Genomic prediction of crossbred performance based on purebred Landrace and Yorkshire

Hadi Esfandyari

P Bijma, M Henryon, O F Christensen, A C Sørensen

AARHUS UNIVERSITY

EGS ABG

European Commission ERASMUS MUNDUS

WAGENINGEN UR
For quality of life

Background

- Crossbreeding
- Breeding goal → Crossbred Performance

- Limited accuracy
    - Selection is based on purebred info
    - PB are housed in high-health environments → G×E
    - Non-additive effects

$r_{pc} < 1$

Background

- Genomic selection

- Training on crossbreds
    - SNP effects may differ between PB and CB
    - Collection of crossbred P and G

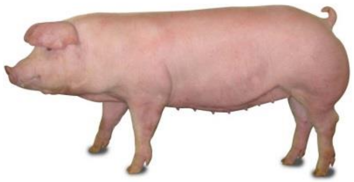
- Training on pure lines
    - $r_{pc} < 1$ is partially due to dominance

Objectives

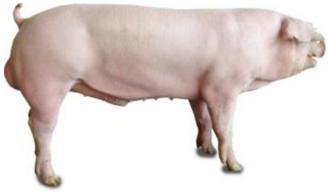
- Compare predictive ability of genomic models
  - Additive (MA)
  - Additive and dominance (MAD)
- Training
  - Separate in both pure lines
  - Combined pure lines

♂ **Landrace**

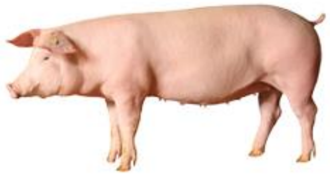
♀ **Yorkshire**



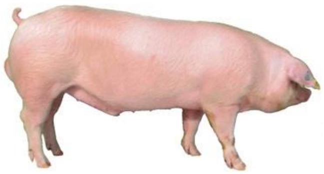
♂ **Yorkshire**



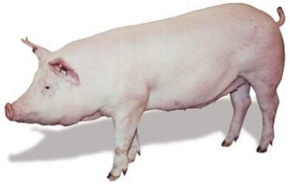
♀ **Landrace**



SEGES



LY



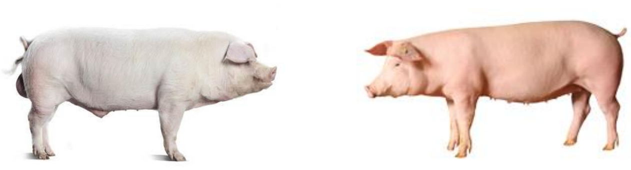
YL

Trait: Litter size

WAGENINGEN UR
For quality of life

AARHUS UNIVERSITY

Landrace



♂ ♀

Full data	**489 000**
Genotyped	**2742 (60K)**
Sow	**2087**
Boar	**655**

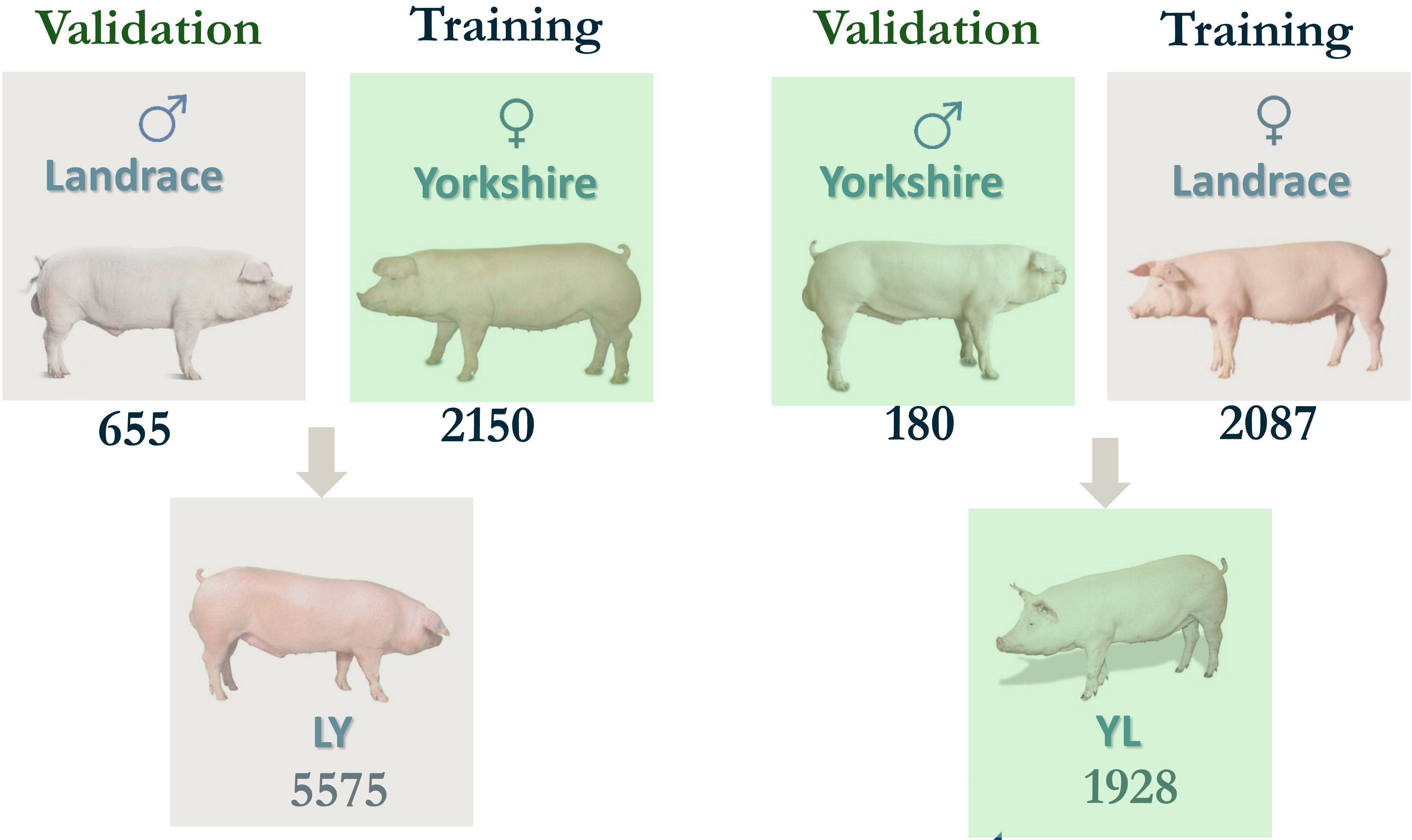
Yorkshire

♀ ♂

Full data	**316 000**
Genotyped	**2330 (60K)**
Sow	**2150**
Boar	**180**

WAGENINGEN UR
For quality of life

AARHUS UNIVERSITY



Validation

♂ Landrace

655

Training

♀ Yorkshire

2150

LY

5575

Validation

♂ Yorkshire

180

Training

♀ Landrace

2087

YL

1928

WAGENINGEN UR
For quality of life

AARHUS UNIVERSITY

Prediction models

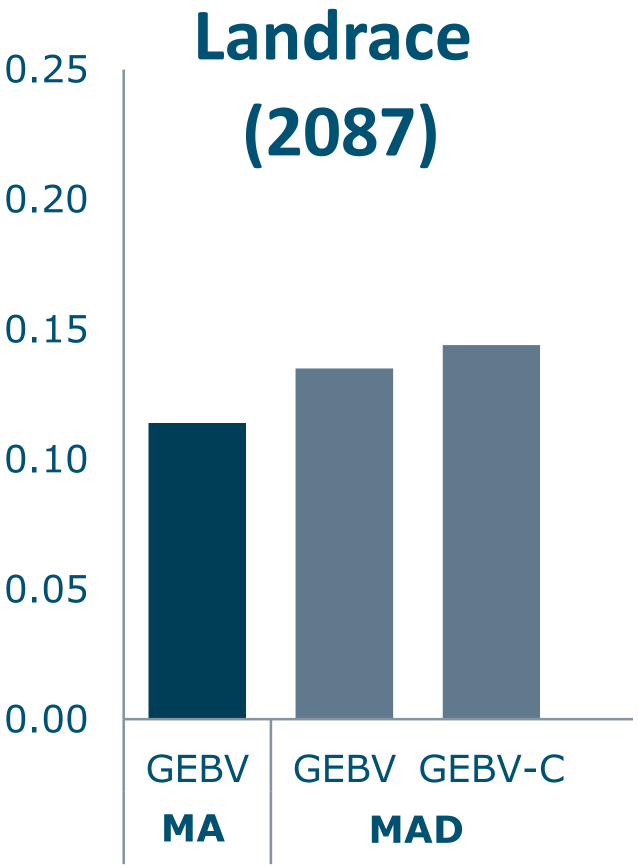
- **Additive model (MA)**
  - **GEBV** → **Expected genotypic value of the offspring of a boar**

- **Additive and dominance model (MAD)**
  - **GEBV**
  - **GEBV-C (SNP allele frequencies from opposite breed)**

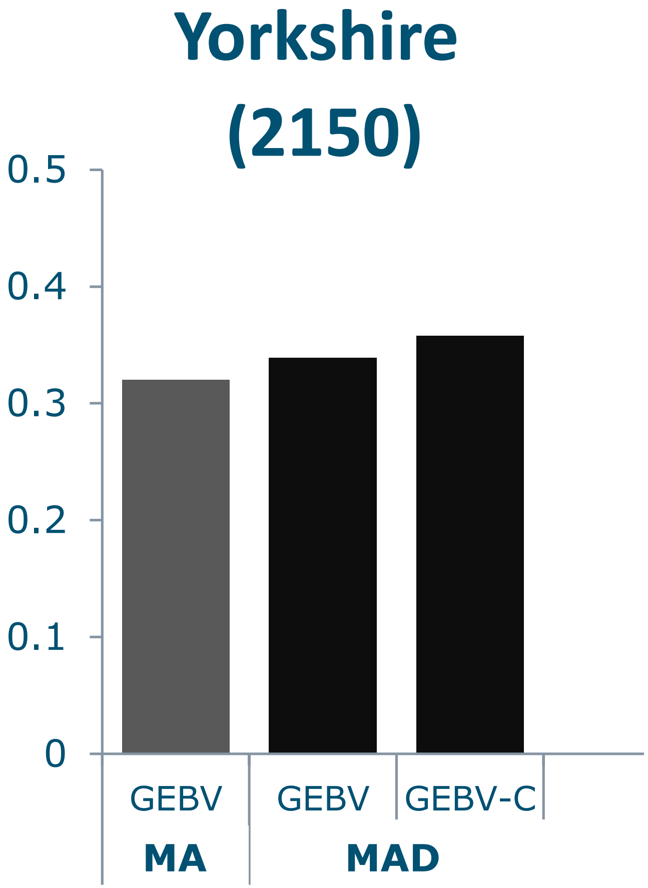
- **Prediction accuracy**

$$cor(GEBV, \overline{P}_{CB})$$

Prediction accuracy (Landrace boars)



Prediction accuracy (Yorkshire boars)



Conclusions

- Prediction accuracy improved by
  - Including dominance
  - Joining two lines

- GEBV for CP based on dominance and allele frequencies was beneficial

AARHUS UNIVERSITY

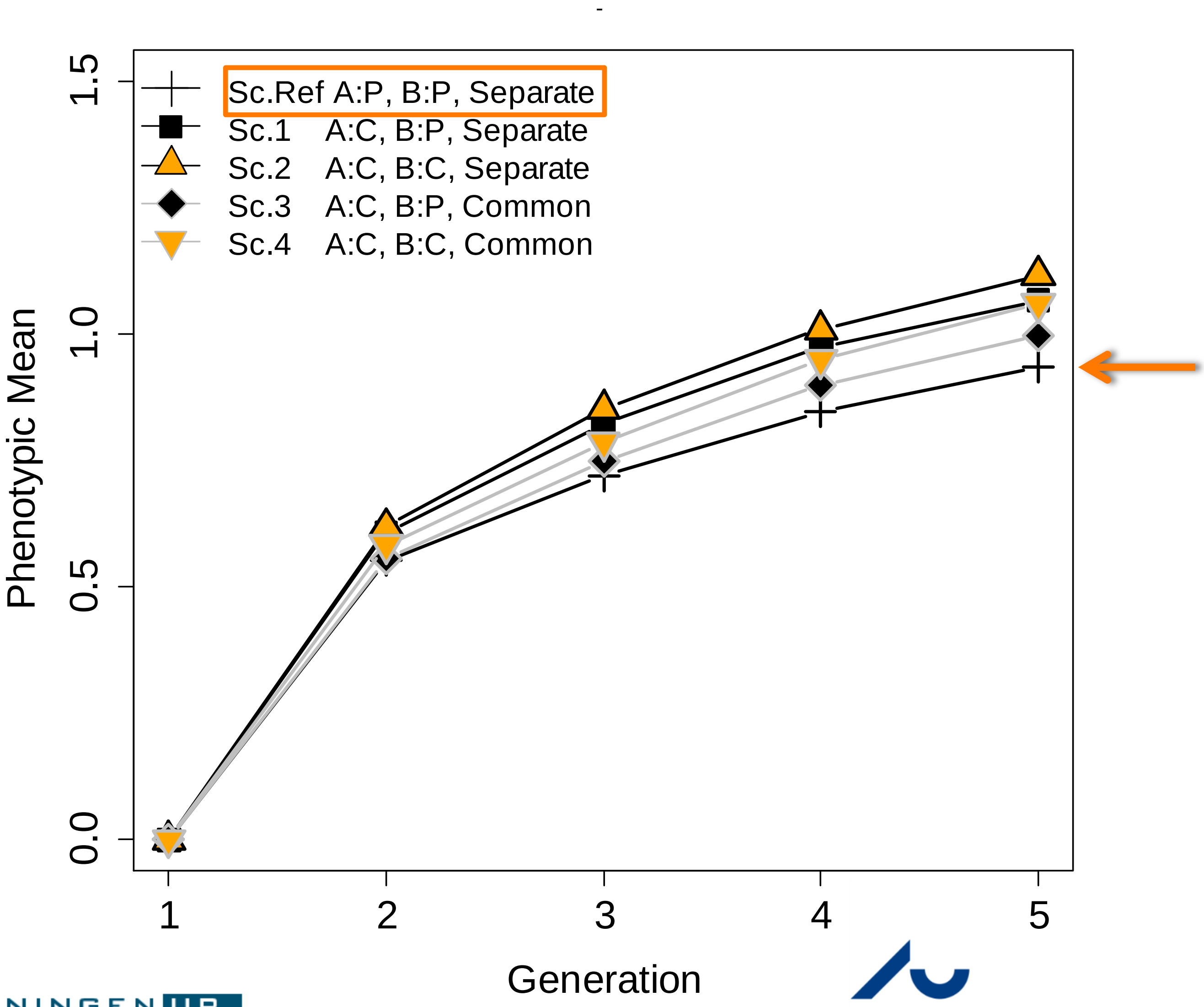
WAGENINGEN UR
For quality of life

European Commission
ERASMUS MUNDUS

EGS ABG

SEGES

GSE 47:16



Sc.Ref A:P, B:P, Separate
Sc.1 A:C, B:P, Separate
Sc.2 A:C, B:C, Separate
Sc.3 A:C, B:P, Common
Sc.4 A:C, B:C, Common
Phenotypic Mean
0.0
0.5
1.0
1.5
Generation
1
2
3
4
5

ACKNOWLEDGEMENTS

EGS
ABG

European Commission
ERASMUS
MUNDUS

LANDBRUG FØDEVARER SEGES

AARHUS UNIVERSITY

WAGENINGEN UR
For quality of life

Conclusions

- Including dominance in genomic prediction models improved prediction accuracy
- GEBV-C based on dominance and allele frequencies was beneficial
- Joining two lines → improved prediction accuracy

Danish data

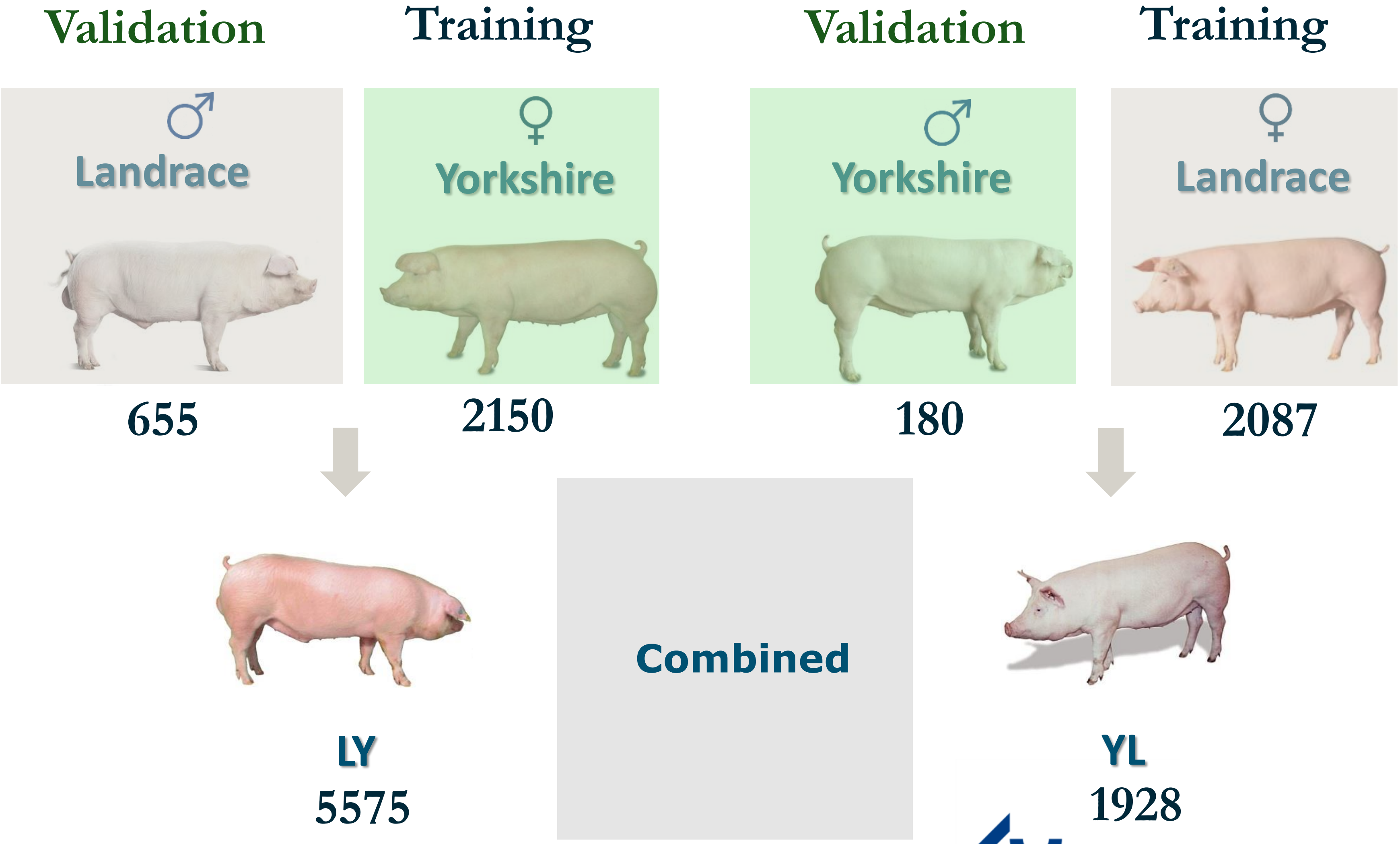
AARHUS UNIVERSITY

WAGENINGEN UR
For quality of life

European Commission
ERASMUS MUNDUS

EGS
ABG

LANDBRUG FØDEVARER
SEGES



Validation

♂ Landrace

655

Training

♀ Yorkshire

2150

Validation

♂ Yorkshire

180

Training

♀ Landrace

2087

LY

5575

Combined

YL

1928

WAGENINGEN UR
For quality of life

AARHUS UNIVERSITY

Prediction models

- **Additive model (MA)**
  - **GEBV**

$$\sum_{j=1}^{s} \textbf{AA} \quad (p_{jr}\hat{a}_j)] + \textbf{Aa} \quad (0.5p_{jr}\hat{a}_j - 0.5q_{jr}\,\hat{a}_j)] + | \; \textbf{aa} \quad -q_{jr}\hat{a}_j)]$$

- **Additive and dominance model (MAD)**
  - **GEBV**
  - **GEBV-C (SNP allele frequencies calculated from the all genotyped sows in opposite breed)**

- **Prediction accuracy** GEBV of the boars and mean phenotype of offspring of the boars

$$cor(GEBV, \overline{P}_{CB})/mean(\sqrt{\frac{n}{n+k}})$$

WAGENINGEN UR
For quality of life

AARHUS UNIVERSITY

Selection criteria

GEBV-P: **Genomic Estimated Breeding value for Purebred Performance**
GEBV-C: **Genomic Estimated Breeding value for Crossbred Performance**

A_1A_2	
Breed A	Breed B
p_Aq_A	p_Bq_B

$A_1A_1 \longrightarrow$ GEBVP: $p_A \times a + q_A \times d$
GEBVC: $p_B \times a + q_B \times d$

Breed A

$A_1A_1 \times \mathbf{A_1A_1}$
$\mathbf{A_1A_2}$
$\mathbf{A_2A_2}$

$\downarrow$

$p_A\,(A_1A_1)$ or $q_A\,(A_1A_2)$

(a) **(d)**

Breed B

$A_1A_1 \times \mathbf{A_1A_1}$
$\mathbf{A_1A_2}$
$\mathbf{A_2A_2}$

$\downarrow$

$p_B\,(A_1A_1)$ or $q_B\,(A_1A_2)$

(a) **(d)**

WAGENINGEN UR
For quality of life

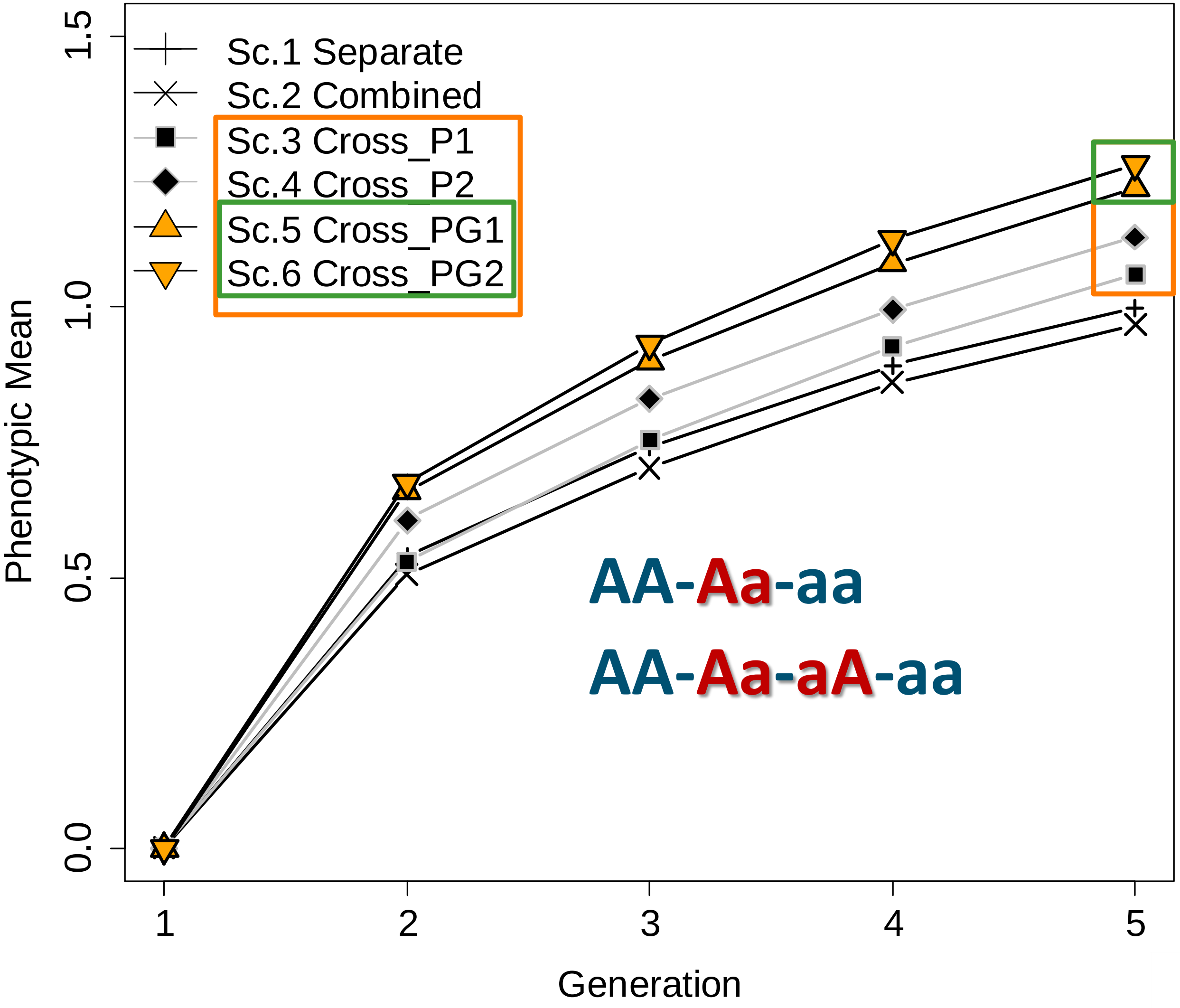
AARHUS UNIVERSITY

Prediction Accuracy

Separate	MA	MAD	
	GEBV	GEBV	GEBV-C
Landrace	0.114	0.125	0.134
Yorkshire	0.321	0.337	0.36

Combined	MA	MAD	
	GEBV	GEBV	GEBV-C
Landrace	0.15	0.16	0.19
Yorkshire	0.36	0.34	0.40

Cumulative response to selection in crossbreds

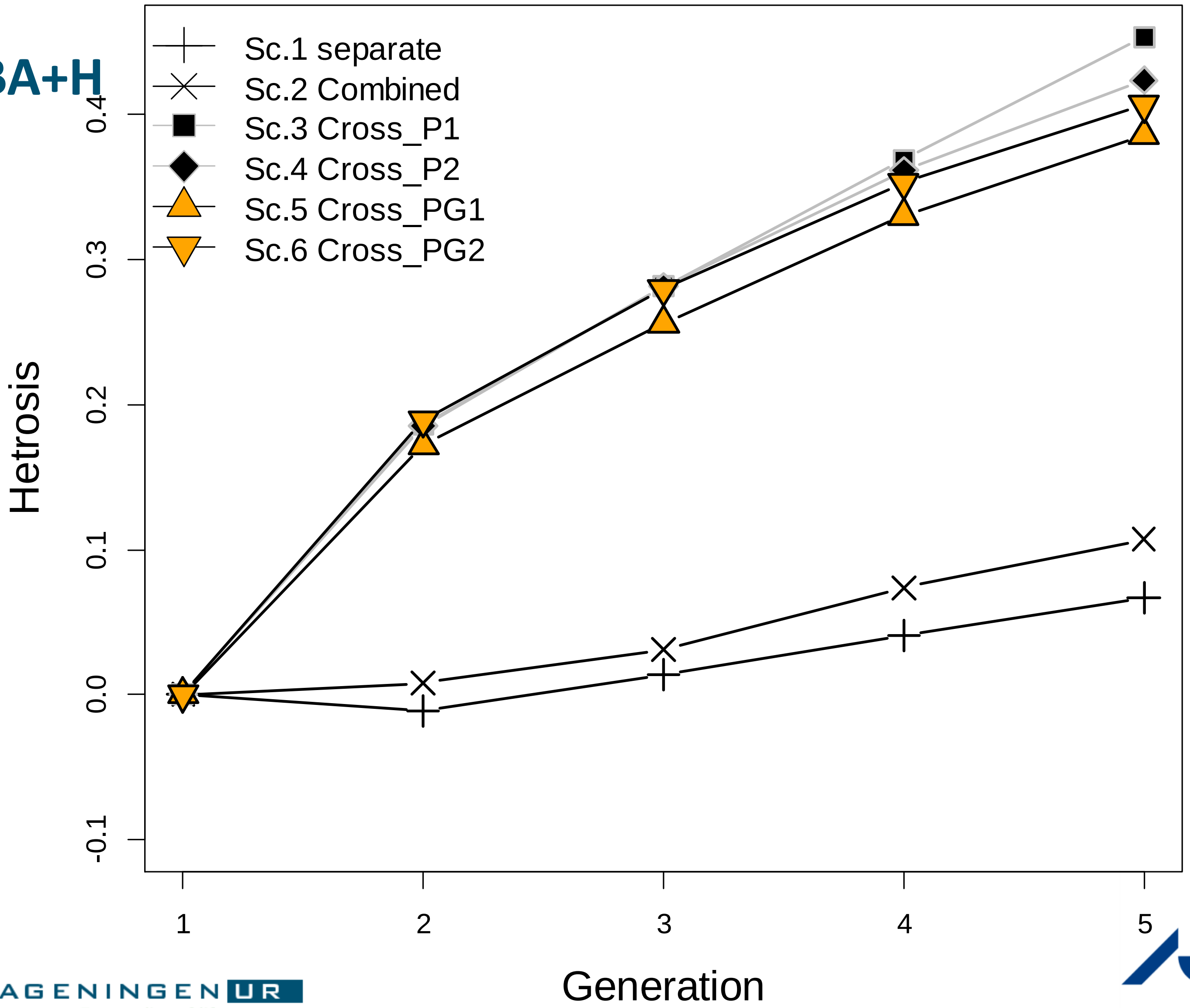


AA-Aa-aa

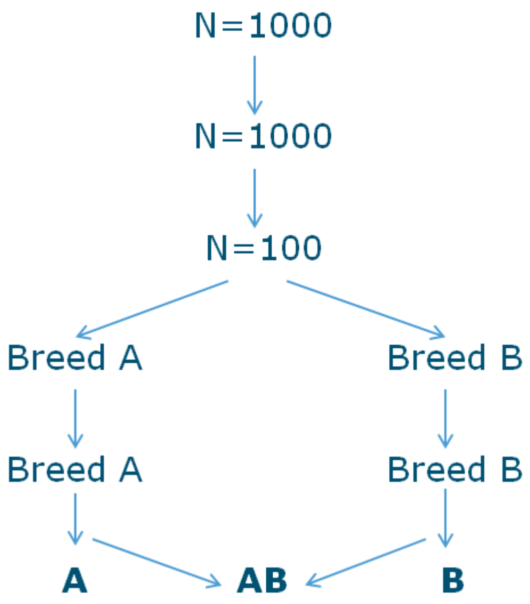
AA-Aa-aA-aa

HETEROSIS

CP=BA+H



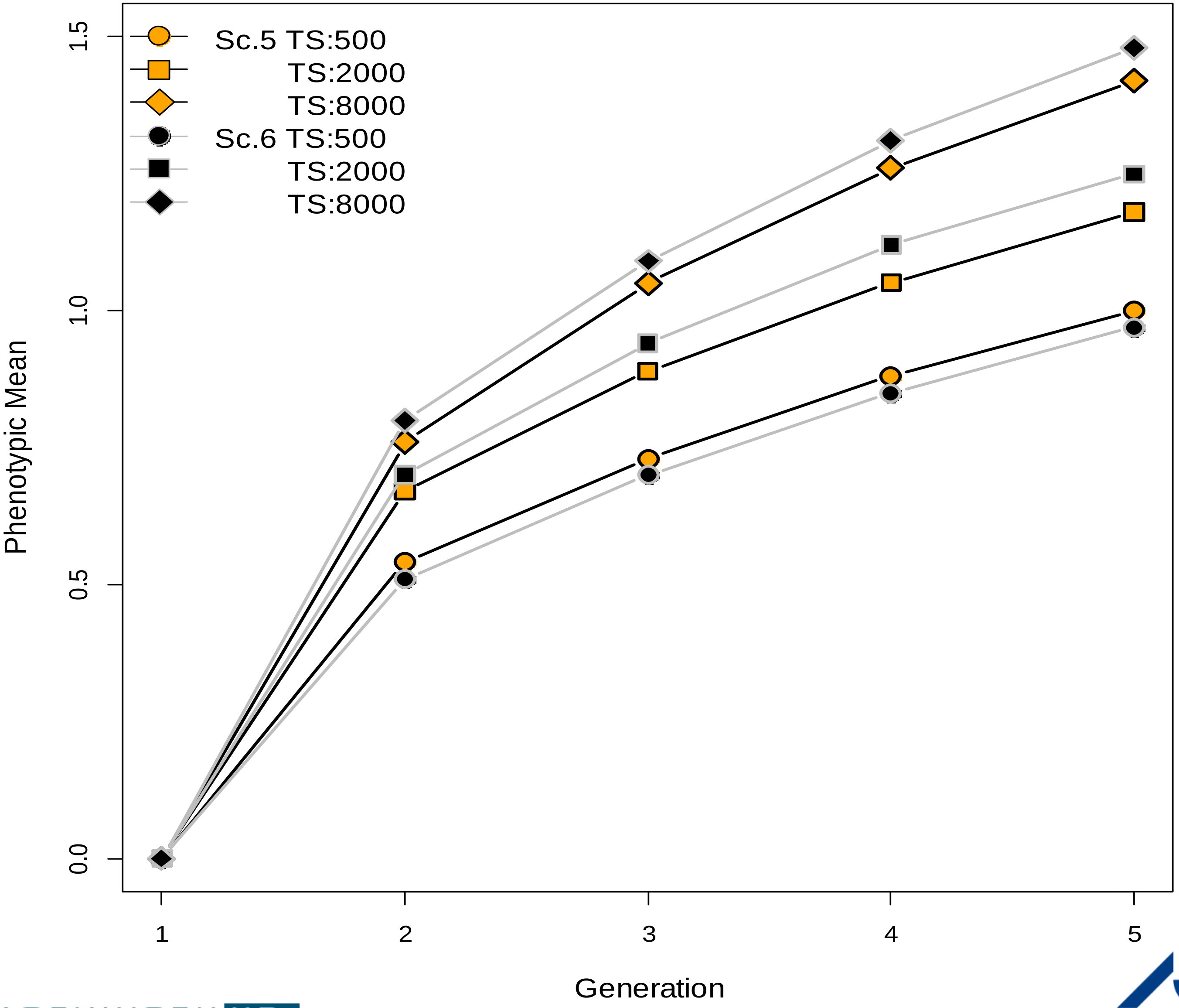
The impact of time since divergence of breeds



300

	Time since divergence				
	1	50	100	200	400
Sc. 5	1.21	1.32	1.33	1.20	0.94
Sc. 6	1.15	1.28	1.30	1.19	0.99

Impact of varying size of the training population



WAGENINGEN UR
For quality of life

AARHUS UNIVERSITY

Prediction models

- **Additive model (MA)**
  - **GEBV** → **Expected genotypic value of the offspring of a boar**

$$\sum_{j=1}^{s} \textbf{AA} \quad (p_{jr}\hat{a}_j)] + \textbf{Aa} \quad (0.5p_{jr}\hat{a}_j - 0.5q_{jr}\,\hat{a}_j)] + \textbf{aa} \quad -q_{jr}\hat{a}_j)]$$

- **Additive and dominance model (MAD)**
  - **GEBV**
  - **GEBV-C (SNP allele frequencies from opposite breed)**

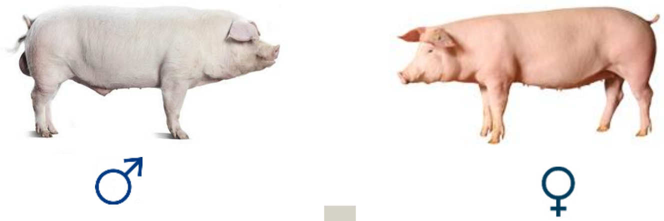
- **Prediction accuracy**

$$cor(GEBV, \overline{P}_{CB})$$

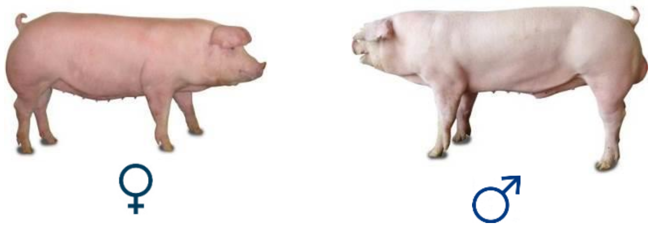
CONCLUSIONS

- Training on CB → Higher response to selection than training on pure lines
- Taking into account the breed origin of allele would increase response to selection **EXCEPT:**
    - Breeds are closely related
    - The reference population is small

Landrace



Yorkshire



Full data	489 000
Genotyped	2742
Sow	2087
Boar	655

Full data	316 000
Genotyped	2330
Sow	0
Boar	

$$\overline{\boldsymbol{P}}_{\boldsymbol{CB}}$$

$$\mathrm{GEBV}-\mathrm{P}=\sum_{j=1}^{s}[(x_{ij})(p_{jr}\hat{a}_j+q_{jr}\hat{d}_j)]+[(y_{ij})(0.5p_{jr}\hat{a}_j+0.5q_{jr}\,\hat{d}_j$$

$$+0.5p_{jr}\hat{d}_j-0.5q_{jr}\,\hat{a}_j)]+[(z_{ij})(-q_{jr}\hat{a}_j+p_{jr}\hat{d}_j)]$$

WAGENINGEN UR
For quality of life