Full pedigree quantitative trait locus analysis in commercial pigs using variance components

Full pedigree quantitative trait locus analysis in commercial pigs using variance components[1]

D. J. de Koning*, R. Pong-Wong*, L. Varona†, G. J. Evans‡, E. Giuffra§[3], A. Sanchez#, G. Plastow‡, J. L. Noguera†, L. Andersson§, and C. S. Haley*

*Roslin Institute, Roslin, Midlothian EH25 9PS, U.K.; † Institut de Recerca i Tecnologia Agroalimentàries, Area de Producció Animal Centre UdL-IRTA, Lleida 25198, Spain; ‡Sygen International PLC, University of Cambridge, Department of Pathology, Cambridge CB2 1QP, U.K.; §Swedish University of Agricultural Sciences, Uppsala S-751 24, Sweden; and Universitat Autònoma de Barcelona, Bellaterra, 08193 Catalunya, Spain.

ABSTRACT: In commercial livestock populations, QTL detection methods often use existing half-sib family structures and ignore additional relationships within and between families. We reanalyzed the data from a large QTL confirmation experiment with 10 pig lines and 10 chromosome regions using identity-by-descent (IBD) scores and variance component analyses. The IBD scores were obtained using a Monte Carlo Markov Chain method, as implemented in the LOKI software, and were used to model a putative QTL in a mixed animal model. The analyses revealed 61 QTL at a nominal 5% level (out of 650 tests). Twenty-seven QTL mapped to areas where QTL have been reported, and eight of these exceeded the threshold to claim confirmed linkage ($P < 0.01$). Forty-two of the putative QTL were detected previously using half-sib analyses, whereas 46 QTL previously identified by half-sib analyses could not be confirmed using the variance component approach. Some of the differences could be traced back to the underlying assumptions between the two methods. Using a deterministic approach to estimate IBD scores on a subset of the data gave very similar results to LOKI. We have demonstrated the feasibility of applying variance component QTL analysis to a large amount of data, equivalent to a genome scan. In many situations, the deterministic IBD approach offers a fast alternative to LOKI.

Key Words: Best Linear Unbiased Prediction, Genomes, Least Squares, Pigs, Quantitative Trait Loci, Variance Components

 J. Anim. Sci. 2003. 81:2155–2163

Introduction

Although methodology to detect or evaluate QTL in arbitrary pedigrees was proposed as early as 1989 by Fernando and Grossman, most QTL detection studies in livestock have been carried out by experimental F_2 and backcross designs (poultry, pigs) or by using existing half-sib family structures (cattle). For successful implementation of marker-assisted selection (**MAS**), segregation of QTL must be confirmed within commercial populations. The method proposed by Fernando and Grossman (1989) is based on a variance components (**VC**) model, where both the allelic QTL effects and the polygenic components are assumed to be normally distributed. The covariance between individuals for a putative QTL is modeled by the probabilities that they share alleles identity by descent (**IBD**), based on linked marker genotypes. George et al. (2000) described a two-step VC approach for arbitrary pedigrees where IBD scores are estimated using a Monte Carlo Markov Chain (**MCMC**) approach (Heath, 1997) and used to detect QTL in a mixed-inheritance model using ASREML (Gilmour et al., 1998). In an attempt to avoid the computational demand and potential convergence problems of the MCMC methods, Pong-Wong et al.

[1]The authors gratefully acknowledge the contributions of the following organizations: Sygen Int. PLC (U.K.), Quality Genetics (Sweden), Institut de Recerca i Tecnologia Agroalimentàries (Spain), and Cooperativa Agricola y Ganadera de Lleida (Spain). This work was partly funded by European Community contract No. BIO4-CT97-2243. Microsatellite primers were supplied by the U.S. Pig Genome Coordinator. D. J. de Koning is financially supported by the Dept. of Environment, Food, and Rural Affairs of the British government through LINK project LK0625. C. S. Haley acknowledges additional support from the BBSRC. G. Davalos (Barcelona), J. Folch (Barcelona), K. Harvey (Sygen), S. Johansson (Uppsala), E. Ramells (COPAGA), and O. Vidal (Barcelona) are gratefully acknowledged for technical assistance.

[2]Correspondence—phone: +44-131-5274460; fax: +44-131-4400434; E-mail: dj.dekoning@bbsrc.ac.uk.

[3]Present address: Centro Ricerche Studi Agroalimentari FPTP-CERSA, LITA, 20090 Segrate, Italy.

Received January 13, 2003.

Accepted April 28, 2003.

Table 1. Overview of populations and their abbreviations

Name	Description
H_1	Hampshire line from Sygen International PLC, Abingdon, U.K.
H_2	Hampshire line from Quality Genetics, Kävlinge, Sweden
L_1	Landrace line from Quality Genetics, Kävlinge, Sweden
L_2	Landrace line from Institut de Recerca i Tecnologia Agroalimentàries, IRTA, Lleida, Spain
LW_1	Large White line from Sygen International PLC, Abingdon, U.K.
LW_2	Large White line from Cooperativa Agricola y Ganadera de Lleida, COPAGA, Lleida, Spain
LW_3	Large White line from Quality Genetics, Kävlinge, Sweden
M	Meishan line from Sygen International PLC, Abingdon, U.K.
P_1	Pietrain line from Sygen International PLC, Abingdon, U.K.
P_2	Pietrain line from Cooperativa Agricola y Ganadera de Lleida, COPAGA, Lleida, Spain

(2001) proposed a deterministic approach (**DET**) to estimate IBD scores, which combines the methods of Wang et al. (1995) and Knott and Haley (1998). Sørensen et al. (2002) subsequently showed a correlation of more than 0.95 between the IBD scores using either MCMC or DET when simulating microsatellite data.

Variance component methods have been promising when tested on simulated data, but there are few reports of their application to real data. In the present study, we tested the applicability of the VC methods to real data by reanalyzing a QTL confirmation experiment and compared this to results obtained by half-sib (**HS**) regression models. We also compared the performance of the MCMC (Heath, 1997) and DET (Pong-Wong et al., 2001) methods to estimate IBD scores with regard to their performance in QTL detection.

Material and Methods

Resource Populations

Data were made available through the European Commission-funded PigQTech confirmation experiment involving three countries, contributing a total of 10 commercial pig breeds (contract No. BIO4-CT97-2243). Animals were from Large White, Landrace, Hampshire, Pietrain, and Meishan synthetic lines that were supplied by PIC International (UK), Quality Genetics (Sweden), IRTA (Spain), and Copaga (Spain). All animals had data on growth and fatness, but the Spanish data also included carcass and meat quality data. Tables 1 and 2 contain an overview of the populations and the traits, whereas Evans et al. (2002; 2003) present details on the experimental design. When the experiment was designed, regions on chromosome (**SSC**) 2, 3, 4, 7, 8, 10, and 13 were chosen as candidate regions because they had published QTL for growth and/or fatness. Three regions on chromosome 1, 6, and 9 were selected as control regions because when the experiment started, no QTL were published for these regions. The Sygen populations were not typed for the candidate region on SSC2 but for an additional candidate region on SSC1q instead. The chromosomal regions are in Table 3. In each of the chromosomal regions (Table 3), two to three microsatellite markers were selected for each population, based on heterozygosity in the sires. Genotypes were available on 71 boars, their mates, and nearly 4,500 offspring for 10 chromosomal regions.

Quantitative Trait Locus Analyses

The data were initially analyzed under a HS approach following Knott et al. (1996) using the QTL express software available at http://qtl.cap.ed.ac.uk/ (Seaton et al., 2002), and as reported by Evans et al. (2002; 2003). These results will be used for comparisons in the present study, but detailed results can be found in Evans et al. (2003).

Quantitative trait locus analyses were performed within breed and company, resulting in 650 region × trait × population analyses. Variance component analyses were performed at 1 to 5 cM intervals along every candidate region. On average, four positions were evaluated for every candidate region. Following the two-step approach proposed by George et al. (2000), the IBD scores were estimated for all positions within the two-generation pedigree of every population. These IBD scores were subsequently used to model the covariance for a putative QTL in a random mixed model. The IBD scores were estimated using an adapted version of the QTL mapping software LOKI (Heath, 1997). This program uses a MCMC approach to obtain IBD scores in arbitrary pedigrees with missing marker data and unknown haplotypes. Thompson and Heath (1999) present a detailed description of the method, whereas George et al. (2000) give an overview of IBD estimation methods in arbitrary pedigrees. We did not evaluate convergence of the MCMC sampler, but instead, we used 10,000 iterations for every position, which is substantially more than the recommended figure of 10 times the number of animals in the pedigree.

For 15 population × region combinations, the estimation of IBD scores was repeated using the DET approach proposed by Pong-Wong et al. (2001). Their approach combines the recursive algorithm of Wang et al. (1995) with the DET approach to estimate IBD between sibs from Knott and Haley (1998). To prevent complicated integration over all possible haplotype phases, the Wang et al. (1995) algorithm is only implemented for nearest phase-known markers.

Table 2. Overview of traits and contributing populations

Name	Description and populations for which data available
Growth	Life daily gain (all populations) and gain during test (H_1, LW_1, and P_1)
Fatness	Ultrasonic back fat thickness (all populations)[a]
Weight	Live weight and carcass weight (L_2, LW_2, P_2)
Length	Carcass length (L_2, LW_2, P_2)
G2	Carcass fatness (L_2, LW_2, P_2)
Fom	Fat on muscle (L_2, LW_2, P_2)
PH45ld	PH, 45 min after slaughter on longissimus dorsi muscle (L_2, LW_2, P_2)
PH45sm	PH, 45 min after slaughter on longissimus semimembranaceus muscle (L_2, LW_2, P_2)
PH24ld	PH, 24 h after slaughter on longissimus dorsi muscle (L_2, LW_2, P_2)
PH24sm	PH, 24 h after slaughter on longissimus semimembranaceus muscles (L_2, LW_2, P_2)
CE45ld	Electric conductivity, 45 min after slaughter on longissimus dorsi muscle (L_2, LW_2, P_2)
CE45sm	Electric conductivity, 45 min after slaughter on longissimus semimembranaceus muscle (L_2, LW_2, P_2)
CE24ld	Electric conductivity, 24 h after slaughter on longissimus dorsi muscle (L_2, LW_2, P_2)
CE24sm	Electric conductivity, 24 h after slaughter on longissimus semimembranaceus muscles (L_2, LW_2, P_2)

[a]All populations have single scores for the P_2 position (average of two measurements at the middle of the loin, 65 mm from the midline), except the Sygen Meishan-derived pigs (M), who have three scores (C: middle of eye muscle, K: end of loin muscle, L: minimum back fat thickness at the loin muscle).

In the second step, an animal model for the quantitative trait, including a random QTL effect, is fitted for every position:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Zu} + \mathbf{Zv} + \mathbf{Wc} + \mathbf{e} \quad [1]$$

where $\mathbf{y}$ is a $(m \times 1)$ vector of phenotypes, $\mathbf{X}$ is a $(m \times s)$ design matrix, β is a $(s \times 1)$ vector of fixed effects (e.g., sex), $\mathbf{Z}$ is an $(m \times q)$ incidence matrix relating animals to phenotypes, $\mathbf{u}$ is a $(q \times 1)$ vector of polygenic effects, $\mathbf{v}$ is a $(q \times 1)$ vector of additive genotypic QTL effects, $\mathbf{W}$ is an $(m \times q)$ incidence matrix relating litters to phenotypes, $\mathbf{c}$ is the $(q \times 1)$ vector of random litter effects, and $\mathbf{e}$ is a residual vector. The random genetic effects $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{c}$ are assumed to be distributed as multivariate normal densities with mean zero and variances $\mathbf{A}\sigma^2_u$, $\mathbf{G}\sigma^2_v$, and $\mathbf{I}\sigma^2_c$, respectively. Matrix $\mathbf{A}$ is the standard additive genetic relationship matrix and $\mathbf{G}$ is the $(q \times q)$ (co)variance matrix for the additive QTL effects, represented by the proportion of alleles IBD (George et al., 2000). The VC analyses were performed using ASREML (Gilmour et al., 1998). A test statistic for a given location was obtained by running an animal model without a QTL effect:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Zu} + \mathbf{Wc} + \mathbf{e} \quad [2]$$

Twice the difference between the logarithms of the likelihood of (1) vs. (2) was used as a log likelihood ratio (**LR**) test. For hypothesis testing, we imposed a nominal threshold of 5% by assuming that the LR would follow a mixture of a χ^2 distribution with 1 df and a peak at zero (Self and Liang, 1987). This may seem anticonservative because we tested multiple positions for every candidate region. To claim significant linkage for new QTL, Lander and Kruglyak (1995) advocated the use of genome-wide thresholds. However, for many regions and traits in the present study, we aimed to confirm published QTL within commercial lines. For this purpose, Lander and Kruglyak (1995) recommended the use of 0.01 nominal *P*-values to claim "confirmed linkage." The same authors also stated that any evidence for QTL exceeding the nominal 5% level should still be reported, even though this is not convincing evidence for the existence of a QTL. To facilitate the comparison between the two methods, we used the 5% nominal level as the basis for detection of a putative QTL. Comparisons were also made at the 1 and 0.1% level to evaluate the effect of statistical stringency on the results.

Results and Discussion

Results of Variance Component Analyses

The VC analyses of all populations and candidate regions for the traits that were available (between 2 and 14 traits/population) revealed 61 QTL exceeding the nominal 5% threshold (Table 3). Electric conductivity 24 h after slaughter in the longissimus dorsi was the only trait showing no QTL. For 27 of the putative trait regions, QTL have been reported in the literature (Bidanel and Rothschild, 2002). Imposing a nominal *P*-value of 0.01, confirmed linkage (Lander and Kruglyak, 1995) can be claimed for eight QTL affecting growth, (carcass) fatness, weight, or pH (Table 3). Nearly all the putative growth- and fatness-related QTL were already reported in the literature, whereas none of the six putative QTL affecting carcass length were published (Table 3, Bidanel and Rothschild, 2002). For other traits, such as pH, electric conductivity, and "fat over meat," comparison to published results was hampered because studies including these traits are sparse and trait definitions vary. Although three regions on SSC1, SSC6, and SSC9 were selected as control regions, Bidanel and Rothschild (2002) summarized several growth, fatness, and meat quality QTL in these regions on SSC1 and SSC6. For the SSC9 region, however, there are no published QTL for the traits considered in this study. With

Table 3. Proportion of phenotypic variance explained by variance component-identified QTL across 10 populations (H_1, H_2, L_1, L_2, LW_1, LW_2, LW_3, M, P_1, and P_2) and selected regions on 10 chromosomes (SSC)[a]

Marker interval Trait[b]	SSC1 CGA-Sw2185	SSC2[b] *IGF2* region	SSC3 Sw72-Sw251	SSC4 Sw35-Sw839	SSC6 Sw316-S0003	SSC7 *MHC* region	SSC8 Sw905-Sw1029	SSC9 Sw983-Sw21	SSC10 S0070-Sw1041	SSC13 S0068-Sw398
Growth				M(0.11)**	LW$_1$(0.13)**	M(0.08)*	M(0.08)*	L$_2$(0.04)*		
			H$_1$(0.08)*		LW$_1$(0.10)*	H$_1$(0.13)**	L$_1$(0.10)*			
			H$_1$(0.09)*			H$_1$(0.12)***				
Fatness	P$_1$(0.17)**	H$_2$(0.08)*	L$_1$(0.13)*	LW$_1$(0.08)*			M(0.07)*		H$_2$(0.27)*	H$_2$(0.12)*
	H$_2$(0.19)*			P$_1$(0.06)*			M(0.07)*			
				H$_2$(0.12)*			M(0.07)*[d]			
Weight		L$_2$(0.07)*		LW$_2$(0.08)*				L$_2$(0.08)*		
Length		L$_2$(0.07)***	L$_2$(0.07)*		LW$_2$(0.06)*			L(0.06)*	L$_2$(0.09)**	L$_2$(0.07)*
G2	LW$_2$(0.26)*	LW$_2$(0.14)***		LW$_2$(0.22)**	LW$_2$(0.05)*		P$_2$(0.10)*			
Fom		LW$_2$(0.10)*							LW(0.18)*	
PH45ld					LW$_2$(0.13)*					
					P$_2$(0.30)**					
PH45sm				P$_2$(0.13)**	LW$_2$(0.06)*					
					P$_2$(0.13)*					
PH24ld	L$_2$(0.10)*	LW$_2$(0.11)**	P$_2$(0.09)*		L$_2$(0.08)*	L$_2$(0.11)**		LW$_2$(0.11)**		
PH24sm		LW$_2$(0.09)*	LW$_2$(0.11)*			P$_2$(0.11)*		LW$_2$(0.06)*		
CE45ld		P$_2$(0.13)*			LW$_2$(0.12)*	LW$_2$(0.07)*				
						P$_2$(0.06)*				
CE24ld					LW$_2$(0.08)*					
CE24sm		LW$_2$(0.09)*						LW$_2$(0.06)*		

*$P < 0.05$.
**$P < 0.01$.
***$P < 0.001$.
[a]When cell contents are underlined, QTL for these traits have already been reported in these regions according to the review by Bidanel and Rothschild (2002).
[b]For trait definitions, see Table 2.
[c]The Sygen populations were not typed for the SSC2 region but for a second candidate region on SSC1 (S0065-Sw1301), which did not show any QTL using the VCA analyses.
[d]QTL for each of three ultrasonic backfat scores.

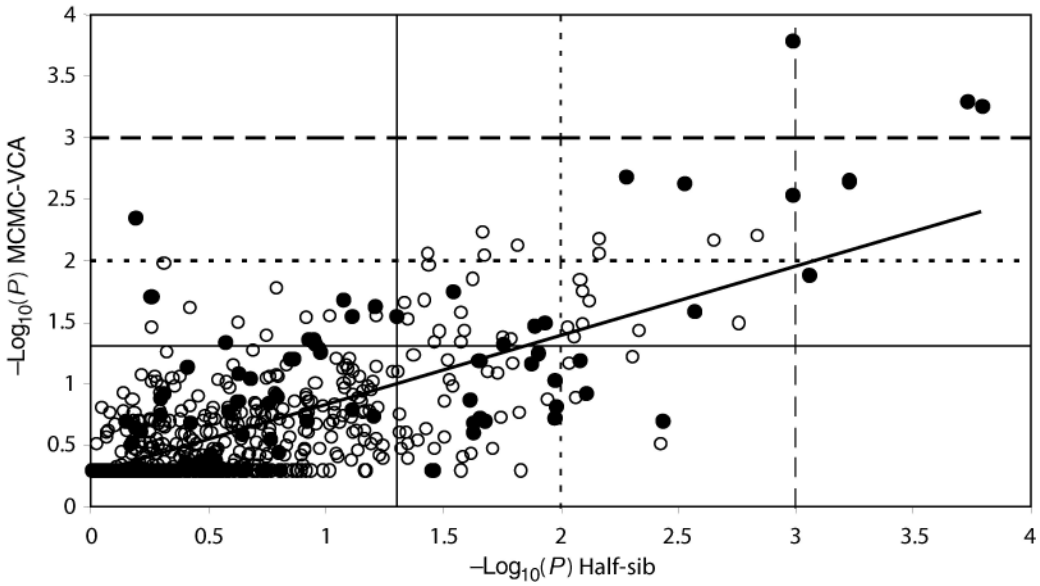

Figure 1. Scatter plot of transformed *P*-values of 650 trait × region evaluations comparing results from the half-sib regression analyses to those of the variance component analyses (VCA) using Monte Carlo Markov Chain (MCMC)-derived identity-by-descent (IBD) scores. The horizontal and vertical lines denote the 0.1% (dashed lines), 1% (dotted lines), and 5% (solid lines) thresholds for variance component and half-sib analyses, respectively. A trend line, which captures approximately 50% of the variation, is added for comparison. Analyses represented by filled circles have been repeated separately using deterministic IBD methods (Figure 2).

only one "control" region remaining, a meaningful comparison between candidate and control regions was not feasible. The most significant QTL ($P < 0.001$) was found for carcass fatness (**G2**) on SSC2 in the Spanish Large White. The QTL with the largest effect, which explained 30% of the phenotypic variance for the pH of the longissimus dorsi muscle 45 min after slaughter, was found on SSC6 in the Spanish Pietrain lines. This QTL is probably due to the Halothane mutation, a major gene that causes porcine stress syndrome and has large pleiotropic effects on carcass traits (Fujii et al., 1991). Evans et al. (2003) detected a Halothane effect in the Spanish Large White population, where one of the sires tested heterozygous for this mutation, but not in the Spanish Pietrain lines. Within breeds, there are no QTL that are present in multiple populations of that breed (Table 3). In contrast, Nagamine et al. (2003) report the same QTL on SSC4 and SSC7 segregating in populations from five different U.K. breeding companies. The lack of correspondence between populations of the same breed in our data could reflect different breeding goals in the three countries where the lines were bred.

Differences Between Half-Sib and Variance Component Results

Figure 1 shows a comparison between the results obtained with the VC analyses and those obtained with the HS analyses by Evans et al. (2003). Note that the slope of the trendline is less than unity, at least in part as a result of a large number of VC analyses giving a LR of zero (Figure 1). Using the nominal 5% significance threshold, the two methods agree in showing 42 QTL as significant. However, 19 QTL are detected only using the VC analyses, whereas 46 QTL, detected by Evans et al. (2003) using the HS analyses, were not confirmed by the VC analyses. When comparing at the 1% level, 11 QTL were detected under both models, 17 under HS only and 6 only under VC. At the 0.1% level, only three QTL are significant under both models, whereas three others are significant only under the HS model. The comparison between the two methods with regard to detection of QTL seems fairly robust to the choice of threshold. This suggests that the discrepancies between the results reflect something more than just differences in Type-I error between the two methods. In both cases, when only one method detected a QTL, the test statistic for the other model varied between nearly significant to completely insignificant (Figure 1). In order to understand the discrepancies between the HS and VC results, it is important to note the methodological differences between the two methods. In the HS model, an allele substitution effect is estimated as a fixed effect for every sire, independent of the other half-sib families. The maternally inherited QTL alleles are assumed to be randomly distributed between half-sibs, and the maternal genotypes are only used to increase the number of offspring that are informative for the inheritance of the sire allele. In the VC model, the variance explained by the QTL is estimated across all animals, assuming segregation of the QTL in both parents. From the 46 QTL that were only detected by the HS analyses, in 23 cases, only a single sire was inferred to be heterozygous for the QTL, whereas in five other cases, only two out

of the nine sires that made up the population were heterozygous for the QTL. When a QTL is segregating at such a low frequency, it could be missed by the VC analyses because the power of detection depends on the variance explained by the QTL across the population.

Alternatively, differences can arise from different allele frequencies in the sire or the dams due to sampling or selection. A QTL could be missed by the HS analyses when none or few sires are segregating. A clear example of this is the QTL for pH 45 min after slaughter in the Spanish Pietrain line, which is probably a consequence of the Halothane mutation. Although none of the five Spanish Pietrain sires carried this mutation, 13 out of 60 dams were heterozygous. This explains why this QTL was detected under the VC analyses, which incorporate information from within dam segregation, but not under the paternal HS analysis. On the other hand, when a QTL is detected in the sires but not segregating in the dams, the QTL effect is "diluted" in the VC analyses and may therefore be missed. We have looked at this hypothesis for 17 putative QTL. This included the Halothane effect in the Spanish Pietrain, whereas the other 16 QTL were picked from the regions that had been analyzed using both MCMC-derived and DET IBD scores and represented one of the three categories: 1) five QTL where identified both by HS and VC analyses, 2) six QTL were only detected under the HS analyses, and 3) five QTL, which were only detected under the VC analyses (using MCMC-derived IBD). These 16 QTL are marked in Figure 2. For these putative QTL, the HS analyses were repeated by fitting a QTL within every maternal full-sib family of sufficient size. Assuming that testing for a maternal QTL effect is independent of testing a paternal QTL effect, we can combine the two *P*-values using Fisher's test:

$$-2\sum_{i=1}^{2}\text{Log}(P_i) \qquad [3]$$

which follows a χ^2_4 distribution. The results for these additional analyses are summarized in Table 4. For the examples where both HS and VC methods detected a QTL, the joint *P*-value is always <0.05, even though there is significant evidence for a QTL in the dams in only one example (Table 4). For the analyses where the QTL is only detected under the HS analyses, none of the maternal HS analyses shows any evidence for a QTL. Although the joint *P*-values are still significant for all but a single example (Table 4), they are larger than those from the paternal HS analyses. For the examples where the HS showed no evidence for a QTL, the maternal HS analyses showed significant evidence in four out of six cases. The joint *P*-values were <0.05 for three cases, whereas all of them were <0.25 (Table 4). The comparison between the joint *P*-values and those from the VC analyses is compromised by the fact that results from the separate optimization of a paternal and maternal model under HS are compared with results from the joint optimization under the VC analyses. Nevertheless, the joint *P*-values provide a better comparison between the HS and VC results and give insight into the mechanisms underlying any discrepancies between the methods. Although we only looked at 17 cases, this provides some evidence that differences in QTL allele frequencies between sexes cause discrepancies between VC and HS analyses. These differences are probably due to sampling but could also reflect effects of selection when the parents originate from specific dam and sire lines. Furthermore, differences between paternal and maternal models could be explained by genomic imprinting, where the allele coming from one parent is silenced in the offspring.

The stability of parameter estimates under the VC analyses might also underlie some of the differences between HS and VC results. The estimates for the three variance components can differ greatly across regions for the same trait. For instance, in the Spanish Large White, a QTL was identified for G2 on SSC1, SSC2, SSC4, and SSC6. Whereas the proportion of variance explained by the QTL varied between 0.05 and 0.26 for these regions, the polygenic heritability varied between 0.00 and 0.15. Even within the same region, VC estimates can vary considerably, and the position with the highest LR is often not the same location that gives the highest variance component for the QTL. This is also reflected by large (i.e., 10-fold) fluctuations in estimates of VC resulting only in minute differences of the LR. It must be noted that the estimate of the QTL variance was more stable than those of litter and polygenic variance. This can be explained by the information underlying the variance components: litter and polygenic variances are proportional to between family variance, whereas QTL variance is estimated on the within-family variance. For a litter, all animals have the same values for the additive genetic relationship, whereas they can have very different values for their IBD relationship at a given genome location.

To our knowledge, only Zhang et al. (1998) have compared the performance of HS and VC analyses on real data. They reported that both methods agree well with regard to the QTL positions, but they did not compare the significance of the QTL under the two methods, although their Table 5 lists several cases where the QTL was only detected under one of the methods. However, such a comparison would be complicated by the fact that they derived the thresholds for the two methods in different ways. It must be noted that Zhang et al. (1998) analyzed a single large granddaughter design, whereas the present study looked at 10 moderately sized HS designs.

Use of Deterministic Identity by Descent Scores

For 15 within-population regions, 127 analyses were repeated with IBD scores obtained by the DET method. These combinations, highlighted in Figure 1, were chosen to represent all populations and to include a least one putative QTL from either the HS or VC analyses

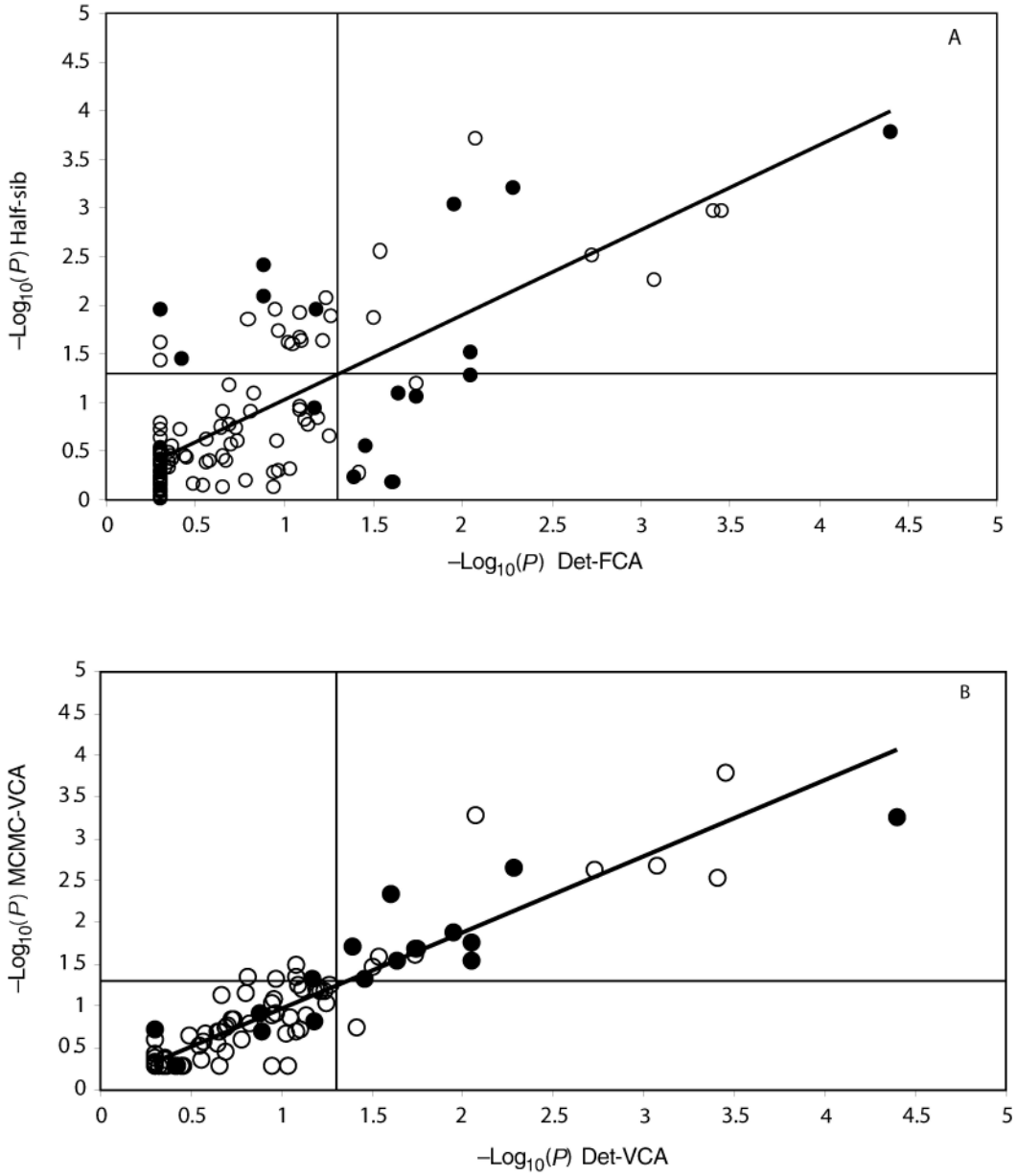

Figure 2. Scatter plots of transformed *P*-values of 127 trait × region evaluations, comparing results of variance component analyses (VCA) using deterministic identity-by-descent methods to half-sib regression analyses (A) and the variance component analyses using identity-by-descent scores that were derived using Monte Carlo Markov Chain methods (B). The horizontal and vertical lines denote the 5% thresholds. Trend lines, which capture 57% (A) and 87% (B) of the variation, are added for comparison. Filled circles indicate positions that were reanalyzed subsequently under a maternal half-sib model (Table 4).

for each within-population region. The 127 analyses represent 15 cases where both VC and HS methods detected a QTL, nine where VC detected a QTL and HS did not, 15 where HS detected a QTL and VC did not, and 88 cases where neither method identified a QTL. The results of the VC analyses using deterministic IBD are compared with both the original HS analyses and the VC analyses using MCMC-derived IBD scores in Figure 2. The VC results using deterministic IBD methods agree very well with those obtained using MCMC methods (Figure 2B). Given the close agreement between the MCMC and DET methods to obtain IBD scores, the latter should be preferred when analyzing large amounts of data with relatively few missing markers. It must be noted that the population structure of the present experiment is still fairly simple (few additional links between families), and may therefore not offer the best comparison between MCMC and DET.

A determining factor for the feasibility of genome scans with VC analyses is the computation time that is required. To analyze five positions in a candidate region for a single trait, the complete analyses (includ-

Table 4. Separate and combined P-values for paternal and maternal half-sib analyses compared to the variance component analyses

Population[a]	SSC[b]	Trait[c]	Type[d]	*P* sire[e]	*P* dam[e]	*P* joint[e]	*P* VC[e]
P_1	7	Growth	C	<0.001	0.121	<0.001	<0.001
LW_2	9	PH24ld	C	0.001	0.217	0.001	0.004
H_2	13	Fatness	C	0.001	0.291	0.002	0.013
P_2	7	PH24sm	C	0.050	0.032	0.012	0.019
LW_2	2	CE24sm	C	0.029	0.291	0.049	0.013
L_2	7	PH24sm	H	0.011	0.109	0.009	0.108
L_2	7	PH45ld	H	0.013	0.170	0.016	0.113
P_2	7	PH45sm	H	0.004	0.667	0.017	0.165
L_2	2	Fom	H	0.011	0.404	0.028	0.345
P_2	7	CE45sm	H	0.008	0.761	0.037	0.126
LW_2	7	CE45sm	H	0.035	0.826	0.131	0.440
P_2	6	PH45ld	V	0.45	0.005	0.016	0.02
P_2	2	CE45ld	V	0.560	<0.001	0.002	0.030
LW_2	2	Fom	V	0.077	0.029	0.016	0.028
P_1	1	Fatness	V	0.650	0.023	0.078	0.015
M	8	Growth	V	0.268	0.086	0.110	0.040
LW_2	2	PH24sm	V	0.085	0.744	0.238	0.019

[a]See Table 1 for details on populations.
[b]SSC = Region on chromosome.
[c]See Table 2 for trait details.
[d]QTL detected under both models (C) or under HS (H) or VC (V) only.
[e]Tabulated *P*-values not adjusted for multiple testing.

ing six ASREML runs) using DET were about nine times faster than those using MCMC (175 and 1,664 central processing unit seconds on a DEC Alpha XP1000 [Hewlett-Packard, Palo Alto, CA] with a 500-MHz processor, respectively). When using DET, the ASREML analyses and the processing of the results became the limiting factor in place of getting the IBD scores. Bayesian analyses for arbitrary pedigrees have also been proposed (Uimari et al. 1996; Bink and Van Arendonk, 1999), but applications to real data are limited because of computational requirements of these methods (Van Kaam et al., 2002). A major advantage of the two-step approach (George et al., 2000), compared with Bayesian methods (Bink and Van Arendonk, 1999), is that once the IBD scores are estimated, a large number of traits or models can be evaluated without the need to repeat the IBD estimation.

The two-step VC approach can accommodate arbitrary pedigrees and, when using ASREML, a wide range of genetic and statistical models. These include multivariate analyses, time series, and random regression models (Gilmour et al., 1998). A prerequisite for using more complicated models is the availability of sufficient amounts of data (i.e., large enough genotyped pedigrees). The possibility of using up to six user-defined covariance matrices in ASREML allows exploration of alternative genetic models in addition to the additive model used in the present study. Hanson et al. (2001) proposed a framework to test for imprinting in sib-pair studies using a VC approach. Shete and Amos (2002) provided a formal derivation for the methods proposed by Hanson et al. (2001) and explored the sensitivity of tests for imprinting to differences in recombination fractions between males and females. The methods of Pong-Wong et al. (2001) allow for parent-specific allelic IBD scores, which could subsequently be used to model separate paternal and maternal QTL effects. However, it is not clear how the sib-pair methodology generalizes to arbitrary pedigrees and how differences between maternal and paternal family sizes affect the power to distinguish Mendelian from imprinted QTL.

We have demonstrated the feasibility of QTL analysis using VC on a large amount of real data, equivalent to a full genome scan. The VC analyses performed well, especially when considering that the marker information was patchy (only 1 to 3 markers per region) and that the number of phenotyped animals was relatively small for reliable estimation of VC. Although the current population structure seems sufficient to detect QTL, larger populations are recommended for more reliable estimation of QTL effects.

Although the VC analyses showed few additional QTL beyond the HS analyses, they provided useful information. The present results increased confidence in those QTL that were detected by both methods and warrant closer scrutiny of the ones that were detected by only one of the methods. The experimental structure was very simple and was designed to be analyzed under a HS model. Advantages of VC methods could be more prominent when phenotypes are also available on the parents, because this information is ignored by the HS methods. As more advanced methodology is becoming available all the time, its usefulness can ultimately only be assessed by the analysis of real data. This will also facilitate further refinement of these methods. When the family structure permits, we recommend HS regression models for the initial analyses of QTL experiments

within commercial lines. The advantages are the computational speed and straightforward interpretation of HS analyses, which can be performed online using QTL Express software (Seaton et al., 2002). Given the effort and money that go into QTL mapping experiments, alternative methods should always be explored to exploit all the information that is present in the experiment. In this context, VC analyses are very useful to reanalyze data because they take all additive genetic relationships into account and provide QTL breeding values for all animals. Any discrepancies between the methods will point to QTL that need closer scrutiny.

Implications

The performance of variance components methods for detecting quantitative trait loci was compared to that of relatively simple half-sib methods using real data from 10 populations of five different pig breeds. Variance component analyses are routinely used in animal breeding for breeding value estimation, and inclusion of quantitative trait loci effects will facilitate marker-assisted selection. It was demonstrated that chromosome regions explaining more than 5% of the phenotypic variation could be detected in pedigrees of 500 offspring. There were some discrepancies between the variance component and half-sib analyses, which could partly be explained by differences underlying the two methods. We believe the present results provide a step forward in robust linkage analyses for outbred pedigrees.

Literature Cited

Bidanel, J. P., and M. Rothschild. 2002. Current status of quantitative trait locus mapping in pigs. Pig News Info. 23:39–54.

Bink, M. C. A. M., and J. A. M. Van Arendonk. 1999. Detection of quantitative trait loci in outbred populations with incomplete marker data. Genetics 151:409–420.

Evans, G., E. Giuffra, A. Sanchez, S. Illán, J. L. Noguera, I. Velander, O. Southwood, C. S. Haley, G. Plastow, and L. Andersson. 2002. PIGQTECH-Transferring QTL technology to the pig breeding industry. Proc. 7th World Congr. Genet. Appl. Livest. Prod., Montpellier, France.

Evans, G., E. Giuffra, A. Sanchez, S. Kerje, G. Davalos, O. Vidal, S. Illán, J. L. Noguera, L. Varona, I. Velander, O. I. Southwood, D.-J. de Koning, C. S. Haley, G. S. Plastow, and L. Andersson. 2003. Identification of quantitative trait loci for production traits in commercial pig populations. Genetics 164:621–627.

Fernando, R. L., and M. Grossman. 1989. Marker assisted selection using best linear unbiased prediction. Genet. Select. Evol. 21:467–477.

Fujii J., K. Otsu, F. Zorzato, S. de Leon, V. K. Khanna, J. E. Weiler, P. J. O'Brien, and D. H. MacLennan. 1991. Identification of a mutation in the porcine ryanodine receptor that is associated with malignant hyperthermia. Science 253:448–451.

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

Gilmour, A. R., B. R. Cullis, S. J. Welham, and R. Thompson. 1998. ASREML: User manual. Orange Agric. Inst., NSW, Australia

Hanson, R. L., S. Kobes, R. S. Lindsay, and W. C. Knowler. 2001. Assessment of parent-of-origin effects in linkage analysis of quantitative traits. Am. J. Hum. Gen. 68:951–962.

Heath, S. C. 1997. Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. Am. J. Hum. Genet. 61:748–760.

Knott, S. A., and C. S. Haley. 1998. Simple multiple marker sib-pair analysis for mapping quantitative trait loci. Heredity 81:48–54.

Lander, E. S., and L. Kruglyak. 1995. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. Nat. Genet. 11:241–247.

Nagamine Y., C. S. Haley, A. Sewalem, and P. M. Visscher. 2003. QTL variation for growth and obesity between and within lines of pigs (Sus scrofa). Genetics 164:629–635.

Pong-Wong, R., A. W. George, J. A. Woolliams, and C. S. Haley. 2001. A simple and rapid method for calculating identity-by-descent matrices using multiple markers. Genet. Select. Evol. 33:453–471.

Seaton G., C. S. Haley, S. A. Knott, M. Kearsey, and P. M. Visscher. 2002. QTL Express: mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics 18:339–340.

Self, S. G., and K. Y. Liang. 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J. Am. Stat. Assoc. 82:605–610.

Shete, S., and C. I. Amos. 2002. Testing for genetic linkage in families by a variance-components approach in the presence of genomic imprinting. Am. J. Hum. Genet. 70:751–757.

Sørensen, A. C., R. Pong-Wong, J. J. Windig, and J. A. Woolliams. 2002. Precision of methods for calculating identity-by-descent matrices using multiple markers. Genet. Select. Evol. 34:557–579.

Thompson, E.A., and S. C. Heath. 1999. Estimation of conditional multilocus gene identity among relatives. Pages 95–113 in Statistics in Molecular Biology. F. Seillier-Moseiwitch, P. Donnely, and M. Watermann, ed. Springer-Verlag IMS Lecture Note Series, New York.

Uimari P., G. Thaller, and I. Hoeschele. 1996. The use of multiple markers in a Bayesian method for mapping quantitative trait loci. Genetics 143:1831–1842.

Van Kaam J. B. C. H. M., M. C. A. M. Bink, H. Bovenhuis, and R. L. Quaas. 2002. Scaling to account for heterogeneous variances in a Bayesian analysis of broiler quantitative trait loci. J. Anim. Sci. 80:45–56.

Wang, T., R. L. Fernando, S. Van der Beek, M. Grossman, and J. A. M. Van Arendonk. 1995. Covariance between relatives for a marked quantitative trait locus. Genet. Select. Evol. 27:251–274.

Zhang, Q., D. Boichard, I. Hoeschele, C. Ernst, A. Eggen, B. Murkve, M. Pfister-Genskow, L. A. Witte, F. E. Grignola, P. Uimari, G. Thaller, and M. D. Bishop. 1998. Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. Genetics 149:1959–1973