Haplotype-assisted genomic evaluations in Nordic Red Dairy Cattle

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

MTT
Agrifood Research
Finland

NAV Nordic Cattle Genetic Evaluation

EAAP 2013, Nantes, France
Wednesday 28th September 2013

Background

- Poor predictive ability of SNP data in admixed populations originating from numerous base breeds like Nordic Red Dairy Cattle
  - inability of SNP to trace genetic relationships (IBS≠IBD) in heterogeneous populations
  - inconsistencies of LD between SNP and QTL
- Haplotypes may be more useful than SNP to recover genetic relationships and LD

Objective of the study

- Improve the prediction accuracy by use of IBD information in haplotypes of chromosomal segments

- Criterium: validation test reliability

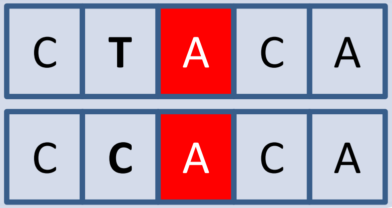
Outline of statistical procedure

Pre-selection of markers

- Genome scan with BayesB to detect QTL signals
- Using SNP genotypes
- Ranking of signals using absolute effect size

Construction of haplotype blocks

- Chromosomal segments around pre-selected SNP
- Using flanking markers



C	**T**	A	C	A
C	**C**	A	C	A

Estimation of haploblock variances

- BayesA
- To achieve correct weighting in multilocus model

Estimation of haplotype effects for prediction

- Solutions of MME
- Inclusion of a polygenic pedigree-based effect

Methods: Estimation of haploblock variances

$$\boxed{\mathrm{DRP} = \mu + \sum_{m=1}^{750/1500} \left(\underbrace{g_{m,1}^{\mathrm{hap}}}_{\mathrm{paternal}} + \underbrace{g_{m,2}^{\mathrm{hap}}}_{\mathrm{maternal}} \right) + \varepsilon}$$

- BayesA used to estimate variance of effects in each haplotype block $\sigma_{gm}^2 = Var(g_m^{\mathrm{hap}})$
- Estimates used to give differing weights in the evaluation model

Methods: Estimation of haplotype effects and prediction

$$\boxed{\mathrm{DRP} = \mu + a + \sum_{m=1}^{750/1500} \left(g_{m,1}^{\mathrm{hap}} + g_{m,2}^{\mathrm{hap}} \right) + \varepsilon}$$

- Solutions of MME with
  - proportion ω of the genetic variance assigned to pedigree $\mathrm{Var}(a) = \omega \hat{\sigma}_a^2 \mathbf{A}$
  - rest $(1-\omega)$ assigned to weighted haploblocks $\mathrm{Var}(g_m^{\mathrm{hap}}) = (1-\omega)\hat{\sigma}_a^2 \cdot \hat{\sigma}_{gm}^2 / S$

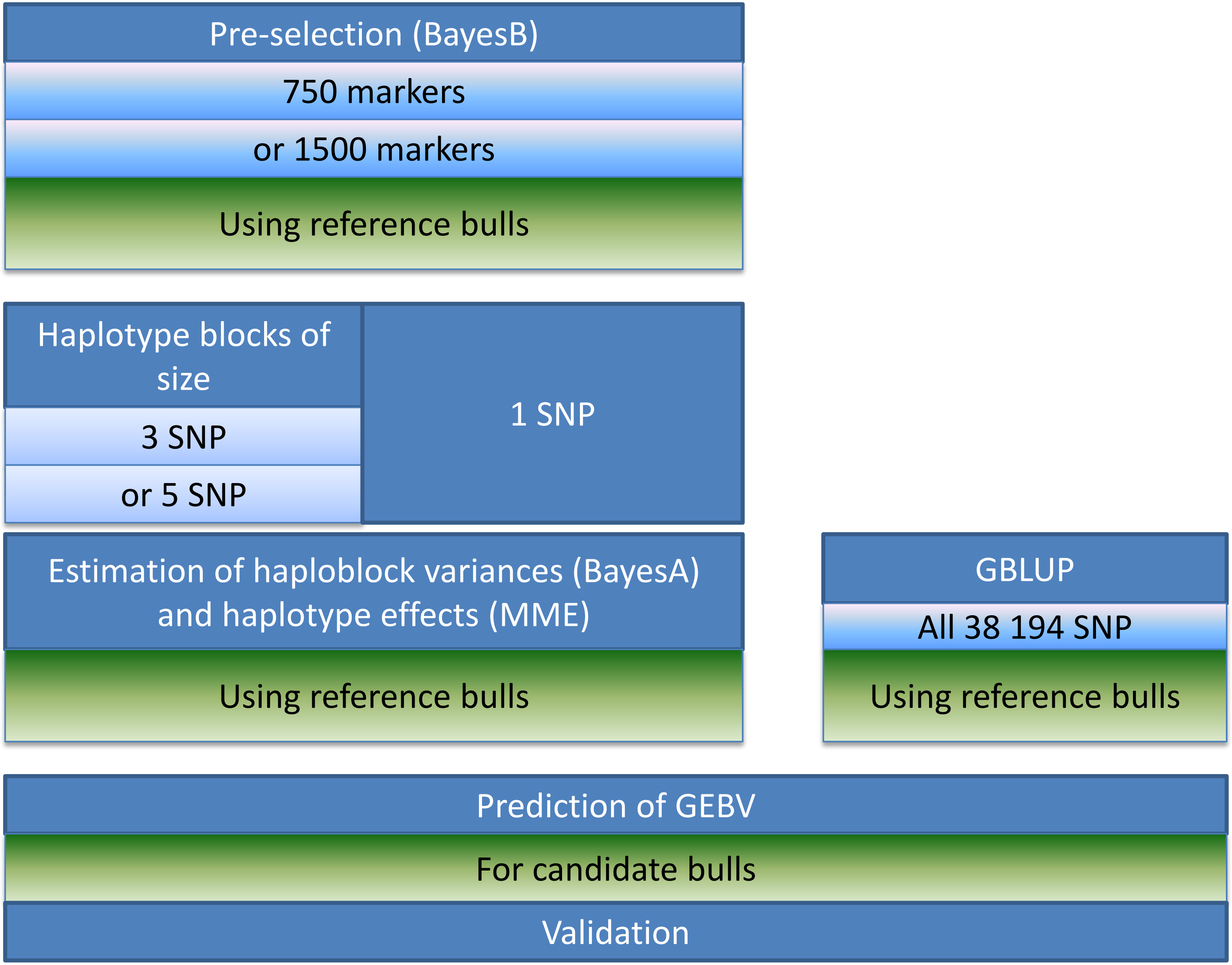
$$\boxed{\mathrm{GEBV} = \hat{a} + \sum_{m=1}^{750/1500} \left(\hat{g}_{m,1}^{\mathrm{hap}} + \hat{g}_{m,2}^{\mathrm{hap}} \right)}$$

Data

- DRP and effective daughter contributions (=weights of observations) from Nordic Cattle Genetic Evalution (February 2013)

- 38 194 SNP (after editing) on autosomes from Illumina Bovine SNP50 BeadChip - imputation of missing genotypes and phasing with BEAGLE v3.3 (Browning & Browning 2009)

	Production traits: milk, protein, fat	Fertility
Reference bulls (born 1971-2005)	4250	4422
Candidate bulls (born 2006-2008)	516	551
Total	**4766**	**4973**



Pre-selection (BayesB)

750 markers

or 1500 markers

Using reference bulls

Haplotype blocks of size

3 SNP

or 5 SNP

1 SNP

Estimation of haploblock variances (BayesA) and haplotype effects (MME)

Using reference bulls

GBLUP

All 38 194 SNP

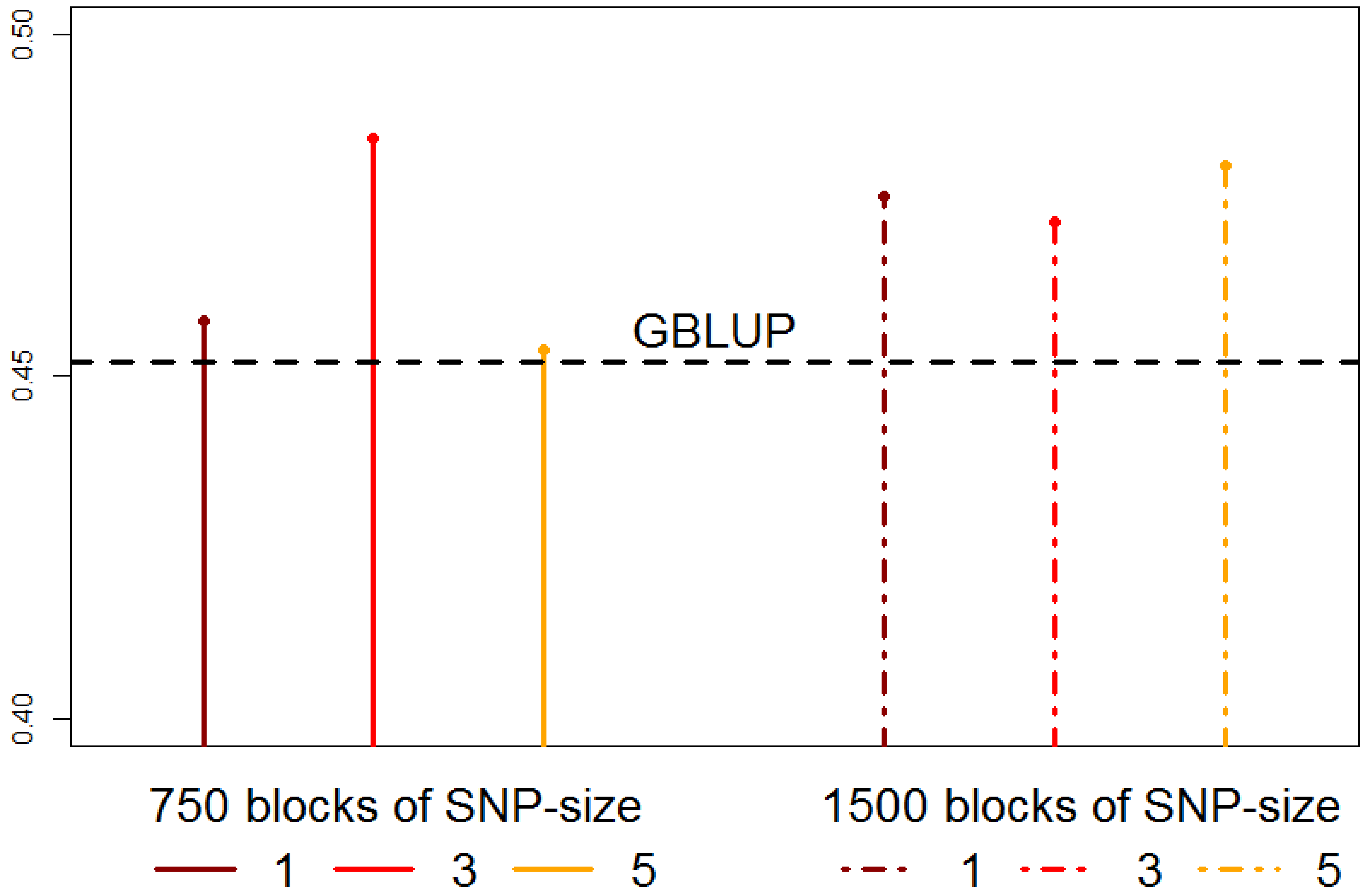
Using reference bulls

Prediction of GEBV

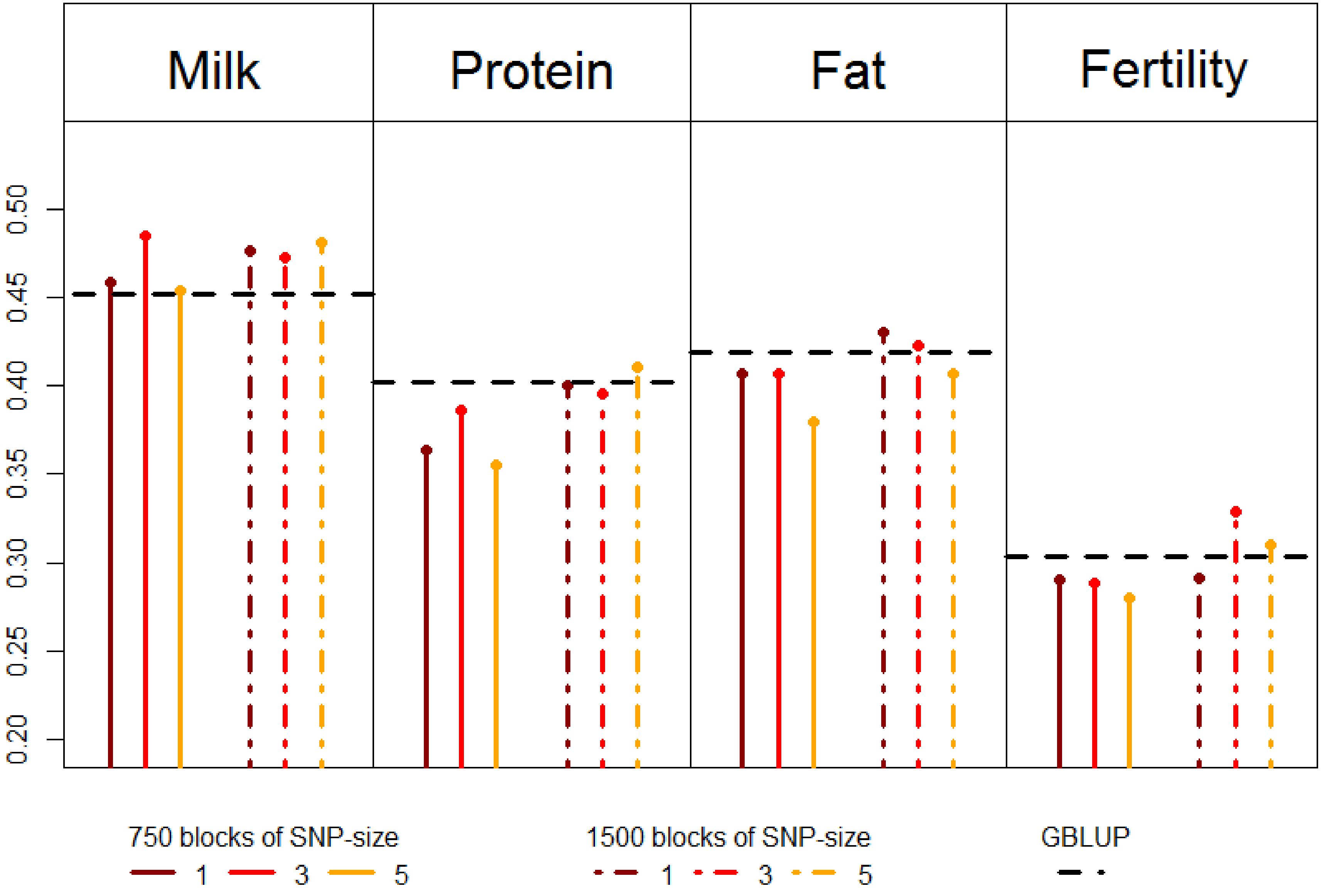
For candidate bulls

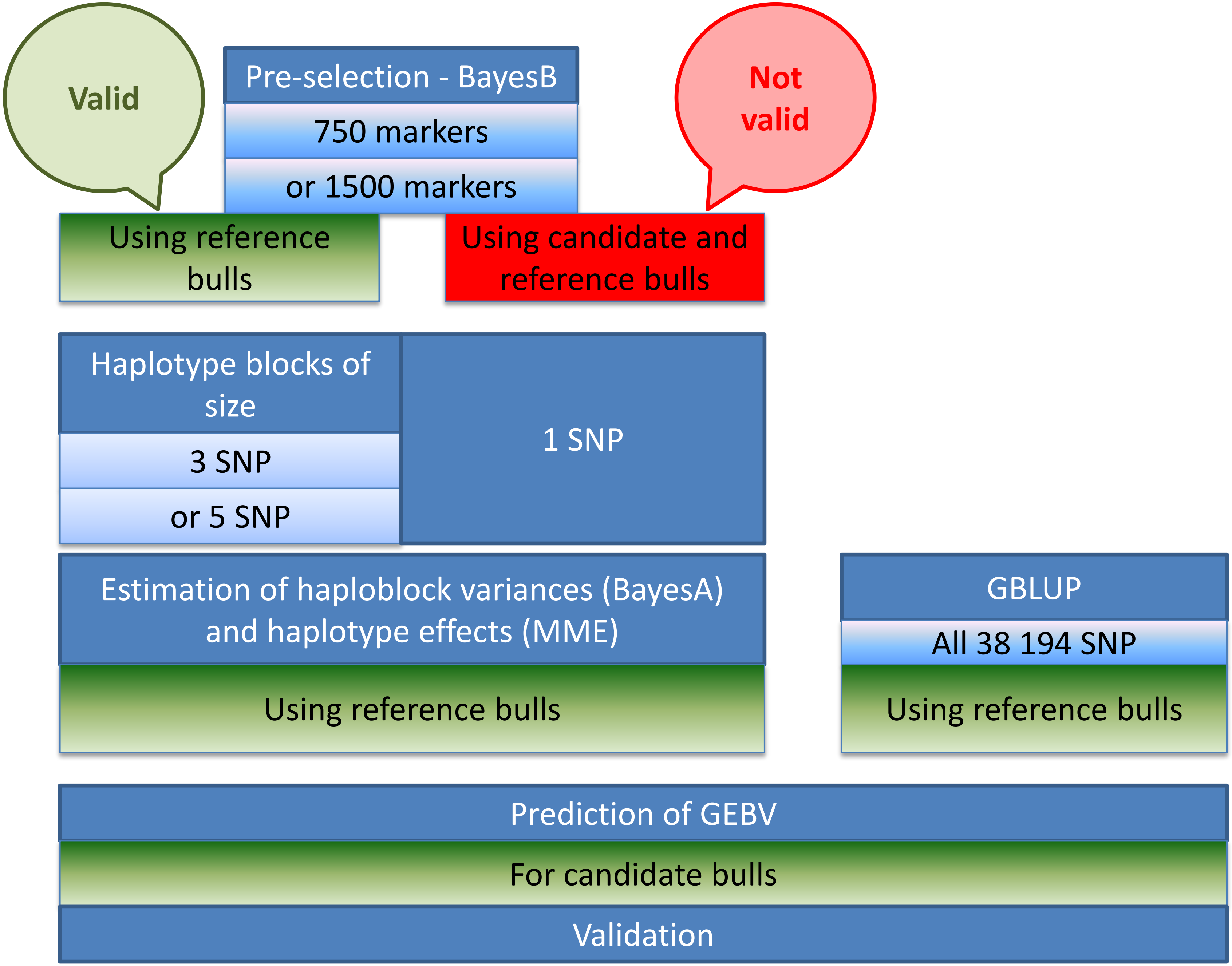
Validation

Results –Validation test reliabilities R^2 for milk



Results – Validation test reliabilities R^2





Valid

Pre-selection - BayesB

750 markers

or 1500 markers

Not valid

Using reference bulls

Using candidate and reference bulls

Haplotype blocks of size

3 SNP

or 5 SNP

1 SNP

Estimation of haploblock variances (BayesA) and haplotype effects (MME)

Using reference bulls

GBLUP

All 38 194 SNP

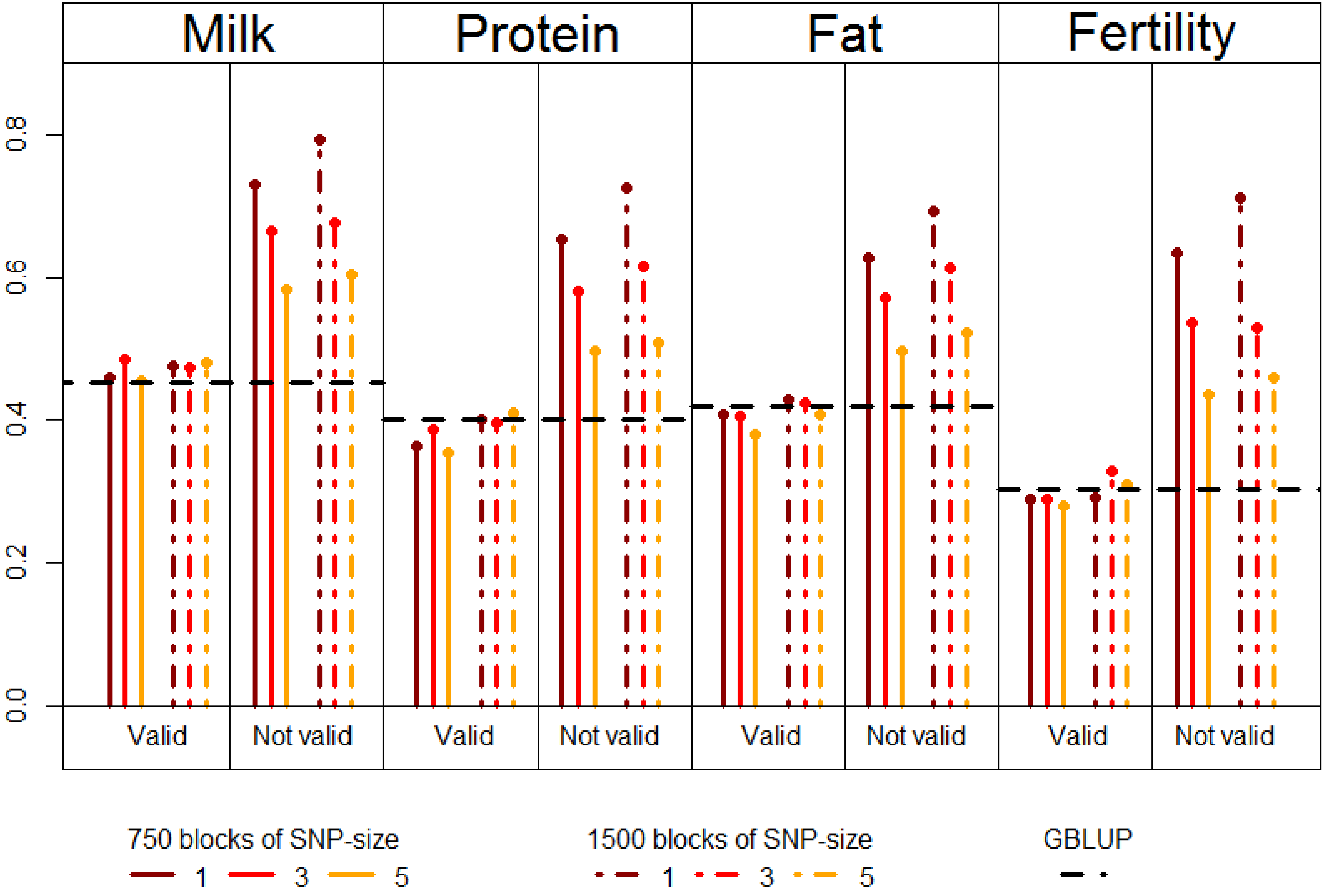
Using reference bulls

Prediction of GEBV

For candidate bulls

Validation

Results – Validation test reliabilities R^2



Conclusions

- With valid marker pre-selection:
  - at best a small advantage over GBLUP
  - number of SNP in haplotype blocks has little impact
  - single SNP perform as well as haplotype blocks (because of using SNP in BayesB for pre-selection?)

- Results for non valid pre-selection suggest that there may be huge potential for improvement:
  - Different pre-selection method needed to pick up ”good” markers/avoid ”bad” markers
  - But: set of ”good”/”bad” markers may change over time

Acknowledgements

- Phenotypic and pedigree data was obtained from the Nordic Cattle Genetic Evaluation NAV

- Genotypes of animals from the Nordic Genomic Selection–project (Viking Genetics, Aarhus University, NAV, FABA, Svensk Mjölk - current Växa Sverige)