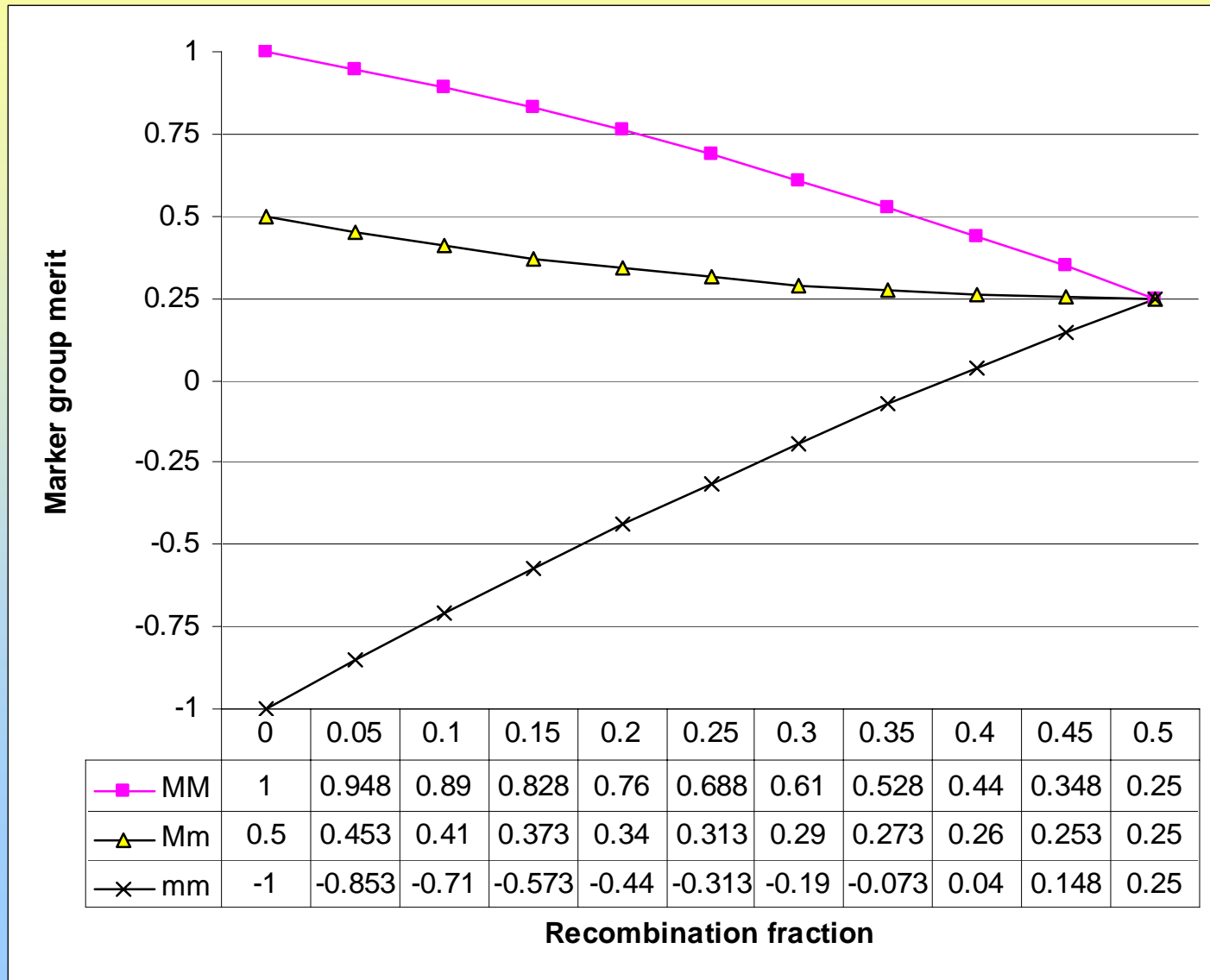| Marker genotype	Predicted merit	Equals
MM	$\dfrac{(\frac{1}{2}(1-r))^2a + 2\frac{1}{2}r\frac{1}{2}(1-r)d + (\frac{1}{2}r)^2(-a)}{\frac{1}{4}}$	$(1-r)^2a + 2r(1-r)d + r^2(-a)$
Mm	$\dfrac{2[(\frac{1}{2}(1-r))^2d + 2\frac{1}{2}r\frac{1}{2}(1-r)(a-a) + (\frac{1}{2}r)^2d]}{\frac{1}{2}}$	$[(1-r)^2 + r^2]d$
mm	$\dfrac{(\frac{1}{2}(1-r))^2(-a) + 2\frac{1}{2}r\frac{1}{2}(1-r)d + (\frac{1}{2}r)^2a}{\frac{1}{4}}$	$(1-r)^2(-a) + 2r(1-r)d + r^2a$

This gives us some sensible predicted merits:

Marker genotype	r = 0	r = ½
MM	a	¼a + ½d - ¼a
Mm	d	¼a + ½d - ¼a
mm	-a	¼a + ½d - ¼a

F2 Cross – inbred parental lines

	0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
MM	1	0.948	0.89	0.828	0.76	0.688	0.61	0.528	0.44	0.348	0.25
Mm	0.5	0.453	0.41	0.373	0.34	0.313	0.29	0.273	0.26	0.253	0.25
mm	-1	-0.853	-0.71	-0.573	-0.44	-0.313	-0.19	-0.073	0.04	0.148	0.25

Marker group merit

Recombination fraction

Using outbred parental lines

With non-inbred parental lines thee can be segregation at both QTL and marker loci in parental lines

Expected differences depend on allele frequency differences

Possibly treat each family different (possibly having a different phase between Q - M

F2 Cross – outbred parental lines

Problem:

Progeny that are heterozygous for the marker locus are not informative
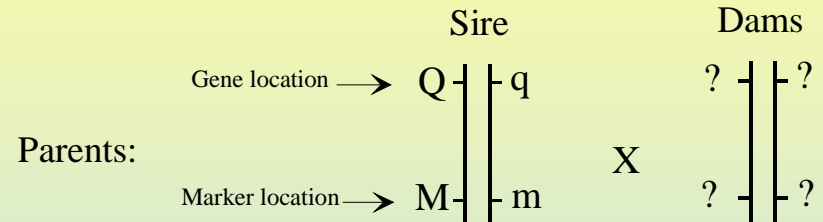
Mm - did this progeny get M or m from the sire?

This example assumes …

Table 2

Sperm ↓	Eggs →	QM	qm	Qm	qM
	Frequency within marker group→	p	(1-p)	p	(1-p)
QM	½(1-r)	a MM			d MM
qm	½(1-r)		-a mm	d mm	
Qm	½r		d mm	a mm	
qM	½r	d MM			-a MM

F2 Cross – outbred parental lines

Table 2

Sperm ↓	Eggs → / Frequency within marker group→	QM / p	qm / (1-p)	Qm / p	qM / (1-p)
QM	½(1-r)	a MM			d MM
qm	½(1-r)		-a mm	d mm	
Qm	½r		d mm	a mm	
qM	½r	d MM			-a MM

Marker group	Expectation	Equals
MM	$\dfrac{\text{½(1-r)pa + ½r(1-p)(-a) + ½rpd + ½(1-r)(1-p)d}}{\text{½(1-r)p + ½r(1-p) + ½rp + ½(1-r)(1-p)}}$	$\dfrac{\text{½(p-r).a + (r.p+½(1-p-r)).d}}{\text{½}}$
mm	$\dfrac{\text{½rpa + ½(1-r)(1-p)(-a) + ½r(1-p)d + ½(1-r)pd}}{\text{½rp + ½(1-r)(1-p) + ½r(1-p) + ½(1-r)p}}$	$\dfrac{\text{½(p+r-1).a + [½(r+p)-rp]d}}{\text{½}}$

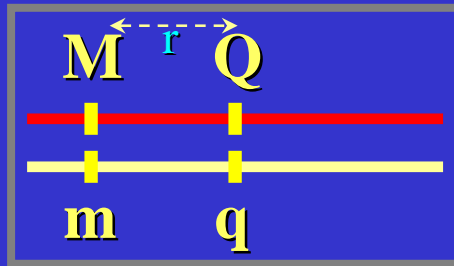# F2 Cross – outbred parental lines

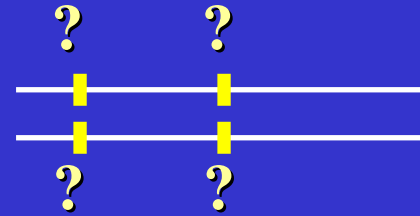| Marker group	Expectation	Equals
MM	$\dfrac{\frac{1}{2}(1-r)pa + \frac{1}{2}r(1-p)(-a) + \frac{1}{2}rpd + \frac{1}{2}(1-r)(1-p)d}{\frac{1}{2}(1-r)p + \frac{1}{2}r(1-p) + \frac{1}{2}rp + \frac{1}{2}(1-r)(1-p)}$	$\dfrac{\frac{1}{2}(p-r).a + (r.p+\frac{1}{2}(1-p-r)).d}{\frac{1}{2}}$
mm	$\dfrac{\frac{1}{2}rpa + \frac{1}{2}(1-r)(1-p)(-a) + \frac{1}{2}r(1-p)d + \frac{1}{2}(1-r)pd}{\frac{1}{2}rp + \frac{1}{2}(1-r)(1-p) + \frac{1}{2}r(1-p) + \frac{1}{2}(1-r)p}$	$\dfrac{\frac{1}{2}(p+r-1).a + [\frac{1}{2}(r+p)-rp]d}{\frac{1}{2}}$
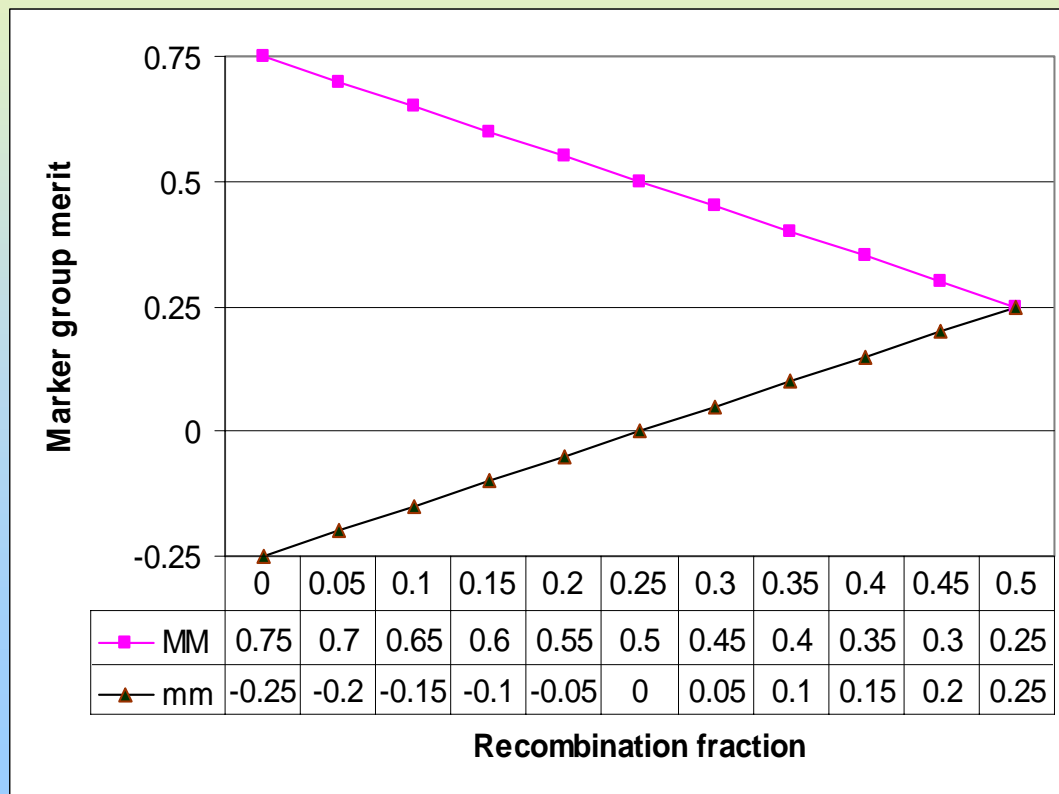
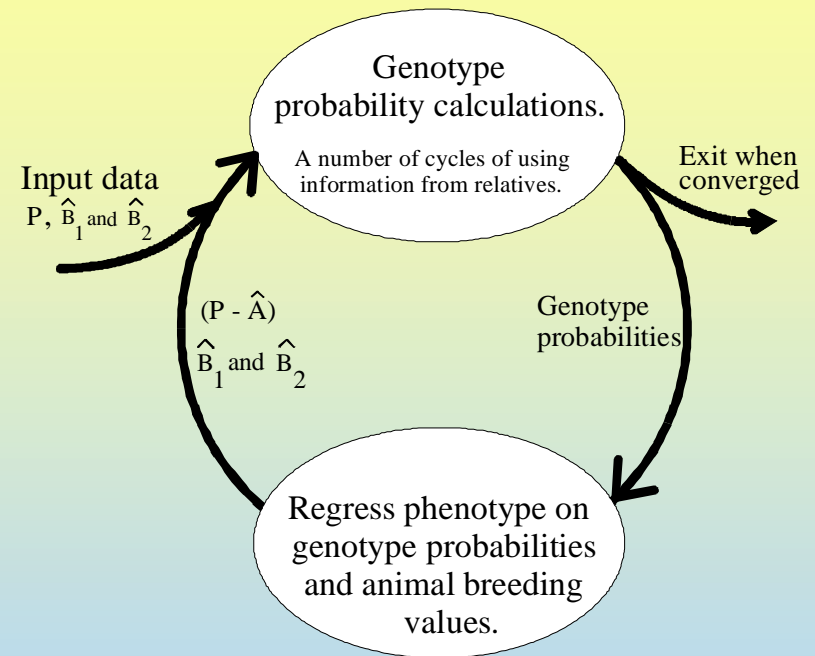|	0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
MM	0.75	0.7	0.65	0.6	0.55	0.5	0.45	0.4	0.35	0.3	0.25
mm	-0.25	-0.2	-0.15	-0.1	-0.05	0	0.05	0.1	0.15	0.2	0.25

Marker group merit

Recombination fraction
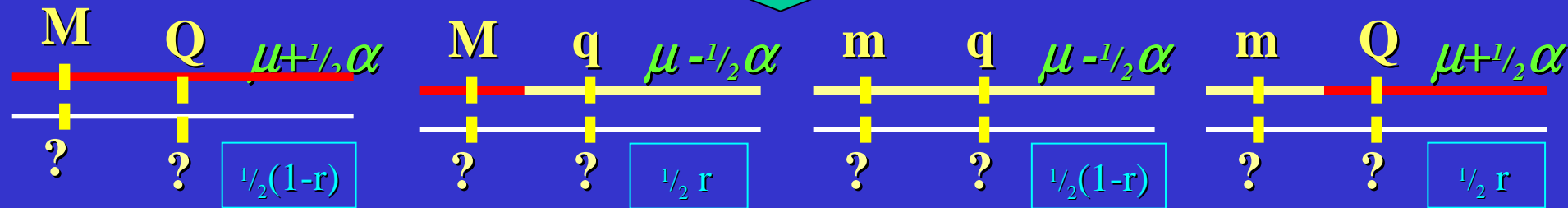
General pedigree outbred populations

- With …
  - richer pedigree
  - more markers
  - better method …
- We can get …
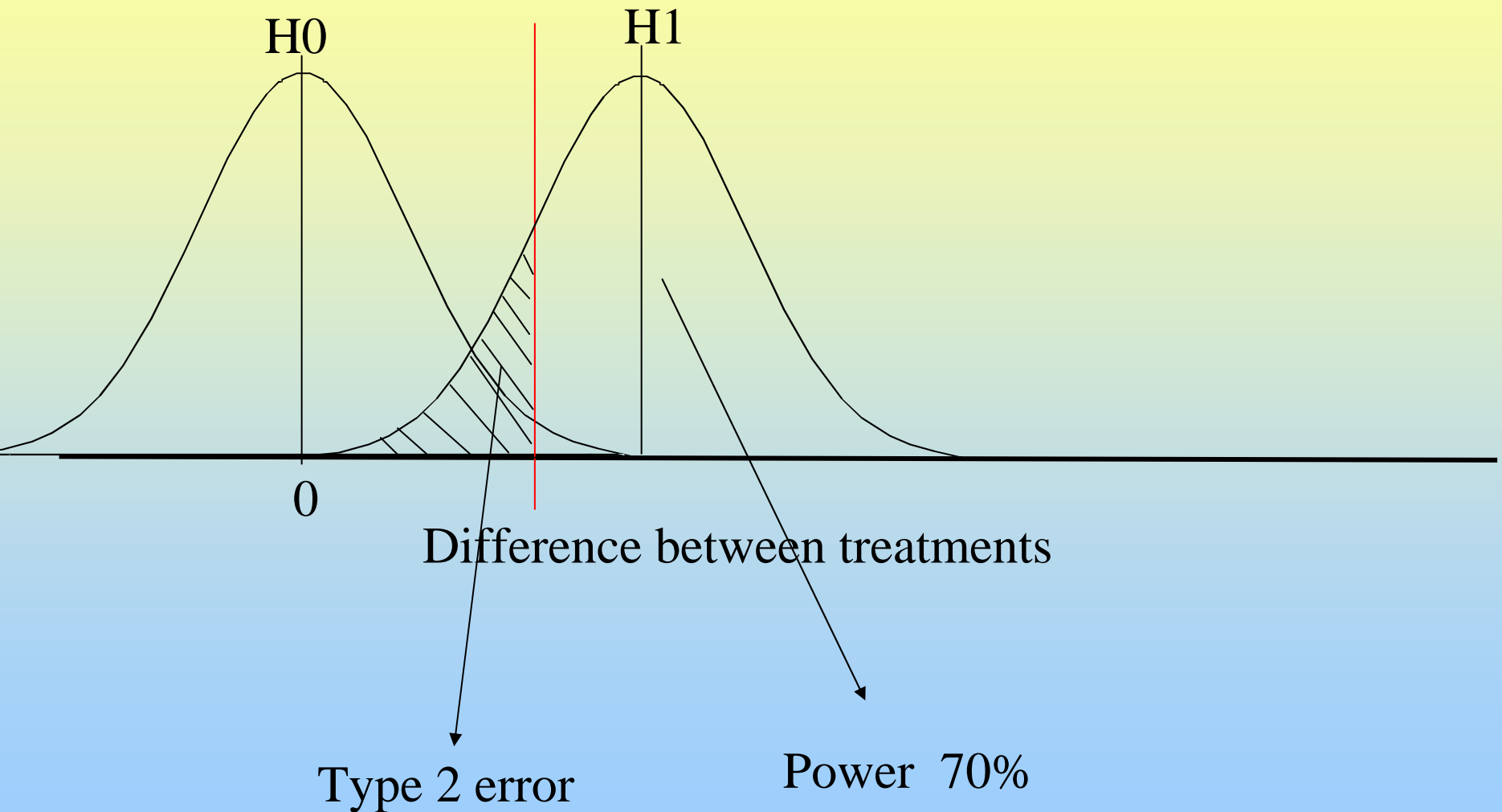 - genotype probabilities for QQ Qq and qq
  - estimates of *a* and *d*.

[Coming]

Input data
P, $\hat{B}_1$ and $\hat{B}_2$

Genotype probability calculations.

A number of cycles of using information from relatives.

Exit when converged

Genotype probabilities

Regress phenotype on genotype probabilities and animal breeding values.

$(P - \hat{A})$
$\hat{B}_1$ and $\hat{B}_2$

How many animals do we need to measure?

H0

threshold

0

Difference between treatments

$SD^2 = 2\sigma^2/n$

Type 1 error

How many animals do we need to measure?

H0

H1

0

Difference between treatments

Type 2 error

Power 70%
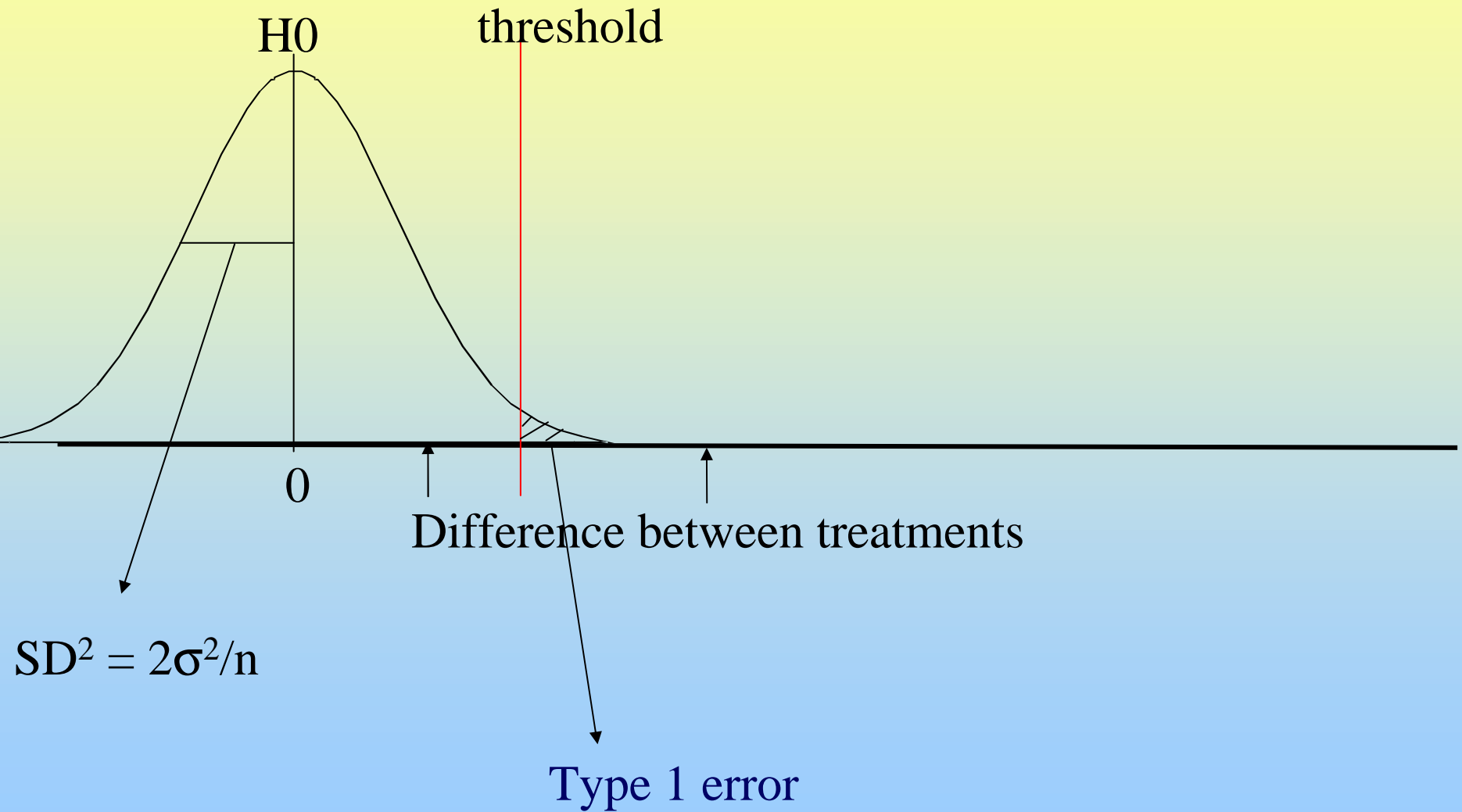
Testing group differences

Diff

SED

0

Difference between treatments

$$95\% CI = D \pm 1.96\ SED$$

How many animals do we need to measure?

H0

H1

Larger n

0

Difference between treatments

Power 99%
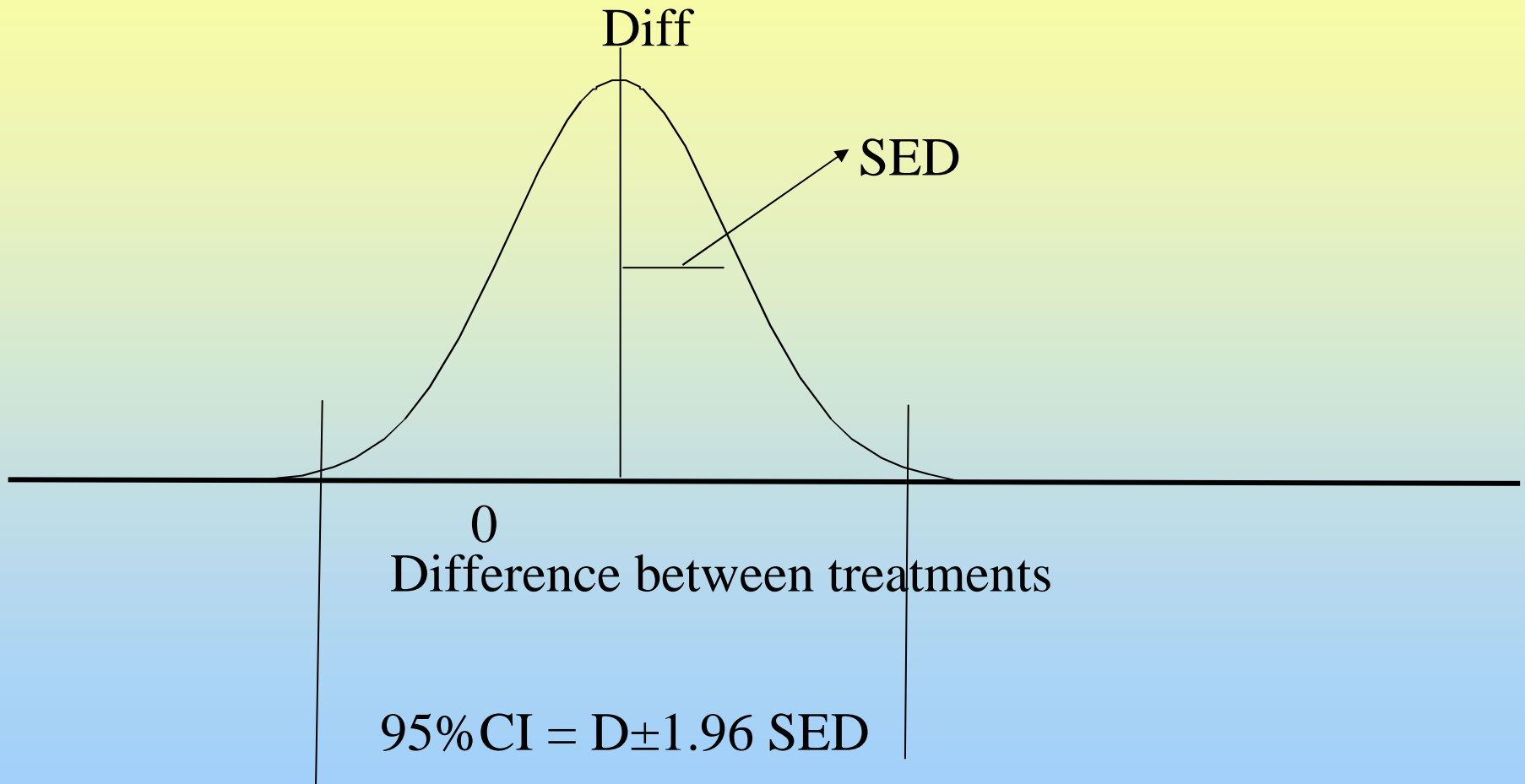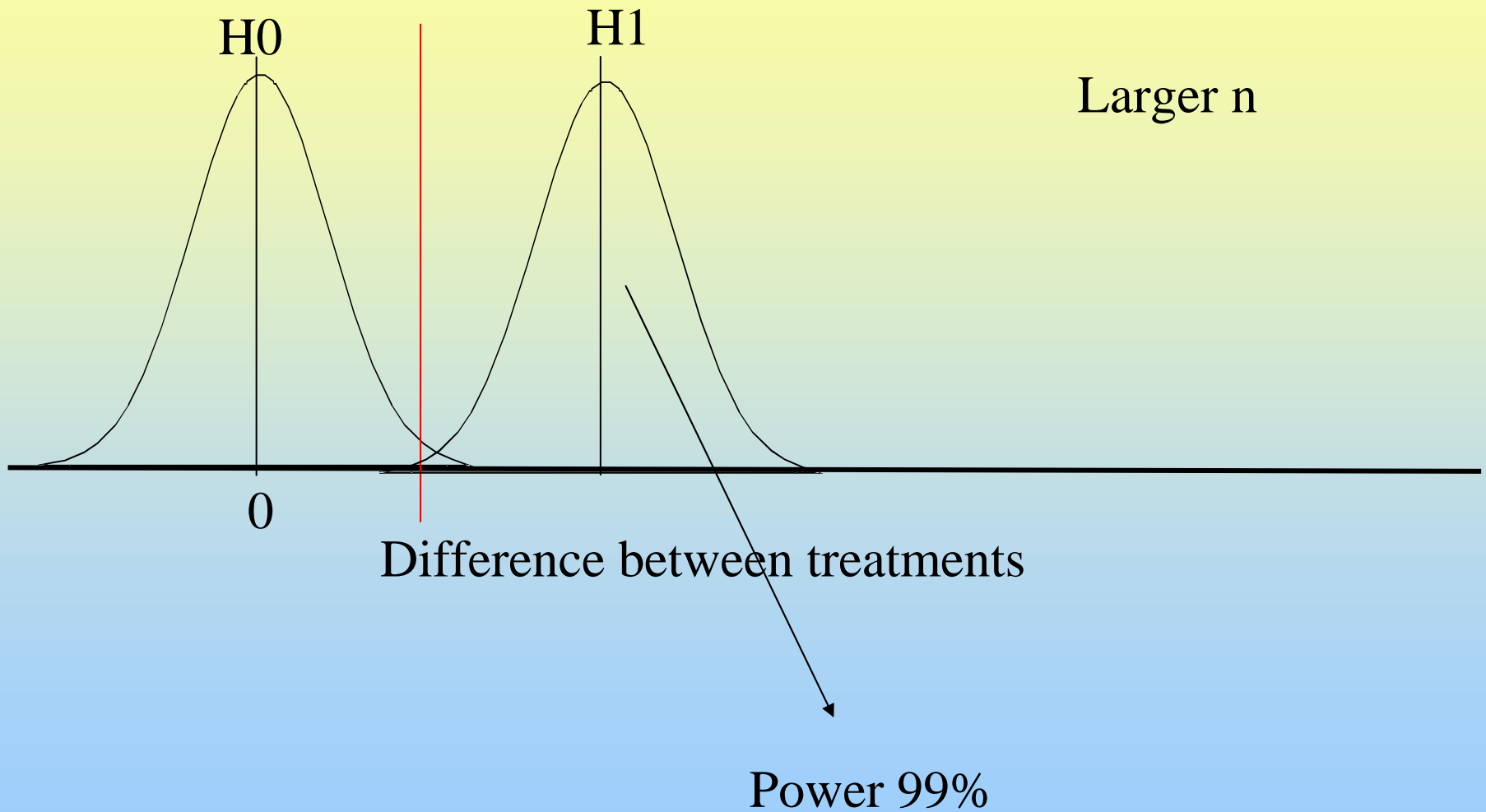
How many offspring needed per sire?
Power of QTL detection?

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