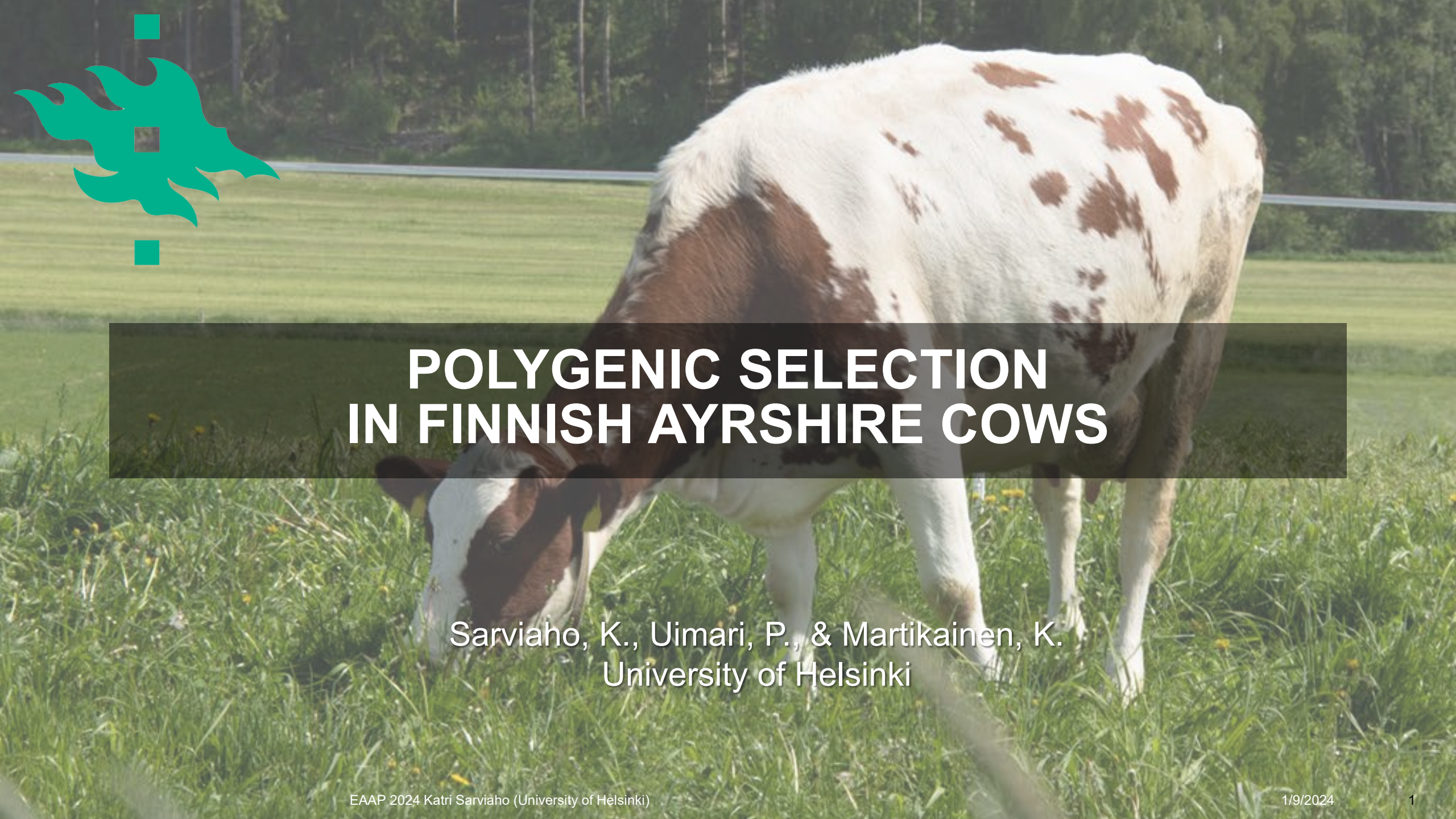

POLYGENIC SELECTION IN FINNISH AYRSHIRE COWS

Sarviaho, K., Uimari, P., & Martikainen, K.
University of Helsinki

INTRODUCTION

FINNISH AYRSHIRE

- Second most common dairy breed in Finland
- Joint evaluation of breeding values in Finland, Sweden, and Denmark by NAV[1]
  - Genomic information included in 2011
  - Genomic estimates for AI-bulls and all genotyped females
  - Selection based on Nordic Total Merit

[1] NAV 2024

GENERATION PROXY SELECTION MAPPING (GPSM)[1]

- Associate SNPs under selection to given proxy, e.g., birth date
  - Detect SNP allele frequency changes
    - Subsequent generations
  - Distinguish polygenic selection from drift
- Identifies SNPs responding to selection
- Accounts for population structure with genomic relationship matrix (GRM)
- Applied e.g., beef cattle and pigs but not in the FAY

[1] Decker et al., 2012; Rowan et al., 2021; Grohmann et al., 2023

AIM

The aim of this study was to discern selection signatures from genomic changes emerging due to random drift in the genome-selected Finnish Ayrshire.

MATERIAL AND METHODS

MATERIAL

- 74,135 FAY genotyped for 45,834 autosomal SNPs
- Animals born in 2009–2020
  - 9,975 bulls
  - 64,140 cows
- SNPs limited to minor allele frequency (MAF) >0.01[1]
  - 43,462 SNPs in bull data
  - 43,641 SNPs in cow data

[1] Purcell & Chang, 2021

METHODS – GPSM (1/2)

- Variable AGE (from Jan 2009 to the birth month of the individual, in full years)
- Proportion of variance (PVE) in AGE explained by genome-wide SNPs[1]
  - y = μ + **Zg** + **e**, where

    y = vector of AGE

    μ = mean AGE

    Z = matrix relating AGE in **y** to **g**

    g = vector of random polygenic effects (uses GRM)[2]

    e = vector of random residual effects
  - $PVE = \frac{\sigma_g^2}{\sigma_g^2+\sigma_e^2}$

[1] Yang et al. 2011; GCTA, 2024 [2] VanRaden, 2008

METHODS – GPSM (2/2)

- Estimate SNP effects[1]
  - y = μ + **xb** + **Zg** + **e**, where
  - …
  - **x** = vector of SNP genotypes
  - **b** = vector of effect size for SNP (related to reference allele, here: the minor allele)
- The SNPs with q-value<0.1 were considered significant

Figure 1. The univariate genome-wide linear mixed model to identify SNPs associated with proxy *AGE*.

[1] Yang et al. 2011; Rowan et al., 2021; GCTA, 2024

RESULTS & DISCUSSION

RESULTS (1/4)

Table 1. Number (N) of animals and markers in the subsets, variance components (σ_g^2, σ_e^2, and σ_P^2), proportion of variance (PVE) and the associated p-value explained by genome-wide SNPs and the related standard errors (SE).

Subset	N_{animal}	N_{SNP}	σ_g^2 (SE)	σ_e^2 (SE)	σ_P^2 (SE)	PVE (SE)	p-value
Bulls	9,975	43,462	2.48 (0.070)	0.78 (0.018)	3.26 (0.066)	0.76 (0.008)	0.000
Cows	64,160	43,641	2.73 (0.043)	0.78 (0.005)	3.52 (0.043)	0.78 (0.003)	0.000

Table 2. Number (N) of observations and markers in the subsets, and number of genome-wide SNPs associated with *AGE*.

Subset	N_{animal}	N_{SNP}	N, significant SNP
Bulls	9,975	43,462	52
Cows	64,160	43,641	54

RESULTS (1/4)

Table 1. Number (N) of animals and markers in the subsets, variance components (σ_g^2, σ_e^2, and σ_P^2), proportion of variance (PVE) and the associated p-value explained by genome-wide SNPs and the related standard errors (SE).

Subset	N_{animal}	N_{SNP}	σ_g^2 (SE)	σ_e^2 (SE)	σ_P^2 (SE)	PVE (SE)	p-value
Bulls	9,975	43,462	2.48 (0.070)	0.78 (0.018)	3.26 (0.066)	0.76 (0.008)	0.000
Cows	64,160	43,641	2.73 (0.043)	0.78 (0.005)	3.52 (0.043)	0.78 (0.003)	0.000

Table 2. Number (N) of observations and markers in the subsets, and number of genome-wide SNPs associated with *AGE*.

Subset	N_{animal}	N_{SNP}	N, significant SNP
Bulls	9,975	43,462	52
Cows	64,160	43,641	54

RESULTS (1/4)

Table 1. Number (N) of animals and markers in the subsets, variance components (σ_g^2, σ_e^2, and σ_P^2), proportion of variance (PVE) and the associated p-value explained by genome-wide SNPs and the related standard errors (SE).

Subset	N_{animal}	N_{SNP}	σ_g^2 (SE)	σ_e^2 (SE)	σ_P^2 (SE)	PVE (SE)	p-value
Bulls	9,975	43,462	2.48 (0.070)	0.78 (0.018)	3.26 (0.066)	0.76 (0.008)	0.000
Cows	64,160	43,641	2.73 (0.043)	0.78 (0.005)	3.52 (0.043)	0.78 (0.003)	0.000

Table 2. Number (N) of observations and markers in the subsets, and number of genome-wide SNPs associated with *AGE*.

Subset	N_{animal}	N_{SNP}	N, significant SNP
Bulls	9,975	43,462	52
Cows	64,160	43,641	54

RESULTS (2/4)

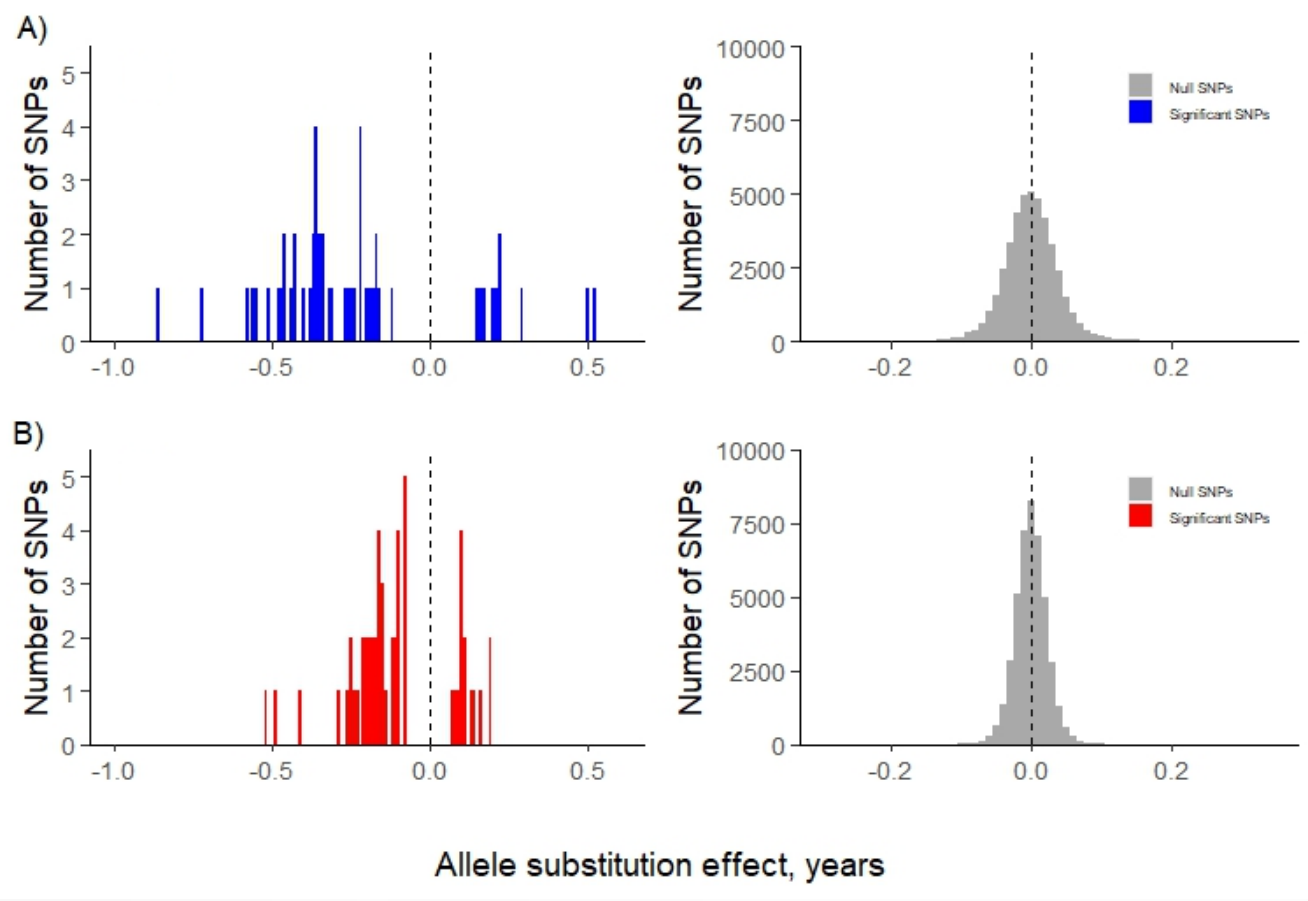

Figure 2. Distribution of allele substitution effects A) for bulls and B) for cows for GPSM significant (left) and null SNPs (right).

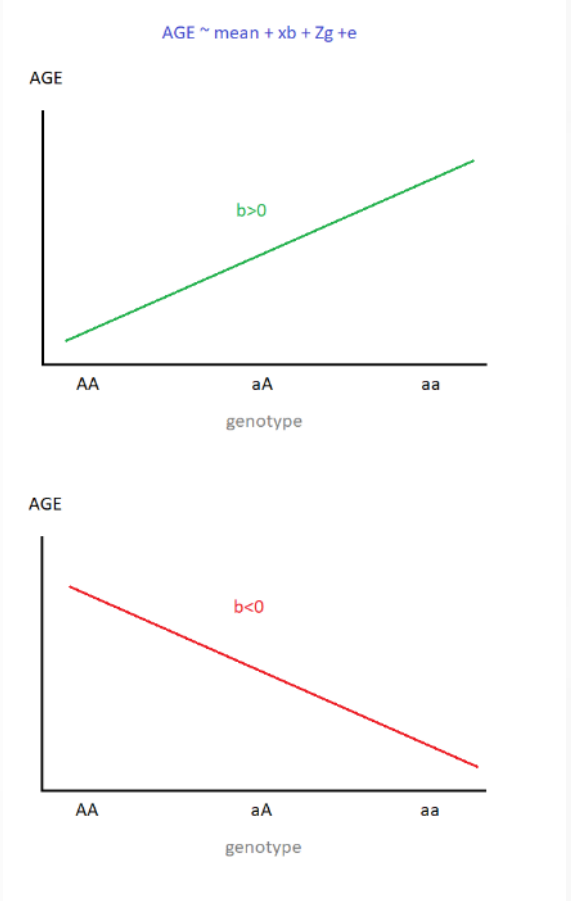

Figure 3. Positive and negative allelic substitution effects.

RESULTS (3/4)

Table 3. Allele substitution effects (b) in the GPSM, their standard errors (SE), and the Q-values associated with the p-value for SNPs in bull genomic data.

	SNP	Chr	Position (bp)	MAF	*b*	*SE*	Q-value
1	BFGL-NGS-111927	18	652,553	0.37	-0.86	0.027	7.13×10^{-214}
2	BFGL-NGS-104227	2	128,074,493	0.28	-0.72	0.030	2.02×10^{-120}
3	BFGL-NGS-18291	8	30,219,965	0.43	-0.58	0.029	1.29×10^{-84}
4	BTB-00495285	12	47,465,304	0.47	-0.56	0.028	2.98×10^{-83}
5	BFGL-NGS-95915	2	56,566,543	0.50	-0.55	0.029	1.31×10^{-75}
6	BTB-00752386	19	40,410,495	0.33	0.52	0.028	1.71×10^{-72}
7	BTB-00885930	24	31,132,914	0.35	-0.51	0.029	5.73×10^{-67}
8	BTB-01884758	1	30,620,433	0.42	0.50	0.029	1.38×10^{-63}
9	BFGL-NGS-28625	2	66,072,100	0.46	-0.48	0.028	1.15×10^{-60}
…	…						
14	UA-IFASA-998	3	67,472,539	0.24	-0.43	0.031	1.30×10^{-41}
…	…						
18	BFGL-NGS-37350	18	48,837,136	0.25	-0.40	0.031	5.46×10^{-35}
…	…						
23	BFGL-NGS-115980	26	20,196,747	0.25	-0.36	0.030	1.65×10^{-28}
…	…						
40	BFGL-NGS-20788	11	709,325	0.16	-0.22	0.036	3.17×10^{-06}
…	…						
52	BFGL-NGS-104720	13	43,896,312	0.19	-0.17	0.042	5.94×10^{-02}

RESULTS (3/4)

Table 3. Allele substitution effects (b) in the GPSM, their standard errors (SE), and the Q-values associated with the p-value for SNPs in bull genomic data.

	SNP	Chr	Position (bp)	MAF	*b*	*SE*	Q-value
1	BFGL-NGS-111927	18	652,553	0.37	-0.86	0.027	7.13×10^{-214}
2	BFGL-NGS-104227	2	128,074,493	0.28	-0.72	0.030	2.02×10^{-120}
3	BFGL-NGS-18291	8	30,219,965	0.43	-0.58	0.029	1.29×10^{-84}
4	BTB-00495285	12	47,465,304	0.47	-0.56	0.028	2.98×10^{-83}
5	BFGL-NGS-95915	2	56,566,543	0.50	-0.55	0.029	1.31×10^{-75}
6	BTB-00752386	19	40,410,495	0.33	0.52	0.028	1.71×10^{-72}
7	BTB-00885930	24	31,132,914	0.35	-0.51	0.029	5.73×10^{-67}
8	BTB-01884758	1	30,620,433	0.42	0.50	0.029	1.38×10^{-63}
9	BFGL-NGS-28625	2	66,072,100	0.46	-0.48	0.028	1.15×10^{-60}
…	…						
14	UA-IFASA-998	3	67,472,539	0.24	-0.43	0.031	1.30×10^{-41}
…	…						
18	BFGL-NGS-37350	18	48,837,136	0.25	-0.40	0.031	5.46×10^{-35}
…	…						
23	BFGL-NGS-115980	26	20,196,747	0.25	-0.36	0.030	1.65×10^{-28}
…	…						
40	BFGL-NGS-20788	11	709,325	0.16	-0.22	0.036	3.17×10^{-06}
…	…						
52	BFGL-NGS-104720	13	43,896,312	0.19	-0.17	0.042	5.94×10^{-02}

RESULTS (3/4)

Table 3. Allele substitution effects (b) in the GPSM, their standard errors (SE), and the Q-values associated with the p-value for SNPs in bull genomic data.

	SNP	Chr	Position (bp)	MAF	*b*	*SE*	Q-value
1	BFGL-NGS-111927	18	652,553	0.37	-0.86	0.027	7.13×10^{-214}
2	BFGL-NGS-104227	2	128,074,493	0.28	-0.72	0.030	2.02×10^{-120}
3	BFGL-NGS-18291	8	30,219,965	0.43	-0.58	0.029	1.29×10^{-84}
4	BTB-00495285	12	47,465,304	0.47	-0.56	0.028	2.98×10^{-83}
5	BFGL-NGS-95915	2	56,566,543	0.50	-0.55	0.029	1.31×10^{-75}
6	BTB-00752386	19	40,410,495	0.33	0.52	0.028	1.71×10^{-72}
7	BTB-00885930	24	31,132,914	0.35	-0.51	0.029	5.73×10^{-67}
8	BTB-01884758	1	30,620,433	0.42	0.50	0.029	1.38×10^{-63}
9	BFGL-NGS-28625	2	66,072,100	0.46	-0.48	0.028	1.15×10^{-60}
…	…						
14	UA-IFASA-998	3	67,472,539	0.24	-0.43	0.031	1.30×10^{-41}
…	…						
18	BFGL-NGS-37350	18	48,837,136	0.25	-0.40	0.031	5.46×10^{-35}
…	…						
23	BFGL-NGS-115980	26	20,196,747	0.25	-0.36	0.030	1.65×10^{-28}
…	…						
40	BFGL-NGS-20788	11	709,325	0.16	-0.22	0.036	3.17×10^{-06}
…	…						
52	BFGL-NGS-104720	13	43,896,312	0.19	-0.17	0.042	5.94×10^{-02}

RESULTS (4/4)

Table 4. Allele substitution effects (b) in the GPSM, their standard errors (SE), and the Q-values associated with the p-value for SNPs in cow genomic data.

	SNP	Chr	Position (bp)	MAF	*b*	*SE*	Q-value
1	BFGL-NGS-111927	18	652,553	0.19	-0.52	0.016	1.84×10^{-222}
2	BFGL-NGS-104227	2	128,074,493	0.17	-0.49	0.017	5.70×10^{-178}
3	BFGL-NGS-102016	10	38,338,318	0.13	-0.41	0.019	1.64×10^{-101}
4	BFGL-NGS-18291	8	30,219,965	0.35	-0.29	0.016	1.45×10^{-67}
5	BTB-00885930	24	31,132,914	0.30	-0.26	0.016	1.53×10^{-52}
6	BFGL-NGS-106992	17	61,506,352	0.32	-0.25	0.016	5.08×10^{-51}
7	BTB-00324256	7	75,268,501	0.34	-0.25	0.016	1.42×10^{-49}
8	BTB-00495285	12	47,465,304	0.40	-0.24	0.016	3.01×10^{-46}
9	BFGL-NGS-37350	18	48,837,136	0.21	-0.23	0.017	8.66×10^{-39}
10	HAPMAP51659-BTA-91205	2	33,135,887	0.28	-0.21	0.016	1.90×10^{-33}
11	BFGL-NGS-95915	2	56,566,543	0.41	-0.20	0.017	4.43×10^{-28}
12	BTB-01458106	18	60,583,784	0.36	-0.19	0.016	7.08×10^{-28}
13	UA-IFASA-998	3	67,472,539	0.20	-0.20	0.017	1.61×10^{-27}
14	HAPMAP38710-BTA-90319	3	29,674,329	0.10	-0.21	0.019	6.73×10^{-27}
15	BFGL-NGS-104977	6	14,005,865	0.29	-0.18	0.017	5.13×10^{-24}
16	BFGL-NGS-115980	26	20,196,747	0.22	-0.19	0.017	8.39×10^{-24}
…	…						
54	BTB-00926636	26	17,857,480	0.01	0.19	0.050	9.14×10^{-02}

RESULTS (4/4)

Table 4. Allele substitution effects (b) in the GPSM, their standard errors (SE), and the Q-values associated with the p-value for SNPs in cow genomic data.

	SNP	Chr	Position (bp)	MAF	*b*	*SE*	Q-value
1	BFGL-NGS-111927	18	652,553	0.19	-0.52	0.016	1.84×10^{-222}
2	BFGL-NGS-104227	2	128,074,493	0.17	-0.49	0.017	5.70×10^{-178}
3	BFGL-NGS-102016	10	38,338,318	0.13	-0.41	0.019	1.64×10^{-101}
4	BFGL-NGS-18291	8	30,219,965	0.35	-0.29	0.016	1.45×10^{-67}
5	BTB-00885930	24	31,132,914	0.30	-0.26	0.016	1.53×10^{-52}
6	BFGL-NGS-106992	17	61,506,352	0.32	-0.25	0.016	5.08×10^{-51}
7	BTB-00324256	7	75,268,501	0.34	-0.25	0.016	1.42×10^{-49}
8	BTB-00495285	12	47,465,304	0.40	-0.24	0.016	3.01×10^{-46}
9	BFGL-NGS-37350	18	48,837,136	0.21	-0.23	0.017	8.66×10^{-39}
10	HAPMAP51659-BTA-91205	2	33,135,887	0.28	-0.21	0.016	1.90×10^{-33}
11	BFGL-NGS-95915	2	56,566,543	0.41	-0.20	0.017	4.43×10^{-28}
12	BTB-01458106	18	60,583,784	0.36	-0.19	0.016	7.08×10^{-28}
13	UA-IFASA-998	3	67,472,539	0.20	-0.20	0.017	1.61×10^{-27}
14	HAPMAP38710-BTA-90319	3	29,674,329	0.10	-0.21	0.019	6.73×10^{-27}
15	BFGL-NGS-104977	6	14,005,865	0.29	-0.18	0.017	5.13×10^{-24}
16	BFGL-NGS-115980	26	20,196,747	0.22	-0.19	0.017	8.39×10^{-24}
…	…						
54	BTB-00926636	26	17,857,480	0.01	0.19	0.050	9.14×10^{-02}

RESULTS (4/4)

Table 4. Allele substitution effects (b) in the GPSM, their standard errors (SE), and the Q-values associated with the p-value for SNPs in cow genomic data.

	SNP	Chr	Position (bp)	MAF	*b*	*SE*	Q-value
1	BFGL-NGS-111927	18	652,553	0.19	-0.52	0.016	1.84×10^{-222}
2	BFGL-NGS-104227	2	128,074,493	0.17	-0.49	0.017	5.70×10^{-178}
3	BFGL-NGS-102016	10	38,338,318	0.13	-0.41	0.019	1.64×10^{-101}
4	BFGL-NGS-18291	8	30,219,965	0.35	-0.29	0.016	1.45×10^{-67}
5	BTB-00885930	24	31,132,914	0.30	-0.26	0.016	1.53×10^{-52}
6	BFGL-NGS-106992	17	61,506,352	0.32	-0.25	0.016	5.08×10^{-51}
7	BTB-00324256	7	75,268,501	0.34	-0.25	0.016	1.42×10^{-49}
8	BTB-00495285	12	47,465,304	0.40	-0.24	0.016	3.01×10^{-46}
9	BFGL-NGS-37350	18	48,837,136	0.21	-0.23	0.017	8.66×10^{-39}
10	HAPMAP51659-BTA-91205	2	33,135,887	0.28	-0.21	0.016	1.90×10^{-33}
11	BFGL-NGS-95915	2	56,566,543	0.41	-0.20	0.017	4.43×10^{-28}
12	BTB-01458106	18	60,583,784	0.36	-0.19	0.016	7.08×10^{-28}
13	UA-IFASA-998	3	67,472,539	0.20	-0.20	0.017	1.61×10^{-27}
14	HAPMAP38710-BTA-90319	3	29,674,329	0.10	-0.21	0.019	6.73×10^{-27}
15	BFGL-NGS-104977	6	14,005,865	0.29	-0.18	0.017	5.13×10^{-24}
16	BFGL-NGS-115980	26	20,196,747	0.22	-0.19	0.017	8.39×10^{-24}
…	…						
54	BTB-00926636	26	17,857,480	0.01	0.19	0.050	9.14×10^{-02}

DISCUSSION

- GPSM detected genome wide associations of genotypes on *AGE*
- Similar PVE in bulls and cows
  - Overall similar selection background (sampling strategy[1]) in the genotyped FAY bulls and cows
- Mean significant SNP effect two-fold in bulls (0.34) compared to cows (0.17)
  - Intense selection of bulls (shorter generation intervals)
- Effect of mixing Nordic Red breeds?
- Association of significant SNPs with traits under selection?

[1] Rowan et al., 2021

CONCLUCIONS

Overall, 52 SNPs in FAY bulls and 54 SNPs in FAY cows associated with birth date.

Selection signatures emerged during genomic selection were identified genome wide.

Selection in bulls more intense compared to cows.

Selection signatures relate to e.g. milk fat yield and stature

REFERENCES

Decker, J. E., Vasco, D. A., McKay, S. D., McClure, M. C., Rolf, M. M., Kim, J., Northcutt, S. L., Bauck, S., Woodward, B. W., Schnabel, R. D., and Taylor, J. F. 2012. A novel analytical method, Birth Date Selection Mapping, detects response of the Angus (Bos taurus) genome to selection on complex traits. BMC Genomics 13:606. https://doi.org/10.1186/1471-2

GCTA. 2024. GCTA-GREML: Estimate variance explained by all the SNPs. Accessed Feb. 12, 2024. https://yanglab.westlake.edu.cn/software/gcta/#GREMLanalysis 164-13-606

Grohmann, C. J., Shull, C. M., Crum, T. E., Schwab, C., Safranski, T. J., and Decker, J. E. 2023. Analysis of polygenic selection in purebred and crossbred pig genomes using generation proxy selection mapping. Genet Sel Evol 55:62. https://doi.org/10.1186/s12711-023-00836-9

NAV 2024. https://nordicebv.info/wp-content/uploads/2023/04/Changes-in-the-official-calculation-of-breeding-values-in-NAV-for-dairy-breeds-1999-2022.pdf

Purcell, S. M. and Chang, C.C. 2021. https://www.cog-genomics.org/plink/1.9/

Rowan, T. N., Durbin, H. J., Seabury, C. M., Schnabel, R. D., and Decker, J. E. 2021. Powerful detection of polygenic selection and evidence of environmental adaptation in US beef cattle. PLoS Genet 17: e1009652. https://doi.org/10.1371/

Sarviaho, K., Uimari, P., and Martikainen, K. 2024. Signatures of positive selection after the introduction of genomic selection in the Finnish Ayrshire population. J. Dairy Sci. 107:4822–4832. https://doi.org/10.3168/jds.2024-24105

VanRaden, P. M. 2008. Efficient Methods to Compute Genomic Predictions. J. Dairy Sci. 91:4414–4423. https://doi.org/10.3168/jds.2007-0980

Yang, J., Manolio, T. A., Pasquale, L. R., Boerwinkle, E., Caporaso, N., Cunningham, J. M., de Andrade, M., Feenstra, B., Feingold, E., Hayes, M. G., Hill, W. G., Landi, M. T., Alonso, A., Lettre, G., Lin, P., Ling, H., Lowe, W., Mathias, R. A., Melbye, M., Pugh, E., Cornelis, M. C., Weir, B. S., Goddard, M. E., & Visscher, P. M. 2011A. Genome partitioning of genetic variation for complex traits using common SNPs. Nat. Genet. 43: 519–525. https://doi.org/10.1038/ng.823