ORIGINAL ARTICLE

www.nature.com/hdy

Controlling false positives in the mapping of epistatic QTL

W-H Wei^{1,2}, S Knott³, CS Haley^{1,2} and D-J de Koning¹

¹Division of Genetics and Genomics, The Roslin Institute and R(D)SVS, University of Edinburgh, Roslin, Midlothian, UK; ²MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, UK and ³Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

This study addresses the poorly explored issue of the control of false positive rate (FPR) in the mapping of pair-wise epistatic quantitative trait loci (QTL). A nested test framework was developed to (1) allow pre-identified QTL to be used directly to detect epistasis in one-dimensional genome scans, (2) to detect novel epistatic QTL pairs in twodimensional genome scans and (3) to derive genome-wide thresholds through permutation and handle multiple testing. We used large-scale simulations to evaluate the performance of both the one- and two-dimensional approaches in mapping different forms and levels of epistasis and to generate profiles of FPR, power and accuracy to inform epistasis mapping studies. We showed that the nested test framework and genome-wide thresholds were essential to control FPR at the 5% level. The one-dimensional approach was generally more powerful than the two-dimensional approach in detecting QTL-associated epistasis and identified nearly all epistatic pairs detected from the two-dimensional approach. However, only the two-dimensional approach could detect epistatic QTL with weak main effects. Combining the two approaches allowed effective mapping of different forms of epistasis, whereas using the nested test framework kept the FPR under control. This approach provides a good search engine for high-throughput epistasis analyses. *Heredity* (2010) **104,** 401–409; doi:10.1038/hdy.2009.129;

published online 30 September 2009

Keywords: epistasis; false positive; multiple testing; nested test framework; power; quantitative trait loci

Introduction

A better knowledge of epistasis would greatly contribute to our understanding of complex trait variation and the evolutionary dynamics of biological systems (Carlborg and Haley, 2004; de Visser and Elena, 2007; Phillips, 2008). With the advances of molecular genetics and genomics, efforts have been made to develop methods and tools for detecting statistical epistasis in different population structures (Malmberg and Mauricio, 2005; Xu and Jia, 2007; Zhang and Liu, 2007; Aylor and Zeng, 2008; Pattin et al., 2008). Research has explored a wide range of topics including statistical modelling (Alvarez-Castro and Carlborg, 2007; Moore et al., 2007; Sung and Wijsman, 2007), model parameterization (Zeng et al., 2005; Wang and Zeng, 2006; Alvarez-Castro et al., 2008), search algorithm (Carlborg et al., 2005; Ritchie and Motsinger, 2005; Kooperberg and Leblanc, 2008; Mechanic et al., 2008), multiple testing (Jannink and Jansen, 2001; Sen and Churchill, 2001; Storey et al., 2005; Stich et al., 2007) and computing efficiency (Ljungberg et al., 2004; Bush et al., 2006). A number of relevant software tools have also been made publicly available (Broman et al., 2003; Hahn et al., 2003; Yandell et al., 2007; Yang et al., 2007; Le Rouzic and Alvarez-Castro, 2008).

There has been, however, a long-standing controversy concerning the importance of non-additive effects including epistasis. Much of the genetic variance within populations seems to be additive, although epistasis can also contribute to the additive variance (Hill et al., 2008). Further, despite many reports of statistical epistasis, there has only been limited success in identifying its functional basis. This has led to scepticism about the function epistasis has in populations, particularly given the difficulty in interpreting and replicating the biological consequences of interactions at just the pair-wise level (Phillips, 2008). In such a situation, one may ask how many of our findings are likely to be false positives. Unfortunately, the answer to that question is not immediately available because investigations on the false positive issue in detecting epistatic loci are surprisingly limited (Storey et al., 2005; Stich et al., 2007; Yang et al., 2007). This could well be because most attention in method development has been focused on increasing the capacity of modelling and the power of detection.

The Bayesian approach (and similar approaches using Markov Chain Monte Carlo algorithms) has been widely adopted to detect epistasis since its first application in 2001 (Sen and Churchill, 2001; Yi and Xu, 2002; Yi *et al.*, 2003; Sung and Wijsman, 2007; Xu and Jia, 2007; Yandell *et al.*, 2007; Yang *et al.*, 2007). Compared with other approaches (for example, regression), the Bayesian approach has some advantages including flexibility in model selection and limited requirement of multiple testing (Yi *et al.*, 2003). However, it also has some disadvantages in computational efficiency (Yi and Shriner, 2008) and repeatability of mapping results.

Correspondence: Dr D-J de Koning, The Roslin Institute and R(D)SVS, University of Edinburgh, Roslin, Midlothian, Scotland EH25 9PS, UK. E-mail: dj.dekoning@roslin.ed.ac.uk

Received 14 April 2009; revised 22 July 2009; accepted 27 July 2009; published online 30 September 2009

Type I error in mapping epistatic loci W-H Wei et a/

High-throughput analysis of many sets of data is required to understand the global gene interaction patterns in different species before the epistasis controversy can be really settled (Phillips, 2008). Limited by the daunting computing demand to cover high-dimensional search space and handle multiple tests, the majority of studies have investigated only a small set of phenotypic traits selected on the basis of personal interest. One way to increase the throughput is to use Grid computing resources (Seaton et al., 2006). The other is to develop fast and effective search algorithms (Ljungberg et al., 2004; Pattin et al., 2008). Using preidentified loci as prior information to detect epistasis in the one-dimensional scale was suggested to reduce the search dimension and, hence, boost speed (Jannink and Jansen, 2001; Carlborg and Haley, 2004; Evans et al., 2006; Kooperberg and Leblanc, 2008). Nevertheless, relying on the pre-identified signals was questionable because that approach would miss out those loci with strong epistatic interactions, but undetectable main effects (Xu and Jia, 2007; Phillips, 2008). The challenge is how to effectively incorporate the one-dimensional scans driven by preidentified loci into a higher-dimensional search.

This study was aimed to address the two questions of how to control false positives and how to effectively use the information of pre-identified loci in the mapping of epistatic quantitative trait loci (QTL) using large-scale simulations. As the issue of balancing false positive rate (FPR) and power is entangled with multiple testing methods and search algorithms, we adopted the following measures to keep the problem tractable: (a) focusing on pair-wise interactions in an F₂ population structure; (b) using the regression approach (Haley and Knott, 1992) to allow easy replication and interpretation; (c) using the conventional epistasis partition given by Jana (1971); (d) using exhaustive search to explore the entire search space and (e) using the nested test framework (Sen and Churchill, 2001) with modification to derive thresholds and perform multiple tests. The objectives are (1) to provide profiles of FPR and power for detecting different forms and levels of epistasis and (2) to find a solution to effectively combine one-dimensional scans with a two-dimensional scan in mapping epistatic loci. Ultimately, we hope to deliver a new application running on the Grid infrastructure to allow high-throughput epistasis analyses and produce results with good quality and known FPR for functional studies.

Materials and methods

Genetic models and epistasis parameterization

The regression method (Haley and Knott, 1992) was extended to map epistatic pairs of QTL (Carlborg and Andersson, 2002; Carapuco *et al.*, 2005). Considering only one pair of loci denoted as L_1 and L_2 , the genetic models have the following simplified forms:

Model 1: $y = \mu + L_1 + L_2 + L_1 + L_2 + e$ (two loci with epistasis)

Model 2: $y = \mu + L_1 + L_2 + e$ (two loci without epistasis) Model 3: $y = \mu + L_1 + e$ (single locus model)

Model 4: $y = \mu + e$ (null model)

where *y* is the trait of interest, μ is the model constant and *e* is the random error term. Additive and dominance genetic effects were modelled for each locus (a_1 and d_1 for L_1 ; a_2 and d_2 for L_2). The interaction between L_1 and L_2 (denoted as L_1L_2) was partitioned as $a_1 \times a_2$, $a_1 \times d_2$, $d_1 \times a_2$ and $d_1 \times d_2$ genetic components following Jana (1971).

Search algorithm

The search algorithm was developed to use effectively the pre-identified QTL with significant marginal effects (marginal-effect QTL) in detecting epistasis. The marginal-effect QTL are detected through a forwards selection approach (ignoring the possibility of epistasis) and tuned by retesting each QTL iteratively, while fitting the remaining QTL as cofactors until the number of QTL and their positions become stable (Wei *et al.*, 2007). The search algorithm is composed of two separate paths for the identification of pairs of loci with epistatic interactions (Figure 1):

- 1D_path—a one-dimensional genome scan for each pre-identified marginal-effect QTL searching for interactions with all other genomic positions. There may be several independent one-dimensional genome scans if more than one marginal-effect QTL has been identified.
- 2D_path—a full two-dimensional genome scan searching for epistatic interactions for loci at all combinations of two positions in the genome irrespective of the positions of pre-identified marginal-effect QTL.

Both paths used an exhaustive search at 1 centiMorgan (cM) intervals to find the best pair of loci (that is with the minimum residual variance or maximum aggregate genetic variance) under Model 1 for each pair of chromosomes and test them for epistasis. Genome-wide thresholds were derived in advance and a specific nested test framework was used for each path to identify epistatic pairs (see below).

Modified nested test framework

Sen and Churchill (2001) suggested a nested test framework that included an overall test (Model 1 vs Model 4, 8 degrees of freedom) for the aggregate effect of a pair of loci (that is the sum of the effects of each locus and their interaction) and an interaction test (Model 1 vs Model 2,



Figure 1 Flowchart of the search algorithm for detecting epistatic QTL in two separate paths (1D_path to the left and 2D_path to the right).

4 degrees of freedom) for the interaction component. The key property of this framework is that the interaction test is nested in the overall test; only pairs that passed the overall test would proceed to the interaction test and the two loci in Model 2 must be the same as those in Model 1. It was suggested to use permutation (Churchill and Doerge, 1994) to derive a genome-wide threshold for the overall test, but a nominal, tabulated, threshold for the interaction test (Sen and Churchill, 2001; Sugiyama *et al.*, 2001).

Using the F ratio test statistics for model comparison, the nested test framework was adopted here to determine the significance of epistasis. An epistatic pair is declared if both tests are significant. Some modification was made to apply the framework to the 1D path in which a marginal-effect QTL was known in advance and fixed in Model 1 when searching for epistasis. The overall test needs to ensure that the aggregate effect of a pair involving the marginal-effect QTL explains significantly more phenotypic variance than the marginaleffect QTL alone. In this case, the overall test is comparing Model 1 against Model 3, where L_1 represents the marginal-effect QTL in both models; this test thus takes only 6 degrees of freedom. The interaction test is still comparing Model 1 against Model 2 with L₁ representing the marginal-effect QTL in both models and this test takes 4 degrees of freedom as in the 2D_path.

Deriving genome-wide thresholds

Permutation (1000 replicates) was used to derive genomewide thresholds in advance. The DIRECT algorithm (DIviding RECTangle), a fast global optimization algorithm that finds the optima through systematically dividing the search space into smaller rectangles, was earlier adopted in QTL mapping studies (Ljungberg *et al.*, 2004). The DIRECT algorithm was applied here to perform fast two-dimensional scans in permutations to derive thresholds for the 2D_path. Exhaustive onedimensional genome scans were performed on permuted data to derive thresholds for each pre-identified marginal-effect QTL in the 1D_path. The procedures are outlined below:

- The 2D_path: For each permuted dataset a DIRECT search was performed to identify the best pair of loci under Model 1. This was tested against Model 4 to calculate and store the F ratio for the overall test. Model 2 was then fitted with the same positions of L_1 and L_2 and used to calculate and store the F ratio for the F ratio for the interaction test. The genome-wide thresholds for both the overall and interaction tests were derived separately from the corresponding lists of stored F ratios.
- For the 1D_path, thresholds were derived for each preidentified marginal-effect QTL. For each round of permutation, the original QTL probabilities were fitted for the identified QTL; then the one-dimensional genome scan was performed with permuted genotypes to find the L_2 giving the minimum residual variance of Model 1; the overall test was performed by comparing the Model 1 against the Model 3 with the QTL fitted to calculate and store an F ratio. The QTL and L_2 were then fitted to Model 2 to calculate and store the F ratio for the interaction test. The rest of the procedure was identical to that of the 2D_path.

Simulation design

An F_2 population was simulated in which the two founder lines contributed 30 individuals each (50% male). Fifteen founder sires from each line were each mated to a different founder dam from the other line (8 progeny per dam) producing 240 F_1 individuals. Thirty F_1 sires were randomly chosen and each mated to 3 F_1 dams (6 progeny per dam) forming an F_2 population with 540 individuals. A genome of 20 chromosomes was simulated with each chromosome carrying 11 microsatellite markers (four alleles with equal allele frequency per marker) evenly spaced at 10 cM with a length of 100 cM giving a total genome length of 2000 cM. The probability of crossovers between markers was generated using Haldane's mapping function.

The total phenotypic variance (Vartot) consisted of four components: additive genetic polygenic variance, nonepistatic marginal-effect QTL variance, epistatic QTL variance and random error variance. To mimic the polygenic effects, 10 bi-allelic loci were simulated with each at 0.5 recombination rate (that is not linked with any of the simulated markers) and allele frequency of 0.5 in the founder lines with additive genetic effects only, accounting for 1% of the Vartot, giving a total polygenic heritability of 10%. All QTL were simulated as bi-allelic and fixed for alternative alleles in the founder lines. Each non-epistatic QTL was assigned additive genetic effects only, accounting for 3% of the Vartot. For each epistatic pair of QTL, the eight genetic effects were simulated and the total genetic variance of the pair (Var_{pair}) and the epistatic variance (Var_{epi}) were calculated using the equations given in Jana (1971). Three groups of simulation scenarios were defined based on the epistatic heritability ($h_{epi}^2 = Var_{epi}/Var_{tot}$), rather than the wholepair heritability $(h_{pair}^2 = Var_{pair}/Var_{tot})$. Each scenario included polygenic effects, except when explicitly stated, and was tested with at least 550 replicates.

The first group was non-epistatic in which six simulation scenarios were designed to test how the algorithm controlled FPR under different genetic backgrounds in which epistasis was not present. These scenarios were noGenetic—neither polygenic nor marginal-effect QTL; 0QTL_noEpi—only polygenic, not QTL; 1QTL_noEpi—one marginal-effect QTL; 2QTL_ noEpi—two marginal-effect QTL; 5QTL_noEpi—five marginal-effect QTL; 8QTL_noEpi—eight marginal-effect QTL. The total QTL variance was the sum of marginal-effect QTL with a variance of 3% of the Var_{tot} each, for example, 24% in the 8QTL_noEpi scenario. Each marginal-effect QTL was simulated at either 85 or 15 cM on one of the chromosomes of 1, 2, 4, 6, 8, 17, 18 and 20 as appropriate.

The second group was the core of the study in which 20 scenarios were defined to test the algorithm in different combinations of epistasis forms and effect sizes to produce profiles of FPR and power in a simplified condition allowing at most one QTL per chromosome. Each of the 20 scenarios had one non-epistatic marginal-effect QTL positioned at 85 cM on chromosome 1 and one epistatic QTL pair in which the first locus of the epistatic pair was positioned at 15 cM on chromosome 2 and the second at 85 cM on chromosome 3. The epistatic QTL pair was simulated to have one of the six forms of epistasis: complementary, duplicate, dominant, recessive, inhibitory (Jana, 1971, 1972; Carlborg *et al.*, 2000)

Type I error in mapping epistatic loci W-H Wei et al

Table 1 Overview and names of the core simulation scenarios							
Epistasis form	$2.5\% h_{epi}^2$	5.0% h_{epi}^2	7.5% h_{epi}^2	$10.0\% h_{epi}^2$	12.5% h _{epi} ²		
Complementary	Com2.5% (17.6)	Com5.0% (35.0)	Com7.5% (52.5)		_		
Dominant	Dom2.5% (6.8)	Dom5.0% (15.6)	Dom7.5% (25.0)		_		
Recessive	Rec2.5% (10.0)	Rec5.0% (20.2)	Rec7.5% (30.5)		_		
Inhibitory	Inh2.5% (10.9)	Inh5.0% (21.7)	Inh7.5% (32.5)		_		
Duplicate	Dup2.5% (4.2)	Dup5.0% (8.3)	Dup7.5% (12.5)		_		
Interaction only	Int2.5% (2.5)	Int5.0% (5.0)	Int7.5% (7.5)	Int10.0% (10.0)	Int12.5% (12.5)		
Dominant Recessive Inhibitory Duplicate Interaction only	Dom2.5% (6.8) Rec2.5% (10.0) Inh2.5% (10.9) Dup2.5% (4.2) Int2.5% (2.5)	Dom5.0% (15.6) Rec5.0% (20.2) Inh5.0% (21.7) Dup5.0% (8.3) Int5.0% (5.0)	Dom7.5% (25.0) Rec7.5% (30.5) Inh7.5% (32.5) Dup7.5% (12.5) Int7.5% (7.5)	 Int10.0% (10.0)	 Int12.5%		

Table 1 Overview and names of the core simulation scenarios^a

^aThe total phenotypic variance (Var_{tot}) was 0.98; the epistatic heritability of a pair of loci (h_{epi}^2 = the epistatic variance of an epistatic pair/Var_{tot}); the whole-pair heritability (h_{pair}^2 = the total genetic variance of an epistatic pair/Var_{tot}) in brackets.

and interaction only (that is additive × additive, without main effects), with an epistatic heritability (h_{epi}^2) of either 2.5, 5 or 7.5% (Table 1).

The third group of scenarios was used to validate the algorithm in two more complicated conditions: (a) the non-epistatic marginal-effect QTL was placed at 85 cM on chromosome 2 in which the first locus of the epistatic pair was simulated at 15 cM (linked loci); (b) in addition to the simplified condition, a second epistatic pair was simulated (with the same form and effect size of epistasis as the first pair) at 85 cM on chromosome 19 and 15 cM on chromosome 20 (two pair). Each epistatic pair was simulated to have 5% h_{epi}^2 with one of the six forms of epistasis giving six scenarios for either the linked-loci or the two-pair condition.

Analysis of simulation results

Power and FPR: For non-epistatic QTL, a QTL was counted as a true positive if mapped to the chromosome in which it was simulated, otherwise it was considered as a false positive. For epistasis, an epistatic pair was regarded as a true positive if both loci each mapped to the chromosome in which they were simulated, otherwise it was considered as a false positive. FPR was calculated per scenario as the percentage of the total number of false positives out of the number of simulation replicates. Power was calculated per non-epistatic QTL or per epistatic pair as the percentage of the total number of detected true positives out of the number of simulation replicates. An average power was given when multiple non-epistatic QTL were present. In the two-pair epistatic scenarios, power was reported for each pair.

Accuracy: The accuracy of locating each simulated nonepistatic QTL or epistatic pair was assessed in two ways: (a) exact (%)—the percentage of the total number occurrence of locating the true positive QTL in the simulated 10 cM marker interval (or the total number of occurrence of locating both loci of the true positive epistatic pair to the corresponding simulated marker intervals) out of the number of simulation replicates and (b) precision (cM)—the average distance between the detected position and the simulated position for each locus of a true positive epistatic pair. An average accuracy was given when multiple non-epistatic QTL were present. The exact (%) accuracy was reported for each pair in the two-pair epistatic scenarios.

Overlap between 2D_path and 1D_path: Overlaps (that is the epistatic pairs detected in both paths) existed because scans conducted in the 1D_path were always covered in a full two-dimensional scan, but tested using different thresholds in the two cases. A QTL zone was defined as a chromosome region around each marginaleffect QTL in which any locus within the zone was regarded as the QTL for sharing similar information. Overlaps can be greatly avoided by skipping any QTL zones searched in the 1D_path from the 2D_path search space. The QTL zone length (for example, 10 cM on either side of a QTL) is critical to determine the amount of overlaps in the 2D_path as well as the overall FPR and power after combining the 1D_path with the 2D_path without QTL zones. Different QTL zone lengths of 10, 20, 30 and 40 cM were compared by re-analysing the simulation results from the 2D_path to determine the most effective length.

Results

Thresholds

The genome-wide thresholds derived for both the 2D_path and 1D_path were fairly consistent across scenarios. For the 2D_path, the averaged genome-wide thresholds were 5.28 and 8.22 (corresponding to a nominal *P*-value of 2.2×10^{-6} and 1.9×10^{-6}) for the overall and interaction tests, respectively. For the 1D_path, the averaged genome-wide thresholds were 4.28 and 5.08 (corresponding to a nominal *P*-value of 3.2×10^{-4} and 5.1×10^{-4}) for the overall and interaction tests, respectively. For the interaction tests, respectively. In both cases, the threshold applied for the interaction tests were dramatically more stringent than the recommended nominal *P*-values of 0.01 (Sugiyama *et al.*, 2001).

FPR and power profiles of mapping marginal-effect QTL At least 1100 replicates were simulated for each nonepistatic scenario to provide clear profiles of FPR and power of detecting marginal-effect QTL used for mapping epistatic loci. On the basis of the 5% genome-wide threshold, the FPR, average power, accuracy as well as the number of marginal-effect QTL detected were calculated for each scenario as appropriate (Table 2). The results showed that given a QTL heritability of 3%, there was around 50% power to detect a marginal-effect QTL and about 30% probability to locate it in the right marker interval.

FPR profile of mapping epistatic pairs in the non-epistatic scenarios

For each non-epistatic scenario, the FPR of epistatic pairs was measured for both the 2D_path and 1D_path using the 5% genome-wide thresholds for the overall and

interaction tests (Figure 2). It was obvious that the FPR values from the 2D_path were quite consistent across scenarios. The 1D_path FPR, however, increased as the average number of marginal-effect QTL increased and departed from the target 5% dramatically when two or more QTL were detected (Table 2). Considering that onedimensional genome scans were performed for each marginal-effect QTL independently, the 1D_path thresholds were adjusted in three steps: (1) the 5% genomewide F ratio threshold was converted to a nominal P-value using appropriate degrees of freedom; (2) the *P*-value was divided by the number of QTL detected and (3) the corrected *P*-value was converted back to obtain an adjusted F ratio threshold. After applying the adjusted thresholds, the resulting 1D path FPR was at the correct level (the 1D_path_corrected series in Figure 2). The adjustment procedure was applied from this point onwards to analyse all the simulation results from the 1D path.

Analysing overlaps between the 1D_path and 2D_path results using different QTL zone lengths (area around a marginal-effect QTL that was not tested for epistasis using the 2D_path) uncovered that increased FPR was a concern in scenarios in which multiple marginal-effect QTL were detected. When skipping QTL zones with a length of 10 cM, the 2D_path FPR reduced from 1.82 to 1.36 in the scenario with five simulated QTL (5QTL_noEpi) and from 3.62 to 2.99 in the scenario with eight simulated QTL (8QTL_noEpi). Increasing the zone length to 20 cM or above made only limited difference in the scenario with eight simulated QTL (further reduced the 2D_path FPR to 2.53) while causing considerable reduction of the 2D_path search space.

The final FPR values for each non-epistatic scenario (the Final series in Figure 2) were calculated by combining the new 1D_path FPR results from using the corrected thresholds with the new 2D_path FPR by skipping QTL zones with a length of 10 cM. These FPR values were below or close to a target level of 5% across scenarios. Thus, this way of integrating results from the 1D_path and 2D_path was applied hereafter to calculate final results.

FPR and power profiles of mapping epistatic pairs in the core scenarios

The FPR and power profiles of mapping epistatic pairs are shown in Table 3 for the 20 core simulation scenarios (Table 1). For each scenario, the FPR and power were calculated for the 1D_path using the corrected 5%

Table 2 False positive rate (FPR), power and accuracy in mapping marginal-effect QTL in the non-epistatic scenarios^a

	QTL^{b}	FPR (%)	Power (%) ^c	Exact (%) ^c
noGenetic	0.1	5.4		_
0QTL noEpi	0.1	6.1	_	_
1QTL noEpi	0.5	5.4	48.2	28.4
2OTL noEpi	1.0	4.5	48.5	27.4
5OTL noEpi	2.6	3.7	51.1	28.1
8QTL_noEpi	4.4	4.2	54.8	35.0

^aEach scenario used at least 1100 simulation replicates.

^bThe average number of marginal-effect QTL detected per replicate. ^cAveraged by the number of non-epistatic QTL simulated in a scenario. genome-wide thresholds ('1D_path' columns, Table 3) and the 2D_path using the 5% genome-wide thresholds ('2D_path' columns). In addition, the 2D_path results after skipping QTL zones with a length of 10 cM and the final integrated results were calculated for each scenario ('2D path no QTL' and 'Final' columns, respectively).

The final FPR values varied around the target 5% level across scenarios (Table 3). Skipping QTL zones from the two-dimensional search space made rather small reduction in FPR values, possibly because the 2D_path FPR values were consistently low across scenarios already. Most of the 1D_path FPR values were below 4% showing that the threshold correction worked well in general. There were, however, a few occasions in which the 1D_path FPR values were slightly >5% (thus the final FPR values slightly above 7% in scenarios with inhibitory or dominant epistasis) given only 550 replicates used per scenario.

The power profiles were more complicated (Table 3). Within a model of epistasis, the final power of detecting epitasis increased as the h_{epi}^2 increased. At the same level of h_{epi}^2 , the complementary (both loci were marginaleffect QTL) scenario had the highest power value; followed by the inhibitory, dominant and recessive (one locus was marginal-effect QTL) scenarios; the duplicate (one locus was marginal-effect QTL at a chance of 1 out 16) and interaction-only (neither locus with main effects) scenarios had the lowest power. However, that order was almost reversed if we based the ranking on the wholepair heritability (that is h_{pair}^2): the interaction-only and duplicate scenarios had higher power than the remaining (Figure 3). For example, at $10\% h_{pair}^2$, the interaction-only scenario had about 80% power, whereas the complementary scenario had a power <10%.

In scenarios with complementary, dominant, recessive or inhibitory epistasis, the 1D_path found nearly every



Figure 2 The FPR profiles of mapping epistatic pairs generated from different search paths in different non-epistatic genetic backgrounds and their controls. Along the *x* axis are six non-epistatic scenarios (from the left) simulating a genetic background with neither polygenic nor marginal-effect QTL effects, polygenic plus either zero, one, two, five or eight marginal-effect QTL effects. Within each non-epistatic scenario, any epistatic QTL pair detected from either the 2D_path (a full two-dimensional genome scan irrespective of pre-identified QTL, using 5% genome-wide thresholds) or the 1D_path (one-dimensional genome scans for each pre-identified QTL, using 5% genome-wide thresholds) counted as false positives. The 1D_path_corrected FPR calculated by testing each false positive pair from the 1D_path against stringent thresholds corrected by the number of pre-identified QTL. The Final FPR calculated by combing the FPR results from the 2D_path and the 1D_path_corrected after skipping any overlaps.

Type I error in	mapping epi	static	loci
	W-H	Wei e	et al

Table 3 False positive rate (FPR) and	power profiles of mappi	ng an epistatic pair in t	he core scenarios ^a
		() 1 1	

	FPR (%)			Power (%)				
	1D_path	2D_path	2D_path no_QTL ^b	Final ^c	1D_path	2D_path	2D_path no_QTL ^b	Final ^c
Com2.5%	3.7	1.2	0.5	4.2	15.6	2.7	0.0	15.6
Com5.0%	2.0	1.6	1.3	3.3	69.3	32.2	0.0	69.3
Com7.5%	1.4	0.7	0.4	1.8	99.1	90.0	0.0	99.1
Dup2.5%	2.0	1.5	1.5	3.5	1.8	1.6	1.2	3.0
Dup5.0%	4.4	2.3	2.0	6.4	14.7	21.8	15.5	30.1
Dup7.5%	2.0	1.6	1.6	3.6	46.6	57.7	29.3	75.9
Dom2.5%	3.3	1.0	0.8	4.1	6.6	2.6	1.7	8.3
Dom5.0%	5.7	2.1	2.0	7.7	55.7	19.8	0.7	56.4
Dom7.5%	5.6	2.0	1.7	7.4	90.8	65.2	0.3	91.1
Rec2.5%	3.3	0.7	0.5	3.8	8.7	1.5	0.2	8.8
Rec5.0%	4.1	1.0	0.8	5.0	51.7	23.3	0.0	51.7
Rec7.5%	4.2	0.5	0.4	4.5	89.7	67.5	0.0	89.7
Inh2.5%	5.1	2.4	2.0	7.1	13.8	2.0	0.0	13.8
Inh5.0%	5.3	2.0	1.8	7.2	61.3	24.8	0.0	61.3
Inh7.5%	5.5	1.9	1.6	7.1	92.3	70.9	0.0	92.3
Int2.5%	1.8	0.4	0.4	2.2	0.0	0.9	0.9	0.9
Int5.0%	1.6	2.0	2.0	3.6	0.0	12.4	12.4	12.4
Int7.5%	0.9	2.9	2.7	3.6	0.0	46.8	46.8	46.8
Int10.0%	0.9	1.3	1.3	2.2	0.2	78.5	78.3	78.5
Int12.5%	2.3	2.0	1.8	4.1	0.2	92.8	92.6	92.8

^aEach scenario with one non-epistatic QTL and one epistatic pair.

^bOmitting regions of 10 cM on either side of the marginal-effect QTL pre-indentified from the 2D_path results.

^cCombined results from the 1D_path and those from the 2D_path after omitting regions of 10 cM on either side of the marginal-effect QTL pre-indentified.



Figure 3 The power profile of detecting epistatic pairs with different forms of epistasis in the context of the whole-pair heritability (the percentage of phenotypic variance explained by the whole-epistatic pair of QTL). Each marked point corresponding to a level of the epistatic heritability of either 2.5, 5 or 7.5% (the

interaction-only form with two extra levels of 10 and 12.5%).

The whole pair heritability(%)

epistatic pairs that could be detected and the 2D_path found essentially the same ones, but fewer of them (Table 3). In scenarios with duplicate epistasis, however, both the 1D_path and 2D_path found around 50% epistatic pairs uniquely (that is there was much less overlap in the 2D_path)—nearly half of those found by the 2D_path were not found by the 1D_path and a third of those found by the 1D_path were not found by the 2D_path. In scenarios with interaction-only epistasis, the 2D_path contributed solely to the final power because the 1D_path had no or rather limited power.

The accuracy in mapping epistasis improved as the $h_{\rm epi}^2$ increased (Table 4). The exact (%) results were in line with the final power results across scenarios, that is the chance of locating both loci into the simulated marker intervals increased as the power increase. The average precision (average distance between the mapped and simulated positions) of each locus of an epistatic pair detected from either the 2D_path or 1D_path was always below 5 cM when h_{epi}^2 was 5.0% or 7.5%. Furthermore, the majority of the QTL were mapped within 10 cM of the simulated location implying that they were mapped to either the simulated or the adjacent marker interval. Almost all remaining QTL were mapped within 20 cM of the simulated position. Such results suggested that the FPR and power profiles would remain unchanged even if the definition of a true positive epistatic pair had included a restriction that the two loci each mapped to the correct chromosome and within 20 cM of the simulated position.

FPR and power profiles in the validation scenarios

Analyses of the 12 validation scenarios showed that the 2D_path and 1D_path behaved almost identically to the core scenarios under the simplified condition. The final FPR and power results from combining the 1D_path and 2D_path for each validation scenario are displayed in Figure 4. Clearly, the FPR values of each validation scenario were similar to those in Table 3 and controlled around the 5% target level. The power profiles for the

Table 4 Accuracy of mapping epistatic pair in the core scenarios^a

	Exact ^b (%)	Locus 1 precision (cM) ^c		Locus 2 precision (cM) ^c	
		1D_path	2D_path	1D_path	2D_path
Com2.5%	10.1	3.0 (96.8, 3.2)	2.7 (100, 0.0)	3.4 (95.7, 4.3)	1.9 (100, 0.0)
Com5.0%	64.2	1.7 (100, 0.0)	1.8 (100, 0.0)	1.7 (99.8, 0.2)	1.7 (100, 0.0)
Com7.5%	96.6	1.3 (100, 0.0)	1.3 (100, 0.0)	1.3 (100, 0.0)	1.3 (100, 0.0)
Dup2.5%	1.5	5.1 (88.9, 11.1)	3.8 (100, 0.0)	4.3 (88.9, 11.1)	3.9 (100, 0.0)
Dup5.0%	15.8	4.2 (85.9, 14.1)	3.8 (93.3, 6.7)	4.9 (85.9, 12.9)	4.0 (89.6, 10.4)
Dup7.5%	41.9	4.8 (89.5, 9.8)	4.0 (92.8, 6.5)	4.7 (90.2, 9.0)	3.5 (94.1, 5.6)
Dom2.5%	3.3	4.3 (88.6, 11.4)	4.0 (87.5, 12.5)	8.8 (80.0, 14.3)	8.1 (75.0, 12.5)
Dom5.0%	35.1	2.8 (98.5, 1.5)	2.3 (98.3, 1.7)	4.6 (91.0, 7.1)	4.2 (93.4, 4.1)
Dom7.5%	66.7	1.7 (99.7, 0.3)	2.0 (99.4, 0.6)	3.5 (95.7, 4.1)	3.7 (94.9, 4.8)
Rec2.5%	4.6	6.4 (81.1, 18.9)	6.4 (88.9, 0.0)	3.2 (97.3, 2.7)	4.9 (77.8, 22.2)
Rec5.0%	32.4	5.2 (84.4, 14.1)	4.8 (86.5, 11.3)	1.9 (100, 0.0)	1.9 (110.0, 0.0)
Rec7.5%	62.9	4.4 (92.1, 6.0)	3.9 (94.4, 4.8)	1.7 (99.6, 0.4)	1.5 (99.7, 0.3)
Inh2.5%	7.3	2.7 (100, 0.0)	3.5 (100, 0.0)	5.2 (87.7, 12.3)	4.3 (100, 0.0)
Inh5.0%	39.8	2.2 (99.4, 0.6)	1.9 (100, 0.0)	4.4 (91.3, 7.0)	3.4 (96.6, 3.4)
Inh7.5%	72.3	1.6 (99.9, 0.1)	1.5 (100, 0.0)	3.4 (95.6, 4.0)	3.3 (96.3, 3.3)
Int2.5%	0.0		5.0 (100, 0.0)	—	8.1 (80.0, 20.0)
Int5.0%	6.7		4.9 (89.9, 10.1)	—	4.2 (91.3, 8.7)
Int7.5%	24.1		4.4 (91.1, 8.5)	—	4.3 (89.9, 10.1)
Int10.0%	45.3		3.9 (92.6, 7.1)		4.0 (93.1, 6.4)
Int12.5%	58.9	—	3.5 (95.5, 4.5)	—	3.6 (94.8, 4.9)

^aEach scenario with one non-epistatic QTL and one epistatic pair.

^bThe percentage of the number of true positive epistatic pairs in which both loci mapped to the simulated marker interval out of the total simulation replicates.

^cThe average distance between the mapped and simulated QTL positions for significant epistatic pairs given for each QTL separately; the first (second) number in brackets is the percentage of distances $\leq 10 \text{ cM}$ (>10 and $\leq 20 \text{ cM}$) out of the total number of true positives.

linked-loci and the two-pair conditions were very similar to each other and to that for the core scenarios with 5% $h_{\rm epi}^2$. The power of each linked-loci scenario was nearly identical to that of the corresponding core scenario, indicating that the algorithm was robust to handle linked-loci situations. Small power reduction was observed in some two-pair scenarios in contrast to the linked-loci scenarios possibly because of the threshold correction as the number of marginal-effect QTL detected in a two-pair scenario almost doubled and consequently the penalty almost doubled as well.

Discussion

In this large-scale simulation study, we examined different aspects of the issue of false positive detection of epistatic loci. We showed performance differences between the 2D_path and 1D_path in mapping different forms and levels of epistasis and produced clear profiles of FPR and power that were validated in more complicated simulation scenarios. In addition, we found that both the 2D_path and 1D_path had strengths and weaknesses in detecting certain forms of epistasis. The strength in one happened to be the weakness of the other. The two paths can be effectively combined by skipping any QTL zones with a length of 10 cM from the 2D_path search space.

It was shown throughout the study that a good control of FPR was achieved by applying the pre-derived genome-wide thresholds to the nested overall and interaction tests based on the modified nested test framework. In a parallel simulation study using the same simulation scenarios and scales, we evaluated an alternative approach that considered only the interaction test in both searching and testing for epistasis, but found hard to control FPR in general (results can be made available on request). The modification of the nested test framework seemed to be very important in increasing power of detection of epistasis associated with preidentified QTL by allowing the one-dimensional scans for epistatic pairs as implemented in the 1D_path and improving the overall search speed because one-dimensional scans were faster than a two-dimensional scan, and the two-dimensional search space would be reduced as a result of the skip of the QTL zones. The framework can be applied to other population structures to help control false positive signals of epistasis. Nevertheless, the theoretical foundation of the framework (for example, the expected FPR level in the 1D_path or 2D_path) requires further work.

The trade-off between FPR and power has always been challenging in multiple testing situations (Rice *et al.*, 2008). Considering the difficulty and expense in identifying the functional consequence of a statistical epistasis (Phillips, 2008), a low FPR level is important to encourage biologists engagement in functional epistasis studies. Using the 5% genome-wide thresholds (with correction for the 1D_path), we controlled the FPR close to the 5% conventional level. Using less stringent thresholds, for example, 5% genome-wide threshold for the overall tests, but 10% genome-wide thresholds for the interaction tests, resulted in a slight increase of power, but considerably higher FPR (results not shown) and thus is not recommended.

The F₂ population size of 540 is common in normal linkage studies, but could be small for epistasis analysis as suggested by Carlborg *et al.* (2006). Our results (Figure 3) showed that for a population of 540 F₂, the power to detect an epistatic pair explaining 10% phenotypic variance was about 80% if the pair had an interaction-only form ($h_{epi}^2 = 10\%$), 40% if it had a duplicate form ($5.0\% < h_{epi}^2 < 7.5\%$), 20% if it had a dominant



Type I error in mapping epistatic loci

Figure 4 The power and FPR profiles of mapping epistatic pairs in the validation scenarios. Along the *x* axis from the left are scenarios with 5% complementary, dominant, recessive, inhibitory, duplicate, interaction-only epistasis. Linked loci: two QTL on chromosome 2, but not interacting with each other, but one of them interacting with a locus on chromosome 3; two pairs: two epistatic pairs simulated each with a same form of epistasis and 5% epistatic heritability in which FPR calculated per scenario, but power calculated per pair per scenario and displayed as two-pair 1 and 2.

form $(2.5\% < h_{epi}^2 < 5.0\%)$ and < 10% if it had either recessive or inhibitory or complementary form ($h_{epi}^2 < 2.5\%$). These results suggested that it was relatively easy to find strong epistasis (for example, $h_{epi}^2 > 7.5\%$), but difficult to detect weak ones given the population size. Our observations are roughly in line with earlier studies: four interaction-only pairs detected from an F2 mouse population (size is 510, 166 females) explained about 36% of the total variation in litter size (that is $h_{epi}^2 = 9\%$ on average) (Peripato *et al.*, 2004); most epistatic pairs detected for obesity-related traits using 513 F₂ mice had a $h_{epi}^2 > 2\%$ despite the relatively relaxed thresholds that were used (Stylianou et al., 2006). Increasing the population size is a good option to increase power (Carlborg et al., 2003; Ma et al., 2009). For example, according to our additional simulation results (Supplementary Figure S1), when increasing the population size from 540 to 840, scenarios with 2.5% h_{epi}^2 at least doubled the power, and scenarios with 5.0% h_{epi}^2 at the population size of 840 had power comparable to those with a h_{epi}^2 of 7.5% at the size of 540.

Even at the population size of 540, the accuracy of mapping epistatic pairs, without any backwards or forwards tuning in place yet, was generally good across all epistasis forms in the context of an F_2 population and increased as epistasis gets stronger. Those accuracy results suggested that the simple regression approach and the non-orthogonal epistasis partition might be sufficient to correctly map epistatic pairs. Considering that the regression models are easy to extend and fast to compute, the modelling approach we adopted here remains competitive in mapping epistasis.

The whole simulation study cost was >10 CPU years, but was computed in a couple of weeks by distributed computing technologies, so is high throughput in itself. On the basis of the work in this study, the task of highthroughput epistasis analyses becomes achievable at least in populations derived from structured crosses. Skipping the QTL zones in the two-dimensional scan would effectively reduce the search space of the scan and, hence, the overall computing time. By combining the 2D_path and 1D_path, the algorithm is balanced between search efficiency and speed thus suitable to be used as the search engine for high-throughput epistasis analyses. Considering that the genome-wide threshold values were consistent across scenarios at a given genome size, it is possible to derive those thresholds in advance for different genome and population sizes and use them directly to save permutation time in multiple epistasis analyses (Broman et al., 2003).

In summary, using the modified nested test framework to perform nested tests for epistasis in the search process controlled by the combined search algorithm (Figure 1) allowed effectively mapping different forms of epistasis while keeping FPR under control. Once integrated with distributed computing resources, the new distributed application can support high-throughput epistasis analysis.

Acknowledgements

This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (http://www.ecdf.ed.ac.uk/). The ECDF is partially supported by the eDIKT initiative (http://www.edikt.org. uk). We are grateful for the funding from BBSRC through the GridQTL project (BBSB1695X) and the EC funded Integrated Project SABRE (EC contract number FOOD-CT-2006-01625). The contributions of WHW, DJdeK and CSH were funded in part by a BBSRC Institute Strategic Programme Grant. SAK is funded by Research Councils UK. WHW and CSH were also funded by MRC core funding at HGU. We acknowledge our gratitude to Kajsa Ljungberg for providing the DIRECT source code, Orjan Carlborg and Gary Churchill for valuable discussions, Jean-Alain Grunchec and Andy Law for assistance in setting up data analysis on ECDF. The authors are grateful for the suggestions by the Heredity subject editor and two anonymous referees who made this paper more accessible to a wider readership.

References

- Alvarez-Castro JM, Carlborg O (2007). A unified model for functional and statistical epistasis and its application in quantitative trait Loci analysis. *Genetics* **176**: 1151–1167.
- Alvarez-Castro JM, Le Rouzic A, Carlborg O (2008). How to perform meaningful estimates of genetic effects. *PLoS Genet* **4**: e1000062.
- Aylor DL, Zeng Z-B (2008). From classical genetics to quantitative genetics to systems biology: modeling epistasis. *PLoS Genet* **4**: e1000029.
- Broman KW, Wu H, Sen S, Churchill GA (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**: 889–890.
- Bush WS, Dudek SM, Ritchie MD (2006). Parallel multifactor dimensionality reduction: a tool for the large-scale analysis of gene-gene interactions. *Bioinformatics* **22**: 2173–2174.

408

- Carapuco M, Novoa A, Bobola N, Mallo M (2005). Hox genes specify vertebral types in the presomitic mesoderm. *Genes Dev* **19**: 2116–2121.
- Carlborg O, Andersson L (2002). Use of randomization testing to detect multiple epistatic QTLs. *Genet Res* **79**: 175–184.
- Carlborg O, Andersson L, Kinghorn B (2000). The use of a genetic algorithm for simultaneous mapping of multiple interacting quantitative trait loci. *Genetics* **155**: 2003–2010.
- Carlborg O, Brockmann GA, Haley CS (2005). Simultaneous mapping of epistatic QTL in DU6i x DBA/2 mice. *Mamm Genome* **16**: 481–494.
- Carlborg O, Haley CS (2004). Epistasis: too often neglected in complex trait studies? *Nat Rev Genet* 5: 618–625.
- Carlborg O, Jacobsson L, Ahgren P, Siegel P, Andersson L (2006). Epistasis and the release of genetic variation during long-term selection. *Nat Genet* 38: 418–420.
- Carlborg O, Kerje S, Schutz K, Jacobsson L, Jensen P, Andersson L (2003). A global search reveals epistatic interaction between QTL for early growth in the chicken. *Genome Res* 13: 413–421.
- Churchill GA, Doerge RW (1994). Empirical threshold values for quantitative triat mapping. *Genetics* **138**: 963–971.
- de Visser JAGM, Elena SF (2007). The evolution of sex: empirical insights into the roles of epistasis and drift. *Nat Rev Genet* 8: 139–149.
- Evans DM, Marchini J, Morris AP, Cardon LR (2006). Two-stage two-locus models in genome-wide association. *PLoS Genet* **2**: e157.
- Hahn LW, Ritchie MD, Moore JH (2003). Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. *Bioinformatics* **19**: 376–382.
- Haley CS, Knott SA (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**: 315–324.
- Hill WG, Goddard ME, Visscher PM (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet* **4**: e1000008.
- Jana S (1971). Simulation of quantitative characters from qualitatively acting genes. I. Nonallelic gene interactions involving two or three loci. *Theor Appl Genet* **41**: 216–226.
- Jana S (1972). Simulation of quantitative characters from qualitatively acting genes. II. Orthogonal subdivision of hereditary variance in two-locus genetic systems. *Theor Appl Genet* **42**: 119–124.
- Jannink J-L, Jansen R (2001). Mapping epistatic quantitative trait loci with one-dimensional genome searches. *Genetics* **157**: 445–454.
- Kooperberg C, Leblanc M (2008). Increasing the power of identifying gene x gene interactions in genome-wide association studies. *Genet Epidemiol* **32**: 255–263.
- Le Rouzic A, Alvarez-Castro JM (2008). Estimation of genetic effects and genotype-phenotype maps. *Evol Bioinform* **4**: 225–235.
- Ljungberg K, Holmgren S, Carlborg O (2004). Simultaneous search for multiple QTL using the global optimization algorithm DIRECT. *Bioinformatics* **20**: 1887–1895.
- Ma J, Ren J, Guo Y, Duan Y, Ding N, Zhou L *et al.* (2009). Genome-wide identification of quantitative trait loci for carcass composition and meat quality in a large-scale White Duroc x Chinese Erhualian resource population. *Anim Genet* **40**: 637–647.
- Malmberg RL, Mauricio R (2005). QTL-based evidence for the role of epistasis in evolution. *Genet Res* 86: 89–95.
- Mechanic LE, Luke BT, Goodman JE, Chanock SJ, Harris CC (2008). Polymorphism interaction analysis (PIA): a method for investigating complex gene-gene interactions. *BMC Bioinformatics* **9**: 146.
- Moore JH, Barney N, Tsai CT, Chiang FT, Gui J, White BC (2007). Symbolic modeling of epistasis. *Hum Hered* **63**: 120–133.

- Pattin KA, White BC, Barney N, Gui J, Nelson HH, Kelsey KT *et al.* (2008). A computationally efficient hypothesis testing method for epistasis analysis using multifactor dimensionality reduction. *Genet Epidemiol* **31**: 31.
- Peripato AC, De Brito RA, Matioli SR, Pletscher LS, Vaughn TT, Cheverud JM (2004). Epistasis affecting litter size in mice. *J Evol Biol* **17**: 593–602.
- Phillips PC (2008). Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet* **9**: 855–867.
- Rice TK, Schork NJ, Rao DC (2008). Methods for handling multiple testing. *Adv Genet* **60**: 293–308.
- Ritchie MD, Motsinger AA (2005). Multifactor dimensionality reduction for detecting gene-gene and gene-environment interactions in pharmacogenomics studies. *Pharmacogenomics* 6: 823–834.
- Seaton G, Hernandez J, Grunchec JA, White I, Allen J, De Koning DJ et al. (2006). Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Vol. 27. Belo Horizonte, Brazil, session 27-07.
- Sen S, Churchill GA (2001). A statistical framework for quantitative trait mapping. *Genetics* **159**: 371–387.
- Stich B, Yu J, Melchinger AE, Piepho HP, Utz HF, Maurer HP et al. (2007). Power to detect higher-order epistatic interactions in a metabolic pathway using a new mapping strategy. *Genetics* 176: 563–570.
- Storey JD, Akey JM, Kruglyak L (2005). Multiple locus linkage analysis of genomewide expression in yeast. *PLoS Biol* **3**: e267.
- Stylianou IM, Korstanje R, Li R, Sheehan S, Paigen B, Churchill GA (2006). Quantitative trait locus analysis for obesity reveals multiple networks of interacting loci. *Mamm Genome* 17: 22–36.
- Sugiyama F, Churchill GA, Higgins DC, Johns C, Makaritsis KP, Gavras H et al. (2001). Concordance of murine quantitative trait loci for salt-induced hypertension with rat and human loci. *Genomics* **71**: 70–77.
- Sung YJ, Wijsman EM (2007). Accounting for epistasis in linkage analysis of general pedigrees. *Hum Hered* **63**: 144–152.
- Wang T, Zeng ZB (2006). Models and partition of variance for quantitative trait loci with epistasis and linkage disequilibrium. *BMC Genet* **7**: 9.
- Wei WH, de Koning DJ, Penman JC, Finlayson HA, Archibald AL, Haley CS (2007). QTLs modulating ear size and aspect in pigs. *Anim Genet* **38**: 222–226.
- Xu S, Jia Z (2007). Genome-wide analysis of epistatic effects for quantitative traits in barley. *Genetics* **175**: 1955–1963.
- Yandell BS, Mehta T, Banerjee S, Shriner D, Venkataraman R, Moon JY et al. (2007). R/qtlbim: QTL with bayesian interval mapping in experimental crosses. *Bioinformatics* 23: 641–643.
- Yang J, Zhu J, Williams RW (2007). Mapping the genetic architecture of complex traits in experimental populations. *Bioinformatics* **23**: 1527–1536.
- Yi N, Shriner D (2008). Advances in Bayesian multiple quantitative trait loci mapping in experimental crosses. *Heredity* **100**: 240–252.
- Yi N, Xu S (2002). Mapping quantitative trait loci with epistatic effects. *Genet Res* **79**: 185–198.
- Yi N, Xu S, Allison DB (2003). Bayesian model choice and search strategies for mapping interacting quantitative trait Loci. *Genetics* **165**: 867–883.
- Zeng ZB, Wang T, Zou W (2005). Modeling quantitative trait Loci and interpretation of models. *Genetics* **169**: 1711–1725.
- Zhang Y, Liu JS (2007). Bayesian inference of epistatic interactions in case-control studies. *Nat Genet* **39**: 1167–1173.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)