

Biology-driven genomic predictions for dry matter intake within and across breeds using WGS data

Renzo Bonifazi, M Heidaritabar, AC Bouwman, LD Barlow, L Chen, P Stothard, J Basarab, C Li, G Plastow, the BovReg consortium, B Gredler-Grandl



X @RenzoBonifazi



Biology-driven genomic predictions

- WGS: millions of variants → pinpoint **causal variants** affecting traits of interest
- Key traits: **biological efficiency**, e.g. **feed efficiency**
- BovReg: catalogue of functionally active **genomic features (GF)** in cattle
- “**multi-omics**” data: genomics (QTL), transcriptomics (eQTL), epigenomics (mQTL), chromatin accessibility (ATAC-seq/ChIP-seq)
- **Functional GF** → SNP prioritization → **Biology-driven genomic predictions**

Aim of the study

Validation of **biology-driven genomic predictions**

using **genomic features** for **dry matter intake**

Data available for genomic predictions

	NLD 	CAN 
Breed	Holstein	Beef crosses
Herds	6	14
Dependent variable	DRP on DMI	Pre-corrected DMI
n. animals DMI & geno (training - validation)	~3k (2.2k and 850)	~5.5k (4k and 1.5k)

- **forward-in-time validation** (SE via bootstrapping)
- 50k to imputed WGS (*Beagle*) → prioritize variants based on **GF**

Genomic features used

GF

Traits / Tissues

QTL

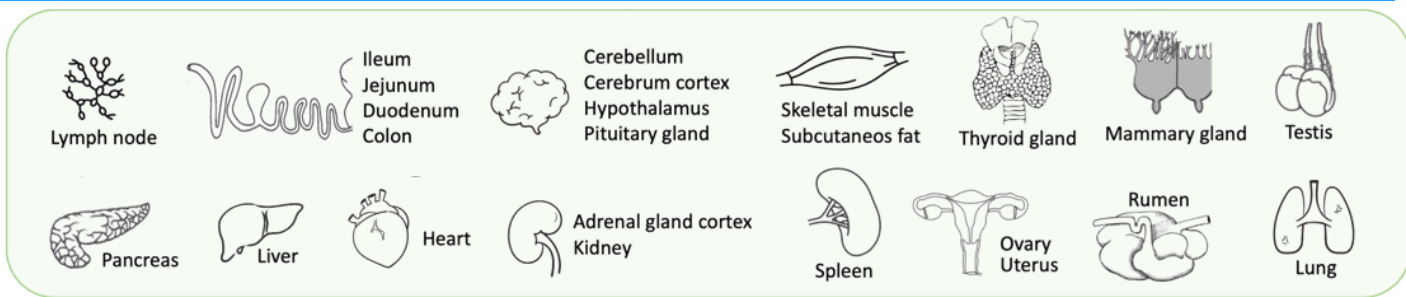
Meat quality, Growth, Milk production, Morphology, Fertility, Health, Feed efficiency, Methane

(Gene, Transcript, Splice)

eQTL

Jejunum, Blood, Liver, Mammary Gland, Adipose, Muscle, Milk, Rumen

ATAC-seq

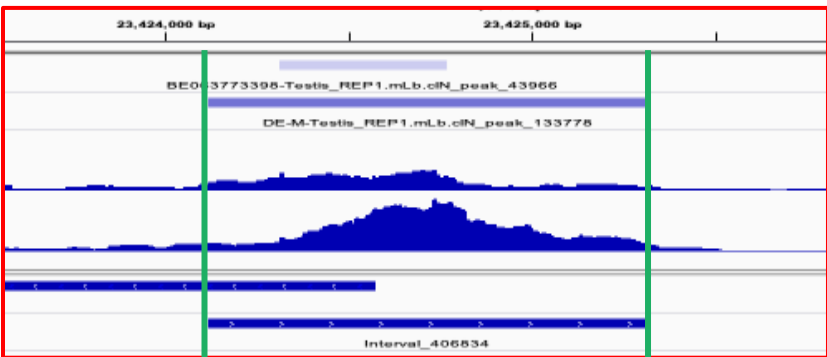
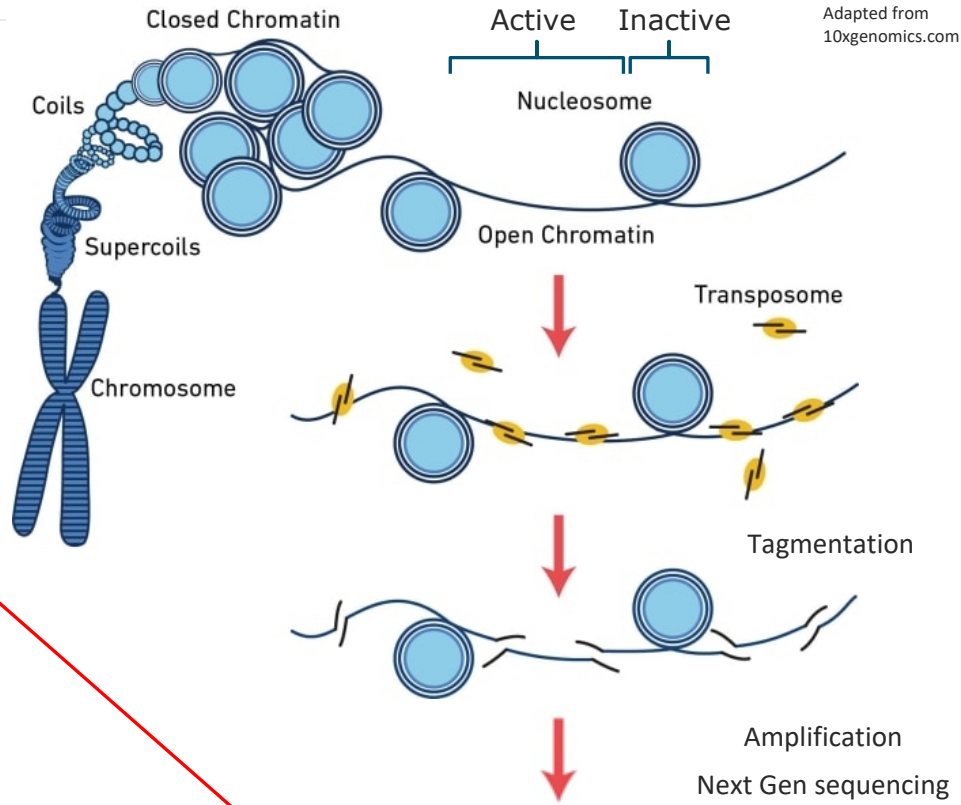


ATAC-seq

Genome-wide mapping of **chromatin accessibility** → **accessible DNA**



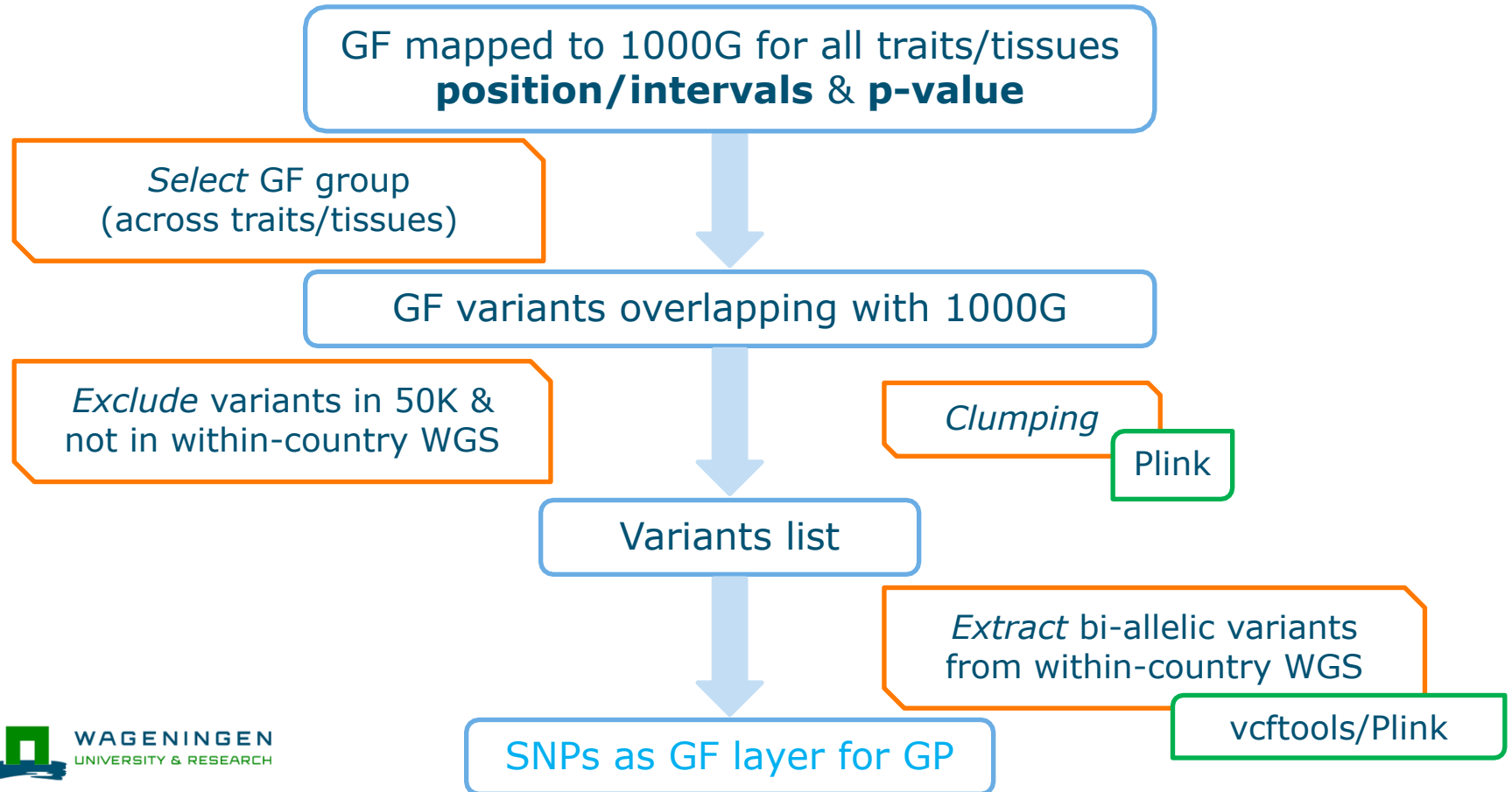
Promotor regions
Enhancers
Regulatory elements



Consensus peaks
across samples



Selection of genomic features

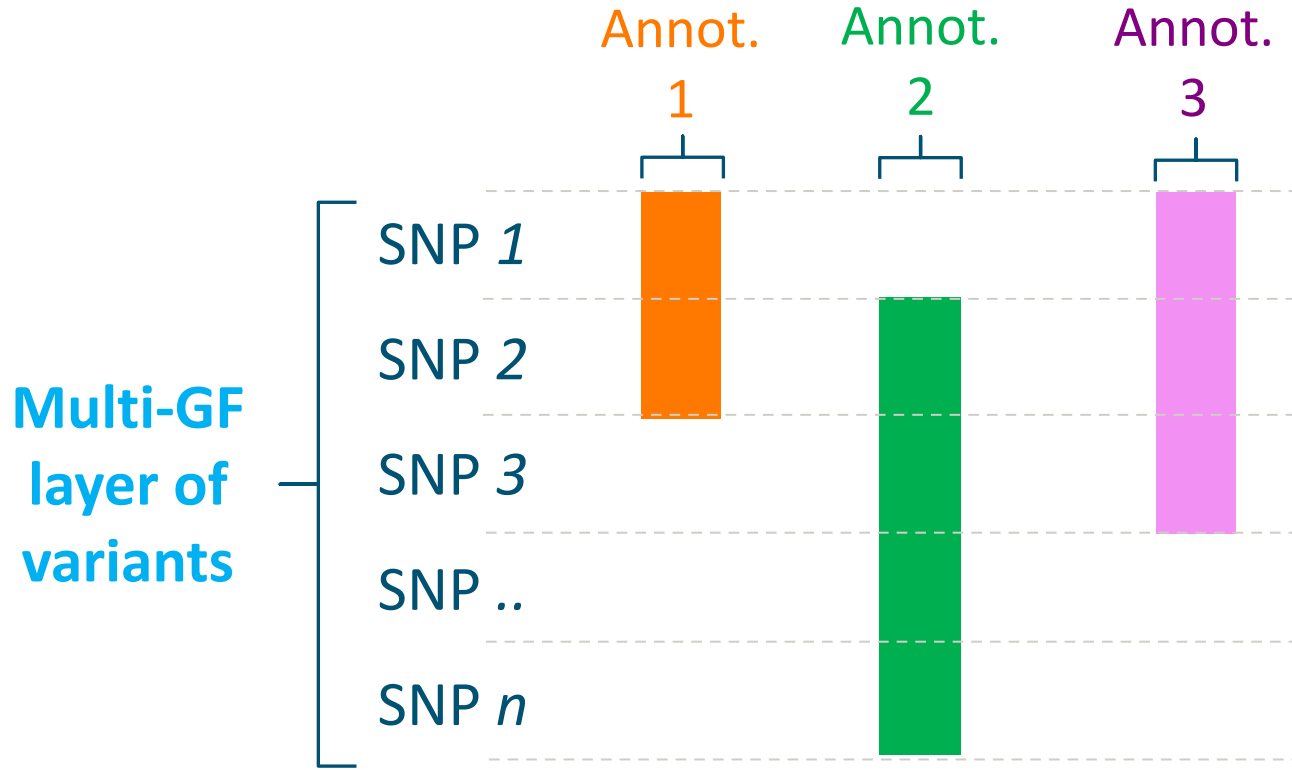


Scenarios and software

Scenario	NLD		CAN	
50K	48K		46K	

- *NextGP.jl* ¹
- Base 50K: SNPBLUP (BayesC₀) common variance across SNPs
- Additional GF layer: SNPBLUP or Bayesian (2 mixture model – no advantage)
- QTL, eQTL, ATAC → Multi-GF → overlapping GF

Overlapping genomic features

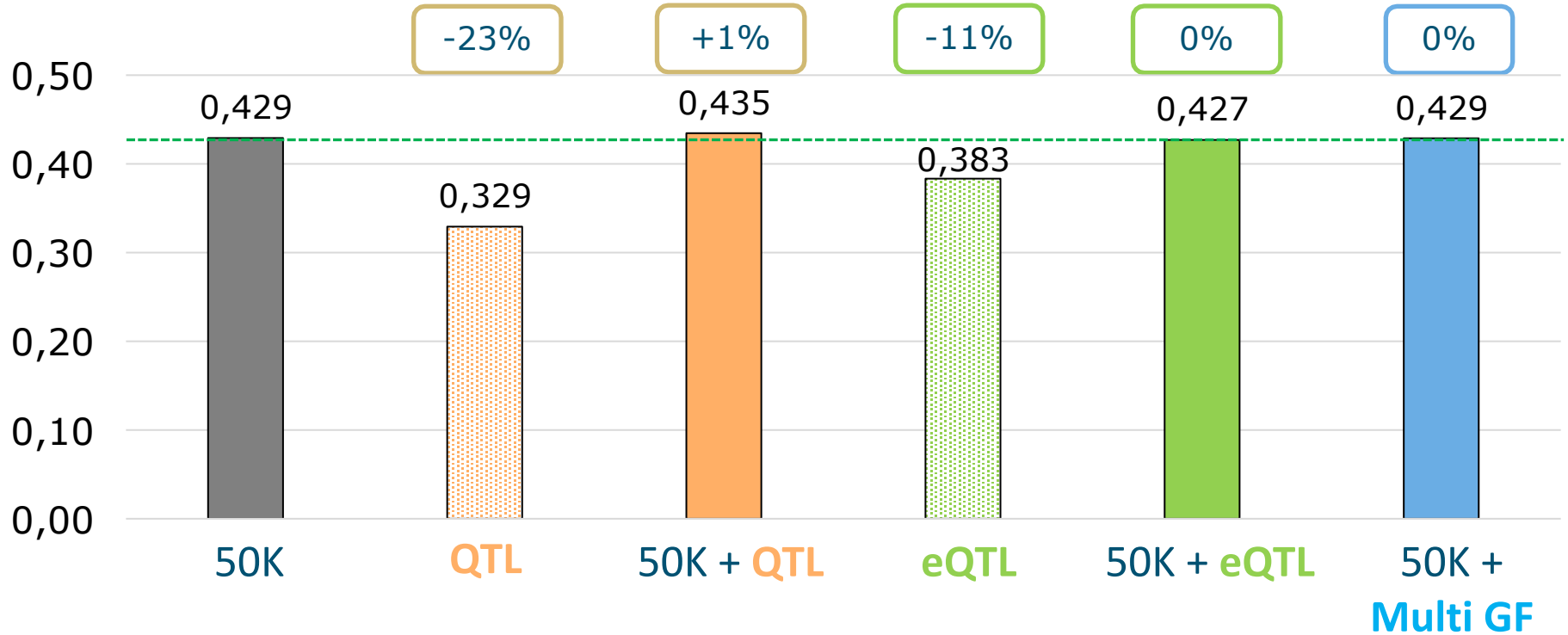


Overlapping genomic features

	QTL	eQTL	ATAC	Variants	%
1	1	0	5	0	
1	1	1	16	0	
1	0	0	2,576	14	
1	0	1	2,819	16	
0	1	0	5,051	28	
0	1	1	7,329	41	
			17,796	100	

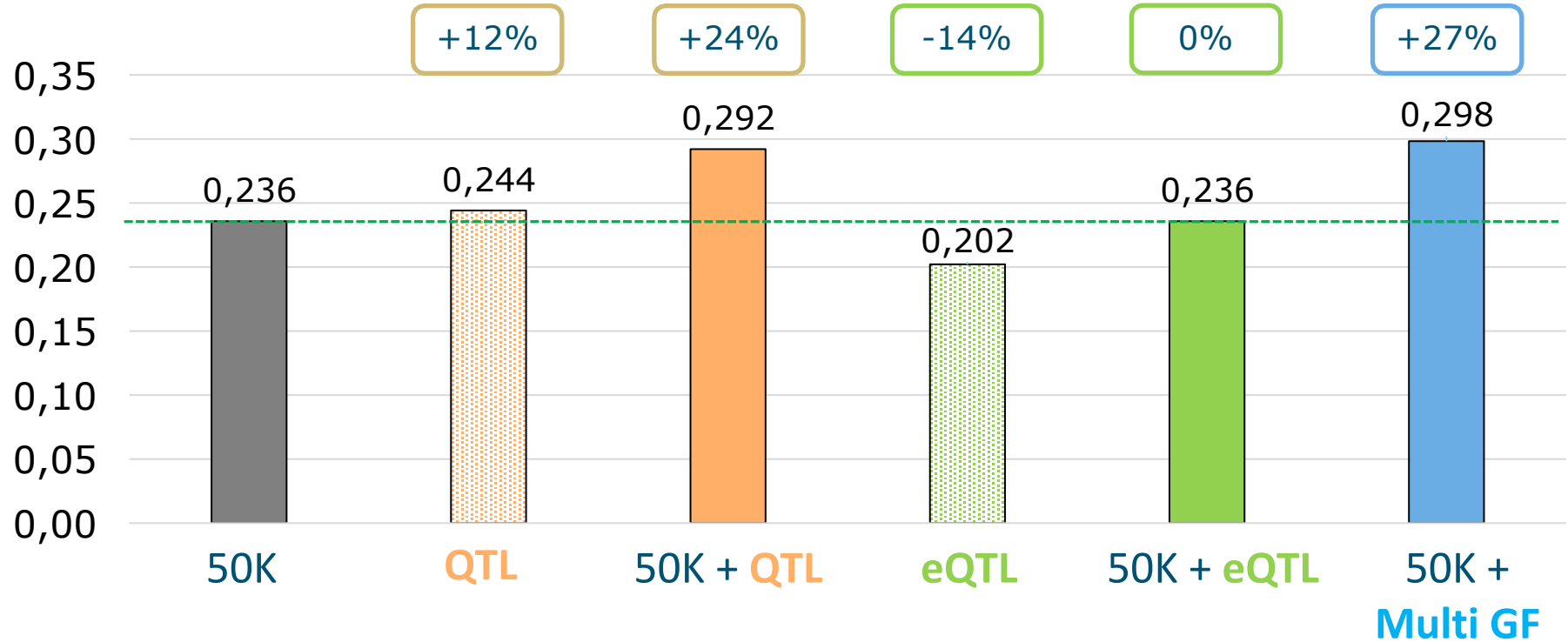
BayesRC π (2 classes)

Results – NLD: prediction accuracy



- SE: ± 0.03
- No impact on dispersion (slope = 0.68 for 50k)

Results – CAN: prediction accuracy



- SE between ± 0.02 and ± 0.03
- Similar pattern for dispersion (slope = 0.56 for 50k)

Conclusions and next steps

- Inclusion of **Genomic Features** could increase **genomic prediction accuracies** for **Dry Matter Intake**
- Results may vary across breeds/datasets → find causal variants (complex traits)
- No advantage using (2 mixture) Bayesian approach over SNPBLUP for GF

Next steps

- focus trait/tissues-specific variants, more detailed annotation modelling (BayesLV), across-/multi-breed NLD-CAN

Acknowledgments

Dr Emre Karaman and Dr Luc Janss

Dr Gabriel Costa M Moreira

The BovReg consortium

This work is carried out within the BovReg project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815668.



Disclaimer: Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.



Thanks for your attention



renzo.bonifazi@wur.nl

X@RenzoBonifazi



Disclaimer: Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

- Inclusion of **Genomic Features** could increase genomic prediction accuracies for **Dry Matter Intake**
- Results may vary across breeds/datasets
- No advantage using Bayesian (2 mix model) over SNPBLUP for **GF**



Thanks for your attention



renzo.bonifazi@wur.nl

X@RenzoBonifazi



Disclaimer: Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Manhattan plots meta-GWAS QTL

ALL

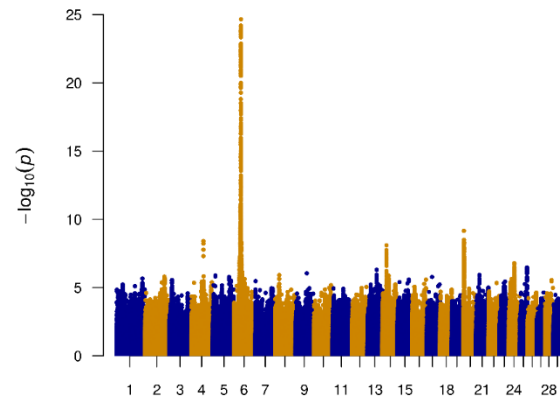
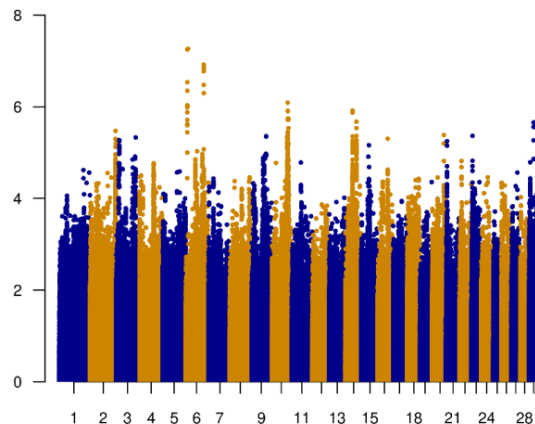
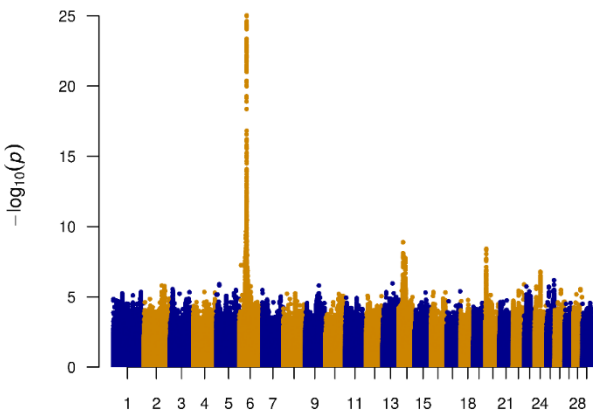
10,539 animals
Beef, Holstein, Finnish Red
MAF > 0.005
30,216,688 variants

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

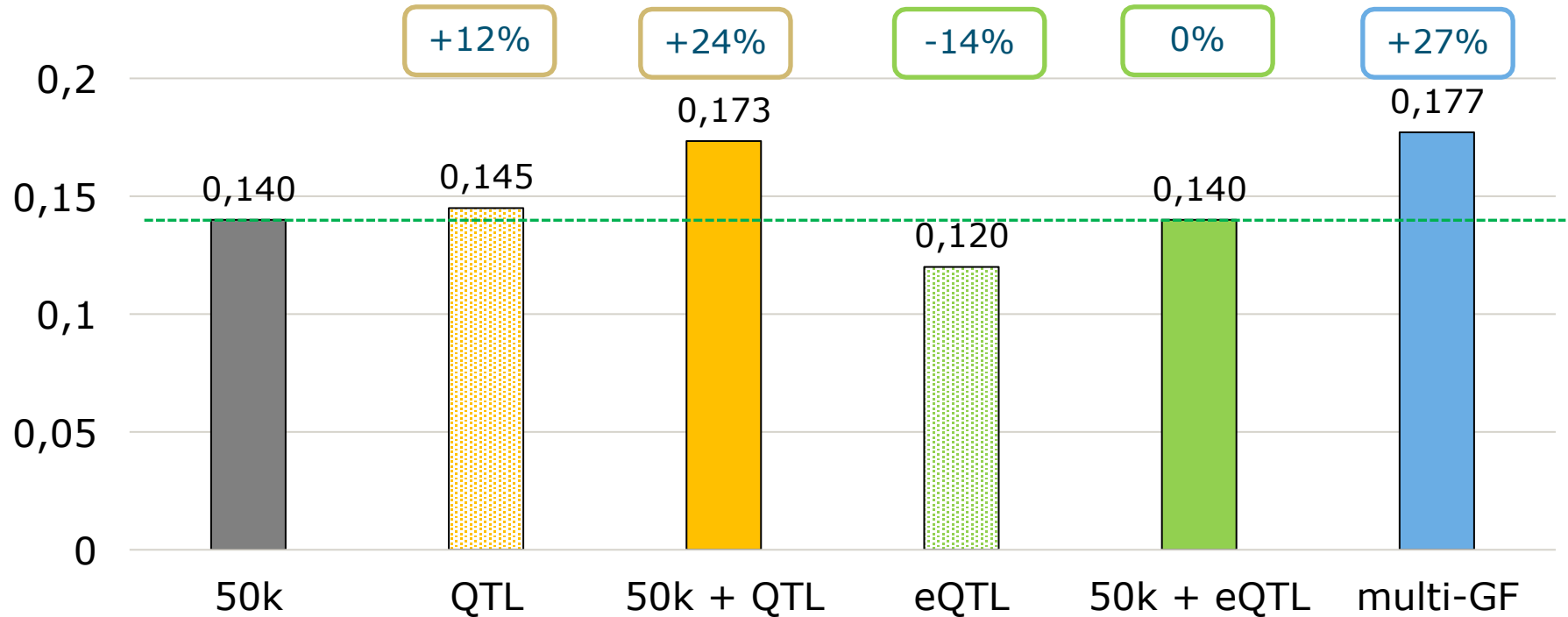


Zhang et al. 2020 Chromosome

Chromosome

Chromosome

Results – CAN: prediction accuracy (not scaled by $\sqrt{h^2}$)



- SE between ± 0.02 and ± 0.03
- No advantage of Bayesian for single GF layers
- Additional GF layers: no impact on dispersion (slope = 0.68 for 50k)