Accuracy of genomic prediction of DMI in Dutch Holsteins using sequence variants from meta-analyses

Birgit Gredler-Grandl and co-authors

BovReg
Understanding cattle genomes

World Congress on Genetics Applied to Livestock Production
Rotterdam | The Netherlands
3 - 8 July 2022

WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018

Thank you to all co-authors (BovReg & gDMI)

AARHUS UNIVERSITET

Coralia Manzanilla-Pech
Zexi Cai
Goutam Sahana

AGRICULTURE VICTORIA

Sunduimijid Bolormaa
Jennie Pryce

UNIVERSITY OF ALBERTA

Yining Wang
ChangXi Li
Dan Hailemariam
Graham Plastow

CONAFE

Noureddine Charfeddine

FBN

Praveen Chitneedi
Christa Kuehn

UNIVERSITY of GUELPH

Tatiana Chud
Flavio Schenkel
Christine Baes

INIA
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria

Beatrix Villanueva
Almudena Fernandez
Oscar González-Recio

Luke
NATURAL RESOURCES INSTITUTE FINLAND

Daniel Fischer
Teri Iso-Touru
Martin Lidauer

WAGENINGEN
UNIVERSITY & RESEARCH

Biaty Raymond
Ghyslaine Schopen
Yvette de Haas
Roel Veerkamp
Aniek Bouwman

WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018

Introduction

- BovReg: Providing a catalogue of functionally active genomic features in cattle
- Key traits: Biological efficiency → Feed efficiency
- Large scale recording of feed intake difficult
  - → size of reference populations limited
  - → accuracy of GEBV is limited

https://hokofarmgroup.com

Objective

Can we improve the accuracy of genomic prediction for DMI in Dutch Holsteins using sequence variants from meta-analyses?

1. meta-GWAS **QTL**

2. meta-GWAS **mQTL**

Meta-GWAS QTL for DMI

- Imputation to sequence level
- GWAS on own dataset by each partner
- Phenotypes: de-regressed BV, direct phenotypes corrected for fixed effects
- GWAS summary statistics to WUR
- Software METAL (Willer et al., 2010)
  - STDERR method

Dairy

Beef

AGRICULTURE VICTORIA

UNIVERSITY of GUELPH

UNIVERSITY OF ALBERTA

WAGENINGEN UNIVERSITY & RESEARCH

FBN

INIA
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria

AARHUS UNIVERSITET

Luke
NATURAL RESOURCES INSTITUTE FINLAND

FBN

WAGENINGEN UNIVERSITY & RESEARCH

100years 1918 — 2018

13,104 animals in GWAS for DMI

Population	n	N variants Imputation r^2>0.6
AU	495	17,158,878
AgVIC	584	17,597,583
FBN dairy	140	14,016,930
FBN beef	253	16,057,514
INIA	561	12,829,538
LUKE	366	18,392,101
UAL	7,552	30,381,524
UoG	588	11,694,898
WUR	2,565	17,817,916

Excluded from meta-GWAS

WAGENINGEN UNIVERSITY & RESEARCH

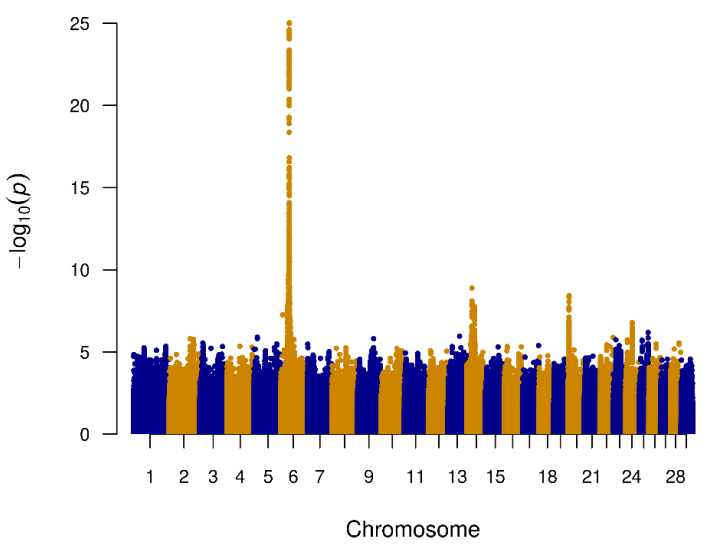
100years 1918 — 2018

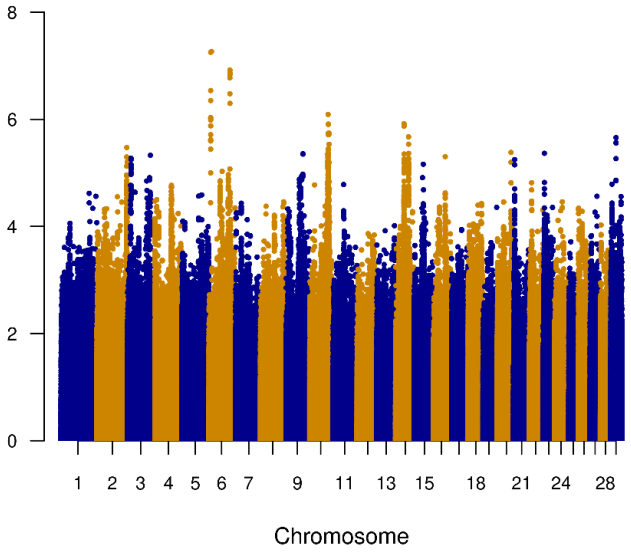
Manhattan plots meta-GWAS QTL

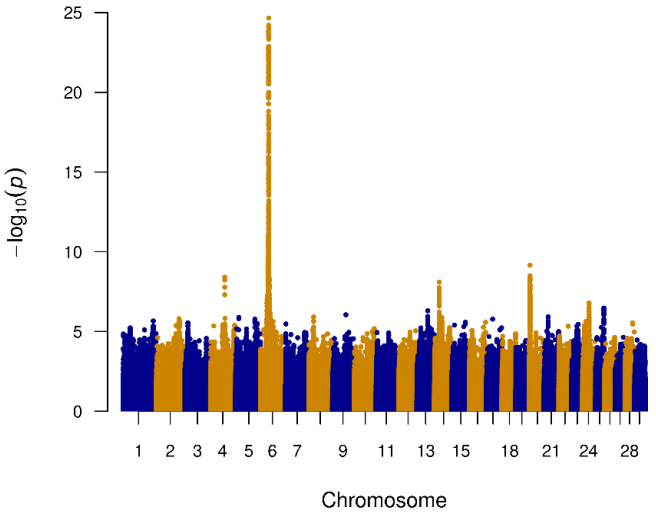
ALL
10,539 animals
Beef, Holstein, Finnish Red
MAF >0.005
30,216,688 variants

HOL
2,368 animals
Only Holstein
MAF >0.005
19,647.876 variants

BEEF
7,805 animals
Only Beef
MAF >0.005
27,839,929 variants







Zhang et al. 2020

WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018

Meta-GWAS mQTL

- Blood plasma metabolomic QTL
- amino acids, short and long-chain fatty acids and compounds from energy and protein metabolism
- Imputation and GWAS on by each partner
- Meta-GWAS by FBN, Germany
    - Software METAL (Willer et al., 2010)
    - 27 metabolites
    - N variants: 14,343,591 and 19,467,841
    - N animals: 241 – 1,103

Dairy

AARHUS UNIVERSITET

Holstein bull calves

Beef

UNIVERSITY OF ALBERTA

Canadian multibreed composite and crossbred

FBN

Holstein x Charolais

WAGENINGEN UNIVERSITY & RESEARCH

100years 1918 — 2018

SNP selection

- Conditional and joint effect method (COJO) to select independently associated variants (Yang et al., 2012)
- Meta-GWAS QTL
    - P-value threshold to declare a genome-wide significant variant:
    - p=5e-3 (ALL3, HOL3, BEEF3)
    - p=5e-5 (ALL5, HOL5, BEEF5)
- Meta-GWAS mQTL:
    - All significant variants $p<10^{-6}$ were selected across all meta-GWAS
- Base scenario: 50k variants from Illumina Bovine snp50 v3 beadchip

Number of variants

Scenario	n
Base 50K	37,179
ALL3	1,810
HOL3	1,834
BEEF3	1,746
ALL5	166
HOL5	63
BEEF5	321
mQTL	17,056

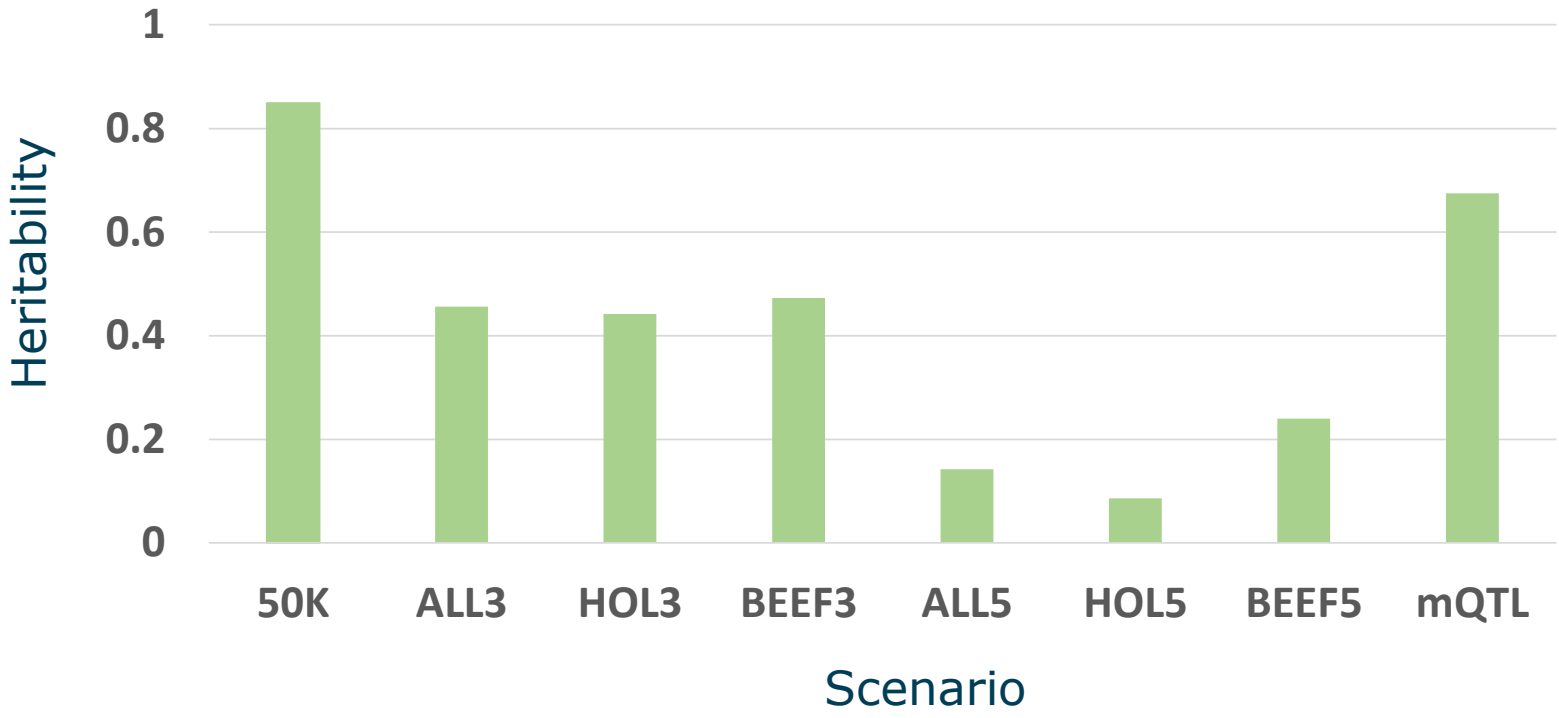
Genomic prediction

- Independent set of 2,162 Holstein cows
- Phenotypes: de-regressed proofs
- 5-fold cross validation:
    - 400 cows randomly selected as validation set
    - 1,762 cows in the training set

- GREML fitting one GRM (mtg2): $\mathbf{y = 1\mu + Wg + e}$

- GREML fitting two GRMs (mtg2): $\mathbf{y = 1\mu + W_1g_1 + W_2g_2 + e}$

Heritability of de-regressed proof – fitting one GRM



Standard errors 0.024 – 0.034

Accuracy of GEBV – fitting one GRM

correlation between GEBV and de-regressed BV, average across 5 CV

Scenario	N variants	Correlation
50k	37,179	0.68
ALL3	1,810	0.56
HOL3	1,834	0.55
BEEF3	1,746	0.57
ALL5	166	0.36
HOL5	63	0.20
BEEF5	321	0.47
mQTL	17,056	0.48

Standard errors 0.027 – 0.048

Accuracy of GEBV – fitting two GRM

correlation between GEBV and de-regressed BV, average across 5 CV

Scenario	N variants	Correlation
50k	37,179	0.68
50k + ALL3	37,179 + 1,810	0.69
50k + HOL3	37,179 + 1,834	0.68
50k + BEEF3	37,179 + 1,746	0.65
50k + ALL5	37,179 + 166	0.68
50k + HOL5	37,179 + 63	0.68
50k + BEEF5	37,179 + 321	0.68
50k + mQTL	37,179 + 17,056	0.68

Standard errors 0.027

Considerations and conclusions

- No clear advantage in accuracy when including sequence variants in GBLUP
    - Genetic architecture of the trait
    - GWAS signals not strong enough
    - Bayesian variable selection methods may perform better

- BEEF sets seem to outperform HOL sets
    - Anomaly
    - Divers phenotypes/diets within HOL
    - Look into covariance between different sets

Thank you for your attention!

www.bovregproject.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815668

WAGENINGEN UNIVERSITY & RESEARCH

100years 1918 — 2018

birgit.gredler-grandl@wur.nl