Multiple Linkage Disequilibrium Mapping Methods to Validate Additive Quantitative Trait Loci in Korean Native Cattle (Hanwoo)

Yi Li and Jong-Joo Kim[1],*

School of Statistics, Shanxi University of Finance and Economics, Taiyuan, Shanxi 030006, China

ABSTRACT: The efficiency of genome-wide association analysis (GWAS) depends on power of detection for quantitative trait loci (QTL) and precision for QTL mapping. In this study, three different strategies for GWAS were applied to detect QTL for carcass quality traits in the Korean cattle, Hanwoo; a linkage disequilibrium single locus regression method (LDRM), a combined linkage and linkage disequilibrium analysis (LDLA) and a BayesCπ approach. The phenotypes of 486 steers were collected for weaning weight (WWT), yearling weight (YWT), carcass weight (CWT), backfat thickness (BFT), *longissimus dorsi* muscle area, and marbling score (Marb). Also the genotype data for the steers and their sires were scored with the Illumina bovine 50K single nucleotide polymorphism (SNP) chips. For the two former GWAS methods, threshold values were set at false discovery rate <0.01 on a chromosome-wide level, while a cut-off threshold value was set in the latter model, such that the top five windows, each of which comprised 10 adjacent SNPs, were chosen with significant variation for the phenotype. Four major additive QTL from these three methods had high concordance found in 64.1 to 64.9Mb for *Bos taurus* autosome (BTA) 7 for WWT, 24.3 to 25.4Mb for BTA14 for CWT, 0.5 to 1.5Mb for BTA6 for BFT and 26.3 to 33.4Mb for BTA29 for BFT. Several candidate genes (i.e. glutamate receptor, ionotropic, ampa 1 [*GRIA1*], family with sequence similarity 110, member B [*FAM110B*], and thymocyte selection-associated high mobility group box [*TOX*]) may be identified close to these QTL. Our result suggests that the use of different linkage disequilibrium mapping approaches can provide more reliable chromosome regions to further pinpoint DNA makers or causative genes in these regions. (**Key Words:** Quantitative Trait Loci, Linkage Disequilibrium, Single Nucleotide Polymorphism, Hanwoo, Carcass Quality)

INTRODUCTION

In beef cattle, growth and carcass quality are very important traits, because of the relevance of these traits to economic profits for beef cattle farmers. Traditional breeding programs based on breeding values by best linear unbiased prediction (BLUP) using pedigree and phenotypic data has achieved significant genetic progress. For example, annual gains for carcass weight (CWT) and *longissimus dorsi* muscle area (LMA) in the 2000's were 8 kg and 2.9 cm^2, respectively, in Hanwoo (NIAS, 2009).

However, genomic prediction at a young stage can further accelerate genetic gain, because prediction errors at

breeding age would be reduced by exploiting information on the value of every chromosomal fragment and on the transmission of chromosome fragments from parents to selection candidates (Garrick, 2011). Additionally, genome-wide prediction has the potential to reduce inbreeding rates because of the increased emphasis on own rather than family information (Daetwyler et al., 2011). Furthermore, quantitative trait loci (QTL) identification provides valuable information for understanding genetic mechanisms underlying beef phenotypes and for detecting causal polymorphisms, which subsequently can be utilized in beef industry through genome selection or gene-base selection (Misztal, 2011).

The availability of genome-wide single nucleotide polymorphism (SNP) panels enables detection of any SNP that is associated with a trait by a genome-wide association study (GWAS), which allows mapping QTL across the genome.

* Corresponding Author: Jong-Joo Kim. Tel: +82-53-810-3027; Fax: +82-53-810-4769, E-mail: kimjj@ynu.ac.kr
[1] School of Biotechnology, Yeungnam University, Gyeongsan 712-749, Korea.

Several QTL mapping methods that exploited linkage disequilibrium between molecular markers and QTL have been suggested and applied (Zhao et al., 2007; Cierco-Ayrolles et al., 2010; Uleberg et al., 2010; Erbe et al., 2011; Sun et al., 2011). These methods are based on the assumption that saturated markers increase the feasibility of QTL detection using historical population-wide linkage disequilibrium, in which a marker allele is in linkage disequilibrium (LD) with the QTL allele across the entire population. The term, LD, implies disequilibrium of two loci due to physical linkage. However, disequilibrium can also exist between two loci that are physically unlinked, impairing reliability of LD-based QTL mapping, because the SNP far from the QTL may generate significant signals, i.e. false positive QTL.

To exclude false positive QTL, combined linkage and LD mapping methods were developed (Blott et al., 2003; Druet and George, 2010), in which linkage information, i.e. co-segregation of a marker and a QTL from parents to progeny, was exploited, so as to detect SNPs that are closely located (physically linked) to the QTL. Bayesian multiple regression methods consider prior information of QTL effect, which was based on previous results about genetic architecture of the traits of interest. Also, the Bayesian method allows a flexible criterion to control false positive QTL, i.e. probability of proportion of false positives among the significant QTL, while frequentist methods set stringent thresholds due to huge number of SNPs markers in the GWAS test (Fernando and Garrick, 2013).

The aim of this study was to detect additive QTL for growth and carcass quality traits in Korean cattle, Hanwoo, by applying three different genome wide association analyses, a linkage disequilibrium single locus regression method (LDRM), a combined linkage and linkage disequilibrium method (LDLA), and the Bayes Cπ method.

MATERIALS AND METHODS

Animals, phenotypes and molecular data

The steers (N = 486) for phenotype and molecular data were chosen among the progeny of candidate bulls for progeny testing in the Hanwoo Improvement Center of National Agriculture Cooperative Federation in Seosan, Chungnam province, Korea. The data set comprised 61 sires and their 486 steers that were born between spring of 2005 and fall of 2007. The number of steers for each of the 61 paternal sire families ranged from 2 to 13 with the average of eight steers. Some details about raising and management of the experimental population and data recording were described in Li et al. (2011).

Six traits for growth and meat quality were chosen: weaning weight (WWT), 365-d yearling weight (YWT), CWT after slaughter, backfat thickness (BFT), LMA, and marbling score (Marb). Details about summary statistics for the observed carcass quality traits were described in Li (2012).

A dense marker map covering the whole genome was embedded in the Illumina Bovine SNP 50K BeadChip (Matukumalli et al., 2009). Details of the SNP genotyping experiments were described in Lee et al. (2008). This quality control resulted in 39,501 useful SNPs which were used for further analysis. This subset of markers covers 2,543.6 Mb of the genome with 70.57±69.0 kb average adjacent marker spacing. For the SNPs analyzed in this study, the average observed heterozygosity was estimated at 0.37±0.12. The average minor allele frequency (MAF) of all SNPs before editing was 0.20 and using the filtered SNPs the averaged MAF increased to 0.27.

Models used

Three complementary approaches were used: i) linkage disequilibrium single locus analysis using the SNP_a option of Qxpak (5.03 version) software (Perez-Enciso and Misztal, 2011); ii) combined linkage disequilibrium and linkage analysis using linkage disequilibrium variance component mapping (LDVCM) software; iii) Bayesian analysis using the BayesCπ option of GenSel (http://bigs.ansci.iastate.edu/bigsgui/login.html). All analysis used appropriate fixed factors or covariates that were fitted in the models (p<0.05) using a general linear model procedure in SAS (SAS 9.1, SAS Institute Inc., Cary, NC, USA). Two fixed effects were fitted in the models: year and season of birth (5 levels) for all the traits and region where the steers were born (41 levels) for WWT, YWT, and Marb, and three covariates, weaning age for WWT, yearling age for YWT, slaughter age for CWT, BFT, and LMA, was also fitted. Pedigrees of the base population animals were traced back for 2 generations to create the numerator relationship matrix, and 1,033 animals were included in the pedigree analysis.

LDRM: First, linkage disequilibrium single locus regression (Grapes et al., 2004; Zhao et al., 2007) were performed.

$$y = \mu + X\alpha + Z_u u + e$$

Where y is the phenotypic record, μ is the average phenotypic performance, X is the design matrix of SNP genotype (e.g. individuals with marker genotypes '11', '12', and '22' are assumed to have genetic values μ_{kAA}, 0, and μ_{kBB}), a is the fixed substitution effect for the SNP, Z_u is the incidence matrix for animal effects, u is the infinitesimal genetic effect, which is distributed as N(0, $A\sigma_u^2$) (the numerator relationship matrix A and the additive genetic variance σ_u^2) and e is a random residual for animal i, which is distributed as N(0, $I\sigma_e^2$) (the identity matrix I and

residual variance σ_e^2). Likelihood ratio tests were performed by removing the single locus SNP genotypic effects, and p-values were obtained assuming a x^2 distribution of the likelihood ratio test with one degree of freedom. Association with false discovery rate (FDR) <0.01 on chromosome-wide level were considered significant.

LDLA: The LDVCM software has been essentially described (Kim and Georges, 2002; Blott et al., 2003). Briefly, the linkage phases of all sires and sons were determined by the approach of Druet and Georges (2010). Then, identity-by-descent (IBD) probabilities (φ_p) at the midpoint of each SNP interval (*p*) were computed for all pairs of haplotypes conditional on the identity-by-state status of flanking markers (Meuwissen and Goddard, 2001). A dendrogram was generated by using the unweighted pair group method with arithmetic mean (UPGMA) hierarchical clustering algorithm with 1-φ_p as the distance measure at QTL location (*p*). Starting at the ancestral node and sequentially descending into the dendrogram, all possible combinations of haplotype clusters were analyzed in place of individual haplotypes. This process identified the set of nodes at which the likelihood of the data were maximized.

To jointly exploit linkage disequilibrium (female meiosis) and linkage (male meiosis) information, the following mixed linear was used:

$$y = \mu + Z_h h + Z_u u + e$$

Where *h* is the vector of random QTL effects corresponding to the defined haplotype clusters. Z_h is an incidence matrix relating maternal haplotypes of sons and sire haplotypes to individual sons. Likelihood ratio tests were performed by removing the haplotype cluster effects, and p-values were obtained assuming a x^2 distribution of the likelihood ratio test with one degree of freedom. Association with FDR <0.01 on chromosome-wide level were considered significant.

BayesCπ: BayesCπ were developed from the BayesB and GBLUP approaches (Meuwissen et al., 2001) and was described in detail by Habier et al. (2011).

$$y = \mu + \sum_{j=1}^{K} X_j \alpha_j \delta_j + e$$

Where K is the number of SNPs, X_j is the vector of genotypes at SNP$_j$, a_j is the random substitution effect for the SNP$_j$, which condition on the variance $\sigma_{\alpha j}^2$, is assumed normally distributed N(0, $\sigma_{\alpha j}^2$) when δ_j=1, while $\sigma_{\alpha j}^2$ =0, when $\delta_j = 0$, $\sigma_{\alpha j}^2$ is a random 0/1 variable indicating the absence (with probability π) or presence (with probability 1–π) of SNP$_j$ in the model. The prior for π was treated as unknown with uniform (0, 1). Gibbs sampling was applied to calculate the posterior means of model parameters μ, a_j, $\sigma_{\alpha j}^2, \sigma_e^2$ and π. The Markov chain Monte Carlo algorithms were run for 50,000 samples, with the first 20,000 samples discarded as burn in. A window size of 10 was used and the variance of each 10-SNP window was used as the criterion to detect QTL. Several windows that shared the same SNP with a large effect were considered to identify the same QTL region. Within each region, because windows were overlapping, the window with the highest variance of genomic estimated breeding value was used and the SNP within this window that explained the largest proportion of genetic variance was used to denote the position of the QTL. To select which positions should be major QTL, we first ranked them by estimated variance. We then chose a cut-off such that the top 5 QTL were detected.

Information about particular genes, located near SNP significantly associated with each trait, was extracted from online sources (http://www.ensembl.org/index.html, http://www.genecards.org/cgi-bin/cardsearch.pl#top, and http://www.uniprot.org).

RESULTS AND DISCUSSION

The GWAS profiles of all additive SNP effects for each trait are divided into the different statistical models (Figure 1). All SNP effects in this report were additive as was expected because predicted transmitting ability predict only additive genetic merit.

A total of 36 signal-trait combinations on 29 chromosomes crossed the chromosome-wide significance threshold by either LDRM or LDVCM, respectively (Tables 1 and 2). These signals in very close proximity to each other were often all significantly associated with a particular trait. This is because sets of SNPs that are physically located near the causal factor will tend to be in linkage disequilibrium. In this study, the same QTL was considered 'detected' if the significant signals were within a 1 Mb bracket. Therefore, the total number of distinct identified QTLs is 17. Of these 17 QTLs, 5 were commonly detected by both methods, while 7 and 5 were solely detected by LDRM and LDVCM, respectively.

Subsequent BayesCπ provides a reference for validation of association results, using a cut-off of the top 5 variances of 10-SNP windows (Table 3). The associated signal plots show in Figure 2, which provide a visualization of whether the consistency QTL were detected by three approaches. The consistency of QTL results obtained from the same dataset using different statistical approaches can provide more reliable chromosome regions to further pinpoint DNA makers or causative genes in these regions.
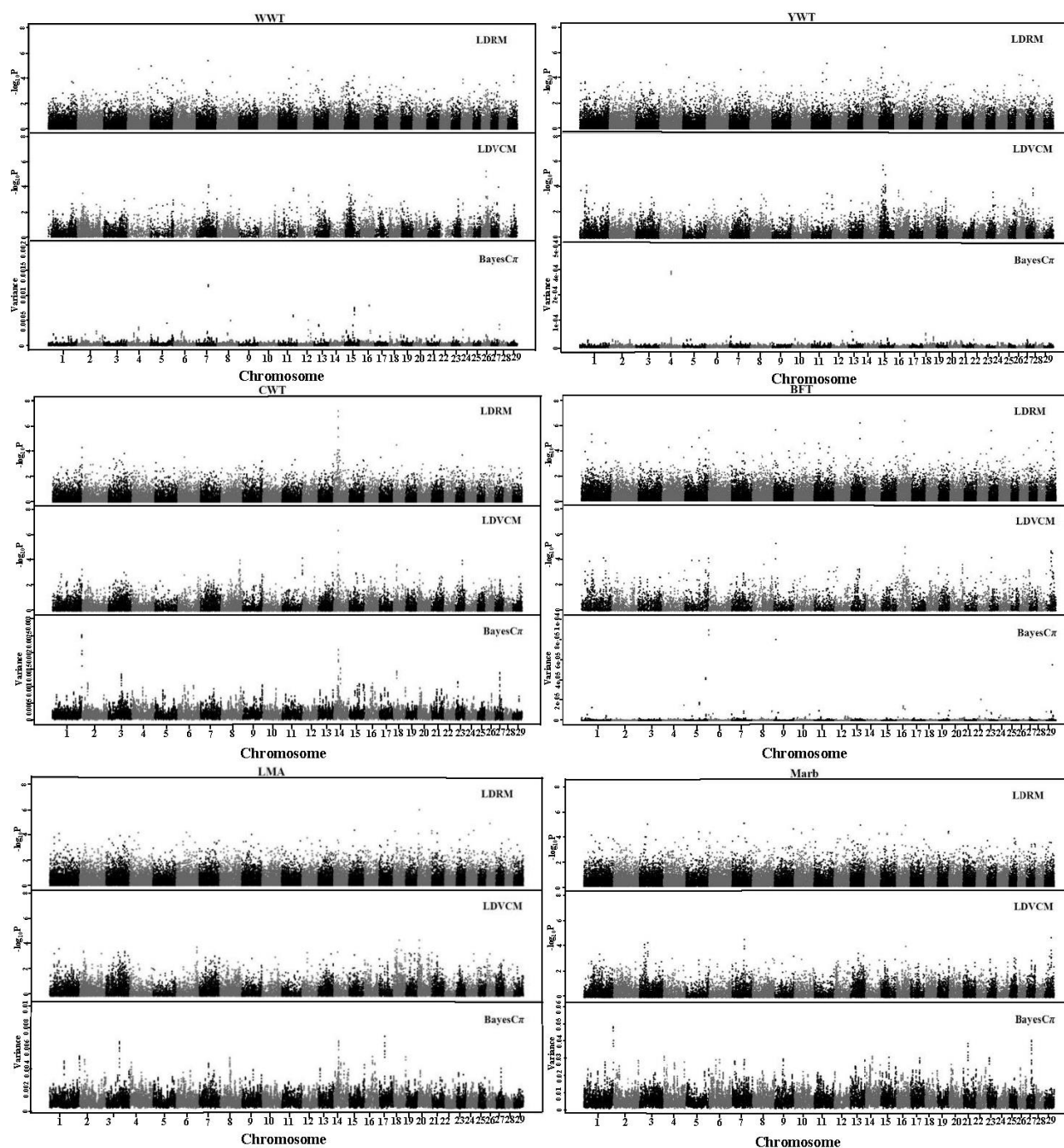
Figure 1. Genome-wide scan for WWT, YWT, CWT, BFT, LMA, and Marb using the three approaches: LDRM, LDVCM, and BayesCπ. WWT, weaning weight; YWT, yearling weight; CWT, carcass weight; BFT, backfat thickness; LMA, *longissimus dorsi* muscle area; LDRM, linkage disequilibrium single locus regression method; LDVCM, linkage disequilibrium variance component mapping.

Quantitative trait loci analysis by LDRM and LDLA

The single marker regression based on LD is shown in Table 1. The LDRM found 17 significant SNP associations, including 1 for WWT, 2 for YWT, 5 for CWT, 8 for BFT, and 1 for LMA. Ten of these SNPs were in agreement with QTL previously reported in the literature, including 1 for WWT, 1 for YWT, 5 for CWT, and 3 for BFT (Mizoshita et al., 2005; McClure et al., 2010). Further, 3 of the 17 detected SNPs are located within the regions of known

genes, and the others are 12,174 to 459,952 bp away from the nearest known genes. Among these 17 SNPs, there are 5, 2 SNPs located within a 2 Mb (between 24.31 to 25.29Mb) for CWT on *Bos taurus* autosome (BTA) 14, and a 0.2Mb (between 52.72 and 52.88Mb) for BFT on BTA13, and this indicated that they are associated with the same QTL.

LDLA results for growth and carcass quality using LDVCM are in Table 2. The LDVCM detected a total of 19 significant signals, including 2 for WWT, 3 for YWT, 2 for

Table 1. The additive QTL and SNP position that were detected for growth and carcass quality traits using LDRM analysis

QTL	SNP[1]	Chr	Position (bp)[2]	Allele[3]	MAF	Effect[4]	Significance (-log₁₀P[5])	Nearest gene[6] Name	Distance (bp)
Weaning weight									
1	BTB-00317489	7	64,536,896	G/A	A(0.27)	−8.10	5.45	*GRIA1*	40,884
Yearling weight									
2	ARS-BFGL-NGS-17747	15	13,517,697	T/C	C(0.35)	7.74	4.82	*SESN3*	21,628
3	ARS-BFGL-NGS-38840	15	31,874,189	T/A	T(0.48)	8.98	6.43	*UBASH3B*	86,347
Carcass weight									
4	BTB-01143619	14	24,315,258	T/C	T(0.06)	22.9	5.17	*FAM110B*	79,759
	BTB-01143580	14	24,383,627	G/A	G(0.06)	22.5	7.20	*UBXN2B*	32,070
	Hapmap30932-BTC-011225	14	24,898,781	T/C	C(0.06)	−24.7	5.81	*TOX*	within
	BTB-01280026	14	25,170,557	T/C	T(0.07)	24.0	6.75	*TOX*	96,111
	Hapmap27934-BTC-065223	14	25,288,714	G/A	A(0.05)	−25.2	5.89	*LOC100140360*	144,727
Backfat thickness									
5	BTB-01384704	6	959,692	A/G	G(0.48)	−0.12	5.57	*LOC100139637*	63,038
6	BTA-28590-no-rs	9	10,295,707	G/A	A(0.24)	−0.14	5.61	*bta-mir-30a*	406,847
7	BTB-01493007	13	52,723,213	T/C	C(0.05)	−0.25	4.92	*PTPRA*	within
	BTB-00534731	13	52,883,767	T/C	C(0.05)	−0.30	6.15	*VPS16*	within
8	BTB-00640968	16	40,694,131	G/C	C(0.26)	0.15	6.31	*CTNNBIP1*	12,174
9	ARS-BFGL-BAC-30052	23	11,964,461	G/T	T(0.45)	0.12	5.53	*ZFAND3*	24,327
10	ARS-BFGL-NGS-39535	29	26,364,496	G/A	A(0.26)	0.13	4.66	*NAV2*	88,899
11	BTB-01020342	29	33,014,346	G/A	G(0.49)	−0.12	5.40	*ETS1*	459,952
Longissimus dorsi muscle area									
12	Hapmap53674-rs29025319	20	8,676,696	T/C	C(0.12)	4.69	5.84	*LOC783819*	132,888

QTL, quantitative trait loci; SNP, single nucleotide polymorphism; LDRM, linkage disequilibrium single locus regression method; MAF, minor allele frequency.

[1,2] SNPs marker annotations and their positions were based on the bovine reference genome (btau4.0).

[3] Nucleotides of substitution. [4] Estimates of additive effects.

[5] Negative logarithm of the comparison-wise p-value of the test-statistic against the null hypothesis of no QTL at the most likely position for the inferred QTL model.

[6] The nearest known gene to the significant SNPs.

CWT, 11 for CWT, and 1 for Marb. Ten of these signals were in agreement with QTL previously reported in the literature, including 2 for WWT, 2 for YWT, 2 for CWT, and 3 for BFT (Mizoshita et al., 2005; McClure et al., 2010). Further, 5 of the 19 detected SNPs are located within the regions of known genes, and the others are 526 to 494,805 bp away from the nearest known genes. In particular, many significant signals (15 out of 19) were located in close distances: *i.e.* 5 SNPs are located within 0.2 Mb (between 26.24 to 26.45 Mb) for BFT on BTA29, which harbors the neuron navigator 2 (*NAV2*) gene.

Quantitative trait loci analysis by BayesCπ

BayesCπ analysis fitted all the SNPs simultaneously. Examination of genetic variance associated with 10-SNPs windows across the genome were chosen by the top 5 (Table 3, Figure 1). The largest differences were an increase in genetic variance associations found per traits: the 64.1 to 64.9 Mb region of BTA1 for WWT, 60.7 to 61.6 Mb region of BTA4 for YWT, 159.4 to 160.2 Mb region of BTA1 for

CWT, 0.5 to 1.5 Mb region of BTA6 for BFT, 28.1 to 29.2 Mb region of BTA17 for LMA, and 159.4 to 160.2 Mb of BTA1 for Marb. Some SNPs windows affected more than one carcass quality trait, including 24.7 to 25.4 Mb for BTA 14 (CWT and LMA), 35.6 to 36.5 Mb for BTA14 (CWT and Marb), 25.2 to 25.9 Mb for BTA27 (CWT and Marb) and 159.4 to 160.2 Mb for BTA1 (CWT, LMA, and Marb) (Stone et al., 1999; Li et al., 2004; Mizoshita et al., 2004; Mizoshita et al., 2005; Morris et al., 2009; McClure et al., 2010).

Comparison of result

The 17 QTL detected by LDRM or LDVCM were integrated using the BayesCπ analysis to remove possible redundancies among those QTL. The purpose of QTL mapping is as a first step towards the identification of QTL locations for further investigation by molecular geneticists. Therefore, agreement of the location of positive signals between the three detection methods was evaluated. The visual agreement of the information generated by these

Table 2. The additive QTL and SNP position that were detected for growth and carcass quality traits using LDVCM analysis

QTL	Chr	Position (bp)[1]	Significance ($-\log_{10}P^2$)	No. of cluster[3]	Flanking interval[4]		Nearest gene[5]	
					Start	End	Name	Distance (bp)
Weaning weight								
1	26	25,742,694	5.25	4	ARS-BFGL-NGS-20281	BTB-01936718	*SORCS3*	within
	26	25,781,274	4.81	5	BTB-01936718	ARS-BFGL-NGS-35579	*SORCS3*	within
Yearling weight								
2	15	21,268,864	5.77	7	BTB-00584953	BFGL-NGS-118595	*PLET1*	286,311
	15	21,298,176	5.42	7	BFGL-NGS-118595	ARS-BFGL-NGS-102765	*PLET1*	315,623
3	15	34,047,460	5.02	7	BTA-36664-no-rs	Hapmap40064-BTA-36665		
Carcass weight								
4	14	24,363,275	6.38	2	UA-IFASA-8415	BTB-01143580	*LOC100295254*	526
	14	25,268,611	4.64	3	Hapmap30734-BTC-065251	Hapmap27934-BTC-065223	*TOX*	130,010
Backfat thickness								
5	9	10,366,027	5.27	3	BTA-28590-no-rs	ARS-BFGL-NGS-14697	*LOC100336080*	440,468
6	16	40,659,543	4.50	3	ARS-BFGL-NGS-13513	BTB-00640968	*NMNAT1*	9,107
	16	40,899,155	5.00	2	BTB-00640968	Hapmap47279-BTA-38917	*PIK3CD*	36,791
7	29	26,242,330	4.69	6	ARS-BFGL-NGS-34288	ARS-BFGL-NGS-24205	*NAV2*	within
	29	26,278,577	4.16	10	ARS-BFGL-NGS-24205	ARS-BFGL-NGS-64656	*NAV2*	within
	29	26,329,748	4.57	5	ARS-BFGL-NGS-64656	ARS-BFGL-NGS-39535	*NAV2*	46,283
	29	26,376,548	3.86	3	ARS-BFGL-NGS-39535	ARS-BFGL-NGS-2475	*NAV2*	93,083
	29	26,453,403	3.90	10	ARS-BFGL-NGS-2475	ARS-BFGL-NGS-91937	*NAV2*	169,938
8	29	29,634,886	3.51	8	ARS-BFGL-NGS-94355	Hapmap54158-rs29026721	*SPA17*	within
9	29	32,987,362	4.52	10	UA-IFASA-5646	BTB-01020342	*ETS1*	494,805
	29	33,043,874	4.51	3	BTB-01020342	ARS-BFGL-NGS-11967	*ETS1*	438,293
Marbling score								
10	29	36,410,597	4.69	11	UA-IFASA-9704	ARS-BFGL-NGS-12285	*HTR1F*	10,414

QTL, quantitative trait loci; SNP, single nucleotide polymorphism; LDVCM, linkage disequilibrium variance component mapping.

[1,4] SNP marker annotations and their positions were based on the bovine reference genome (btau4.0).

[2] Negative logarithm of the comparison-wise p-value of the test-statistic against the null hypothesis of no QTL at the most likely position for the inferred QTL model.

[3] The number of clusters in the haplotype dendrogram that yields the highest LRT scores.

[4] Markers flanking the position of the maximum test statistic.

[5] The nearest known gene to the significant SNP.

methods is evident from Figure 2. Four QTL from these three methods had the high concordance locations and statistics found in 64.1to 64.9 Mb for BTA7 for WWT, 24.3 to 25.4 Mb for BTA14 for CWT, 0.5 to 1.5 Mb for BTA6 for BFT and 26.3 to 33.4 Mb for BTA29 for BFT. As shown in Figure 2 and Table 1, 2, and 3, the QTL for WWT was positioned within 64.1 to 64.9 Mb, one of the gene glutamate receptor, ionotropic, ampa 1 (*GRIA1*), that encodes glutamate receptor 1. The QTL for CWT was detected in the proximal region (24.3 to 25.4 Mb) of BTA14, which encompasses family with sequence similarity 110, member B (FAM110B), ubx domain protein 2B (UBXN2B), and thymocyte selection-associated high mobility group box (TOX) as positional and functional candidate genes for the CWT QTL in cattle. A QTL for BFT was detected in the proximal region (0.5 to 1.5 Mb) of the BTA 6, in which Locus (*LOC*) *100139637* gene was located. Another QTL was detected for Marb in the proximal region of the BTA29 (26.3 to 33.4 Mb), around which *NAV2* gene was located.

These QTL locations that were found by each method were as expected because the chromosomes were suspected to contain QTL.

Previous work has shown that these three methods belong to three distinct analytical methods, *i.e.* LDRM tests a single marker at a time and regards the markers as independent of all other markers (Grapes et al., 2004; Zhao et al., 2007; MacLeod et al., 2010), LDVCM tests the mid-point of the marker brackets, which corresponded best when the QTL was masked between analyzed markers(Kim and Georges, 2002; Blott et al., 2003), and BayesCπ test the effects of all markers, which are fitted simultaneously. However, QTL with large effects can be detected by both the Bayesian shrinkage and linear regression mapping methods. By specifying proper prior distributions for SNP effects, the ignorable small SNP effects are coerced to zero and only SNPs with larger effects on the phenotype are fitted in the model, hence Bayesian shrinkage analysis could reduce possible spurious QTL effects by adjusting all

Table 3. The additive QTL and SNP position that were detected for growth and carcass quality traits using BayesCπ analysis

QTL	Chr	Start[1]	End[2]	Most highest variance SNP					Nearest gene[6]	
				SNP	Position[3]	Allele[4]	MAF	Effect[5]	Name	Distance (bp)
Weaning weight										
1	7	64,120,586	64,934,489	BTB-00317489	64,536,896	G/A	A(0.27)	−0.05	*GRIA1*	40,884
2	16	47,191,344	48,227,271	ARS-BFGL-NGS-39382	47,648,555	A/G	G(0.30)	0.04	*PLCH2*	within
3	15	51,974,721	53,245,382	Hapmap60878-rs29012474	52,710,098	A/C	C(0.40)	0.04	*RAB6A*	within
4	11	82,506,706	83,918,981	ARS-BFGL-BAC-11115	82,938,260	T/C	C(0.28)	−0.04	*MSGN1*	11,224
5	8	72,487,663	73,067,269	ARS-BFGL-NGS-38434	72,487,663	C/G	G(0.33)	0.03	*NUDT18*	within
Yearling weight										
1	4	60,684,925	61,594,363	Hapmap43661-BTA-70769	61,087,638	T/C	C(0.46)	0.03	*LOC100141099*	162,878
2	13	20,842,651	22,328,437	ARS-BFGL-NGS-90391	21,411,575	C/T	T(0.31)	−0.01	*NEBL*	within
3	18	10,115,183	11,038,104	ARS-BFGL-NGS-56801	10,472,709	C/T	T(0.48)	−0.01	*MGC140224*	37,170
4	7	7,153,567	8,549,081	Hapmap59438-rs29012637	7,215,131	A/G	G(0.36)	0.01	*LOC520939*	51,772
5	18	51,626,670	52,162,143	ARS-BFGL-NGS-56579	51,887,350	C/T	T(0.42)	−0.01	*ZNF226*	142
Carcass weight										
1	1	159,392,014	160,186,915	ARS-BFGL-NGS-87492	159,738,525	C/A	A(0.43)	0.05	*LOC100139888*	260,980
2	14	24,675,437	25,364,673	Hapmap27935-BTC-065354	25,031,801	A/G	G(0.48)	0.03	*TOX*	within
3	14	35,637,412	36,497,204	BTB-01640837	36,095,273	G/T	T(0.43)	0.04	*LOC509345*	443,377
4	18	18,069,608	18,938,560	ARS-BFGL-NGS-39866	18,183,364	A/G	G(0.39)	−0.05	*CYLD*	within
5	27	25,498,209	25,760,531	BTB-00119427	25,651,351	G/A	A(0.48)	−0.02	*DLC1*	61,060
Backfat thickness										
1	6	545,275	1,496,086	BTB-01384704	959,692	A/G	A(0.48)	−0.01	*LOC100139637*	63,038
2	9	9,439,530	11,271,075	BTA-28590-no-rs	10,295,707	G/A	A(0.24)	−0.01	*bta-mir-30a*	406,847
3	29	32,524,535	33,401,150	BTB-01020342	33,014,346	A/G	G(0.49)	−0.01	*ETS1*	459,952
4	5	108,980,836	111,370,382	BTB-00236217	110,984,106	T/C	C(0.41)	0.01	*CD9*	2,769
5	22	18,261,833	19,187,625	ARS-BFGL-NGS-84780	18,637,206	T/C	T(0.48)	0.01	*LMCD1*	134,899
***Longissimus dorsi* muscle area**										
1	17	28,104,077	29,183,585	Hapmap39898-BTA-18764	28,708,255	G/T	T(0.39)	−0.07	*LOC784783*	909,550
2	14	24,675,437	25,364,673	Hapmap27935-BTC-065354	25,031,801	A/G	G(0.48)	0.07	*TOX*	within
3	3	71,841,366	72,372,751	BTB-01097190	72,372,751	T/C	C(0.41)	−0.06	*ST6GALNAC5*	within
4	1	159,468,529	159,767,156	ARS-BFGL-NGS-87492	159,738,525	C/A	A(0.43)	0.05	*LOC100139888*	260,980
5	16	40,624,954	41,447,343	BTB-00640892	41,392,873	C/T	T(0.46)	0.05	*MIR34A*	33
Marbling score										
1	1	159,392,014	160,152,519	ARS-BFGL-NGS-87492	159,738,525	C/A	A(0.43)	0.18	*LOC100139888*	260,980
2	27	25,240,234	25,858,456	ARS-BFGL-NGS-11088	25,677,257	G/A	G(0.40)	0.14	*SEMA5A*	53426
3	21	24,546,255	25,415,220	BTB-00811262	24,753,371	C/T	T(0.42)	−0.15	*BTBD1*	within
4	4	4,247,357	4,746,954	BFGL-NGS-112064	4,485,873	A/C	A(0.49)	−0.14	*COBL*	109,664
5	14	35,637,412	36,200,511	BTB-01640837	36,095,273	G/T	T(0.43)	0.15	*LOC509345*	443,377

QTL, quantitative trait loci; SNP, single nucleotide polymorphism; MAF, minor allele frequency.

[1,2,3] SNP marker annotations and their positions were based on the bovine reference genome (btau4.0).

[4] Nucleotides of substitution. [5] Estimates of additive effects. [6] The nearest known gene to the significant SNP.

other QTL effects. This was also explained by Xu (2003) and Sun (2011). Therefore, here we attempted to minimize the number of false positives by combined these three methods result.

The establishment of a threshold is a complicated issue and has a profound influence on the results, especially when methods with different nature of scores are compared. False discovery rate derived threshold takes into account the number of tests that are performed as well as how significant one test is relative to the others in multiple comparison procedures. Two levels of significant controls were used in LDRM and LDVCM analysis based on genome-wide (FDR<0.05) and chromosome-wide (FDR<0.01) type I errors, which were computed for all SNP

(Fernando et al., 2004). Usually, significance tests are not required in Bayesian analysis; only frequentists emphasize significance tests. However, to facilitate comparison with other methods, we accumulated the effects of adjacent SNPs together into a genomic window and then chose the top 5 as cut-off the criterion to detect QTL in BayesCπ. Figure 2 clearly shows the major peaks in BayesCπ coincide with LDRM and LDVCM. Therefore, these present findings together with those of Sun et al. (2011) suggest that the genetic variances estimated by BayesCπ including all markers without polygenic effects can exploit the generated gene discovery knowledge.

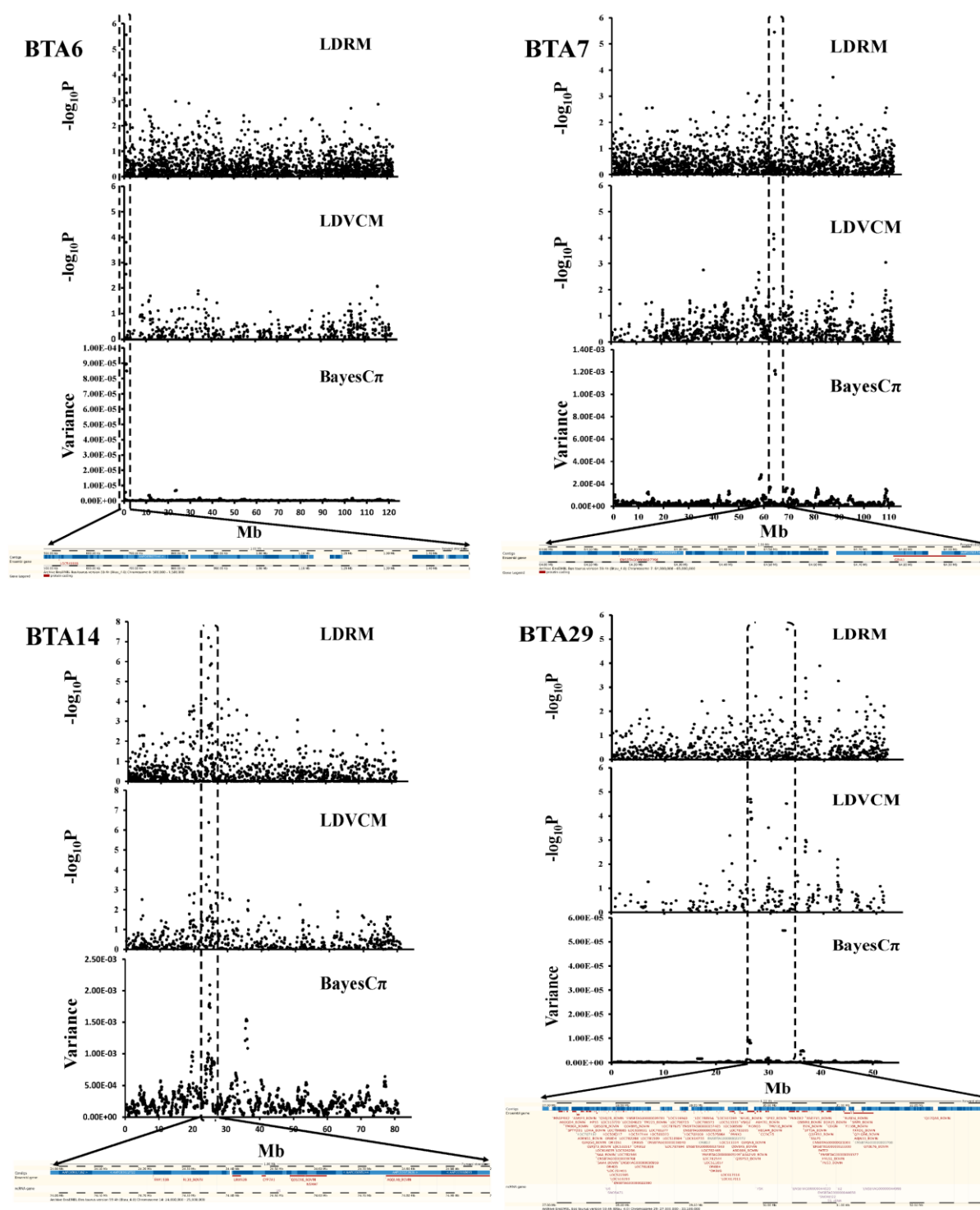Table 1 and 3 show that the SNP effects explained by BayesCπ were smaller than those by LDRM. These result

Figure 2. QTL profile (upper) and annotation (lower) of the detected SNPs associated with BFT on BTA6, WWT on BTA7, CWT on BTA14 and BFT on BTA29. QTL, quantitative trait loci; SNPs, single nucleotide polymorphisms; BFT, backfat thickness; BTA, *Bos taurus* autosome; WWT, weaning weight; CWT, carcass weight.

were not surprising because we performed LDRM in the way of SNP by SNP individually via regressing the traits on the genotypes of a SNP and thus ignored all possible SNP-covariate or SNP-SNP pairs for interaction effect, leading to

the model effect acting as the effect of the marker, particularly with small mapping population (Xu, 2003; Hoggart et al., 2008). Additionally, the SNP effect can be overestimated in LDRM, especially when the effect of a

SNP is small, this is due to the so called Beavis effect (Xu, 2003). Therefore, the work in this study is more related to QTL position rather than the precise estimation of their effects.

Our study successfully uncovered many variants related to QTLs. However, the traits YWT, CWT, and BFT measured in Hanwoo have QTL with larger additive effects than WWT, LMA, and Marb. Several explanations have been proposed for our scant outcomes from genetic markers. First, the currently identified SNPs might not fully describe genetic diversity. For instance, these SNPs may not capture some forms of genetic variability that are due to copy number variation. Second, genetic mechanisms might involve complex interactions among genes and between genes and environmental conditions, or epigenetic mechanisms which are not fully captured by additive models. However, opportunities may exist for improving predictions by exploiting additive genetic variation. A third explanation—the one we focus on here—lies in the limitations posed by the genetic models and statistical methods.

CONCLUSION

No linkage disequilibrium mapping method tested in this study outperformed the others, but it is interesting that the positions of the SNP selected by the different models (LDRM, LDLA, BayesCπ) were close. Some well-known candidate genes for the traits of interest were located close to these QTL. Four major additive QTL in each analysis had high concordance including 64.1 to 64.9 Mb for BTA7 for WWT, 24.3 to 25.4 Mb for BTA14 for CWT, 0.5 to 1.5 Mb for BTA6 for BFT and 26.3 to 33.4 Mb for BTA29 for BFT. We suggest the use of combined different LD mapping approaches can provide more reliable chromosome regions to further pinpoint DNA makers or causative genes in these regions.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

REFERENCES

Blott, S., J. J. Kim, S. Moisio, A. Schmidt-Kuntzel, A. Cornet, P. Berzi, N. Cambisano, C. Ford, B. Grisart, D. Johnson, L. Karim, P. Simon, R. Snell, R. Spelman, J. Wong, J. Vilkki, M. Georges, F. Farnir, and W. Coppieters. 2003. Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. Genetics 163:253-266.

Cierco-Ayrolles, C., S. Dejean, A. Legarra, H. Gilbert, T. Druet, F. Ytournel, D. Estivals, N. Oumouhou, and B. Mangin. 2010. Does probabilistic modelling of linkage disequilibrium evolution improve the accuracy of QTL location in animal pedigree? Genet. Sel. Evol. 42:38.

Daetwyler, H. D., R. Pong-Wong, B. Villanueva, and J. A. Woolliams. 2010. The impact of genetic architecture on genome-wide evaluation methods. Genetics 185:1021-1031.

Druet, T. and M. Georges. 2010. A hidden markov model combining linkage and linkage disequilibrium information for haplotype reconstruction and quantitative trait locus fine mapping. Genetics 184:789-798.

Erbe, M., F. Ytournel, E. C. G. Pimentel, A. R. Sharifi, and H. Simianer. 2011. Power and robustness of three whole genome association mapping approaches in selected populations. J. Anim. Breed. Genet. 128:3-14.

Fernando, R. L. and D. Garrick. 2013. Bayesian methods applied to GWAS. Methods Mol. Biol. 1019:237-274.

Fernando, R. L., D. Nettleton, B. R. Southey, J. C. M. Dekkers, M. F. Rothschild, and M. Soller. 2004. Controlling the proportion of false positives in multiple dependent tests. Genetics 166:611-619.

Garrick, D. J. 2011. The nature, scope and impact of genomic prediction in beef cattle in the United States. Genet. Sel. Evol. 43:17.

Grapes, L., J. C. M. Dekkers, M. F. Rothschild, and R. L. Fernando. 2004. Comparing linkage disequilibrium-based methods for fine mapping quantitative trait loci. Genetics 166:1561-1570.

Habier, D., R. L. Fernando, K. Kizilkaya, and D. J. Garrick. 2011. Extension of the bayesian alphabet for genomic selection. BMC Bioinformatics 12:186.

Hoggart, C. J., J. C. Whittaker, M. De Iorio, and D. J. Balding. 2008. Simultaneous analysis of all SNPs in genome-wide and re-sequencing association studies. PLoS Genet. 4(7):e1000130.

Kim, J. J. and M. Georges. 2002. Evaluation of a new fine-mapping method exploiting linkage disequilibrium: A case study analysing a QTL with major effect on milk composition on bovine chromosome 14. Asian Australas. J. Anim. Sci. 15:1250-1256.

Lee, Y. S., J. H. Lee, J. Y. Lee, J. J. Kim, H. S. Park, and J. S. Yeo. 2008. Identification of candidate SNP (single nucleotide polymorphism) for growth and carcass traits related to QTL on chromosome 6 in Hanwoo (Korean Cattle). Asian Australas. J. Anim. Sci. 21:1703-1709.

Li, C., J. Basarab, W. M. Snelling, B. Benkel, J. Kneeland, B.

Murdoch, C. Hansen, and S. S. Moore. 2004. Identification and fine mapping of quantitative trait loci for backfat on bovine chromosomes 2, 5, 6, 19, 21, and 23 in a commercial line of *Bos taurus*. J. Anim. Sci. 82:967-972.

Li, Y. 2012. Genome-wide association study to identify QTLfor growth and carcass quality traits in Korean native cattle, Hanwoo. Ph.D. Thesis, Yeungnam University, Gyeongsan, Gyeongbuk, Korea.

Li, Y., J. H. Lee, Y. M. Lee, and J. J. Kim.2011. Application of linkage disequilibrium mapping methods to detect QTL for carcass quality on chromosome 6 using a high density SNP map in Hanwoo. Asian Australas. J. Anim. Sci. 24:457-462.

MacLeod, I. M., B. J. Hayes, K. W. Savin, A. J. Chamberlain, H. C. McPartlan, and M. E. Goddard. 2010. Power of a genome scan to detect and locate quantitative trait loci in cattle using dense single nucleotide polymorphisms. J. Anim. Breed. Genet. 127:133-142.

Matukumalli, L. K., C. T. Lawley, R. D. Schnabel, J. F. Taylor, M. F. Allan, M. P. Heaton, J. O'Connell, S. S. Moore, T. P. L. Smith, T. S. Sonstegard, and C. P. Van Tassell. 2009. Development and characterization of a high density SNP genotyping assay for cattle. PloS One 4(4):e5350.

McClure, M. C., N. S. Morsci, R. D. Schnabel, J. W. Kim, P. Yao, M. M. Rolf, S. D. McKay, S. J. Gregg, R. H. Chapple, S. L. Northcutt, and J. F. Taylor. 2010. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. Anim. Genet. 41:597-607.

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819-1829.

Meuwissen, T. H. and M. E. Goddard. 2001. Prediction of identity by descent probabilities from marker-haplotypes. Genet. Sel. Evol. 33:605-634.

Misztal, I. 2011. FAQ for genomic selection. J. Anim. Breed. Genet. 128:245-246.

Mizoshita, K., A. Takano, T. Watanabe, A. Takasuga, and Y. Sugimoto. 2005. Identification of a 1.1-Mb region for a carcass weight QTL onbovine Chromosome 14. Mamm. Genome 16:532-537.

Mizoshita, K., T. Watanabe, H. Hayashi, C. Kubota, H. Yamakuchi, J. Todoroki, and Y. Sugimoto. 2004. Quantitative trait loci analysis for growth and carcass traits in a half-sib family of purebred Japanese Black (Wagyu) cattle. J. Anim. Sci. 82:3415-3420.

Morris, C. A., W. S. Pitchford, N. G. Cullen, A. K. Esmailizadeh, S. M. Hickey, D. Hyndman, K. G. Dodds, R. A. Afolayan, A. M. Crawford, and C. D. K. Bottema. 2009. Quantitative trait loci for live animal and carcass composition traits in Jersey and Limousin back-cross cattle finished on pasture or feedlot. Anim. Genet. 40:648-654.

NIAS. 2009. Annual Research Report http://www.nias.go.kr/ Accessed August 24, 2011.

Perez-Enciso, M. and I. Misztal. 2011. Qxpak.5: Old mixed model solutions for new genomics problems. BMC Bioinformatics 12:202.

Sun, X., D. Habier, R. L. Fernando, D. J. Garrick, and J. C. M. Dekkers. 2011. Genomic breeding value prediction and QTL mapping of QTLMAS2010 data using Bayesian Methods. BMC Proceedings 5(Suppl 3):S13.

Stone, R. T., J. W. Keele, S. D. Shackelford, S. M. Kappes, and M. Koohmaraie. 1999. A primary screen of the bovine genome for quantitative trait loci affecting carcass and growth traits. J. Anim. Sci.77:1379-1384.

Uleberg, E. and T. H. E. Meuwissen. 2010. Fine mapping and detection of the causative mutation underlying quantitative trait loci. J. Anim. Breed. Genet. 127:404-410.

Xu, S. 2003. Estimating polygenic effects using markers of the entire genome. Genetics 163:789-801.

Xu, S. 2003. Theoretical basis of the Beavis effect. Genetics 165:2259-2268.

Zhao, H. H., R. L. Fernando, and J. C. M. Dekkers. 2007. Power and precision of alternate methods for linkage disequilibrium mapping of quantitative trait loci. Genetics 175:1975-1986.