Accuracy of genomic predictions using subsets of SNP markers in dual purpose Fleckvieh cattle

Birgit Gredler[1,2], Mahlet A. Teka[2], Hermann Schwarzenbacher[3], Christa Egger-Danner[3], Johann Sölkner[2]

[1]University of Natural Resources and Life Sciences Vienna, Austria
[2]Qualitas AG, Switzerland
[3]ZuchtData EDV Dienstleistungen GmbH, Austria

Introduction

- Genomic selection can be more cost effective if low density SNP panels are available at a lower price
  - Genotyping of cows
- Low density panels deliver high proportion of accuracy of a high density panel (e.g. Moser et al., 2010)
- Selected SNPs on the low density panel may work for single traits but not across traits and not across breeds
- Objective was to assess the predictive ability of subsets of SNP in dual purpose Fleckvieh

Data

- 5,556 dual purpose Fleckvieh bulls
- Bulls born between 1969 – 2005
- Genotyped with Illumina 50k Chip
- Quality checking of SNP – data:
  - Call rate
  - MAF
  - HWE
  - Pedigree checking
- After filtering 41,008 SNP

Phenotypes

- Traits:
  - Total merit index (TMI)
  - Milk yield (MY)
  - Protein yield (PY)
  - Fat percentage (F%)
  - Somatic cell score (SCS)
  - Longevity (LO)
  - Female fertility (FE)
  - Male calving ease (CE)

- Deregressed breeding values (Garrick et al. 2009)

Reference and validation set

- Reference set: bulls born before 2003
- Validation set: bulls born 2003, 2004, 2005

Trait	N Reference	N Validation	Trait	N Reference	N Validation
TMI	3,731	1,662	SCS	3,757	1,676
MY	3,731	1,662	LO	3,751	1,010
PY	3,731	1,549	FE	3,749	902
F%	3,731	1,549	CE	3,756	1,800

Methods

- **SNP-BLUP** and **BayesB** (Meuwissen, 2009)
- Statistical model: $$y = \mu + \sum_{j=1}^{Nm} X_j a_j + e$$
  - y = vector of phenotypes
  - Nm = number of markers
  - μ = overall mean
  - a_j = effect of marker j
  - X_j = vector allocating genotype of bulls for marker j
  - e = vector of residuals – e ~ N(0,I σ_e^2 / w)
  - σ_e^2 = error variance
  - w = weighting factor (bull ‘s reliability from daughters with parent information excluded)
- Predictive ability: correlation between direct genomic breeding values and conventional breeding values

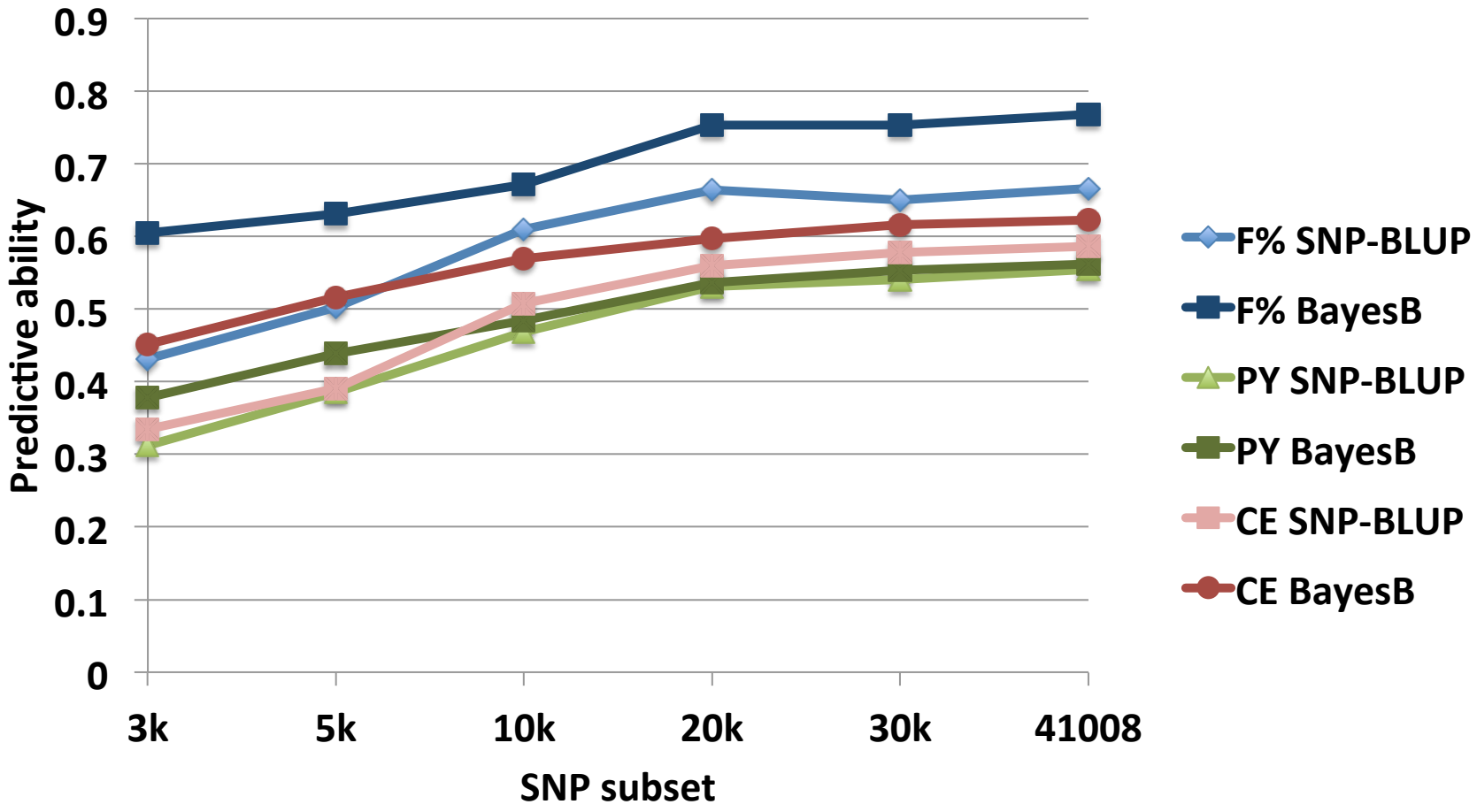
SNP selection

- Full set of SNP (n = 41,008)
- Illumina 3k Chip (n = 2,889)
- Random selection (PY, F%, CE)
  - 5k: 3k + 2k randomly sampled
  - 10k: 3k + 7k randomly sampled
  - 20k: 3k + 17k randomly sampled
  - 30k: 3k + 27k randomly sampled

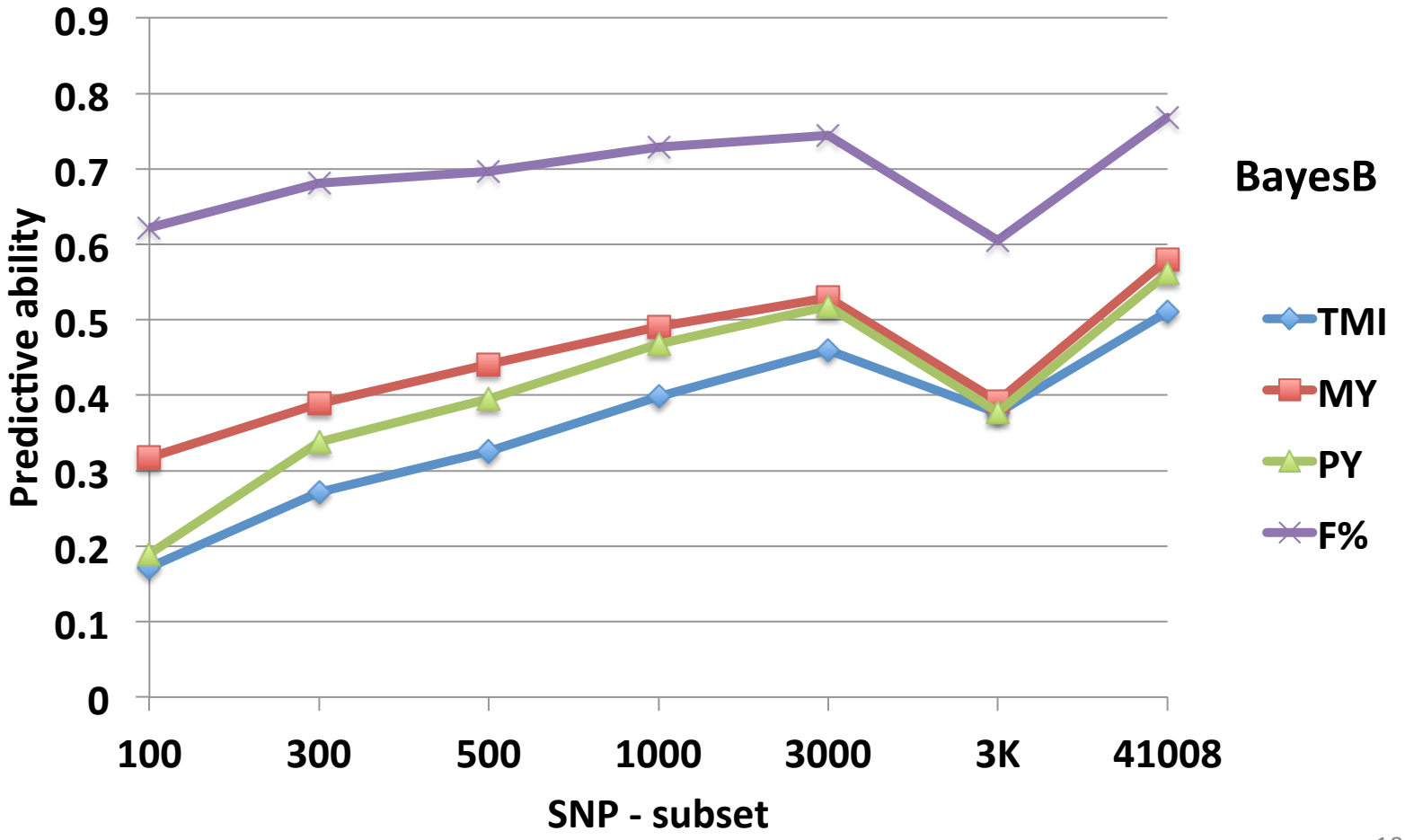
SNP selection

- SNP selection according to absolute size of BayesB SNP-effect within trait
  - 100 SNPs
  - 300 SNPs
  - 500 SNPs
  - 1,000 SNPs
  - 3,000 SNPs

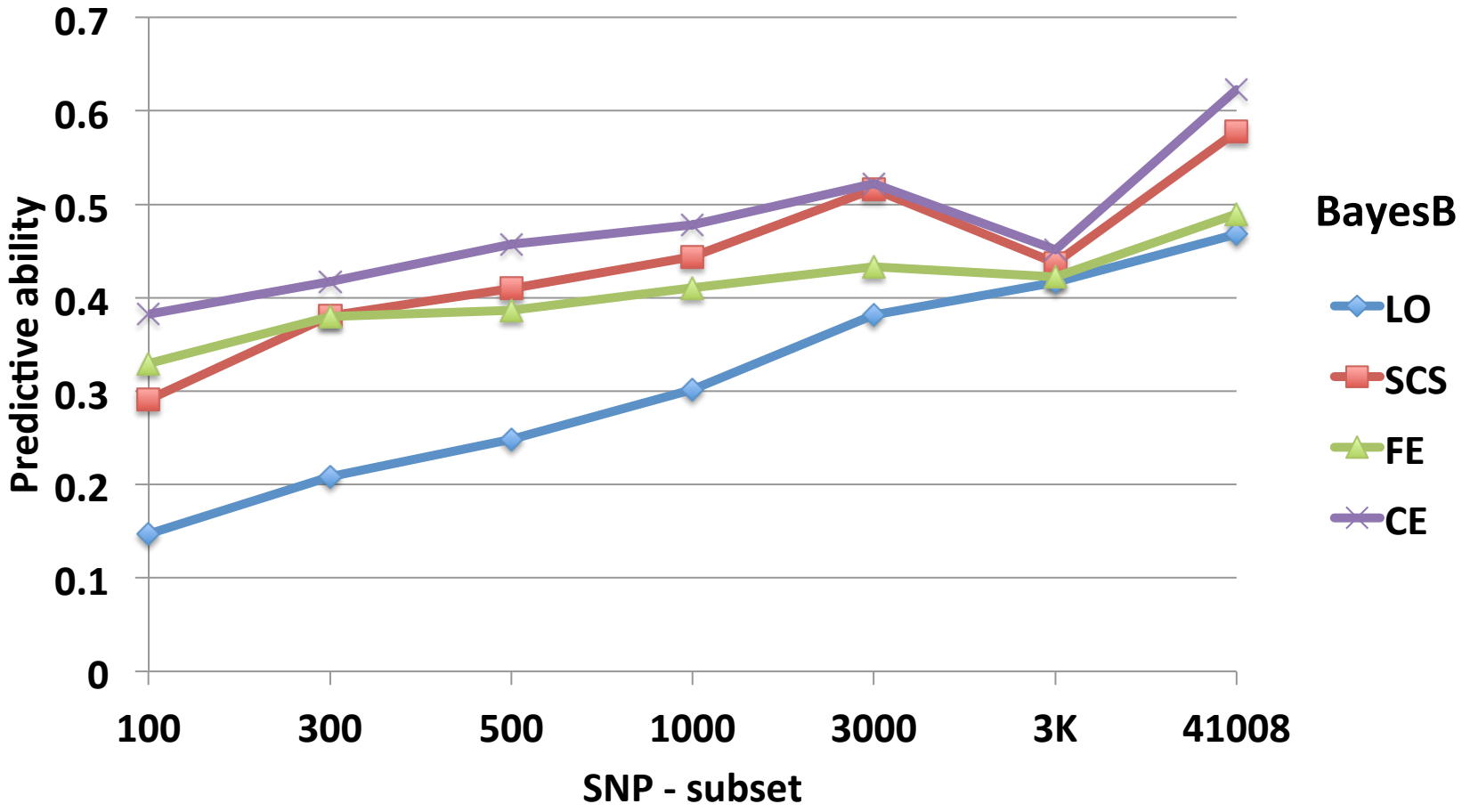
Predictive ability – random SNP selection



Predictive ability
selection according to BayesB effect size



Predictive ability
selection according to BayesB effect size

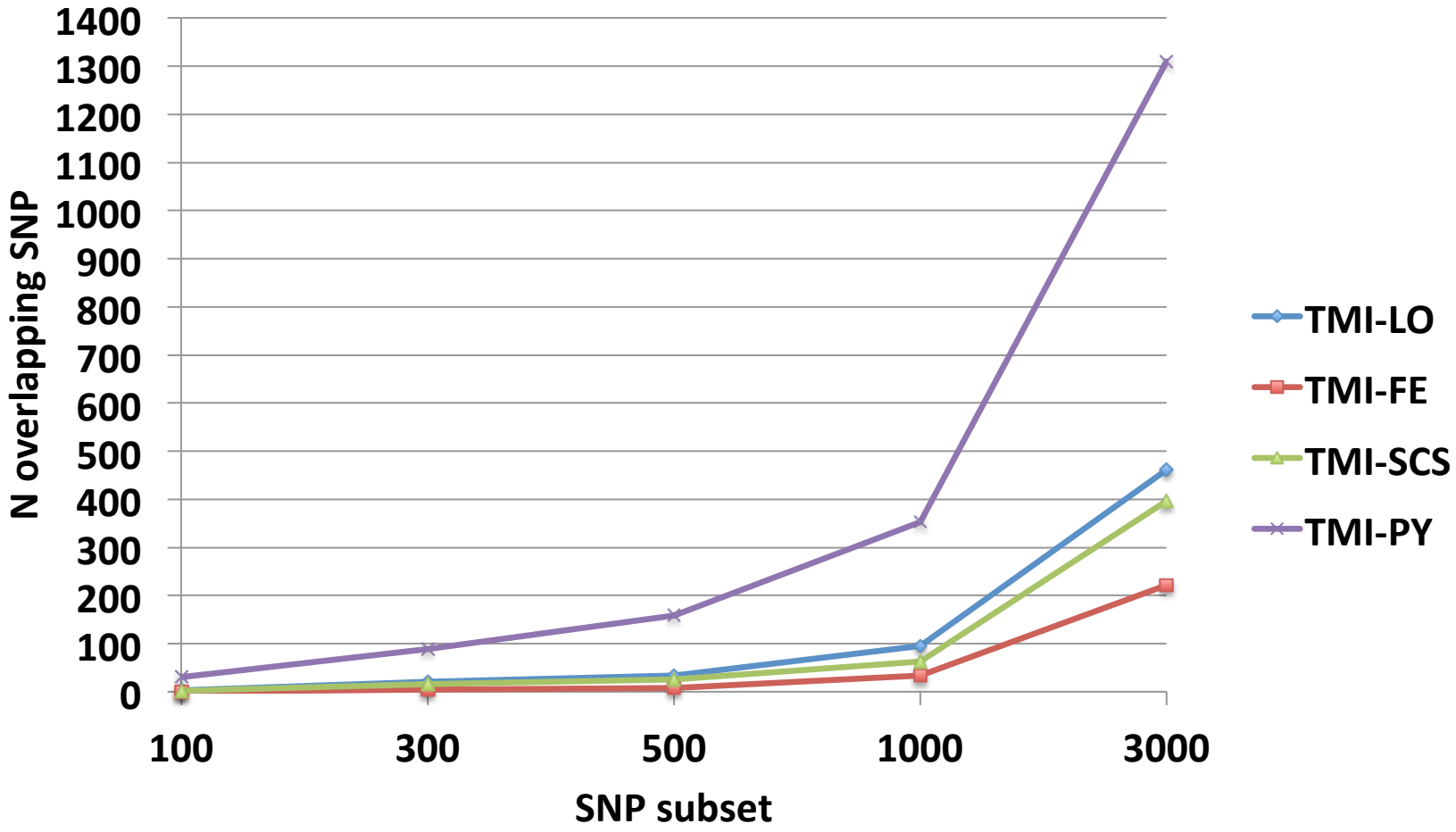


Predictive ability
selection according to BayesB effect size vs. 3k

Percentage of predictive ability with 3k and 3,000 SNP with highest effect compared to full set

Trait	3k	3,000	Trait	3k	3,000
TMI	73.8 %	89.9 %	LO	88.9 %	81.4 %
MY	67.5 %	91.2 %	SCS	75.7 %	89.3 %
PY	67.2 %	92.2 %	FE	86.2 %	88.4 %
F%	78.8 %	96.9 %	CE	72.4 %	83.6 %

Number of SNP overlapping across traits



Summary

- Predictive ability of Illumina 3k chip varies between 67 % and 89 % of Illumina 50k chip
- Trait specific subsets give higher accuracy compared to Illumina 3k
- Highest accuracy reached with full set of SNP for all traits
- If low density chips are used, imputation from low to high density seems to be more effective in Fleckvieh
- Use of new Illumina 6k chip (?) or Illumina HD (777k) chip

Acknowledgements

Genotype pool Fleckvieh bulls in Austria and Germany

Federation of Austrian Fleckvieh Cattle Breeders (AGÖF)

Federation of Cattle Breeders in Southern Germany (ASR)

Förderverein Biotechnologieforschung (Germany)

Institute of Animal Breeding (Bavarian State Research Center for Agriculture)

Institute of Animal Breeding and Husbandry (Christian Albrecht University Kiel)

ZuchtData EDV-Dienstleistungen GmbH

KeLeKi