

ZB MED - Informationszentrum Lebenswissenschaften

12th Day of the Doctoral Student

Seyfert, Hans-Martin

Dummerstorf, 2011

Bayes-method for mapping imprinted QTL in an inbred F2-design - Ulrike
Borchardt and Norbert Reinsch

urn:nbn:de:hbz:38m:1-30093

Bayes-method for mapping imprinted QTL in an inbred F₂-design

Ulrike Borchardt¹ and Norbert Reinsch¹

¹Leibniz Institut for Farm Animal Biology, Dummerstorf

1 Introduction

Most QTL studies in pigs involve additive and dominant effects but imprinting is often ignored. Genomic imprinting refers to a phenomenon in which the same genes are expressed differently, depending on their parental origin. It has been shown; that between 5% and 19% for traits connected to growth and body components of the total additive genetic variance was controlled by imprinted loci [Neugebauer]. To find additive, dominance and imprinted QTL we used a multiple QTL model estimating all effects simultaneously in a single model. We implemented an efficient and adapted Bayesian method with genomic imprinting to estimate all effects (based on [Xu]). This method is also an essentially model-selection-free method. Small effects get a small weight so that their inclusion should have negligible effects on the final results.

2 Materials and methods

Statistical methods:

For an F₂ population the mixed model is:

$$y_i = \mu + V_i d + \sum_{j=1}^p (X_{ij}a_j + W_{ij}b_j + Z_{ij}c_j) + e_i,$$

Let y_i for $i = 1, \dots, n$ be the phenotypic value of the i th animal, μ is the population mean, p is the number of possible QTL. V_i is the design matrix of the fixed effects d . X_{ij} , W_{ij} and Z_{ij} are dummy variables, which coded the genotype of the j th QTL and i th individual. The QTL effects of QTL j are a_j (additive), b_j (dominant) and c_j (imprinting). E_j is the residual error, distributed as $N(0, \sigma^2)$.

To estimate the additive, dominant and imprinting effects, we applied the model-selection-free method of (Xu). The multiple marker analysis and the linear model were extended by imprinting and fixed effects. Xu's method uses that each marker effect gets a normal prior

with mean zero and a unique variance. The effect-specific prior variance was estimated from the data.

X_{ij} , W_{ij} and Z_{ij} are defined for genotype A_1A_1 : $\sqrt{2}$, -1 and 0 . For a genotype A_1A_2 : 0 , 1 , $\sqrt{2}$, for A_2A_1 : 0 , 1 , $-\sqrt{2}$ and for A_2A_2 : $-\sqrt{2}$, -1 and 0 . The coded variables are independent and have a zero expectation and a unity variance. The regression coefficients are defined as $a_j = G_{11} - G_{22}$ for additive effects, $b_j = G_{12} + G_{21} - G_{11} - G_{22}$ for dominant effects and $c_j = G_{12} - G_{21}$ for imprinting effects, where G_{11} , G_{12} , G_{21} and G_{22} are the four genotypic values.

The likelihood function is the distribution of the observables (phenotypes and marker information) and the function of the unobservables (effects and variances). The prior distribution is the distribution of the unobservables. In the MCMC-implemented Bayesian analysis, we sample the unobservables from the above joint posterior distribution.

Animals and analyzed traits

In this study, we used a total of 645 F_2 pigs, derived from Duroc and Pietrain founder animals. All animals were kept at the Frankenhorst experimental research farm at the university of Bonn (Germany). The phenotypes of all F_2 pigs were recorded in an abattoir, according the guidelines of German performance stations. In total 26 traits were analyzed. On the one hand, the traits describe carcass compositions and on the other hand the meat quality.

125 microsatellites and 6 SNP markers were used to genotype the animals and covered 89% of these. The CRI-MAP 2.4 software recalculates the gender-average linkage map. Gencheck and Pedcheck (version 1.1) checked genotypic errors. More detailed information can be found in (Liu) and (Grosse-Brinkhaus).

3 Results

In this study, we have 131 markers on the 18 autosomes. We calculated genotype probabilities for 244 possible QTL positions. A large part of these are uncertain genotypes. For the mixed model we use different fixed effects like birth or slaughter months, family, gender, carcass weight and age. As default initial values all QTL effects were set equal zero and all variances equal to 0.5. The length of Markov chain is 110000 sweeps. 10000 sweeps were deleted (burn-in period). The observations were added and the posterior mean and variances of QTL effects are calculated.

In the Bayesian analysis markers with small effects are shrunken towards zero. We performed several analyses to check the results. We see that the peaks swing around special points. To have significance thresholds to distinguish QTL from non-QTL, we randomly shuffled the phenotypes (permutation test) [Sillanpää].

Table 1: Significant effects

Chromosome	Markers	Phenotype	Additive effect	Dominance effect	Imprinting effect
SSC 1	SW1851	Ph _{24h} ham	-0.021		
	SW2166	dressing			0.207
	S0113 – SWR2182	Ph ₄₅ loin		-0.005	
		Ph decline		-0.019	
		dressing			-0.260
		Ph _{24h} ham			0.014
SSC 3	S0164	Thawing loss (g)		0.0513	
SSC 5	SW1134 – S0005	Conductivity 24h ham			1.970
	S0005	Conductivity 24h ham			-1.940
SSC 6	S0087 – SW1067	Estimated belly lean content (%)	0.148		
		Ratio of fat to muscle area	-0.002		
		Loin eye area above M. long. dorsi (cm ²)	1.056		
		Bonner formula (%)	0.547		
	S0300 - CKM	Estimated belly lean content (%)			-0.147
		Loin eye area above M. long. dorsi (cm ²)			-0.078
		Bonner formula (%)			-0.478
		Backfat loin (cm)			0.060
	S0220 – S0059	Backfat shoulder (cm)	-0.081		
	S0059 – S0003	Cooking loss	0.509		
SSC 7	S0102 – SW175	Carcass length (cm)	-0.028		
SSC 13	SW344	Conductivity 45 min loin		-0.012	
SSC 14	SWC27	Drip loss		0.437	-0.659
SSC 17	SW840 – SW2431	Side fat thickness (cm)	0.0288		
		Ratio of fat to muscle area	0.0009		
		Backfat loin (cm)		-0.143	
		Estimated belly lean content (%)	-0.414		

Table 1 is sorted by chromosome and position within each chromosome. Additionally, the significant additive, dominance, and imprinting effects of each QTL are listed. A total of 26 QTL for 17 of the 26 traits were found to be significant at 5% genome-wise levels.

Additive effects are often on chromosome 6 and 17. Dominant effects have a cluster on chromosome 5 and 16. But many possible QTL effects are under the threshold (especially carcass composition traits). Similar to the dominant effects the imprinting effects on chromosome 2 are under the significant threshold. Especially the fat components between markers SW1879 and SW1844 are present. The second cluster on imprinting effects is on chromosome 6. The phenotypes dressing and Conductivity 24h ham have opposite direction effect in a small interval and either indicate two QTL with imprinting effects in repulsion are an artefact due to large monte carlo errors.

4 Discussion

Simultaneous QTL-mapping is very time expansive, because imprinting genotypes are uncertain. In every iteration many uncertain genotypes have to be imputed. The comparison of results with other studies is not easy, because the special F₂ population of Duroc and Pietrain and the genome-wise level (most paper have chromosome-wise level and detected more QTL). In direct comparison with the same population and single-QTL-models we detected the same QTL on chromosome 1 ph 24h ham and loin, on chromosome 6 the estimated belly lean content and on chromosome 7 the carcass length [Liu]. We see that our analysis yields similar QTL position in comparison with other studies (additive and dominance effects). Therefore we can accept the imprinting effects (positions), especially the cluster on chromosome 2 (not significant) and 6. Genomic imprinting can be interpreted as an interaction effect between alleles and parental origin with locus. Identification of interactions between different loci – additive-additive, additive-dominant, dominant-additive, dominant-dominant – is even more demanding from a computational point of view.

5 Literature

Liu G, Jennen DG, Tholen E, Juengst H, Kleinwächter T, Hölker M, Tesfaye D, Un G, Schreinemachers HJ, Murani E, Ponsuksili S, Kim JJ, Schellander K, Wimmers K. : A genome scan reveals QTL for growth, fatness, leanness and meat quality in a Duroc-Pietrain resource population. *Animal Genetics* 38; 241-52.

Grosse-Brinkhaus C, Jonas E, Buschbell H, Phatsara C, Tesfaye D, Jüngst H, Looft C, Schellander K, Tholen E. : Epistatic QTL pairs associated with meat quality and carcass composition traits in a porcine Duroc × Pietrain population. *Genet Sel Evol.* 26;42:39.

Neugebauer N., Luther H., Reinsch N. : Parent-of-origin effects cause genetic variation in pig performance traits. *Animal*

Sillanpää, M.A., Arjas E., 1998: Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. *Genetics* 148; 1373-1388

Xu, S., 2002: Estimating polygenic effects using markers of the entire genome. *Genetics* 163; 789-801

Vergleich der Güte verschiedener Pooling-Designs bei Tests auf differentiell exprimierte Gene

Henrik Rudolf, Norbert Reinsch

Leibniz-Institut für Nutztierbiologie Dummerstorf, Forschungsbereich Genetik und Biometrie

1 Einleitung

In Genexpressionsstudien ist Pooling, das Mischen der mRNA mehrerer Individuen, eine weit verbreitete Methode. Wissenschaftler greifen auf dieses Mittel zurück, um unzureichende RNA-Mengen vor der Hybridisierung auszugleichen, oder die Kosten des Versuchs durch eine Verringerung der Anzahl der Arrays zu senken. Peng *et al.* (2003) schlussfolgern, dass optimale, gepoolte Designs wählbar sind, so dass die Einhaltung statistischer Voraussetzungen einhergeht mit einer Reduktion der Kosten der Versuche. Zhang *et al.* (2007) gelang es, die theoretisch positive Differenz zwischen den zu erwartenden Expressionswerten eines Pools und einem Mittel der beteiligten Individuen (Jensensche Ungleichung) auf der Log-Skala zu approximieren, wie auch die Varianz von Pools gleicher Größe. Bisher wurde empfohlen, während eines Versuches die Poolgröße nach Möglichkeit nicht zu verändern. Erstens, weil dies zu Verzerrungen in den Parametern führt und außerdem die Varianz-Homogenität in den Daten verletzt.

Bei Versuchen in denen nach Tieren mit seltenen Eigenschaften gesucht wird, ist jedoch vor Beginn nicht klar wie viele Subjekte zu dieser Gruppe gehören. Daher ist bei fest vorgegebener Anzahl von Arrays flexibles Pooling mit Größen entsprechend der gefundenen Individuen je Versuchsdurchgang wünschenswert. Wir analysieren diese Pooling Experimente und behandeln die bereits erwähnten Nachteile. Dazu wird im Design jedem Pool einer Behandlung ein gleichgroßer Pool von Kontrollen gegenüber gestellt. Es lässt sich zeigen, dass dann die Bedingung für unverfälschte Kontraste bei bestimmten Strukturen der Kovarianzmatrix erfüllt ist. Zur Untersuchung der neuen Methode in der Analyse haben wir Experimente, die die beschriebenen Voraussetzungen an Daten und Kovarianz umsetzen, simuliert. Im Folgenden werden erste Ergebnisse präsentiert, welche anhand der Güte der Tests Rückschlüsse auf Einsparmöglichkeiten bei der Zahl der Arrays gestatten.