

## Introduction

- APY algorithm can be used to compute  $\mathbf{G}^{-1}$  in GBLUP and ssGBLUP for large-scale genomic evaluations
- Do we need to include all genotyped animals?
- Indirect Predictions (IP) can be used for:
  - Interim evaluations
  - Predictions for animals not in the main evaluation
- There is need for a measure of accuracy to release along with IP

## Objectives

Compute accuracy for IP based on prediction error covariance (PEC) of SNP effects from ssGBLUP using direct inversion of  $\mathbf{G}$  or the APY algorithm

## Materials and Methods

- American Angus Association data

### Complete dataset (Benchmark GEBV and accuracy)

35k post-weaning gain (PWG) phenotypes up to 2013  
192k animals in the pedigree up to 2013  
60k genotyped animals born up to 2014  
GEBV accuracy was calculated based on PEV

### Reduced dataset (SNP effects and PEC)

35k post weaning gain phenotypes up to 2013  
192k animals in the pedigree up to 2013  
58k genotyped animals born up to 2013

### Validation animals (IP and accuracy for validation animals)

2k genotyped animals born in 2014

### ssGBLUP with APY

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{W} \\ \mathbf{W}'\mathbf{X} & \mathbf{W}'\mathbf{W} + \mathbf{H}^{-1}\lambda \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{W}'\mathbf{y} \end{bmatrix} \quad \mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{cc} & \mathbf{G}_{cn} \\ \mathbf{G}_{nc} & \mathbf{G}_{nn} \end{bmatrix} \quad \begin{array}{l} c = \text{Core} \\ n = \text{Non-core} \end{array}$$

$$\mathbf{G}_{APY}^{-1} = \begin{bmatrix} \mathbf{G}_{cc}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{G}_{cc}^{-1}\mathbf{G}_{cn} \\ \mathbf{I} \end{bmatrix} \mathbf{M}_{nn}^{-1} \begin{bmatrix} -\mathbf{G}_{nc}\mathbf{G}_{cc}^{-1} & \mathbf{I} \end{bmatrix}$$

$$m_{nn,i} = g_{ii} - G_{ic}\mathbf{G}_{cc}^{-1}G_{ci}$$

### SNP effects

$$\hat{\mathbf{a}} = \alpha\mathbf{b} \frac{1}{2\sum pq} \mathbf{Z}'\mathbf{G}^{-1}\hat{\mathbf{u}}_{\text{reduced}} \quad \text{IP} = \mathbf{Z}\hat{\mathbf{a}}$$

### PEC and accuracy of IP

Duarte et al. 2014  
Liu et al. 2017  
Legarra et al. 2018  
Aguilar et al. 2019

$$\text{var}(\hat{\mathbf{a}}) = \alpha\mathbf{b} \frac{1}{2\sum pq} \mathbf{Z}'\mathbf{G}^{-1}(\mathbf{G}\sigma_u^2 - \mathbf{C}^{22})\mathbf{G}^{-1} \mathbf{Z} \frac{1}{2\sum pq} \alpha\mathbf{b}$$

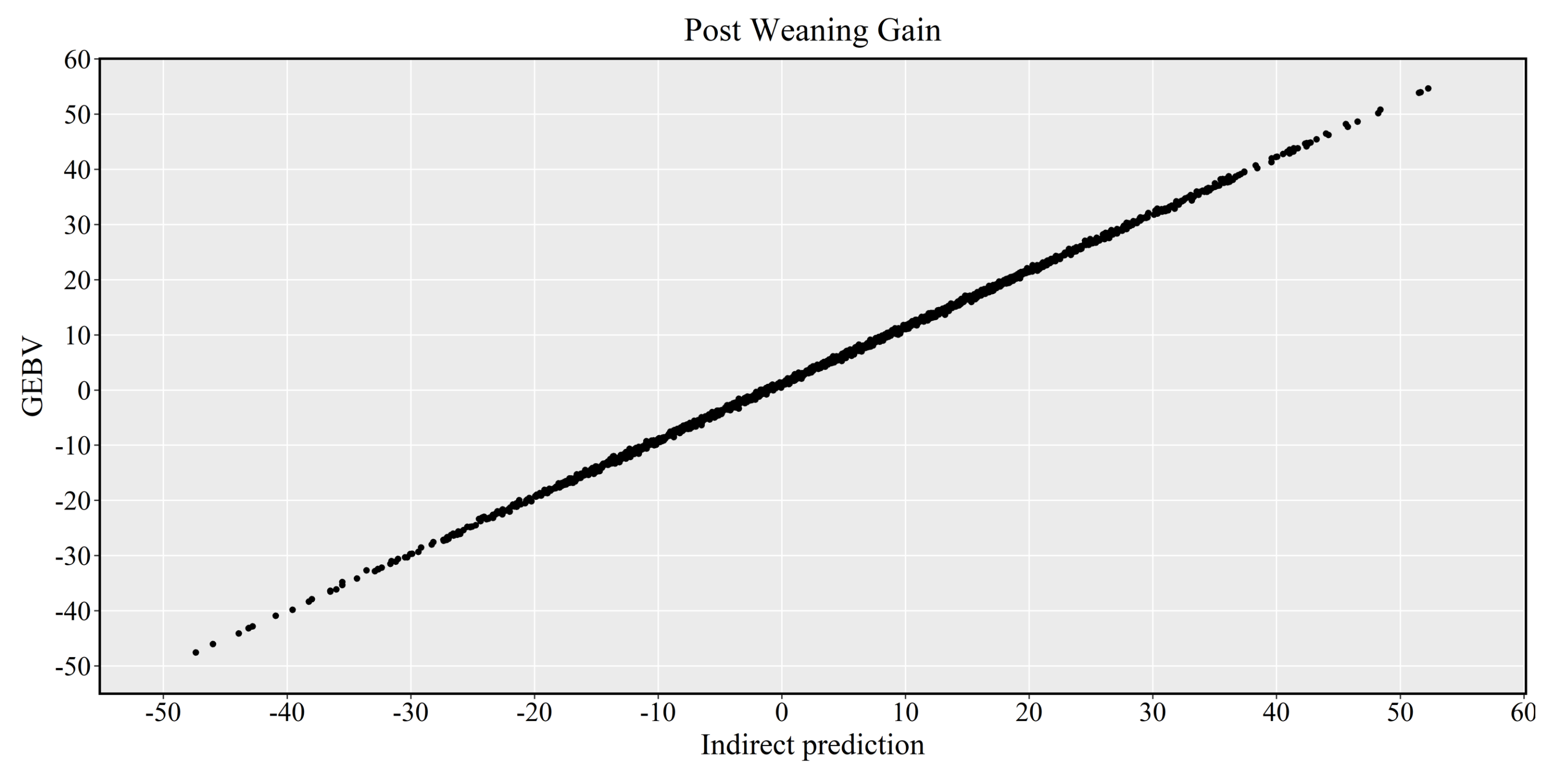
$$\text{IP}_{\text{acc}_i} = \sqrt{\alpha\mathbf{b} \left( 1 - \frac{z_i \text{var}(\hat{a}_i) z_i'}{\sigma_u^2} \right)} \quad \begin{array}{l} \alpha = \text{Blending (VanRaden 2008)} \\ \mathbf{b} = \text{Tuning (Vitezica et al., 2011)} \end{array}$$

- SNP effects and PEC were calculated using  $\mathbf{G}^{-1}$  or  $\mathbf{G}_{APY}^{-1}$
- Correlation between benchmark (GEBV and accuracy) and validation (IP and accuracy of IP) were calculated for 2k validation animals

## Results

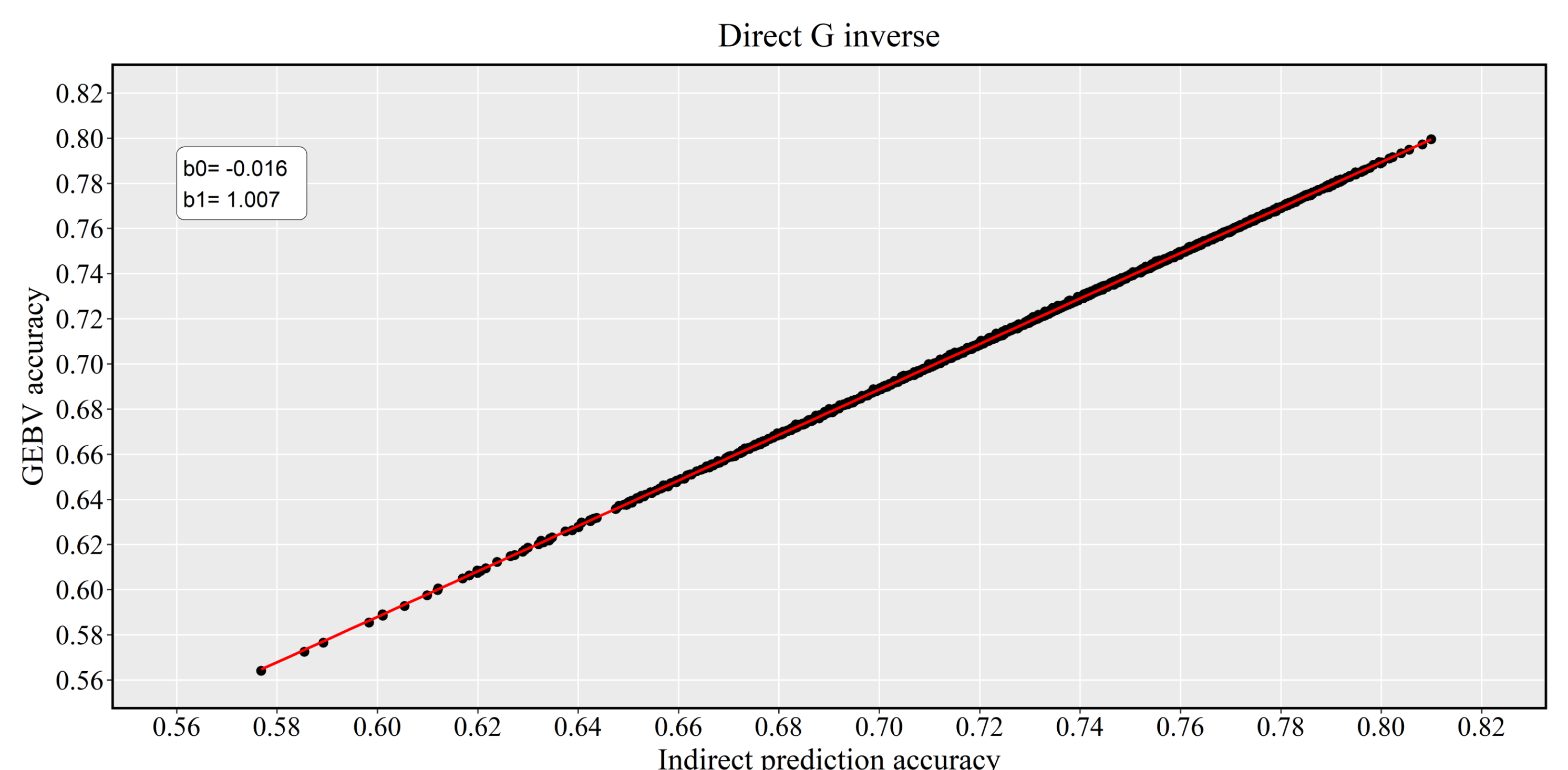
### Correlation between GEBV and IP

- $\geq 0.99$  using either  $\mathbf{G}^{-1}$  or  $\mathbf{G}_{APY}^{-1}$

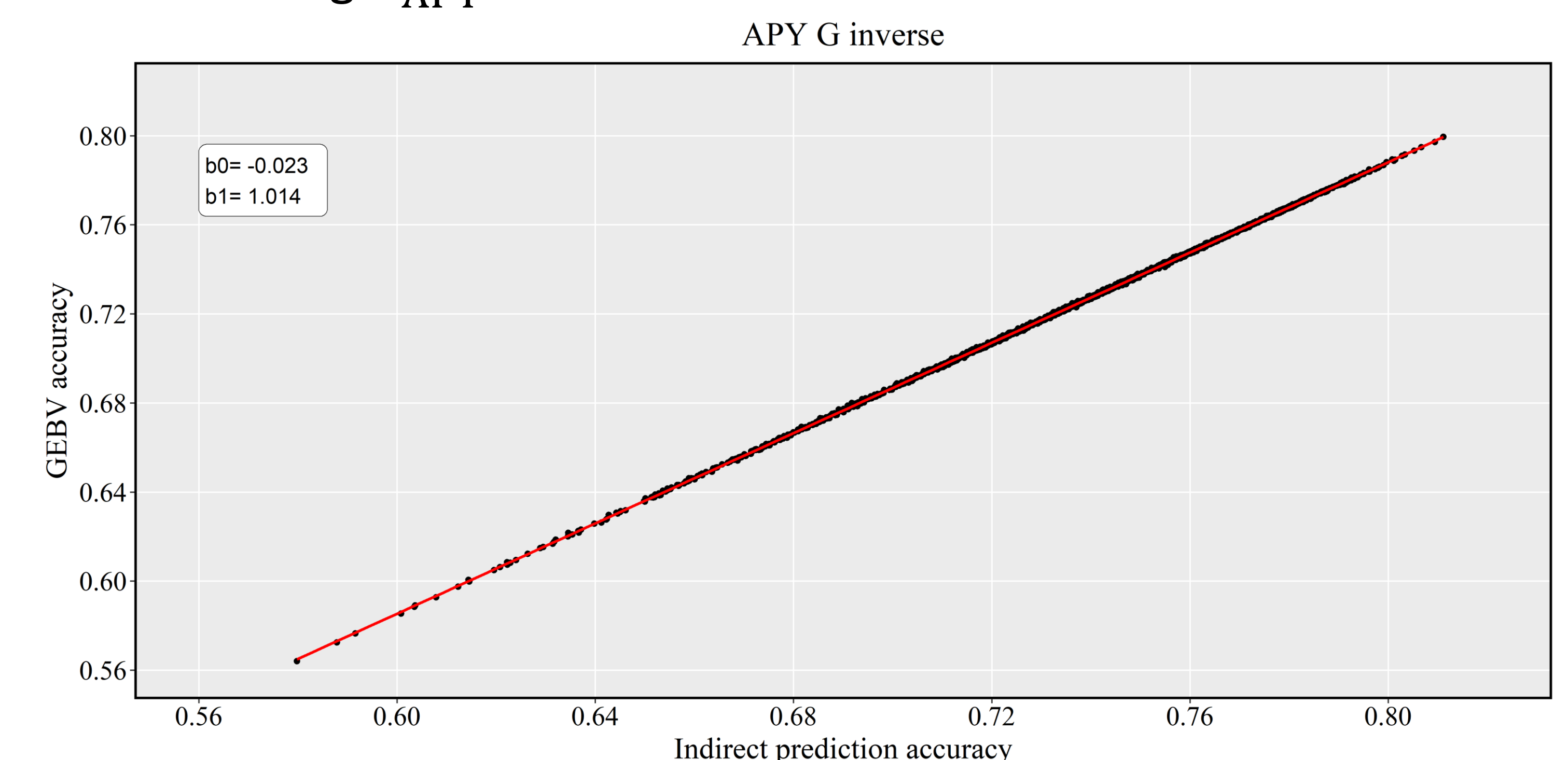


### Accuracy correlations

- $\geq 0.99$  using direct  $\mathbf{G}^{-1}$



- $\geq 0.99$  using  $\mathbf{G}_{APY}^{-1}$



### Considerations for large datasets

- IP accuracies based on prediction error (co)variance depend on the inversion of the LHS of MME, which limits its use for large datasets

## Conclusions

- IP from ssGBLUP can be used as genomic prediction
- Accuracies for IP are available with either direct inversion of  $\mathbf{G}$  or using APY
- Ongoing studies aim to extend the ideas to larger genotyped populations