Genet. Res., Camb. (2008), **90**, *pp.* 363–374. © 2008 Cambridge University Press
doi:10.1017/S0016672308009336 Printed in the United Kingdom

Detecting dominant QTL with variance component analysis in simulated pedigrees

SUZANNE J. ROWE[1,2,*], RICARDO PONG-WONG[1], CHRISTOPHER S. HALEY[1,3], SARA A. KNOTT[2] AND DIRK-JAN DE KONING[1,3]
[1] *Genetics and Genomics, Roslin Institute, Midlothian, Edinburgh EH25 9PS, UK*
[2] *Institute of Evolutionary Biology, University of Edinburgh, Kings Buildings, Edinburgh EH9 3JT, UK*
[3] *Royal (Dick) School of Veterinary Studies, University of Edinburgh, Summerhall, Edinburgh EH9 1QH, UK*

(*Received 30 October 2007 and in revised form 21 April 2008*)

Summary

Dominance is an important source of variation in complex traits. Here, we have carried out the first thorough investigation of quantitative trait locus (QTL) detection using variance component (VC) models extended to incorporate both additive and dominant QTL effects. Simulation results showed that the empirical distribution of the test statistic when testing for dominant QTL effects did not behave in accordance with existing theoretical expectations and varied with pedigree structure. Extensive simulations were carried out to assess accuracy of estimates, type 1 error and statistical power in two-generation human-, poultry- and pig-type pedigrees each with 1900 progeny in small-, medium- and large-sized families, respectively. The distribution of the likelihood-ratio test statistic was heavily dependent on family structure, with empirical thresholds lower for human pedigrees. Power to detect QTL was high (0·84–1·0) in pig and poultry scenarios for dominance effects accounting for $>7\%$ of phenotypic variance but much lower (0·42) in human-type pedigrees. Maternal or common environment effects can be partially confounded with dominance and must be fitted in the QTL model. Including dominance in the QTL model did not affect power to detect additive QTL effects. Also, detection of spurious dominance QTL effects only occurred when maternal effects were not included in the QTL model. When dominance effects were present in the data but not in the analysis model, this resulted in spurious detection of additive QTL or inflated estimates of additive QTL effects. The study demonstrates that dominance can be included routinely in QTL analysis of general pedigrees; however, optimal power is dependent on selection of the appropriate thresholds for pedigree structure.

1. Introduction

Historically, dominance has often been ignored or treated as a nuisance parameter, for example in genetic evaluations of livestock and quantification of variance components (VCs). The importance of detection and quantification of dominance effects underlying complex traits, however, is underlined by an increasing body of evidence for dominant quantitative trait locus (QTL) with major effects on human disease and agricultural traits of economic importance. Duong *et al.* (2006) found eight completely dominant QTL associated with hypertension in congenic rat lines, and in agriculture, examples of dominant QTL include fertility and production traits in cattle (Cohen-Zinder *et al.*, 2005), chicken (Ikeobi *et al.*, 2002; Hocking, 2005), tomatoes (Semel *et al.*, 2006) and maize (Zhang *et al.*, 2006). Liu *et al.* (2007) performed a genome-wide scan on an F_2 Duroc $\times$ Pietrain cross and found 40 additive QTL and 31 QTL showing overdominance effects. Although definitions vary, overdominance is a phenomenon for which there is increasing evidence in plants as the underlying mechanism for heterosis (Xiao *et al.*, 1995; Frascaroli *et al.*, 2007). Lippman & Zamir (2007) review detection and characterization of heterosis, overdominance and pseudo-overdominance.

* Corresponding author. Genetics and Genomics, Roslin Institute, Midlothian, Edinburgh EH25 9PS, UK. Tel: +44 (0)131 527 4462. e-mail: suzanne.rowe@bbsrc.ac.uk

To date, the detection of these dominant QTL effects has predominantly involved model species or experimental crosses requiring inbred or genetically divergent populations. Reproductive constraints render these test crosses impractical for many agricultural species, while for human and natural populations they are unethical and/or untenable. In commercial livestock populations, it is often more relevant, practical and cost-effective to explore QTL segregating within a population, particularly if the objective is to facilitate selection within that population. There is evidence to suggest that much of the variation found between lines is segregating within lines (De Koning *et al*., 2004) and, furthermore, most evolutionarily important variation appears to occur within lines (Erickson *et al*., 2004).

There is, therefore, an increasing need for QTL methodology to routinely account for genetic interactions such as dominance within any population structure. Independently developed within human and livestock research, VC-based linkage methods (Fernando & Grossman, 1989; Goldgar, 1990; Amos, 1994; Grignola *et al*., 1997; Almasy & Blangero, 1998; Allison *et al*., 1999; George *et al*., 2000) have the advantage of simultaneously locating and estimating genetic effects within arbitrary pedigrees. Genetic parameters associated with the polygenic effect, and at specific loci using marker and pedigree information, can be estimated simultaneously. The incorporation of many alleles or allelic effects and all relationships within a pedigree has been shown to increase the power to detect QTL over sib-based methods (Williams & Blangero, 1999; Sham *et al*., 2000; Kolbehdari *et al*., 2005; Rowe *et al*., 2006). Furthermore, linkage disequilibrium and haplotype information can be incorporated to provide greater accuracy (Meuwissen *et al*., 2002; Lee & Van der Werf, 2006). Most importantly, there is the potential for flexibility to incorporate random effects and their interactions, for example, dominance, epistasis and maternal effects, limited only by the size and structure of the experimental population.

Although undeniably an important source of variation, non-additive effects are notoriously difficult to estimate due to confounding with other sources such as common maternal environment (Gengler *et al*., 1997; Misztal, 1997). Computational complexity combined with the more generic problems of setting appropriate thresholds to account for multiple testing and lack of suitable data have inhibited the extension of VC methodology to incorporate interactions such as dominance and epistasis. Although extensions to VC QTL linkage models to incorporate dominance are widely discussed (Sham *et al*., 2000; Diao & Lin, 2005), they have rarely been implemented, indicating a need for further investigation before the full potential of these methods can be unleashed.

In the present study, extensive simulations have been used to explore the power and potential for partitioning QTL effects into additive and dominant components using VC methods for linkage analysis. Varying full-sib and half-sib population structures have been used to evaluate accuracy and the power to detect additive and dominant genetic effects in pedigrees that are representative of commercial livestock and human scenarios.

2. Materials and methods

(i) *Statistical genetic models for VC analysis*

Population-wide linkage equilibrium between QTL and marker alleles was assumed for all analyses. Following the two-step approach described by George *et al*. (2000), for each putative QTL position, marker information was used to estimate identity-by-descent (IBD) coefficients for all relationships in the pedigree. In the second step, different QTL models were fitted for given genome locations using the following models:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{e} \quad \text{(null or polygenic)} \tag{1}$$

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{Za} + \mathbf{e} \quad \text{(null + additive QTL)} \tag{2}$$

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{Za} + \mathbf{Zd} + \mathbf{e} \quad \text{(null + additive QTL + dominance QTL)} \tag{3}$$

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{Wm} + \mathbf{Za} + \mathbf{Zd} + \mathbf{e} \quad \text{(null + maternal + additive QTL + dominance QTL)} \tag{4}$$

where $\mathbf{y}$ is a vector of phenotypic observations, $\boldsymbol{\beta}$ is a vector of fixed effects, $\mathbf{u}$, $\mathbf{a}$, $\mathbf{d}$, $\mathbf{m}$ and $\mathbf{e}$ are vectors of random additive polygenic effects, additive and dominance QTL effects at the putative QTL position, non-genetic maternal effects and residuals, respectively, and $\mathbf{X}$, $\mathbf{Z}$ and $\mathbf{W}$ are incidence matrices relating to records of fixed and random genetic and maternal effects, respectively.

Variances for polygenic and QTL effects are distributed as follows: $\mathrm{Var}(\mathbf{a}) = \mathbf{G}\sigma^2_{\mathrm{q}}$, $\mathrm{Var}(\mathbf{d}) = \mathbf{D}\sigma^2_{\mathrm{d}}$, $\mathrm{var}(\mathbf{e}) = \mathbf{I}\sigma^2_{\mathrm{e}}$ and $\mathrm{var}(\mathbf{u}) = \mathbf{A}\sigma^2_{\mathrm{a}}$. For the non-genetic maternal effect: $\mathrm{Var}(\mathbf{m}) = \mathbf{I}\sigma^2_{\mathrm{m}}$.

https://doi.org/10.1017/S0016672308009336 Published online by Cambridge University Press

Here, **A** is the standard additive relationship matrix based on pedigree data only, **G** is the QTL additive genetic relationship matrix based on marker information and **D** is the QTL dominance genetic relationship matrix representing the probability that two individuals have the same pair of alleles in common based on marker information. VCs for each model were estimated using REML (Patterson & Thompson, 1971) implemented in the ASReml package (Gilmour *et al*., 1995).

(ii) *Calculating the relationship matrices A, G and D needed for the mixed model analysis*

The relationship matrices **G** and **D** for a given QTL position are calculated from the gametic IBD matrix as outlined by Liu *et al*. (2002). The gametic IBD matrix is a $2n \times 2n$ matrix containing the probability of identity of descent between either of the two gametes of an individual with the gametes of the remaining individuals in the pedigree. In contrast to George *et al*. (2000), who used a Monte Carlo method, the gametic IBD matrix was estimated with the recursive method of Pong-Wong *et al*. (2001), which uses the two first-available fully informative or phase-known flanking markers. The **G** and **D** matrices are conditional on flanking marker information and therefore unique for each position evaluated for a QTL; hence, the calculation of **G** and **D** requires the prior calculation of the gametic IBD matrix conditional on linked marker information at the position of the putative QTL. Here, the matrices were calculated every 5 cM. In order to estimate the VCs for the different models, ASReml requires the inverse of the relationship matrices **A**, **G** and **D**. The version of ASReml used calculates the inverse of the **A** matrix directly from pedigree data, but the inverse for **G** and **D** were calculated from the gametic matrix, inverted using a separate routine and then passed into ASReml.

(iii) *Test statistic*

A test statistic for a given location was obtained by comparing the likelihood of the full *vs*. the null model. The log likelihood-ratio test (LRT) statistic was calculated as twice the difference between the log likelihood of the full model and the reduced model. Power was estimated both empirically using thresholds derived from 1000 chromosome-wise replicates, and using tabulated values assuming that the test statistic is χ^2 distributed with degrees of freedom equal to the number of extra parameters estimated in the full model compared with the reduced model. This is conservative for a test at a single location in the genome as the test statistic under the null hypothesis is likely to be distributed as a complex mixture of distributions (Self & Liang, 1987; Stram & Lee, 1994; Allison *et al*., 1999; Visscher, 2006). For QTL mapping, it has been suggested that the most straightforward method of achieving the critical null value is to halve the P value obtained for χ^2_k, where k is the number of extra VCs in the full model (Visscher, 2006). In practice, this mixture of distributions χ^2_{0-1} is equal to using the 10% critical threshold for a 5% type 1 error rate.

Three tests were carried out: (i) additive QTL (2) *vs*. null (1) to test significance of the QTL VC under a purely additive model (denoted 1v0); (ii) additive QTL + dominance QTL (3) *vs*. null (1) to test significance of QTL VCs under a model including additive and dominance effects (denoted 2v0); and (iii) additive QTL + dominance QTL (3) *vs*. additive QTL (2) to test the significance of the dominance VC (denoted 2v1). To estimate the effect of common environment, the model including additive and dominance QTL effects was further extended to incorporate a random dam effect (4), representing a maternal or full-sib family effect.

Table 1. *Population parameters for simulated pedigrees*

	Sires	Dams per sire	No. of half-sibs per sire	No. of full-sibs per dam
Chicken	19	5	100	20
Pig	10	19	190	10
Human	633	1	–	3

(iv) *Population structure*

The method was implemented in three simulated populations, representative of poultry, pig and human pedigrees (Table 1). The parental generation was simulated by random sampling without replacement from an unrelated base population. Under each scenario, random mating of parents was simulated to obtain a second generation of 1900 progeny.

A 20 cM chromosome was simulated with five markers spaced at 5 cM intervals and a bi-allelic QTL between the second and third markers at 7·5 cM. To simulate polygenic variance, ten unlinked additive effects of 0·2 were simulated each with an allele frequency of 0·5 following Alfonso & Haley (1998). The phenotypes generated under a polygenic model were normally distributed, indicating that these unlinked QTL were sufficient to provide a reasonably structured polygenic variance.

Dominance effects were simulated ranging from partial to overdominance over a range of additive effects. These are summarized in Table 2. The variances of the QTL additive (a) and dominance (d) effects

https://doi.org/10.1017/S0016672308009336 Published online by Cambridge University Press

Table 2. *Summary of scenarios*

Scenario	QTL effect		Heritability (h^2)		
	Additive (a)	Dominant (d)	Additive QTL $\sum^2_q/\sigma^2_P$	Dominant QTL σ^2_d/σ^2_P	Total[a] $\sigma^2_{a+}\sigma^2_{q+}\sigma^2_d/\sigma^2_P$
1	0·00	0·00	0·00	0·00	0·11
2	0·10	0·00	0·00	0·00	0·12
3	0·20	0·00	0·01	0·00	0·13
4	0·30	0·00	0·03	0·00	0·14
5	0·00	0·50	0·00	0·04	0·15
6	0·40	0·00	0·04	0·00	0·16
7	0·00	0·60	0·00	0·05	0·16
8	0·00	0·70	0·00	0·07	0·18
9	0·50	0·00	0·07	0·00	0·18
10	0·00	0·80	0·00	0·09	0·19
11	0·50	0·50	0·07	0·03	0·21
12	0·60	0·60	0·09	0·05	0·24
13	0·80	0·40	0·16	0·02	0·27
14	0·70	0·70	0·12	0·06	0·27
15	0·80	0·50	0·15	0·03	0·28
16	0·80	0·60	0·15	0·04	0·29
17	0·80	0·70	0·15	0·06	0·30
18	0·80	0·80	0·15	0·07	0·31

σ^2_P is the phenotypic variance.
[a] Total heritability includes polygenic variance (σ^2_a) of 0·2, residual variance 0·75, expected additive QTL variance (σ^2_q) estimated by ($a^2/2$) and expected dominant QTL variance (σ^2_d) estimated by ($d^2/4$).

were calculated as 0·5 a^2 and 0·25d^2, respectively, because allele frequencies were set to 0·5.

For each individual, a residual effect was sampled from a normal distribution with mean 0 and a variance of 1·5. As the error variance was constant, phenotypic variance and overall heritability increased with genetic effects. In the base scenario with no QTL simulated, polygenic heritability was 0·11. Total heritability (polygenic and QTL) ranged from 0·1 to 0·31 with dominance QTL effects ranging from 0 to 9% of the phenotypic variance. For each scenario where QTL were simulated, 100 replicates were analysed and the test statistics described above calculated at 2, 7, 12 and 17 cM.

(v) *Maternal effect*

Common environment or direct maternal effects are often, at least partially, confounded with dominance. To explore the effect on detection of dominance QTL, random maternal effects were simulated in the pig population. A maternal effect was simulated by sampling for each full-sib family from a normal distribution with variance of 0·1 and assigning this value to each full-sib offspring. A residual effect was sampled from a normal distribution with mean 0 and a variance of 0·75. The implication of potential maternal effects was evaluated in three different ways: a maternal effect was simulated with a range of dominance QTL effects. Firstly, the maternal effect was not fitted in the model to test for spurious detection of dominance. Secondly, a maternal effect was included in the linear model to test whether the model correctly accounts for the maternal variance. Finally, no maternal effect was simulated but a maternal component was included in the linear model to test whether the dominance variance was correctly identified or incorrectly estimated as a maternal effect.

(vi) *Null distribution*

Chromosome-wise type 1 error rates were determined empirically for all three population structures. A total of 1000 replicates were used to explore 1, 5 and 10% thresholds under the null scenario (both additive and dominance QTL effects set to zero). Point-wise test statistics were determined with 1000 replicates at the QTL position. Empirical distributions for point-wise tests were compared with tabulated values for χ^2_k, where k is equal to the number of extra VCs estimated using P values for 1, 5 and 10%, and under the assumption that halving the P value accounts for a mixture of distributions using P values of 2, 10 and 20% for 1, 5 and 10% thresholds. A 5% threshold was determined for empirical power calculations and comparisons based on the analysis of 1000 replicates.

The empirical distribution of the LRT under the null scenario appeared to vary across pedigrees, in particular, differing between human and livestock. This could be due to the difference in number of full

https://doi.org/10.1017/S0016672308009336 Published online by Cambridge University Press

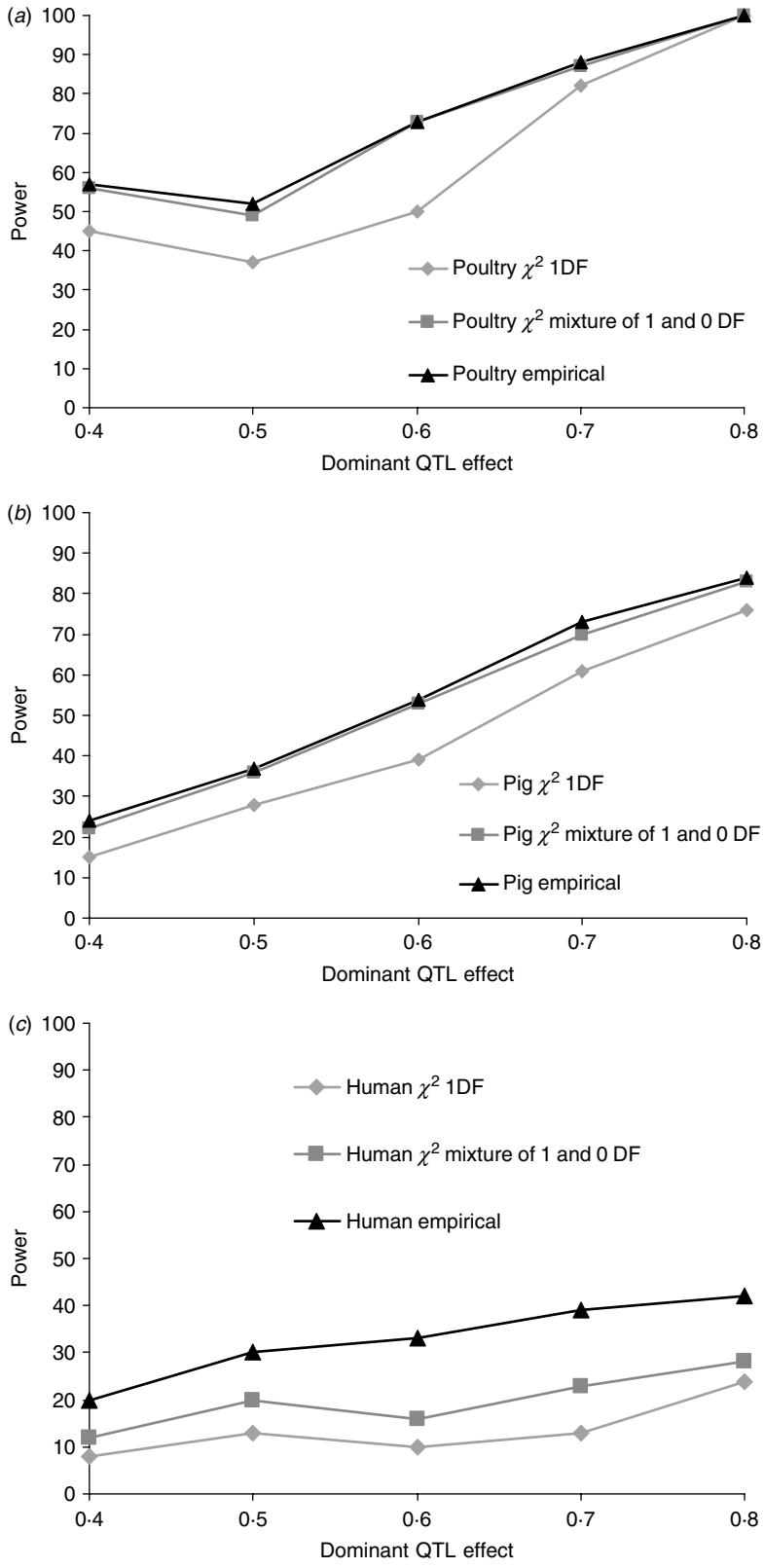

Fig. 1. Proportion of replicates where test for dominance (2v1) is significant ($P<0{\cdot}05$) when comparing the full model and the additive model. A total of 100 chromosome-wise replicates in (top to bottom) (*a*) poultry, (*b*) pig and (*c*) human pedigrees under partial to complete dominance. Simulated additive effect fixed at 0·8 comparing tabulated 5% χ^2_1, χ^2_{1-0} thresholds and 5% empirical threshold. Mixture threshold is estimated by using tabulated 10% χ^2_1 threshold.

https://doi.org/10.1017/S0016672308009336 Published online by Cambridge University Press

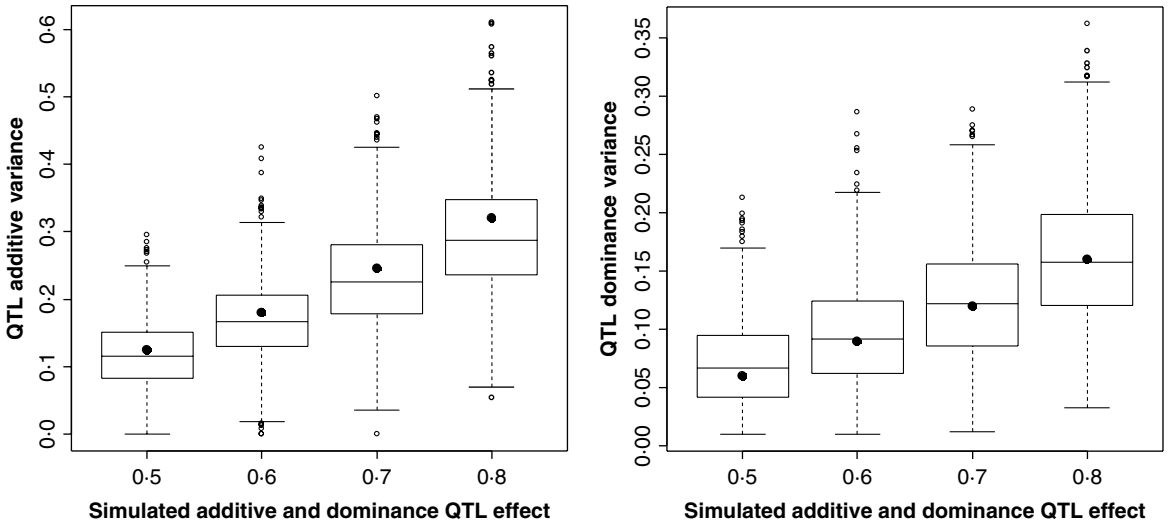

Fig. 2. Estimates of VCs from simulated poultry data. Box plots showing the range of variance estimates. Full dominance is simulated. Variance estimates for single marker position (for 1000 replicates of each scenario) for additive and dominant QTL effects. The black circles indicate the expected VCs. All replicates were significant for a QTL when testing under the full model (additive and dominance QTL effects *vs.* null).

sibs per family (three for humans compared with 10 or 20 for livestock) or the lack of half-sib relationships in the human pedigree. To explore this, chromosome-wise null LRT distributions were determined for five additional pedigrees. The number of offspring remained constant at 1900 but pedigrees with 1 or 2 dams per sire and 3, 5 or 10 offspring per dam were compared with pig and poultry pedigrees to explore the effects of family structure on the distribution of the null test statistic.

3. Results

All results shown are based on 5% empirical thresholds from 1000 chromosome-wise replicates. Full results for all populations and scenarios can be found in Supplementary Table S1 for power to detect additive and dominant QTL effects, S2 for power in pig populations with maternal or common environment effects and S3 for estimates of VCs.

(i) *Power to detect dominance effects*

Figures 1*a*, *b* and *c* give the proportion of replicates detecting dominance using 5% empirical and tabulated thresholds for χ^2_1 and a mixture χ^2_{1-0}. Testing for dominance involved comparing the additive QTL model with the full model incorporating both an additive QTL and a dominant QTL effect (denoted 2v1). For the pig and poultry pedigrees, power under empirical and the χ^2_{1-0} mixture of distributions was in close agreement. For human pedigrees, both χ^2_1 and χ^2_{1-0} were conservative when compared with empirical thresholds. Power under the empirical 5% threshold was ~100% in the poultry scenario, ~84% for the pig scenario and ~42% for humans when the QTL dominance variance was about 7% of the phenotypic variance (dominant effect = 0·8, i.e. complete dominance). Under χ^2_1 thresholds this dropped to ~95, ~75 and ~25. Power to detect dominance was greater for all pedigrees using empirical thresholds. Although the ranking did not change when using tabulated values, power to detect dominance, particularly in humans, was much lower and differences between models greater. When comparing the full model with the null model (denoted 2v0), all replicates detected a QTL for the pig and poultry scenarios, and 96% (84% under tabulated thresholds) of replicates detected a QTL for the human scenario (Supplementary Table S1).

Figure 2 shows the estimates for the additive and dominance QTL VCs from the comparison of the full model with the null model (2v0). In all replicates, a QTL was detected at the 5% significance level. Scenarios shown were for complete dominance with effects ranging from 0·5 to 0·8 (also given in Supplementary Table S2). These show that although estimates are wide-ranging, they appear to accurately estimate the mean.

(ii) *Overdominant, spurious additive and dominant QTL effects*

Figure 3 shows power to detect simulated additive effects ranging from 0·1 to 0·5, or 1–7% of phenotypic variance. Power reached 90% when the additive variance amounted to >4% of the phenotypic variance. In this case, no dominance effect was simulated and there was little spurious dominance detected. Furthermore, power to detect an additive effect was

https://doi.org/10.1017/S0016672308009336 Published online by Cambridge University Press

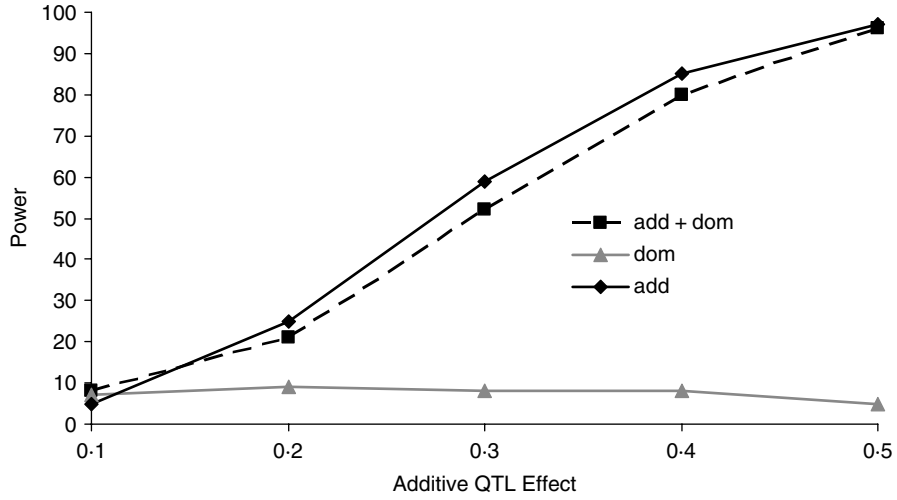

Fig. 3. Percentage of replicates detecting additive QTL effects ($P<0{\cdot}05$) using the full model (add + dom) and the additive model (add) and testing the difference between the two (dom) in a simulated pig population. A dominance effect of zero is simulated.

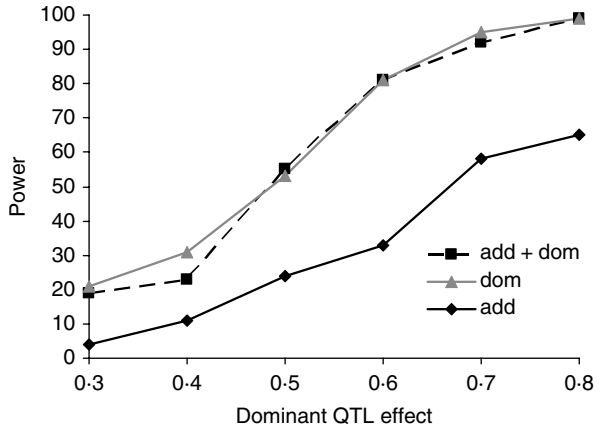

Fig. 4. Overdominance: percentage of replicates detecting overdominant QTL effects ($P<0{\cdot}05$) using the full model (add + dom) and the additive model (add) and testing the difference between the two (dom) in a simulated poultry population over a range of dominant QTL effects when an additive effect of zero is simulated.

similar whether or not the extra dominance component was included in the analysis. This shows that a routine scan including dominance would not result in too great a loss of power even in the absence of any dominant effects. Although results are shown only for the pig population, the same pattern was seen for poultry and human scenarios. Figure 4, however, shows that spurious additive QTL effects were found when dominant QTL effects were not fitted. With dominant QTL effects ranging from 0·5 to 0·8 and simulated additive effects of zero, i.e. overdominance, there is both spurious detection of an additive QTL effect if dominance was not included in the model and inflated estimates of additive variance (Table 3).

Table 3. *Estimates of variance due to additive QTL and additive and dominant QTL effects under overdominance when the additive QTL effect of zero is simulated*

Expectation		Additive QTL *vs.* null model (1v0)	Additive + dominant QTL *vs.* null model (2v0)	
Add	Dom	Add	Add	Dom
0	0·062	0·020	0·008	0·067
0	0·090	0·036	0·009	0·098
0	0·122	0·037	0·008	0·122
0	0·160	0·067	0·010	0·161

Variances are mean estimates at the highest chromosome-wise test statistic across 100 iterations in a simulated pig population.

(iii) *Maternal effects and dominance in the pig scenario*

Figure 5 clearly shows that a simulated maternal effect can masquerade as a dominant QTL effect. In the most extreme case when a dominance effect of 0 was simulated, ignoring the common environment resulted in a type 1 error of 67%. When a maternal effect was fitted in the model, maternal effects and dominance appear to be successfully separated with little or no loss of power (Table 4). If a maternal effect was fitted when not present, there was little loss of power (for results see supplementary Tables S2 and S3), indicating that a maternal component fitted in the absence of a maternal effect should not prevent detection of variance due to dominance.

(iv) *Null distribution*

Table 5 shows that the point-wise test statistic was conservative for all models when compared with

https://doi.org/10.1017/S0016672308009336 Published online by Cambridge University Press

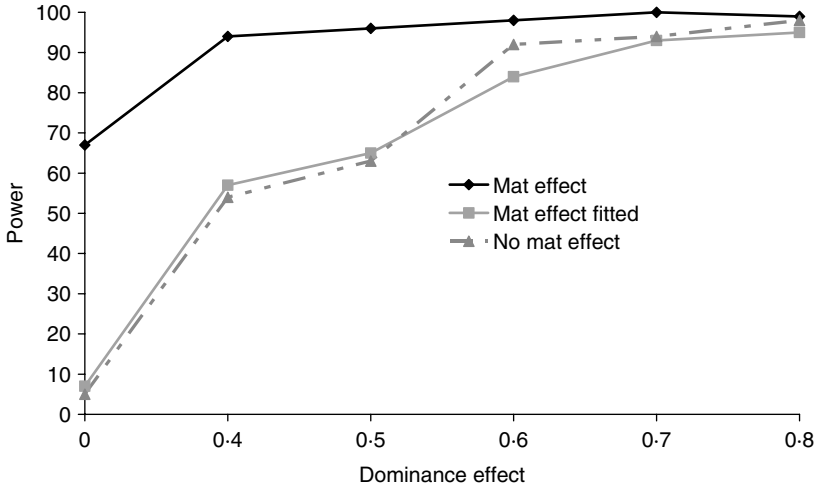

Fig. 5. Effects of simulating and/or fitting direct maternal effects on proportion of replicates where test for dominance (2v1) is significant ($P<0{\cdot}05$) when comparing the full model and the additive model. A total of 100 chromosome-wise replicates in the pig population under partial to complete dominance (additive QTL effect fixed at 0·8). 'No mat effect', no maternal effect simulated or fitted; 'mat effect', maternal variance of 0·1 simulated but not fitted; 'mat effect fitted', maternal variance of 0·1 simulated and fitted.

Table 4. *Variance estimates for dominant QTL effect of 0·8 and maternal effects*

Maternal effect simulated	Maternal effect fitted	Dom variance (0·16)	Maternal variance (0·10)
N	N	0·17	–
Y	N	0·25	–
Y	Y	0·20	0·09
N	Y	0·17	0·02

Values in parentheses are simulated or expected variances. Variances are mean estimates at the highest chromosome-wise test statistic across 100 iterations in a simulated pig population.

tabulated χ^2 values. This was also true if the mixture of distributions was taken into account by assuming a P value of 0·1 to derive a 5% critical threshold. Table 5 shows that for the test for dominance the equivalent P value under the χ^2_1 distribution to a 5% empirical threshold was actually 0·21 for pig and poultry and 0·35 for human pedigrees. The test for additive QTL effects although still conservative was much closer to tabulated values, particularly 10% thresholds with equivalent P value under the χ^2_1 distributions of 0·14, 0·15 and 0·11 for poultry, pig and human pedigrees, respectively.

Chromosome-wise type 1 error rates were close to tabulated thresholds for the additive model, for all three simulated pedigrees. Tabulated type 1 error rates for the full model and for dominance, however, remained conservative. None of these tabulated values were corrected for multiple testing.

Figure 6 compares distributions of the empirical test statistic with χ^2 distributions. In particular, for the test for dominance the empirically derived null statistic appears to vary according to pedigree structure, i.e. is lower in humans. This is apparent to a lesser extent in the model testing for both additive and dominance and somewhat reversed in the additive test.

Figure 7/Table 6 compares distributions of the test statistic under the null hypothesis of no QTL for eight pedigree structures. These vary from human families with three full-sibs to more complex structures such as poultry with 20 full-sibs and 200 half-sibs. It is apparent that the three human pedigrees, i.e. single dam families, have very similar distributions seemingly regardless of the number of full-sibs. Similarly, the three pedigrees with two dam families have similar distributions to each other but clearly different from those of the human or larger livestock families. The pig and poultry distributions are similar to each other although the pig distribution appears slightly more conservative. The χ^2_{1-0} distribution appears to be comparable with the pig and poultry distribution, although again as these were chromosome-wise tests that are uncorrected for multiple testing, they cannot be directly compared with nominal values.

4. Discussion

This study provides a comprehensive evaluation of the performance of VC analysis over a range of dominant QTL effects. The method was successfully used to estimate and apportion variances due to polygenic, additive, dominant and non-genetic effects. Power >95% was achieved when testing for dominant QTL

https://doi.org/10.1017/S0016672308009336 Published online by Cambridge University Press

Table 5. *Empirical 5% thresholds for LRT test statistic* (*and the corresponding P value under* χ^2 *distribution*). *A total of 1000 replicates simulated for single point-wise and multiple chromosome-wise testing under the null scenario of no QTL effects*

	Point-wise						Chromosome-wise					
	Poultry		Pig		Human		Poultry		Pig		Human	
2v0	3·18	(0·20)	3·30	(0·19)	3·00	(0·22)	4·86	(0·09)	4·94	(0·08)	4·58	(0·10)
2v1	1·60	(0·21)	1·58	(0·21)	0·86	(0·35)	2·62	(0·11)	2·70	(0·10)	1·48	(0·22)
1v0	2·20	(0·14)	2·08	(0·15)	2·62	(0·11)	3·78	(0·05)	3·46	(0·06)	3·78	(0·05)

The 2v0 testing model including an additive QTL and a dominant QTL effect *vs.* a null model with *P* value in parentheses assuming χ^2_2; the 2v1 testing model including an additive and a dominant QTL effect *vs.* an additive QTL model with *P* value in parentheses assuming χ^2_1; and the 1v0 testing model including an additive QTL *vs.* a null model with *P* value in parentheses assuming χ^2_1.

effects accounting for 7% of the total variance for a simulated poultry population. Power to detect additive QTL was also high (~97% for an additive effect explaining 7% of phenotypic variance. Although the upper range of the simulated QTL effects is high, these values are plausible in terms of the published literature.

Simulation results showed that, unlike the test for additive QTL effects, the empirical distribution of the test statistic when testing for dominant QTL effects did not behave in accordance with existing theoretical expectations. Theoretically, the asymptotic distribution of the LRT is a mixture of χ^2 with different degrees of freedom when testing VCs under the null hypothesis that they are zero (Self & Liang, 1987; Stram & Lee, 1994). For example, with one extra VC, the null distribution is a mixture of $\frac{1}{2}\chi_0$ (i.e. variance is zero) and $\frac{1}{2}\chi_1$ (i.e. variance is non-zero). With a model including two VCs, such as additive and dominant QTL effects, the expectation of the distribution would be a mixture of $\frac{1}{4}\chi_0$ (both VCs are zero), $\frac{1}{2}\chi_1$ (one is non-zero) and $\frac{1}{4}\chi_2$ (both are non-zero). Visscher (2006) provides a thorough review.

Problems with incorrect assumptions about the distribution include inflated type II errors leading to reduced power. Extending the linear model to include a dominance component resulted in a conservative test when imposing a χ^2_1 distribution for the LRT statistic. The test remained conservative even when thresholds were halved under the assumption of a mixture of distributions. One explanation might be that additive and dominant QTL effects are not entirely independent. Furthermore, the null distribution for the dominance test varies with family structure, in particular, with the number of dam families per sire. Distributions of the test statistic in Figure 7/Table 6 appear to group by number of dam families, with human-type pedigrees with a single dam per sire most conservative, regardless of family size. The number of full-sibs within dam families did not appear to affect the distribution. It is possible that the lack of half-sib families might result in confounding of additive and dominance effects at the QTL. Theoretically, if both components need to be estimated within dam, lack of information might lead to a higher probability of VCs being zero.

Results showed that power is also affected by population structure. Power to detect dominance at the QTL was similar for pig and poultry populations but much lower for humans.

This was unsurprising as the human population consisted of many small families with low numbers of full-sibs, making it difficult to detect dominance. Increased power might be achieved in human studies from a pedigree with more generations providing information from relationships such as grandparents and cousins, but this needs to be explored further.

It is anticipated that further correction for multiple testing for large linkage groups or genome-wide testing would be necessary. The distribution, however, of H_0 when testing for multiple linked positions is unresolved and authors have used different approximations (Xu & Atchley, 1995; Pratt *et al*., 2000; Piepho, 2001; Nagamine *et al*., 2004). Procedures such as permutation and bootstrapping enable the setting of empirical thresholds and circumvent problems associated with failure of distributional assumptions and independence of multiple tests (Churchill & Doerge, 1994; Visscher *et al*., 1996), although computational complexity can restrict their use within the VC framework.

The method described by Piepho relies on the gradient of change in likelihood. However, this method still assumes that the test statistic for a single test follows a standard χ^2 distribution under the null hypothesis and therefore does not address the issue of mixture distributions that is apparent for these types of analyses. It is difficult to ascertain whether the method is appropriate here, when the test statistic follows a mixture of distributions and likelihoods under the null scenario are very flat.

There is strong evidence to suggest that a common environment effect should be routinely evaluated in

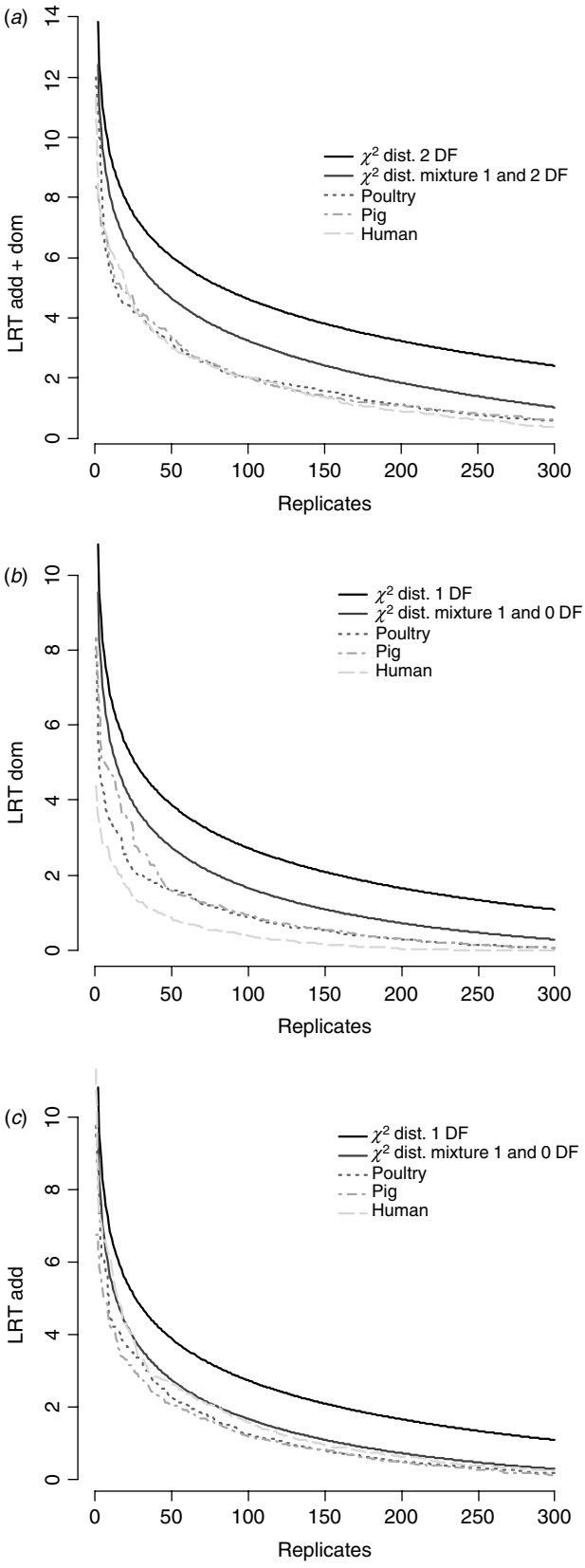

Fig. 6. Distribution of empirical point-wise test statistic in pig, poultry and human pedigrees for (from top to bottom) (*a*) additive and dominance effects, (*b*) dominance effects and (*c*) additive effects compared with χ^2_1 and χ^2_{1-0} distributions. The top 300 values of 1000 replicates are displayed.

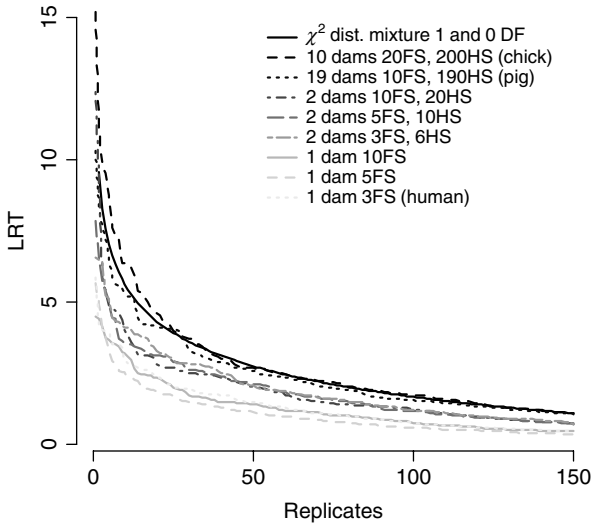

Fig. 7. Comparison of distribution of empirical chromosome-wise test statistic for dominance effects under null hypothesis of no QTL in pedigrees with varying full-sib (FS) and half-sib (HS) structures. χ^2_{1-0} is also plotted for comparison. All pedigrees have 1900 total offspring. The top 150 values of 1000 replicates are displayed for clarity.

Table 6. *Empirical 5% thresholds for LRT test statistic when testing for dominance and corresponding P value under* χ^2_1 *distribution. A total of 1000 replicates simulated for chromosome-wise testing under the null scenario of no QTL effects*

Pedigree	Sires	Dams per sire	Progeny per dam	LRT 5% empirical threshold
1 (human)	633	1	3	1·46
2	380	1	5	1·14
3	190	1	10	1·38
4	317	2	3	2·1
5	190	2	5	2·02
6	195	2	10	2·06
7 (chick)	19	5	20	2·62
8 (pig)	10	19	10	2·70

all VC QTL models as, if unaccounted for, most of the variation due to common environment masquerades as dominance. We have shown that the presence of common environmental effects has little effect on false-negative rates but a potentially huge impact on false-positive rates.

We demonstrated that incorporating a dominance effect in a genome scan has very limited detrimental effect on the power to detect purely additive QTL. Detection of spurious dominance was also rare, suggesting that dominance could be routinely included in genome scans. We have also shown by simulation that, if not fitted in the analysis model, dominance may be detected as spurious additive effects or inflated

https://doi.org/10.1017/S0016672308009336 Published online by Cambridge University Press

estimates of additive genetic variance. This suggests that dominant QTL effects can be detected as additive QTL when additive-only models are used; see also Misztal *et al*. (1998) and Pante *et al*. (2002) for similar effects with polygenic dominance. This has important implications for predicting response to selection as the success of any selection programme is dependent on correctly identifying the mode of inheritance and proportion of variance explained by the QTL. For example, Hayes & Miller (2000) have shown that including dominance effects in mate selection can be a powerful tool for exploiting previously untapped genetic variation, whereas Dekkers & Chakraborty (2004) have discussed maximization of crossbred performance by incorporating information from overdominant QTL.

A further confounding factor not studied here might be polygenic dominance. It is, however, unlikely to have affected our results as most of the information for polygenic dominance would have come from the covariance of full-sibs and should have been accounted for by the common environment effect. This might not be the case within other relationships in deeper, more complex, pedigrees, suggesting that the inclusion of a polygenic dominance effect may be valuable when examining such data structures.

5. Conclusions

VC methods were implemented to detect dominant QTL. Type 1 error rates and power were explored using extensive simulations. Results indicate that if the mixture of distributions is taken into account, nominal χ^2 thresholds were appropriate when testing for additive QTL but conservative when testing for dominant QTL in all pedigrees and particularly in the case of the populations with sires mated to only one or two dams. Ascertaining the correct null distribution is a difficult issue, but one that merits revisiting. Here, we have shown that although theoretically the tabulated χ^2 values are fairly robust, the expected probability of non-zero variances varies with population structure; thus there are instances when greater power is achieved by empirically deriving the correct distribution of the test statistic. Power to detect dominant QTL effects was high in livestock pedigrees with little spurious dominance and could be successfully routinely employed under the proviso that common environment or direct maternal effects are accounted for.

Effects of extra generations or extended pedigrees are yet to be explored but may provide greater power for structures with few dam families. Future aims are to extend the linear model to include other non-additive effects such as epistasis and to simulate behaviour of the test statistic under varying modes of inheritance, size and structure of pedigree.

We thank the reviewers for their helpful comments. We also thank the Biotechnology and Biological Sciences Research Council, the Research Councils UK, Genesis Faraday and Cobb Vantress for funding.

References

Alfonso, L. and Haley, C. S. (1998). Power of different F-2 schemes for QTL detection in livestock. *Animal Science* **66**, 1–8.

Allison, D. B., Neale, M. C., Zannolli, R., Schork, N. J., Amos, C. I. & Blangero, J. (1999). Testing the robustness of the likelihood-ratio test in a variance-component quantitative-trait loci-mapping procedure. *American Journal of Human Genetics* **65**, 531–544.

Almasy, L. & Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *American Journal of Human Genetics* **62**, 1198–1211.

Amos, C. I. (1994). Robust variance-components approach for assessing genetic linkage in pedigrees. *American Journal of Human Genetics* **54**, 535–543.

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

Cohen-Zinder, M., Seroussi, E., Larkin, D. M., Loor, J. J., Everts-Van Der Wind, A., Lee, J. H., Drackley, J. K., Band, M. R., Hernandez, A. G., Shani, M., Lewin, H. A., Weller, J. I. & Ron, M. (2005). Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Research* **15**, 936–944.

De Koning, D. J., Haley, C. S., Windsor, D., Hocking, P. M., Griffin, H., Morris, A., Vincent, J. & Burt, D. W. (2004). Segregation of QTL for production traits in commercial meat-type chickens. *Genetical Research* **83**, 211–220.

Dekkers, J. C. M. & Chakraborty, R. (2004). Optimizing purebred selection for crossbred performance using QTL with different degrees of dominance. *Genetics, Selection, Evolution* **36**, 297–324.

Diao, G. & Lin, D. Y. (2005). A powerful and robust method for mapping quantitative trait loci in general pedigrees. *American Journal of Human Genetics* **77**, 97–111.

Duong, C., Charron, S., Xiao, C., Hamet, P., Menard, A., Roy, J. & Deng, A. Y. (2006). Distinct quantitative trait loci for kidney, cardiac, and aortic mass dissociated from and associated with blood pressure in Dahl congenic rats. *Mammalian Genome* **17**, 1147–1161.

Erickson, D. L., Fenster, C. B., Stenoien, H. K. & Price, D. (2004). Quantitative trait locus analyses and the study of evolutionary process. *Molecular Ecology* **13**, 2505–2522.

Fernando, R. L. & Grossman, M. (1989). Marker assisted selection using best linear unbiased prediction. *Genetics, Selection, Evolution* **21**, 467–477.

Frascaroli, E., Cane, M. A., Landi, P., Pea, G., Gianfranceschi, L., Villa, M., Morgante, M. & Pe, M. E. (2007). Classical genetic and quantitative trait loci analyses of heterosis in a maize hybrid between two elite inbred lines. *Genetics* **176**, 625–644.

Gengler, N., Van Vleck, L. D., MacNeil, M. D., Misztal, I. & Pariacote, F. A. (1997). Influence of dominance relationships on the estimation of dominance variance with sire-dam subclass effects. *Journal of Animal Science* **75**, 2885–2891.

George, A. W., Visscher, P. M. & Haley, C. S. (2000). Mapping quantitative trait loci in complex pedigrees: a two-step variance component approach. *Genetics* **156**, 2081–2092.

Gilmour, A. R., Thompson, R. & Cullis, B. R. (1995). Average information REML, an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* **51**, 1440–1450.

Goldgar, D. E. (1990). Multipoint analysis of human quantitative genetic variation. *American Journal of Human Genetics* **47**, 957–967.

Grignola, F. E., Zhang, Q. & Hoeschele, I. (1997). Mapping linked quantitative trait loci via residual maximum likelihood. *Genetics, Selection, Evolution* **29**, 529–544.

Hayes, B. J. & Miller, S. P. (2000). Mate selection strategies to exploit across- and within-breed dominance variation. *Journal of Animal Breeding and Genetics* **117**, 347–359.

Hocking, P. M. (2005). Review of QTL mapping results in chickens. *World's Poultry Science Journal* **61**, 215–226.

Ikeobi, C. O. N., Woolliams, J. A., Morrice, D. R., Law, A., Windsor, D., Burt, D. W. & Hocking, P. M. (2002). Quantitative trait loci affecting fatness in the chicken. *Animal Genetics* **33**, 428–435.

Kolbehdari, D., Jansen, G. B., Schaeffer, L. R. & Allen, B. O. (2005). Power of QTL detection by either fixed or random models in half-sib designs. *Genetics, Selection, Evolution* **37**, 601–614.

Lee, S. H. & Van der Werf, J. H. (2006). Using dominance relationship coefficients based on linkage disequilibrium and linkage with a general complex pedigree to increase mapping resolution. *Genetics* **174**, 1009–1016.

Lippman, Z. B. & Zamir, D. (2007). Heterosis: revisiting the magic. *Trends in Genetics* **23**, 60–66.

Liu, Y., Jansen, G. B. & Lin, C. Y. (2002). The covariance between relatives conditional on genetic markers. *Genetics Selection Evolution* **34**, 657–678.

Liu, G., Jennen, D. G., Tholen, E., Juengst, H., Kleinwachter, T., Holker, M., Tesfaye, D., Un, G., Schreinemachers, H. J., Murani, E., Ponsuksili, S., Kim, J. J., Schellander, K. & Wimmers, K. (2007). A genome scan reveals QTL for growth, fatness, leanness and meat quality in a Duroc–Pietrain resource population. *Animal Genetics* **38**, 241–252.

Meuwissen, T. H. E., Karlsen, A., Lien, S., Olsaker, I. & Goddard, M. E. (2002). Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. *Genetics* **161**, 373–379.

Misztal, I. (1997). Estimation of variance components with large-scale dominance models. *Journal of Dairy Science* **80**, 965–974.

Misztal, I., Varona, L., Culbertson, M. S., Bertrand, J. K., Mabry, J. W., Lawlor, T. J., Van Tassell, C. P. & Gengler, N. (1998). Studies on the value of incorporating the effect of dominance in genetic evaluations of dairy cattle. *Biotechnologie, Agronomie, Société et Environnement* **2**, 227–233.

Nagamine, Y., Visscher, P. M. & Haley, C. S. (2004). QTL detection and allelic effects for growth and fat traits in outbred pig populations. *Genetics, Selection, Evolution* **36**, 83–96.

Pante, M. J. R., Gjerde, B., McMillan, I. & Misztal, I. (2002). Estimation of additive and dominance genetic variances for body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. *Aquaculture* **204**, 383–392.

Patterson, H. D. & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**, 545–554.

Piepho, H. P. (2001). A quick method for computing approximate thresholds for quantitative trait loci detection. *Genetics* **157**, 425–432.

Pong-Wong, R., George, A. W., Woolliams, J. A. & Haley, C. S. (2001). A simple and rapid method for calculating identity-by-descent matrices using multiple markers. *Genetics, Selection, Evolution* **33**, 453–471.

Pratt, S. C., Daly, M. J. & Kruglyak, L. (2000). Exact multipoint quantitative-trait linkage analysis in pedigrees by variance components. *American Journal of Human Genetics* **66**, 1153–1157.

Rowe, S. J., Windsor, D., Haley, C. S., Burt, D. W., Hocking, P. M., Griffin, H., Vincent, J. & De Koning, D. J. (2006). QTL analysis of body weight and conformation score in commercial broiler chickens using variance component and half-sib analyses. *Animal Genetics* **37**, 269–272.

Self, S. G. & Liang, K. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* **82**, 605–610.

Semel, Y., Nissenbaum, J., Menda, N., Zinder, M., Krieger, U., Issman, N., Pleban, T., Lippman, Z., Gur, A. & Zamir, D. (2006). Overdominant quantitative trait loci for yield and fitness in tomato. *Proceedings of the National Academy of Sciences of the USA* **103**, 12981–12986.

Sham, P. C., Cherny, S. S., Purcell, S. & Hewitt, J. K. (2000). Power of linkage versus association analysis of quantitative traits, by use of variance-components models, for sibship data. *American Journal of Human Genetics* **66**, 1616–1630.

Stram, D. O. & Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics* **50**, 1171–1177.

Visscher, P. M. (2006). A note on the asymptotic distribution of likelihood ratio tests to test variance components. *Twin Research and Human Genetics* **9**, 490–495.

Visscher, P. M., Thompson, R. & Haley, C. S. (1996). Confidence intervals in QTL mapping by bootstrapping. *Genetics* **143**, 1013–1020.

Williams, J. T. & Blangero, J. (1999). Comparison of variance components and sibpair-based approaches to quantitative trait linkage analysis in unselected samples. *Genetic Epidemiology* **16**, 113–134.

Xiao, J., Li, J., Yuan, L. & Tanksley, S. D. (1995). Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. *Genetics* **140**, 745–754.

Xu, S. & Atchley, W. R. (1995). A random model approach to interval mapping of quantitative trait loci. *Genetics* **141**, 1189–1197.

Zhang, Z. M., Zhao, M. J., Ding, H. P., Rong, T. Z. & Pan, G. T. (2006). QTL mapping analysis of plant height and ear height of maize (*Zea mays* L.). *Genetika* **42**, 391–396.