

# Genomic selection for new traits – optimal prediction and reference population design

Mario Calus

Yvette de Haas

Marcin Pszczola

Roel Veerkamp



---

# Background

---

- Closing the phenomic gap
    - Linking genotype and phenotype information
  
  - Methods, data collection and experiments to select for 'new' traits
  
  - From a breeders' perspective:
    - Use genotypic information for selection for new traits
- => Genomic selection for new traits

---

# Outline

---

- Why genomic selection (GS) for new traits?
- Predicted accuracy and selection response with small reference populations
- Ways to increase selection accuracy and response

=> Focus on dairy cattle, some principles also apply to other species

---

# Why GS for new traits?

---

- 'New' traits:
  - Difficult or expensive to measure
  - E.g. feed efficiency & methane emission in dairy cattle
  
- Measuring large daughter groups per bull inhibited
  
- Instead: GS based on small-scale reference populations

---

# When phenotyping is (very) expensive

---

- A fixed budget may yield the highest GS accuracy by:

- Phenotyping and genotyping the same animals

- 1,000s  instead of 10,000s 

- Fewer records; lower  $h^2$  individual records

- 'Optimal' reference population may still yield low accuracy

- Breeding value accuracy below usual standards
- 'Phenomic gap' bigger for new compared to conventional traits
- Is genetic progress possible?

---

# Predicted selection response based on cow reference populations (RP)

---

Deterministic simulations (using SelAction (Rutten et al., 2002))

- Breeding goal trait;  $h^2 = 0.3$ ; RP 6,250 bulls (acc. = 0.8)
- New trait introduced:
  - $h^2 = 0.3$  (or 0.05)
  - $r_g$  with breeding goal trait = -0.5, 0, 0.5
  - Economic value is 10,50,100% relative to breeding goal trait

*Animal*, page 1 of 9 © The Animal Consortium 2012  
doi:10.1017/S1751731112001450



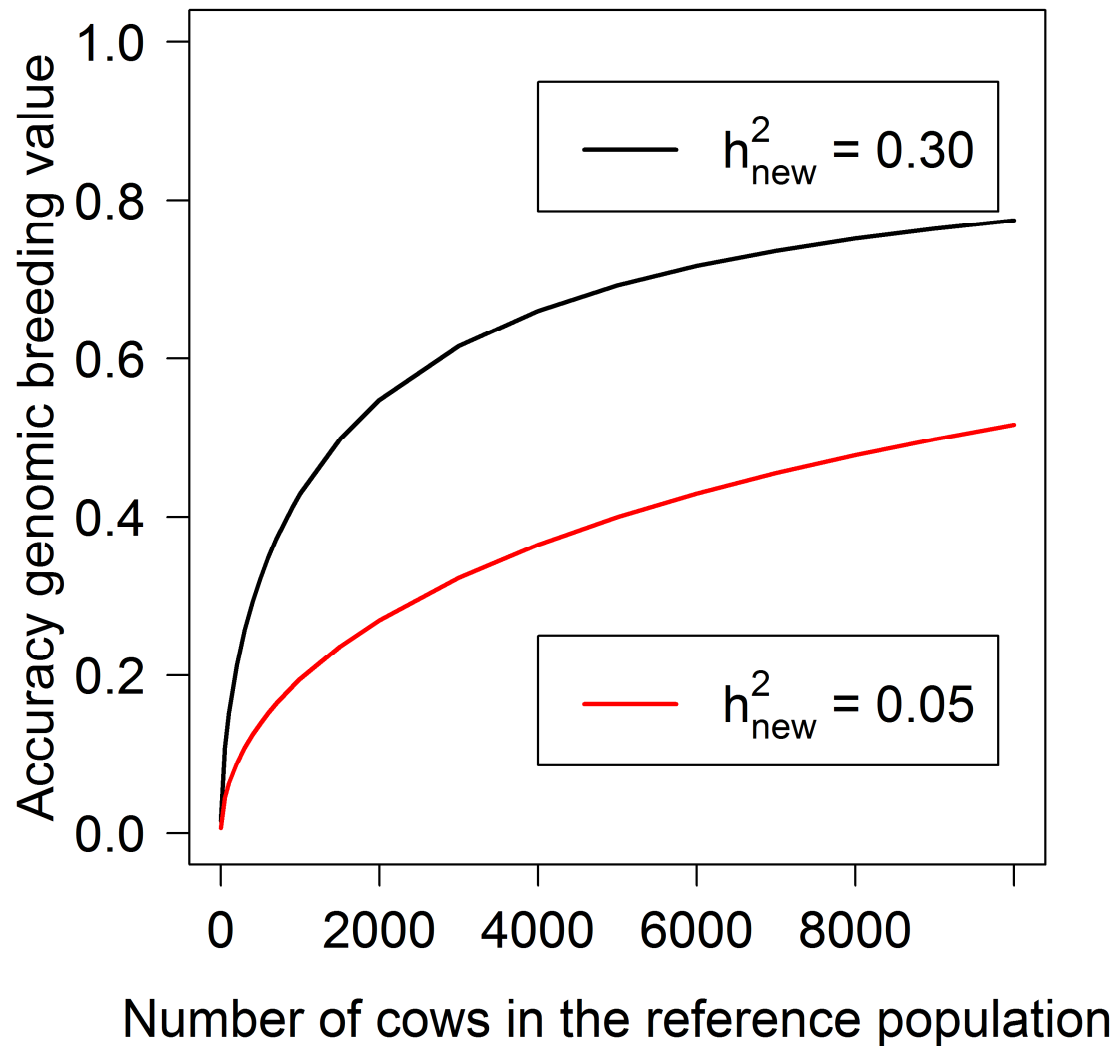
**Predicted accuracy of and response to genomic selection for new traits in dairy cattle**

M. P. L. Calus<sup>1†</sup>, Y. de Haas<sup>1</sup>, M. Pszczola<sup>1,2,3</sup> and R. F. Veerkamp<sup>1</sup>

Animal Breeding &  
Genomics Centre

# Accuracy genomic breeding values

## Daetwyler-formula

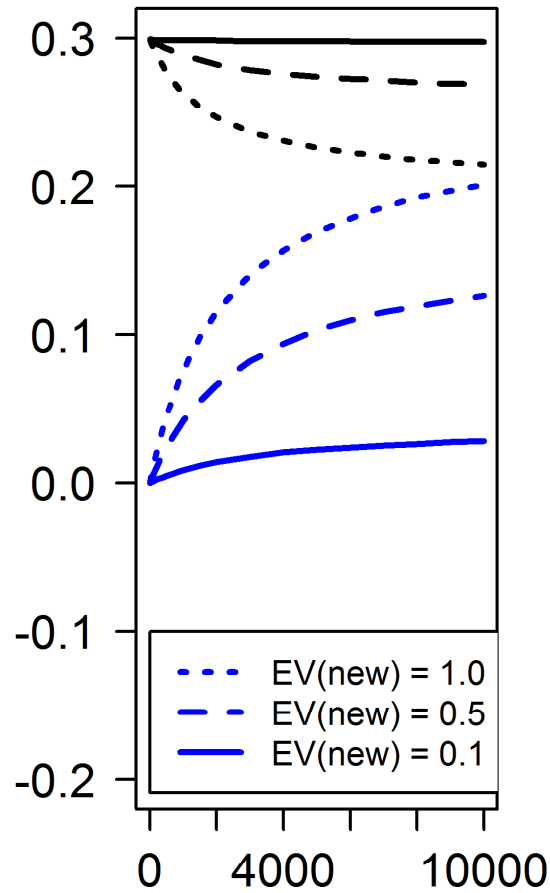


$h^2 = 0.30$

Black (upper) = breeding goal; blue (lower) = new trait

$r_g = 0.0$

Response (genetic SD per year)



Number of cows in the reference population

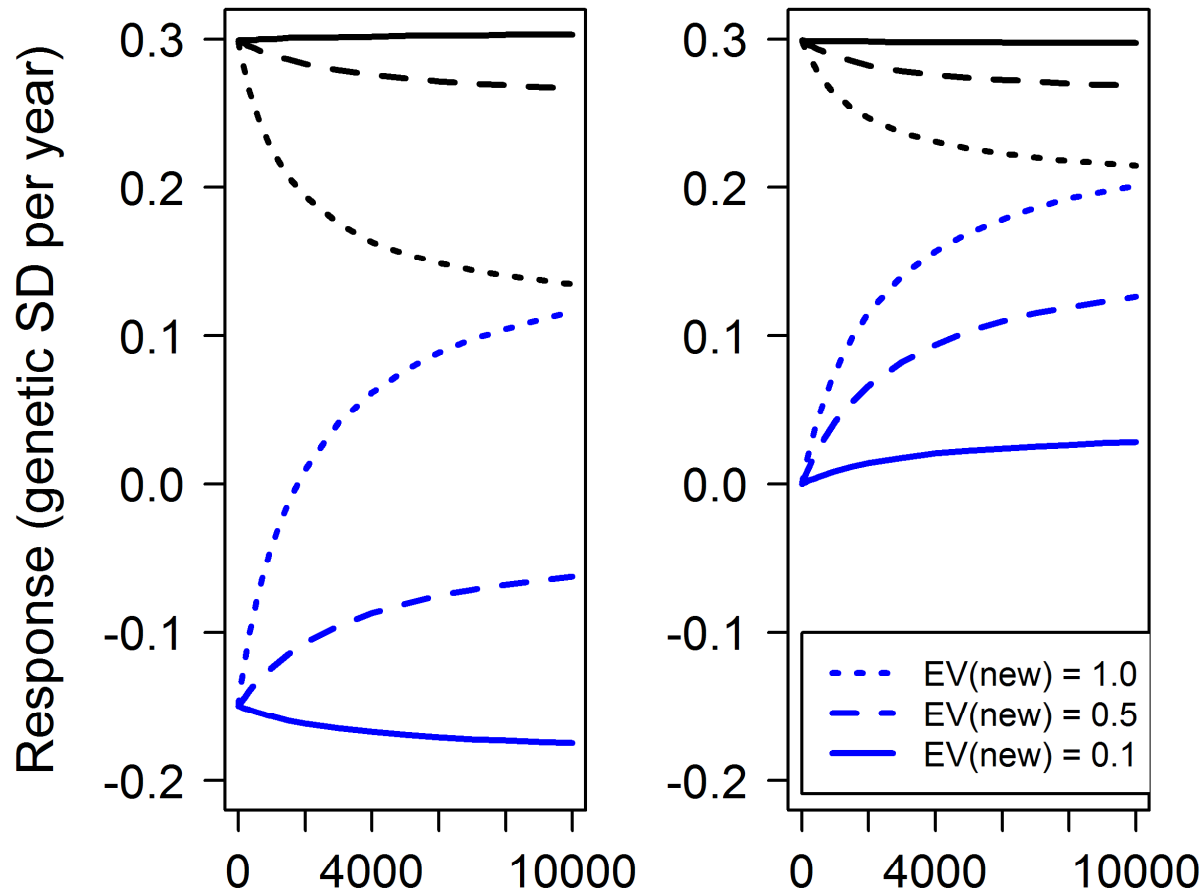


$h^2 = 0.30$

Black (upper) = breeding goal; blue (lower) = new trait

$r_g = -0.5$

$r_g = 0.0$



Number of cows in the reference population

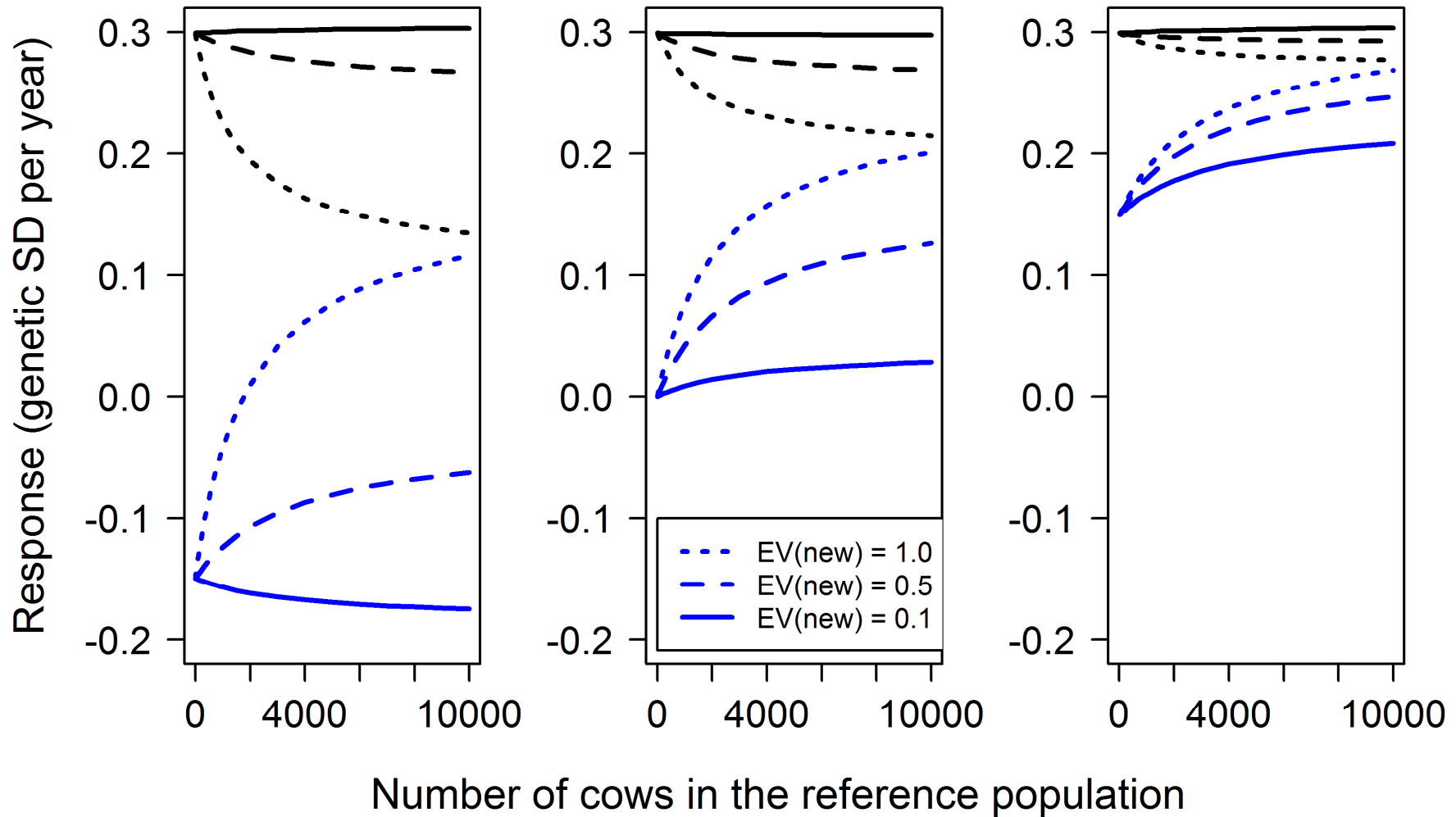
$h^2 = 0.30$

Black (upper) = breeding goal; blue (lower) = new trait

$r_g = -0.5$

$r_g = 0.0$

$r_g = 0.5$



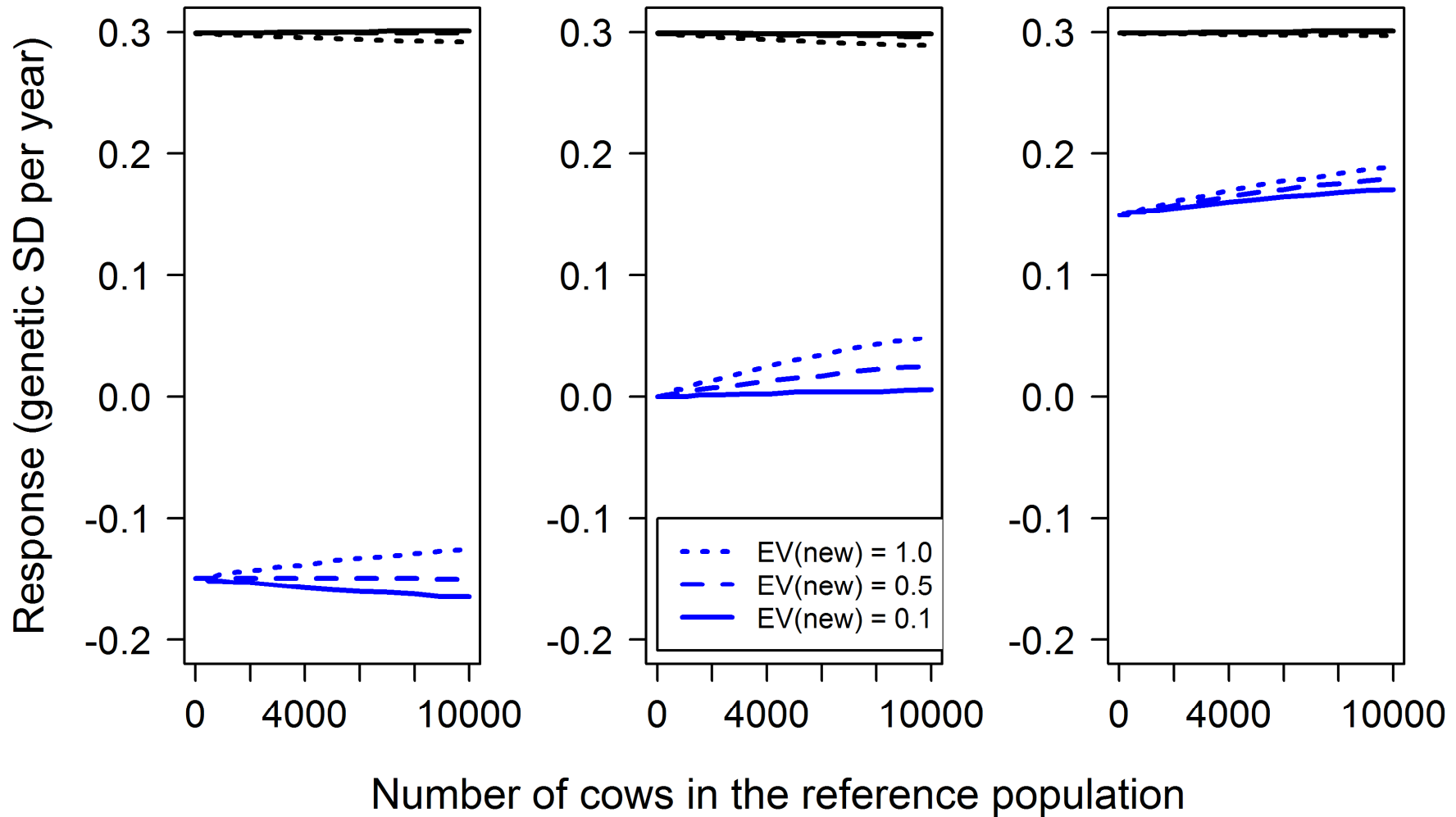
$h^2 = 0.05$

Black (upper) = breeding goal; blue (lower) = new trait

$r_g = -0.5$

$r_g = 0.0$

$r_g = 0.5$



---

# Results summarized

---

Required RP size for new trait depends on new trait:

- Heritability ( $h^2$ )
- Economic value relative to current breeding goal
- Genetic correlation with current breeding goal
  
- 2,000 cows yields considerable genetic response for a new trait with:
  - ⇒ small or positive genetic correlation with the index
  - ⇒ high positive economic value
- $\geq 10,000$  cows required with negative genetic correlation

---

---

How to further increase  
accuracy & selection response  
cost effectively?

---

# How to further increase accuracy & selection response cost effectively?

---

- Use of indicator traits
  - Similar to traditional selection based on pedigree
  
- Combining reference populations
  - Cow and bull RP
  - (international) cooperation
  
- Optimal reference population design

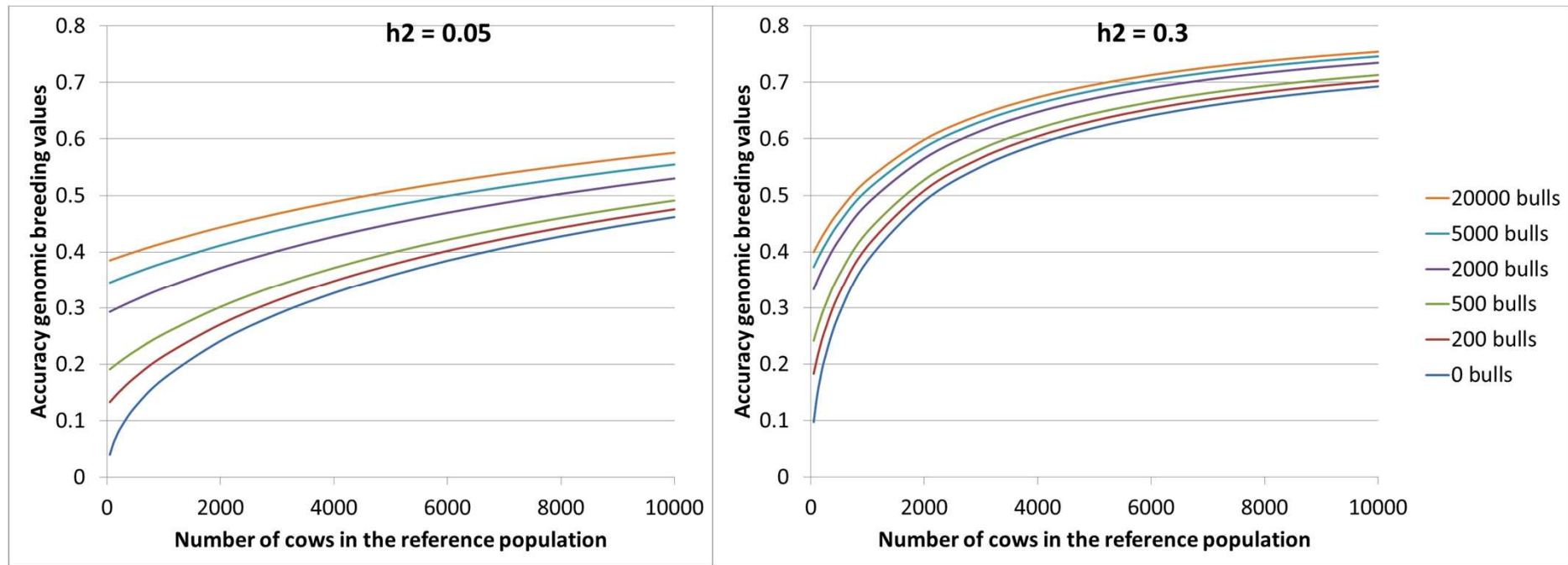
---

# Use of indicator traits - strategies

---

- Use available related traits on:
  - Reference animals
  - Other (i.e. bull) reference populations
  
- Develop new 'proxies'
  - Accept lower accuracy for low cost phenotypes
  
- Examples:
  - Measuring methane *De Haas et al., Session 25, Wednesday 8.30*
  - Use mid-infrared (MIR) profiles to predict e.g. energy balance *Poster McParland et al., Session 12*

# Combining cow and bull reference populations – $r_g(\text{cow};\text{bull}) = 0.5$



Bull RP on correlated trait increases accuracy especially:

- Low heritability traits
- Small cow reference populations



# Use of indicator traits – prediction model

- Requires multi-trait (MT) genomic prediction model:



- Conclusions:

- Developed Bayesian MT model can deal with indicator trait (i.e. missing records for evaluated trait)
- Genotyping more important than phenotypes for indicator trait, when  $r_g < 0.5$

[Predicting dry matter intake using live weight and milk yield](#)

*Poster Pszczola et al., Session 12*

# Combining cow and bull reference populations – real data I

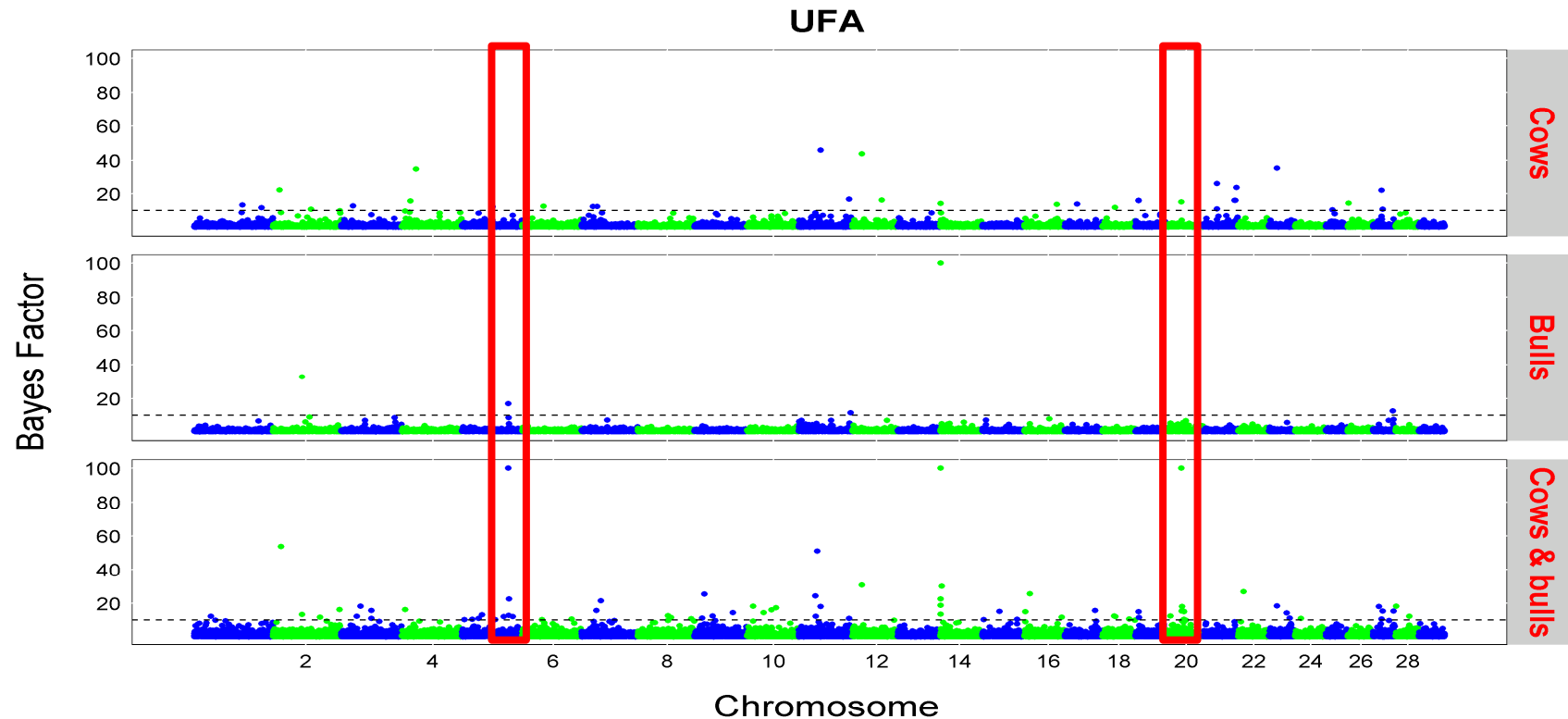
- Genomic prediction – fat & protein yield
- 1,609 cows with phenotypes; 296 bulls with DYDs

Reference population	Trait			
	Fat		Protein	
	<b>Accuracy</b>	SE	<b>Accuracy</b>	SE
Cows	0.328	0.068	0.189	0.050
Bulls & cows	0.347	0.063	0.206	0.055

Poster Calus et al., Session 12

# Combining cow and bull reference populations – real data II

- GWAS – (unsaturated) fatty acids (FA)
- 370 cows with predicted phenotype; 226 bulls with dEBV



Results for more FA groups; see Poster Bastin et al., Session 12

# Pooling reference populations (internationally)

## ■ RobustMilk project

- Pooling of research herd data across countries
- See: Veerkamp et al., Session 50, Thursday 15.15



## ■ gDMI collaboration

- Genomic breeding values for dry matter intake (DMI)
- Global initiative using data across >10 countries

---

# Issues encountered when combining data

---

- Connections between populations
  - Problems due to poor connectedness may be alleviated by using SNP instead of pedigree data<sup>1</sup>
  
- Traits may be different due to:
  - GxE interaction
  - Differences in trait definitions / measuring procedures
  - Use of indicator traits

=> Flexible use of combined data requires multi-trait approach<sup>2</sup>

---

# Reference population design

---

## Selective genotyping:

- RP should reflect (i.e. be related to) the selection candidates

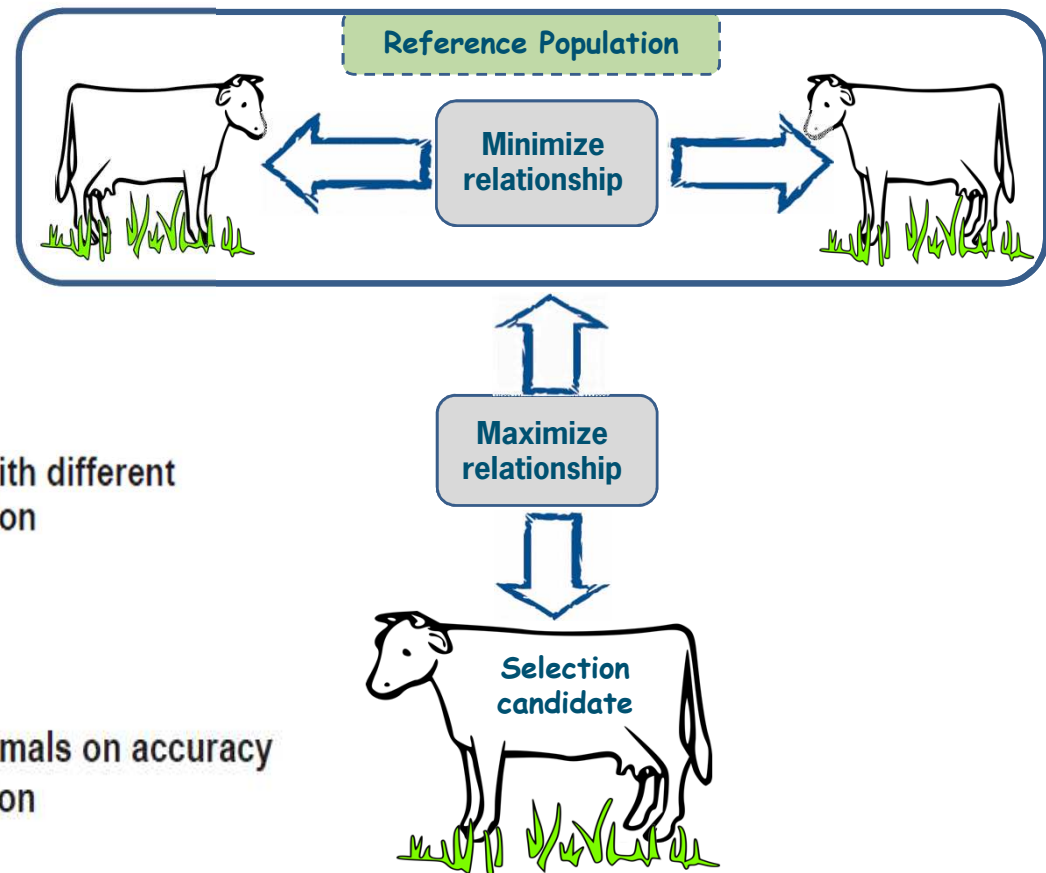
## Selective phenotyping:

- Sampling extreme phenotypes
  - From both tails – Jimenez-Montero et al., *Animal*, 2012

# Reference population design *(Pszczola et al. 2012)*

If it cannot be big, let it be well designed

## Optimal design of the reference population



J. Dairy Sci. 95:389–400  
doi:10.3168/jds.2011-4338  
© American Dairy Science Association®, 2012.

Reliability of direct genomic values for animals with different relationships within and to the reference population



J. Dairy Sci. 95:5412–5421  
<http://dx.doi.org/10.3168/jds.2012-5550>  
© American Dairy Science Association®, 2012.

The impact of genotyping different groups of animals on accuracy when moving from traditional to genomic selection

---

# GS for new traits – Many open questions

---

- Are (moderate) relationships to RP still required for GS in future applications?
  - Future developments: (imputed) sequence data
  
- How can we directly strengthen the phenotype-genotype associations when RP are of limited size?
  - Closing the phenomic gap for new traits
  
- ???



---

# Conclusions

---

GS for new traits:

- Small reference populations can yield considerable genetic response, depending on
    - $h^2$  & economic value
    - $r_g$  with current breeding goal
  
  - Accuracy and selection response can be further increased by:
    - Reference population design
    - Use of indicator traits
    - Combining RP (bull-cow; international)
- } Requires Multi-trait approach

---

# Acknowledgements

---



Animal Breeding &  
Genomics Centre