Increasing long term response by selecting for favorable minor alleles

Chuanyu Sun

Paul VanRaden

National Association of Animal breeders, USA
Animal Improvement Programs Laboratory, USA

Introduction

- **Genomic selection has become a standard tool in dairy cattle breeding**

- **Response to genomic selection can continue for many generations or decline rapidly, depending on the number of QTLs, their frequencies, linkage with markers, and effects on the trait or index selected, etc.**
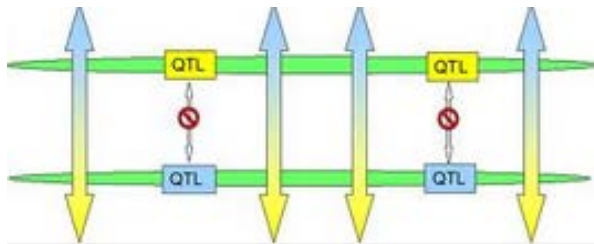
Introduction

- Based on simulations (Muir, 2007) or deterministic predictions (Goddard 2009)

  Long-term gains from genomic selection can be less than from phenotypic selection or from selection on pedigree and phenotypes

- But long term response can be improved by modifying the selection pressure applied to a QTL as its allele frequency changes, as demonstrated

  For 1 QTL in combination with phenotypic selection (Dekkers et al., 1998)

  For multiple QTL using index selection (Jannink, 2010; Goddard 2009)

Introduction

- The weight for each marker or QTL is adjusted according to its current frequency, giving markers where the ***favorable allele has low frequency more weight*** in the index.

- Such methods can improve long term response and will be referred to as ***favorable minor allele (FMA) selection***.

Introduction

Objective

◆ **Propose simple and improved formulas for**

    ✓ **weighting favorable minor alleles**

    ✓ **increasing the long term progress from genomic selection**

    ✓ **with less reduction of short term progress.**

◆ **These formulas are applied to both simulated and real data**

Materials and Methods

- **With standard genomic selection, estimated breeding values were calculated using**

$$\hat{u}_i = \sum_j \hat{\beta}_j z_{ij}$$

- **With FMA selection**

$$\hat{u}_i = \sum_j \beta'_j z_{ij}$$

Materials and Methods

- Two new formulas to implement FMA selection were derived

$$\beta'_j = \begin{cases} \hat{\beta}_j * (1 + (0.5 f_j^{-0.5} - 1) * \delta) & if\,(\hat{\beta}_j \geq 0) \\ \hat{\beta}_j * (1 + (0.5 * (1 - f_j)^{-0.5} - 1) * \delta) & if\,(\hat{\beta}_j < 0) \end{cases}$$

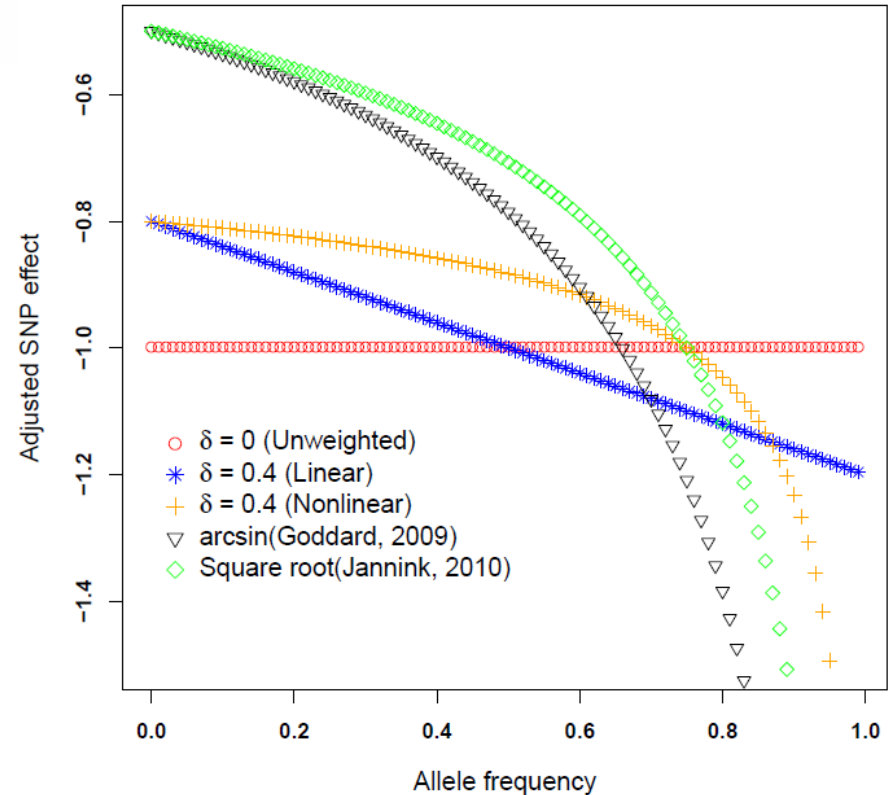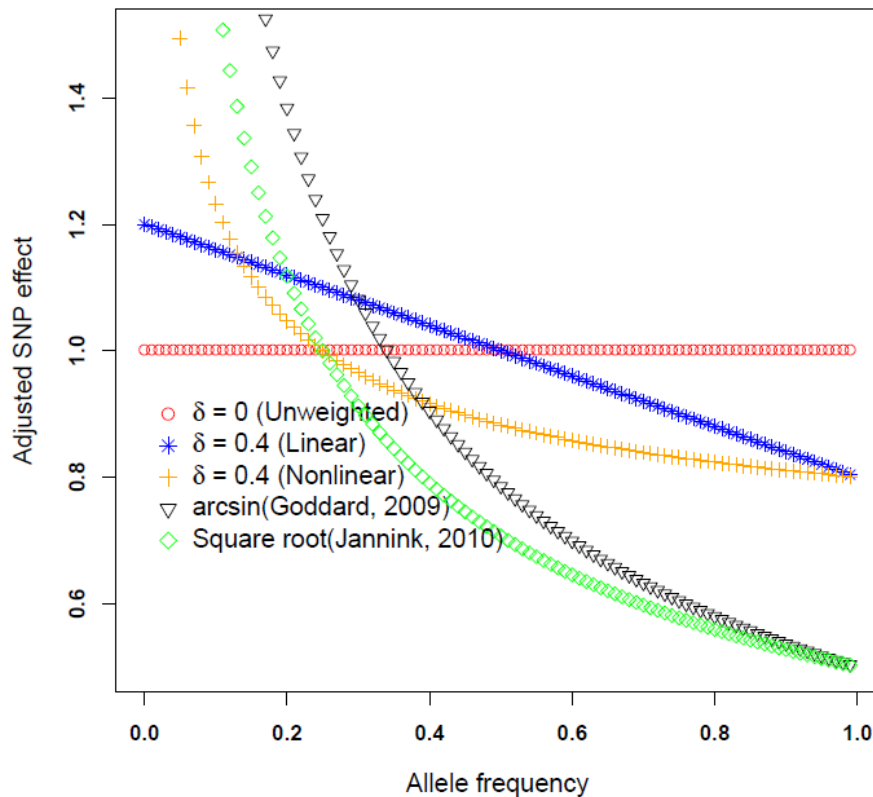$$when \quad 0 < f_j < 1 ; \quad otherwise \quad \beta'_j = \hat{\beta}_j$$

$$f_j \text{ is allele frequency for } j^{th} \text{ SNP}$$

Identical to formula of Jannink (2010) if $\delta = 1$

$$\beta'_j = \begin{cases} \hat{\beta}_j * (1 + (0.5 - f_j) * \delta) & if\,(\hat{\beta}_j \geq 0) \\ \hat{\beta}_j * (1 + (f_j - 0.5) * \delta) & if\,(\hat{\beta}_j < 0) \end{cases}$$

Materials and Methods

◆ **Both of them include parameter *δ* that can vary from 0 to 1 to balance long and short term progress.**

◆ **For both the nonlinear and linear formulas, *δ* = 0 corresponds to official, unweighted genomic selection.**

Materials and Methods

Simulation

➢ 20 generations

➢ Genotypes and BV were simulated using real pedigree for first generation

➢ In each subsequent generation, the top 100 males and top 1,000 females were selected and mated to produce 1,500 males and 1,500 females for the next generation

| QTLs	QTL distribution	Number of replicates	Largest QTL variance[1]
3000	Normal	100	0.599 (0.43 –1.1)
	Heavy tail	100	4.76 (1.89 – 24.14)

Materials and Methods

- Selection acted directly on 3,000 QTL effects (100 per chromosome) instead of indirectly on estimated marker effects.

- Responses to 20 generations of selection were tested using linear and nonlinear weighting formulas with δ parameters ranging from 0 to 1

QTL distribution	Number of replicates	Method	δ				
Normal	100	Linear	0	0.2	0.4	0.6	–
		Non linear	–	0.2	0.4	0.6	1.0
Heavy tail	100	Linear	0	0.2	0.4	0.6	–
		Non linear	–	0.2	0.4	0.6	1.0

The best parameter value for *δ were used to real data (HO, JE, BS)*

Results - response

The ratio of genetic progress based on adjusted genomic breeding values using Jannink (2010) formula over unadjusted genetic progress.
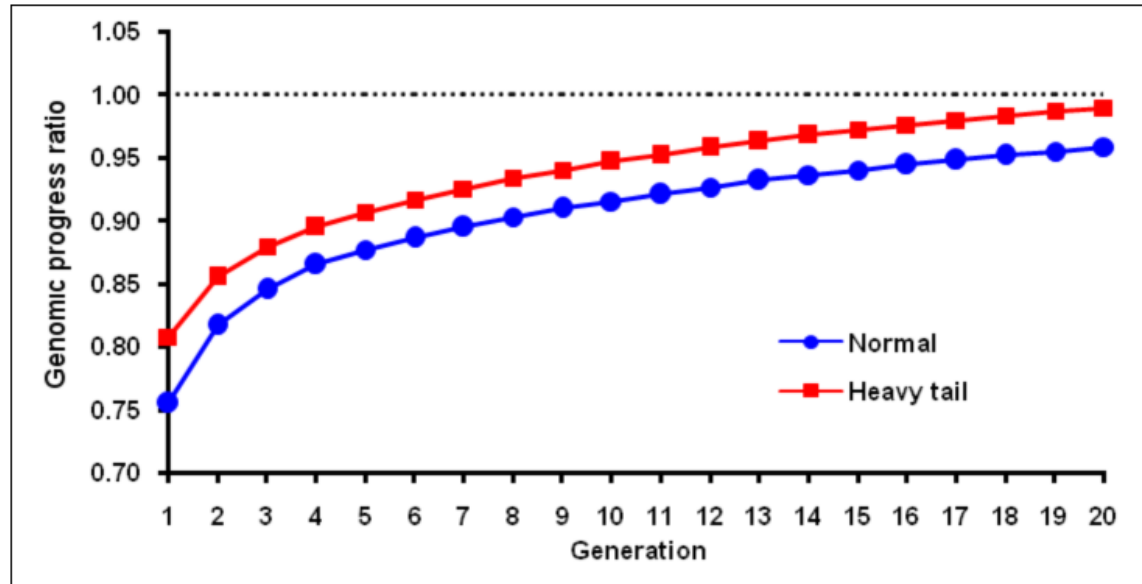

Large losses in early generations and did not recover these losses within 20 generations

Therefore the remaining results focused on optimizing δ to balance long term and short term progress

Results - response

1. *Normal distribution*
2. maximum response was achieved when $\delta = 0.2$ (or 0.1) using FMA selection
3. losses were larger in the first few generations with non linear formula
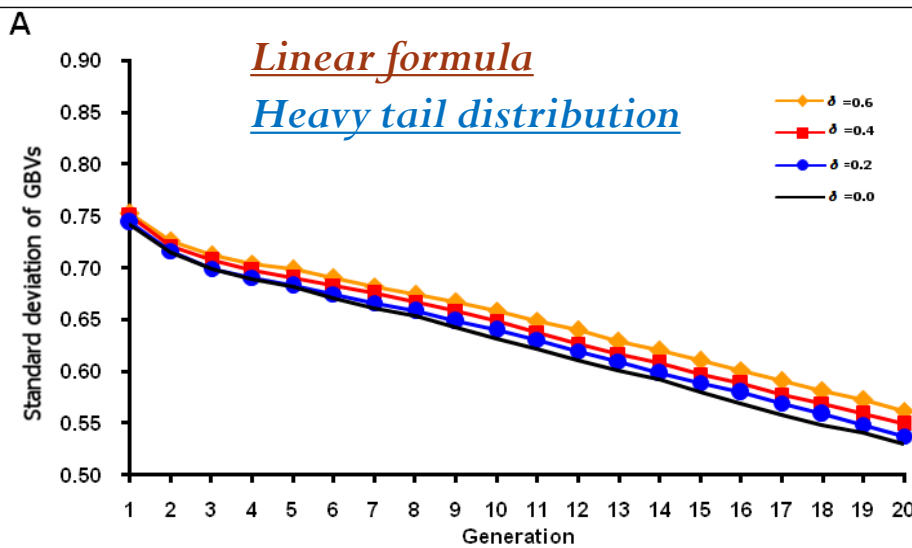

A — *Linear formula*
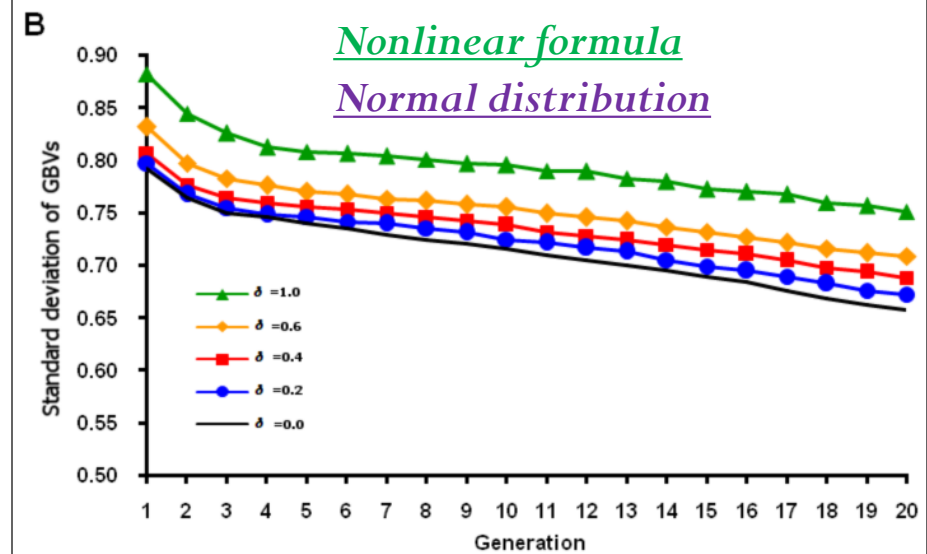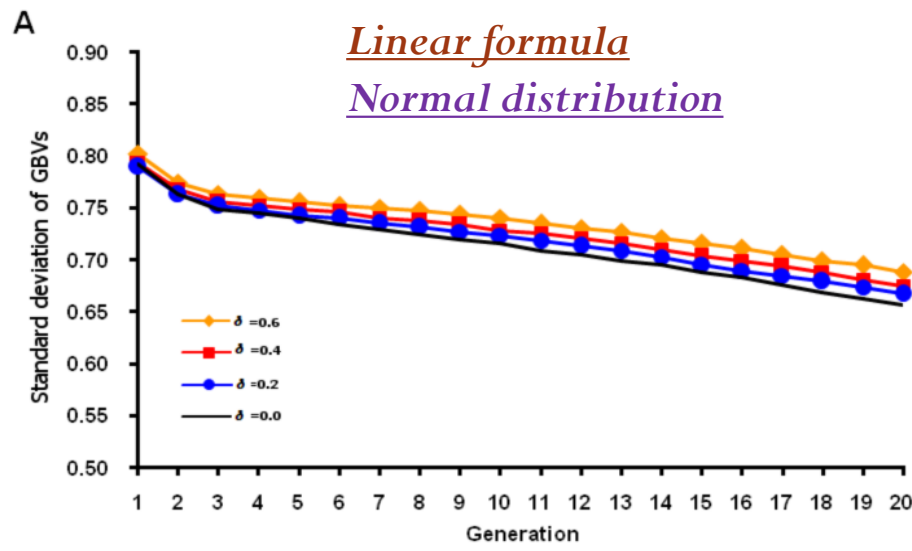B — *Nonlinear formula*
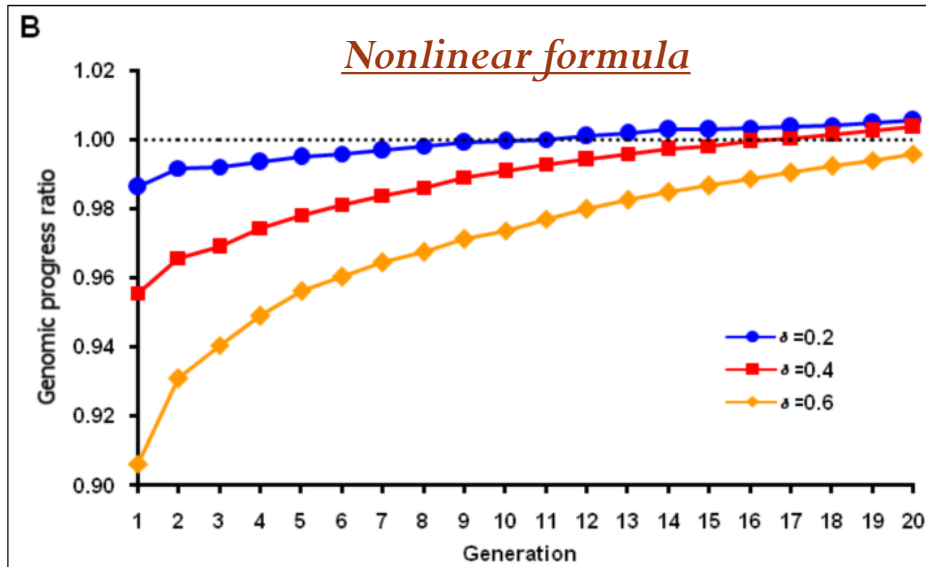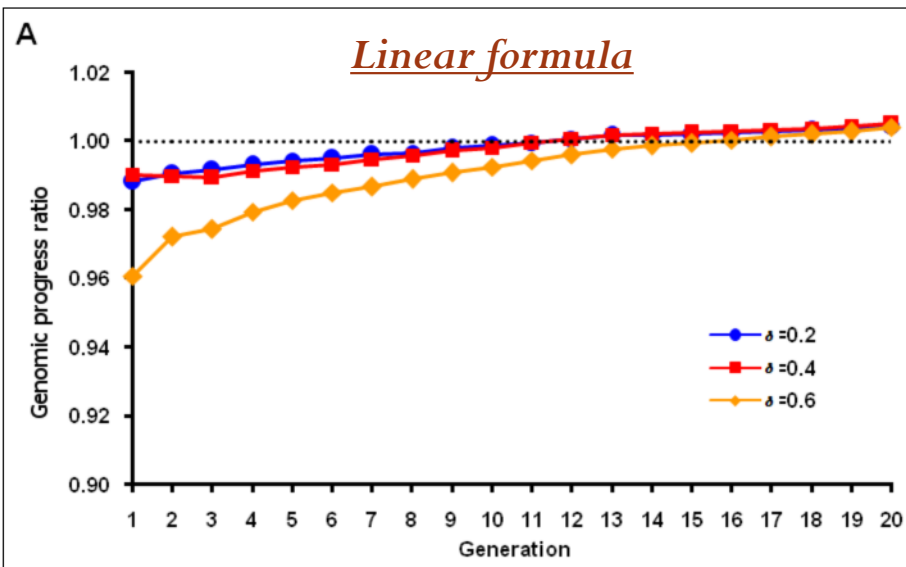
Results - response

1. *Heavy tail distribution*
2. maximum response by generation 20 was achieved with $\delta = 0.4$ or $\delta = 0.6$ using FMA selection
3. More losses in the first few generations with $\delta = 0.6$
4. More losses in the first few generations with non linear formula

A — *Linear formula*

B — *Nonlinear formula*
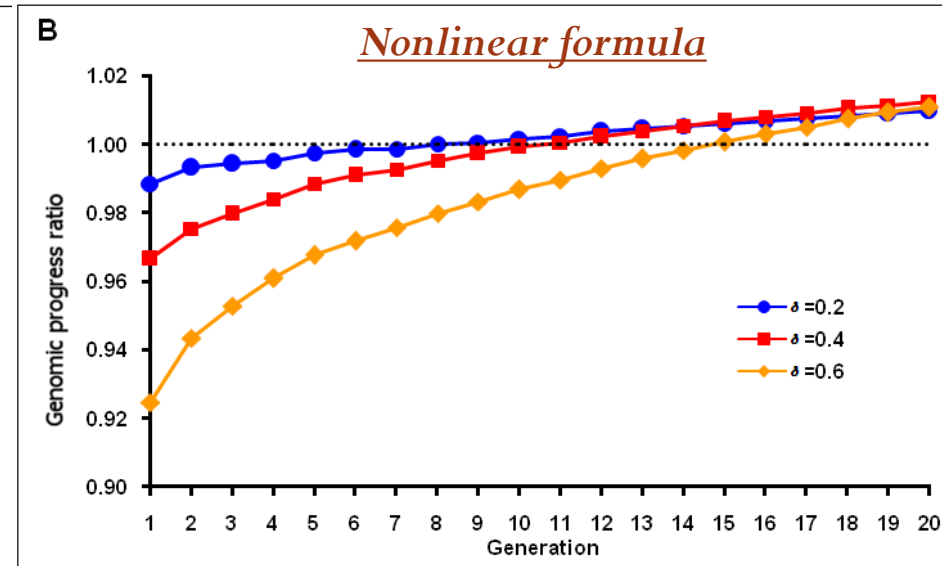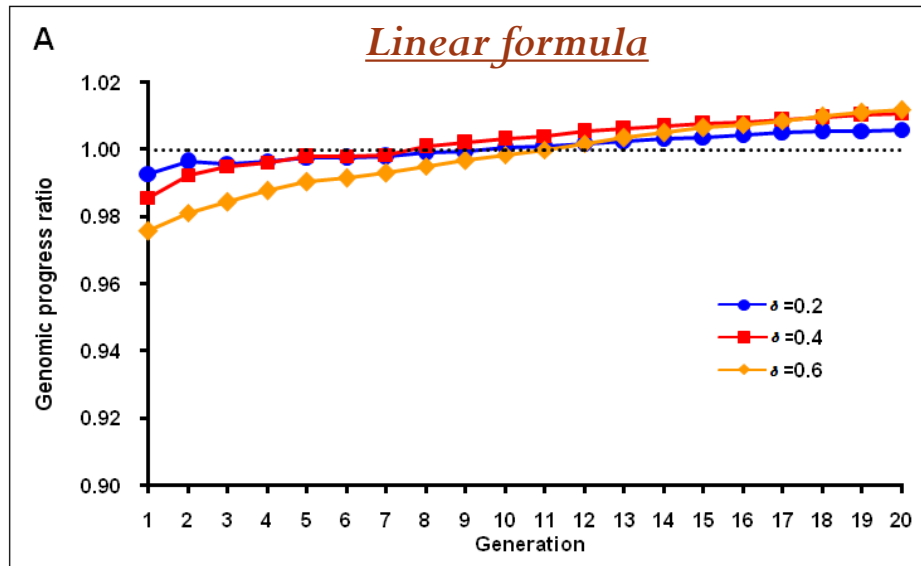
Results –variance
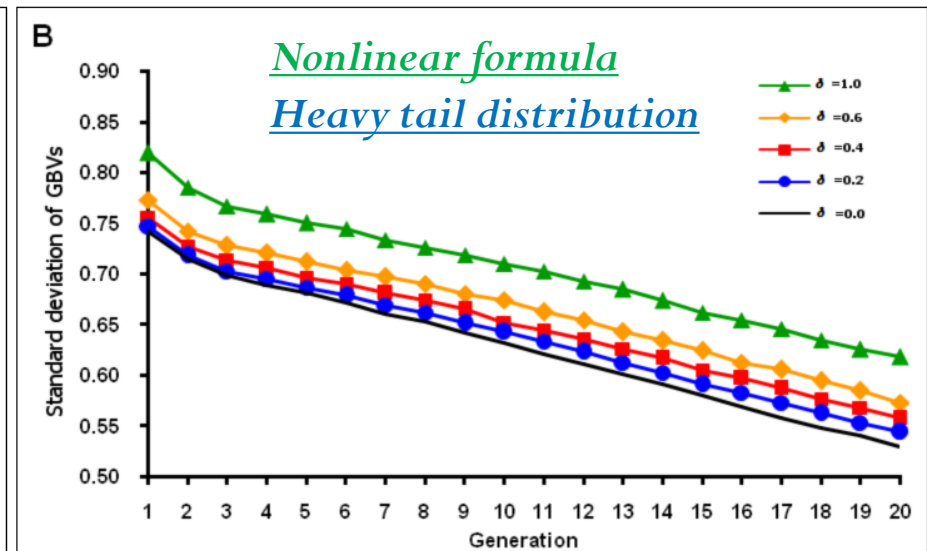
Results – inbreeding

◆ Larger values of δ showed a little higher genomic inbreeding when true allele frequency was used but showed a little lower inbreeding when using 0.5 as the allele frequency for each locus or using pedigree inbreeding.

◆ Genomic selection can increase genomic inbreeding rapidly

Method	δ	Normal distribution			Heavy tail distribution		
		0.500	true	Ped	0.500	true	Ped
Linear	0.0	0.457	0.261	0.095	0.456	0.256	0.093
	0.2	0.447	0.268	0.093	0.448	0.260	0.092
	0.4	0.437	0.272	0.091	0.441	0.264	0.091
	0.6	0.427	0.278	0.089	0.433	0.269	0.089
Nonlinear	0.2	0.443	0.265	0.092	0.445	0.258	0.092
	0.4	0.428	0.269	0.089	0.433	0.262	0.090
	0.6	0.413	0.273	0.087	0.422	0.266	0.088
	1.0	0.384	0.283	0.083	0.399	0.275	0.085

Results – real data

- The linear and nonlinear formulas were both applied with the parameter value for δ set to 0.4 based on the optimum from simulated data

- The correlation between official and FMA evaluation

Method	Holstein	Jersey	Brown Swiss
Linear formula	0.994	0.994	0.989
Nonlinear formula	0.991	0.986	0.978

Results

◆ **Expected future inbreeding (EFI)**

  Half the animal's average pedigree relationship to its breed

◆ **Genomic future inbreeding (GFI)**

  Half the animal's average genomic relationship to its breed

◆ **Correlation of the difference between FMA and official evaluation with GFI and EFI, using linear weighting and δ = 0.4**

Inbreeding measure	Holstein	Jersey	Brown Swiss
GFI	-0.85	-0.94	-0.85
EFI	-0.45	-0.59	-0.27

Conclusions

- *Short term and long term progress were balanced using new formulas for FMA selection*.

- Previous formulas put too much emphasis on rare favorable alleles and resulted in less progress than standard genomic selection over 20 simulated generations.

- *Optimum values of δ differed depending on QTL distribution, with lower values favored if QTL effects are small because allele frequencies will change more slowly with selection*.

- FMA selection using linear formula could increase long term response and less losses in the first few generations, and can be used for routine evaluation.

- *The differences between FMA and standard genomic selection were highly correlated to the animal's GFI. Thus, strategies to reduce genomic inbreeding could achieve almost as much long term progress as FMA selection*.

THANK YOU
FOR YOUR ATTENTION

QUESTIONS?