Genomic Selection Using Best Linear Unbiased Prediction with a Trait Specific Relationship Matrix

Zhe Zhang

Supervisor: Prof. Qin Zhang (CAU)

Dr. DJ de Koning (RI)

China Agricultural University (CAU), Beijing, China

Roslin Institute (RI), University of Edinburgh, Edinburgh, UK

Genomic selection (GS)

- Marker assisted selection on whole genome high density markers.
  - ☐ Require high LD between QTL and markers
- Many advantages over traditional methods
  - ☐ Explain more genetic variance,
  - ☐ Higher accuracy
  - ☐ Shorter generation interval
- Cost effective and widely implemented in dairy cattle genetic evaluation.
  - ☐ May be not cost effective for other livestock species

Genomic selection: models

- Marker (segment) effects model:

$$y = Xb + \sum_{i=1}^{m} Z_i g_i + e \qquad (1)$$

- Total genetic effect model:

$$y = Xb + Zu + e \qquad (2)$$

Assumption on the genetic architecture of trait of interest	Model 1	Model 2
Infinitesimal	RR-BLUP	G-BLUP
Non-infinitesimal	BayesA, B,C,…	**TA-BLUP**

Relationship matrix in MME

$$y = Xb + Zu + e$$

$$Vu = \begin{cases} A\ \sigma_u^2 & \rightarrow \text{traditional BLUP} \leftarrow \text{pedigree} \\ \mathbf{G}\ \sigma_u^2 & \rightarrow \text{G-BLUP} \leftarrow \text{markers} \\ \mathbf{T}_a \sigma_u^2 & \rightarrow \text{TA-BLUP} \leftarrow \text{markers \& weights} \end{cases}$$

A: numerator relationship matrix

G: realized relationship matrix

TA: trait specific relationship matrix

Trait specific relationship matrix (TA)

$$S_{ijk} = \sum_{m=1}^{2} \sum_{n=1}^{2} I_{mn} \bigg/ 4$$

$$S_{ij} = \sum_{k=1}^{N} S_{ijk} w_k \bigg/ \sum_{k=1}^{N} w_k$$

S_{ijk}	AA	Aa	aa
AA	1	0.5	0
Aa	0.5	0.5	0.5
aa	0	0.5	1

G-BLUP: $W_k=1 \rightarrow$ G

TA-BLUP: $W_k=|g_k| \rightarrow$ TA

Aim of this study

- Assess the performance of TA-BLUP using the common dataset of this workshop.

  - Trait Q only.

- Implement TA-BLUP in low density marker chips.

  - Low density marker were selected based on their size of estimated effects.

Methods

- ## BayesB
  - □ Bayesian variable selection method B (*Meuwissen et al. 2001*)

- ## RRBLUP
  - □ Ridge regression BLUP (*Meuwissen et al. 2001*)

- ## TABLUP
  - □ TA**B**: TABLUP with weights from Bayes**B**
  - □ TA**P**: TABLUP with weights from RRBLU**P**.

Methods

Step 1: marker effect estimation

  BayesB, RRBLUP → GEBV

Step 2: construct TA matrix

  All markers + w → TA → TABLUP → GEBV

  Subset markers + w → TA → TABLUP → GEBV

Distribution of estimated marker effects
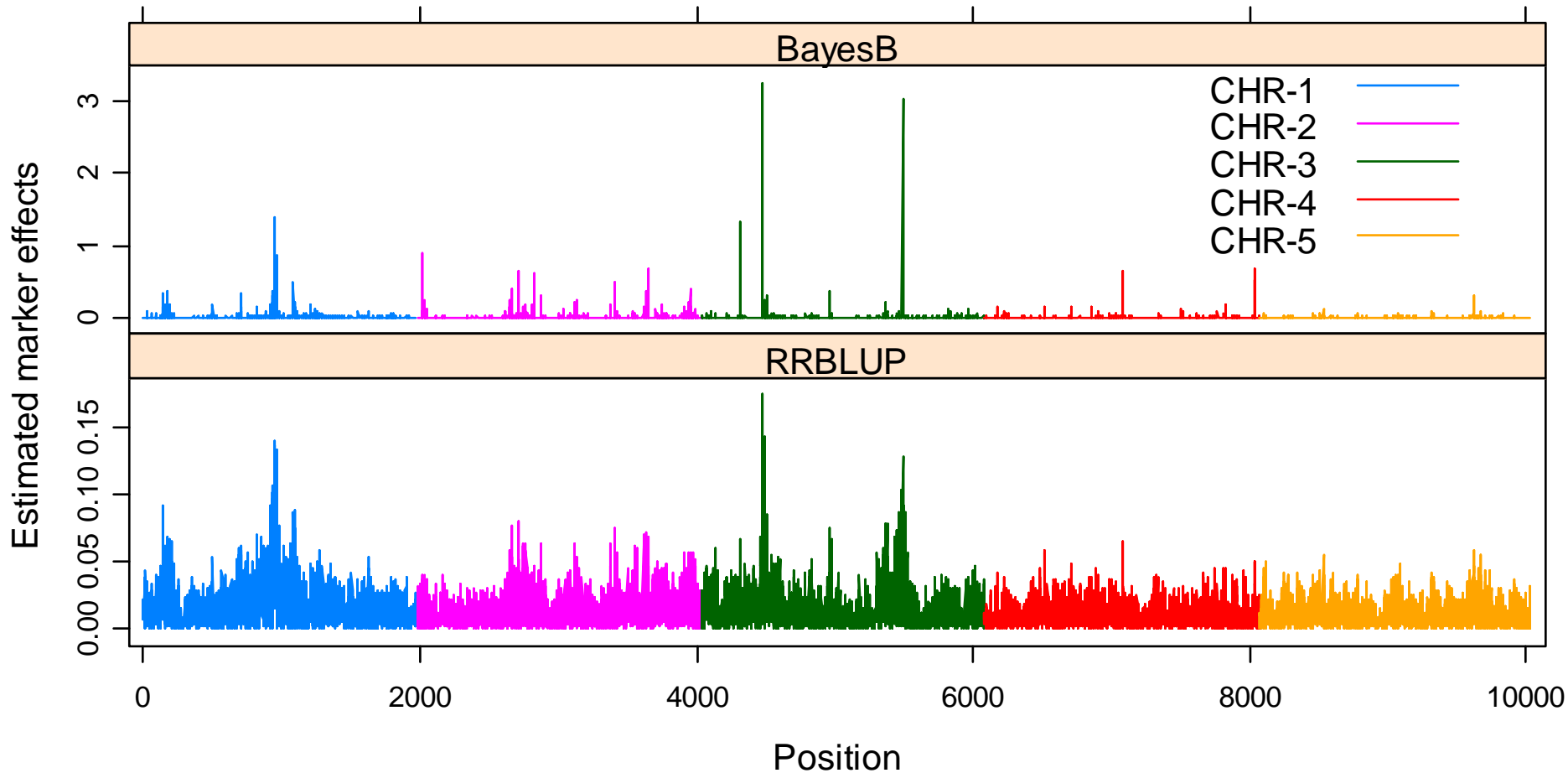
Correlations of GEBVs between methods

	RR-BLUP	BayesB	TA-BLUP	
			TA-P	TA-B
RR-BLUP		0.980	0.977	0.969
BayesB	0.938		0.969	0.981
TA-P	0.985	0.959		0.992
TA-B	0.942	0.999	0.962	

Upper triangle: correlation in reference population (F0 - F3)
Lower triangle: correlation in the young individuals (F4)

Correlation of GEBVs between TA-BLUP using low density markers and RR-BLUP or BayesB using all markers
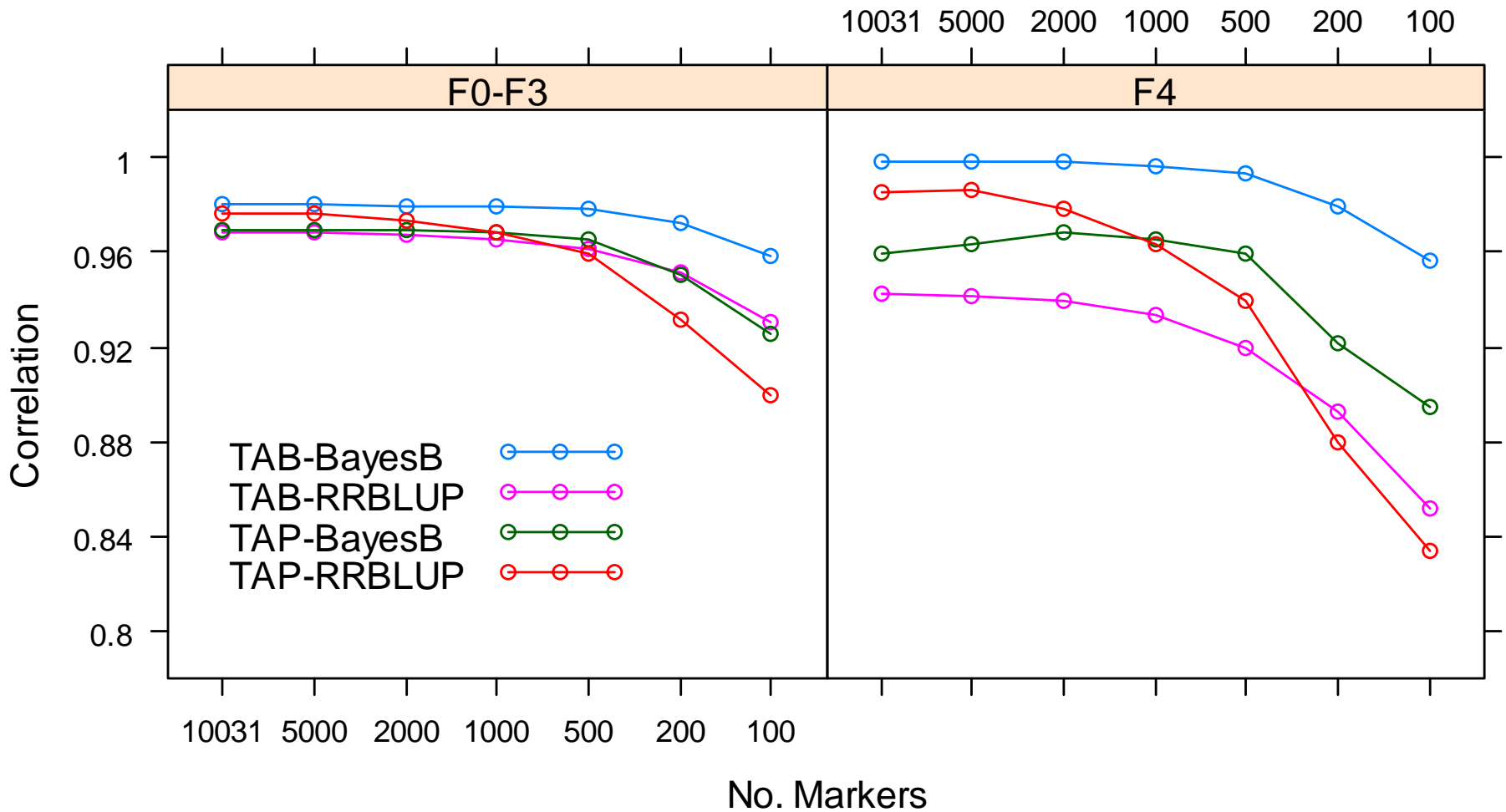
Conclusions

- The trait (Q) simulated in the common dataset obviously departed from the infinitesimal model.

  - Bayesian method should show advantage over RR-BLUP.

- TA-BLUP performed as well as BayesB with all markers. (BayesB, TA-BLUP)

- With only a proportion of all markers, TA-BLUP can perform nearly equivalent to HD-BayesB.

  - GS with low density markers might be cost-effective by using TA-BLUP. (TA-BLUP 500)

Acknowledgements