Modeling QTL Effect On BTA6 In Dairy Cattle Using A Random Regression Test Day Model

T. Suchocki[*], Q. Zhang[†] and J. Szyda[*]

Introduction

The genetic variation of quantitative traits is attributed to a combined effect of several loci with large effects (quantitative trait loci, QTL) and many loci with a small effect each, which in statistical models are jointly treated as a polygenes effect. Models exist which incorporate effects of both, a QTL and polygenes as constant in time (Fernando and Grossman (1989)). However, for most of the dairy populations detailed information of the observed yield variation across a whole lactation is available in a form of test day yields. Many statistical models for test day yields have been proposed. One of the first test day models for dairy cattle was developed by Ptak and Schaeffer (1993). In this model, a so called fixed regression model, an additive genetic variance is assumed to be constant across time. Further development comprised a random regression model, which assumes that an additive genetic variance is variable across lactation (Schaeffer and Jamrozik (1996)). However, none of those models considers effects of quantitative trait loci (QTL), and genetic background is modeled only by the polygenic effect. The major goal of this study is to incorporate a QTL effect into the model, as well as to estimate its position and effect on milk yield based on test day records sampled from Chinese Holstein-Friesian dairy cows.

Material and methods

The real data originates from Chinese Holstein-Friesian dairy cattle population and consists of 26 paternal halfsib families (1 231 daughters of 26 sires). The number of daughters per sire varies between 5 and 119, with the average of 43. All animals were selected from nine herds in the region of Beijing. These herds represent the largest dairy cattle farms in the area with an average 305–day milk yield of about 8 500 kg (Chen et al. (2006)). All individuals have marker information consisting of fourteen microsatellite markers on chromosome 6 covering 63.5 cM. For each individual test day yields for milk yield are available. For each animal the number of test day records varies between five and ten.

In this study two models were considered. The simplest model was applied to 305–day yields and both random effects - an additive polygenic and a QTL are modeled as constant in time:

$$y_i = X_i\beta + \alpha_i + q_i + \epsilon_i$$

---

[*] Institute of Animal Genetics, Wroclaw University of Environmental and Life Sciences, Kozuchowska 7, 51 - 631 Wroclaw, Poland

[†] Department of Animal Genetics and Breeding, China Agricultural University, Beijing 100193, China

where y_i is a 305–day lactation yield of cow i; β is a vector of fixed nongenetic effects comprising an overall mean, herd and calving year; $\alpha = [\alpha_1, \ldots, \alpha_n] \sim \mathcal{N}(0, A\sigma^2_\alpha)$ is a random additive polygenic effect of cows with A representing an additive polygenic relationships among cows and σ^2_α being a component of the additive genetic variance attributed to polygenes; $q = [q_1, \ldots, q_n] \sim \mathcal{N}(0, Q\sigma^2_q)$ is a random QTL effect of a cows with Q representing IBD (Identity By Descent) matrix estimated using Citius (Szydlowski and Gengler (2008)) and σ^2_q being the QTL variance; $\epsilon = [\epsilon_1, \ldots, \epsilon_n] \sim \mathcal{N}(0, I\sigma^2_\epsilon)$ is a residual with σ^2_ϵ denoting the error variance; I representing an identity matrix and n is a total number of animals.

The second model was applied to test day yields. Consequently, an additive polygenic effect is now variable in time and a permanent environmental effect describing variation in environment during lactation is added. Legendre orthogonal polynomials of order three were used for modeling of both time dependent effects. Model has the form:

$$y_{ij} = X_i\beta + \sum_{m=0}^{3} \alpha_{im}\phi_m(\tau_{ij}) + q_i + \sum_{m=0}^{3} p_{im}\phi_m(\tau_{ij}) + \epsilon_{ij}$$

where y_{ij} is the jth test day record of ith animal; $\alpha \sim \mathcal{N}(0, A \otimes G_\alpha)$ with G_α representing the covariance matrix of random regression coefficients for an additive polygenic term and $p \sim \mathcal{N}(0, I \otimes G_p)$ is a permanent environmental effect, with G_p representing the covariance matrix of random regression coefficients for a permanent environmental effect; $\phi_m(\tau)$ is a Legendre polynomial of order m at time point τ, while other effects are the same as above. Since the orthogonal Legendre polynomials are defined for an interval $[-1, 1]$, the observed days in milk t were rescaled to τ using:

$$\tau = 2 \times \frac{t - t_{min}}{t_{max} - t_{min}} - 1$$

where t_{min} is the lowest (5) and t_{max} the highest value (305) of days in milk t.

For each of the two models likelihood profiles were generated in steps every 1 cM. For the assessment of the most likely QTL position the likelihood ratio test (Λ) was used. The corresponding hypotheses were $H_0 : \sigma^2_q = 0$ and $H_1 : \sigma^2_q > 0$. The test statistic is given by:

$$\Lambda = -2\frac{\ln L(\widehat{\theta}_0)}{\ln L(\widehat{\theta})}$$

where $L(\widehat{\theta})$ is the likelihood function obtained at $\widehat{\theta} = [\widehat{\beta}, \widehat{\sigma}^2_\alpha,\ \widehat{\sigma}^2_q, \widehat{\sigma}^2_\epsilon]$ and $L(\widehat{\theta}_0)$ is the likelihood function obtained at $\widehat{\theta}_0 = [\widehat{\beta}, \widehat{\sigma}^2_\alpha, 0, \widehat{\sigma}^2_\epsilon]$ for model one and respectively at $\widehat{\theta} = [\widehat{\beta}, \widehat{G}_\alpha, \widehat{\sigma}^2_q,\ \widehat{G}_p, \widehat{\sigma}^2_\epsilon]$ and at $\widehat{\theta}_0 = [\widehat{\beta}, \widehat{G}_\alpha, 0, \widehat{G}_p, \widehat{\sigma}^2_\epsilon]$ for model two. Under H_0, Λ follows asymptotically a mixture of chi square distributions: $0.5\chi^2_0 + 0.5\chi^2_1$. To address a multiple testing issue, Bonferroni corrections were applied to nominal p values.

Results and Conclusion

Figure 1 shows p values corresponding to Λ obtained by model one. A peak corresponding to $p = 0.059$ after a multiple testing correction is observed near marker BMS470 at 32 cM.

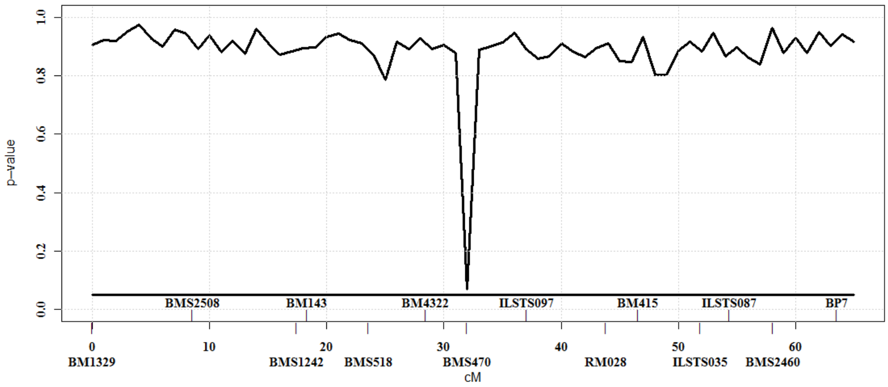

Figure 1: **p values corresponding to Λ for model one, corrected for multiple testing.**

Figure 2 represents four peaks in the vicinity of markers BMS470 (33 cM), ILSTS097 (37 cM), RM028 (43 cM) and ILSTS035 (52 cM) respectively. But only the first peak is statistically significant after Bonferroni correction.

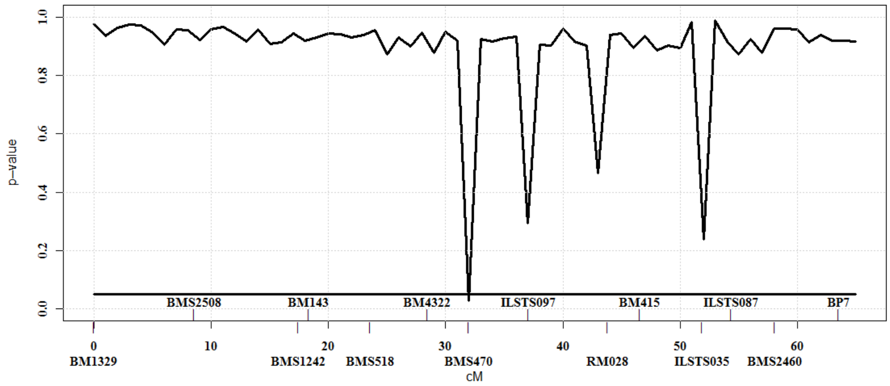

Figure 2: **p values corresponding to Λ for model two, corrected for multiple testing.**

A QTL was detected by each of the models at a similar position near marker BMS470 (Table 1). However, for model two one QTL was statistically significant after Bonferroni correction and three other QTLs were significant without any correction, while for model one it only reached borderline significance of 0.0590.

Table 1: Estimated QTL positions and p - values corresponding to models one and two.

model 1			model 2		
QTL position (cM)	Corrected p - value	Nominal p - value	QTL position (cM)	Corrected p - value	Nominal p - value
32	0.0590	0.0009	33	0.0290	0.0004
			37	0.2939	0.0045
			43	0.4733	0.0072
			52	0.3116	0.0047

QTLs for milk yield in the vicinity of markers BMS470, ILSTS097 and RM028 were already reported in previous studies (Table 2.).

Table 2: QTL on milk yield in the interval between markers BMS470 and RM028, reported by other studies.

p - value	References	Peak
< 0.05	Chen et al. (2006)	BMS470
< 0.01	Gao et al. (2009)	BMS470
0.00041	Szyda et al. (2005)	BM4322 (28.47cM) - ILSTS097
< 0.01	Zhang et al. (1998)	BM143 (18.33cM) - RM028

Comparing the two used models we find that more complex structure of parameters (model 2) gave better results in QTL detection. We found one significant QTL and three probable positions of QTLs, which may be found using other models (QTL effect modeled as variable in time). In model one only one possible QTL position was found.

References

Chen, H., Zhang, Q., Yin, C. C., Wang, C. K., Gong, W. J., , and Mei, G. (2006). *J. Dairy Sci.*, 89:782–790.

Fernando, R. L. and Grossman, M. (1989). *Genet. Sel. Evol.*, 21:467–477.

Gao, H., Fang, M., LIU, J., and Zhang, Q. (2009). *Heredity*, 103:368–376.

Ptak, E. and Schaeffer, L. R. (1993). *Livest. Prod. Sci.*, 34:23–34.

Schaeffer, L. R. and Jamrozik, J. (1996). *J. Dairy Sci.*, 79:2044–2055.

Szyda, J., Liu, Z., Reinhardt, F., and Reents, R. (2005). *J. Dairy Sci.*, 88:356–367.

Szydlowski, M. and Gengler, N. (2008). *Journal of Animal Breeding and Genetics*, 125:330–338.

Zhang, Q., Boichard, D., Hoeschele, I., Ernst, C., Eggen, A., Murkve, B., M., P.-G., Witte, L. A., Grignola, F. E., Uimari, P., Thaller, G., and Bishop, M. D. (1998). *Genetics*, 149:1959–1973.