



Using RNAseq data to improve genomic selection in dairy cattle

Thomas Lopdell, Kathryn Tiplady, Mathew Littlejohn

INNOVATION

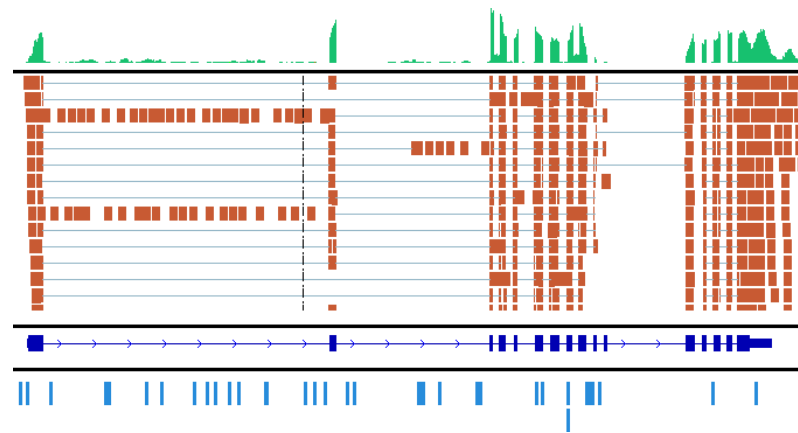
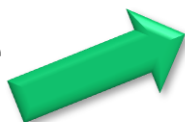


Introduction – Genomic Selection

- Using variants in DNA to predict animal performance
- Advantage: animals with no phenotypes

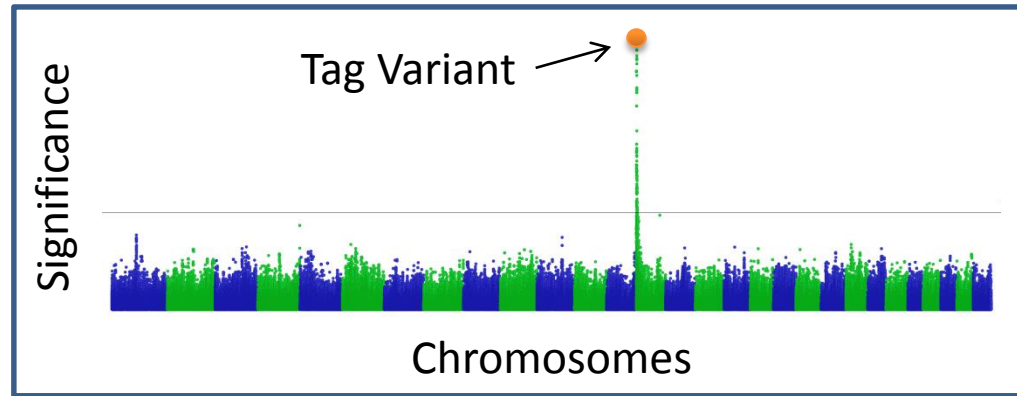
Introduction

- Adding RNAseq
 - Measure gene expression
 - Enrich variants in genes



Introduction

- Adding RNAseq
 - Measure gene expression
 - Enrich variants in genes
 - Identify eQTL



Aims

- Identify tag variants for regulatory loci
- Test predictive ability of variant set

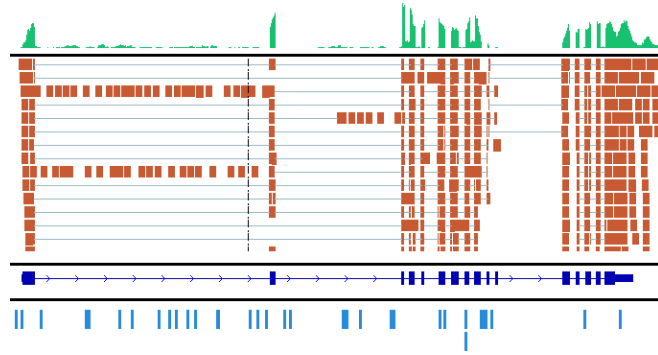
Methods (1)



Cows (x373)



Sequence,
map reads



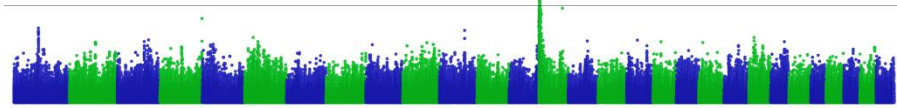
Gene expressions



477,531 variants

Methods (2)

Gene expressions



eQTL (x3,738)

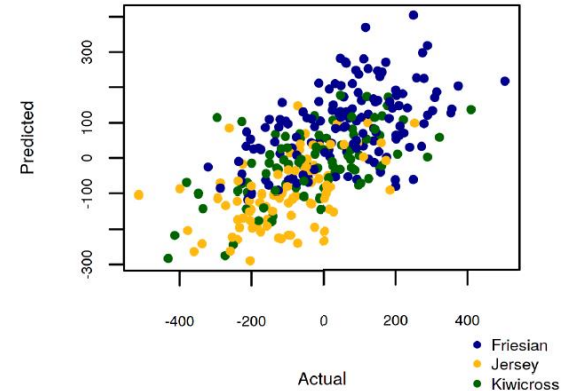


Tag variants (x3,695)



Predictions

Predicted vs actual: ydmilk



1,093,581 variants (including SNP chip)



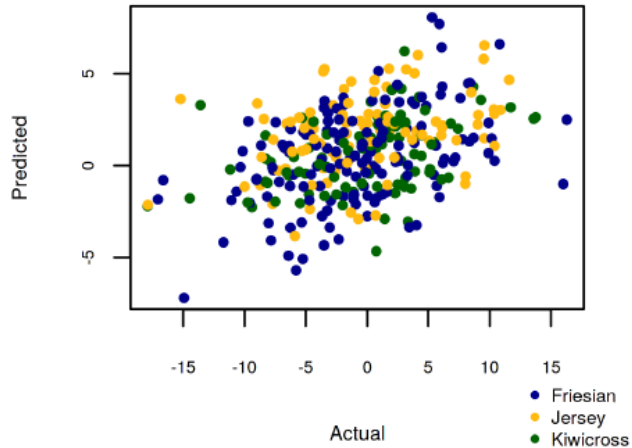
Results – eQTL

- Replicated known eQTL associated with milk production and composition

Gene	P-value	%Var Fat	%Var Milk	%Var Protein
MGST1	1.90×10^{-54}	2.06	2.34	2.68
DGAT1	3.77×10^{-53}	22.49	23.54	45.86
GPAT4	7.19×10^{-21}	1.36	0.55	2.82
PLAG1	1.47×10^{-14}	3.70	0.83	0.31

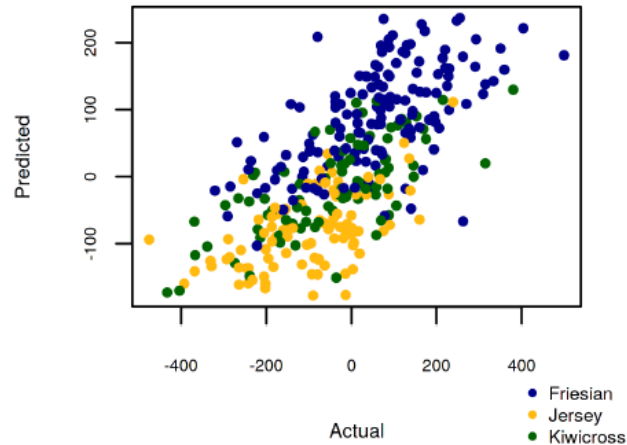
Results – Predictions

Fat Yield



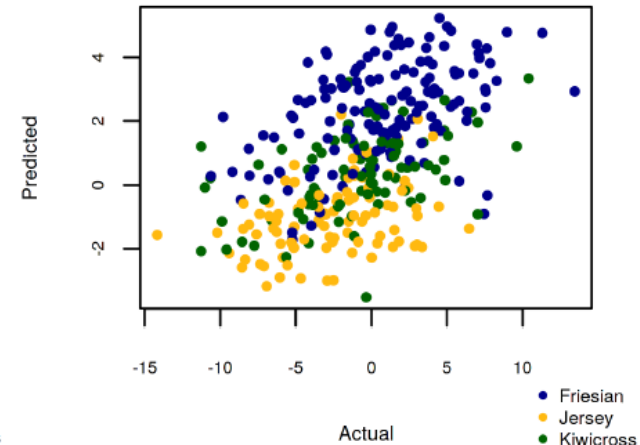
R = 0.389

Milk Yield



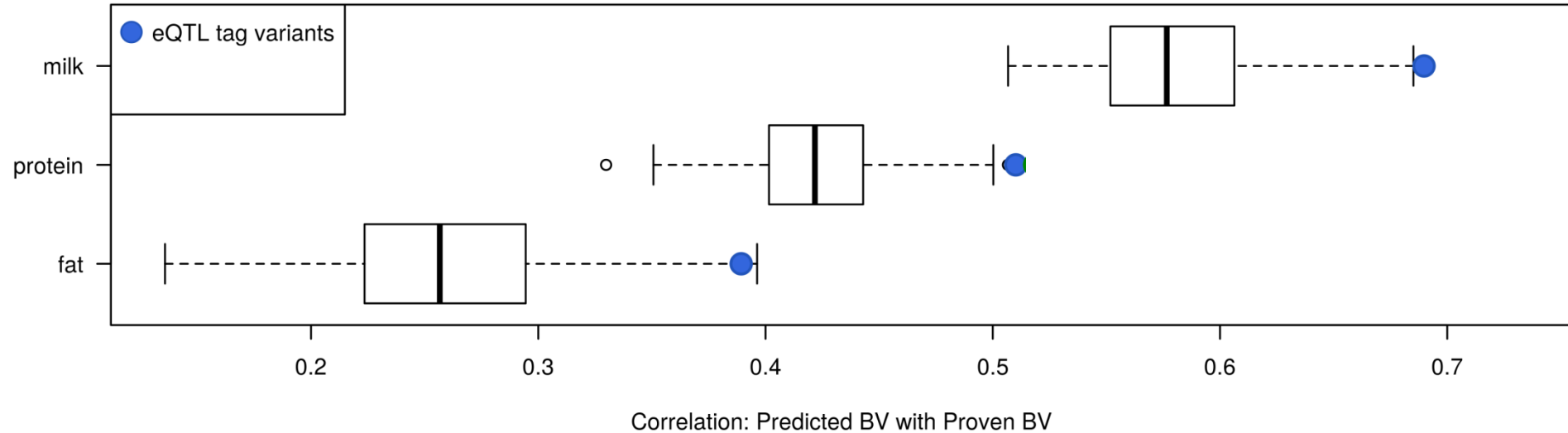
R = 0.690

Protein Yield

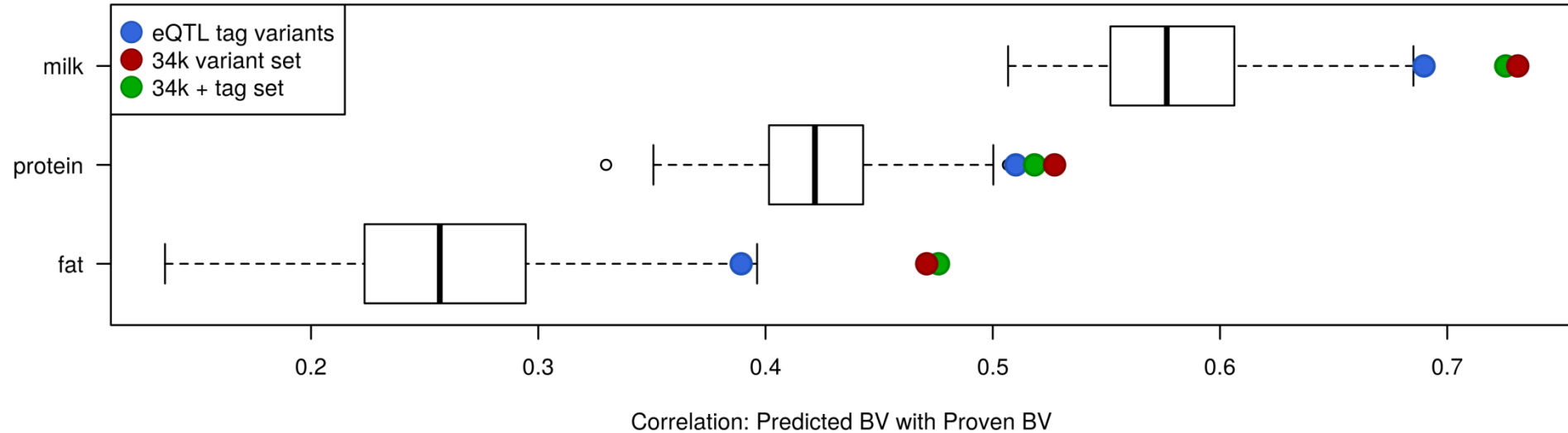


R = 0.510

Results – Predictions



Results – Predictions



Conclusion

- eQTL tag variants predict better on average than SNP chip genotypes
 - Likely enriched for causative variants
- Accuracy may be improved further by including protein-coding variants