

Pset 1

Timothy Mulderrig

2025-02-03

Q1

As a plant breeder I want to fix an allele that provides resistance to an important crop disease. Unfortunately, I do not have the ability to directly genotype the causative mutation, but I have identified a genetic marker that is closely linked the causative mutation. Based on mapping experiments the recombination frequency between the marker (alleles: A & a) and the QTL (alleles: Q – resistant & q – susceptible) is $r = .25$

I have crossed a homozygous resistant line to a homozygous susceptible line to get an offspring with the genotype: [A/Q] [a/q]

I will self this individual to generate offspring segregating for resistance.

1A)

If I select only lines with the genotype AA, what expected proportion of the selected lines that will be homozygous for the favorable QTL (QQ)? (2 pts)

Frequency of possible gametes

```
r=0.25
freq_AQ <- (1-r)*0.5
freq_aq <- (1-r)*0.5
freq_Aq <- r*(0.5)
freq_aQ <- r*(0.5)
```

Frequency of Genotype combinations in F2

```
#multiply by 2 in cases where unique gametes could originate from the pollen or egg donor

AAQQ <- freq_AQ*freq_AQ
AAQq <- freq_AQ*freq_Aq*2
AAqq <- freq_Aq*freq_Aq
AaQQ <- freq_AQ*freq_aQ*2
AaQq <- freq_AQ*freq_aq*2 + freq_Aq*freq_aQ*2
Aaqq <- freq_Aq*freq_aq*2
aaQQ <- freq_aQ*freq_aQ
aaQq <- freq_aQ*freq_aq*2
aaqq <- freq_aq*freq_aq

AAQQ+AAQq+AAqq+AaQQ+AaQq+Aaqq+aaQQ+aaQq+aaqq
```

```
## [1] 1
```

Selected genotypes

```
sel_genos <- AAQQ + AAQq + AAqq  
QQ_sel <- AAQQ / sel_genos  
print(QQ_sel)
```

```
## [1] 0.5625
```

Answer: The expected proportion of AA selected lines that will be homozygous for the resistance QTL is 56.25%.

Q2

Similar to Q1 I have identified markers linked to a QTL of interest and I now want to practice marker assisted selection to fix the favorable allele in a population. In this scenario I have 2 markers linked on either side of the QTL. Marker 1 (alleles A, a) has a recombination frequency $r = .4$ with the QTL and Marker 2 (alleles B, b) has a recombination frequency $r = .3$ with the QTL. As before I've crossed a line that is homozygous for the beneficial allele (Q) and a line that is homozygous for the unfavorable allele (q):

$[A/Q/B] \times [a/q/b]$

I selfed the heterozygous line above to generate many offspring. The QQ genotype has a value of $a = 5$ for the trait of interest, the unfavorable genotype qq has a value of $-a = -5$ for the trait of interest. The trait is additive (no dominance)

2A)

As a breeder I want to fix the resistance allele. If I select only lines with the genotype AABB, what is the expected proportion of the selected lines that will be homozygous for the favorable QTL (QQ)? (2 pts)

Frequency of possible gametes

```
r1 <- 0.4  
r2 <- 0.3  
  
freq_AQB <- 0.5*(1-r1)*(1-r2)  
freq_AQb <- 0.5*(1-r1)*(r2)  
freq_Aqb <- 0.5*(r1)*(1-r2)  
freq_AqB <- 0.5*(r1)*(r2)  
freq_aQB <- 0.5*(r1)*(1-r2)  
freq_aQb <- 0.5*(r1)*(r2)  
freq_aqB <- 0.5*(1-r1)*(r2)  
freq_aqb <- 0.5*(1-r1)*(1-r2)  
  
freq_AQB + freq_AQb + freq_Aqb + freq_AqB + freq_aQB + freq_aQb + freq_aqB + freq_aqb
```

```
## [1] 1
```

Frequency of genotypes

```

AAQQBB <- freq_AQB*freq_AQB
AAQQBb <- freq_AQB*freq_Aqb*2
AAQQbb <- freq_Aqb*freq_Aqb
AAQqBB <- freq_AQB*freq_AqB*2
AAQqBb <- freq_AQB*freq_Aqb*2 + freq_Aqb*freq_AqB*2
AAQqbb <- freq_Aqb*freq_Aqb*2
AAqqBB <- freq_AqB*freq_AqB
AAqqBb <- freq_AqB*freq_Aqb*2
AAqqbb <- freq_Aqb*freq_Aqb
AaQQBB <- freq_AQB*freq_aQB*2
AaQQBb <- freq_AQB*freq_aqb*2 + freq_Aqb*freq_aQB*2
AaQQbb <- freq_Aqb*freq_aqb*2
AaQqBB <- freq_AQB*freq_aqB*2 + freq_AqB*freq_aQB*2
AaQqBb <- freq_AQB*freq_aqb*2 + freq_Aqb*freq_aqB*2 + freq_Aqb*freq_aQB*2 + freq_aqb*freq_AqB*2
AaQqbb <- freq_Aqb*freq_aqb*2 + freq_Aqb*freq_aqb*2
AaqqBB <- freq_AqB*freq_aqB*2
AaqqBb <- freq_AqB*freq_aqb*2 + freq_aqb*freq_AqB*2
Aaqqbb <- freq_Aqb*freq_aqb*2
aaQQBB <- freq_aQB*freq_aQB
aaQQBb <- freq_aQB*freq_aqb*2
aaQQbb <- freq_aqb*freq_aqb
aaQqBB <- freq_aQB*freq_aqB*2
aaQqBb <- freq_aQB*freq_aqb*2+freq_aqb*freq_aqB*2
aaQqbb <- freq_aqb*freq_aqb*2
aaqqBB <- freq_aqB*freq_aqB
aaqqBb <- freq_aqB*freq_aqb*2
aaqqbb <- freq_aqb*freq_aqb

sum(c(AAQQBB, AAQQBb, AAQQbb, AAQqBB, AAQqBb, AAQqbb, AAqqBB, AAqqBb, AAqqbb,
      AaQQBB, AaQQBb, AaQQbb, AaQqBB, AaQqBb, AaQqbb, AaqqBB, AaqqBb, Aaqqbb,
      aaQQBB, aaQQBb, aaQQbb, aaQqBB, aaQqBb, aaQqbb, aaqqBB, aaqqBb, aaqqbb))

```

```
## [1] 1
```

Selected genotypes

```

sel_genos_b <- AAQQBB + AAQqBB + AAqqBB
QQ_sel_b <- AAQQBB / sel_genos_b
print(QQ_sel_b)

```

```
## [1] 0.6049383
```

Answer: The expected proportion of selected lines that will be homozygous for the resistance QTL is 60.49 %

2B)

What is the expected value for offspring with the genotype AABB? (3 pts)

```

AABB_e <- AAQQBB*5 + AAQqBB*0 + AAqqBB*-5
print(AABB_e)

```

```
## [1] 0.2025
```

Answer: The expected value for offspring with AAbb is 0.25

2C)

What is the expected value for lines with the genotype AA? (3 pts)

```
AAQQBB*5 + AAQQBb*5 + AAQqbb*5 + AAQqBB*0 + AAQqBb*0 + AAqqbb*0 + AAqqBB*(-5) + AAqqBb*(-5) + AAqqbb*(-5)
```

```
## [1] 0.25
```

Answer: The expected value for lines with AA is 0.25

2D)

What is the expected value for lines with the genotype BB? (3 pts)

```
AAQQBB*5 + AaQQBB*5 + aaQQBB*5 + AAQqBB*0 + AaQqBB*0 + aaQqBB*0 + AAqqBB*(-5) + AaqqBB*(-5) + aaqqBB*(-5)
```

```
## [1] 0.5
```

Answer: The expected value for lines with BB is 0.5

2E)

How close would a single marker need to be in order for marker assisted selection to as effective and the two marker example above? (2 pts)

Assume there is a marker c that is linked to qtl q with a recombination frequency of rc between c and q in the gametes [C/Q][c/q].

The potential gametes appear with the frequencies: $* [C/Q] = 0.5(1-rc)$ $[C/q] = 0.5(rc)$ $[c/Q] = 0.5(rc)$ $[c/q] = 0.5(1-rc)$

The resulting genotypes appear with the frequencies:

- $CCQQ \leftarrow \text{freq_CQ} * \text{freq_CQ}$
- $CCQq \leftarrow \text{freq_CQ} * \text{freq_Cq} * 2$
- $CCqq \leftarrow \text{freq_Cq} * \text{freq_Cq}$
- $CcQQ \leftarrow \text{freq_CQ} * \text{freq_cQ} * 2$
- $CcQq \leftarrow \text{freq_CQ} * \text{freq_cq} * 2 + \text{freq_Cq} * \text{freq_cQ} * 2$
- $Ccqq \leftarrow \text{freq_Cq} * \text{freq_cq} * 2$
- $ccQQ \leftarrow \text{freq_cQ} * \text{freq_cQ}$
- $ccQq \leftarrow \text{freq_cQ} * \text{freq_cq} * 2$
- $ccqq \leftarrow \text{freq_cq} * \text{freq_cq}$

If we select only the lines with CC marker scores, the expected genotypic value would be as follows:

- Expected Value = $CCQQ * 5 + CCQq * 0 + CCqq * -5$ or
- Expected Value = $(0.5 * (1-rc)) * (0.5 * (1-rc)) * 5 + (0.5 * rc)(0.5 * rc) * (-5)$
- Expected Value = $(5)(0.25rc^2 - 0.5rc + 0.25) + (-5)(0.25rc^2)$
- Expected Value = $1.25rc^2 - 2.5rc + 1.25 - 1.25rc^2$
- Expected Value = $-2.5rc + 1.25$

```
rc <- (AABB_e-1.25) / (-2.5)
print(rc)
```

```
## [1] 0.419
```

Answer: A marker would need to have a recombination frequency between itself and the QTL of 0.419 for the selected lines to have the same expected value as the lines selected with our flanking markers