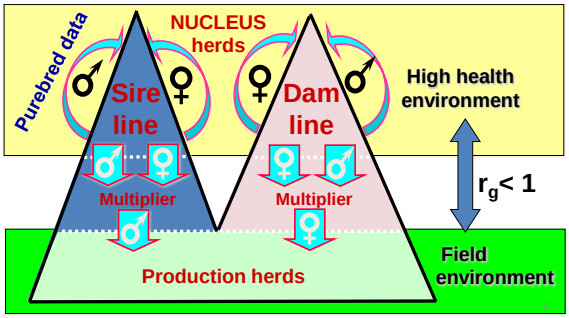
Genomic Selection in Crossbreeding Programs

Jack Dekkers

1

Current Pyramid Selection Programs

Limitations: - limited selection for performance in the field
- no selection for traits not recorded in nucleus - disease
- no selection for heterosis

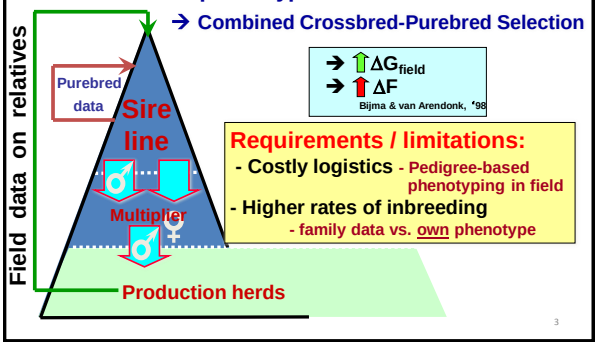


Selection for Performance in Field

'Traditional' Breeding Solution:

Collect phenotypes on relatives in field

→ Combined Crossbred-Purebred Selection



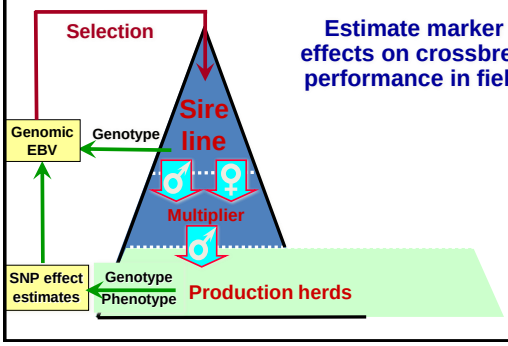
→ ⇧ΔG_{field}
→ ⇧ΔF
Bijma & van Arendonk, '98

Requirements / limitations:
- Costly logistics - Pedigree-based phenotyping in field
- Higher rates of inbreeding
  - family data vs. own phenotype

3

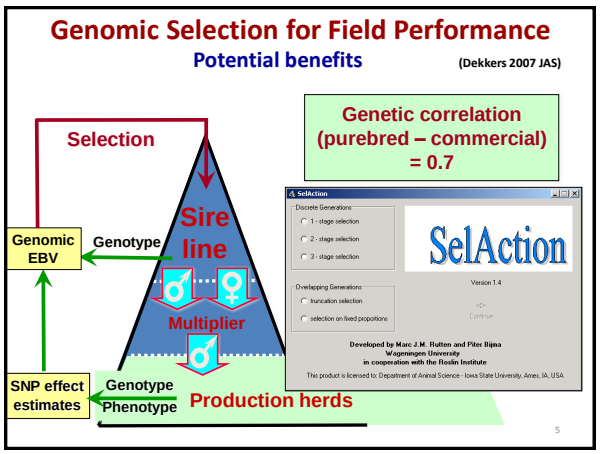
Selection for Performance in Field

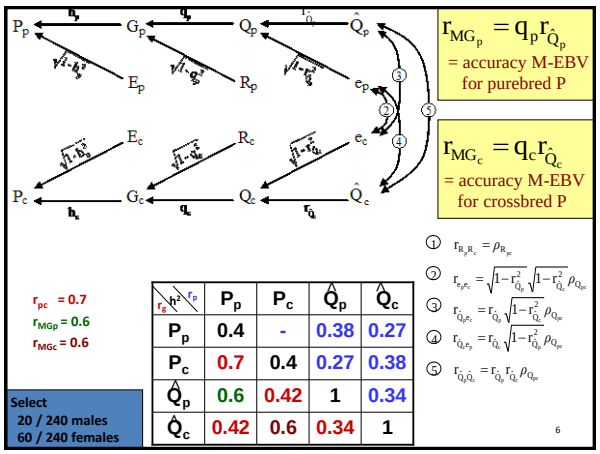
Possible Genomic Selection Solution:



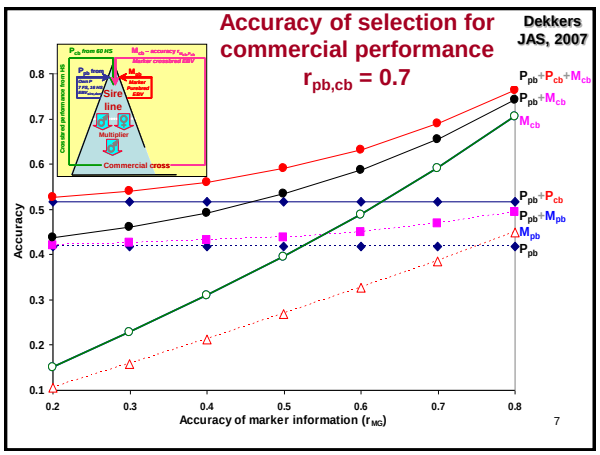
Estimate marker effects on crossbred performance in field

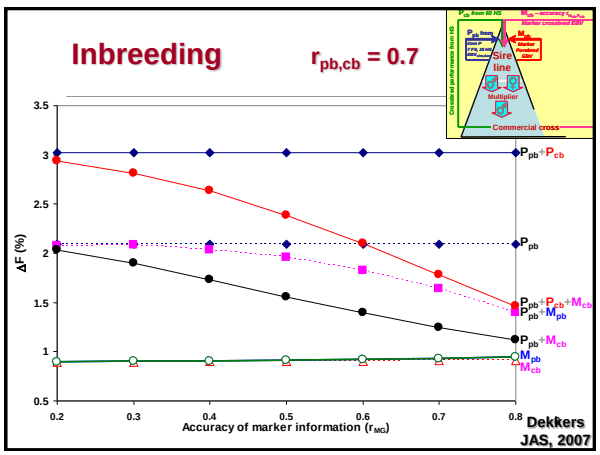
4





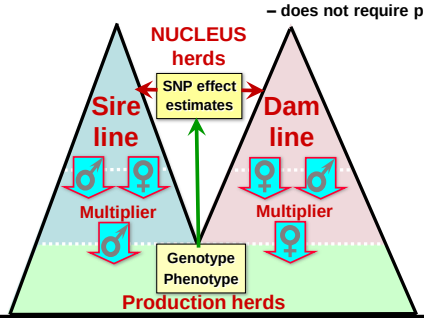
	P_p	P_c	$\hat{Q}_p$	$\hat{Q}_c$
P_p	0.4	-	0.38	0.27
P_c	0.7	0.4	0.27	0.38
$\hat{Q}_p$	0.6	0.42	1	0.34
$\hat{Q}_c$	0.42	0.6	0.34	1





GS training on Crossbred data

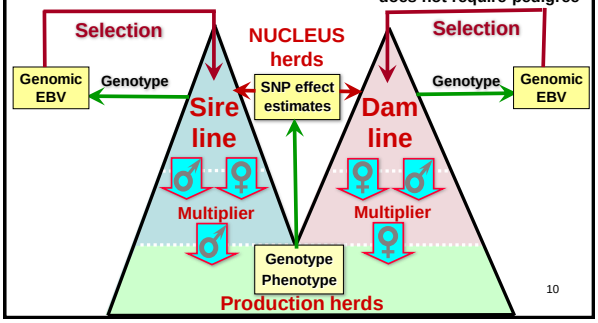
1. Genotype parents and train on crossbred progeny performance – requires pedigree
2. Genotype crossbreds and train on own phenotype – does not require pedigree



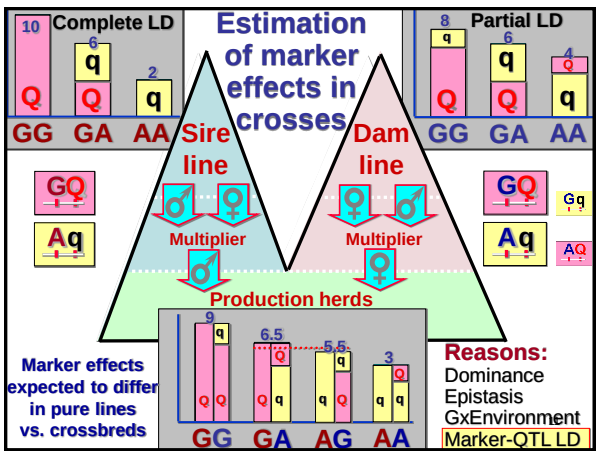
9

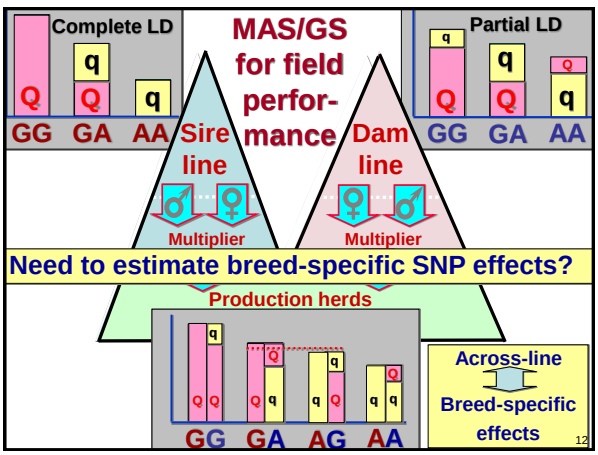
GS training on Crossbred data

1. Genotype parents and train on crossbred progeny performance – requires pedigree
2. Genotype crossbreds and train on own phenotype – does not require pedigree



10

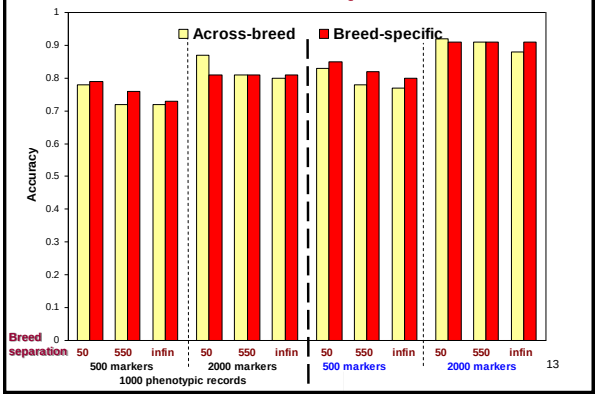




12

Noelia Ibanez-Escriche, Rohan Fernando, Ali Toosi, Jack Dekkers

GSE 2009 Additive QTL model



13

Impact of Dominance

p_S / q_S	Sire Breed BB	Bb	bb
Freq.	p_S^2	$2p_Sq_S$	q_S^2
Value	a	d	-a

p_D / q_D	Dam Breed BB	Bb	bb
Freq.	p_D^2	$2p_Dq_D$	q_D^2
Value	a	d	-a

Purebred allele substitution effects

$\alpha_{S,p} = a + (q_S - p_S)d$ $\alpha_{D,p} = a + (q_D - p_D)d$

Crossbred	Dam allele B (p_D)	b (q_D)
S allele B (p_S)	a	d
S allele b (q_S)	d	-a

Sire allele substitution effect in crossbred

$\alpha_{S,c} = (p_Da+q_Dd)-(p_Dd-q_Da) = a + (q_D-p_D)d$

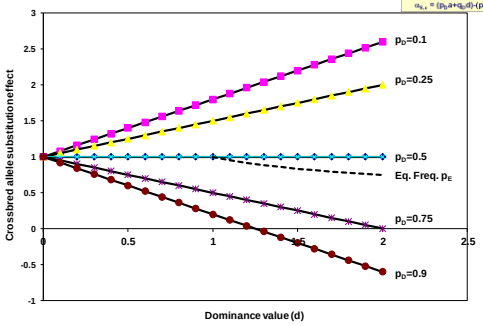
➔ Allele substitution effect in one parental breed depends on allele freq.s in other parental breed

➔ Breed-specific allele effects

14

Allele substitution effect in sire breed for varying dam breed frequencies p_D and degrees of dominance

a = 1



15

Impact of Dominance

p_S / q_S	Sire Breed BB	Bb	bb
Freq.	p_S^2	$2p_Sq_S$	q_S^2
Value	a	d	-a

p_D / q_D	Dam Breed BB	Bb	bb
Freq.	p_D^2	$2p_Dq_D$	q_D^2
Value	a	d	-a

Purebred allele substitution effects

$\alpha_{S,p} = a + (q_S - p_S)d$ $\alpha_{D,p} = a + (q_D - p_D)d$

Crossbred	Dam allele B (p_D)	b (q_D)
S allele B (p_S)	a	d
S allele b (q_S)	d	-a

Sire allele substitution effect in crossbred

$\alpha_{S,c} = (p_Da+q_Dd)-(p_Dd-q_Da) = a + (q_D-p_D)d$

➔ Allele substitution effect in one parental breed depends on allele freq.s in other parental breed

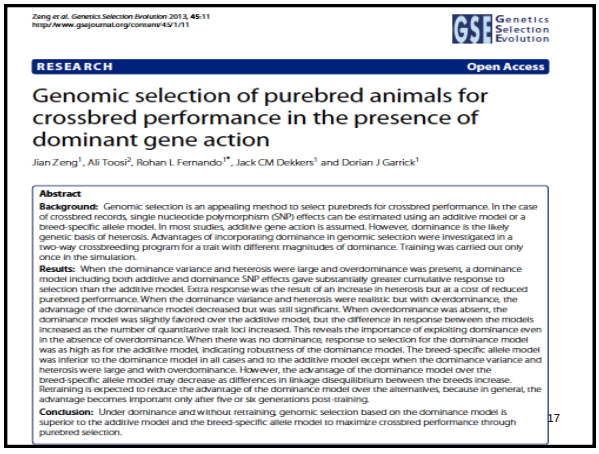
➔ Breed-specific allele effects

➔ Effects change over time

➔ Estimate 'a' and 'd' *vs* allele substitution effects

16

Zeng et al. Genetics Selection Evolution 2013, 45:11
http://www.gsejournal.org/content/45/1/11



RESEARCH Open Access

Genomic selection of purebred animals for crossbred performance in the presence of dominant gene action

Jian Zeng[1], Ali Toosi[2], Rohan L Fernando[1*], Jack CM Dekkers[1] and Dorian J Garrick[1]

Abstract

Background: Genomic selection is an appealing method to select purebreds for crossbred performance. In the case of crossbred records, single nucleotide polymorphism (SNP) effects can be estimated using an additive model or a breed-specific allele model. In most studies, additive gene action is assumed. However, dominance is the likely genetic basis of heterosis. Advantages of incorporating dominance in genomic selection were investigated in a two-way crossbreeding program for a trait with different magnitudes of dominance. Training was carried out only once in the simulation.

Results: When the dominance variance and heterosis were large and overdominance was present, a dominance model including both additive and dominance SNP effects gave substantially greater cumulative response to selection than the additive model. Extra response was the result of an increase in heterosis but at a cost of reduced purebred performance. When the dominance variance and heterosis were realistic but with overdominance, the advantage of the dominance model decreased but was still significant. When overdominance was absent, the dominance model was slightly favored over the additive model, but the difference in response between the models increased as the number of quantitative trait loci increased. This reveals the importance of exploiting dominance even in the absence of overdominance. When there was no dominance, response to selection for the dominance model was as high as for the additive model, indicating robustness of the dominance model. The breed-specific allele model was inferior to the dominance model in all cases and to the additive model except when the dominance variance and heterosis were large and with overdominance. However, the advantage of the dominance model over the breed-specific allele model may decrease as differences in linkage disequilibrium between the breeds increase. Retraining is expected to reduce the advantage of the dominance model over the alternatives, because in general, the advantage becomes important only after five or six generations post-training.

Conclusion: Under dominance and without retraining, genomic selection based on the dominance model is superior to the additive model and the breed-specific allele model to maximize crossbred performance through purebred selection.

17

Breed-Specific SNP Allele Mode (BSAM)

$$y_i = \mu + \sum_{j=1}^{k}(X_{ij}^A \alpha_j^A + X_{ij}^B \alpha_j^B) + e_i$$

- X_{ij}^A or $X_{ij}^B \in \{0,1\}$ is the copy number of a given allele at SNP j of breed origin A or breed origin B
- α_j^S or α_j^D is the breed-specific substitution effect for the allele of breed origin A or B
- The SNP allele effects have breed-specific variance $\sigma^2_{\alpha^A}$ and $\sigma^2_{\alpha^B}$, and breed-specific π parameter π_{α^A} and π_{α^B}

18

Dominance Model

$$y_i = \mu + \sum_{j=1}^{k}(X_{ij}a_j + W_{ij}d_j) + e_i$$

- $W_{ij} \in \{0,1\}$ is the indicator for heterozygous genotype
- a_j is the additive effect and d_j the dominance effect
- $d_j|\mu_d, \sigma_d^2 = \begin{cases} 0 & \text{with probability } \pi_d \\ \sim N(\mu_d, \sigma_d^2) & \text{with probability } 1-\pi_d \end{cases}$
- $\mu_d|\sigma_d^2 \sim N(\eta, \sigma_d^2/\phi)$, where $\eta = \frac{H}{\sum_j (p_j^S - p_j^D)^2}$ and ϕ measures the strength of the prior belief in terms of σ_d^2

19

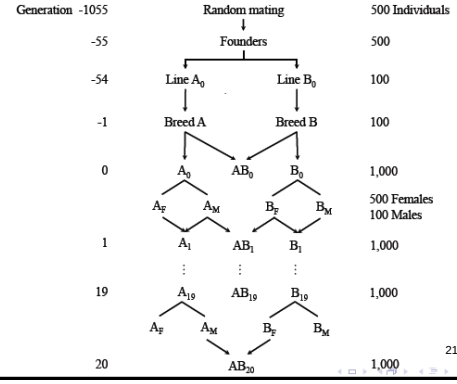
Design of Study

Scenario	Gene Action	h^2_{BS}	h^2_{NS}	Heterosis
1	Overdominance	0.5	0.33	31%
2	Overdominance	0.5	0.4	9.1%
3	Incomplete Dominance	0.5	0.35	8.4%
4	Additive	0.5	0.5	0

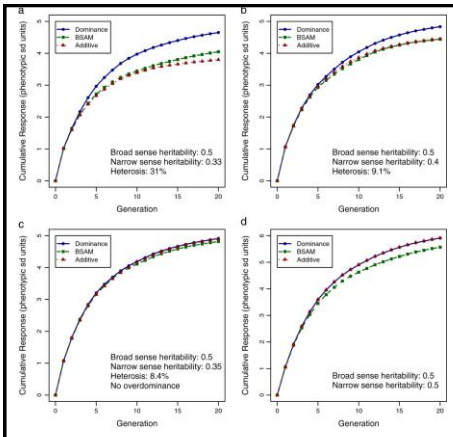
- Other parameters, such as LD, allele frequencies, QTL positions, etc, were kept constant across scenarios
- Training was carried out only once

20

Breed Formation and Selection Program



21



Cumulative response to genomic selection one chromosome, 100 QTL and 1000 SNPs

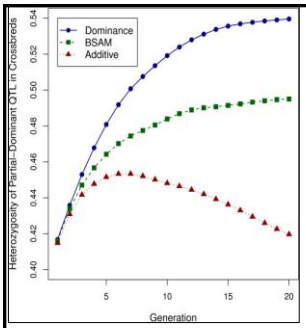
(a) large V_D with overdominance

(**b**) Realistic V_D with overdominance

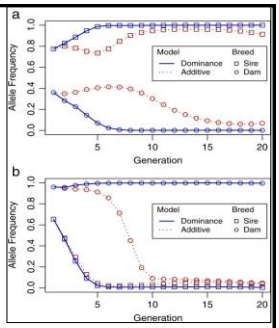
(**c**) realistic V_D without overdominance

(**d**) No dominance

22



Changes in heterozygosity for over-dominant QTL in crossbreds over generations.
Plotted values are changes in heterozygosity for over-dominant QTL in crossbreds over generations of selection, under the dominance model, BSAM and the additive model, with one chromosome and large dominance



Changes in allele frequencies of two over-dominant QTL in the parental breeds, with one chromosome and large dominance

(**a**) shows alternate alleles approaching fixation in the sire and dam breeds more rapidly with the dominance than the additive model

(**b**) shows the same allele approaching fixation in both parental breeds with the additive model, in contrast to the dominance model

23

Conclusions

Genomic Selection provides great opportunities for genetic improvement of commercial crossbred performance

- **Can substantially increase rates of response for commercial crossbred performance**
- **Reduces rate of inbreeding - Compared to CCPS**
- **Requires statistical methods for estimation of marker-effects in crossbred populations**
  - **Breed-specific allele models**
  - **Dominance models**