How to remove bias in genomic predictions ?

Zulma G. Vitezica[1] & Andrés Legarra[2]

PROBLEM: How selection is accounted for in genomic evaluation is unclear and cause bias

- Unbiased predictions are of paramount importance in selection
  - Accurate estimates of the genetic trend
  - Comparison of animals across generations

OBJECTIVE: propose a method to remove bias of genomic prediction

- Predictions by a single-step method
  - Based on genomic (**G**) and pedigree-based relationship (**A**) matrices
- For **G** to be correct
  - Base allele frequencies would be required. Unfeasible in practice.
  - Two corrections **α** and **Fst** to refer **G** and **A** to the same base population

SOLUTION: Correction factors

- **α** is simply the difference between means for $\mathbf{A}_{22}$ and **G**

$$\alpha = \frac{1}{n^2}\left[\sum_i \sum_j \mathbf{A}_{22\,(i,j)} - \sum_i \sum_j \mathbf{G}_{i,j}\right]$$

- **α** accounts for the fact that genotyped animals are related through pedigree more than **G** is able to reflect

BLUP$_{ALPHA}$

$$\mathbf{G}^{*} = \mathbf{G} + \mathbf{11}'\alpha$$

- Wright's **Fst** can be defined as the mean relationship between gametes in a recent population with respect to an older base population

$$F_{ST} = \frac{1}{2} mean\left(\mathbf{A}_{22} - \mathbf{G}\right) = \frac{1}{2}\alpha$$

- Powell et al. (2010) suggested to use **Fst** to correctly compare relationships

BLUP$_{Fst}$

$$\mathbf{G}^{**} = \left(1 - \frac{1}{2}\alpha\right)\mathbf{G} + \mathbf{11}'\alpha$$

Simulation RESULTS:

Means of EBV (h^2=0.30) **for selection candidates**

Prediction method	Low selection	High selection
	TBV=0.53 (0.03)	TBV=2.01 (0.15)
Pedigree BLUP	**0.54**	**2.05**
Single Step BLUP$_{ALPHA}$	**0.52**	**2.10**
Single Step BLUP$_{Fst}$	**0.52**	**2.10**
Single Step BLUP, no correction	**0.29**	**1.41**

Conclusion

- Single-step method with correction (either **BLUP$_{ALPHA}$** or **BLUP$_{Fst}$**) was a preferred method for accounting for bias in genomic predictions

[1]INRA, UMR 1289 TANDEM, 31326 Castanet-Tolosan, France, zulma.vitezica@ensat.fr

[2]INRA, UR 631, 31326 Castanet-Tolosan, France, andres.legarra@toulouse.inra.fr