Genet. Res., Camb. (2003), **82**, *pp.* 139–149. *With 5 figures.*
DOI: 10.1017/S0016672303006396 Printed in the United Kingdom

Inferences regarding the numbers and locations of QTLs under multiple-QTL models using interval mapping and composite interval mapping

THEODORE W. CORNFORTH*† AND ANTHONY D. LONG
Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697-2525, USA

(*Received 3 March 2003 and in revised form 24 June 2003*)

Summary

This paper examines the properties of likelihood maps generated by interval mapping (IM) and composite interval mapping (CIM), two widely used methods for detecting quantitative trait loci (QTLs). We evaluate the usefulness of interpretations of entire maps, rather than only evaluating summary statistics that consider isolated features of maps. A simulation study was performed in which traits with varying genetic architectures, including 20–40 QTLs per chromosome, were examined with both IM and CIM under different marker densities and sample sizes. IM was found to be an unreliable tool for precise estimation of the number and locations of individual QTLs, although it has greater power for simply detecting the presence of QTLs than CIM. The ability of CIM to resolve the correct number of QTLs and to estimate their locations correctly is good if there are three or fewer QTLs per 100 centiMorgans, but can lead to erroneous inferences for more complex architectures. When the underlying genetic architecture of a trait consists of several QTLs with randomly distributed effects and locations likelihood profiles were often indicative of a few underlying genes of large effect. Studies that have detected more than a few QTLs per chromosome should be interpreted with caution.

1. Introduction

An understanding of the nature of genes affecting quantitative traits has been a fundamental goal of research in quantitative genetics at least since the work of Sax (1923) but, for the most part, it remains elusive. However, approaches that take advantage of modern high-throughput genotyping technologies have brought this understanding and its attendant benefits to medicine and agriculture nearer than ever (Mackay, 2001). Two methods are widely used because of their power and computational efficiency: interval mapping (IM) (Lander & Botstein, 1989) and composite interval mapping (CIM) (Zeng, 1994). The advantages of these approaches include increased statistical power over single marker analyses (Thoday, 1961), greater algorithmic simplicity than multiple interval mapping (MIM) (Kao *et al.*, 1999) and lower computational burden than Bayesian interval mapping methods (Sillanpää & Arjas, 1998).

* Corresponding author. Tel: +1 541 346 4556. Fax: +1 541 346 4548. e-mail: cornforth@uoneuro.uoregon.edu
† Present address: Institute of Neuroscience, 1254 University of Oregon, Eugene, OR 97403-1254, USA.

In both IM and CIM, molecular marker data and phenotype data are collected from a number of individuals and are used to calculate the likelihood, L_1, that a QTL is present at some given location in the genome covered by the linkage map. The likelihood of no QTL being present, L_0, is also determined. In this study, the likelihood information is summarized with a likelihood ratio (LR), which equals $2\ln(L_1/L_0)$. The LR statistic is usually calculated across the entire genome at positions 1 cM apart and then visualized by plotting LR against map location. Ideally, the resulting likelihood map leads to correct inferences regarding the number and location of factors affecting the trait of interest by identifying those portions of the genome showing higher, statistically significant LRs than other portions.

Over the past decade, hundreds of studies have used IM and CIM to reach conclusions regarding the locations and effects of genes contributing to variation in a wide variety of traits in many species (reviewed in Lynch & Walsh, 1998). Thus, it is of great interest to investigate the power and accuracy of these methods in order to develop a clearer picture

Table 1. *Simulation parameters summary*

Number of QTLs[a]	Conditions[a]	Heritability	Locations of QTLs (cM)	Relative effects of QTLs[b]
1	64	0·05	1 48	+1·0
2	288	0·10	18, 78 38, 58 45, 50	+1·0, +1·0 +0·5, +2·0 −1·0, +1·0
3	32	0·15	18, 48, 78	+1·0, +1·0, +1·0
5[c]	25	0·45	10·0, 29·0, 47·5, 69·5, 90·0	+3·3, +2·7, +5·5, +2·8, +1·9
20	128	0·50	Equal[d] Random[f]	Equal[e] Random[g]
40	128	0·50	Equal Random	Equal Random

[a] For each QTL number, simulations were performed for each possible combination of QTL location and QTL effect. Each such combination was in turn performed for each combination of mapping method, marker density and sample size (as described in the text) to give the total number of conditions.
[b] The effects correspond to the QTL locations in order from left to right.
[c] Data estimated from Long *et al.* (1995).
[d] For 'equal spacing' in 20 QTL conditions, the QTLs are spaced every 5 cM and, for equal spacing in 40 QTL conditions, 40 QTLs are evenly spaced every 2·5 cM.
[e] In 'equal effects' conditions, all QTLs have the same relative effect.
[f] For 'random spacing' conditions, 20 or 40 QTL locations were randomly chosen and all maps with the same number of QTLs used the same randomly chosen set.
[g] For 'random effects' conditions, all QTLs had random uniform effects and all maps with the same number of QTL used the same randomly chosen set.

of the appropriate confidence or skepticism with which conclusions drawn from mapping results should be met.

Beginning with Lander & Botstein (1989), many studies have used computer simulation and other techniques to investigate the properties of IM (Van Ooijen, 1992; Carbonell *et al.*, 1993; Darvasi *et al.*, 1993; Zeng, 1994; Wright & Kong, 1997; Ronin *et al.*, 1999; Walling *et al.*, 2002). CIM has not been as extensively examined (Zeng, 1994; Visscher & Haley, 1996; Goffinet & Mangin, 1998; Visscher *et al.*, 2000). Unfortunately, it is often difficult to compare and summarize results across such studies owing to the range of approaches and simulation parameters that are used. Also, it is well known that interval mapping can fail to distinguish between one QTL of large effect and several QTLs of small effect (Visscher & Haley, 1996; Liu & Dekkers, 1998), yet most simulation studies have investigated relatively simple genetic architectures. A third shortcoming of previous studies is that they rely on summary statistics that take into account only one or two aspects of likelihood maps.

In the present study, all previously examined genetic architectures, as well as several more complicated architectures, are analyzed in a unified, consistent framework for both IM and CIM. Furthermore, an attempt was made to investigate, quantify and summarize the properties of whole likelihood maps and not to rely solely on summary statistics. Thus, we place an emphasis on evaluating the usefulness of IM and CIM in a 'real world' context and address their utility to experimenters who frequently use likelihood maps as the primary evidence for the existence and location of QTLs. In this spirit, we also provide an extensive downloadable database of likelihood maps and other data generated under known genetic architectures.

2. Methods

To investigate the performance of IM and CIM under conditions representative of those likely to be encountered in QTL mapping studies, quantitative traits were mapped in simulated backcross (BC_1) populations whose members were each represented by a linkage group ('chromosome') of length 100 cM (cf Van Ooijen, 1992). Phenotypes were simulated as the sum of the additive genetic effects of any QTLs present, plus a random normal environmental effect with variance chosen to give the traits the heritabilities listed in Table 1.

For CIM, a cofactor selection method was used whereby all markers more than 10 cM away from the interval in question were considered (cf Zeng 1994). Among these, markers at least 10 cM from all other markers were chosen as cofactors, starting with the marker farthest away from the interval on the 'left' side. When the interval was reached, cofactors were

chosen, starting with the marker farthest away from the interval on the right. This method was found to give results similar to the commonly used method of selecting as cofactors the five or so markers of greatest effect outside a window around the interval with stepwise regression (Basten *et al*., 1999), yet is less computationally intensive (data not shown).

A total of 665 mapping conditions were simulated, each with 500 replications. The parameter sets for each of these conditions can be organized into six categories (Table 1). Unless otherwise noted, simulations were performed for each combination of the following parameters in addition to each combination of QTL location and QTL effect described in Table 1: mapping method {IM, CIM}, marker density {1 cM, random, 7 cM, 20 cM} and sample size {100, 200, 500, 1000}. Random marker density is a condition with markers located at 0·0, 1·1, 3·6, 13·8, 24·2, 35·5, 52·5, 62·9, 73·1, 73·4, 84·9, 89·7, 91·2, 94·6 and 100·0 cM from the left end of the chromosome, giving an average spacing of 7·1 cM.

20 further mapping conditions were simulated with parameters chosen based on the mapping conditions and positions and effects of QTLs mapped by Long *et al*. (1995), who detected five QTLs affecting abdominal bristle number on the third chromosome of *Drosophila melanogaster* (Table 1). (QTLs were simulated to have only additive effects, although Long *et al*. (1995) found evidence of epistasis.) Unlike with the other five condition categories, the simulations were performed with CIM only for each combination of marker density {Long, 1 cM, random, 7 cM, 20 cM}, and sample size {66, 100, 200, 500, 1000}. 'Long' marker density approximates the location of markers used and 66 was the sample size used the analysis by Long *et al*. (1995) of chromosome three. Because these markers spanned a linkage group 108 cM long, the standard 1 cM, 7 cM, random and 20 cM marker densities also spanned this length instead of the usual 100 cM. (See results in web supplement p. 48 (power and significance proportion data), pp. 89–93 (fractile curves) and pp. 1374–1423 (sample plots).)

For each mapping condition, the 10th, 25th, 50th, 75th and 90th percentile LRs for the 500 generated likelihood maps at each position along the chromosome were plotted (cf Visscher & Haley, 1996). For reference, the appropriate empirical 5% significance threshold was plotted along with the five fractile curves. These thresholds were calculated based on the maximum LR values obtained from 1000 likelihood maps generated under conditions identical to the corresponding mapping condition except that no QTL was present (cf Churchill & Doerge, 1994; Doerge & Churchill, 1996).

For one, two and three QTL conditions, an additional means of summarizing and visualizing mapping performance for the 500 replicates was used. This method views the chromosome as a histogram with bin widths spanning 5 cM. The height of a bin in these histograms is equal to the proportion of the time out of the 500 replicates that the maximum LR is present at the location spanned by that bin (cf Liu and Dekkers, 1998). This method was not informative for conditions with more than three QTLs owing to the relationship between QTL spacing and the width of the bins. When appropriate, however, it is particularly useful for visualizing the precision of the estimates of QTL location.

Two methods of numerically summarizing the results were also used. The 10th, 25th, 50th, 75th and 90th percentile proportions of the chromosome that were reported as significant for each condition were calculated. For a given map, this 'significance proportion' is simply the number of times the LR is above the significance threshold divided by the number of positions at which the LR was calculated. This method was found to be a useful supplementary means of quantifying likelihood profile shape and results are provided in the web supplement. The power of QTL detection (Van Ooijen, 1992) was also calculated for all conditions.

Some discussion and clarification of terms is necessary. In the presentation of results below, the resolving ability of a map is called 'high' if the peaks corresponding to putative QTLs can reasonably be called narrow and if the map can distinguish multiple nearby QTLs. Resolving ability is distinct from 'localization ability', which is the average difference between the true location of the QTL and the location of the maximum LR in a given peak (Fig. 1*a*, *b*). The concept of a peak itself is subjective; our working definition of a peak is a region of the likelihood profile that is a local maximum and that can reasonably be interpreted to suggest that factors in the region affect the trait in question (Fig. 1*c*). The characterization of resolving ability in the presentation of results below takes into account the shape of peaks even if they are below the significance threshold. Such an analysis, together with the power data, gives a more complete picture of the properties of maps than would be given by an analysis that considered only maps, whose interpretations would certainly result in a conclusion that a QTL is present (Fig. 1*d*).

3. Results

Owing to space constraints, only a representative sample of the data is presented here. Fractile curves for all conditions, maximum LR histograms for all one, two, and three QTL conditions, and complete significance proportion and power data for all conditions are available at http://hjmuller.bio.uci.edu/~labhome/work-im1.html. This web supplement is a

https://doi.org/10.1017/S0016672303006396 Published online by Cambridge University Press

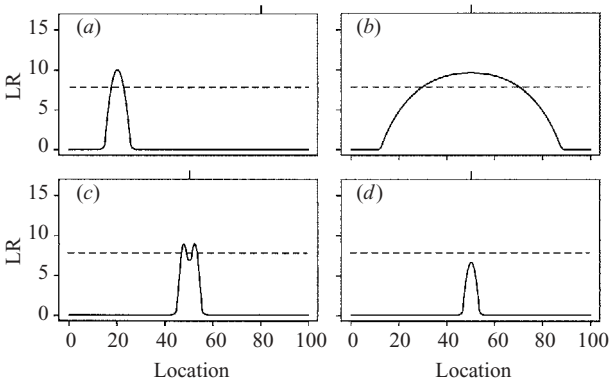

Fig. 1. Idealized depictions of likelihood profile properties. QTL locations are indicated by a hash mark at the top border. The height of the mark is a measure of the relative QTL effect. Negative effects are indicated by marks extending below the top border. The likelihood map in (a) has good resolving ability but poor localization ability; one would be strongly led to believe that there is a QTL at ~20 cM when, in reality, it is at 80 cM. By contrast, the likelihood map in (b) has poor resolving ability and good localization ability, because the peak correctly identifies the location of the QTL as 50 cM. However, one would not be confident that the QTL is near 50 cM. In the commonly seen situation depicted in (c), a standard definition of a peak (a local LR score maxima separated by LR values below the significance threshold) would lead to the conclusion that there are two closely linked QTLs. However, a reasonable interpretation would be that there is a single QTL near 50 cM. Similarly, although the maximum in (d) is slightly below the significance level, the resolving ability of the map is clearly excellent and there is suggestive evidence for a single QTL near 50 cM.

large archival file that may be appropriate for experiments with marker densities, sample sizes and hypothesized genetic architectures matching one of the simulated conditions.

(i) *IM vs CIM for a small number of QTLs*

IM and CIM maps for conditions with three or fewer QTLs of uniform effect both display good localization ability. For a typical individual map, a peak's maximum LR tends to be within 10 cM of the true QTL location across sample size and marker density conditions (Fig. 2*a*, *b*, web supplement pp. 4–27 (max. LR histograms)). Both IM and CIM localization ability increase somewhat with sample size and marker density. In addition, a lack of informative recombination events in sample sizes of less than 100 used to detect three QTL gives poor localization ability regardless of mapping method used.

Unlike localization ability, there are significant differences in resolving ability between IM and CIM for uniform effects maps. This is especially noticeable when two or more QTLs are present, because multiple peaks must be distinguished for an accurate map interpretation because the maximum LR across the chromosome does not suffice to indicate QTL location. The very broad patterns in the IM fractile curves reflect the fact that, for individual IM maps with two or more QTLs, peaks are often uninformative owing to the large portion of the chromosome they encompass (Fig. 3*a*; web supplement pp. 49–60 (fractile curves) and pp. 94–477 (sample plots)). In fact, most IM maps show a single wide peak. The CIM peaks, by contrast, are fairly narrow on average (<20 cM) (Fig. 3*b*; web supplement pp. 38–43 (significance proportion data), pp. 69–80 (fractile curves) and pp. 734–1117 (sample plots)) and CIM maps with one to three QTLs usually show clearly discernible peaks in the appropriate regions, although, more often than not, the correct number are not present (there are often more and sometimes less), especially when QTLs are less than 30 cM apart or when a sample size of 100 is used and three QTLs are present. For both IM and CIM, resolving ability is unaffected by sample size but increases with increasing marker density.

CIM LR values tend to increase as marker density decreases and as sample size increases, whereas IM LR values increase with sample size but are unrelated to marker density for uniform effects maps. This difference is probably due to the higher number of CIM cofactors being used at higher marker densities in our study, which reduces the power of CIM. The power of QTL detection was observed to be considerably higher for all IM maps with three or fewer QTLs than for corresponding CIM maps (Table 2; web supplement pp. 28–33 and pp. 38–43 (power data)).

The properties of CIM maps with two QTLs of effects −1·0 and +1·0 are similar to those with QTLs of uniform effects, except with somewhat better localization ability and substantially better resolving ability. Two clear, narrow peaks in the correct location are generally seen (Figs 2*d*, 3*d*; web supplement pp. 24–26 (max LR histograms), pp. 42–43 (significance proportion data), pp. 77–79 (fractile curves) and pp. 990–1085 (sample plots)). Localization ability for IM maps with QTLs of effects −1·0 and +1·0 is similar to localization ability for corresponding CIM maps, and resolving ability is surprisingly good given IM trends observed elsewhere (Fig. 3*c*, *d*). For example, when the QTLs are 60 cM apart and have effects of −1·0 and +1·0 in IM maps, there are generally two peaks that localize the QTL to within 5–10 cM, are roughly 30–40 cM wide and have similar shapes across sample size/marker density conditions. As the QTL distance decreases to 20 cM and even 5 cM apart, the maps have properties similar to the maps with QTLs 60 cM apart, except that the inside portions of the peaks tend to be truncated; that is, there tends to be a deep 'valley' that often leads to

https://doi.org/10.1017/S0016672303006396 Published online by Cambridge University Press

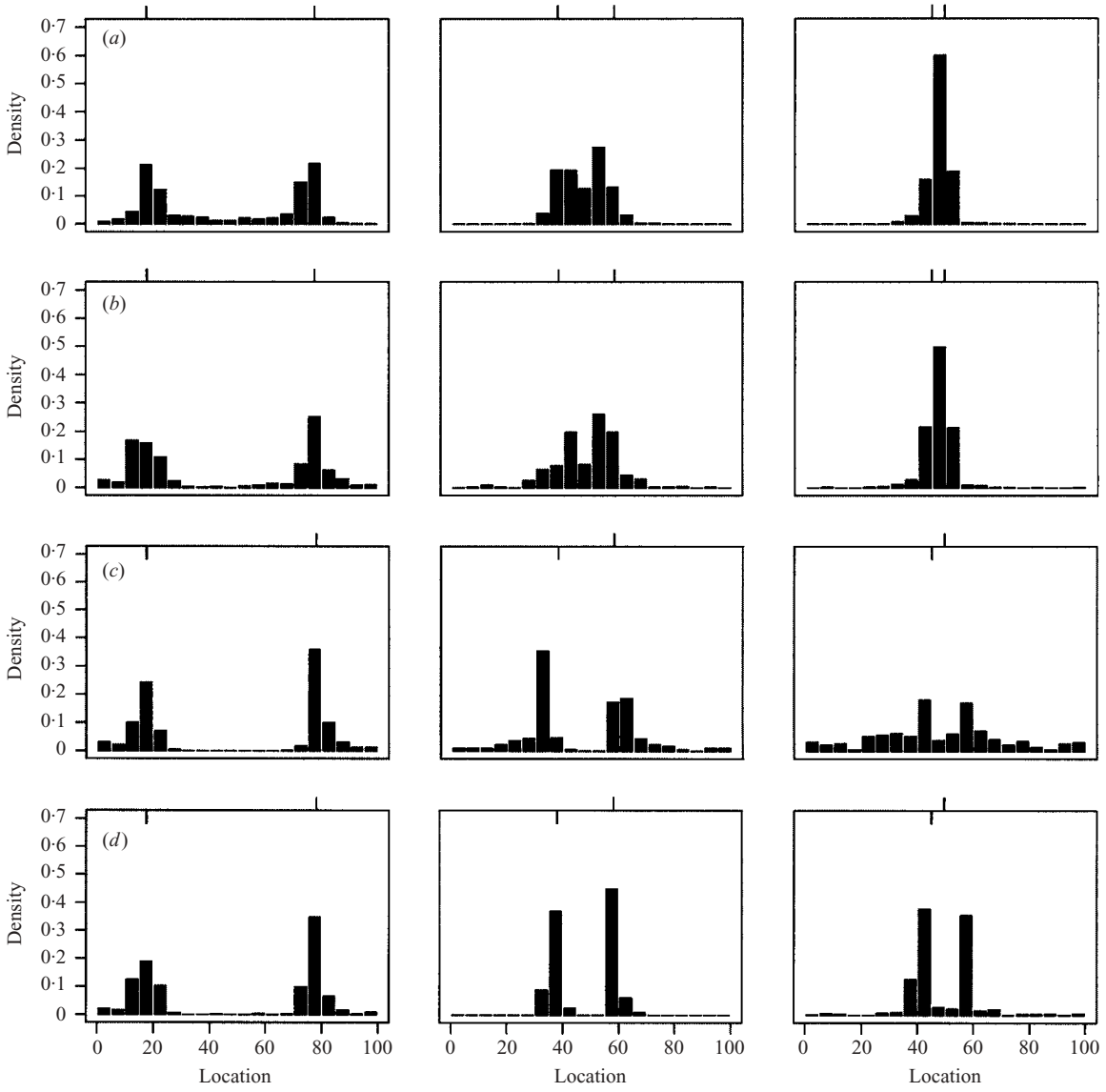

Fig. 2. Maximum LR histograms for selected parameter sets. The histograms are as follows. QTL locations are indicated by a hash mark at the top border. The height of the mark is a measure of the relative QTL effect. Negative effects are indicated by marks extending below the top border. (a) IM with a marker density of 7 cM, a sample size of 500 and two QTLs with the same relative effects and with locations of (left to right) 18 cM and 78 cM, 38 cM and 58 cM, and 45 cM and 50 cM. (b) CIM with a marker density of 7 cM, a sample size of 500 and two QTLs with the same relative effects and with locations of 18 cM and 78 cM, 38 cM and 58 cM, and 45 cM and 50 cM. (c) IM with a marker density of 7 cM, a sample size of 500 and two QTLs with effects of -1 and $+1$, and with locations of 18 cM and 78 cM, 38 cM and 58 cM, and 45 cM and 50 cM. (d) CIM with a marker density of 7 cM, a sample size of 500 and two QTLs with effects of -1 and $+1$, and locations of 18 cM and 78 cM, 38 cM and 58 cM, and 45 cM and 50 cM.

an interpretation of two peaks in the appropriate regions. Despite this excellent overall resolving ability, roughly one-quarter to one-third of the time, one or the other of the QTLs is virtually undetected (Fig. 2*c*, Fig. 3*c*; web supplement pp. 12–14 (max LR histograms), pp. 32–33 (significance proportion data), pp. 57–59 (fractile curves) and pp. 350–445 (sample plots)).

The resolving ability and localization ability of IM and CIM maps with QTLs of effects 0·5 and 2·0 are essentially the same as those of corresponding maps with a single QTL, because the presence of the QTL with effect 0·5 is almost never suggested by the maps. This is perhaps not surprising given that the QTL with a relative effect of 0·5 accounts for only 1/17th of the genotypic variance.

All IM and CIM nonuniform QTL effect maps display the same LR increase and power trends as corresponding uniform QTL effect maps. Similarly, IM power is higher than CIM when two QTL of effects 0·5 and 2·0 are modeled. By contrast, IM power is generally lower than CIM power for traits controlled by two QTLs of opposite effect (Table 2; web supplement pp. 30–33 and pp. 40–43 (power data)).

https://doi.org/10.1017/S0016672303006396 Published online by Cambridge University Press

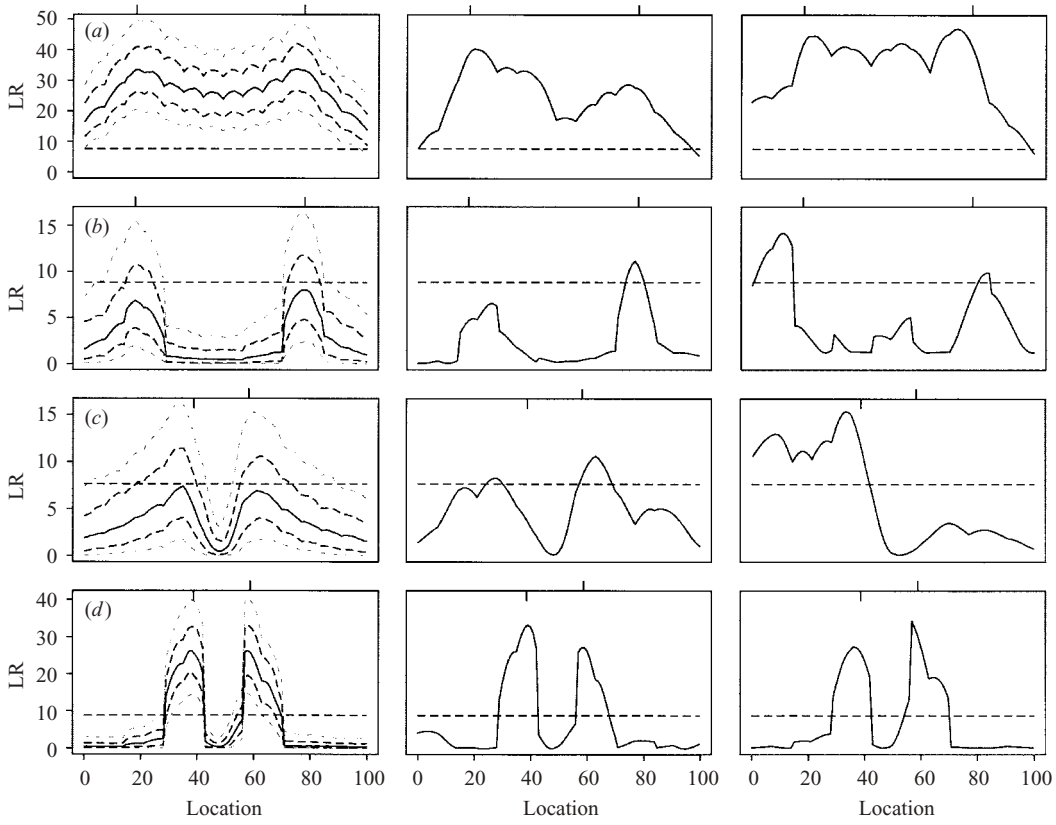

Fig. 3. Fractile curves and example individual likelihood maps for selected parameter sets with a small number of QTLs. QTL locations are indicated by a hash mark at the top border. The height of the mark is a measure of the relative QTL effect. Negative effects are indicated by marks extending below the top border. The first column contains the fractile curves and the other two plots in a row are representative example individual likelihood maps. (a) IM with a marker density of 7 cM, a sample size of 500 and two QTLs with the same relative effects and locations of 18 cM and 78 cM. (b) CIM with a marker density of 7 cM, a sample size of 500 and two QTLs with the same relative effects and locations of 18 cM and 78 cM. (c) IM with a marker density of 7 cM, a sample size of 500 and two QTLs with effects of -1 and $+1$ at 38 cM and 58 cM. (d) CIM with a marker density of 7 cM, a sample size of 500 and two QTLs with effects of -1 and $+1$ at 38 cM and 58 cM. Notice that, although IM and CIM can perform well on average, individual realizations of the simulations can be misleading with regards to the underlying genetic architecture. For example, in (b) column two, the left-hand QTL would probably not be identified, just as the right-hand QTL would not in (c) column three. In (b) column three, the left-hand QTL would be identified, although the likelihood at the true location of the QTL is significantly lower than at the peak.

(ii) *CIM and a several-QTL model*

The simulated traits controlled by 20 and 40 QTLs are meant to approximate a polygenic trait rather than a trait with relatively few factors of large effect. At its extreme of an infinite number of genetic factors (Bulmer, 1980), many would consider the common observation that a small number of QTLs are mapped for many traits as being inconsistent with a polygenic model. IM 20 and 40 QTL maps are, almost without exception, large domelike shapes with inflections in the likelihood profile owing only to the notches at marker locations with power always 1·0. Thus, results are here presented only for CIM 20 and 40 QTL maps. Furthermore, CIM 20 and 40 QTL maps have very similar properties and so statements made about CIM 20 QTL maps should be taken to hold for CIM 40 QTL maps unless stated otherwise.

Maps with uniform spacing of the QTLs and uniform effects tend to display the whole-chromosome encompassing profiles of the IM maps across sample size and marker density conditions (Fig. 4*a*; web supplement pp. 81, 85 (fractile curves), pp. 1118–1149 and pp. 1246–1277 (sample plots), and pp. 44, 46 (significance proportion data)). When a marker density of 1 cM was used for these maps, a regular pattern of peaks approximately every 10–15 cM was observed that do not seem to correspond to QTL locations.

For CIM conditions with uniform QTL spacing but random effects, resolving ability and localization ability are similar to the uniform spacing and uniform effects conditions, except that the general pattern is the production of broad peaks or entire regions with relatively high LR that correspond roughly to the places where there are clusters of QTL with high effects, which are on the far left side and the far right side of the chromosome (Fig. 4*b*; web supplement

https://doi.org/10.1017/S0016672303006396 Published online by Cambridge University Press

Table 2. *Power data for selected parameter sets*

QTL location and effect		18/+1, 78/+1		38/–1, 58/+1		CIM (20 QTLs of equal location and effect)
Marker density	Sample size	IM	CIM	IM	CIM	
1	100	0·75	0·11	0·21	0·37	0·38
	200	0·97	0·25	0·47	0·73	0·79
	500	1·00	0·64	0·94	1·00	1·00
	1000	1·00	0·98	1·00	1·00	1·00
7	100	0·77	0·20	0·20	0·47	0·75
	200	0·98	0·35	0·44	0·81	1·00
	500	1·00	0·82	0·89	1·00	1·00
	1000	1·00	1·00	1·00	1·00	1·00
Random	100	0·74	0·14	0·18	0·33	0·54
	200	0·97	0·19	0·46	0·57	0·86
	500	1·00	0·62	0·90	0·98	1·00
	1000	1·00	0·91	1·00	1·00	1·00
20	100	0·74	0·26	0·17	0·53	0·95
	200	0·97	0·56	0·35	0·86	1·00
	500	1·00	0·97	0·89	1·00	1·00
	1000	1·00	1·00	1·00	1·00	1·00

pp. 82, 86 (fractile curves), pp. 1150–1181 and pp. 1278–1309 (sample plots), and pp. 44, 46 (significance proportion data)).

CIM 20 QTL random location, uniform effects maps have basically the same properties as uniform location, random effects maps except that they show the opposite pattern, with the high region located in the center of the chromosome instead of at the edges (Fig. 4*c*; web supplement pp. 83, 87 (fractile curves), pp. 1182–1213 and pp. 1310–1341 (sample plots), and pp. 45, 47 (significance proportion data)). Because the QTLs all have the same effect, the fact that QTL locations are concentrated towards the center (there is only one QTL to the left of 22·8 cM and none to the right of 86·7) accounts for this observation.

CIM 20 QTL random location, random effects maps have profile shapes determined primarily by the QTLs with high effects on the edges. As is the case with the other 20 and 40 QTL maps, the interpretation of any given map would probably bear little resemblance to the actual genetic architecture; a likely interpretation might take the form of 'one or more factors affecting the trait in the 10–30 cM region and 70–90 cM region', which would probably correspond to the QTL at 22·8 cM, 23·4 cM and 24·2 cM with effects 2·1, 3·9 and 4·9 and to the QTLs at 77·8 cM, 78·0 cM and 82·9 cM with effects of 4·6, 3·3 and 3·0.

A similar resolving ability across marker densities is observed for all CIM 20 and 40 QTL maps, as well as the above-mentioned 10–15 cM valley pattern observed in maps with 1 cM density, is illustrated by three fractile curves and six individual maps from the CIM 20 QTL with random spacing and random effects conditions in Fig. 5*d–f* (web supplement pp. 84, 88 (fractile curves), 1214–1245 and 1342–1373 (sample plots), and 45, 47 (significance proportion data)). Power was higher for all CIM 20 and 40 QTL maps than for CIM one, two and three QTL maps, and was never less than 0·38 (Table 2; web supplement pp. 45, 47 (power data)).

4. Discussion

(i) *IM vs CIM for a small number of QTLs*

For practical purposes, IM should be used primarily as a tool for identifying whole chromosomes or linkage groups with effects on a trait. This is because IM typically does not reliably resolve multiple QTLs but can almost always identify the location of groups of QTL to within 40 cM or less. Even if the trait in question is not known to be controlled by one or fewer QTL per 100 cM, the location estimate provided by IM can be useful preliminary data for higher resolution analysis of a region or for purposes other than positional cloning, such as marker-assisted selection or candidate gene identification. Our results demonstrate the greater power that IM generally has compared with CIM, which is related to the use of conditioning markers (Zeng, 1993).

Nevertheless, CIM can frequently be used to obtain accurate information about QTL number and location for traits that are controlled by three or fewer QTLs per 100 cM, even when trait heritabilities are fairly low, assuming that QTL effects are additive and a reasonable sample size (at least 200) is used. Within these general trends for CIM, significant benefits of increased marker density were observed for QTL resolution and, to a lesser extent, localization. The shape of an LR profile is almost entirely independent

https://doi.org/10.1017/S0016672303006396 Published online by Cambridge University Press

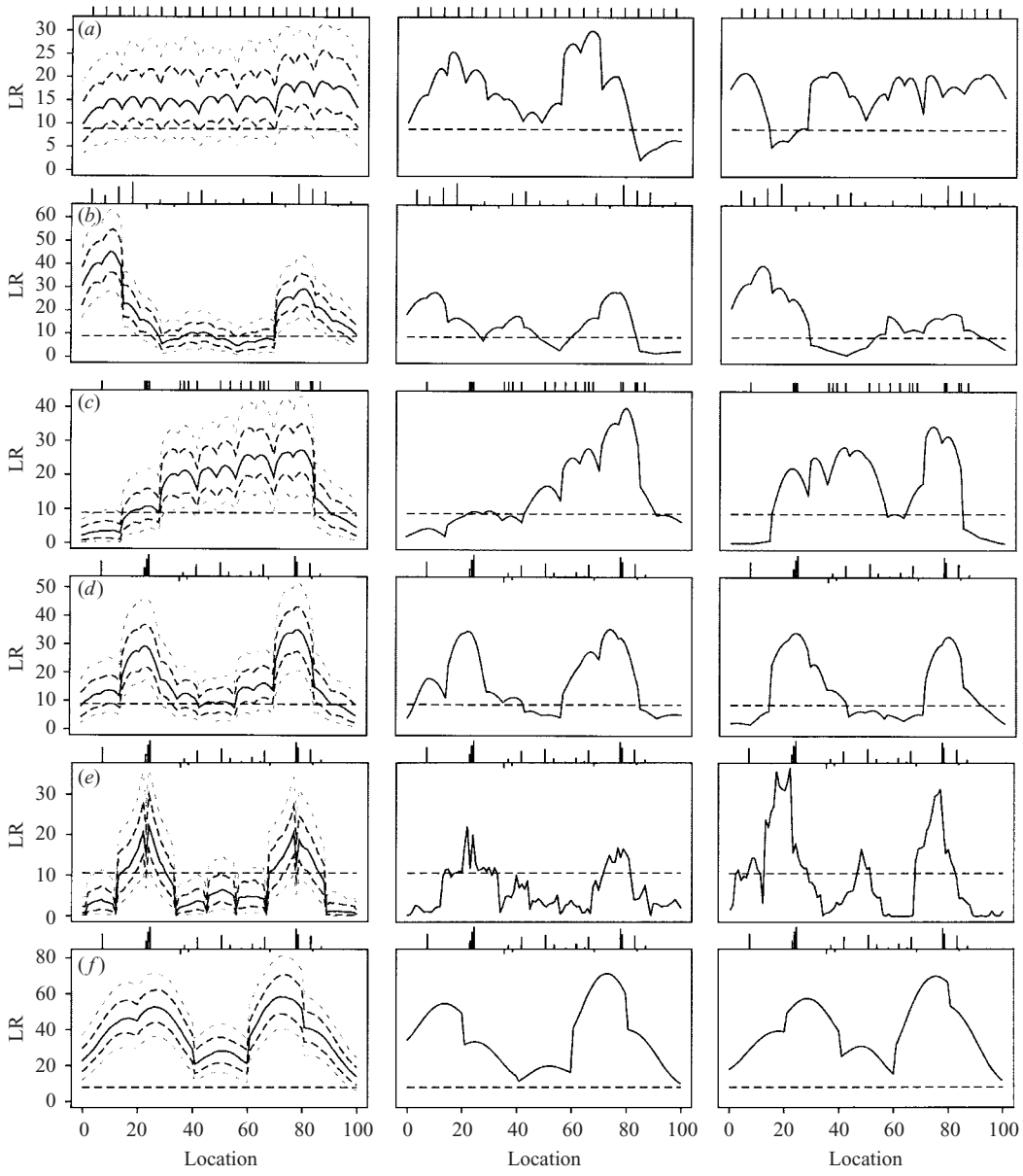

Fig. 4. Fractile curves and example individual likelihood maps for selected parameter sets with 20 QTLs. QTL locations are indicated by a hash mark at the top border. The height of the mark is a measure of the relative QTL effect. Negative effects are indicated by marks extending below the top border. (a) CIM with a marker density of 7 cM, a sample size of 500 and equally spaced QTLs of equal effects. (b) CIM with a marker density of 7 cM, a sample size of 500 and equally spaced QTLs of random effects. (c) CIM with a marker density of 7 cM, a sample size of 500 and randomly spaced QTLs of equal effects. (d) CIM with a marker density of 7 cM, a sample size of 500 and randomly spaced QTLs of random effects. (e) CIM with a marker density of 1 cM, a sample size of 500 and randomly spaced QTLs of random effects. (f) CIM with a marker density of 20 cM, a sample size of 500 and randomly spaced QTLs of random effects.

of sample size, although the magnitude of the LR values increases with sample size.

Other studies investigating the properties of IM and CIM for a small number of QTL are consistent with the results presented here. For example, Van Ooijen's conclusion (Van Ooijen, 1992) that a single QTL explaining 5% of the phenotypic variance can be mapped only to roughly the nearest 40 cM is based on stricter criteria for determining confidence in location estimates than in the present study. However, as the maximum LR histograms attest, if IM is to be used for rough mapping only, localization ability is good by any standard. Wright & Kong (1997) found IM localization ability to be good if one QTL is present, unless marker density is low. In general, for three or fewer QTLs, our results show good IM and CIM localization ability across all marker density conditions. For these three or fewer QTL conditions, however, a decline in localization ability related to low marker density (an appreciable bias of location estimates toward the center when the marker density was 20 cM) was observed, as predicted by Wright & Kong (1997).

https://doi.org/10.1017/S0016672303006396 Published online by Cambridge University Press

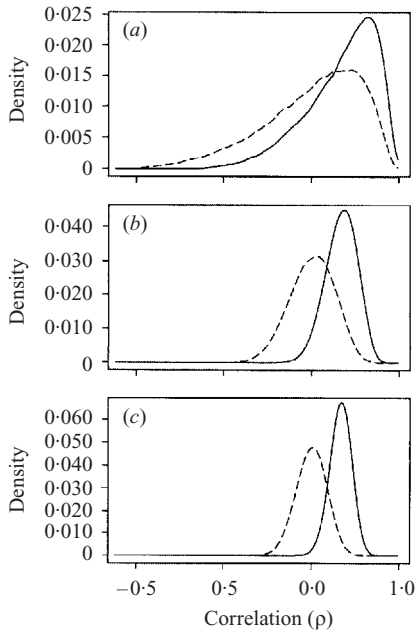

Fig. 5. Distribution of location-wise correlations between simulated maps for traits with simple and complex genetic architectures. Solid lines are the distribution of the correlation coefficient (ρ) or average ρ for LR values at each location obtained from 500 simulated CIM maps with a marker density of 7 cM and a sample size of 500 for a trait with two QTLs at 25 cM and 91 cM, of relative effects 12·4 and 13·8, and a trait heritability of 0·24. The broken line is the distribution of average ρ obtained from 500 simulated CIM maps with a marker density of 7 cM and a sample size of 500 for the 20 QTLs, random locations and effects condition. In all cases, the distribution is based on a random sampling of the 500 simulations to obtain 100,000 estimated ρ values. The histogram's bin widths are 0·01 and it was smoothed using the Nadaraya–Watson kernel regression method with a bandwidth of 0·05 (*ksmooth* function in R). (A) The distribution of ρ when a QTL mapping experiment is replicated once. (B) The distribution of the average ρ based on six replicate experiments (i.e. $(6 \times 5)/2 = 15$ estimates of ρ). (C) The distribution of the average ρ based on 12 replicate experiments (i.e. $(12 \times 11)/2 = 66$ estimates of ρ).

For a small number of QTLs, Darvasi *et al.* (1993) found that IM power is not substantially influenced by marker spacing if QTLs are not present at marker locations. This is consistent with our results, because we made an effort to ensure that QTLs were not present at the markers.

Unlike with IM, CIM marker density affects power. One possible explanation for this is that, with CIM, marker density is directly related to the number of cofactors available for selection: with 20 cM density, there are at most six cofactors, whereas the other densities have nine and, as previously mentioned, the number of cofactors used affects CIM power. Darvasi *et al.* (1993) also report that, as sample size and trait heritability decrease, the effect of marker density on localization ability decreases. For example, marker density is unlikely to have a significant effect on localization ability for sample sizes of 100 or 200, but is likely to at higher sample sizes, which we observe to some extent for both IM and CIM. Zeng (1994) examined both IM and CIM with a heritability of the simulated trait of 0·70 and observed localization ability to be good for both IM and CIM (usually within the 5–15 cM range), although, in most circumstances CIM had an advantage over IM. We observe similar trends for the condition of a small number of QTLs that Zeng simulated (Fig. 2).

(ii) *CIM and a several-QTL model*

Many of the simulated maps observed when the underlying genetic architecture consists of a large number of QTLs at random locations resemble the likelihood profiles generated in real-life mapping studies (cf Long *et al.*, 1995; Macdonald & Goldstein, 1999). Thus, it is possible that the true genetic architectures of many traits could resemble those present in the 20 or 40 QTL random location, random effects conditions. Of course, mapping studies do not claim to identify all QTLs and their locations, only large segregating genetic factors that affect the trait. The results of the present study underscore the fact that these identified factors might be regions a few centi-Morgans in size harboring multiple closely linked QTLs of high effect relative to other regions in the genome. When more than a few QTLs are mapped on a chromosome, caution should be exercised in equating mapped factors or peaks to single QTLs/genes. As demonstrated in Fig. 4*d*, even a map with two well defined peaks is not inconsistent with the presence of 20 QTLs. This study thus confirms the important results of Visscher & Haley (1996) and Liu & Dekkers (1998): it is difficult to distinguish between a single QTL of large effect and many QTLs (with small and/or large effects). However, these previous studies used a genetic model consisting of 101 QTLs of equal effect and equal spacing, and it is valuable to ask whether their conclusions are robust to relaxing the assumption of equal effects and spacing. Here, we extend these earlier results to a fairly plausible genetic architecture in which a trait is determined by more than a handful of factors whose effects are not equally spaced at random intervals. The inability to distinguish between a single QTL of large effect and a several-QTL trait holds to a substantial extent under a variety of our more realistic models. The conjecture of Visscher & Haley (1996) that greater sample size and/or marker density would eliminate this problem is not born out in our study.

(iii) *Discussion of methods*

The characteristics of the mapping conditions in the present study were varied in such a way that the

importance of factors likely to affect the performance of IM and CIM could be easily evaluated. In particular, the effect of the genetic architecture of the theoretical quantitative trait was presumed to be of utmost importance. The choices of simulated marker densities were intended to reflect realistic lower (20 cM) and upper (1 cM) boundaries for marker densities used in modern studies. The use of regularly spaced markers for most of the analysis in this and other studies (cf Charmet, 2000) is likely to be a valid simplification, because we found that differences between the effects of irregular and regular marker spacing on QTL mapping properties to be minimal for realistic sample sizes and heritabilities.

Various methods have been proposed and used to quantify the power to detect QTLs and the precision of QTL mapping procedures. Many of these focus on the properties of expected LR curves (Lander & Botstein, 1989; Wright & Kong, 1997; Ronin *et al*., 1999). Although expected LR curves are useful, they can be misleading, because any single QTL mapping experiment is but one realization of the random process that gives rise to the expected LR curve. Thus, a mapping method might perform well on average yet still give potentially misleading results for any given realization (e.g. the sample plots in Figs 3*a*, 4*d*). A method used in some studies that is perhaps the most direct means of quantifying mapping precision is the calculation of the mean location estimate (or deviation from location estimate) based on the maximum test statistic value in a relevant region of the likelihood profile over multiple simulated mapping populations (Carbonell *et al*., 1993; Zeng, 1994; Charmet, 2000). However, the occasional lack of discernible relevant peaks across simulations makes such an analysis difficult for multiple QTLs (Zeng, 1994), especially if it is to be automated and if maps with more than two QTLs per 100 cM are being investigated.

Although methods for quantifying the accuracy of location estimates are conceptually and computationally straightforward, they can be removed from the type of inference commonly drawn from a mapping study. For this reason, we attempted to summarize the nature of the inferences that would be made by looking at individual likelihood maps, which largely determine the direction of future experiments and the conclusions drawn from them. We found developing a novel summary statistic that would summarize the quality of inference drawn from such maps to be more problematic than expected, because experimental geneticists take into account a range of difficult-to-quantify factors when interpreting a likelihood profile. These include the shapes and heights of peaks as they relate to other nearby peaks and to the overall shape of the likelihood profile in a nearby region, and to the overall shape across the entire chromosome, as well as intuition about the general characteristics of rational likelihood map interpretations, and possibly even *a priori* knowledge about the underlying genetic architecture of the trait such as knowledge of candidate gene locations. Thus, the development of a useful definition of a single, numerical 'score' for the informativeness of QTL likelihood map data (or similar data that is interpreted in a somewhat subjective fashion) for now remains unworkable.

One potential approach to such a problem takes advantage of unused 'information' in the likelihood profiles. We simulated replicate QTL mapping experiments under a complex genetic architecture of 20 QTLs with random locations and random effects, and under a simple genetic architecture of two QTLs that might reasonably be inferred from more complex architecture. In Fig. 5, we show the distribution of the average correlation coefficient between likelihood profiles that would be observed if the QTL mapping experiment were replicated two, six or 12 times. It can be seen that replicating a QTL mapping experiment many times results in likelihood profiles that are more correlated over experiments under a two-QTL model than under a several-QTL model. We are not suggesting that QTL mapping experiments be replicated to this level but that there is information that is not presently being used that could perhaps be extracted under some sort of data re-sampling scheme.

(iv) *Conclusion and future directions*

We have shown that it would frequently be difficult to reach the conclusions of studies such as Long *et al*. (1995) regarding the numbers, locations and effects of third chromosome bristle QTLs given the experiment carried out (with respect to sample size and marker density). Our goal in so doing is not to criticize or reinterpret the conclusions of such work, but rather to demonstrate that the lack of correspondence between the architecture described in a study and the likelihood of detecting such an architecture given the experimental design used could be turned into an advantage. In future studies, it might be of value to simulate QTL mapping experiments under the observed genetic architecture (and marker densities and sample sizes used) and to ask how likely the observed results are under the 'uncovered' architecture. This is possible within the framework of currently used software (cf Basten *et al*., 1999) but is rarely carried out as a test of consistency of the reported architecture with the study design that detected that architecture.

In only a few instances has the QTL mapping community identified the actual causative variants that were originally identified with IM or CIM (Long *et al*., 1998; Frary *et al*., 2000; Long *et al*., 2000; Pasyukova *et al*., 2000; Thornsberry *et al*., 2001). Despite this, and the fact that the statistical models

underlying the methods are based on the assumption that one or a small number of QTLs per linkage group affect the trait, the information regarding the number of QTLs and their locations and effects that these genome-wide mapping methods provide is often held to be accurate, even when many QTLs are identified (Wright & Kong, 1997). Our results provide some justification for this attitude but also highlight reservations. Inasmuch as it is a goal in modern quantitative genetics to distinguish between underlying inheritances of a 'few genes of large effect' versus models that are more polygenic in nature (cf Orr & Coyne, 1992), it will be of use to develop hypothesis testing frameworks that explicitly test the support for a polygenic versus small number of genes model.

We thank M. Drapeau and S. Macdonald for discussions and sharing unpublished results, and members of the Long laboratory for suggestions and encouragement. We also thank two anonymous reviewers for comments that improved the manuscript. This project was supported by National Institutes of Health grant HG-02107.

References

Basten, C. J., Weir, B. S. & Zeng, Z.-B. (1999). *QTL Cartographer: A Reference Manual and Tutorial for QTL Mapping*. North Carolina State University, Raleigh, North Carolina (http://statgen.ncsu.edu/).

Bulmer, M. G. (1980). *The Mathematical Theory of Quantitative Genetics*. Oxford: Clarendon Press.

Carbonell, E. A., Asins, M. J., Baselga, M., Balansard, E. & Gerig, T. M. (1993). Power studies in the estimation of genetic parameters and the localization of quantitative trait loci for backcross and doubled haploid populations. *Theoretical and Applied Genetics* **86**, 411–416.

Charmet, G. (2000). Power and accuracy of QTL detection: simulation studies of one-QTL models. *Agronomie* **20**, 309–323.

Churchill, G. A. & Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963–971.

Darvasi, A., Weinreb, A., Minke, V., Weller, J. I. & Soller, M. (1993). Detecting marker–QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics* **134**, 943–951.

Doerge, R. W. & Churchill, G. A. (1996). Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**, 285–294.

Frary, A., Nesbitt, T. C., Frary, A., Grandillo, S., van der Knaap, E., Cong, B., Liu, J. P., Meller, J., Elber, R., Alpert, K. B. & Tanksley, S. D. (2000). *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science* **289**, 85–88.

Goffinet, B. & Mangin, B. (1998). Comparing methods to detect more than one QTL on a chromosome. *Theoretical and Applied Genetics* **96**, 628–633.

Kao, C. H., Zeng, Z.-B. & Teasdale, R. D. (1999). Multiple interval mapping for quantitative trait loci. *Genetics* **152**, 1203–1216.

Lander, E. S. & Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185–199.

Liu, Z., Dekkers, J. C. M. (1998). Least squares interval mapping of quantitative trait loci under the infinitesimal genetic model in outbred populations. *Genetics* **148**, 495–505.

Long, A. D., Mullaney, S. L., Reid, L. A., Fry, J. D., Langley, C. H. & Mackay, T. F. C. (1995). High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. *Genetics* **139**, 1273–1291.

Long, A. D., Lyman, R. F., Langley, C. H. & Mackay, T. F. C. (1998). Two sites in the *Delta* gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* **149**, 999–1017.

Long, A. D., Lyman, R. F., Morgan, A. H., Langley, C. H. & Mackay, T. F. C. (2000). Both naturally occurring insertions of transposable elements and intermediate frequency polymorphisms at the achaete-scute complex are associated with variation in bristle number in *Drosophila melanogaster*. *Genetics* **154**, 1255–1269.

Lynch, M. & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, Massachusetts: Sinauer Associates.

Macdonald, S. J. & Goldstein, D. B. (1999). A quantitative genetic analysis of male sexual traits distinguishing the sibling species *Drosophila simulans* and *D. sechellia*. *Genetics* **153**, 1683–1699.

Mackay, T. F. C. (2001). Quantitative trait loci in *Drosophila*. *Nature Reviews Genetics* **2**, 11–20.

Orr, H. A. & Coyne, J. A. (1992). The genetics of adaptation – a reassessment. *American Naturalist* **140**, 725–742.

Pasyukova, E. G., Vieira, C. & Mackay, T. F. C. (2000). Deficiency mapping of quantitative trait loci affecting longevity in *Drosophila melanogaster*. *Genetics* **156**, 1129–1146.

Ronin, Y., Korol, A. B. & Nevo, E. (1999). Single- and multiple-trait mapping analysis of linked quantitative trait loci: some asymptotic analytical approximations. *Genetics* **151**, 387–396.

Sillanpää, M. J. & Arjas, E. (1998). Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. *Genetics* **148**, 1373–1388.

Thoday, J. M. (1961). Location of polygenes. *Nature* **191**, 368–370.

Thornsberry, J. M., Goodman, M. M., Doebley, J., Kresovich, S., Nielsen, D. & Buckler, E. S. (2001). *Dwarf8* polymorphisms associate with variation in flowering time. *Nature Genetics* **28**, 286–289.

Van Ooijen, J. W. (1992). Accuracy of mapping quantitative trait loci in autogamous species. *Theoretical and Applied Genetics* **84**, 803–811.

Visscher, P., Whittaker, J. & Jansen, R. (2000). Mapping multiple QTL of different effects: comparison of a simple sequential testing strategy and multiple QTL mapping. *Molecular Breeding* **6**, 11–24.

Visscher, P. M. & Haley, C. S. (1996). Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models. *Theoretical and Applied Genetics* **93**, 691–702.

Walling, G. A., Haley, C. S., Perez-Enciso, M., Thompson, R. & Visscher, P. M. (2002). On the mapping of quantitative trait loci at marker and non-marker locations. *Genetical Research* **79**, 97–106.

Wright, F. W. & Kong, A. (1997). Linkage mapping in experimental crosses: the robustness of single-gene models. *Genetics* **146**, 417–425.

Zeng, Z.-B. (1993). Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proceedings of the National Academy of Science USA* **90**, 10972–10976.

Zeng, Z.-B. (1994). Precision mapping of quantitative trait loci. *Genetics* **136**, 1457–1468.