Reliability of genomic prediction using low density chips

D.Segelke* †, J.Chen*, Z.Liu*, F.Reinhardt*, G.Thaller †, R.Reents*

Vereinigte Informationssystem Tierhaltung vit, Heideweg 1, 27283 Verden (Aller), Germany
† Institute of Animal Breeding and Husbandry, Christian-Albrechts-University, 24098 Kiel, Germany

Conclusions:
The accuracy of genotype imputation from the low-density chips Illumina 3K & BovineLD to the BovineSNP50 BeadChip was investigated for 3 different data sets using the imputation programs Findhap & Beagle. Allele error rate of Findhap was almost 3 times higher than Beagle. The 6K chip was twice as accurate as the 3K chip. Using Findhap & the 3K Chip reduce the reliability of genomic prediction about 5%. Beagle and the 6K chip had a reliability loss of round 1%. The low-density chip BovineLD 6K could be used routinely for large-scale genotyping.

Introduction:
With the availability of dense SNP marker chips, genomic evaluation was implemented in many countries. But for an average dairy breeder genotyping his whole herd is still too expensive. Therefore there was a demand to develop low density chips to reduce genotyping costs.

Materials & Methods:
To study the accuracy of low density chips, Illumina Bovine3K BeadChip (3K) and BovineLD BeadChip (6K), were simulated using Illumina marker maps. Accuracy of imputing was investigated by using the software packages Beagle and Findhap. Three different genotype data sets were used: EuroGenomics data set with 14,405 reference bulls, smaller EuroGenomics data set with 11,670 older reference bulls and a data set containing 31,597 genotyped German Holstein animals. For validation, 1,374 bulls for the EuroGenomics data set, 534 bulls for the smaller EuroGenomics data set and 2,205 animals of the genotyped German Holstein data set, their imputed genotypes were compared to their original genotypes to calculate allele error rate. Additionally reliability loss by using imputed instead of original genotypes was computed. Furthermore the difference in combined genomic-enhanced breeding values (GEBV) between original and imputed genotypes were calculated.

Results:
Allele imputation error rates depended on the size of the reference population, imputation programs, chip density and relationship between reference and validation animals. Allele error rate for the EuroGenomics data set II was 3.3% for Findhap 3K, 1.6% for Beagle 3K, 1.7% for Findhap 6K and 0.6% for Beagle 6K, respectively.

Reliability of the genomic prediction decreased on average about 5.3% for Findhap 3K, 2.6% for Beagle 3K, 1.9% for Findhap 6K and 1.0% for Beagle 6K averaged over 12 selected traits, in comparison to those of using original genotypes. Differences in GEBV of original and imputed genotypes were largest for Findhap 3K, whereas Beagle 6K had the smallest difference.

Acknowledgements: The organization Förderverein Biotechnologieforschung e.V. (FBF) is thanked for financial support. The EuroGenomics consortium is kindly acknowledged for providing genomic data.