Paper presented at the 61st Annual Conference of the
European Association for Animal Production
Heraklion, Greece, 23.-27.8. 2010, Session code G38.03

Genomic metrics of individual autozygosity, applied to a cattle population

Johann Sölkner, Maja Ferenčaković,
Hermann Schwarzenbacher, Birgit Gredler,
Christian Fürst, Ino Čurik
Correspondence: johann.soelkner@boku.ac.at

BOKU

ZuchtData
EDV-DIENSTLEISTUNGEN GMBH

UNIVERSITAS STUDIORUM ZAGRABIENSIS
MDCLXIX

Background

- Use of SNP Chip genotypes to predict levels of inbreeding
- Which way ?

- Evaluation: Correlate values of genomic metrics of autozygosity with pedigree inbreeding coefficients

Metrics of Autozygosity

1. Values based on variance of genotype values
   a) Diagonal element of the genomic relationship matrix **G**, as implemented in genomic selection procedures (Van Raden et al., 2008)
   b) Diagonal element of the genomic relationship matrix **G**, as implemented by group of Peter Visscher (Yang et al., 2010)
2. Values based on SNP homozygosity
   a) Level of homozygosity: number of homozygous loci/number of all loci.
   b) Runs of homozygosity: overall length of homozygous segments

Data

- 910 Fleckvieh bulls
  - born 1999 – 2004
- Illumina Bovine SNP50^{TM} Beadchip
- 42,198 SNP
  - >1% MAF;
  - Exclude SNP with extreme deviation from HWE
  - Autosomal SNP

Diagonal **G** (Van Raden et al., 2008)

$$G_{jj} = \sum_i \frac{(x_{ij} - 2p_i)^2}{2p_i(1-p_i)}$$

$\mathrm{x_{ij}} = 0, 1, 2 \quad \textit{for}$
homozygous allele 1,
heterozygous,
homozygous allele 2

1. Base population frequencies, p_i
   - Combine genotypes and pedigrees
   - Algorithm by Gengler (2007)
2. Simple frequency estimates, $p_i = \sum_j x_{ij} / 2n$

Diagonal **G** (Yang et al., 2010)

$$G_{jj} = 1 + \frac{1}{N} \sum_i \frac{x_{ij}^2 - (1 - 2p_i)x_{ij} + 2p_i^2}{2p_i(1 - p_i)}$$

- Average of 2 estimates
  - variance of additive genotype values
  - homozygosity excess
    [O(# hom) – *E*(# hom)] / [1 – *E*(# hom)]

Level of homozygosity

$$\frac{\text{Homozygous loci}}{\text{All loci}}$$

Runs of homozygosity

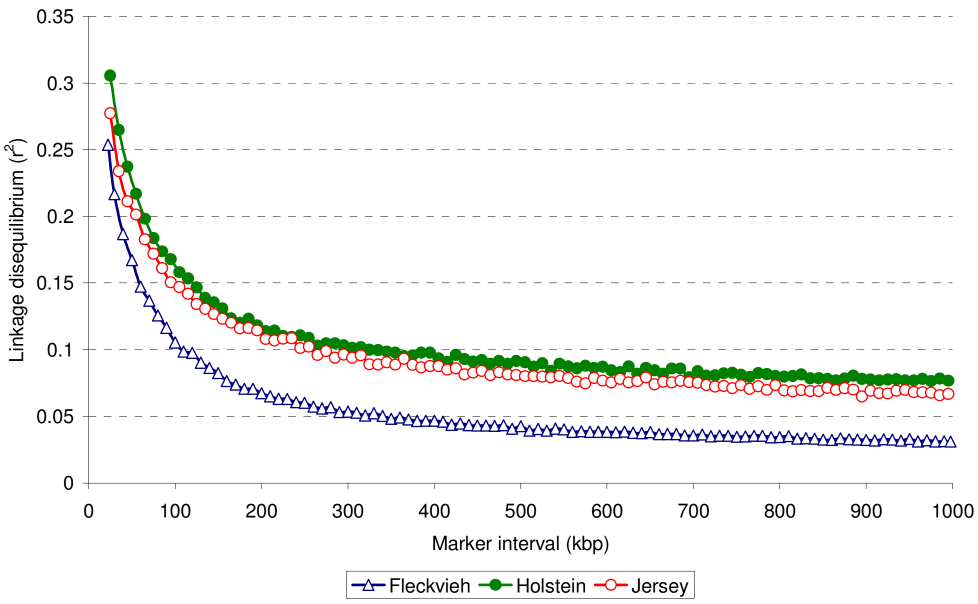
- Find and add the length of homozygous segments in the genome
  - Define minimum length (2 Mb, 4 Mb, 6 Mb, 8 Mb, 10 Mb)
  - Allow maximum number of heterozygous SNP in a segment (0 in this case)
  - Total length (Mb) of homozygous segments:

2 Mb	4 Mb	6 Mb	8 Mb	10 Mb
819	257	123	50	40

Pedigree inbreeding (F_ped)

- Complete pedigrees
  - > 6 complete generation equivalents

- Low levels
  - Average: 0.0144
  - Range: 0 – 0.0906



Correlations with F_ped

G VR_base	0.37
G VR_simple	0.13
G Yang	0.41

Hom_tot	0.57
ROH_2Mb	0.58
ROH_4Mb	0.66
ROH_6Mb	0.69
ROH_8Mb	0.68
ROH_10Mb	0.67

Correlations of **G** based values

	G VR_simple	**G** Yang
G VR_base	0.80	0.83
G VR_simple		0.79

Correlations of homozygosity based values

	ROH 2 Mb	ROH 4 Mb	ROH 6 Mb	ROH 8 Mb	ROH 10 Mb
Hom tot	0.86	0.83	0.77	0.74	0.70
ROH 2 Mb		0.81	0.76	0.73	0.70
ROH 4 Mb			0.95	0.91	0.88
ROH 6 Mb				0.96	0.93
ROH 8 Mb					0.97

Correlations of **G** based and homozygosity based values

	Hom tot	ROH 2 Mb	ROH 4 Mb	ROH 6 Mb	ROH 8 Mb	ROH 10 Mb
G VR_base	0.33	0.31	0.44	0.49	0.51	0.50
G VR_simple	0.07	0.07	0.19	0.23	0.25	0.24
G Yang	0.62	0.53	0.59	0.58	0.58	0.55

Discussion

- 3 metrics based on variance of genotype values relatively highly correlated (~0.80)
- Correlations of these to homozygosity based metrics very different
    - 0.07-0.25 for **G** VR_simple
    - 0.31-0.51 for **G** VR_base
    - 0.53-0.62 for **G** Yang
- Correlations of **G** VR_base and **G** Yang to F_ped similar and moderate (0.37, 0.41)

Discussion

- Correlations of homozygosity based measures to F-ped are much higher than those of variance of genotypes based measures (>0.60 vs. ~0.40)
- Is F_ped a good reference?
  - Simulations potentially useful
- Other **G** based values to come
  - Theo Meuwissen, personal communication

Validation of formulas (in preparation)

- Analysis of inbreeding depression for sperm quality traits
- Already validated via pedigree inbreeding coefficients
- Fleckvieh, Brown Swiss, Tyrol Grey

Conclusion

- **If** homozygosity based values of inbreeding are substantially more accurate in predicting level of autozygosity and levels of inbreeding depression than variance of genotype values based ones, we advocate use of these for computing the diagonal of **G**

Acknowledgments

- AGÖF (Association of Austrian Fleckvieh breeders) for the genotypes
- Peter Visscher, Jian Yang, Mike Goddard for provision of the derivation of formula (5) in Yang et al., 2010