AARHUS
UNIVERSITY

April 2010

Genomic selection – importance of size of reference population and SNP density

Aarhus University:
Mogens Sandø Lund

SLU:
Freddy Fikse

Behind the scenes

AARHUS UNIVERSITET

- Mogens Lund
- Bernt Guldbrandtsen
- Rasmus Froberg Brøndum
- Guosheng Su

SLU

- Freddy Fikse
- Elisenda Rius Vilarrasa

MTT

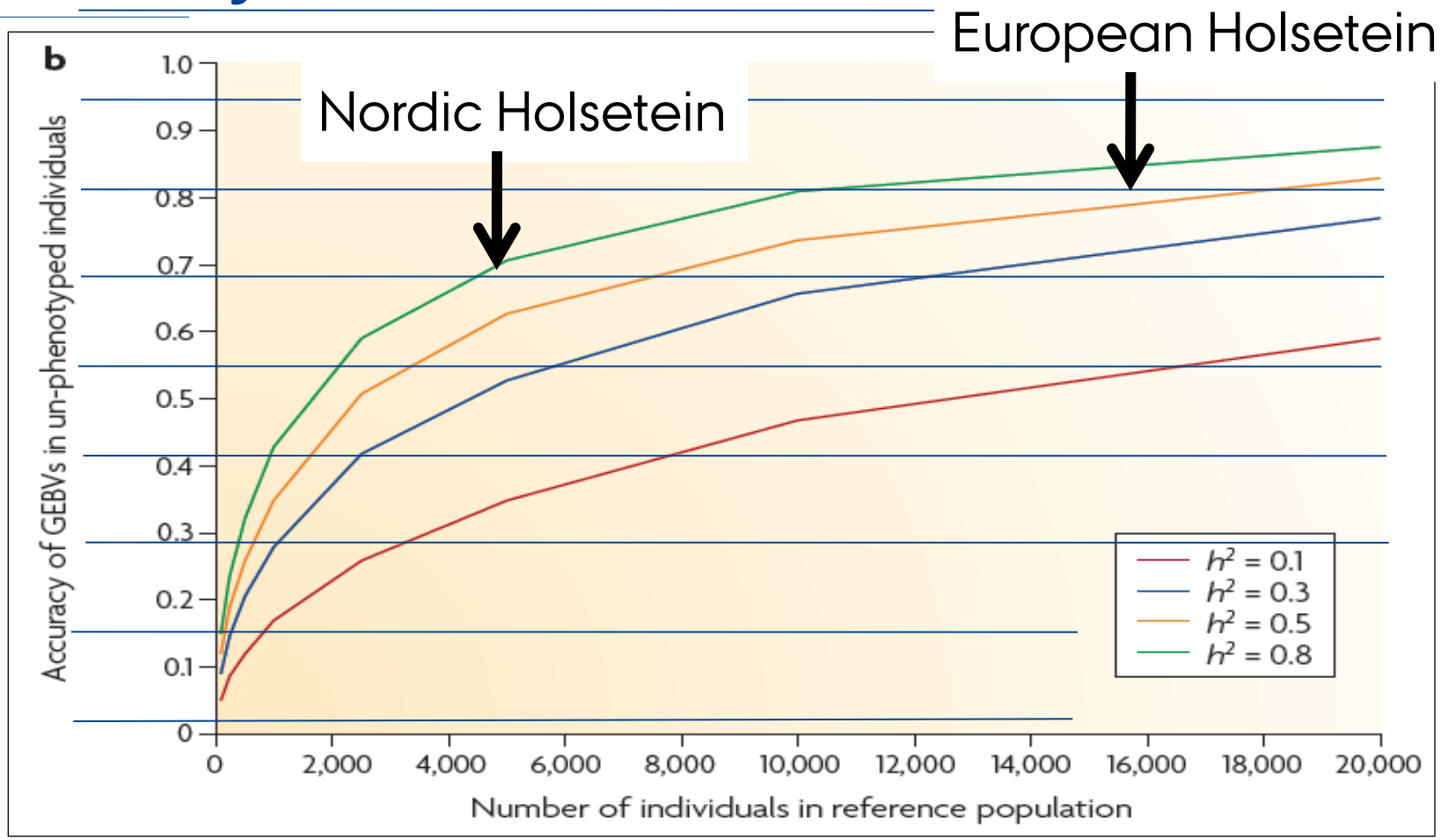
- Ismo Strandén
- Nina Schulman

Genomic selection

- Using dense SNPset på predict breeding values
  - Markers/haplotypes in close linkage disequilibrium with QTL
  - Many ”phenotypes” for each marker/haplotype to estimate effects accurately

- Critical factors
  - SNP density
  - Size of reference population (and effective population size)
  - Interplay between these factors in our breeds?

Accuracy increase with reference size



Goddard & Hayes, 2008

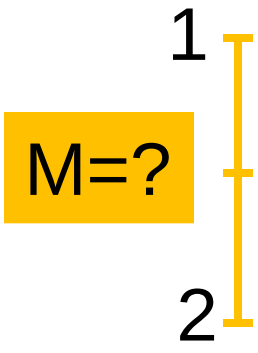
Improving genomic prediction by EuroGenomics collaboration

- Exchange genotypes between VikingGenetics, CRV, UNCEIA, DHV-VIT
- Increase reference from 4500 to 16.000
- Three different SNPsets used
- Initial imputation step to predict unobserved SNPs for CRV bulls

Imputation using LD and LA information

USDA chip

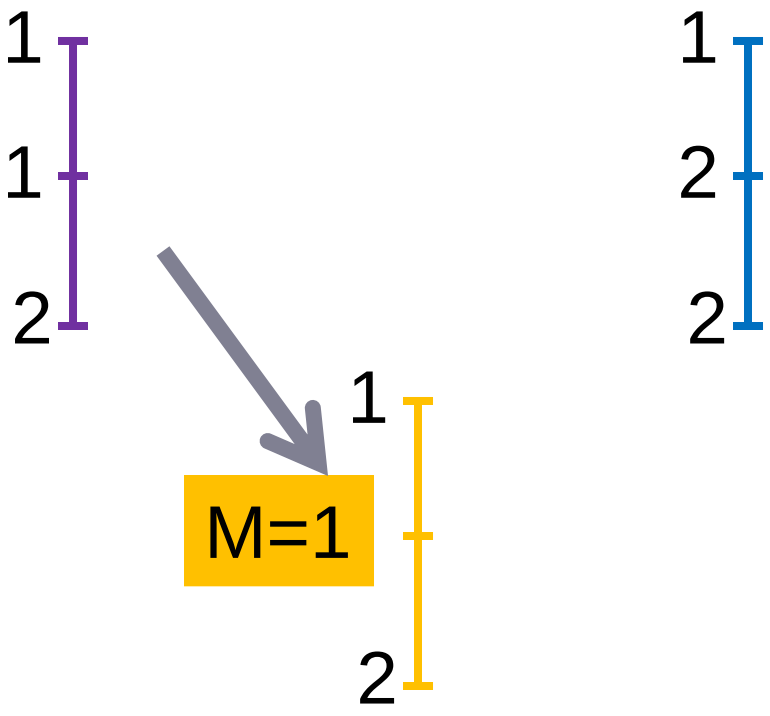
CRV chip



Imputation using LD and LA information

USDA chip

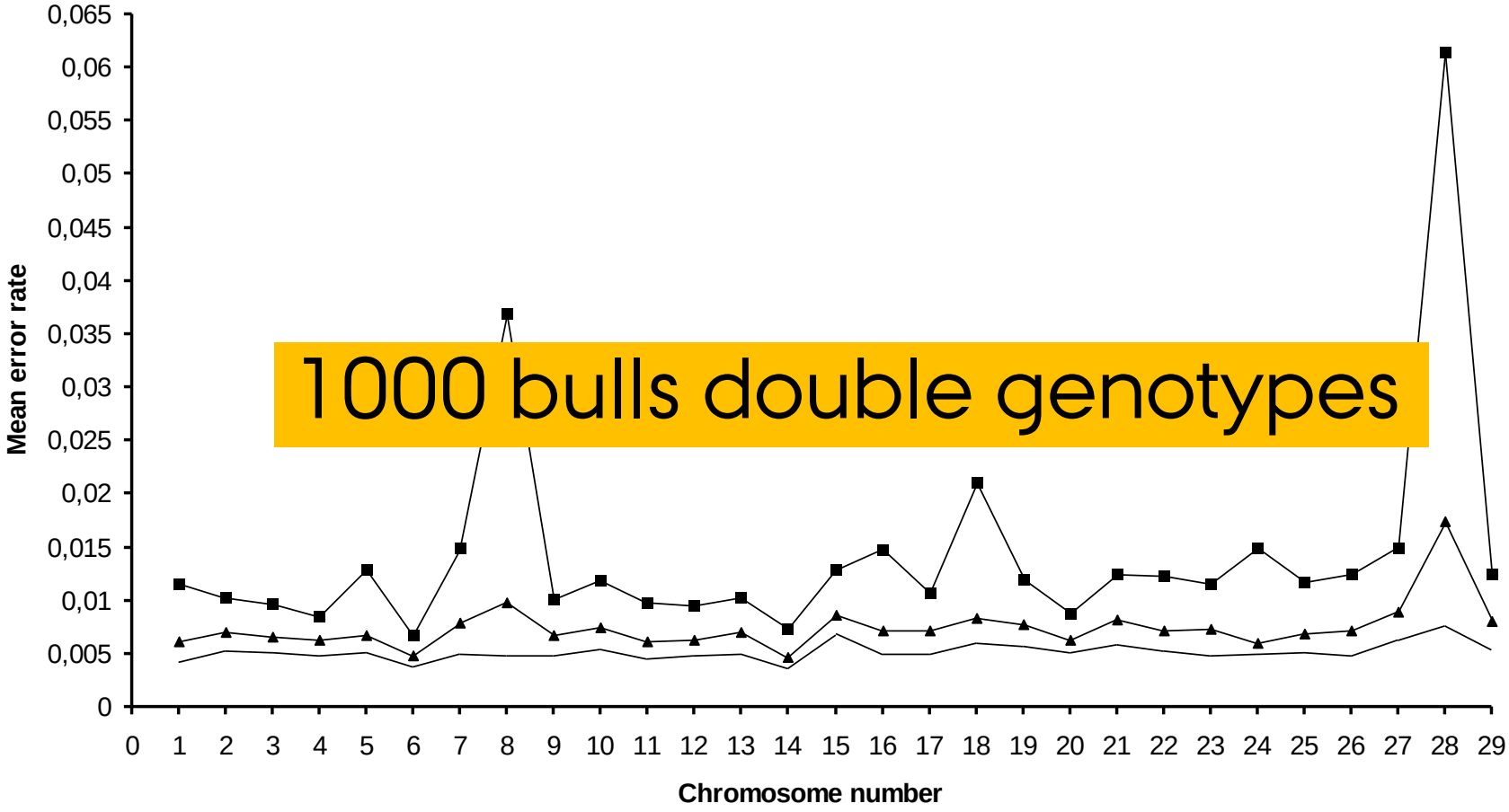
CRV chip



Imputation of USDA chip from CRV (Tom Druet)

› To study imputation (prediction of missing genotypes) across CRV 60K SNP panel and Illumina Bovine SNP50™ SNP chip («USDA» chip)

Double genotyping 500, 1000, or 2000 bulls



Improving genomic prediction by EuroGenomics collaboration

- Objective: Increase reliabilities of genomic breeding values by combining reference populations

- Hypothesis: Reliability increase 10-15% by using combined european reference compared to a purely Nordic reference

Joint EuroGenomics dataset

- 15,966 EuroGenomics bulls
- 19.4 million daughters
- 939 bulls with daughters in multiple countries
- Median number of daughters were:
  - DHV-VIT(117), UNCEIA (85), VikingGenetics (117), CRV (153)

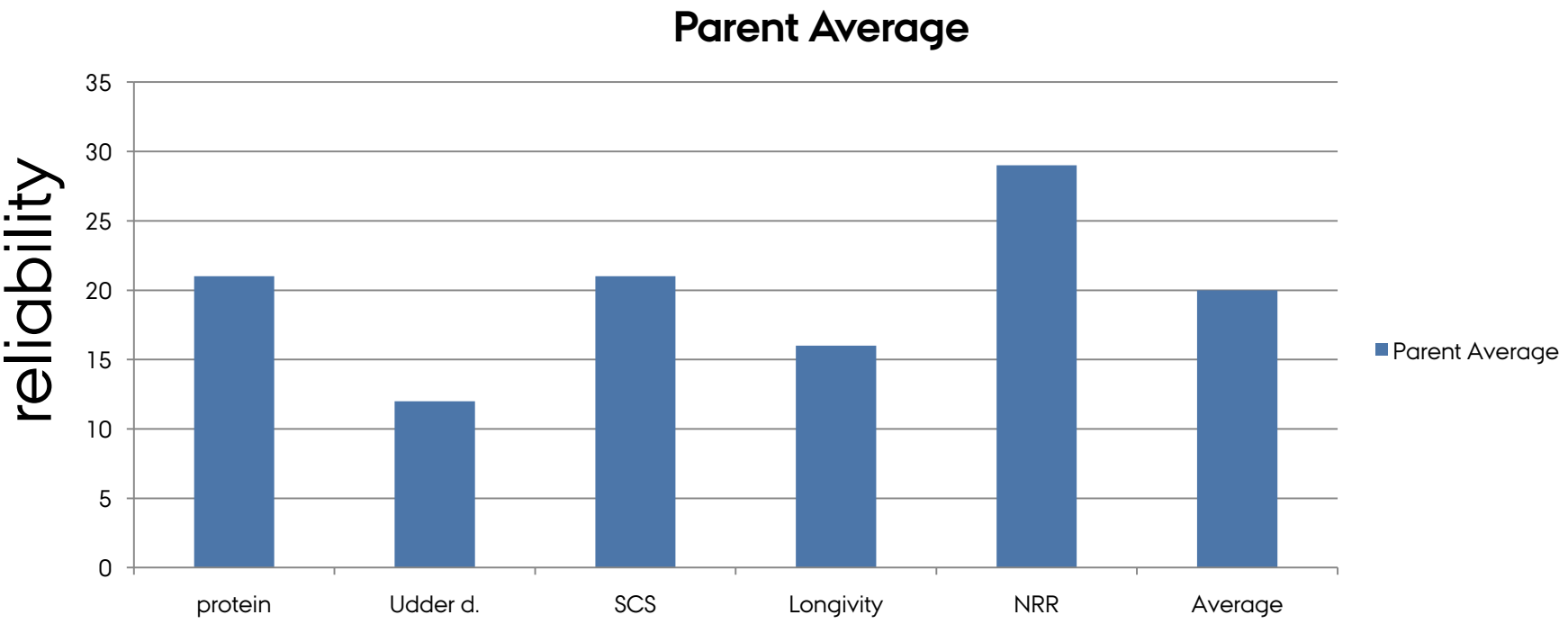
Reference and test population

- Nordic reference population
  - 75% of domestic population (oldest bulls)
- EuroGenomics reference population
  - all bulls born before same date as for domestic
- Test population
  - 25% yongest bulls
  - Sire in reference
  - EDC min. 20

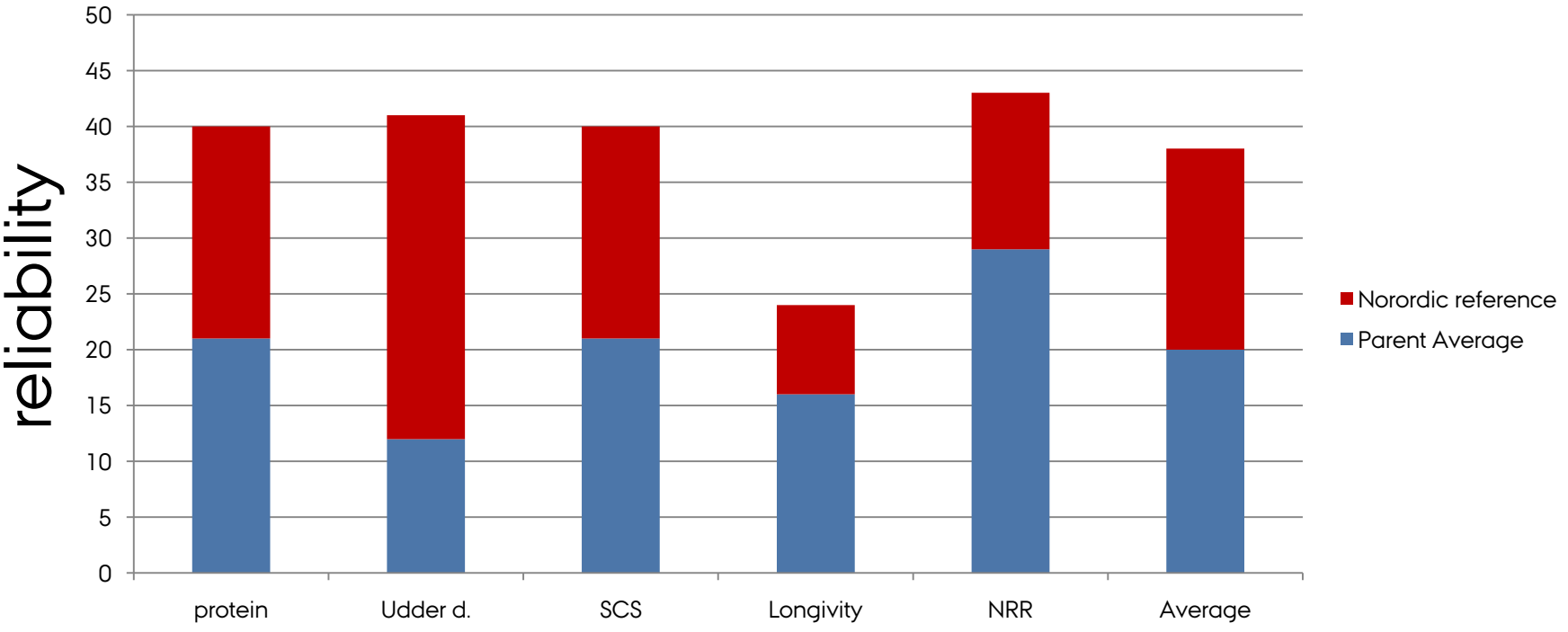
Data

- SNP data
  - DSF, Fr, Deu: USDA Illumina 50K
  - Holland: Imputed from costum 50K Illumina chip
- Phenotypes
  - Deregressed proofs of Interbull jan. 2010 (domestic scale)
- Traits
  - protein, udder deapth, SCS, Longivity, NRR

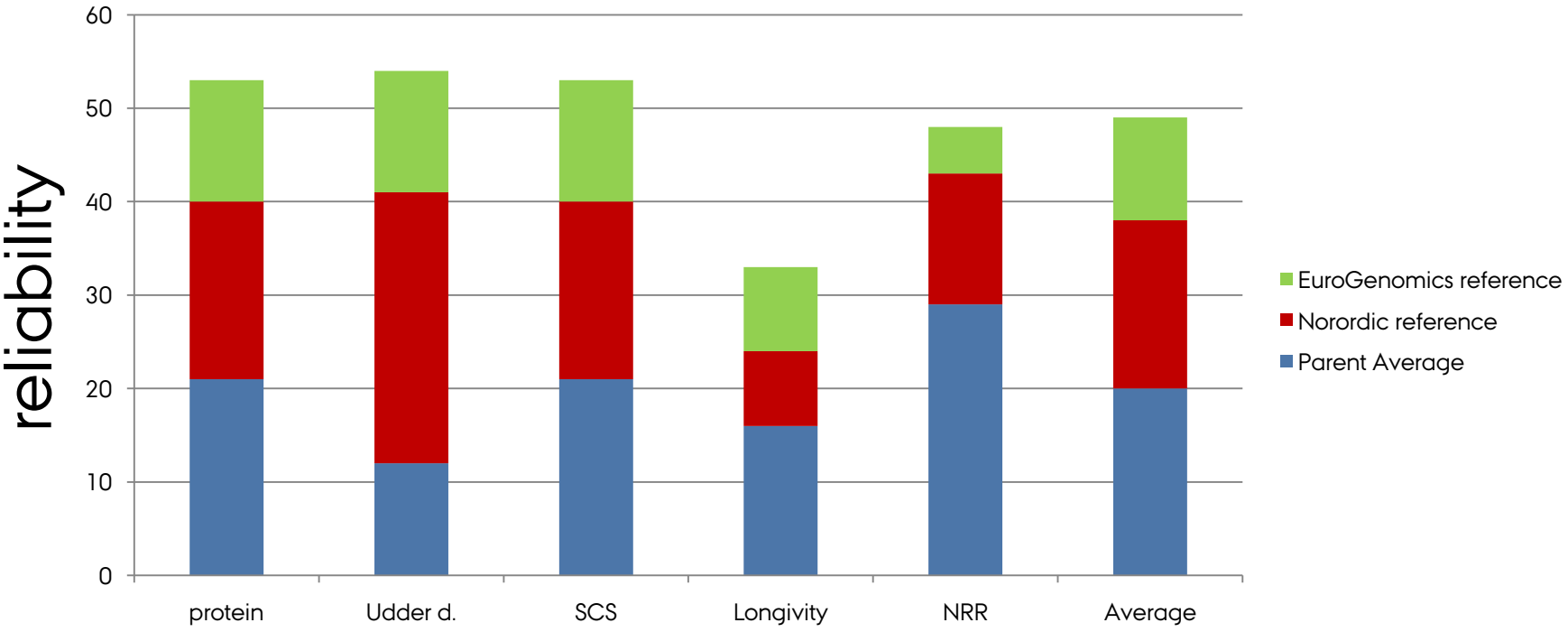
Reliability of parent average



Reliability of GEBVs using Nordic reference



Reliability of GEBVs using European reference



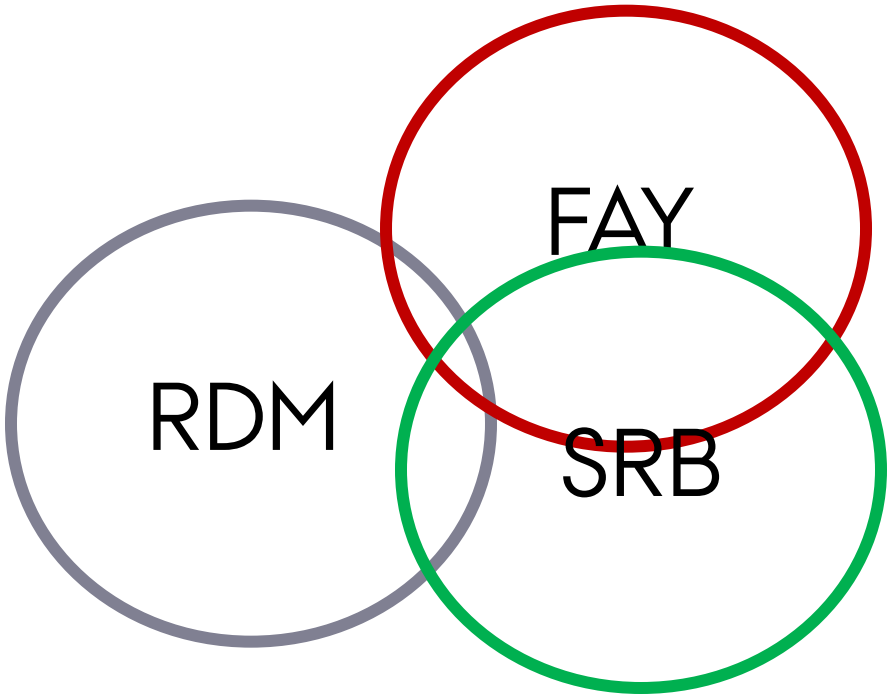
Conclusions

- Reliability increased by 11% (less than expected?)
- Starting point higher than expected
- Genetic correlation between traits less than 1 over the 4 countries

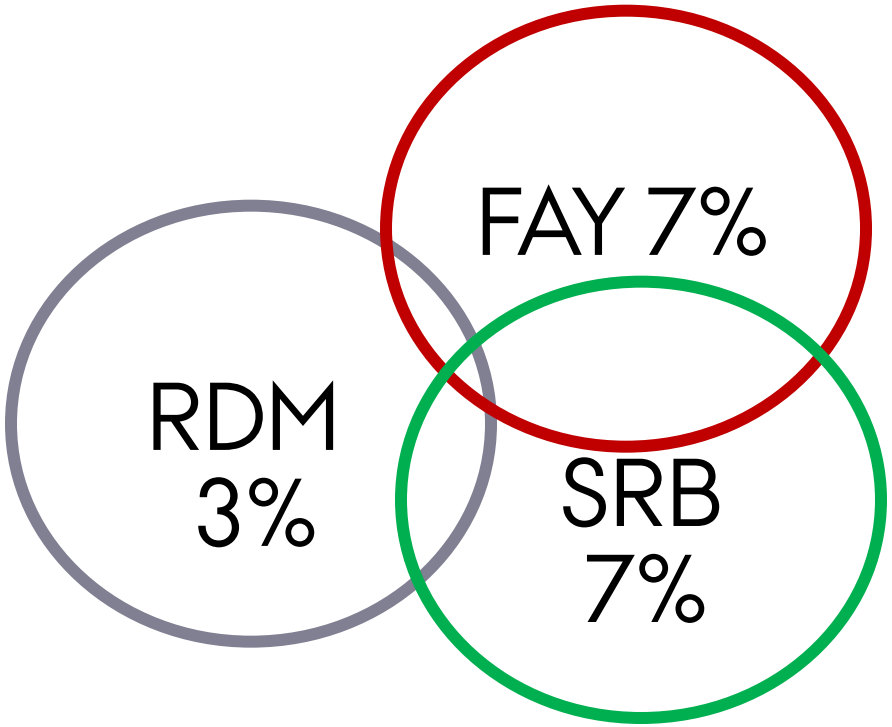
What about the other breeds?

- Jersey
- Red Nordic breeds
  - RDM
  - SRB
  - FAY

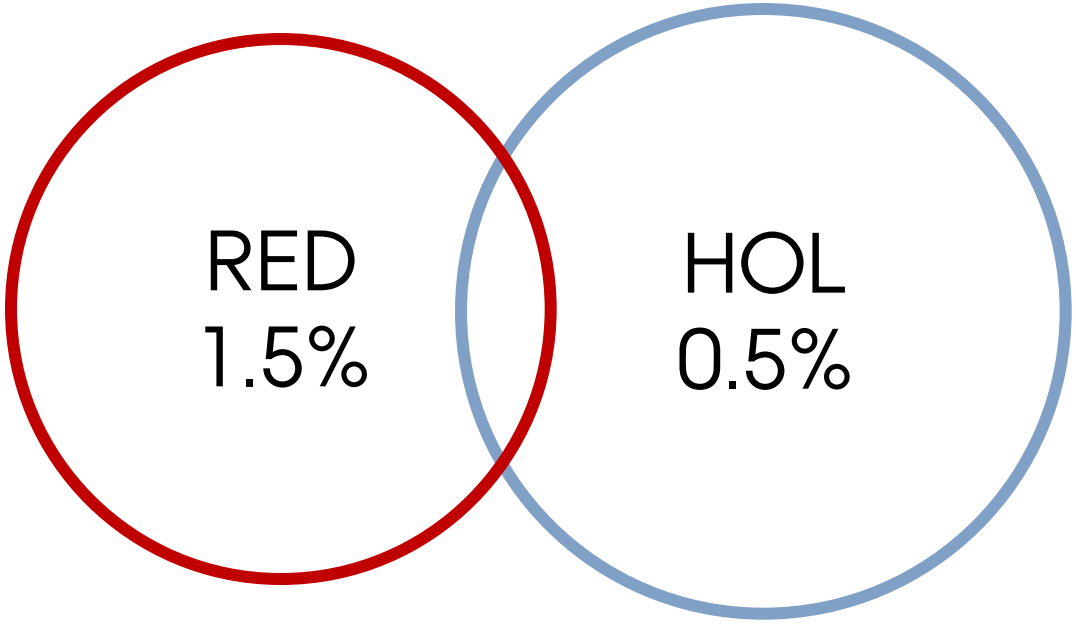
Clustering of the red breeds?



Increase in accuracy by common reference (SLU,AU, MTT)



Increase in accuracy by common reference with Holstein



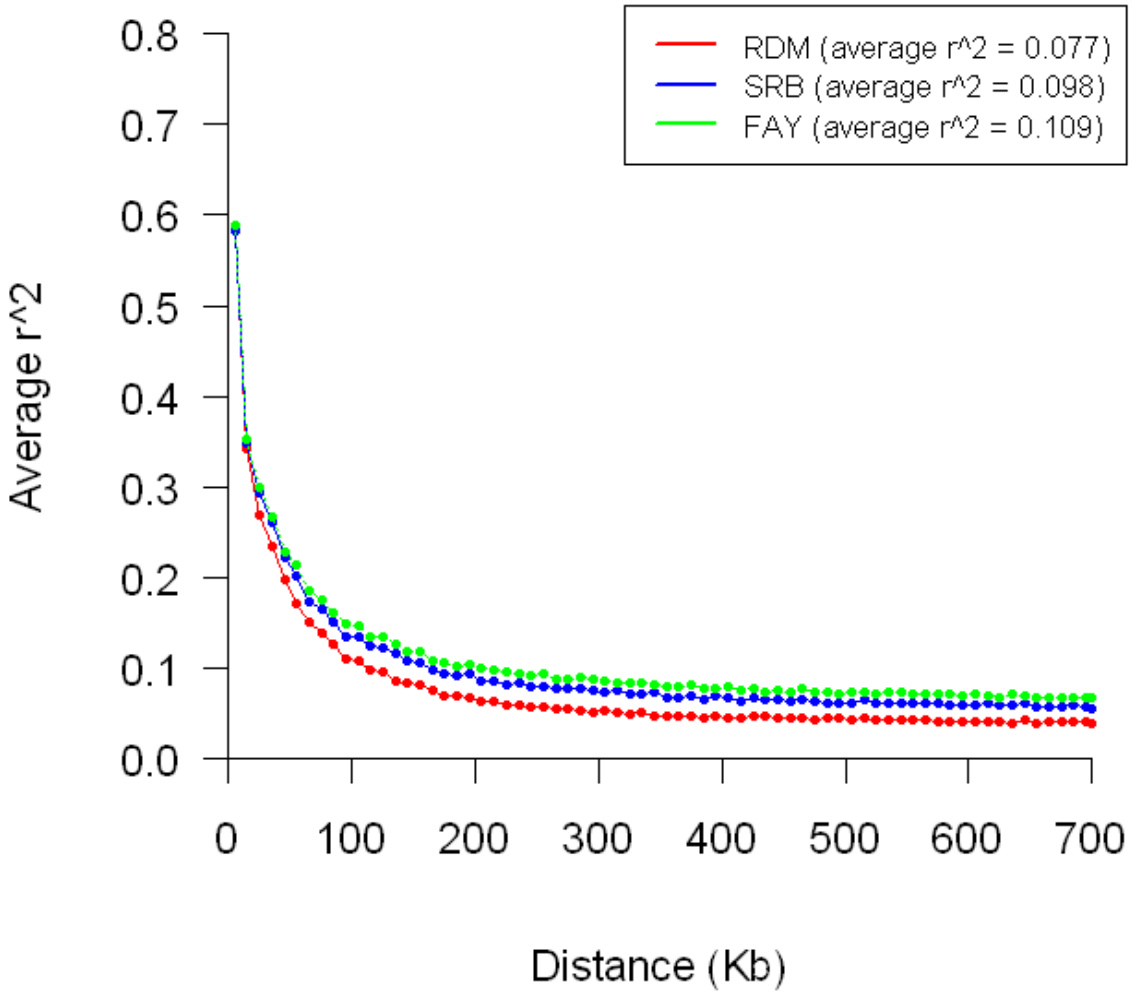
Reliability affected by:

✚ More bulls with sire in reference
✚ Better estimation of SNP effects

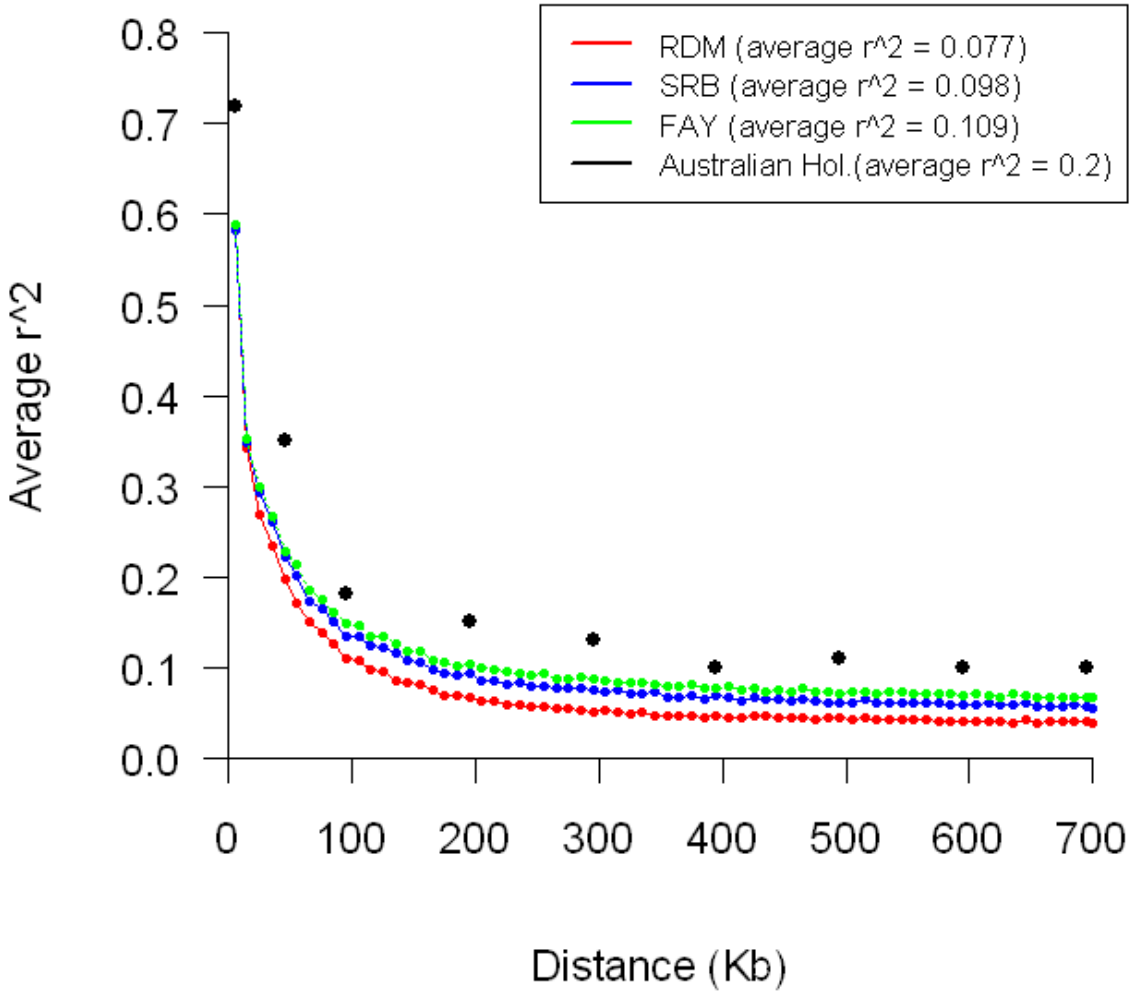
− Several (related) breeds
− Relationships within breed

Some effects cancel out !

Strength of Associations

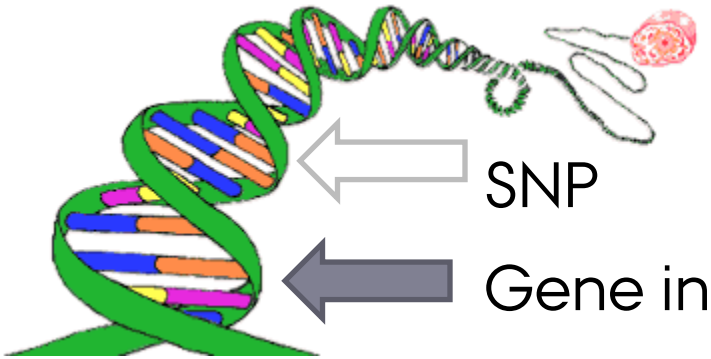


Strength of Associations



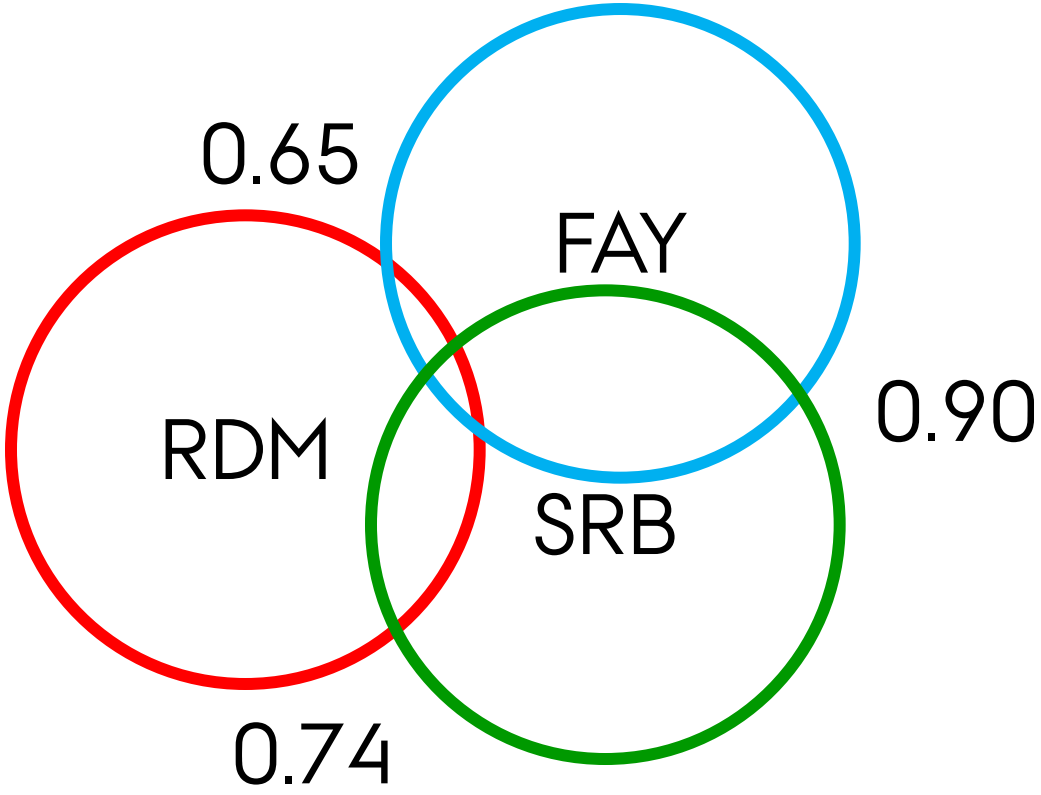
“Direction” of Association

Persistency of Phase

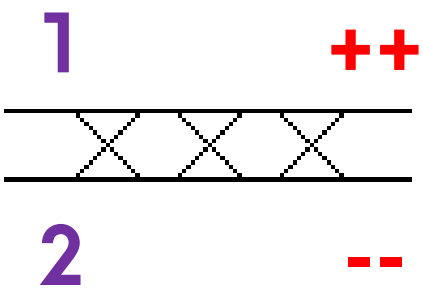


SNP Allele	Effect on Mastitis Resistance
1	++
2	--

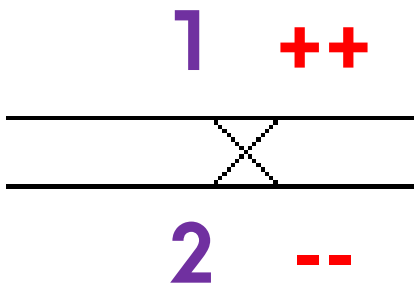
Persistency of Phase



SNP Density and Associations



More crossing over possible for genes far apart



Fewer crossing over possible for genes close together

Improvements with BovineHD

- Distance between SNP and QTL becomes much shorter
  - Stronger associations within breed

    (Marginal effect)
  - Direction of association more similar across breed

    (Large effect, especially for RDM ⇔ SRB&FAY and HOL ⇨ RED)
- Benefits greater when more data is added

Imputation to Combine

› <1% error rate for Bovine SNP50K & CRV50K

› ~3% error rate for Bovine SNP50K & 3k

? Error rates for Nordic Red

Available SNP-chips

Illumina

- Bovine SNP50K
- Bovine HD
- Bovine 3K panel
- Customized chips (CRV)

} Subsets of each other

Affymetrix

- Bovine 25K SNP Kit

…more to come …

Upgrade SNP50K ⇨ BovineHD

- Need to genotype *some* bulls in the reference population with BovineHD
  - Only influential bulls
- Holstein
  - Reduce costs through EuroGenomics
    - 500+500
- Nordic Red
  - Evaluate how much re-genotyping is needed
  - Second half of 2010

Opportunities with 3K

- Large scale screening of potential bull dams
  - Contract matings
  - Embryo flushing

- Cows comprise the future reference population ?!
  - Tendency to reduce number of bulls

New data sources and traits

- Recording devices
  - Milk flow meters
  - Voluntary Milking systems
  - Activity meters
  - Herd navigator
  - …

- Collect phenotypic data
- Genotype cows with 3K panel

} Genomic selection for new traits

Summary

- Reliability of GEBVs depend strongly on reference size
- Large increase for HOL in EuroGenomics
- Smaller increase for Nordic Red breeds
- High density SNP set will be important
- Low density SNP set for screening and reference
- Imputation techniques are crucial