



Genome-wide association studies using BayesC and estimation of genetic parameters for perinatal sucking reflex in Brown Swiss calves

C. Dreher¹, P. Stratz¹, R. Wellmann¹, S. Preuß¹, H. Hamann², J. Bennewitz¹

¹ Institute of Animal Science, University of Hohenheim, Germany ² State Office for Spatial Information and Land Development Baden-Wuerttemberg, Germany

> WCGALP 2018 Biology- Growth and Development





Introduction

- Calves need to ingest colostrum shortly after birth (<6h) to gain important immunoglobulines for passive immunisation
- ~ 10 % of Italian Brown Swiss calves show weak sucking reflex (SR) (Maltecca et al., 2007) \rightarrow $h^2 = 0.12 0.26$
- Project in cooperation with local breeding organisation in South Germany:
- ➔ Quantitative genetic and genomic analyses of weak sucking reflex (SR) in South-German Brown Swiss calves





Data collection

- 06/2015 07/2016
- Phenotype documentation in categorial traits by farmers
 - > 10,000 calves on ~ 240 farms
- Target trait SR recorded in 4 categories (none, weak, normal, strong)
- Genotype collection (50K SNP chip) through tissue samples from eartags by staff members of the breeding organisation
 - > 5,000 calves





Genotype imputation

- 777 K SNP chip genotype data: **192** calves
 - Pair of at least 2 calves from same farm, ideally same sire, divergent SR
- Imputation of 50 K SNP data to 777 K SNP data
 - High imputation accuracies





WCGALP 2018 Biology- Growth and Development





Trait encoding and incidence for datasets

	Sucking reflex	None	Weak	Normal	Strong	Sum
Encoding	Model 1	1	2	3	4	
	Model 2	1	0	0	0	
	Model 3	1	1	0	0	
Number of	Phenotypes	373	511	2,448	518	3,850
calves	Genotypes	420	467	1,320	354	2,561



Environmental factors

- Herd
- Season/ Halfyear
- Gestation length (linear and squared)
- Sex (only for genotyped individuals)









Variance component estimation

Sire-Threshold-Model using R-package MCMCgImm (Hadfield, 2010)

l = Xb + Za + e

- *l*: Latent variable for SR *y* using probit-link function
- X, Z: Designmatrices related to b, a
- *b*: Vector of fixed effects with flat prior (herd, halfyear, gestation length)
- *a*: Random effects of sire $a | \sigma_a^2 \sim N(0, \sigma_a^2 A)$ with $\sigma_a^2 \sim inv gamma(0.01, 0.01)$
- e: Residual effects $e \sim N(0, I)$
- 100,000 Iterations; Burn-in 10,000; thinning interval 100





Results variance component estimation

Model	Encoding of sucking reflex	h^2	I-95% CI	u-95% CI
Model 1	1 (none), 2 (weak), 3 (normal), 4 (strong)	0.12	0.012	0.261
Model 2	1 (none), 0 (weak, normal, strong)	0.10	0.015	0.223
Model 3	1 (none, weak) , 0 (normal, strong)	0.08	0.011	0.178





Single-Marker-GWAS using GCTA (Yang et al., 2011)

$$y = Xb + g + e$$

- *y:* Precorrected SR, encoded in 3 different models
- *X:* Designmatrix related to *b*
- b: Fixed additive effect of tested SNP
- *g*: Random cumulative effects of SNPs, with $g \sim N(0, \sigma_g^2 \mathbf{G})$
- e: Residual effects





GWAS Results (Single-Marker-Model)



- Most significant SNPs found in ordinal model 1
- Candidate genes in comparison with ensembl.org data base on BTA 6 and 19





Multi-Marker-GWAS (BayesC)

$$y = \mu \mathbf{1} + Xa + e$$

- *y*: Precorrected SR
- μ: Intercept
- **a**: Random SNP effects, with $a_j | \gamma_j \sim \gamma_j t(0, v, s^2) + (1 \gamma_j) \varepsilon t(0, v, s^2)$

$$P(\gamma_j = 1) = 0.2$$

- X: Designmatrix related to a
- *e*: Residual effects, with $e \sim N(0, \sigma_e^2 I)$





GWAS BayesC

- R package BayesDsamples (Wellmann and Bennewitz, 2012)
- 50,000 Iterations, 25,000 Burn-In
- Calculation of window posterior probability of association (WPPA)

(Fernando and Garrick, 2013) from every 25th sample

- Windowsize 1 Mb
- Degrees of freedom 2.5





GWAS Results (WPPA)



 \rightarrow Less than 40 % probability that a window shows association





Discussion & Conclusion

• Low h² (0.8 - 0.12) for sucking reflex

→ long-term process and continuous phenotype recording is necessary for a promising breeding strategy against a weak sucking reflex

- Polygenic nature of sucking reflex
 - \rightarrow genetic tests are not expedient
 - \rightarrow routine phenotype recording and genomic selection is essential





Discussion & Conclusion

- Very low WPPA results, but window on BTA 19 explains 10 % of genetic variance
- No stringent correction of multiple testing for Single-Marker-GWAS, but still identification of 2 candidate genes related to metabolic processes and nerve cell growth
- Effects of farm have a huge impact on SR
 - \rightarrow more research needed on particular influence on sucking reflex





Take home message

- SR in Brown Swiss calves is a new trait with polygenic nature and low h²
- Breeding against weak SR is a long-term process wich implies routine phenotype recording and genomic selection
- Phenotypes for SR should be captured in ordinal scaling
- Despite a promising breeding strategy, the effects of farm should not be neglected, meaning a good management practice on farms is most essential





Acknowledgements

We thank the *Förderverein Bioökonomieforschung e.V.*, the breeding organization *Rinderunion Baden-Württemberg e.V.* for good cooperation. The project was funded by the *Landwirtschaftliche Rentenbank*, Frankfurt, Germany.













Thank you for your attention!





References

Fernando, R. L., and D. Garrick. 2013. Bayesian methods applied to GWAS. Methods in molecular biology (Clifton, N.J.) 1019:237–274.

Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: The MCMCgImm R package. Journal of Statistical Software 33(2):1–22.

Maltecca, C., A. Rossoni, C. Nicoletti, E. Santus, K. A. Weigel, and A. Bagnato. 2007. Estimation of genetic parameters for perinatal sucking behavior of Italian Brown Swiss calves. Journal of dairy science 90(10):4814–4820.

Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011. GCTA: A Tool for Genomewide Complex Trait Analysis. The American Journal of Human Genetics 88(1):76–82.





Herd effects



herd





Quality control criteria for genotype data

- Call Rate ≥ 0.99
- SNPs must be in HWE
- Minor Allele Frequency ≥ 0.03
- HD-SNPs with divergent genotype to 50K SNP data were excluded



23



UNIVERSITY OF

Candidate genes

BTA 6

•FRYL (FRY like transcription coactivator)

 \rightarrow Protein coding gene

 \rightarrow Plays key role in cell morphogenesis and

in patterning sensory neuron dendritic fields

 \rightarrow May function as a transcriptional activator









Candidate genes

BTA 19

•NPEPPS (Aminopeptidase puromycin sensitive)

→Protein coding gene

 \rightarrow Plays a role in the innate immune system

 \rightarrow May be involved in proteolytic events

regulating the cell cycle







VCE of 777K SNP data using GCTA (Yang et al., 2011)

 $y = Xb + g + \varepsilon$

- y: trait of interest (SR), encoded in 3 different ways (ordinal, binary*)
- b: vector of fixed effects (sex, herd, season)
- X: Designmatrix related to b
- g: vector of total additive genetic effects of the individuals, with $g \sim N(0, A\sigma_G^2)$
- ε : vector of residual effects, with $\varepsilon \sim N(0, I\sigma_e^2)$

*Transformation of estimate of variance from observed scale to underlying scale Biology- Growth and Development





Results VCE

Model	Encoding of sucking reflex	h^2	I-95% CI	u-95% Cl	$h_g^2 \; (se_g)$
Model 1	1 (no), 2 (weak), 3 (normal), 4 (strong)	0.12	0.012	0.261	0.07 (0.02)
Model 2	1 (no), 0 (weak, normal, strong)	0.10	0.015	0.223	0.10 (0.05)
Model 3	1 (no, weak) , 0 (normal, strong)	0.08	0.011	0.178	0.09 (0.04)



Plots Model 1







Plots Model 2





Density of var1



Plots Model 3









Cross validation results for imputation of SNP data







Farm phenotype sheet







Distribution of SNPs for 50K and 777K Data





BTA





Data collection

- Farm questionnaires (~ 200 farms)
 - Feeding practice
 - \rightarrow Mineral contents for cows in lactation and dry period
 - \rightarrow Kind of colostrum offering
 - Housing practice
 - \rightarrow Access to grassland
 - \rightarrow Size and kind of calving pen





Data editing

- Phenotype data: 10,637 calves
 - no crossbreed calves
 - no calves with missing information for SR, birth date, sex or gestation length of mother or missing sire
 - only calves from farms with more than 50 animals in the data and calves in every trait category
 - > 3850 calves left from 52 farms and 176 sires





Data editing

- Genotype data:
- 50 K SNP chip genotype data: 3,072 calves
 - no calves with missing or incorrect phenotype information
 - no calves with malformations or abnormalities
- ➤ 2,775 calves





Datasets

	phenotype	50K	777K
	data	genotype data	genotype data
number of calves left	3,850	2,775	2,561





A-priori-Verteilung

SNP-Effect α_m

• Kommt aus einer von zwei unterschiedlich skalierten t-Verteilungen.

 $V_G = V_A$, mit V_A (Schätzer aus GCTA)

→ Skalierungsparameter (s²): $s_{small}^2 = 0.01 * s_{big}^2$ und $E(V_G) = V_{\widehat{G}}$

- Wahrscheinlichkeit, dass ein SNP im LD mit einem QTL ist (aus der Verteilung mit der größer skalierten Verteilung) : *pLD* = 0,2
- Freiheitsgrade(v): v = 2.5





Sire-Threshold-Model using MCMCgImm

- uninformative prior
 - Prior distribution of variance: inverse-Gamma distribution, parametrized by parameters nu and V (nu=0.02,V=1)

• 250,000 Iterations, 10,000 Burn-In, thinning interval 100