

Genomics to avoid inbreeding and increase diversity

Theo Meuwissen

Norw. Univ. Life Sci. & Nofima

Ås, Norway



Introduction

- GS adopted in many countries
- Expectations high:
 - accurate GEBV for:
 - Young genotyped animals
 - Nonrecorded animals (trait nor pedigree)
 - Short generation interval
 - Difficult /costly traits
 - Animals living in a different environment (GxE)
 - Perhaps also to manage inbreeding

Inbreeding:

- Expectation:
 - GS reduces ΔF because it increases $r_{\text{within_fam}}$
 - Less selection from same families (Daetwyler et al. 2007)
- But:
 - GS reduces generation interval (in cattle)
 - More rounds of selection / time period
 - Fewer parents per generation
 - $r_{\text{within_fam}}$ was very good in progeny test

AIM:

- What happens with ΔF when:
 - Changing from progeny test to GS scheme
 - Use genomic data to manage inbreeding
 - And how to do this
- How to use genomics to increase diversity
 - Counteract influx from foreign genetics
 - Counteract 'holsteinisation'

Moving from progeny test scheme to GS

A simulation study



Simulation of base population

- $N_e=200$ (Fisher-Wright idealised pop.)
- 2,000 ($=N_e * 10$) discrete generations
- 30 chroms of 1 Morgan each (10^6 bp)
- Mutation $10^{-8}/\text{bp}$ (infinite sites mutation mod)
- Recombination $10^{-8}/\text{bp}$
- 3,000 random SNPs \Rightarrow QTL
 - QTL effects from double exponential distrib.
- 15,000 SNPs with highest MAF \Rightarrow markers
 - Marker \neq QTL

Simulation of breeding scheme

- Not possible to simulate entire cattle population
- Reduced size of simulated population
 - Number of selected males the same (in SD and SS)
 - Selected selection intensities identical
 - Conventional scheme: similar ΔG and ΔF
 - progeny test: keep test population outside breeding pop
 - Simulate progeny test results by DYDs:

$$DYD_i = \frac{1}{2}TBV_i + \sqrt{\frac{\frac{3}{4}V_g + V_e}{N_{dghtrs}}} * r_i$$

TBV & GEBV

$$TBV_i = \sum_{j=1}^{3000} x_{ij1} g_{j1} + x_{ij2} g_{j2}$$

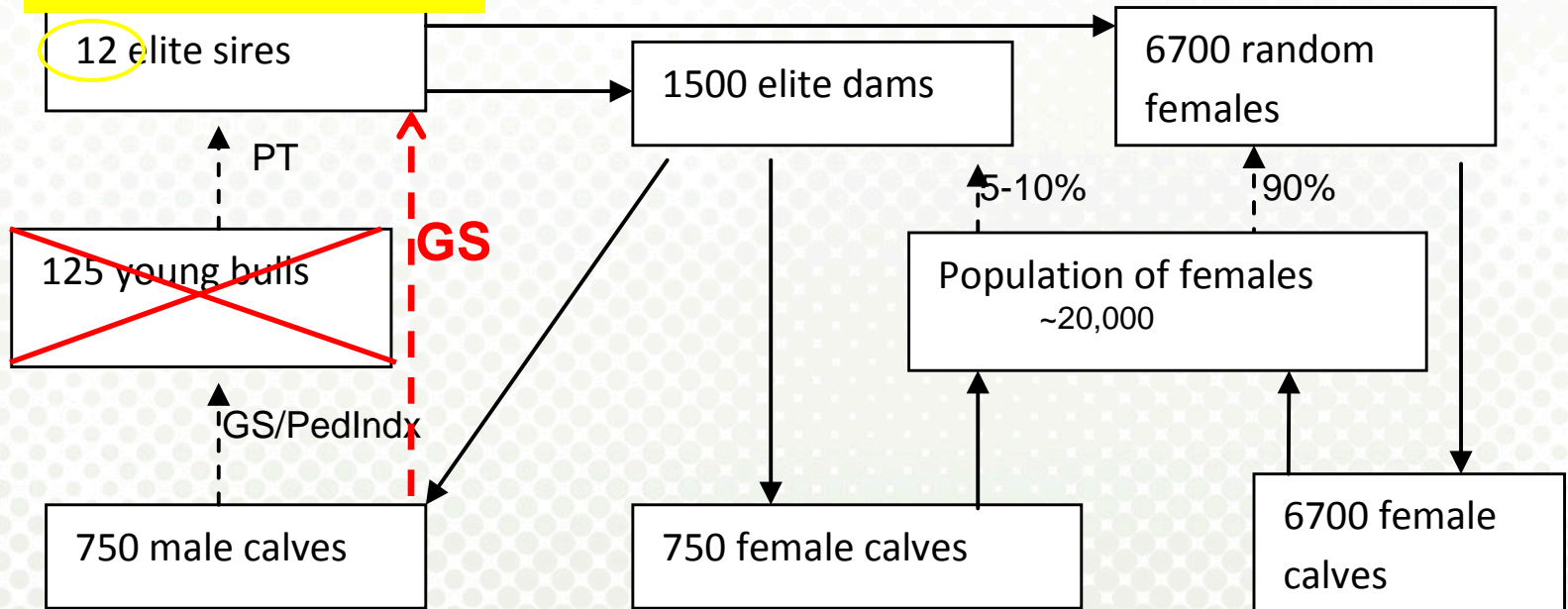
- $Y_i = TBV_i + e_i$
- $e_i \sim N(0, V_e)$
- V_e is adjusted so that h^2 is .1, **.15** or .30
- GS: only applied to young-bulls;
 - GBLUP (BLUP of marker effects; no blending):

$$y_i = \mu + \sum_{j=1}^n X_{ij} a_j + e_i$$

$$GEBV_i = \sum_{j=1}^n x_{ij} a_j$$

Genomic selection scheme (GS)

• 20 / 30 / 40



-----> Selection (TBLUP, unless stated otherwise)

GS schemes

	ΔG	ΔF	Acc
Conv	1	1	xx
GS_12	1.33	0.98	0.61
GS_30	1.25	0.47	0.63
GS_40	1.2	0.36	0.63

Conv: $\Delta G=0.22 \sigma_g/\text{yr}$; $\Delta F=.0025/\text{yr}$

Lillehammer et al, 2011

Effect of h^2 on ΔG and ΔF

	ΔG		ΔF	
h^2	0.1	0.3	0.1	0.3
Conv	1	1	1	1
PS_125	1.15	1.11	0.69	0.57
GS_12	1.4	1.29	0.93	1.14
GS_40	1.25	1.17	0.33	0.35

Conclusion: effect GS on ΔF in cattle

- $r_{\text{within_fam}}$ is already high
- ΔF increases due to shorter generation int.
- Need to select some more elite sires
 - To counter inbreeding
 - Maintain a good reference population size
- Turn-over rate of elite sires higher with GS
 - Reduces ΔF

Genomic management of inbreeding



Introduction

- Past: pedigree relations used for ΔF control
 - Measures inbreeding at unlinked, neutral loci
 - Do these exist?
- Currently GBLUP: more accurate Gmatrix
- Optimum contribution selection
 - With genomic control of inbreeding ?

Principles of OC selection

- $V_{gt} = (1 - F_t)Vg_0$ \Rightarrow need control of F or ΔF
- $F_t = C_{t-1}$ \Rightarrow need control of Coancest.
- $C_t = \frac{1}{2}\bar{A}_t$ \Rightarrow need control of \bar{A}

- \bar{A} equals average relationship of parents
 - Weighted by number of offspring
 - Including self relationships
- Control of \bar{A} controls F_t and V_{gt}

Optimum Contribution Selection

(Meuwissen, '97)

- Maximises ΔG
- While restricting

$$\bar{A}_t = 2C_t = 2[C_{t-1} + \Delta F(1 - C_{t-1})]$$

$$\bar{A}_t = 2C_t \approx 2(C_{t-1} + \Delta F)$$

- Maximises genetic gain and controls ΔF
- Average relationship :
 - Pedigree: **A**
 - Genomics /SNP genotypes: **G**

Genomic relationship estimation

- In matrix notation $\mathbf{G} = \mathbf{X}^* \mathbf{X}' / n$
 - \mathbf{X} = matrix of standardised SNP genotypes X_{ij}
 - If genotypes coded 0,1,2 then mean is subtracted and divided by the standard deviation (mean = 0; sd = 1)
 - n = number of SNPs
 - $G_{ik} = \sum_{j=1:n} X_{ij}^* X_{kj} / n$
 - is correlation estimate

OC with ΔF constraint 0.005

	ΔG	ΔF -ped	ΔF -genom
ΔFA -TEBV	2.26	0.005	0.007
ΔFA -GEBV	3.08	0.005	0.021
ΔFG -GEBV	1.91	0.004	0.005

Sonesson et al., GSE, 2012

Conclusions ΔF manag.:

- Traditional selection acts on Mendelian sampling terms (MST)
- ΔF management: constrain $\text{Var}(\text{MST})$
 - OC acts on pedigree inbreeding
- Genomic selection acts on SNPs
- ΔF management should constrain $\text{Var}(\Delta q)$
i.e. variance of freq. change of SNPs
 - OC acts on G matrix based on SNPs
- If not OC finds ways to increase ΔF that are undetected by the A matrix

Example: ΔF genomic \gg ΔF pedigree

- ΔF A: 2 sibs / family is still OK
- What ΔF A does not see:
 - GBLUP selects sibs that inherited region X IBD from their sire
 - In other families, also this region X is selected
 - Generation after generation the frequency of a haplotype at region X is increased
 - ΔF A assumes relationships over generations is 0.5
 - But GBLUP focusses on the same haplo generation after generation

Genomics to increase between breed diversity

Rescue-ing a breed



Problem description

- Local breed that is no longer fashionable
- Semen of large commercial breed introduced
 - Assume Holstein semen => Holsteinisation
- Breed will be lost (will become Holstein)
 - Diversity will be lost
- How to rescue the breed ?
 - Using genomics
 - Assuming we can manage the selections in part of the breed

OC type of approach

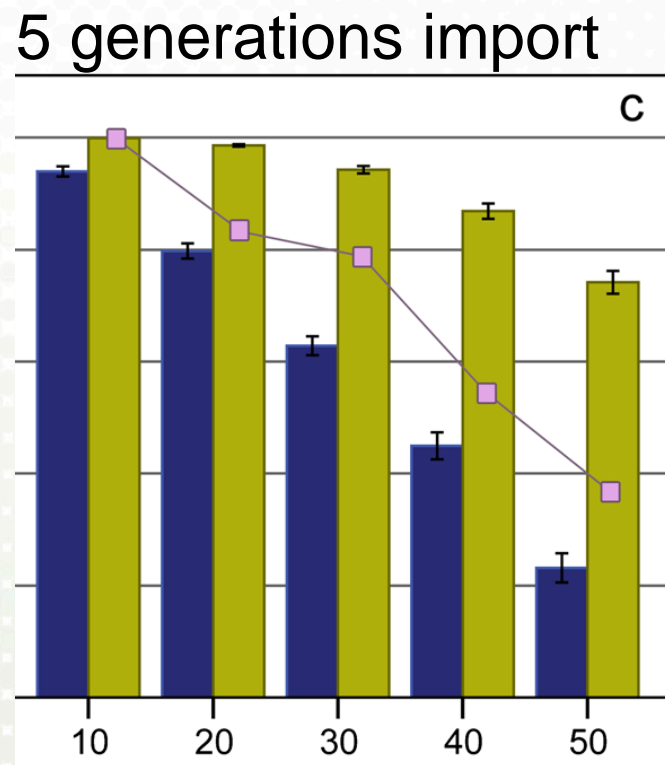
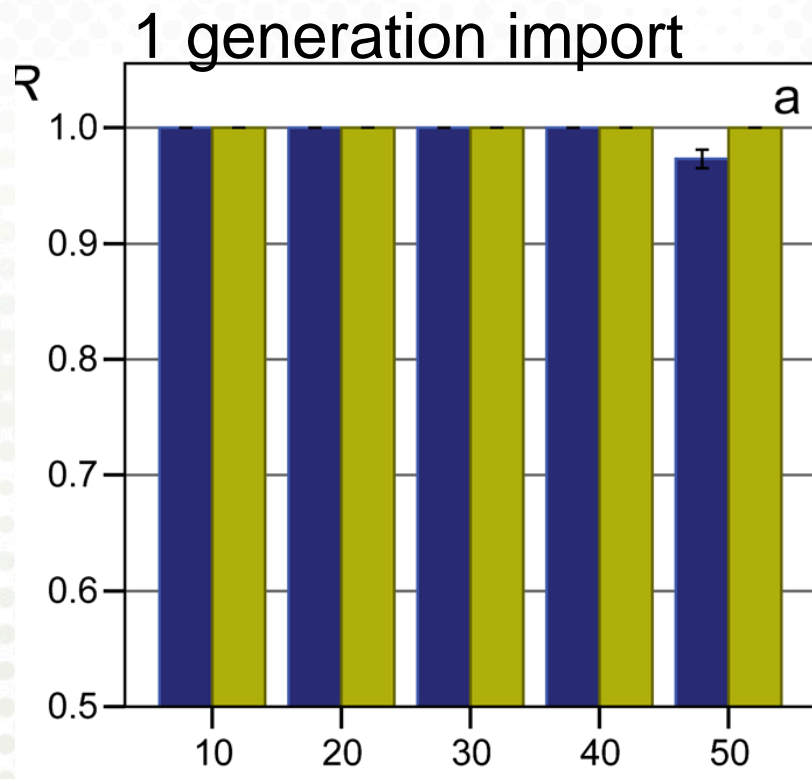
- Minimise:

$$\sum c_i \bar{G}_{i,hol}$$

- c_i is the optimal contribution of the animal
- $\bar{G}_{i,hol}$ is the average genomic relationship of i with the introduced holstein bulls

Amador et al. GSE 2013

Recovery of local breed genome



- 1 generation of OC-G management
- 10 generations of OC-G management
- 1 generation of OC-A management

Amador et al. 2013

Conclusions: recovery of local genetics

- OC_goal: minimise relat. with Holsteins
- Recovery was almost 100%
 - If introgression was not high ($<30\%$)
 - Not extending over many generations (≤ 3)
- Genomic relationships were effective tool
 - When used over >1 generation of management
 - Pedigree was effective when only 1 generation
- Inbreeding was increased due to recovery
 - Focusses on limited part of genome

Overall Conclusions

- PT \Rightarrow GS scheme may increase ΔF
 - Counteract by selection more elite sires
 - Or control the genomic inbreeding
- Use of GS requires genomic control of ΔF
 - GEBV increase freqs of same haplos over generat's
 - ΔFA control: ΔFG 4-fold too high
- Genomics to enhance diversity
 - Example: recovery from recent introgressions
 - Genomic OC was often very successful