# Use of high density SNP in genomic evaluation in Holstein-Friesians

Ghyslaine Schopen
Team Research

28 August 2012

# Introduction

- BovineHD Beadchip (~777k SNP) of Illumina.
- Use of HD SNP increases computing time.
- Identification of obsolete SNP.
- Haplotype scores (i.e. ancestral haplotypes).
- No recombination from one haplotype score to another.

# Aim

To reduce the amount of HD SNP based
on haplotype scores and to investigate the effect on
the reliability of direct genomic values.

# Materials & Methods (1/2)

- Animals
  - 548 BovineHD bulls (EuroGenomics consortium).
  - Holstein-Friesians genotyped with 50k and imputed to HD.
  - 30,483 animals

- Deregressed proofs for 9 traits that are part of the TMI.

- Genotypes
  - After editing ~600k SNP.
  - Edits applied for:
    - call rate > 95%
    - minor allele frequency <0.025%
    - Hardy-Weinberg equilibrium

# Materials & Methods (2/2)

- Imputation with Beagle.
- Reduction of HD loci
  - 2 datasets:
    - 38,355 SNP (small dataset)
    - 115,690 SNP (large dataset)
- SNP from routine genomic evaluation used (standard dataset).
- Genomic validation based on Gibbs sampling (BayesB)

# Results (1/2)

- Average # of haplotype score per locus was 45.
- Imputation error averaged 0.748%
- Percentage of obsolete loci
  - small dataset 93.4%
  - large dataset 80.2%

# Results (2/2)

## Increase in reliability of DGV

| Trait | n_Ref | small | large | standard |
|---|---|---|---|---|
| feet & legs | 17,296 | 20.6 | 21.0 | 19.1 |
| somatic cell count | 18,536 | 35.8 | 35.5 | 30.3 |
| kg protein | 18,485 | 28.7 | 30.7 | 29.2 |
| kg milk | 18,484 | 35.0 | 36.1 | 35.2 |
| kg fat | 18,483 | 44.3 | 44.2 | 43.1 |
| longevity | 18,405 | 13.5 | 14.2 | 12.1 |
| non-return 56d | 17,729 | 7.0 | 10.2 | 9.5 |
| calving interval | 18,482 | 26.9 | 27.7 | 27.7 |
| udder | 18,281 | 31.7 | 31.6 | 29.4 |
| Average | | 27.1 | 27.9 | 26.2 |

# Results (2/2)

## Increase in reliability of DGV

| Trait | n_Ref | small | large | standard |
|---|---|---|---|---|
| feet & legs | 17,296 | **20.6** | 21.0 | 19.1 |
| somatic cell count | 18,536 | **35.8** | 35.5 | 30.3 |
| kg protein | 18,485 | 28.7 | **30.7** | 29.2 |
| kg milk | 18,484 | 35.0 | **36.1** | 35.2 |
| kg fat | 18,483 | **44.3** | 44.2 | 43.1 |
| longevity | 18,405 | 13.5 | **14.2** | 12.1 |
| non-return 56d | 17,729 | 7.0 | **10.2** | 9.5 |
| calving interval | 18,482 | 26.9 | **27.7** | **27.7** |
| udder | 18,281 | **31.7** | 31.6 | 29.4 |
| Average | | 27.1 | 27.9 | 26.2 |

# Discussion

- Average # of haplotype score per locus was 45:
  - low # of animals
  - difficult to estimate effect
  - effect becomes (close to) zero

- Reliabilities of DGV may be underestimated.

- Not optimal HD subset
  - addition to reliability of DGV low

# Conclusions

- Identification of obsolete loci eliminates 80 - 93% of HD loci.

- The more HD loci used in a genomic evaluation, the higher the increase in reliability of DGV.

- Using HD loci increase reliability of DGV with 1 – 2% compared to SNP currently used for routine genomic evaluation

# Thank You for your attention!

Genomic evaluation with HD loci results in a higher reliability of DGV than genomic evaluation with routinely used SNP !!!