Exclusion or inclusion of information of culled bulls on the single-step genomic evaluation

Minna Koivula, Ismo Strandén, Gert Pedersen Aamand and Esa A. Mäntysaari

Minna.Koivula@luke.fi

Luke
NATURAL RESOURCES
INSTITUTE FINLAND

Introduction

- Number of genotyped animals has increased rapidly creating computational challenges for genomic evaluation
- In animal model BLUP, candidate animals without progeny and phenotype do not contribute information to the evaluation and can be discarded
- In theory, genotyped candidate animal without progeny can bring information into single-step BLUP (ssGBLUP), and affect estimation of other breeding values

Aims

- To study the effect of inclusion or exclusion of genomic information of culled bulls on
  - the GEBVs of genotyped bulls with own information
  - GEBVs of young selected bulls.

- ssGBLUP was computed using Nordic test-day (TD) model and TD data for the Nordic Red Dairy Cattle (RDC).

Materials and methods

- Nordic NAV RDC test-day (TD) data from February 2015
  - 3.9 million cows with a total of 86.5 million records
- Multiple trait multi-lactation models:
  - Production evaluation
  - Milk, protein and fat 305d yield (G)EBVs for all animals

Genotype data

- Marker data from February 2015
  - Genotyped RDC animals with 46914 SNPs
    - I. 20276 genotyped animals with records or information in the TD data (reference animals)
      - 14580 cows and 5696 bulls
      - Called as “Normal”
    - II. + 1140 young genotyped bulls bought for service (candidate bulls)
      - Called as “Medium”
    - III. + 8770 **culled** young genotyped bulls
      - Called as “Big”

Pedigree data

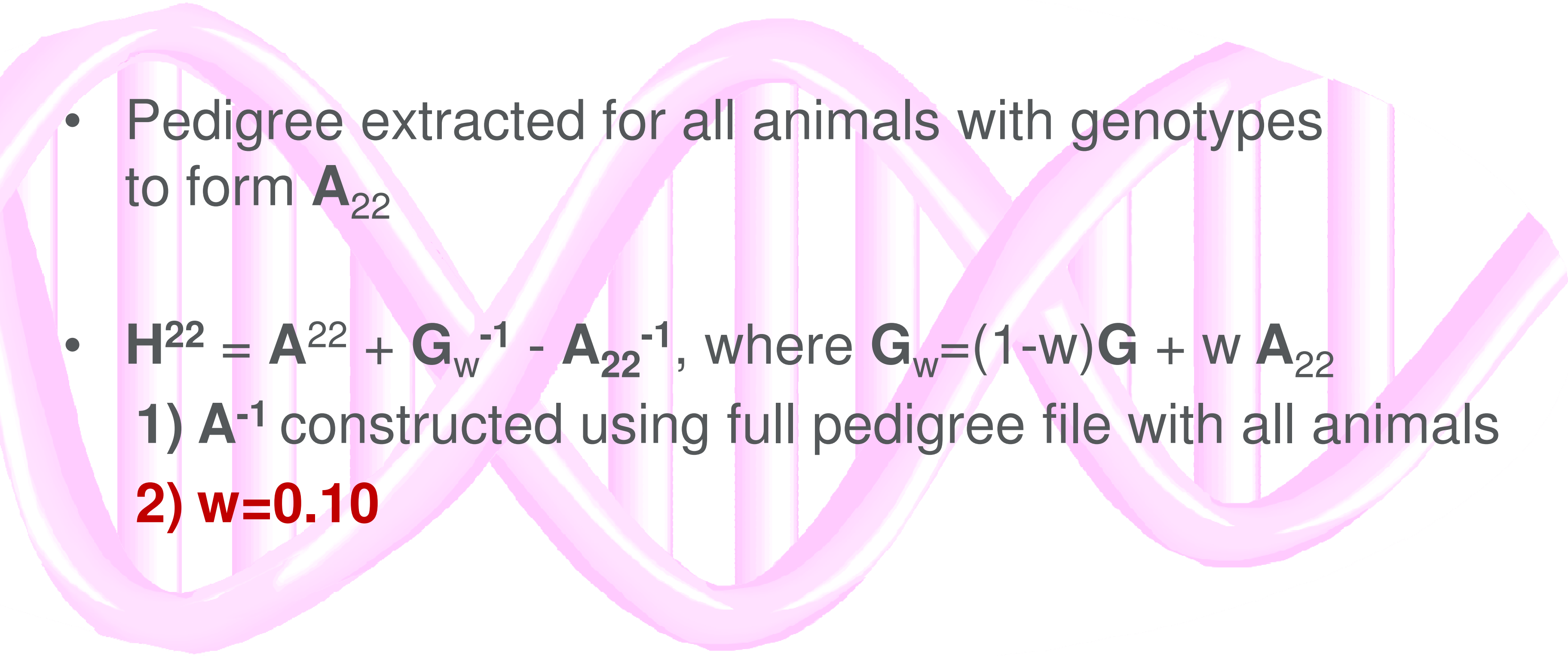
- Pedigree information in the TD model and TD ssGBLUP
    I. Nordic RDC pedigree from February evaluation

    II. Nordic RDC pedigree
    \+ pedigree information for 1140 genotyped young candidate RDC bulls that had been bought for service but without information

    III. Nordic RDC pedigree
    \+ pedigree information for 1140 selected candidate & 8770 culled genotyped RDC bulls

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

H-matrix



- Pedigree extracted for all animals with genotypes to form $\mathbf{A}_{22}$
- $\mathbf{H}^{22} = \mathbf{A}^{22} + \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1}$, where $\mathbf{G}_w = (1-w)\mathbf{G} + w\,\mathbf{A}_{22}$
  1) $\mathbf{A}^{-1}$ constructed using full pedigree file with all animals
  2) w=0.10

Different TD runs

- Three TD model runs for EBV:
  each using different pedigree information
- Three TD ssGBLUP runs: each using different pedigree
  + genomic information

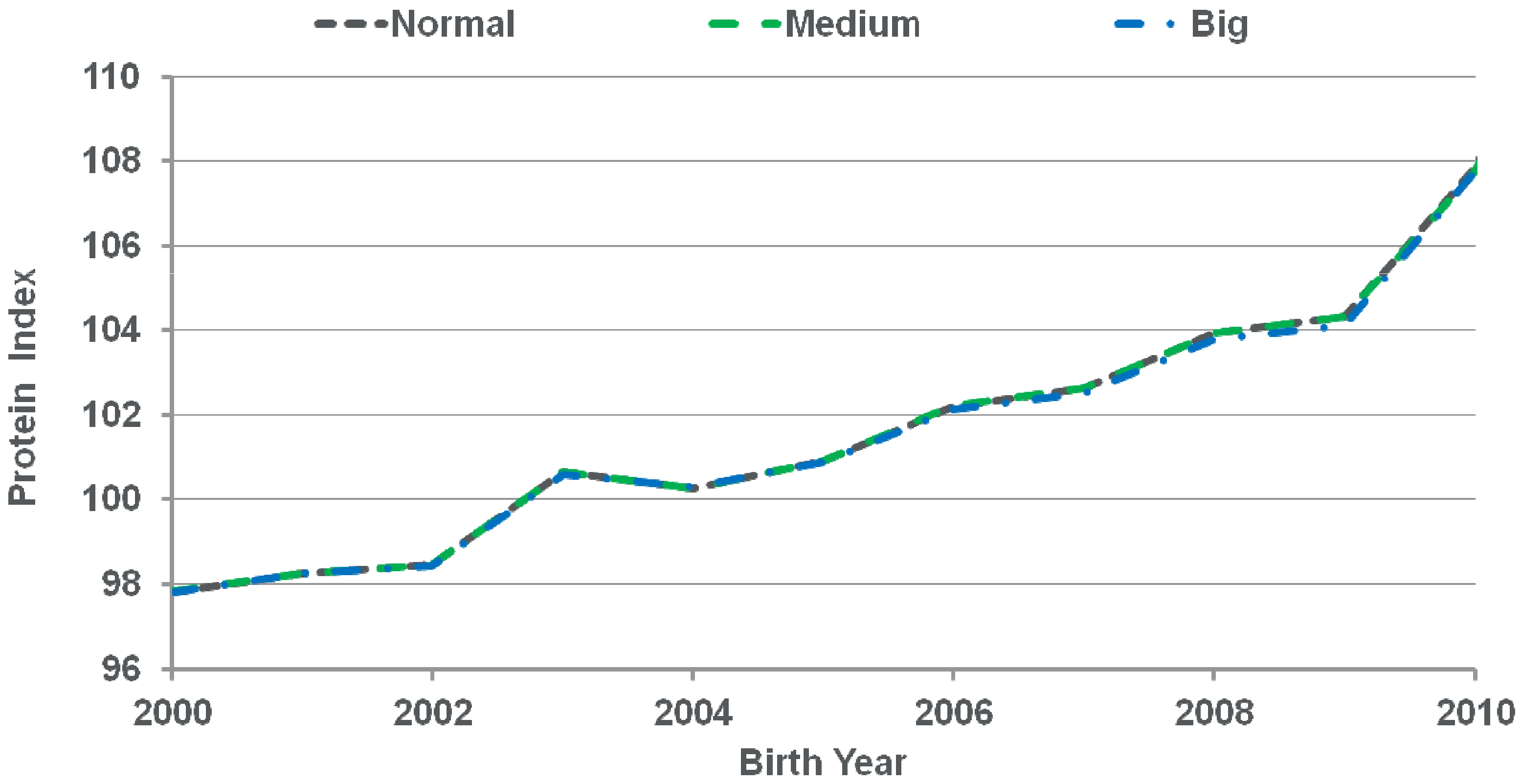
TD run	**TD model**	**TD ssGBLUP**
	Number of animals in the Pedigree	Number of Genotyped animals in the $\mathbf{H}^{22}$
Normal	5172229	20276
Medium	5173381	21416
Big	5182461	30186

Correlations among protein EBVs and GEBVs for genotyped bulls common in all TD runs (5696 bulls)

	Correlations among EBVs			**Correlations among GEBVs**		
	Normal	Medium	Big	Normal	Medium	Big
Normal		1.00	1.00		1.00	1.00
Medium			1.00			1.00

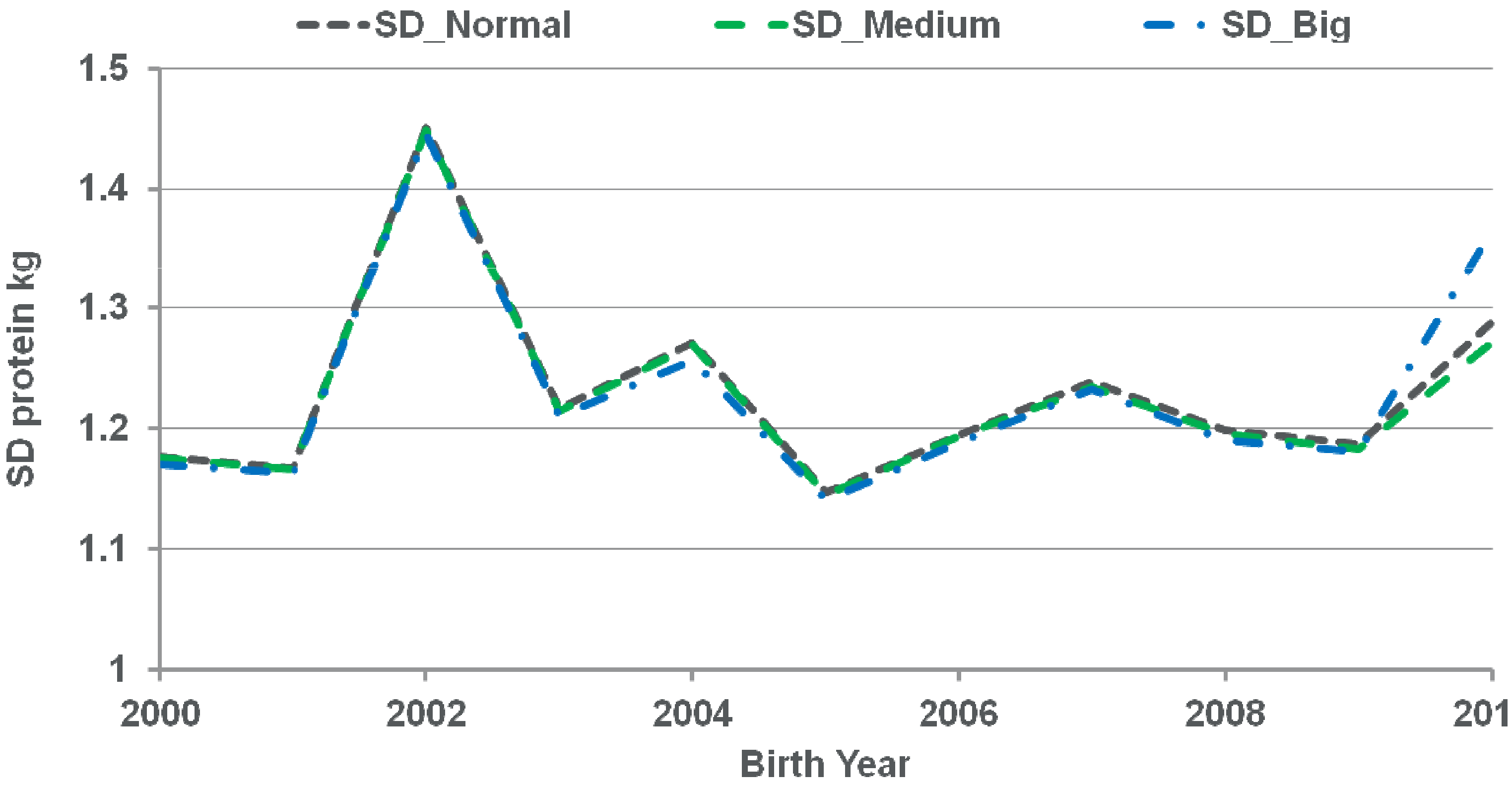
Neither EBVs, nor GEBVs do not differ if the ”extra” genotyped bulls are included into the analyses → culled bulls and young bulls without own information can be left out from the TD ssGBLUP

Trends for protein GEBVs from different TD ssGBLUP runs. Genotyped bulls common in all TD runs (5696 bulls)

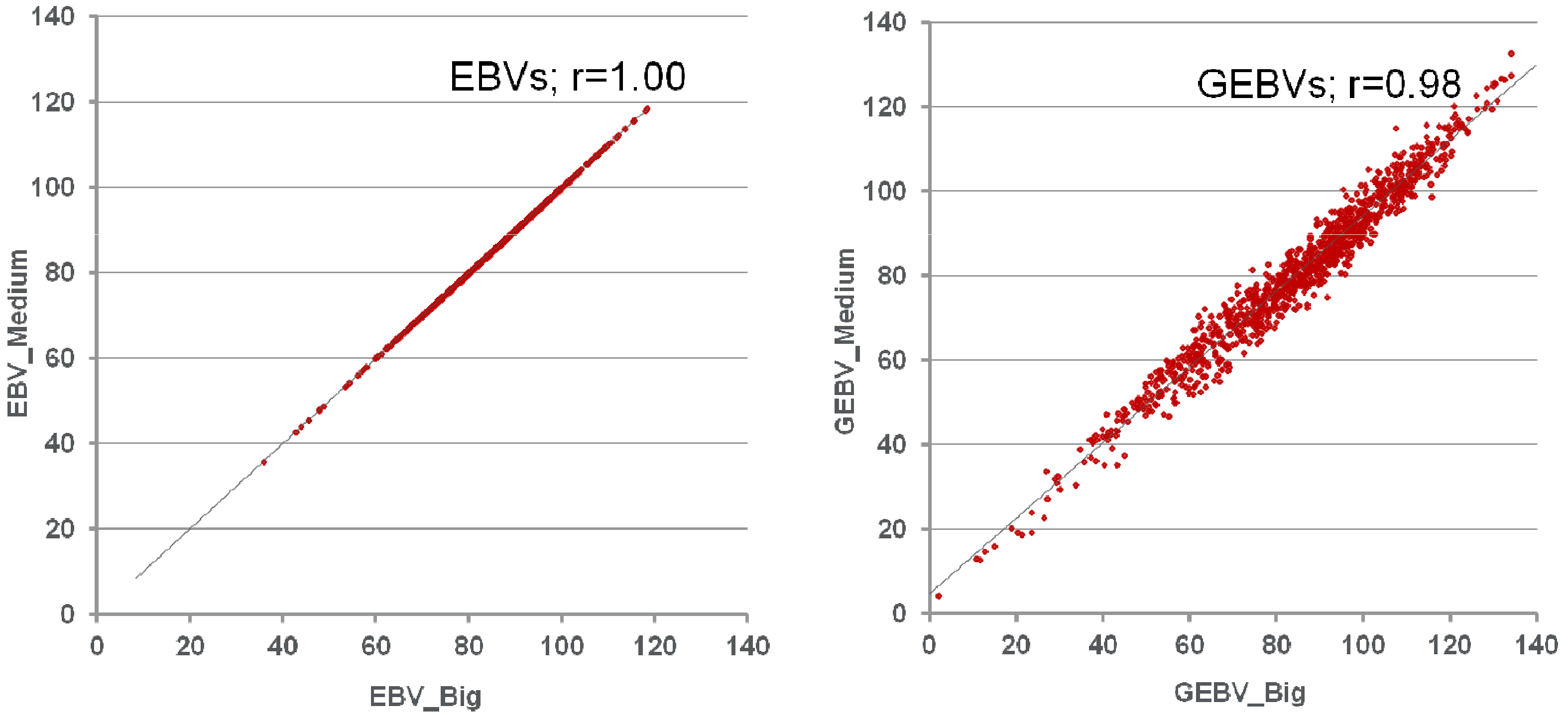


GEBVs are expressed as standardized breeding values with SD of 10 units for bulls born 2005 - 2007.

SDs for protein (kg) GEBVs from different TD ssGBLUP runs. Genotyped bulls common in all TD runs (5696 bulls)

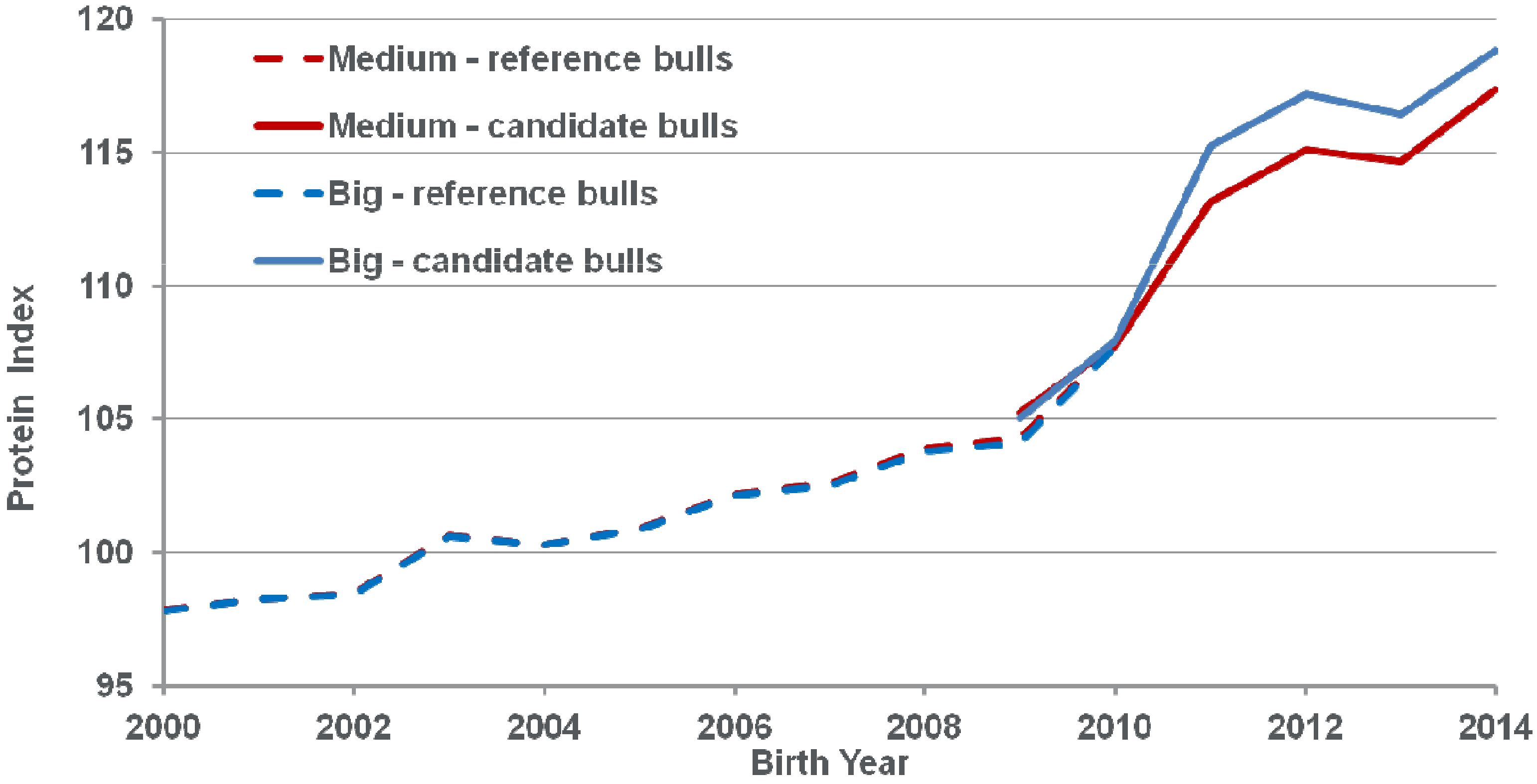


Scatter plot of protein (kg) EBVs and GEBVs for genotyped young bulls bought for service (1140 young candidate bulls)



Trends for protein GEBVs from **Medium and Big** TDssGBLUP runs.

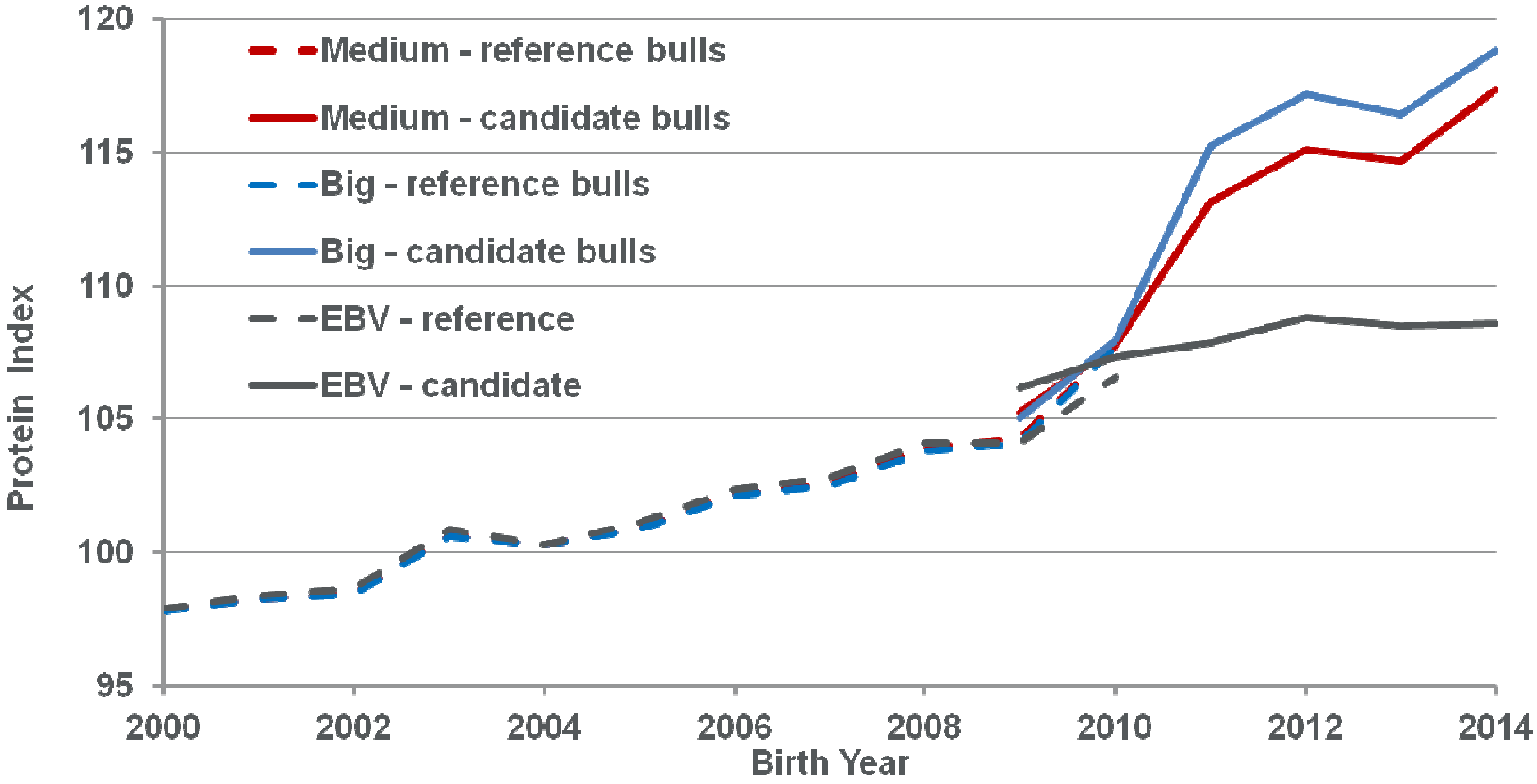
Trends separately for genotyped reference bulls common in all TD runs (5696 bulls) and for young candidate bulls (1140)



GEBVs are expressed as standardized breeding values with SD of 10 units for bulls born 2005 - 2007.

Trends for protein GEBVs from **Medium and Big** TDssGBLUP runs.

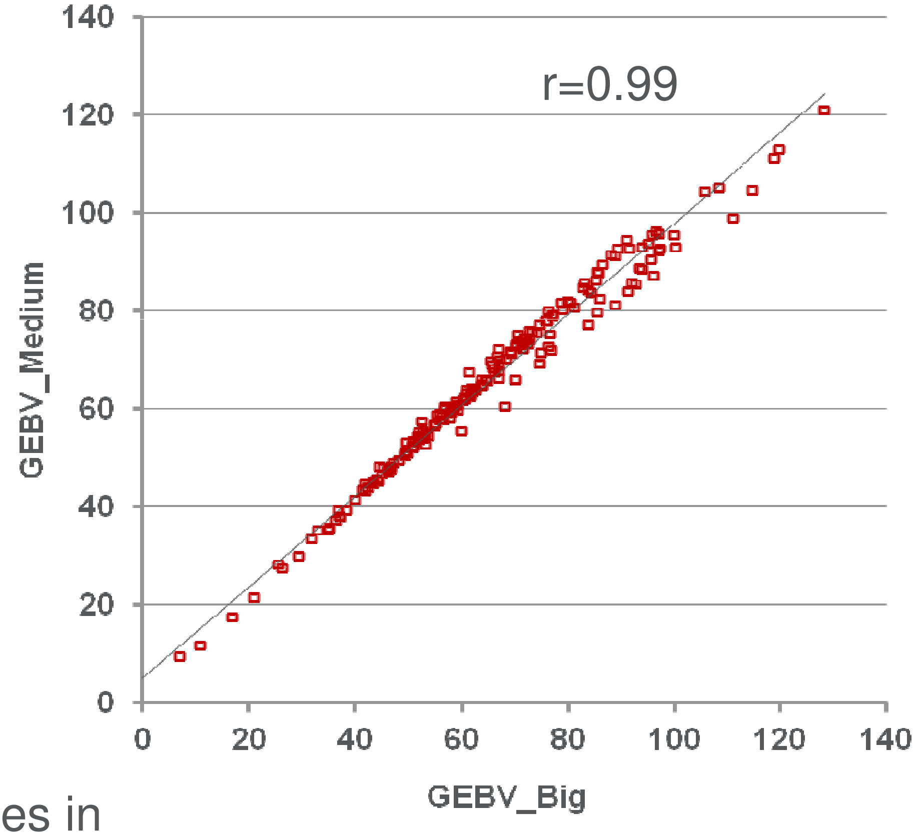
Trends separately for genotyped reference bulls common in all TD runs (5696 bulls) and for young candidate bulls (1140)



Luke
NATURAL RESOURCES
INSTITUTE FINLAND

Protein EBVs and GEBVs **for 187 sires** of young genotyped bulls bought for service

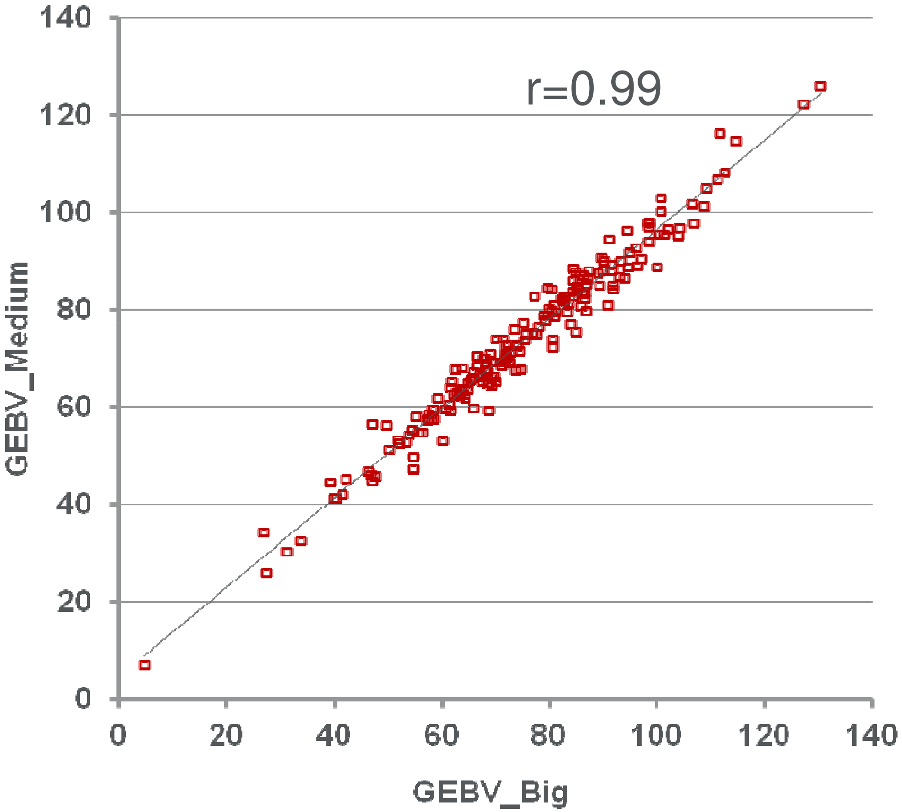
	EBV Big	**GEBV Big**
EBV Medium	**1.00**	
GEBV Medium		**0.99**



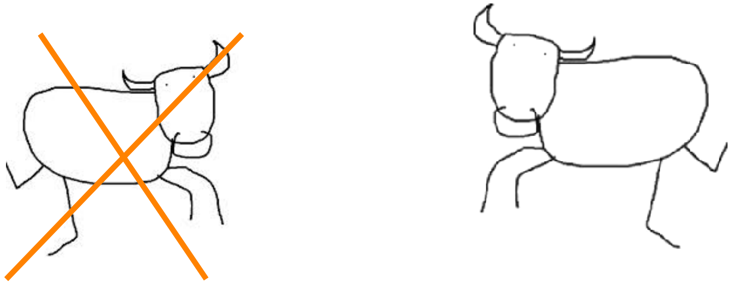
150 sires common in all TD runs, 187 sires in candidate set (there are 37 sires that have no information in the reference data, their correlations were ~0.98)

Protein EBVs and GEBVs **for 1116 bull dams** of young genotyped bulls bought for service

	EBV Big	**GEBV Big**
EBV Medium	**1.00**	
GEBV Medium		**0.99**



Discussion



- It seems that culled bulls can be safely excluded from the TD ssGBLUP without negative effect on the genomic evaluation of bulls with information (reference bulls) or young candidate bulls.
  - GEBVs (and EBVs) of reference bulls were the same whether genomic information of candidate or culled bulls were used in the TD model
  - Genetic trends did not change if information of culled bulls were used.

However,

- Information of culled bulls seem to affect a little to the genomic evaluation of young candidate bulls
  - If culled half sibs are taken into the TD ssGBLUP - genetic level of the candidate bull is overestimated
- Bull sire or bull dam GEBVs do not change notably with full genotypic data – but
  - There is a tendency that especially young sires and dams of the candidate bulls get higher GEBVs if also information of high number of culled sons is used
- It seems that to avoid overestimation of GEBVs of candidate bulls it could be better to discard genomic information of culled bulls

Acknowledgements

Phenotypic and Genomic data

NAV Nordic Cattle Genetic Evaluation,

Support:

Ministry of Agriculture and Forestry Finland,
NAV, Viking Genetics, FABA Coop and Valio Ltd

Thank you!

NATURAL RESOURCES
INSTITUTE FINLAND

NATURAL RESOURCES
INSTITUTE FINLAND