

Prediction of the Confidence Interval of Quantitative Trait Loci Location

Peter M. Visscher^{1,3} and Mike E. Goddard²

Received 4 Sept. 2003—Final 28 Nov. 2003

In 1997, Darvasi and Soller presented empirical predictions of the confidence interval of quantitative trait loci (QTL) location for dense marker maps in experimental crosses. They showed from simulation results for backcross and F_2 populations from inbred lines that the 95% confidence interval was a simple function of sample size and the effect of the QTL. In this study, we derive by theory simple equations that can be used to predict any confidence interval and show that for the 95% confidence interval, they are in good agreement with the empirical results given by Darvasi and Soller. A general form of the confidence interval is given that also applies to other population structures (*e.g.*, collections of sib pairs). Furthermore, the expected shape of the likelihood-ratio-test around the true QTL location is derived, which is shown to be extremely leptokurtic. It is shown that this shape explains why confidence intervals from the Log of Odds (LOD) drop-off method and bootstrap results frequently differ for real data sets.

KEY WORDS: Bootstrap; confidence interval; likelihood-ratio-test; QTL; QTL mapping.

INTRODUCTION

Genome scans to detect quantitative trait loci (QTL) or targeted scans to attempt to replicate previously reported QTL are now commonplace, both in experimental and outbred populations. Estimation of the position of a QTL is usually achieved with low precision, even in controlled crosses from inbred lines, but the accuracy of the position is important, either for subsequent introgression or fine-mapping (*e.g.*, Lynch and Walsh, 1998). Darvasi and Soller (1997) presented simulation results and empirical predictions of the 95% confidence interval (CI) of QTL locations for crosses between (inbred) lines in the case of an infinitely dense marker map. A summary of their simulation results gave

a remarkably simple prediction of the 95% CI (CI95, in cM) for backcross (BC) and F_2 populations,

$$CI95_{BC} = 3000/[n(\alpha + \delta)^2]$$

$$CI95_{F_2} = 1500/(n\alpha^2)$$

with α and δ the additive and dominance effect of the QTL in residual standard deviation units, respectively, and n the sample size. They also presented a general prediction relating the CI to the proportion of variance explained by the QTL (q^2),

$$CI95 = 530/(nq^2).$$

These results were based on testing for an additive effect only (*i.e.*, the test statistic for QTL detection was based on one degree of freedom). In addition, the empirical predictions were obtained for cases when the observed CI was <20 cM (*i.e.*, for large sample sizes and/or large QTL effects). The average proportion of variance explained by the QTL in the scenarios (sample size and QTL effects) investigated by Darvasi and Soller (1997) corresponding to an observed CI of less than 20 cM was approximately 30% and 40% for backcross and F_2 populations, respectively.

¹ Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh, UK.

² Department of Primary Industry, Attwood, and University of Melbourne, Melbourne, Australia.

³ To whom correspondence should be addressed at Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK. Tel: +44 131 650 7702. Fax: +44 131 650 6564. e-mail: peter.visscher@ed.ac.uk

The results from Darvasi and Soller (1997) can be used in combination with power studies before an experiment to assess the expected confidence region of a QTL given its effect and the sample size of the experiment. After an experiment is conducted and a QTL is detected, either the LOD drop-off method (Lander and Botstein, 1989) or a bootstrap procedure (Talbot *et al.*, 1999; Visscher *et al.*, 1996) are used to estimate the confidence interval of a QTL.

In this study, we derive simple theoretical formulae for the confidence interval and show that they agree with the empirical results obtained by Darvasi and Soller (1997) when the confidence interval is small. In addition, we derive the shape of the likelihood profile around the true QTL location and demonstrate why the asymptotic theory does not hold for dense marker maps, and why bootstrap CI are often larger than CI calculated from a 1-LOD (Log of Odds) support interval.

THEORY

Prediction of Confidence Intervals

We consider a backcross or F_2 population of size n and a segregating QTL with an additive and dominance effect of α and δ , respectively. The variance due to the QTL is $\sigma_q^2 = \frac{1}{4}(\alpha - \delta)^2$ for the backcross population and $\sigma_q^2 = \frac{1}{2}\alpha^2 + \frac{1}{4}\delta^2$ for the F_2 population. The residual variance, which is due to environmental effects and unlinked genetic effects, is $\sigma_e^2 = 1$, so that the QTL effects are expressed in residual standard deviation units. A dense marker map is assumed so that, for any position on the chromosome, we can determine whether an offspring inherited the paternal or maternal allele of its mother and father. Assume that an F_1 individual, or, for example, a common parent in a half-sib design, has genotype MQ/mq, where M is a marker allele on the same chromosome as QTL allele Q, and m and q are the marker and QTL allele on the other chromosome, respectively. We denote Y_M , Y_m , Y_Q , and Y_q the progeny mean of individuals inheriting marker allele M and m, and QTL allele Q and q, respectively. Similarly, Y_{MQ} , Y_{mq} , Y_{Mq} , and Y_{mQ} are the means of the progeny inheriting haplotypes MQ, mq, Mq, and mQ, respectively. The last two are the recombinant haplotypes. The recombination fraction between the QTL and marker is r . It follows that,

$$\begin{aligned} Y_M &= (1-r)Y_{MQ} + rY_{Mq} \\ Y_m &= (1-r)Y_{mq} + rY_{mQ} \\ Y_Q &= (1-r)Y_{MQ} + rY_{mQ} \\ Y_q &= (1-r)Y_{mq} + rY_{Mq} \end{aligned}$$

The normal approach to mapping a QTL is to calculate either a likelihood ratio test (LRT) or an F statistic for a putative QTL located at regular points along the chromosome, such as every cM. The position with the highest test statistic is the estimated position of the QTL. Results from simulation studies and analyses of data sets have shown that LRT and F tests give very similar results in terms of type-I and type-II errors and parameter estimates (e.g., Lynch and Walsh, 1998). We will use a squared t -statistic, which is equivalent to an F-test and distributed as a chi-square if we ignore error in estimating the denominator mean square. That is,

$$T_{\text{marker}} = (Y_M - Y_m)^2 / \text{var}(Y_M - Y_m),$$

with $\text{var}(Y_M - Y_m) \approx 4\sigma_e^2/n$, and n the progeny sample size. This approximation ignores within-marker segregation due to the QTL and is valid for either a small recombination fraction between marker and QTL or for a small effect due to the QTL. This is frequently assumed in theoretical calculations for QTL mapping studies (e.g., Lander and Botstein, 1989; Lynch and Walsh, 1998; Soller *et al.*, 1976). Similarly, the test statistic for a marker at the QTL is $T_{\text{QTL}} = (Y_Q - Y_q)^2 / (4\sigma_e^2/n)$. For the estimated position of the QTL to be at the marker, the test statistic at the marker must be greater than that at the QTL, that is

$$D = (Y_M - Y_m)^2 - (Y_Q - Y_q)^2 > 0.$$

Decomposing the allelic means into haplotype means shown above gives an equivalent expression,

$$D' = D/[4r(1-r)] = (Y_{MQ} - Y_{mq})(Y_{Mq} - Y_{mQ}) > 0.$$

Note that this equation implies that the recombinant offspring give the “wrong” estimate of the sign of the QTL effect, or, usually much less likely, that the non-recombinant mean difference is in the opposite direction from the true value. For an additive QTL, the mean and variance of the statistic D' are,

$$\begin{aligned} E(D') &= -\alpha^2 \text{ and } \text{Var}(D') = 4\alpha^2/[nr(1-r)] \\ &= 16\alpha^2/[(1-e^{-4d})n], \end{aligned}$$

if we assume Haldane’s mapping function and d is map distance in Morgan. A full derivation of these formulae is given in Appendix A.

The boundary of the $(1-\beta)$ confidence interval is the position on the chromosome at which $P(D' > 0) = \beta/2$. The factor of two (in $\beta/2$) is necessary because we wish to calculate the confidence interval to the left and right of the QTL. Asymptotically, $Z = [D' - E(D')]/\sigma[D']$ is normally distributed as a standard normal variate, so $D' > 0$ is equivalent to $Z > -E[D']/\sigma[D']$ and the probability of these can be

calculated from standard normal tables. Hence, the $1 - \beta$ confidence interval ends at a distance d from the true QTL position such that

$$\begin{aligned} \text{Prob}(Z > -E[D']/\sigma[D']) &= \beta/2, \quad \text{or} \\ \frac{1}{4}\alpha[n(1 - e^{-4d})]^{\frac{1}{2}} &= k_{\beta/2}, \end{aligned}$$

with $k_{\beta/2}$ the $\beta/2$ cut-off point of the normal distribution. For example, for $\beta = 0.05$, $k_{0.025} = 1.96$. Solving for d gives

$$d = -0.25 \log(1 - 16 k_{\beta/2}^2 / n\alpha^2)$$

If the confidence interval is small, using $\log(1 - x) \approx -x$ for small x , gives

$$d \approx 4X_{(1-\beta)}/(n\alpha^2),$$

with $X_{(1-\beta)}$ the threshold of a central χ^2 distribution with 1 degree of freedom corresponding to a cumulative density of $(1 - \beta)$. The confidence interval in cM is,

$$\begin{aligned} \text{CI}(1 - \beta) &= 2(100d) \\ &= -50 \log[1 - 16X_{(1-\beta)}/(n\alpha^2)] \\ &\approx 800X_{(1-\beta)}/(n\alpha^2). \end{aligned}$$

For F_2 populations, the two homozygous genotypes (on average half of the number of individuals in the population) are used in the test for an additive effect but the effect (i.e., the difference between the two homozygotes) is twice as large as in a backcross. The exact expression for F_2 populations is,

$$\begin{aligned} \text{CI}(1 - \beta) &= 2(100d) \\ &= -50 \log[1 - 8X_{(1-\beta)}/(n\alpha^2)] \\ &\approx 400X_{(1-\beta)}/(n\alpha^2). \end{aligned}$$

Combining the results for backcross (or a single sire in a halfsib design) and F_2 populations gives,

$$\text{CI}(1 - \beta) \approx (200x)X_{(1-\beta)}/(n\alpha^2),$$

with $x = 2$ for F_2 and $x = 4$ for backcross populations. For example, for BC and F_2 population the 95% confidence interval are predicted as,

$$\text{CI}_{95\text{BC}} \approx (200)(4)(3.84)/(n\alpha^2) = 3073/(n\alpha^2), \quad \text{and}$$

$$\text{CI}_{95\text{F}_2} \approx (200)(2)(3.84)/(n\alpha^2) = 1537/(n\alpha^2).$$

These are very close to the empirical prediction of Darvasi and Soller (1997) for 95% confidence intervals. The prediction of the CI as a function of the proportion of the variance explained by the QTL is,

$$\text{CI} \approx 200X_{(1-\beta)}(1 - q^2)/(nq^2)$$

For example, the 95% CI for a QTL that explains 35% of the variation in either a backcross or F_2 population, the average QTL heritability on which the Darvasi and Soller (1997) prediction was based, is,

$$\text{CI}_{95} \approx (200)(3.84)(1 - 0.35)/(nq^2) = 499/(nq^2),$$

again very close to the empirical results from Darvasi and Soller (1997).

A general form of the prediction of the CI for dense marker maps that also applies to other population structures is,

$$\text{CI} \approx 200X_{(1-\beta)}/\lambda,$$

where λ is the non-centrality parameter (NCP) of a χ^2 test for the presence of a QTL at the true QTL location. For example, for collections of n sib pairs, the approximate NCP for linkage using either maximum likelihood or a combination of sibling squared differences and squared sums is, $\lambda = nq^4/8$ (e.g., Visscher and Hopper, 2001). Hence, the predicted 95% CI for sib pairs is,

$$\text{CI}_{95} \approx 1600X_{(1-\beta)}/(nq^4) = 6144/(nq^4).$$

For 1000 sib pairs and a QTL explaining 50% of the phenotypic variance, the predicted 95% CI is about 25 cM, confirming that sib pair designs are not powerful for precise QTL localization, unless the sample size is very large (10,000s) and the effect of the QTL is large.

Likelihood Curvature Around the QTL

In this section, we derive the expected log-likelihood value at distance d from the QTL. Consider a backcross population from inbred lines, with a dense set of fully informative markers (with genotypes MM and Mm). The mean phenotype of the marker genotypes for location d are $Y_{M(d)}$ and $Y_{m(d)}$, respectively. At any marker location, the full model of analysis is,

$$y = \alpha_d c + e$$

with α_d the additive effect at location d and c an indicator variable. For marker genotypes MM and Mm, c is $\frac{1}{2}$ and $-\frac{1}{2}$, respectively. Hence, $E(c) = 0$, $E(c^2) = \frac{1}{4}$ and $\Sigma c^2 = n \text{var}(c) = \frac{1}{4}n$. The log-likelihood function is,

$$L = \text{constant} - \frac{1}{2}n \log(\sigma^2) - \frac{1}{2}\Sigma(y - \alpha_d c)^2/\sigma^2.$$

At location d , the maximum likelihood value is proportional to,

$\text{ML}(d) = -\frac{1}{2}n \log(\hat{\sigma}^2) - \frac{1}{2}n$, with maximum likelihood estimators (MLE)

$$\text{MLE}(\hat{\alpha}_d) = 2[Y_{M(d)} - Y_{m(d)}], \quad \text{and}$$

$$\text{MLE}(\hat{\sigma}^2) = \Sigma(y - \hat{\alpha}_d c)^2/n.$$

The expectation of the ML estimate of the variance at location d is,

$$\begin{aligned} E[\hat{\sigma}_d^2] &= E[\Sigma(y^2)] - E[\hat{\alpha}_d^2] \Sigma c^2 \\ &= n[\sigma_e^2 + \sigma_q^2] - (1 - 2r)^2 \alpha^2 \left(\frac{1}{4}n\right) \\ &= n[\sigma_e^2 + (1 - e^{-4d})\sigma_q^2] \end{aligned}$$

Hence, to a first-order approximation, the expected value of the maximum log-likelihood value pertaining to location d is,

$$\begin{aligned} E[ML(d)] &\approx -\frac{1}{2}n \log(n[\sigma_e^2 + (1 - e^{-4d})\sigma_q^2]) - \frac{1}{2}n \\ &\approx -\frac{1}{2}n \log[\sigma_e^2 + (1 - e^{-4d})\sigma_q^2] + \text{constant} \\ &\approx -\frac{1}{2}n \log(1 - e^{-4d}q^2) + \text{constant} \end{aligned}$$

with $q^2 = \sigma_q^2/(\sigma_e^2 + \sigma_q^2)$, the QTL “heritability.” If we make the further assumption that the QTL variance is small, then $E[ML(d)] \approx \frac{1}{2}ne^{-4d}q^2$. Finally, the expected likelihood-ratio test to test locations d against location 0 (the QTL position) is,

$$\begin{aligned} E(\Delta) &= E\{2[ML(0) - ML(d)]\} \\ &\approx n \log[(1 - e^{-4d}q^2)/(1 - q^2)] \\ &= n \log[1 + (1 - e^{-4d})q^2/(1 - q^2)] \\ &\approx n(1 - e^{-4d})q^2/(1 - q^2) \\ &\approx n(1 - e^{-4d})q^2 \end{aligned}$$

Hence, to a first-order approximation, the expected drop-off in the test statistic is proportional to $[1 - (1 - 2r)^2] = (1 - e^{-4d})$. Around the location of the QTL (small r and $r = d$), the decline is linear whereas at greater distances it asymptotes (to a value of nq^2).

RESULTS AND DISCUSSION

We have shown that the empirical results obtained by Darvasi and Soller (1997) can be derived using large sample theory and therefore it can be generalized to predict the confidence interval for any given coverage probability. Unless the confidence interval is small it is better to use the formula with $1 - \exp(-4d)$ than $4d$. Table I gives examples of the predicted 90%, 95%, and 99% confidence intervals in F_2 populations for a range of parameters, calculated using the exact expression $\{CI(1 - \beta) = -50 \log[1 - 8X_{(1-\beta)}/(n\alpha^2)]\}$. The CI for backcross populations are approximately twice those reported in Table I. Unless the QTL effect is large ($\alpha = 1$) or the sample size is large ($n = 1000$), predicted confidence intervals are typically large,

Table I. Prediction of 90%, 95%, and 99% Confidence Intervals (in cM) in F_2 Populations

		CI90	CI95	CI99
α	n			
0.25	100	*	*	*
	500	59	205	*
	1000	21	34	95
0.50	100	100	*	*
	500	10	14	28
	1000	5	7	12
1.00	100	12	18	38
	500	2	3	6
	1000	1	2	3

*No solution, confidence interval is infinite.

spanning 9 to >100 cM. For a number of combinations of parameters, the confidence interval is infinite because even at a very large distance from the QTL, there is a large enough probability for the test statistic to be greater than the test statistic at the true QTL position. Figure 1 shows the predicted 95% confidence interval as a function of the proportion of variance explained by the QTL, q^2 , for population sizes of 100, 500, and 1000. The confidence interval was calculated as $CI95 = -50 \ln[1 - (4)(3.84)(1 - q^2)/(nq^2)]$, with n the sample size and 3.84 the 95% threshold of a χ^2 distribution with 1 degree of freedom. Note that because the parameterization is in the proportion of variance explained by the QTL, the results are identical for either a backcross or F_2 population. For an additive model this implies that the size of the QTL effect (in SD units)

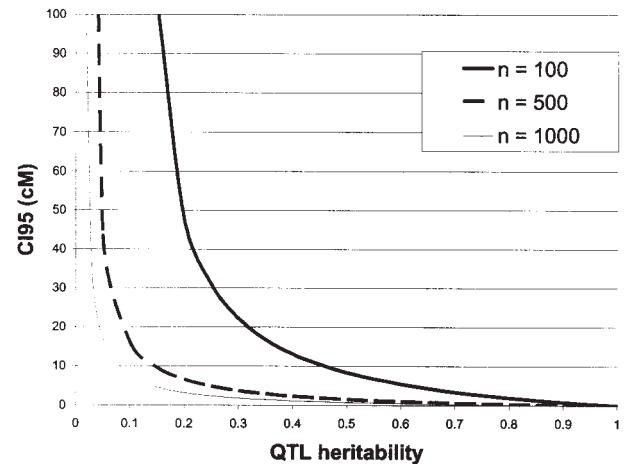


Fig. 1. Predicted 95% confidence of QTL location (in cM) in a backcross or F_2 population as a function of sample size (n) and the QTL heritability.

in the backcross is $\sqrt{2}$ larger than the effect in an F_2 population.

Roberts *et al.* (1999) also considered the confidence interval of trait locus position based on simulation of an affected sib pair design. They summarized their simulation results by an empirical formula for the standard error of trait locus position (SEL), $SEL = 7.02(ELOD)^{-0.59}$, where ELOD is the expected LOD score based on repeated simulations of an experiment. Assuming the 95% CI is four times SEL, and using our result for $E(\log L)$, we can write the Roberts *et al.* formula as,

$$\begin{aligned} CI95(\text{Roberts}) &= (4)(7.02)[(1 + \lambda)/4.6]^{-0.59} \\ &= 69.1(1 + \lambda)^{-0.59}, \end{aligned}$$

where λ is the non-centrality parameter of a non-central χ^2 distribution with 1 degree of freedom. This can be compared with our approximate formula for a sib pair design which is $CI95 = 768/\lambda$. At low values of λ , the Roberts *et al.* (1999) formula yields much lower CI95. For instance, with 500 sib-pairs and a QTL heritability of 0.1, $\lambda = 0.65$, the ELOD is 0.35 and $CI95(\text{Roberts}) = 51$ cM, whereas our formula gives $CI95 = 1229$ cM. This discrepancy is largely due to the measure of SEL preferred by Roberts *et al.* (1999). They preferred a measure of SEL based on the curvature of the ELOD near the true QTL position. This approach gave SEL several times smaller than the empirical variation in estimated position of the QTL across replicate simulations. Roberts *et al.* (1999) prefer the method based on ELOD curvature because it is not affected by the length of the chromosome. However, we do not think this is an adequate reason for using a formula for SEL that clearly underestimates the empirical variation in the estimated position of the QTL. Roberts *et al.* Figure 2(A) clearly shows the long tails on the distribution of estimated QTL position, so that the 95% CI would be almost the whole chromosome. Our formula for CI95 assumes an infinitely long chromosome but the $CI95 = 1229$ cM leads to the conclusion that, for a real chromosome of length, say, 100 cM, the CI95 occupies the whole chromosome. We simulated the above sib pair design (500 pairs, fully informative marker, $q^2 = 0.10$) and obtained an empirical power of 0.22 and 0.07 at a type-I error rate of 0.05 and 0.01, respectively, consistent with predictions of Visscher and Hopper (2001). The extremely low power of this design therefore also suggests that an expected 95% confidence interval of 51 cM is too small.

After a QTL mapping experiment has been completed, the CI for the position of the QTL can be

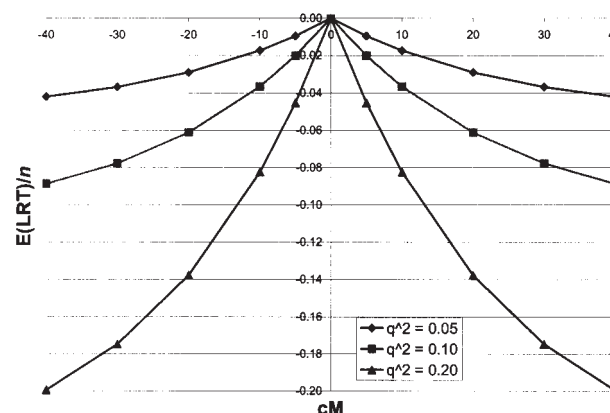


Fig. 2. Expected curvature of the likelihood profile around the true location of a QTL. The x-axis denotes the distance (cM) from the QTL. The y-axis is twice the difference between the expected log-likelihood values at the marker and QTL, scaled by the sample size n . The three curves are for QTL that explain 5%, 10%, and 20% of the variance, respectively.

calculated either by the LOD drop-off method or by bootstrap. In theory, the 1-LOD support interval approximately corresponds to a 97% confidence interval, because 1 LOD is equivalent to a likelihood-ratio-test statistic of 4.605, which corresponds to an asymptotic χ^2 p -value of 0.032 (Mangin *et al.*, 1994). However, simulation studies (Van Ooijen, 1992; Visscher *et al.*, 1996) have shown that the LOD support interval is relatively too short, that is, the proportion of CI97 that contains the QTL is less than 0.97. Van Ooijen (1992) recommended that a 2-LOD support interval should be used. Bootstrap procedures (Lebreton and Visscher, 1998; Visscher *et al.*, 1996) are generally slightly conservative. Unless the sample size is very large (1000s) or the QTL effect is very large (> 1 standard deviation), the length of the confidence intervals calculated from LOD drop methods or bootstrap methods can be widely different, sometimes by a factor of two. One of the reasons for this is that the LOD drop method is based on asymptotic theory, which states that the curvature of the log-likelihood function around the true value is quadratic. Figure 2 shows that the expected log-likelihood, scaled by the sample size, is extremely leptokurtic and is clearly not quadratic in d . A quadratic approximation of the likelihood function around the maximum will ignore the long tails of the leptokurtic distribution. The bootstrap method will sample these long tails because it is based on observed data. In our opinion the asymptotic theory of the likelihood function does not apply here despite the large total sample size (n) because it requires a large number of recombinants between adjacent markers. As the formula for D' and its variance

make clear, it is the number of recombinants that is critical. For instance, $\text{var}(D')$ is proportional to $1/(nr(1-r))$, and nr is the number of recombinants. As the sample size n increases so does the number of recombinants between any two fixed points, but the CI decreases and so the number of recombinants in smaller and smaller intervals becomes relevant. Consequently, the likelihood function based on a dense marker map never approaches the quadratic predicted by asymptotic theory. With markers spaced every 10 or 20 cM and a large sample size (i.e., many recombinants between adjacent markers), the expected log likelihood is quadratic in the interval between the markers flanking the QTL (e.g., Roberts *et al.*, 1999). If the sample size is large enough, the leptokurtic tails are very thin, and the LOD drop and bootstrap methods give similar confidence intervals (Walling *et al.*, 2000). However, as the density of markers is increased, Figure 2 shows that the expected log-likelihood does not remain quadratic. Therefore, our results imply that asymptotic results from dense marker maps should be interpreted with caution. However, a denser map should still give a smaller confidence interval than a sparser marker map, because more information is used to detect recombination events.

ACKNOWLEDGMENTS

This work was supported by the UK Biotechnology and Biological Sciences Research Council (PMV) and the Department of Primary Industries in Victoria, Australia (M.E.G.).

APPENDIX A

Derivation of the Mean and Variance of D'

The definition of D and D' are,

$$D = (Y_M - Y_m)^2 - (Y_Q - Y_q)^2 \text{ and}$$

$$D' = D/[4r(1-r)] = (Y_{MQ} - Y_{mq})(Y_{Mq} - Y_{mQ})$$

As in the main text, we assume that $\text{var}(Y_M - Y_m) \approx 4\sigma_e^2/n = \text{var}(Y_Q - Y_q)$.

$$\begin{aligned} E(D) &= \text{var}(Y_M - Y_m) + E^2(Y_M - Y_m) \\ &\quad - \text{var}(Y_Q - Y_q) - E^2(Y_Q - Y_q) \\ &= 4\sigma_e^2/n + (1-2r)^2\alpha^2 - 4\sigma_e^2/n - \alpha^2 \\ &= -\alpha^2[4r(1-r)], \quad \text{and } E(D') = -\alpha^2 \end{aligned}$$

$$\begin{aligned} \text{var}(D') &= \text{var}(Y_{MQ} - Y_{mq})(Y_{Mq} - Y_{mQ}) \\ &= \text{var}(Y_1 Y_2) \end{aligned}$$

Using the expression $\text{var}(Y_1 Y_2) \approx E^2(Y_1)\text{var}(Y_2) + E^2(Y_2)\text{var}(Y_1) + 2E(Y_1)E(Y_2)\text{cov}(Y_1, Y_2)$ and that here the covariance is zero, that $E(Y_1) = \alpha = -E(Y_2)$ and that $\text{var}(Y_1) = 4\sigma_e^2/[(1-r)n]$ and $\text{var}(Y_2) = 4\sigma_e^2/(rn)$, gives,

$$\begin{aligned} \text{var}(D') &= \alpha^2 4\sigma_e^2/(rn) + (-\alpha)^2 4\sigma_e^2/[(1-r)n] \\ &= [4\alpha^2 \sigma_e^2/n]/[r(1-r)]. \end{aligned}$$

In Haldane's mapping function, $r = \frac{1}{2}(1 - e^{-2d})$, so $r(1-r) = \frac{1}{4}(1 - e^{-4d})$. Using this in the above expression, and letting $\sigma_e = 1$ gives the equation from the main text,

$$\text{var}(D') = 16\alpha^2/[(1 - e^{-4d})n].$$

REFERENCES

- Darvasi, A., and Soller, M. (1997). A simple method to calculate resolving power and confidence interval of QTL map location. *Behav. Genet.* **27**:125-132.
- Lander, E. S., and Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, **121**:185-199.
- Lebreton, C. M., and Visscher, P. M. (1998). Empirical nonparametric bootstrap strategies in quantitative trait loci mapping: conditioning on the genetic model. *Genetics* **148**:525-535.
- Lynch, M., and Walsh, B. (1998). *Genetics and analysis of quantitative traits*, Sunderland: Sinauer Associates, MA.
- Mangin, B., Goffinet, B., and Rebai, A. (1994). Constructing confidence intervals for QTL location. *Genetics* **138**:1301-1308.
- Roberts, S. B., MacLean, C. J., Neale, M. C., Eaves, L. J., and Kendler, K. S. (1999). Replication of linkage studies of complex traits: an examination of variation in location estimates. *Am. J. Hum. Genet.* **65**:876-884.
- Soller, M., Brody, T., and Genizi, A. (1976). On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theor. Appl. Genet.* **47**:35-39.
- Talbot, C. J., Nicod, A., Cherny, S. S., Fulker, D. W., Collins, A. C., and Flint, J. (1999). High-resolution mapping of quantitative trait loci in outbred mice. *Nat. Genet.* **21**:305-308.
- Van Ooijen, J. W. (1992). Accuracy of mapping quantitative trait loci in autogamous species. *Theor. Appl. Genet.* **84**:803-811.
- Visscher, P. M., and Hopper, J. L. (2001). Power of regression and maximum likelihood methods to map QTL from sib-pair and DZ twin data. *Ann. Hum. Genet.* **65**:583-601.
- Visscher, P. M., Thompson, R., and Haley, C. S. (1996). Confidence intervals in QTL mapping by bootstrapping. *Genetics* **143**:1013-1020.
- Walling, G. A., Visscher, P. M., Andersson, L., Rothschild, M. F., Wang, L., Moser, G., Groenen, M. A., Bidanel, J. P., Cepica, S., Archibald, A. L., Geldermann, H., de Koning, D. J., Milan, D., and Haley, C. S. (2000). Combined analyses of data from quantitative trait loci mapping studies. Chromosome 4 effects on porcine growth and fatness. *Genetics* **155**:1369-1378.

Edited by Stacey Cherny