57th Annual Meeting of the European Association for Animal Production
Antalya (Turkey), 17th to 20th September 2006 – Session 3 (G3), Abstract G3.16
Kathrin-Friederike.Stock@tiho-hannover.de

Estimation of genetic parameters for categorical, continuous and molecular genetic data in multivariate animal threshold models using Gibbs sampling

K.F. Stock[1], O. Distl[1] and I. Hoeschele[2]*

[1]Institute for Animal Breeding and Genetics, University of Veterinary Medicine Hannover, 30559 Hannover, Germany, [2]Virginia Bioinformatics Institute and Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg 24061, Virginia, USA

Abstract

Simulated populations included 7 generations and 40000 animals per generation. Fixed effects, residual and additive genetic variances for one continuous trait (T1) and liabilities for four categorical traits (T2 to T5) were simulated. QTL effects were simulated for T2 with recombination rates (r) of 0.00 or 0.01 and polymorphism information content (PIC) of markers of 0.9 or 0.7. Simulated heritabilities (h^2) were 0.50 (T1), 0.25 (T3, T5) and 0.10 (T2, T4). Simulated additive genetic correlations (r_g) were ±0.20. After dichotomization trait prevalences were 0.25 (T2, T5) and 0.10 (T3, T4). Phenotypes of 10000 animals from one generation (P1) or phenotypes and genotypes of 5000 animals and their parents (G2) were used for multivariate estimations. Most biased parameters were r_{g12} and r_{g14} in P1 (-33% to -55%) and h^2_2 and r_{g45} in G2 (+36% to +52%). Correlations between true and predicted breeding values (BV) for the categorical traits did mostly not differ significantly between P1 and G2. Selection on the basis of BV from P1 was significantly less effective than selection on the basis of genotype and BV from G2. Selection response was significantly lower with PIC=0.7 and r=0.00 than with PIC=0.9 and r=0.00 or r=0.01.

Introduction

Many important traits in animal breeding are categorical and often binary, violating the basic assumptions for genetic analysis in mixed linear models. Transformation which can be used to compensate for underestimation of linear estimates of heritabilities and residual correlations of non-linear traits may introduce relevant bias, especially in case of extreme prevalences (e.g. Abdel-Azim and Berger, 1999; Mäntysaari et al., 1991; Van Vleck and Gregory 1992). The magnitude of bias depends on data and pedigree structure and needs to be checked individually for any analyzed dataset. Transformation and check of transformation

reliability can be avoided by (co)variance component estimation in threshold or mixed linear-threshold models. Implementation of such models is possible via Markov chain Monte Carlo methods like Gibbs sampling. As opposed to sire models, animal models fully use all available pedigree information, but implementation of animal threshold models is not always straightforward. Particularly with low trait prevalences and few observations per animals, problems with accuracy of (co)variance component estimates and convergence of the Gibbs chain may arise (Hoeschele and Tier, 1995; Luo et al., 2001; Moreno et al., 1997).

In the horse, binary coding has been used for radiographic health traits, and high prevalences of radiologically visible alterations, mostly in the range of 10 to 25 percent, have been determined in the limbs of young Warmblood riding horses (Stock and Distl, 2006a, b; Willms et al., 1999; Winter et al., 1996). Because strength and soundness of the locomotory system is of great importance in all sectors of the horse industry and genetics were found to play a significant role for the development of radiographic findings, inclusion of radiographic health traits in the current breeding schemes has been suggested (Stock and Distl, 2005a, b). Reliably estimated genetic parameters provide the basis to do so. Relevant genetic correlations between the categorical traits and between the categorical traits on the one hand and linear body measures on the other hand imply genetic analyses in mixed linear-threshold animal models. Selection for radiographic health of the equine limbs should benefit from genetic evaluation in the threshold model (Meijering and Gianola, 1985; Matos et al., 1997). Furthermore, increasing knowledge on the molecular genetic determination of radiographic findings in the equine limbs (Böneker et al., 2006; Dierks and Distl, 2006) gives rise to the question how to optimally use phenotype and genotype information on radiographic health traits in genetic analyses and for selection.

The aim of this study was to characterize the properties of multivariate estimation of genetic parameters and prediction of breeding values for categorical, continuous and molecular genetic data using linear-threshold animal models and Gibbs sampling. On the basis of simulated data the impact of data structure and quality of molecular genetic marker information on the accuracy of genetic parameter estimates and predicted breeding values and the expected response to selection was investigated in the context of important radiological health traits in the Warmblood horse.

Material and methods

Data simulation

Simulated data were used for this study, with simulation parameters chosen according to the results of previous studies on radiographic findings in the limbs of Warmblood riding horses.

Simulation included fixed, residual and additive genetic effects for one continuous trait (T1) and liabilities of four categorical traits (T2 to T5), and QTL effects for the liability of one of the categorical traits (T2). Additive genetic effects (a) were normally distributed with $a \sim N(0, A\sigma^2_{ai})$ and i = 1, 2, ..., 5, genetic variances (σ^2_a) were set to 4.8 (T1) or 1.0 (T2 to T5), and genetic covariances were chosen such that additive genetic correlations (r_g) were $r_{g12} = r_{g13} = 0.20$ and $r_{g14} = r_{g45} = -0.20$. Additive genetic effects of offspring ($a_{offspring}$) were derived from additive genetic effects of their parents (a_{sire}, a_{dam}) and Mendelian sampling term (m) as $a_{offspring} = 0.5\ (a_{sire} + a_{dam}) + m$ with $m \sim N(0, 0.5A\sigma^2_{ai})$. Residual effects were normally distributed $e \sim N(0, I\sigma^2_{ei})$ with residual variances (σ^2_e) such that heritabilities (h^2) were $h^2_1 = 0.50$, $h^2_2 = h^2_4 = 0.10$, and $h^2_3 = h^2_5 = 0.25$. A fixed contemporary group effect was simulated with five levels per generation and two levels each represented in two subsequent generations.

For T2 two QTL and two flanking markers per QTL with five alleles each were simulated, with one of the marker alleles being linked to the unfavorable QTL allele, i.e. the allele increasing the probability of T2. Marker alleles were randomly distributed and equally prevalent, and total QTL variance was set equal to the additive genetic variance. In order to study the effects of different quality of genetic marker information on the estimation of genetic parameters, three scenarios were simulated: no recombination between genetic markers and QTL, polymorphism information content (PIC) of 0.9 of all markers (r0p9); recombination rate of 0.01 between markers and QTL, polymorphism information content (PIC) of 0.9 of all markers (r1p9); no recombination between genetic markers and QTL, polymorphism information content (PIC) of 0.7 of all markers (r0p7). After simulation on the linear scale liabilities of the categorical traits were dichotomized to obtain trait prevalences of 0.25 (T_2, T_5) or 0.10 (T_3, T_4).

Each of the three simulated populations included 280,000 animals, evenly distributed over 7 generations and with a male to female ratio of 1:1 in each generation. Per generation 9,000 females and 400 males were randomly chosen as parents of the next generation. Each dam was randomly mated to five sires, and each sire was randomly mated to five (200 sires), 150 (160 sires) or 500 dams (40 sires). Replicates (n = 10) were generated by drawing random samples of 10,000 animals from the fourth generation and tracing their pedigree back over three generations. For the genetic analyses, two different datasets were created within each replicate. Dataset P1 included all 10,000 animals with records for the continuous trait and the four binary traits, information on the fixed effects of sex and contemporary group, and pedigree information over three generations. Dataset G2 included 5,000 animals, randomly chosen from the animals included in dataset P1, plus their parents with respective information

on traits, sex, contemporary group and pedigree, and additional information on the marker genotype of the animals.

Genetic analyses

Genetic parameters were estimated and breeding values were predicted using Gibbs sampling with the threshold version of the Multiple Trait Gibbs Sampler for Animal Models (MTGSAM) (Van Tassell and Van Vleck, 1996), a software which supports multivariate genetic analyses of any combination of continuous and categorical traits. Random and residual effects are assumed to be normally distributed, and flat priors are used for the fixed effects. For our analyses, we chose a starting value of one for all additive genetic variances and a starting value of zero for all additive genetic covariances, and we fixed the residual covariances between all traits to zero. For the binary traits, residual variances were fixed to one and thresholds were fixed to zero in order to ensure identifiability of the model (Harville and Mee, 1984). A proper prior using an inverse Wishart distribution with minimum shape parameter (i.e. $\nu_{IW} = 7$) was adopted for the genetic (co)variance matrix in order to ensure posterior propriety. The fixed effects of sex and contemporary group were considered in all analyses. The fixed effect of marker genotype was considered in the analysis of dataset G2 only, distinguishing between individuals homozygous negative for the unfavorable alleles of all genetic markers, individuals heterozygous for the unfavorable allele of at least one of the genetic markers, individuals homozygous for the unfavorable allele of at least one of the four genetic markers.

$y_{ijlm} = \mu + SEX_i + CONT_j + a_l + e_{ijlm}$ (P1), and

$y_{ijklm} = \mu + SEX_i + CONT_j + QTL_k + a_l + e_{ijklm}$ (G2),

with y_{ijlm} (y_{ijklm}) = observation on trait T1 (continuous) or on trait T2, T3, T4 or T5 (binary) for the l^{th} animal,

μ = model constant,

SEX_i = fixed effect of the sex of the animal (i = 1-2),

$CONT_j$ = fixed effect of the contemporary group (j = 1-5 for P1; j = 1-8 for G2),

QTL_k = fixed effect of the QTL marker genotype (n = 1-3),

a_l = random additive genetic effect of the l^{th} animal (l = 1-30533 to 30766 for P1; l = 1-26253 to 26664 for G2), and

e_{ijlm} (e_{ijklm}) = random residual.

The total length of the Gibbs chain was set to 205000 in all analyses, and all samples after 5000 rounds of burn-in were saved. Convergence of the Gibbs chain and the need for additional rounds of burn-in to be discarded was checked by visual inspection of sample plots.

Effective sample size (ESS) and Monte Carlo error (MCE) was calculated for all (co)variance estimates by the times series method implemented in the post-Gibbs analysis program POSTGIBBSF90 (Tsurata, 2005) with a thinning rate of ten. Un-thinned chains were used to calculate posterior means of additive genetic (co)variance, heritability and additive genetic correlation estimates. Bias of heritability and additive genetic correlation estimates was calculated as the mean relative deviation of the estimated values (par_{est}) from the true, i.e. simulated, values (par_{true}).

$$bias = (par_{est} - par_{true}) / par_{true}$$

Breeding values and selection

True and predicted breeding values (BV) were compared in the sires of those 5,000 animals with trait records, which were included in both datasets of each replicate, using Pearson correlation coefficients and the procedure CORR of the Statistical Analysis System (SAS), version 9.1.3 (SAS Institute, Cary, NC, USA, 2005). True and predicted relative breeding values (RBV) were derived from true and predicted breeding values (BV) by standardization on a relative scale with a mean of 100 and a standard deviation of 20, using the sires of animals with trait records as the reference population for the standardization. Larger RBV for the continuous trait indicate genetic predisposition for higher values, lower RBV genetic predisposition for lower values of this trait. RBV for the binary traits were transformed so that larger RBV will mean that the animals are less likely and lower RBV will mean that the animals are more likely to transmit a predisposition for the particular trait.

The response to selection with focus on the QTL trait was studied using either the RBV for T2 or the marker genotype or the RBV for T2 and the marker genotype as selection criteria. Response to selection was defined as relative decrease of the prevalences of the binary traits in the offspring of the selected sires compared to the offspring of all sires. Selected sires needed to have an above-average RBV for T2 ($RBV_{T2} > 100$) and/or to be homozygous negative for the unfavorable allele of all genetic markers. Only sires which were represented by at least 10 offspring with trait records were considered for selection. In each case and within each replicate, the expected response to selection was assessed by comparing the prevalences of the binary traits in the offspring of the selected sires and in all 5,000 animals with trait records, which were included in both datasets.

Evaluation

The influence of data structure and quality of genetic marker information on ESS, bias and correlation between true and predicted breeding values and the influence of the selection

criteria on the response to selection was tested via analysis of variance using the procedure GLM of Statistical Analysis Systems, (SAS), version 9.1.3 (SAS Institute, Cary, NC, USA, 2005). Effective sample size, bias, correlation between true and predicted breeding values or response to selection was considered as dependent variable, and dataset (P1, G2) and quality of genetic marker information (r0p9, r1p9, r0p7) or selection criteria (RBV_{T2}, marker genotype, RBV_{T2} and marker genotype) were considered as fixed effects.

Results

Convergence of the Gibbs chains was achieved after maximally 30,000 rounds of Gibbs sampling, leaving 175,000 to 200,000 rounds for analyses of posterior distributions. Mean, minimum and maximum ESS of heritabilities and selected additive genetic correlations by dataset and quality of genotype information on T2 are given in Table 1. ESS was significantly larger in the analyses of dataset G2 than in the analyses of dataset P1 for heritabilities of traits T1, T2, T4 and T5 and additive genetic correlation between T1 and T3 ($P \leq 0.02$). The opposite was true with respect to the additive genetic correlation between T1 and T2 ($P < 0.01$). No significant influence on ESS was determined for the quality of the genetic marker information. Mean MCE was 0.002 to 0.010 for the heritabilities and 0.001 to 0.004 for the additive genetic correlations in all analyses.

Mean, minimum and maximum bias of heritabilities and additive genetic correlations by dataset and quality of genotype information on T2 are given in Table 2. Mean bias was largest for additive genetic correlations between T1 and T2 and between T1 and T4 in analyses of dataset P1 (-0.33 to -0.55) and for heritability of T2 and additive genetic correlation between T4 and T5 in analyses of dataset G2 (0.36 to 0.52). The analyzed dataset had a significant influence on the bias of the heritability estimates for T1 and T2 ($P < 0.001$) and of the estimated additive genetic correlations between T1 and T2 and between T1 and T4 ($P < 0.01$). Bias of heritability estimate for T2 was further significantly dependent on the quality of genetic marker information ($P < 0.01$), with lower means in scenario r0p7 than in scenarios r0p9 and r1p9.

Correlations between true and predicted BV for traits T1 to T5 by dataset and quality of genotype information on T2 and for all sires of offspring with trait records and sires with ten or more offspring with trait records are given in Table 3. For all traits correlation coefficients increased with increasing number of informative offspring per sire. The dataset used for the genetic analyses had neither in all sires with informative offspring nor in the sires with ten or more informative offspring a significant influence on BV correlations for T2. In most cases BV correlations were higher in analyses of dataset P1 than in analyses of dataset G2.

However, BV correlations for T2 were significantly influenced by the quality of the genetic marker information, with significantly higher correlation coefficients in scenarios r0p9 and r1p9 than in scenario r0p7 ($P < 0.001$).

Relative changes of prevalences of binary traits T2 to T5 after selection of sires for T2 by dataset and quality of genotype information on T2 are given in Table 4. Response to selection with respect to T2 was larger after selection on the basis of RBV for T2 from analyses of dataset P1 (-0.14 to -0.12) than after selection on the basis of polygenic RBV for T2 from analyses of G2 (-0.09) or marker genotype (-0.07 to -0.02). The prevalence of T2 was most effectively lowered, if both RBV for T2 from analyses of dataset G2 and marker genotype served as selection criteria ($P < 0.001$). Selection response for T2 was significantly influenced by the quality of genetic marker information, with larger prevalence decreases in scenarios r0p9 and r1p9 than in scenario r0p7 ($P < 0.001$).

Conclusions

Feasibility of multivariate estimation of genetic parameters and prediction of breeding values in mixed linear-threshold animal models using Gibbs sampling with data and pedigree structures similar to those encountered in the Warmblood horse has been shown. If genes or genome regions which do not fit in the polygenic model have been identified, combined use of phenotype and genotype information can increase the reliability of genetic analyses and the response to selection.

References

Abdel-Azim GA, Berger PJ (1999). Properties of threshold model predictions. Journal of Anim Sci 77: 582-590.

Böneker C, Löhring K, Wittwer C, Distl O (2006). Genome-wide search for markers associated with osteochondrosis in Hanoverian Warmblood horses. Submitted.

Dierks C, Distl O (2006). Identification of a quantitative gene locus on equine chromosome 4 responsible for osteochondrosis in hock joints of Hanoverian Warmblood horses. Submitted.

Harville, D.A., Mee, R.W. (1984). A mixed-model procedure for analyzing ordered categorical. Biometrics 40: 393-408.

Hoeschele I, Tier B (1995). Estimation of variance components of threshold characters by marginal posterior modes and means via Gibbs Sampling. Genet Sel Evol 27: 519-540.

Luo MF, Boettcher PJ, Schaeffer LR, Dekkers JCM (2001). Bayesian inference for categorical traits with an application to variance component estimation. J Dairy Sci 84: 694-704.

Mäntysaari EA, Quaas RL, Gröhn YT (1991). Simulation study on covariance component estimation for two binary traits in an underlying continuous scale. J Dairy Sci 74: 580-591.

Matos, C. A. P., Thomas, D. L., Gianola, D., Tempelman, R. J. and Young, L. D. (1997). Genetic analysis of discrete reproductive traits in sheep using linear and nonlinear models. I. Estimation of genetic parameters. J Anim Sci 75: 76-87.

Meijering, A. and Gianola, D. (1985). Linear versus nonlinear methods of sire evaluation for categorical traits: a simulation study. Genet Sel Evol 17: 115-132.

Moreno C, Sorensen D, García-Cortés LA, Varona L, Altarriba, J (1997). On biased inferences about variance components in the binary threshold model. Gen Sel Evol 29: 145-160.

Stock KF, Distl O (2005a). Prediction of breeding values for osseous fragments in fetlock and hock joints, deforming arthropathy in hock joints and pathologic changes in navicular bones of Hanoverian Warmblood horses. Livest Prod Sci 92: 77-94.

Stock KF, Distl O (2005b). Expected response to selection when accounting for orthopedic health traits in a population of Warmblood riding horses. Am J Vet Res 66: 1371-1379.

Stock KF, Distl O (2006). Genetic correlations between osseous fragments in fetlock and hock joints, deforming arthropathy in hock joints and pathologic changes in the navicular bones of Warmblood riding horses. Livest Prod Sci, in press (available online since 24 July 2006, doi:10.1016/j.livsci.2006.04.027).

Stock KF, Distl O (2006). Genetic analyses of radiographic appearance of the navicular bones in the Warmblood horse. Am J Vet Res 67, 1013-1019.

Van Tassell, C. P. and Van Vleck, L. D. (1996). Multiple-trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood-based (co)variance component inference. J Anim Sci 74: 2586-2597.

Tsurata, S. (2005). POSTGIBBSF90 [http://nce.ads.uga.edu/~ignacy/newprograms.html].

Van Vleck LD, Gregory KE (1992). Multiple-trait restricted maximum likelihood for simulated measures of ovulation rate with underlying multivariate normal distributions. J Anim Sci 70: 57-61.

Willms F, Röhe R, Kalm E (1999). Genetische Analyse von Merkmalskomplexen in der Reitpferdezucht unter Berücksichtigung von Gliedmaßenveränderungen. 1. Mitteilung: Züchterische Bedeutung von Gliedmaßenveränderungen. Züchtungskunde 71: 330-345.

Winter D, Bruns E, Glodek P, Hertsch B (1996). Genetische Disposition von Gliedmaßenerkrankungen bei Reitpferden. Züchtungskunde 68: 92-108.

Table 1. Mean (minimum, maximum) effective sample size of estimated heritabilities (h^2) and selected additive genetic correlations (r_g) for the continuous trait T1 and the binary traits T2 to T5 by dataset (P1, G2) and for different scenarios with respect to recombination rate (r) between markers and QTL and polymorphism information content (PIC) of markers; values for r = 0.00 and PIC = 0.9 in the first line, for r = 0.01 and PIC = 0.9 in the second line, and for r = 0.00 and PIC = 0.7 in the third line.

Genetic parameter		Dataset	
		P1	G2
Heritabilities	h^2_1	641.7 (533.5, 778.6)	1620.5 (1417.3, 1753.0)
		652.5 (504.5, 829.9)	1640.8 (1330.0, 1881.2)
		628.1 (371.6, 858.4)	1655.6 (1487.9, 1949.0)
	h^2_2	93.8 (36.4, 117.9)	102.2 (62.8, 129.2)
		86.5 (59.6, 119.8)	105.2 (83.6, 120.9)
		69.8 (42.7, 100.7)	100.1 (75.5, 120.5)
	h^2_3	79.0 (25.9, 113.0)	84.8 (58.9, 102.4)
		68.7 (30.7, 105.5)	81.2 (42.3, 117.4)
		83.9 (49.6, 113.7)	91.6 (54.1, 141.7)
	h^2_4	53.4 (31.5, 77.2)	82.0 (46.2, 115.2)
		57.2 (34.3, 108.0)	82.8 (64.6, 114.6)
		76.2 (50.8, 102.5)	70.9 (37.0, 91.5)
	h^2_5	131.1 (71.7, 179.7)	207.1 (159.8, 284.2)
		121.8 (79.5, 167.8)	198.4 (100.8, 289.6)
		129.3 (95.9, 174.1)	227.1 (141.5, 274.9)
Additive genetic correlations	r_{g12}	435.7 (306.8, 639.8)	314.6 (240.9, 451.0)
		464.0 (191.0, 701.5)	337.0 (172.1, 445.7)
		347.2 (222.0, 482.4)	346.4 (257.6, 474.2)
	r_{g13}	383.8 (221.8, 497.4)	430.9 (201.5, 650.8)
		384.6 (75.2, 662.9)	459.5 (238.8, 652.6)
		353.4 (181.9, 523.0)	449.4 (194.5, 723.8)
	r_{g14}	257.1 (127.9, 371.0)	281.5 (158.0, 389.1)
		291.2 (191.9, 514.9)	284.2 (111.2, 419.9)
		278.3 (170.7, 470.7)	245.5 (172.3, 334.4)
	r_{g45}	164.7 (83.0, 236.4)	145.9 (112.3, 182.1)
		159.8 (91.2, 221.1)	164.5 (117.2, 223.2)
		163.6 (123.8, 241.2)	156.8 (104.1, 205.1)

Table 2. Mean (minimum, maximum) relative bias of estimated heritabilities (h^2) and selected additive genetic correlations (r_g) for the continuous trait T1 and the binary traits T2 to T5 by dataset (P1, G2) and for different scenarios with respect to recombination rate (r) between markers and QTL and polymorphism information content (PIC) of markers; values for r = 0.00 and PIC = 0.9 in the first line, for r = 0.01 and PIC = 0.9 in the second line, and for r = 0.00 and PIC = 0.7 in the third line.

Genetic parameter		Dataset	
		P1	G2
Heritabilities	h^2_1	0.061 (-0.013, 0.169)	-0.064 (-0.089, -0.034)
		0.060 (-0.014, 0.165)	-0.064 (-0.087, -0.034)
		0.059 (-0.010, 0.168)	-0.064 (-0.089, -0.033)
	h^2_2	0.068 (-0.321, 0.460)	0.506 (0.233, 0.706)
		0.078 (-0.330, 0.490)	0.519 (0.208, 0.696)
		-0.134 (-0.414, 0.171)	0.467 (0.219, 0.672)
	h^2_3	0.013 (-0.218, 0.176)	-0.057 (-0.218, 0.067)
		0.019 (-0.205, 0.167)	-0.055 (-0.177, 0.049)
		0.024 (-0.185, 0.180)	-0.045 (-0.178, 0.082)
	h^2_4	0.160 (-0.143, 0.434)	0.087 (-0.135, 0.238)
		0.141 (-0.162, 0.390)	0.080 (-0.140, 0.240)
		0.150 (-0.201, 0.418)	0.078 (-0.131, 0.242)
	h^2_5	-0.018 (-0.143, 0.088)	0.068 (-0.065, 0.147)
		-0.010 (-0.156, 0.087)	0.068 (-0.061, 0.157)
		-0.015 (-0.143, 0.081)	0.066 (-0.071, 0.156)
Additive genetic correlations	r_{g12}	-0.503 (-0.958, 0.040)	-0.082 (-0.417, 0.637)
		-0.488 (-0.925, 0.014)	-0.066 (-0.381, 0.624)
		-0.331 (-0.908, 0.131)	-0.101 (-0.447, 0.501)
	r_{g13}	-0.056 (-0.335, 0.360)	-0.017 (-0.449, 0.438)
		-0.060 (-0.320, 0.366)	-0.019 (-0.436, 0.404)
		-0.065 (-0.330, 0.315)	-0.021 (-0.440, 0.376)
	r_{g14}	-0.494 (-0.904, 0.026)	0.008 (-0.416, 0.410)
		-0.484 (-0.916, -0.007)	0.010 (-0.378, 0.470)
		-0.473 (-0.825, 0.031)	-0.004 (-0.405, 0.456)
	r_{g45}	0.127 (-0.598, 0.753)	0.364 (-0.813, 0.923)
		0.140 (-0.567, 0.760)	0.369 (-0.834, 1.033)
		0.163 (-0.652, 0.795)	0.376 (-0.806, 0.960)

Table 3. Correlations between true and predicted breeding values for the continuous trait T1 and the binary traits T2 to T5, in all sires with offspring with trait records and sires with 10 or more offspring with trait records for different scenarios with respect to recombination rate (r) between markers and QTL and polymorphism information content (PIC) of markers in the analyses of datasets A1 to C2; values for r = 0.00 and PIC = 0.9 in the first line, for r = 0.01 and PIC = 0.9 in the second line, and for r = 0.00 and PIC = 0.7 in the third line.

Trait	All sires (n = 273-303, n_{off} = 17.43)		Sires with ≥10 offspring (n = 193-198, n_{off} = 24.78)	
	A1	C2	A1	C2
T1	0.849	0.870	0.935	0.914
	0.849	0.870	0.934	0.914
	0.849	0.870	0.935	0.914
T2	0.514	0.484	0.587	0.535
	0.513	0.483	0.585	0.531
	0.417	0.479	0.490	0.532
T3	0.635	0.626	0.728	0.671
	0.634	0.625	0.728	0.670
	0.634	0.623	0.727	0.667
T4	0.481	0.469	0.525	0.476
	0.480	0.469	0.524	0.476
	0.481	0.471	0.525	0.478
T5	0.664	0.645	0.752	0.700
	0.664	0.645	0.752	0.700
	0.665	0.646	0.753	0.700

n: number of sires; n_{off}: average number of offspring per sire.

Table 4. Relative changes of prevalences of binary traits T2 to T5 after selection of sires for T2 based on the relative breeding value for this trait (RBV_{T2},), on genotype or on genotype and RBV_{T2}, for different scenarios with respect to recombination rate (r) between markers and QTL and polymorphism information content (PIC) of markers using predicted breeding values from analyses of datasets P1 and G2; values for r = 0.00 and PIC = 0.9 in the first line, for r = 0.01 and PIC = 0.9 in the second line, and for r = 0.00 and PIC = 0.7 in the third line.

Prediction basis	Selection basis	Trait			
		T2	T3	T4	T5
P1	RBV_{T2}	-0.135	-0.065	+0.029	+0.032
		-0.134	-0.071	+0.030	+0.037
		-0.119	-0.070	+0.018	+0.053
G2	RBV_{T2}	-0.085	-0.124	-0.004	+0.034
		-0.090	-0.127	+0.003	+0.042
		-0.092	-0.123	-0.006	+0.045
	RBV_{T2}, marker genotype	-0.167	-0.146	-0.021	+0.103
		-0.164	-0.150	-0.013	+0.106
		-0.120	-0.144	-0.024	+0.119
	marker genotype	-0.072	+0.030	+0.008	+0.021
		-0.071	+0.030	+0.010	+0.021
		-0.022	+0.030	+0.008	+0.021