The potential of genotyping pooled DNA to leverage commercial phenotypes for genetic improvement of beef cattle

R.Ventura, L. F. Brito,G. Vandervoort, M. Sargolzaei, A. Hess, M. McMorris, T. Caldwell, A. Canovas, K. G. Dodds, S. Clarke, S. P. Miller

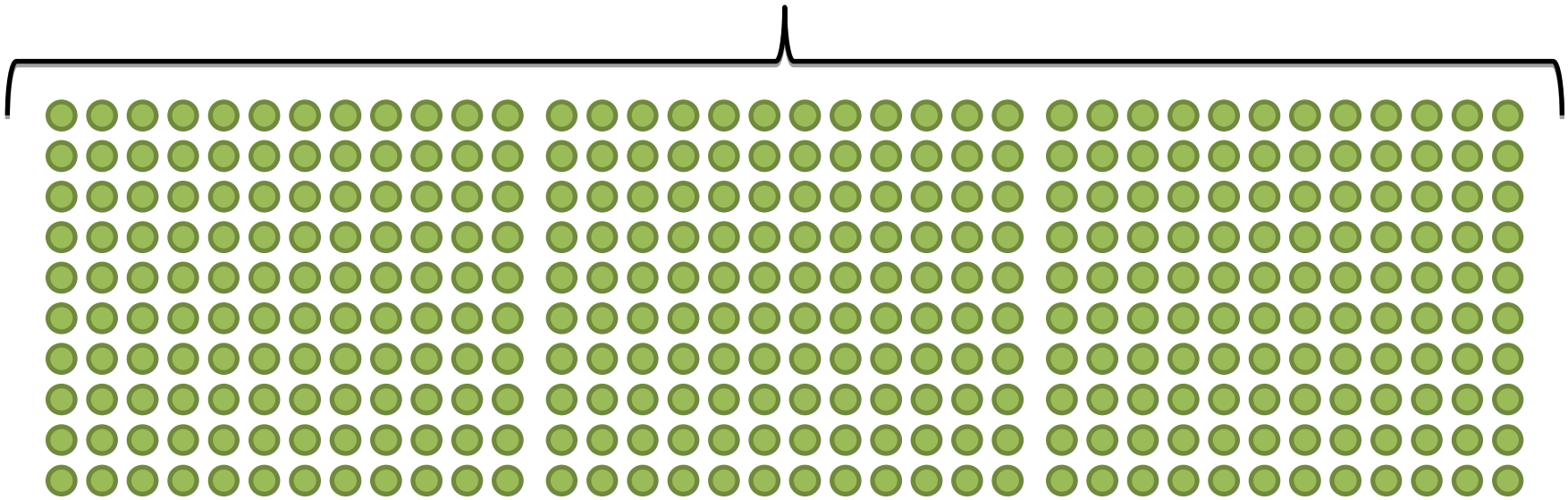
Genomic selection is now a reality in beef cattle,
and require a large group of animals to be individually genotyped

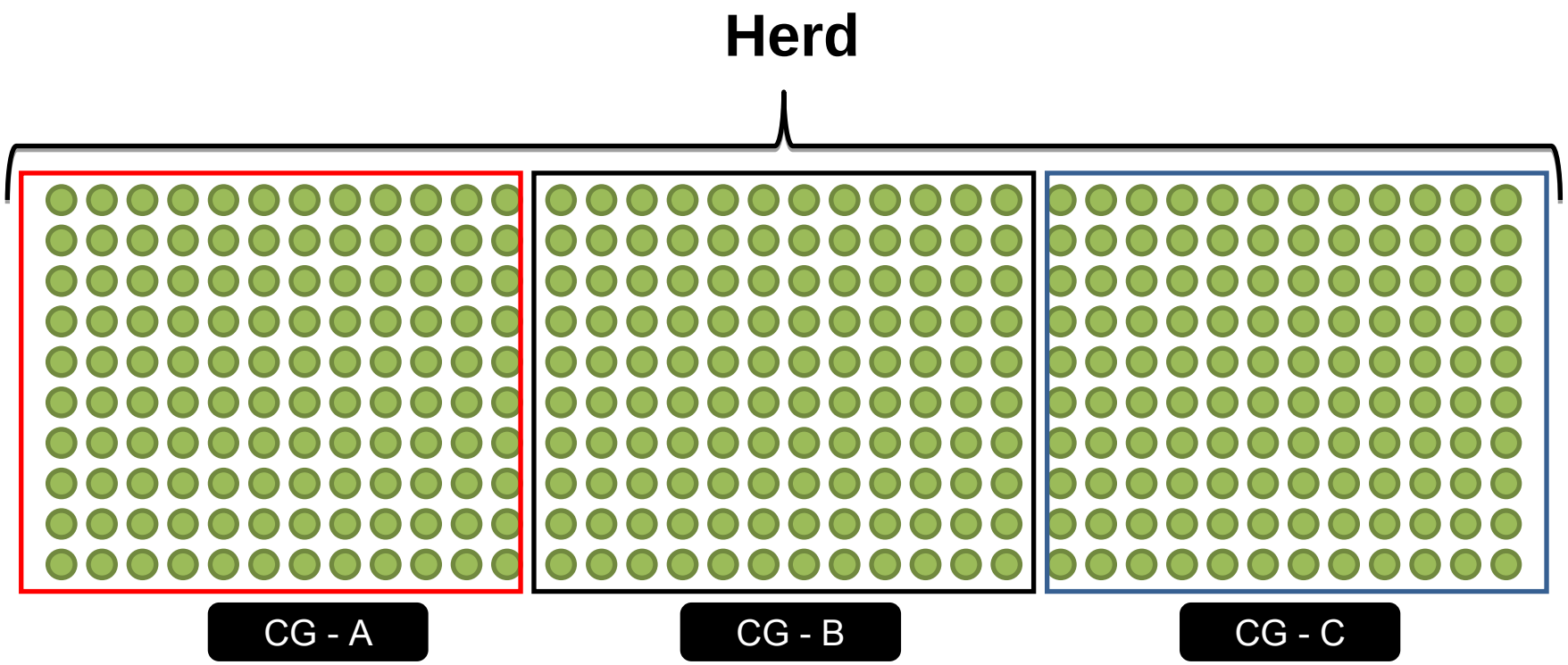
Genotyping pooled DNA :

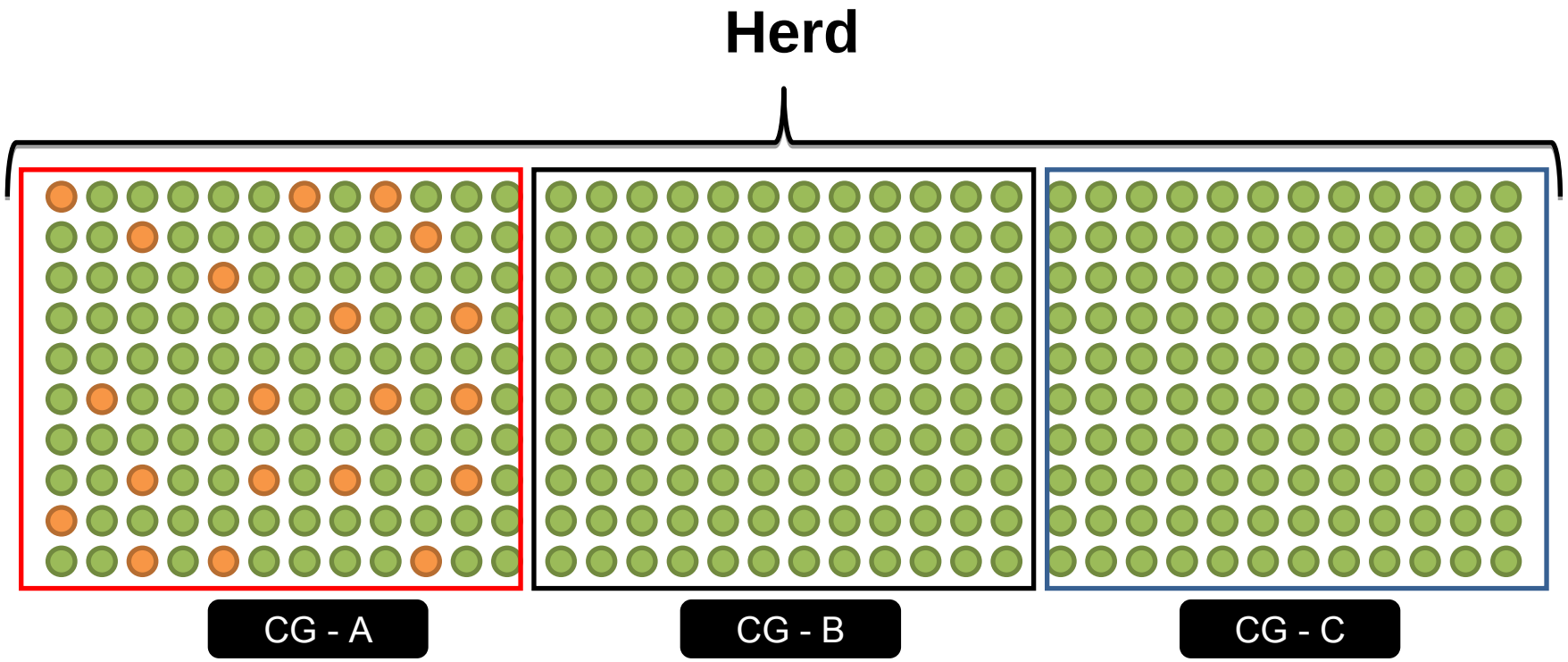
Group animals based on the Phenotype similarity, mix the biological material and genotype the entire group only once

Could we build a reference population
in a economical way?
(first look on traits only recorded in the commercial sector)

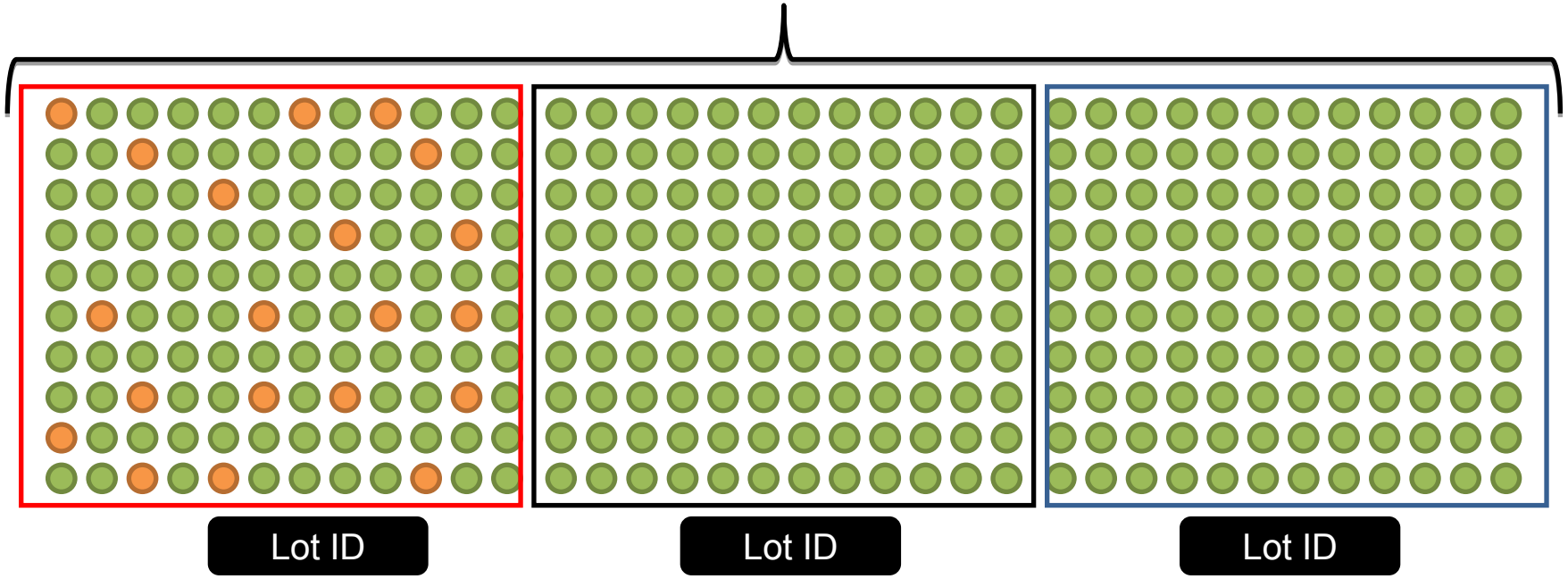
Herd




Herd
CG - A
CG - B
CG - C


Herd
CG - A
CG - B
CG - C
Extremely -
Extremely +
25 individuals in each pool
sorted based on the phenotype
25 individuals in each pool
sorted based on the phenotype

Packing plant



Extremely -

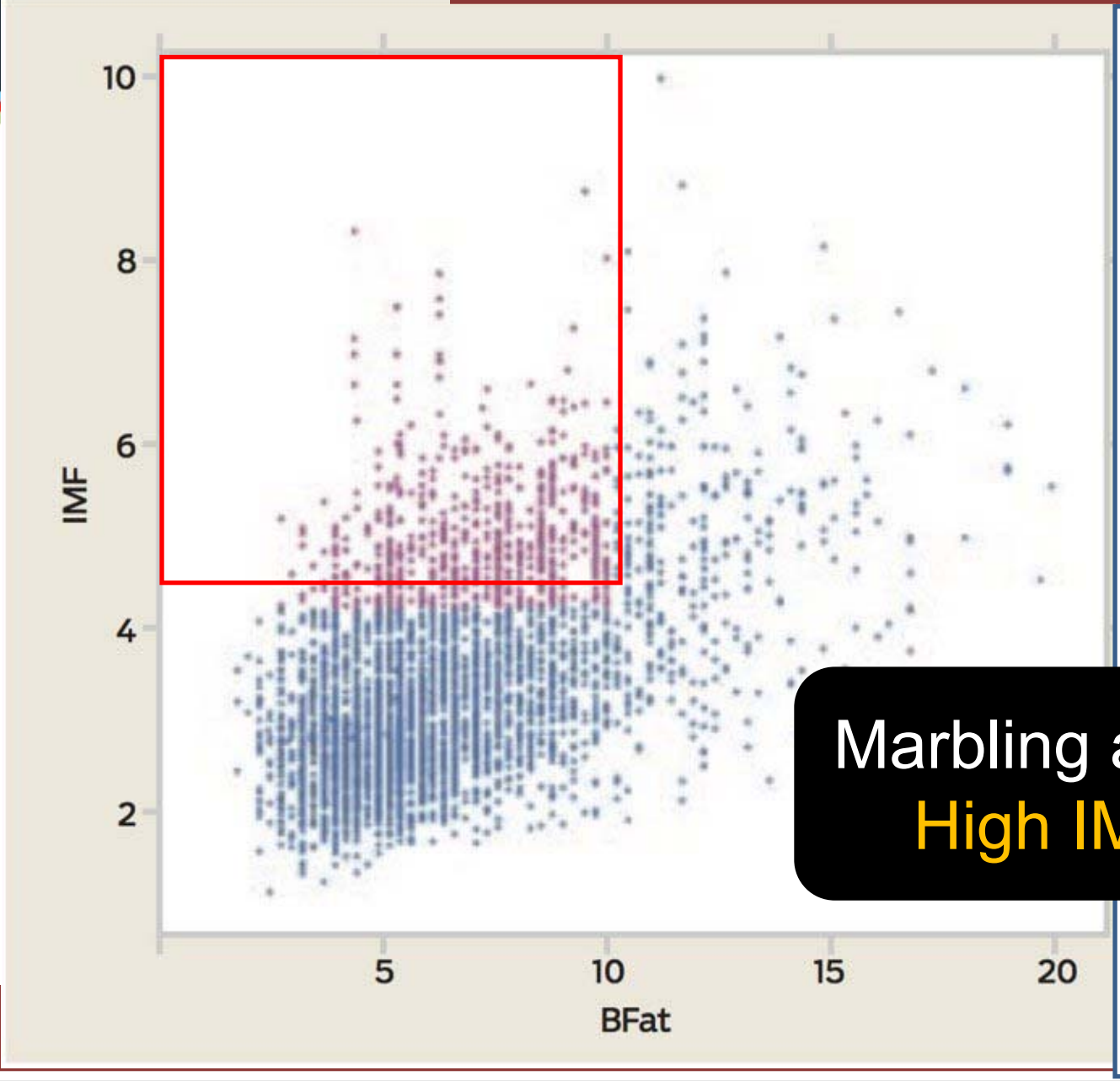
Extremely +

25 individuals in each pool
sorted based on the phenotype

25 individuals in each pool
sorted based on the phenotype

GENETIC POTENTIAL

The real cost of fat (Mike McMorris, 2013)



Marbling and Yield grade

High IMF / Low IMF

https://www.canadiancattlemen.ca/2016/07/13/the-real-cost-of-fat-in-beef-cattle/

Material and Methods

157,870 registered American Angus animals:
Ultrasound intramuscular fat phenotypes
Another set of economically important traits (ongoing)
87K imputed genotypes

Pools

Simulated the pools by averaging the allele frequencies on individually genotyped animals in a group

Material and Methods

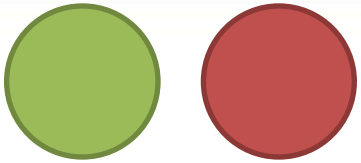
157,870 registered American Angus animals:
Ultrasound intramuscular fat phenotypes
Another set of economically important traits (ongoing)
87K imputed genotypes

Genotypes of pooled DNA samples:

(Reverter et al., 2016)

SNP genotypes in the pools were categorized into to “0”, “1”, and “2” genotypes based on their B-allele frequencies

1) if the B-allele frequency was ≤0.3, then SNP genotype was assigned to a “0”
2) if the B-allele frequency was >0.3 and ≤0.7, then SNP genotype was assigned to “1”
3) if the B-allele frequency was >0.7 and ≤1.0, then SNP genotype was assigned to “2.”



(1000 Low and 1000 High IMF)

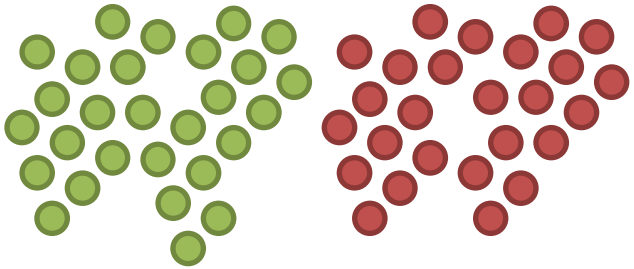
157,870 original phenotypes

Top 1,000 PCTIMF

Bottom 1,000 PCTIMF

Each set of 1,000 animals subdivided into 40 groups of 25 individuals

Each set of 1,000 animals subdivided into 40 groups of 25 individuals

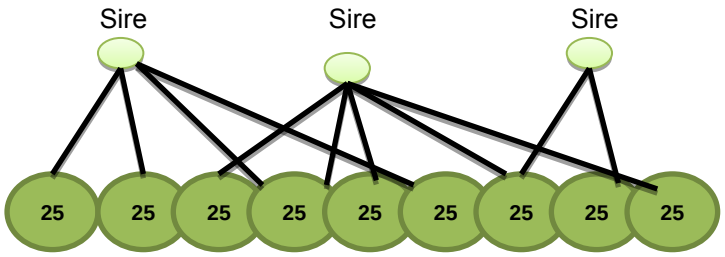


40 Pools
25 animals Low IMF

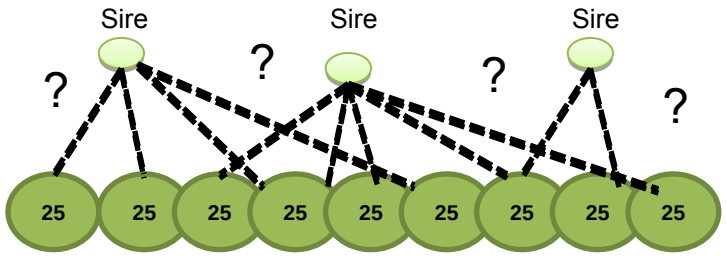
40 Pools
25 animals High IMF

40 H + 40 L as reference population to obtain predictions for the sires

All sires genotyped
Sire-progeny relationship



Sires genotyped
No relationship previously identified



J Anim Sci. 2016 Oct;94(10):4096-4108. doi: 10.2527/jas.2016-0675.

Genomic analyses of tropical beef cattle fertility based on genotyping pools of Brahman cows with unknown pedigree.

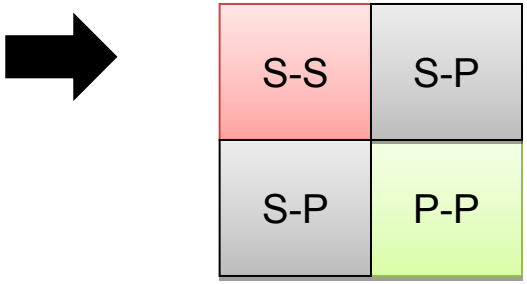
Reverter A, Porto-Neto LR, Fortes MR, McCulloch R, Lyons RE, Moore S, Nicol D, Henshall J, Lehnert SA.

Abstract

We introduce an innovative approach to lowering the overall cost of obtaining genomic EBV (GEBV) and encourage their use in commercial extensive herds of Brahman beef cattle. In our approach, the DNA genotyping of cow herds from 2 independent properties was performed using a high-density bovine SNP chip on DNA from pooled blood samples, grouped according to the result of a pregnancy test following their first and second joining opportunities. For the DNA pooling strategy, 15 to 28 blood samples from the same phenotype and contemporary group were allocated to pools. Across the 2 properties, a total of 183 pools were created representing 4,164 cows. In addition, blood samples from 309 bulls from the same properties were also taken. After genotyping and quality control, 74,584 remaining SNP were used for analyses. Pools and individual DNA samples were related by means of a "hybrid" genomic relationship matrix. The pooled genotyping analysis of 2 large and independent commercial populations of tropical beef cattle was able to recover significant and plausible associations between SNP and pregnancy test outcome. We discuss 24 SNP with significant association (< 1.0 × 10) and mapped within 40 kb of an

GBLUP

“Hybrid G matrix”



Material and Methods

157,870 original phenotypes

Top 1,000 PCTIMF

Bottom 1,000 PCTIMF

Validation Step 1

AAA weekly National cattle evaluation
IMF EPD

Highest 30 sires
IMF EPD

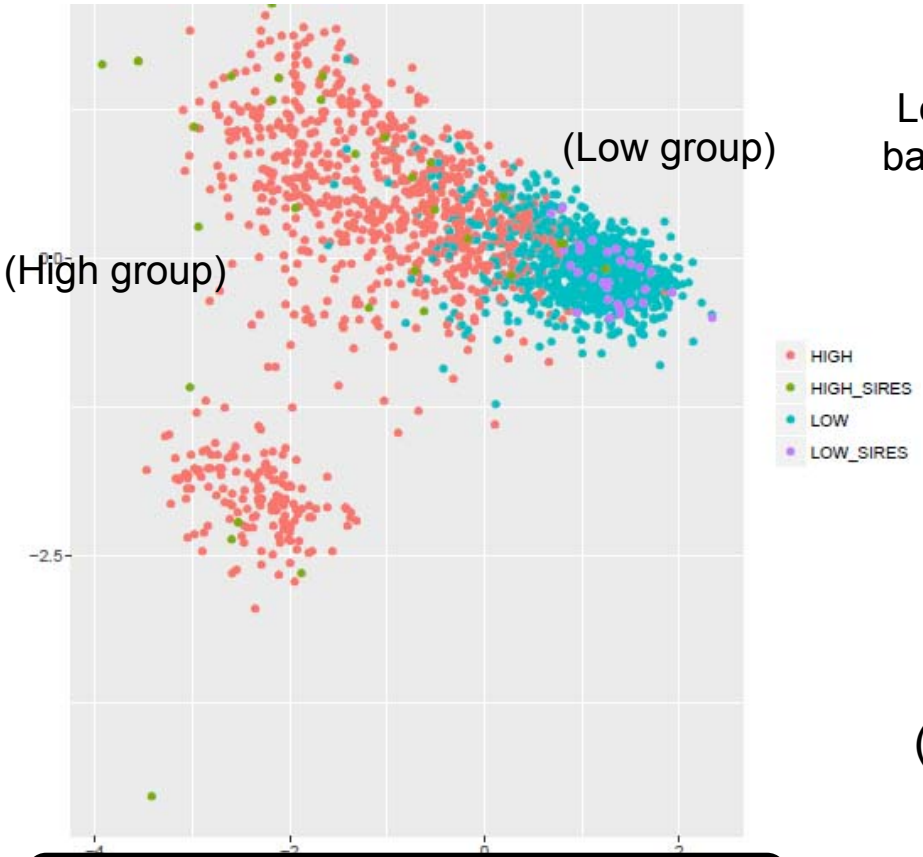
Lowest 30 sires
IMF EPD

Cor (DGV from GBLUP-pool, EPDs at Angus)

Clustered the genotypes before and after pooling

Individual genotypes

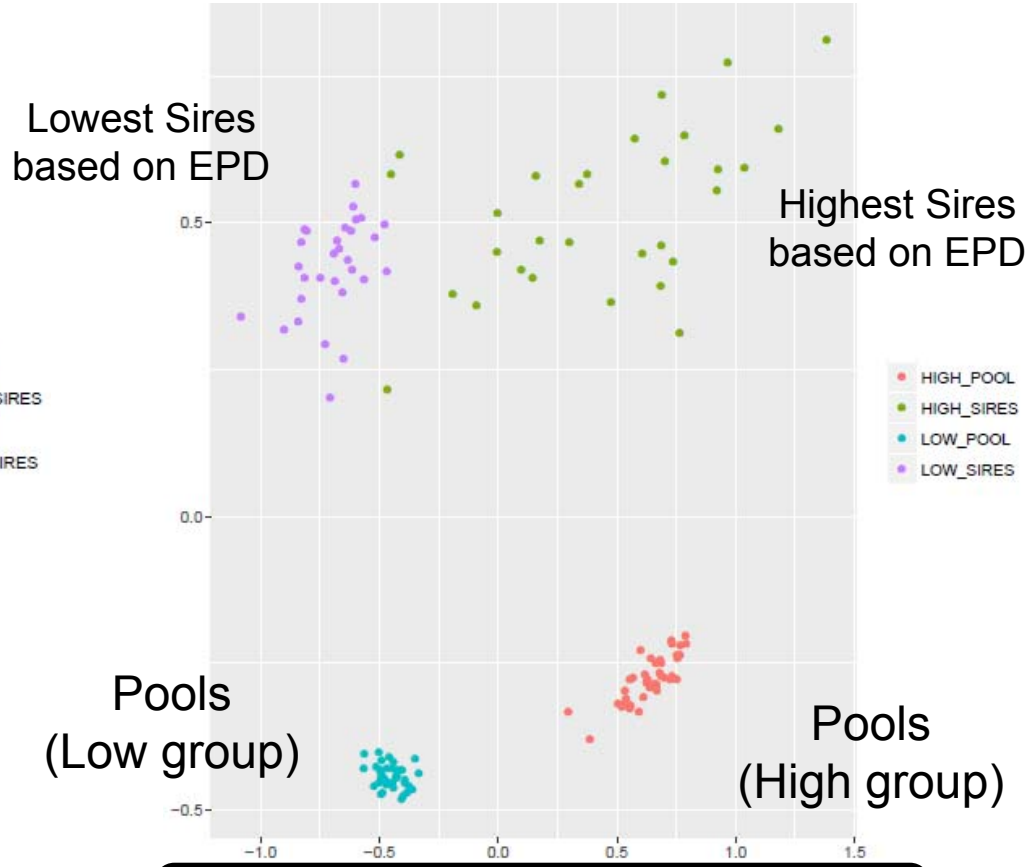
(1000 Low and 1000 High IMF)



1000 H, 1000 L, 30H, 30L

Pool (40 High + 40 Low)

(25 individuals each)



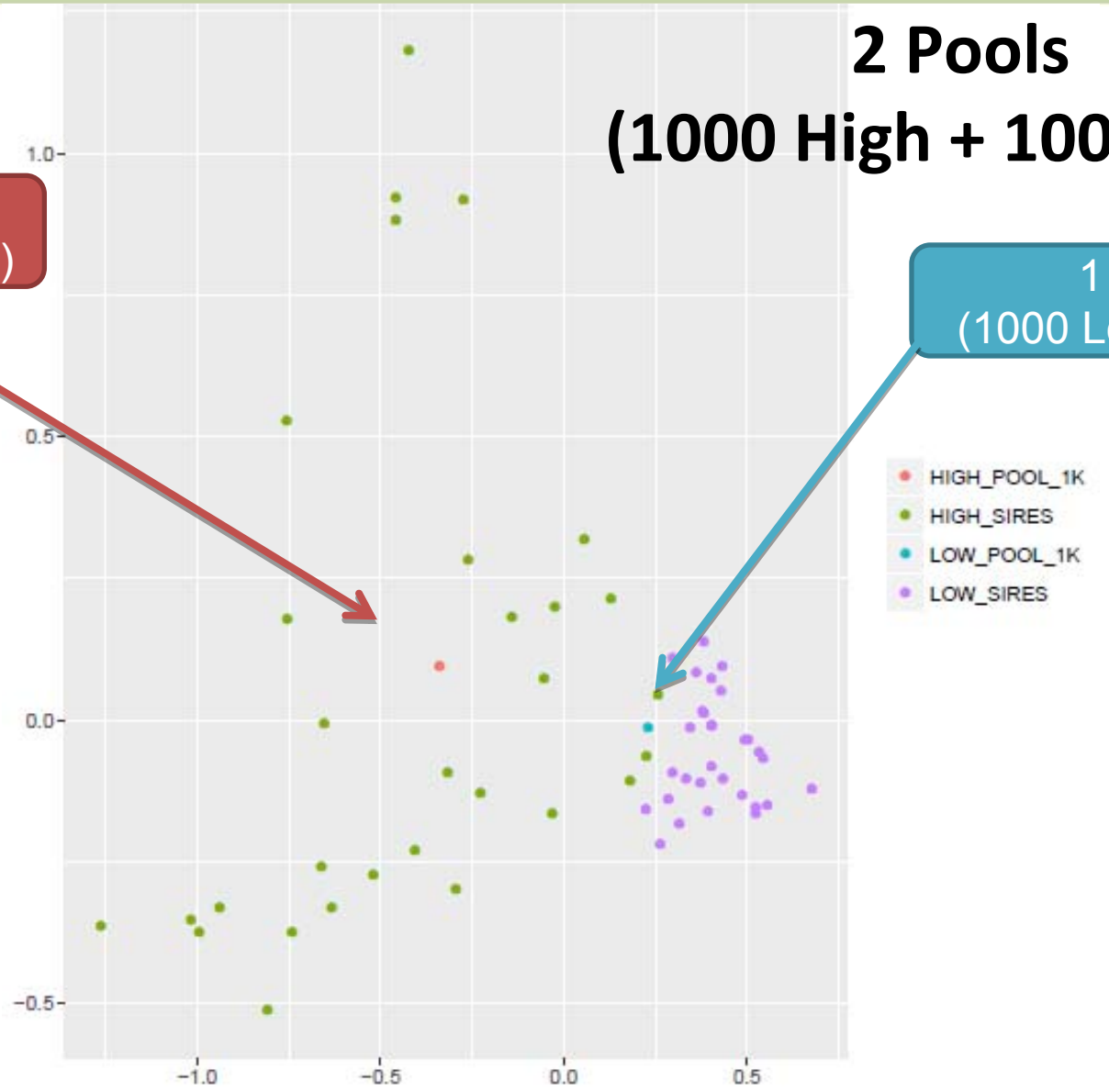
40 Pools H, 40 Pools L, 30H, 30L

Clustered the genotypes before and after pooling

2 Pools
(1000 High + 1000 Low)

1 Pool
(1000 Highest IMF)

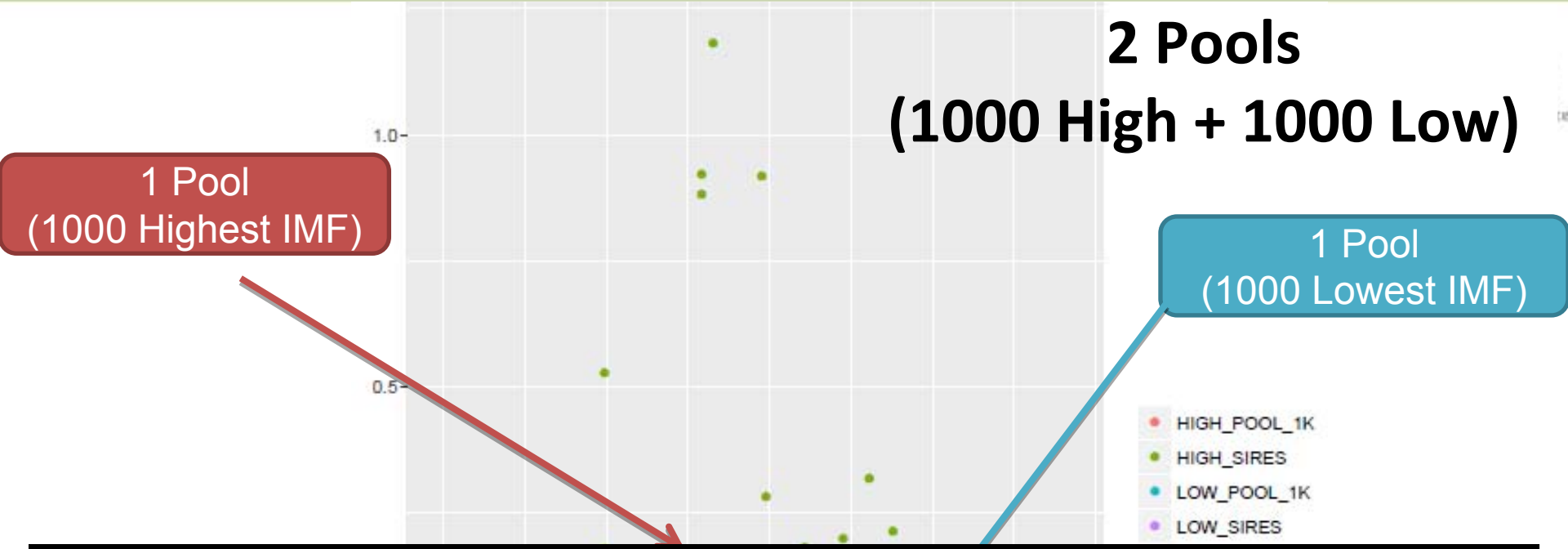
1 Pool
(1000 Lowest IMF)



2 Pools
(1000 High + 1000 Low)

1 Pool
(1000 Highest IMF)

1 Pool
(1000 Lowest IMF)



2 pools approach:
cor(difference in sires genomic relationship with the high and the low groups, EPDs) = 0.858

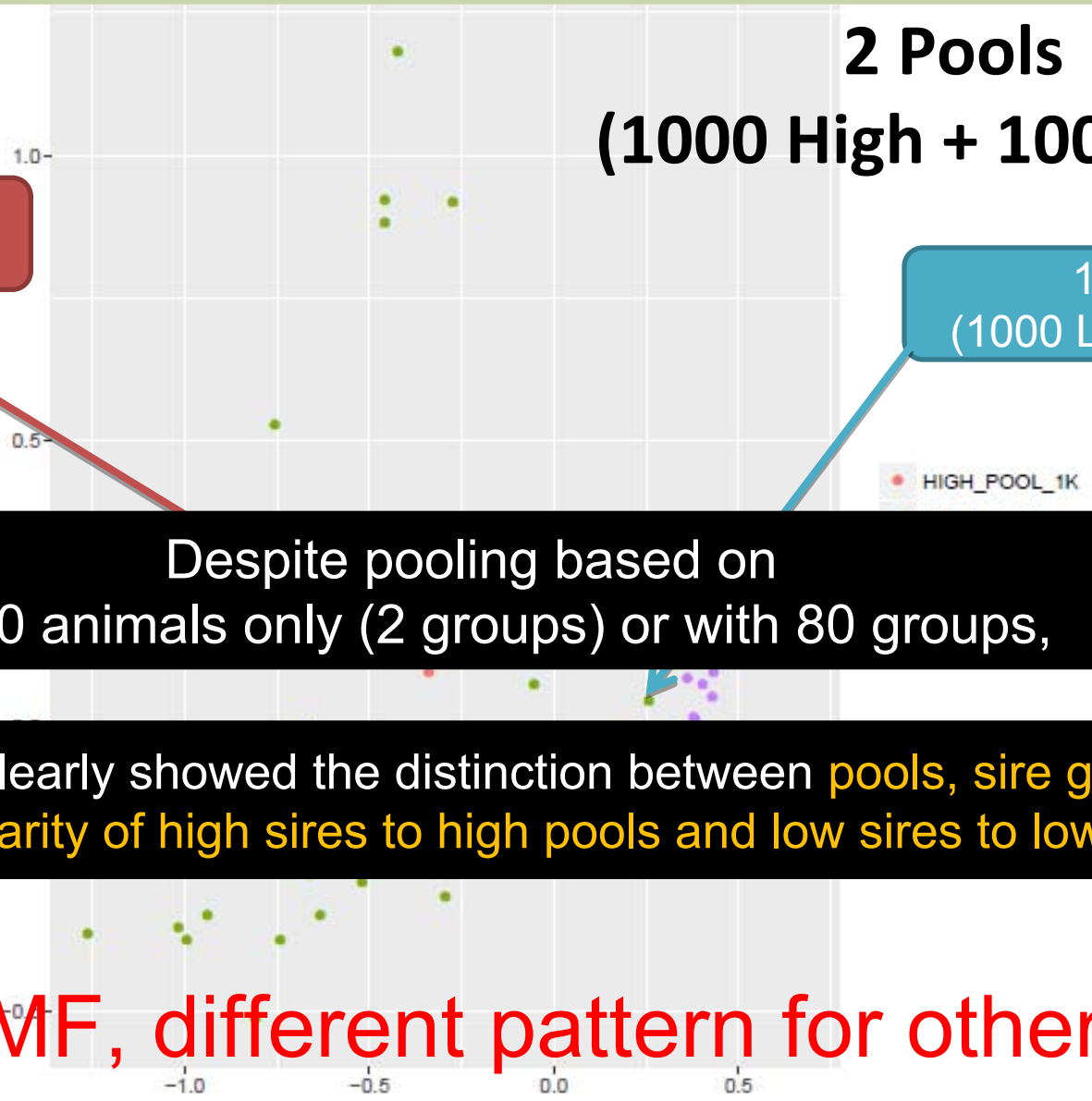
```
cor((dataResults$HIGH - dataResults$LOW) ,dataResults$EPD_Bull)
```

80 pools approach - GBLUP: Cor(DGVs,EPDs) = 0.895

2 Pools
(1000 High + 1000 Low)

1 Pool
(1000 Highest IMF)

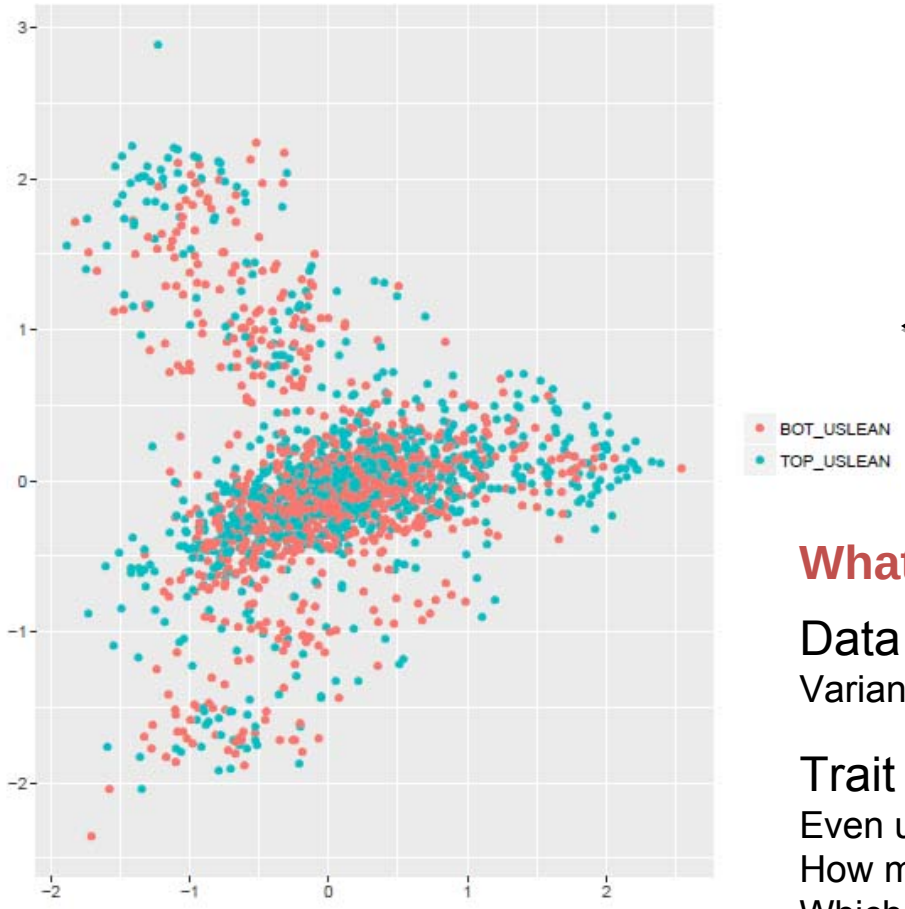
1 Pool
(1000 Lowest IMF)



Despite pooling based on
1,000 animals only (2 groups) or with 80 groups,

PCA plots clearly showed the distinction between pools, sire groups and the similarity of high sires to high pools and low sires to low pools

For PCTIMF, different pattern for other traits!



REA

RIBFAT

USLEAN

What are the factors causing this issue?

Data issue? (NO)

Variance components re-estimated at Angus and at Project level

Trait structure (80K vs SNPs associated)

Even under similar h2, cluster pattern could be different

How many markers affecting the trait, how many Chrs?

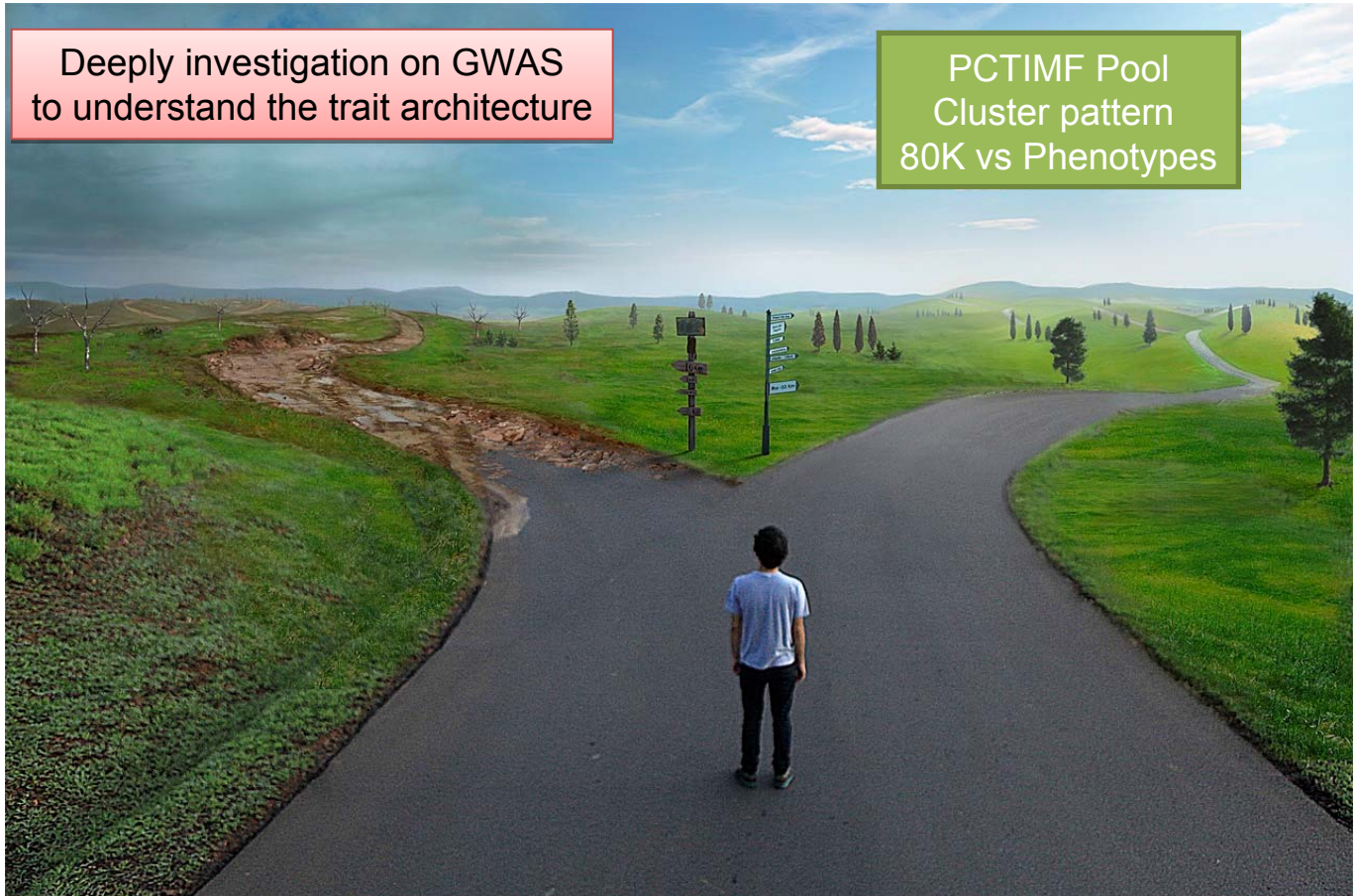
Which GWAS method to select markers?

What is the information available for each GWAS study?

How many animals? Extremes?

% Phenotypic variance explained by the markers?

Switching the project focus …



GWAS using different reference sizes (extremes or not for each trait)

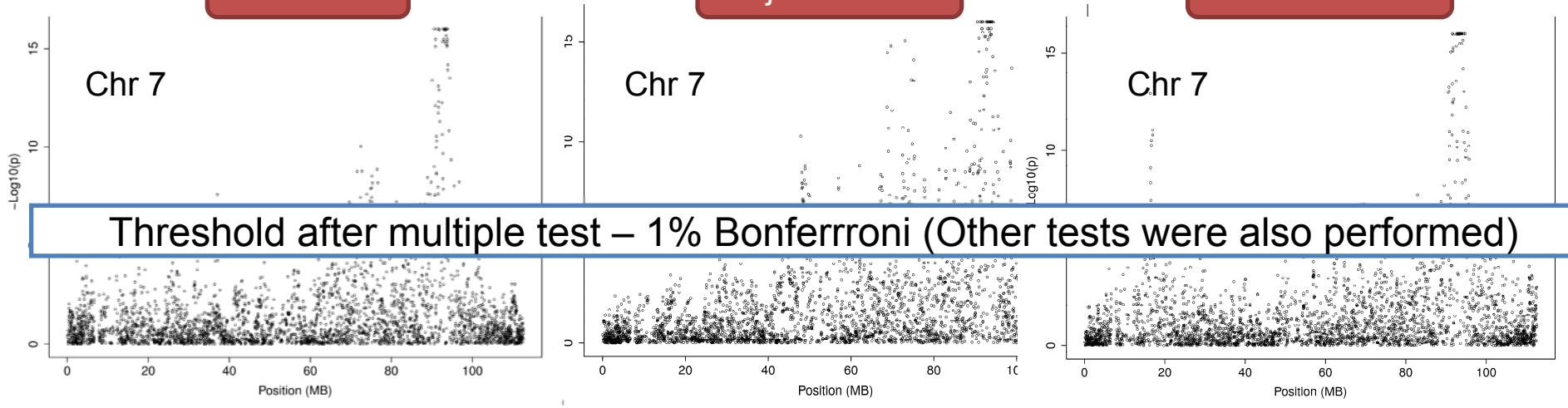
157,870 original phenotypes

Normal selection	Extremes selection
1,000 animals	1,000 animals
2,000 animals	2,000 animals
5,000 animals	5,000 animals
10,000 animals	10,000 animals

Raw REA

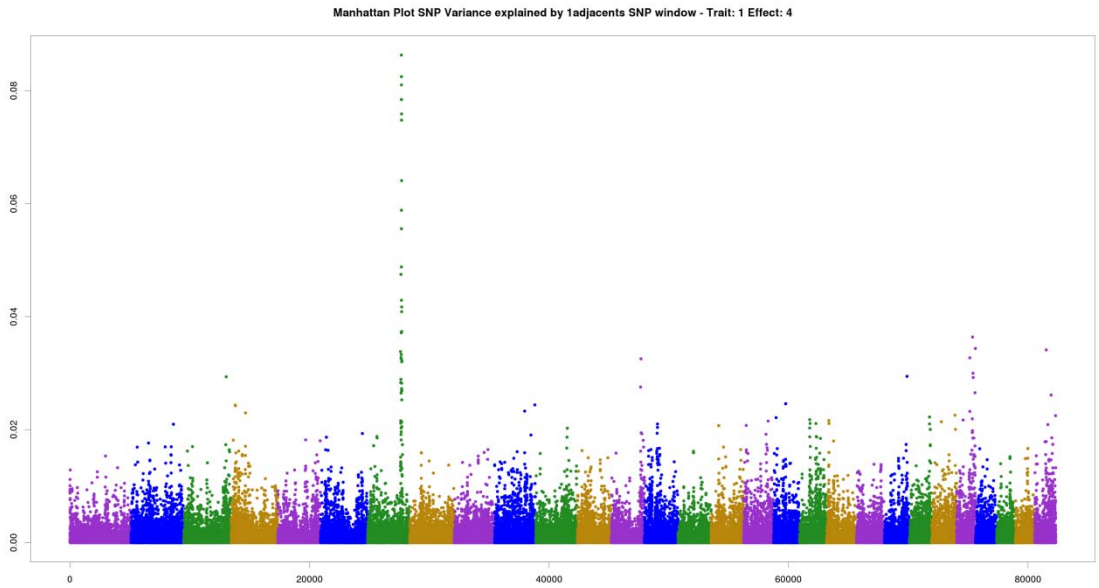
Adjusted REA

EBV REA



Threshold after multiple test – 1% Bonferrroni (Other tests were also performed)

Adjusted REA ssGWAS



Take home message

Pooling is an effective way to investigate new traits at a very low cost

Pooling based on all the medium/high density markers may require the identification of SNPs after GWAS

Trait architecture must be carefully investigated even with similar heritability before applying DNA pooling

.: Acknowledgments

Support since August 2017 – JP 16/19514-2