Metabolomic phenotypic prediction of growth in pigs

Pernille Sarup

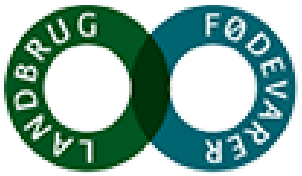
J. Jensen, T. Ostersen, O.F. Christensen, P. Sørensen



How to increase EBV accuracies by 20-50%

Pernille Sarup

J. Jensen, T. Ostersen, O.F. Christensen, P. Sørensen



SEGES
Svineproduktion

GenSAP

InnovationsFonden
FORSKNING, TEKNOLOGI & VÆKST I DANMARK

Background

Accurate EBVs need lots of phenotypic measurements

But

Some phenotypes are expensive or difficult to measure

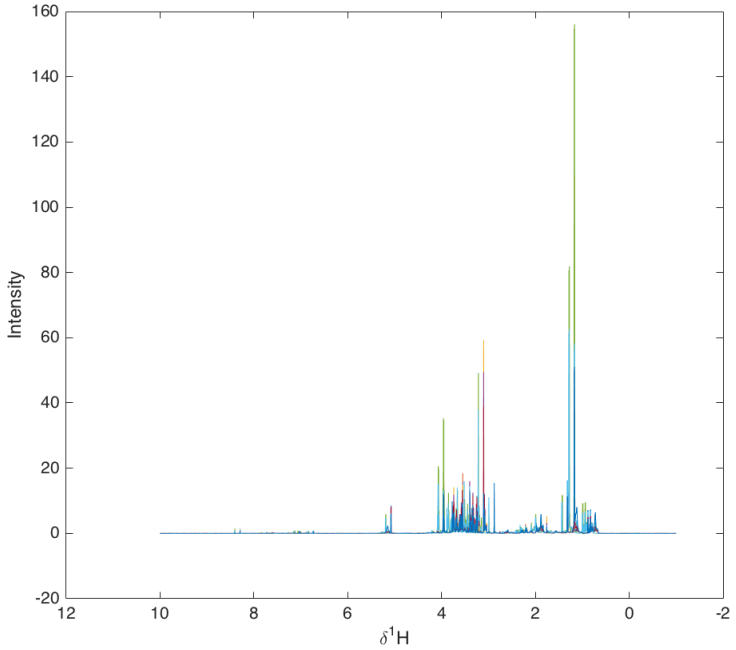
Current solution:

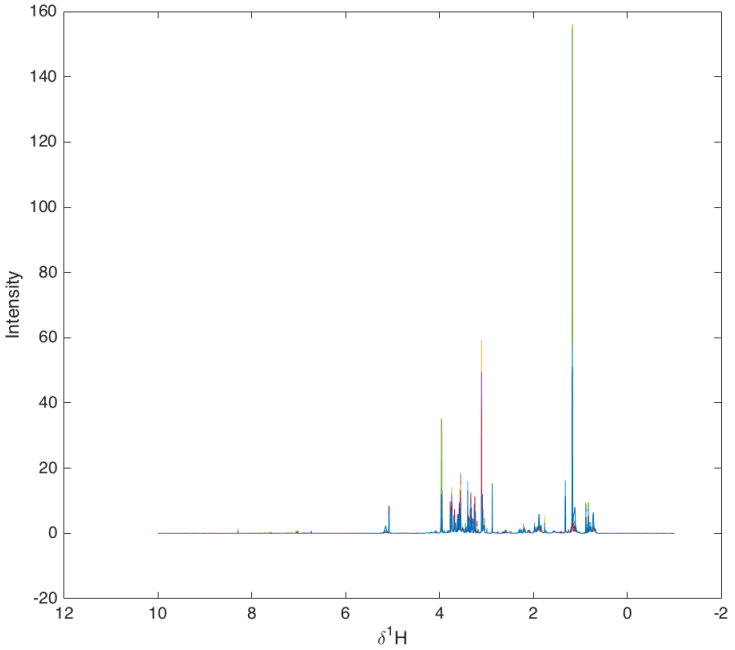
- Breeding values using information on correlated phenotypes and additive genetic relationships with tested animals

Background

New option

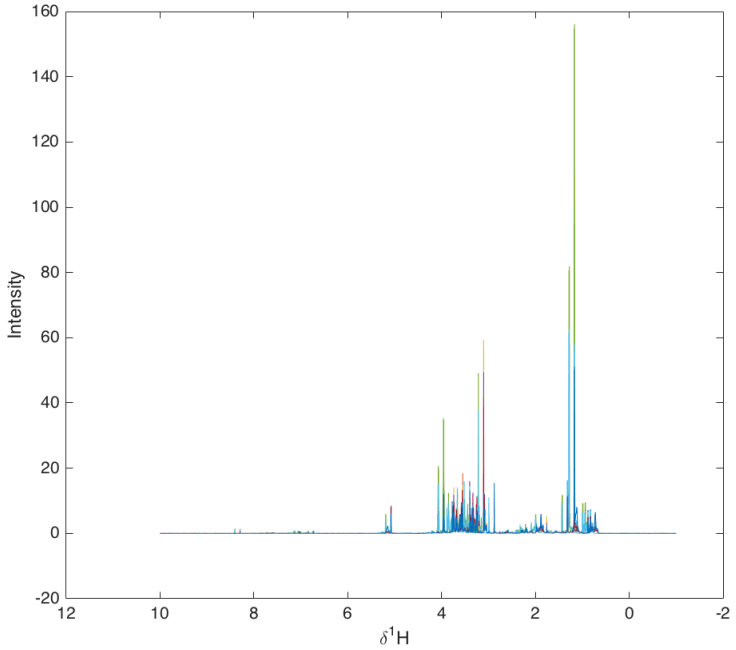
Use large scale molecular phenotypes (e.g. H^1 NMR metabolomics) for prediction of relevant phenotypes in large cohorts.

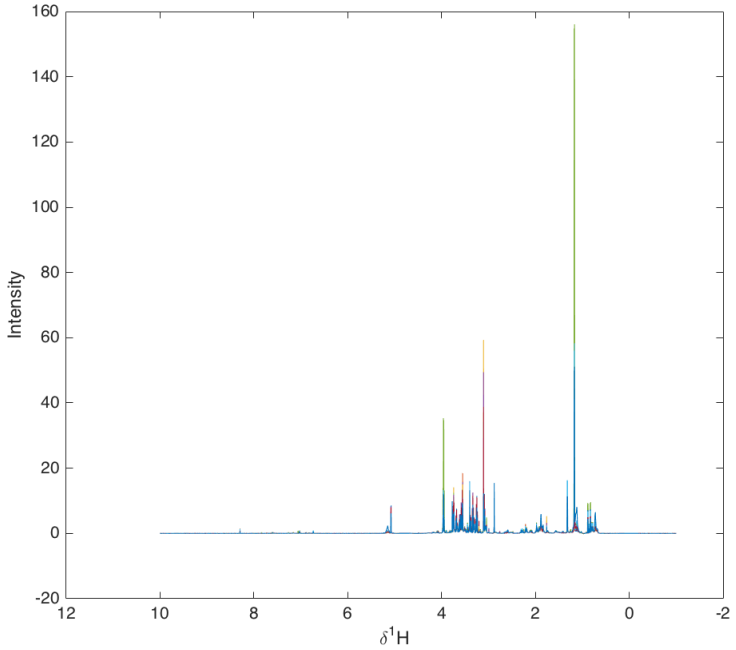




Background

NMR spectra are affected by both the genotype and the environment, but only the genetic component is useful for prediction of breeding values.





Aim

Can we use these NMR predicted phenotypes to increase accuracy of EBVs?

Data

2320 finishing Duroc x(Yorkshire x Landrace)

- Genotypes (6K)
- NMR – metabolomics on blood serum
- Total Feed Intake
- Average Daily Gain
- Back Fat

Adjusted phenotypes

Growth phenotypes were adjusted for fixed effects of stable-year-month and sex

$$y = X_{sec}sec + X_{sex}sex + e$$

$$y_{adj} = \hat{e}$$

Models

GBLUP

$$\mathbf{y} = \mathbf{Z}_l\mathbf{l} + \mathbf{Z}_g\mathbf{g} + \mathbf{e}$$

$$\mathbf{g} \sim \mathrm{N}\left(0, \mathbf{G}\sigma_g^2\right)$$

$$\mathbf{G} = (\mathbf{W}\mathbf{W}'/\mathrm{n})$$

Models

MBLUP

$$\mathbf{y} = \mathbf{Z}_l\mathbf{l} + \mathbf{Z}_\mathrm{m}\mathbf{m} + \mathbf{e}$$

$$\mathbf{m} \sim \mathrm{N}(0, \mathbf{M}\sigma_m^2)$$

$$\mathbf{M} = (\mathbf{X}\mathbf{X}'/\mathrm{n})$$

$\mathbf{X}$ = **scaled and centered NMR intensities**

Cross validation schemes

Prediction across environments and family:

- Oldest half of stable-year-month in the training population (across)

Prediction within environments and family:

- Alternate stable-year-month in the training population (within)

Cross validation schemes

GBLUP

$$\mathbf{y} = \mathbf{Z}_l\mathbf{l} + \mathbf{Z}_g\mathbf{g} + \mathbf{e}$$

Cross-validated:

Univariate $\hat{\mathbf{g}}_{cv}$

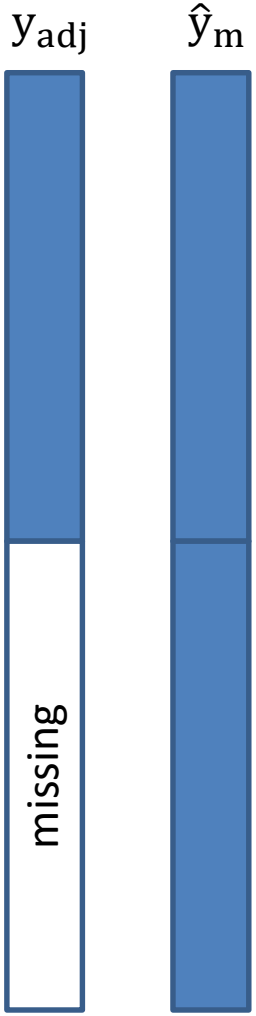
Cross validation schemes

MBLUP

$$\mathbf{y} = \mathbf{Z}_l\mathbf{l} + \mathbf{Z}_\mathrm{m}\mathbf{m} + \mathbf{e}$$

$$\hat{\mathbf{y}}_\mathrm{m} = \hat{\mathbf{m}}$$

Cross validation schemes



Bivariate GBLUP

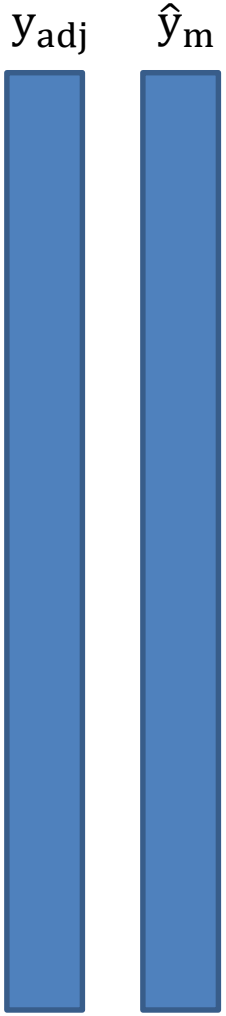
$$\mathbf{y} = \mathbf{Z}_l\mathbf{l} + \mathbf{Z}_g\mathbf{g} + \mathbf{e}$$

$$\hat{\mathbf{y}}_m = \mathbf{Z}_l\mathbf{l} + \mathbf{Z}_g\mathbf{g} + \mathbf{e}$$

Cross-validated:

Bivariate $\hat{\mathbf{g}}_{cv}$

Proxies for "True" genetic values



The adjusted observations themselves:

$$\mathbf{y}$$

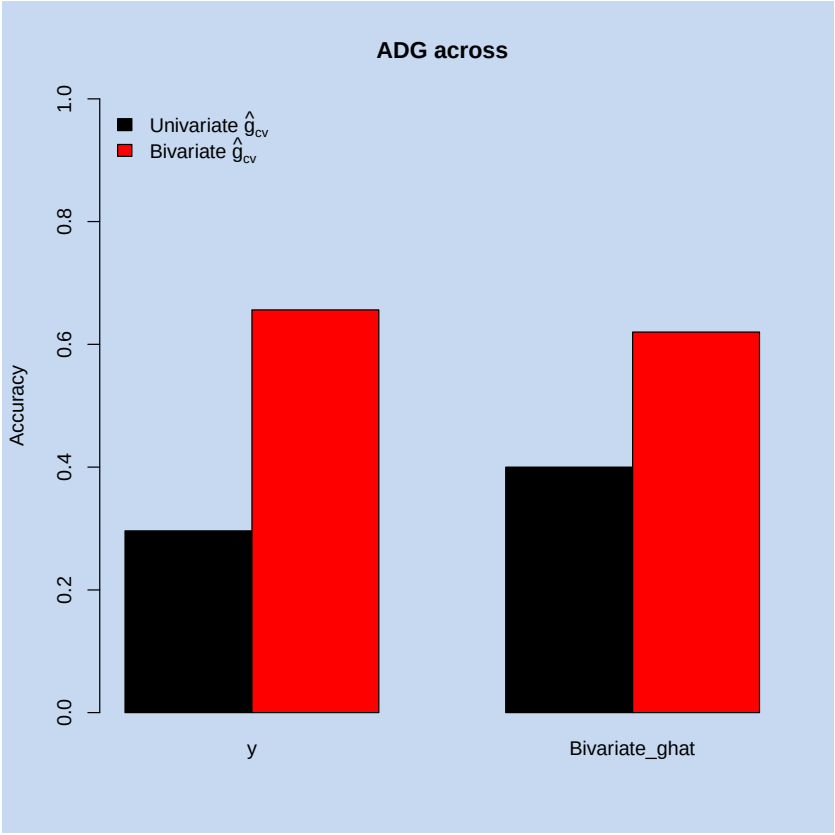
accuracy = cor($\mathbf{y}$, $\hat{\mathbf{g}}_{cv}$)/h

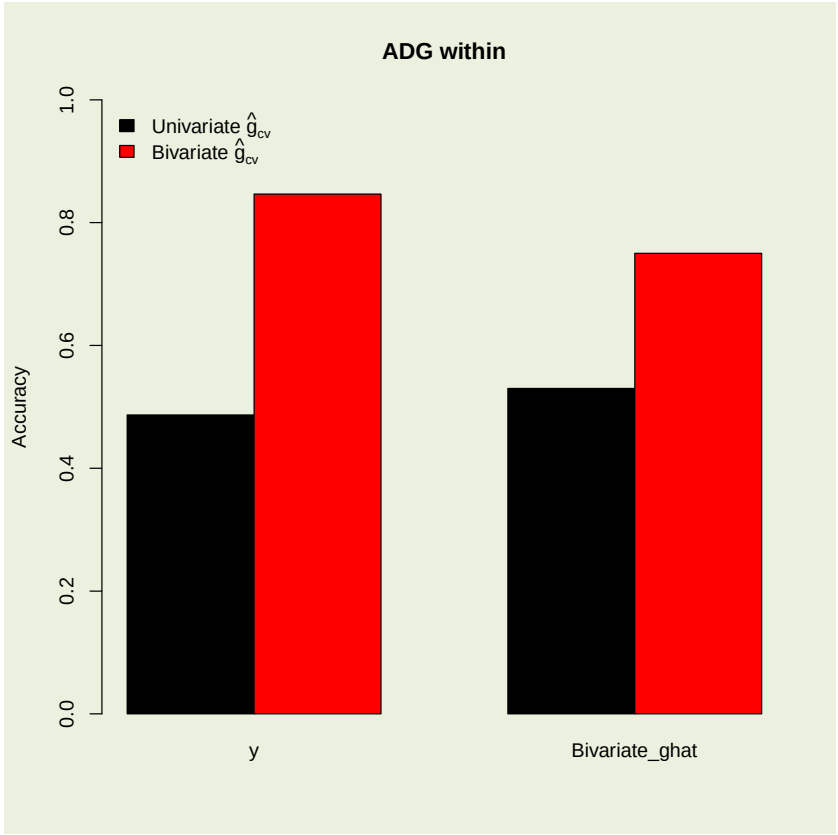
Bivariate GBLUP with all information:

$$\text{Bivariate } \hat{\mathbf{g}}$$

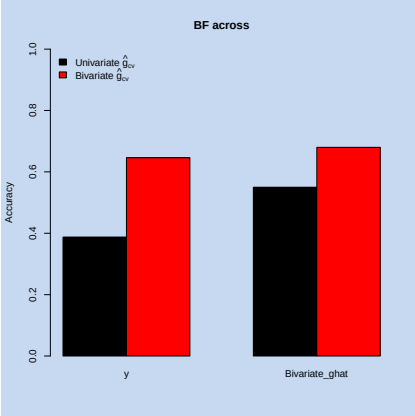
accuracy = cor(Bivariate $\hat{\mathbf{g}}$, $\hat{\mathbf{g}}_{cv}$)

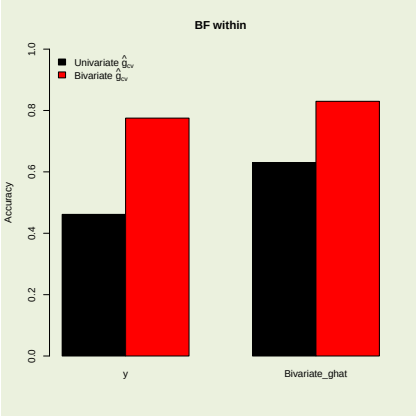
Accuracy ADG

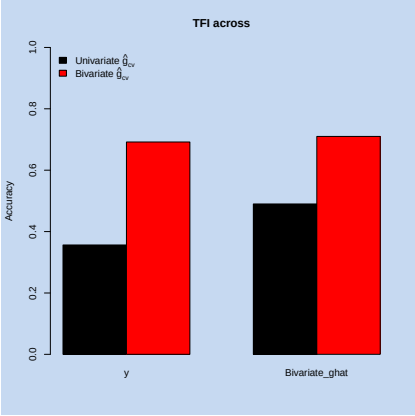


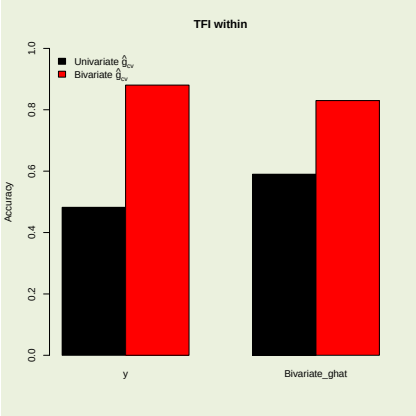


Accuracy


BF across
Univariate $\hat{g}_{cv}$
Bivariate $\hat{g}_{cv}$
Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
y
Bivariate_ghat


BF within
Univariate $\hat{g}_{cv}$
Bivariate $\hat{g}_{cv}$
Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
y
Bivariate_ghat


TFI across
Univariate $\hat{g}_{cv}$
Bivariate $\hat{g}_{cv}$
Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
y
Bivariate_ghat


TFI within
Univariate $\hat{g}_{cv}$
Bivariate $\hat{g}_{cv}$
Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
y
Bivariate_ghat

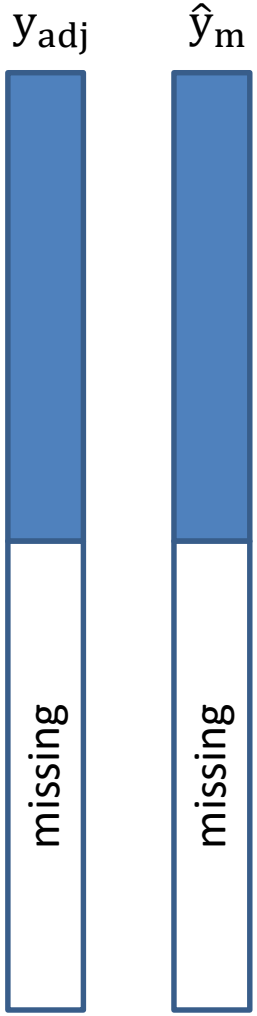
Conclusions

20-50% increased accuracy of EBVs for all three traits with added information from NMR metabolomics.

Increased accuracy of EBVs for all three traits when predicting within environment and family (within) – Benefit from NMR information remains!

Questions?

Only NMR data in training population



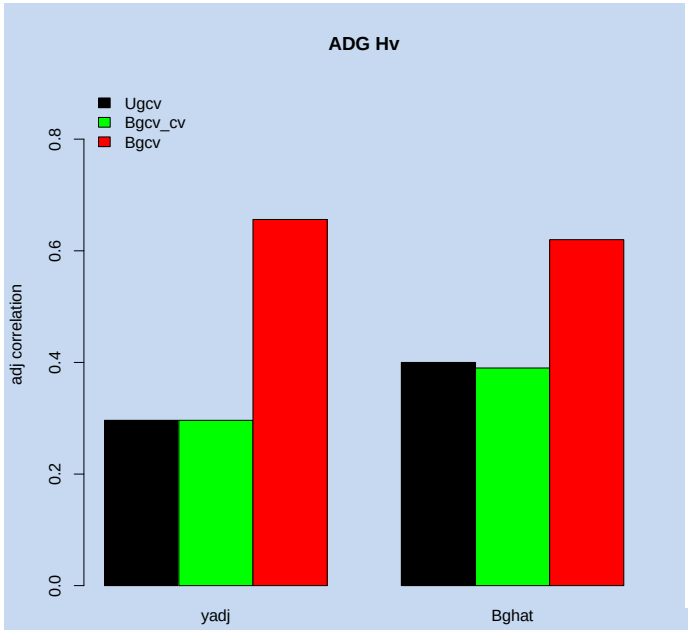
Bivariate GBLUP

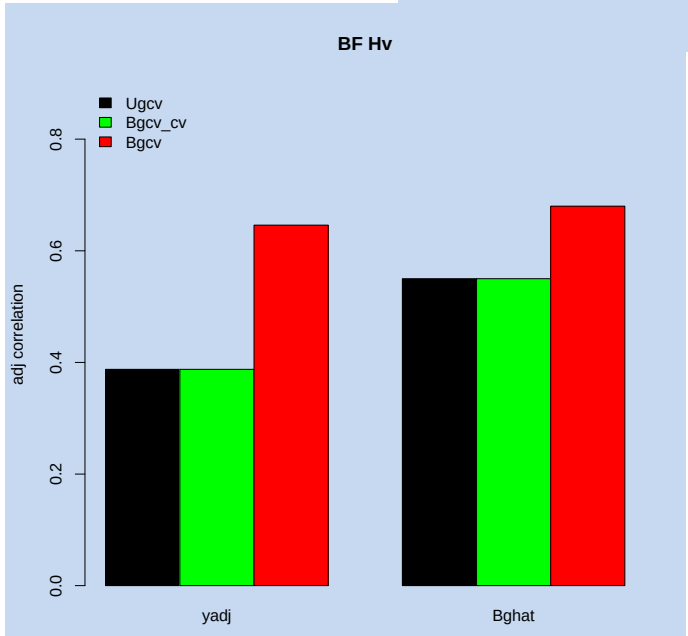
$$\mathbf{y}_{adj} = \mathbf{Z}_l\mathbf{l} + \mathbf{Z}_g\mathbf{g} + \mathbf{e}$$

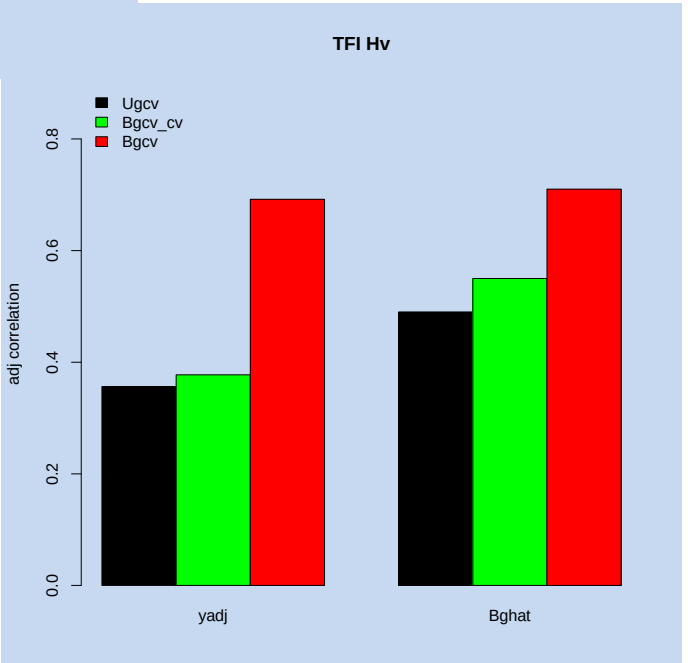
$$\hat{\mathbf{y}}_m = \mathbf{Z}_l\mathbf{l} + \mathbf{Z}_g\mathbf{g} + \mathbf{e}$$

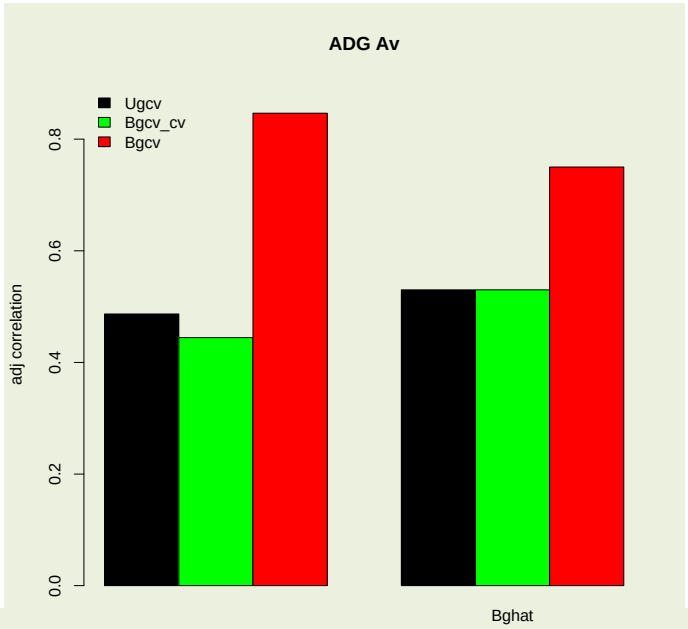
Cross-validated:

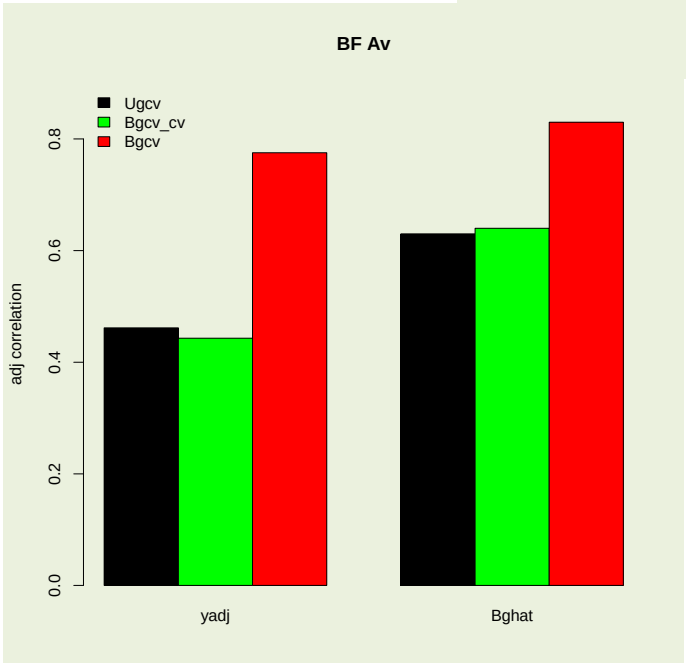
$$B\hat{\mathbf{g}}_{adj,cv,cv}$$

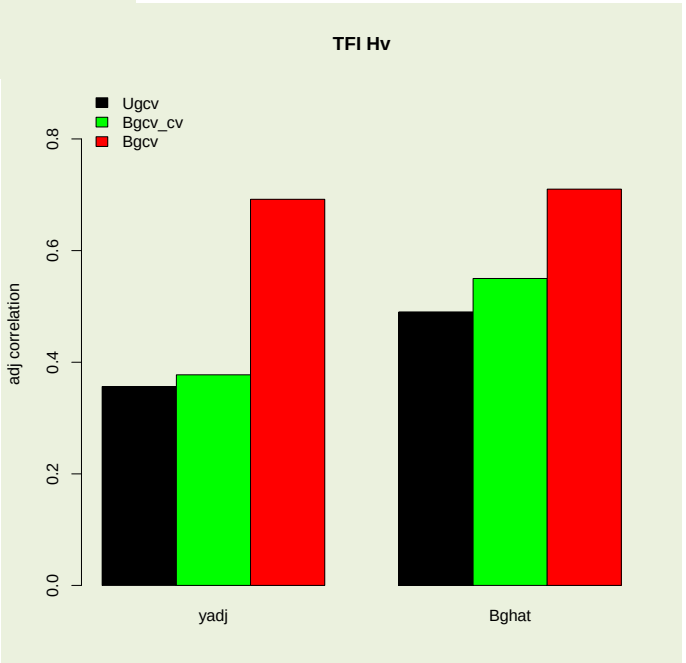

ADG Hv
Ugcv
Bgcv_cv
Bgcv
adj correlation
0.0
0.2
0.4
0.6
0.8
yadj
Bghat


BF Hv
Ugcv
Bgcv_cv
Bgcv
adj correlation
0.0
0.2
0.4
0.6
0.8
yadj
Bghat


TFI Hv
Ugcv
Bgcv_cv
Bgcv
adj correlation
0.0
0.2
0.4
0.6
0.8
yadj
Bghat


ADG Av
Ugcv
Bgcv_cv
Bgcv
adj correlation
0.0
0.2
0.4
0.6
0.8
Bghat


BF Av
Ugcv
Bgcv_cv
Bgcv
adj correlation
0.0
0.2
0.4
0.6
0.8
yadj
Bghat


TFI Hv
Ugcv
Bgcv_cv
Bgcv
adj correlation
0.0
0.2
0.4
0.6
0.8
yadj
Bghat