2. Estimation of power in crosses between inbred lines

For QTL bracketed by two markers, the effect measured will not be reduced by the marker bracket, therefore, will be reduced by this factor and (1-R), respectively, where R is the recombination frequency. Recombinant individuals will be deleted. The proportion of recombinants for the F-2 and BC designs will be (1-R). For QTL effects, the substitution effect measured will decrease proportional to 1-2r, as compared to complete linkage. Thus, to achieve power equal to the case of complete linkage, it will be necessary to increase the number of offsprings by a factor of 1/(1-2r)

4.2 Optimization of experimental designs

The distribution of QTL phenotypic variance in the F-2 population. But these are lower bounds. A few studies have also empirically estimated confidence intervals by repeat simulation, and these are generally close to the theoretical values for large samples. Examples will be presented.

Table 4.1. The expectation of the contrasts, and required numbers of offsprings to obtain statistical power of 0.9 for the BC and F-2 designs. (2a = 0.282, = 0.1, and r = 0, are given in Table 4.1.) Power, 1-

Table 4.1. The expectation of the contrasts, and required numbers of offsprings to obtain statistical power of 0.9 for the BC and F-2 designs. (2a = 0.282, = 0.1, and r = 0, a-d are the standard normal distribution values where: n = 0.50, Z = 1.96, and Z = 2.57.

Table 4.1. Cross Contrast Sample size a p 0.025 1050 1030 1200 2200 1050 0.05 2100 2100 2500 2700 2500

Chapter 4: Statistical power to detect QTL, parameter confidence intervals, and optimization of experimental designs...
Power for the granddaughter design is assumed to be the same as for the BC or F-2. However, the power for the granddaughter design is not as great as for the BC or F-2. Soller and Genizi (1978) estimated power for the granddaughter design assuming a nested ANOVA with codominance. The expectation of the mean squares for A for the F-2 design with codominance at one locus is given by

\[ \text{MSE}_A = \sigma^2 + \frac{\sigma_G^2}{n} + \frac{\sigma_e^2}{n} \]

where \( \sigma^2 \) is the residual variance, \( \sigma_G^2 \) is the genetic variance between lines, and \( n \) is the number of individuals.

### Estimation of Power for Segregating Populations

The power for segregating populations can be estimated using the following equation:

\[ \text{Power} = 1 - \Phi(\frac{2 \sqrt{h^2(1-h^2)/n}}{\sqrt{\frac{h^2+(1-h^2)/2}{n}}} \times 1) \]

where \( h^2 \) is the heritability and \( n \) is the number of lines.

The variance between lines will be \( h^2 \) for all the replicate progeny designs considered. The variance within lines will be \( \frac{h^2+(1-h^2)/2}{n} \) for all the replicate progeny designs.

### Table 4.2. Between and Within Progeny Group Variance Components, and the Required Number of Lines

<table>
<thead>
<tr>
<th>Progeny Type</th>
<th>Variance Component</th>
<th>Between Lines</th>
<th>Within Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-3</td>
<td></td>
<td>( h^2(1-h^2) )</td>
<td>( h^2+(1-h^2)/2 )</td>
</tr>
<tr>
<td>F-2</td>
<td></td>
<td>( h^2 )</td>
<td>( h^2+(1-h^2)/n )</td>
</tr>
<tr>
<td>BC</td>
<td></td>
<td>( h^2(1-h^2) )</td>
<td>( h^2+(1-h^2)/n )</td>
</tr>
</tbody>
</table>

### 4.3 Replicate Progeny in Crosses between Inbred Lines

The power for the granddaughter design is assumed to be the same as for the BC or F-2. However, the power for the granddaughter design is not as great as for the BC or F-2. Soller and Genizi (1978) estimated power for the granddaughter design assuming a nested ANOVA with codominance. The expectation of the mean squares for A for the F-2 design with codominance at one locus is given by

\[ \text{MSE}_A = \sigma^2 + \frac{\sigma_G^2}{n} + \frac{\sigma_e^2}{n} \]

where \( \sigma^2 \) is the residual variance, \( \sigma_G^2 \) is the genetic variance between lines, and \( n \) is the number of individuals.

The variance between lines will be \( h^2 \) for all the replicate progeny designs considered. The variance within lines will be \( \frac{h^2+(1-h^2)/2}{n} \) for all the replicate progeny designs.

### Table 4.2. Between and Within Progeny Group Variance Components, and the Required Number of Lines

<table>
<thead>
<tr>
<th>Progeny Type</th>
<th>Variance Component</th>
<th>Between Lines</th>
<th>Within Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-3</td>
<td></td>
<td>( h^2(1-h^2) )</td>
<td>( h^2+(1-h^2)/2 )</td>
</tr>
<tr>
<td>F-2</td>
<td></td>
<td>( h^2 )</td>
<td>( h^2+(1-h^2)/n )</td>
</tr>
<tr>
<td>BC</td>
<td></td>
<td>( h^2(1-h^2) )</td>
<td>( h^2+(1-h^2)/n )</td>
</tr>
</tbody>
</table>

### 4.3 Replicate Progeny in Crosses between Inbred Lines

The power for the granddaughter design is assumed to be the same as for the BC or F-2. However, the power for the granddaughter design is not as great as for the BC or F-2. Soller and Genizi (1978) estimated power for the granddaughter design assuming a nested ANOVA with codominance. The expectation of the mean squares for A for the F-2 design with codominance at one locus is given by

\[ \text{MSE}_A = \sigma^2 + \frac{\sigma_G^2}{n} + \frac{\sigma_e^2}{n} \]

where \( \sigma^2 \) is the residual variance, \( \sigma_G^2 \) is the genetic variance between lines, and \( n \) is the number of individuals.

The variance between lines will be \( h^2 \) for all the replicate progeny designs considered. The variance within lines will be \( \frac{h^2+(1-h^2)/2}{n} \) for all the replicate progeny designs.

### Table 4.2. Between and Within Progeny Group Variance Components, and the Required Number of Lines

<table>
<thead>
<tr>
<th>Progeny Type</th>
<th>Variance Component</th>
<th>Between Lines</th>
<th>Within Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-3</td>
<td></td>
<td>( h^2(1-h^2) )</td>
<td>( h^2+(1-h^2)/2 )</td>
</tr>
<tr>
<td>F-2</td>
<td></td>
<td>( h^2 )</td>
<td>( h^2+(1-h^2)/n )</td>
</tr>
<tr>
<td>BC</td>
<td></td>
<td>( h^2(1-h^2) )</td>
<td>( h^2+(1-h^2)/n )</td>
</tr>
</tbody>
</table>

### 4.3 Replicate Progeny in Crosses between Inbred Lines

The power for the granddaughter design is assumed to be the same as for the BC or F-2. However, the power for the granddaughter design is not as great as for the BC or F-2. Soller and Genizi (1978) estimated power for the granddaughter design assuming a nested ANOVA with codominance. The expectation of the mean squares for A for the F-2 design with codominance at one locus is given by

\[ \text{MSE}_A = \sigma^2 + \frac{\sigma_G^2}{n} + \frac{\sigma_e^2}{n} \]

where \( \sigma^2 \) is the residual variance, \( \sigma_G^2 \) is the genetic variance between lines, and \( n \) is the number of individuals.

The variance between lines will be \( h^2 \) for all the replicate progeny designs considered. The variance within lines will be \( \frac{h^2+(1-h^2)/2}{n} \) for all the replicate progeny designs.

### Table 4.2. Between and Within Progeny Group Variance Components, and the Required Number of Lines

<table>
<thead>
<tr>
<th>Progeny Type</th>
<th>Variance Component</th>
<th>Between Lines</th>
<th>Within Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-3</td>
<td></td>
<td>( h^2(1-h^2) )</td>
<td>( h^2+(1-h^2)/2 )</td>
</tr>
<tr>
<td>F-2</td>
<td></td>
<td>( h^2 )</td>
<td>( h^2+(1-h^2)/n )</td>
</tr>
<tr>
<td>BC</td>
<td></td>
<td>( h^2(1-h^2) )</td>
<td>( h^2+(1-h^2)/n )</td>
</tr>
</tbody>
</table>
4.5 Confidence intervals for QTL parameters

As shown in Equations (3.16) to (3.19), for the estimation of QTL parameters in a linear model, the confidence interval (CI) for the QTL effect can be estimated from the inverse of the maximum likelihood (ML) matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The CI for QTL effects measured relative to the genetic SD can be estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

As mentioned in Chapter 2, very large samples will be required in the SNP design of the QTL. The power of the SNP design increases with the number of genotypes. The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.

The power of the SNP design is estimated by dividing the QTL effect by the standard error of the QTL effect estimate. The standard error of the QTL effect estimate is estimated from the inverse of the ML matrix of second differentials. This is also the case for QTL effects measured relative to the genetic SD.
unlinked. For RIL, optimum spacing will be about 50 cM. Even if the cost of obtaining trait
the number of individuals genotyped. Rather than genotyping a sample large enough to obtain
the desired statistical power, a smaller sample is genotyped. Further genotyping will not be
between inbred lines or half-sib families. Of course at these distances the markers will be
unlimited, and costs of phenotyping are low relative to genotyping costs, then marker spacing
of close to 80 cM between will give maximum statistical power per unit cost for crosses

Finally, Motro and Soller (1993) suggested sequential sampling as a further tool to reduce
There is interest of QTL identified for a given trait, which can be combined into a panel of
microsatellites it is now possible to develop virtually unlimited numbers of markers.
individuals genotyped times the number of markers genotyped for individual. Darvasi and
methods is most useful if the number of quantitative traits of interest is relatively low. Since
determine accurately the number of individuals of each genotype in a pool from the band

In this case a linked QTL is detected by band intensity when pools of individuals with "high"
individuals with similar phenotypes for the quantitative trait can be combined prior to assay.
That is, instead of genotyping each individual separately, genetic material from several
at the cost of each marker genotype, optimum marker spacing is still close to 30 cM and

deterministically computed the 95% confidence intervals, as described above. In addition to

4.6 Optimization of Experimental Designs

Optimization of experimental designs to obtain maximum power per unit cost will depend on the
the number of genotype assays performed: selective genotyping, sample pooling, and

The major cost elements of QTL detection are producing the individuals for analysis, scoring
produced and scored for quantitative traits, as compared to designs in which all individuals

with 2000 individuals. For the BC design and a marker bracket of 50 CM, a similar SE was
Discrepancies increased with decrease in sample size. Confidence intervals were largest for

For a complete genome search, the total number of genotypes will be the number of
producing the individuals for analysis, scoring of these traits, and data analysis. Optimization of
produced and scored for quantitative traits are also genotyped. Unlike replicate progeny, these other

Deterministic estimates based on assuming that all other parameters were fixed tended to
differentials. As for the BC design with a single marker, the matrix of second differentials

The standard error for r with a substitution effect of 0.5 was about 0.1

R2, and the residual variance, they also estimated the QTL allele frequencies. Thus, if the sample of individuals recorded for the quantitative trait is

interval" method, despite the deficiencies considered above, may be the method of choice for

In addition to replicate progeny, which we considered above, several other techniques

The disadvantages of this method for increasing the number of genotypes are larger size of the

Weller and Wyler, 1992). The opposite should occur. It should be noted though, that even for very large samples, the

Third, estimates of the QTL effect will be biased by a
deterministically computed the 95% confidence intervals, as described above. In addition to

techniques are trait specific.

scored for the quantitative traits are also genotyped. Unlike replicate progeny, these other

The likelihood function can behave marked differently for other
differences. As for the BC design with a single marker, the matrix of second differentials

interval" method, despite the deficiencies considered above, may be the method of choice for

All of these techniques require increasing the number of individuals in the analysis.

The major cost elements of QTL detection are producing the individuals for analysis, scoring
produced and scored for quantitative traits, as compared to designs in which all individuals

with 2000 individuals. For the BC design and a marker bracket of 50 CM, a similar SE was

Discrepancies increased with decrease in sample size. Confidence intervals were largest for

For a complete genome search, the total number of genotypes will be the number of
producing the individuals for analysis, scoring of these traits, and data analysis. Optimization of
produced and scored for quantitative traits are also genotyped. Unlike replicate progeny, these other

Deterministic estimates based on assuming that all other parameters were fixed tended to
differentials. As for the BC design with a single marker, the matrix of second differentials

The standard error for r with a substitution effect of 0.5 was about 0.1

R2, and the residual variance, they also estimated the QTL allele frequencies. Thus, if the sample of individuals recorded for the quantitative trait is

interval" method, despite the deficiencies considered above, may be the method of choice for

In addition to replicate progeny, which we considered above, several other techniques

The disadvantages of this method for increasing the number of genotypes are larger size of the

Weller and Wyler, 1992). The opposite should occur. It should be noted though, that even for very large samples, the

Third, estimates of the QTL effect will be biased by a
required for those markers that clearly show no significant effect or that show a significant effect. Additional individuals will be genotyped only for those markers that display "borderline" significance. By this method it is possible to reduce the total number of genotypings required by nearly half for a single trait. However, if the number of traits is large, nearly every marker will have borderline significance for at least one quantitative trait. Thus, like selective genotyping and sample pooling, this method is useful only if the number of traits under consideration is small.

4.7 Summary

Numerous misconceptions with respect to the power of QTL detection and experiment design optimization are prevalent. In most cases power to detect a segregating QTL of a magnitude likely to be segregating in the population will require genotyping at least 500 individuals, and often many more. Most experiments have been too small to find effects of the magnitude that could be reasonably expected. Unless the phenotyping costs are very high relative to genotyping costs, experimental designs with very wide marker spacing are optimum, and decreasing marker intervals below 20 cM will have virtually no effect for most experimental designs. Power per individual genotype can be dramatically increased by replicate progeny, selective genotyping, sample pooling and sequential sampling, and the effect of these techniques are cumulative. Except for replicate progeny, these other techniques are trait specific, and are therefore most appropriate for experiments that consider only a few traits.

References


