

Practical implementation of genomic evaluation across species

Daniela Lourenco

University of Georgia

EAAP 2015, Warsaw

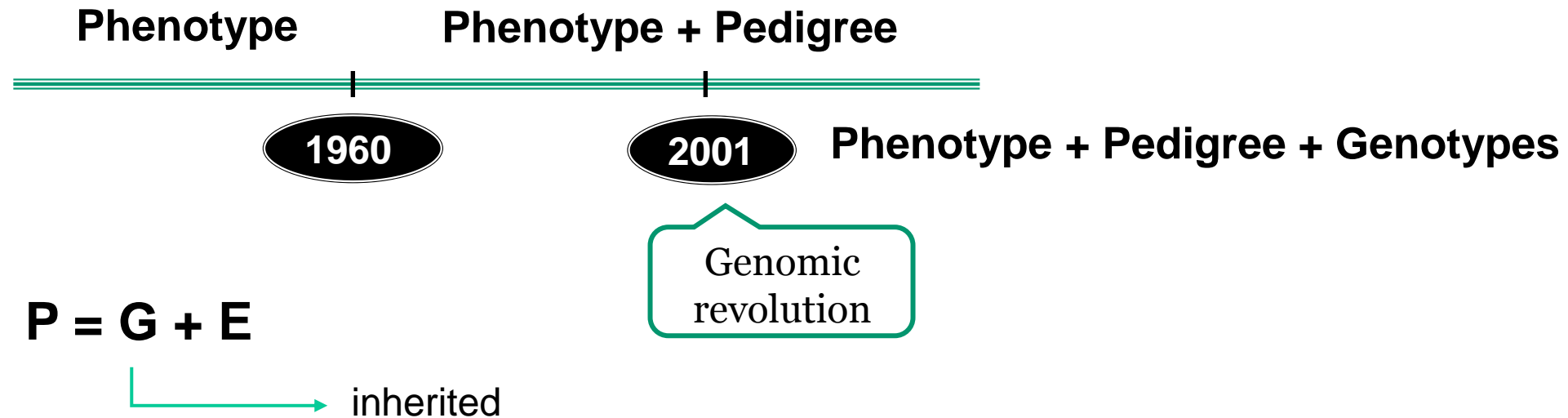
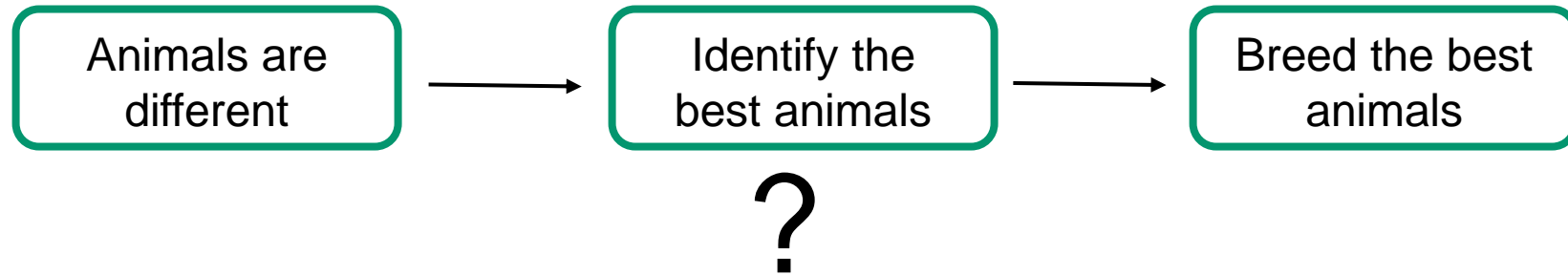
Why is the industry interested in Genomic Selection?

- Because everybody is talking about it?
- Because
 - More accurate evaluation
 - Reduces generation interval
 - Increases genetic gain



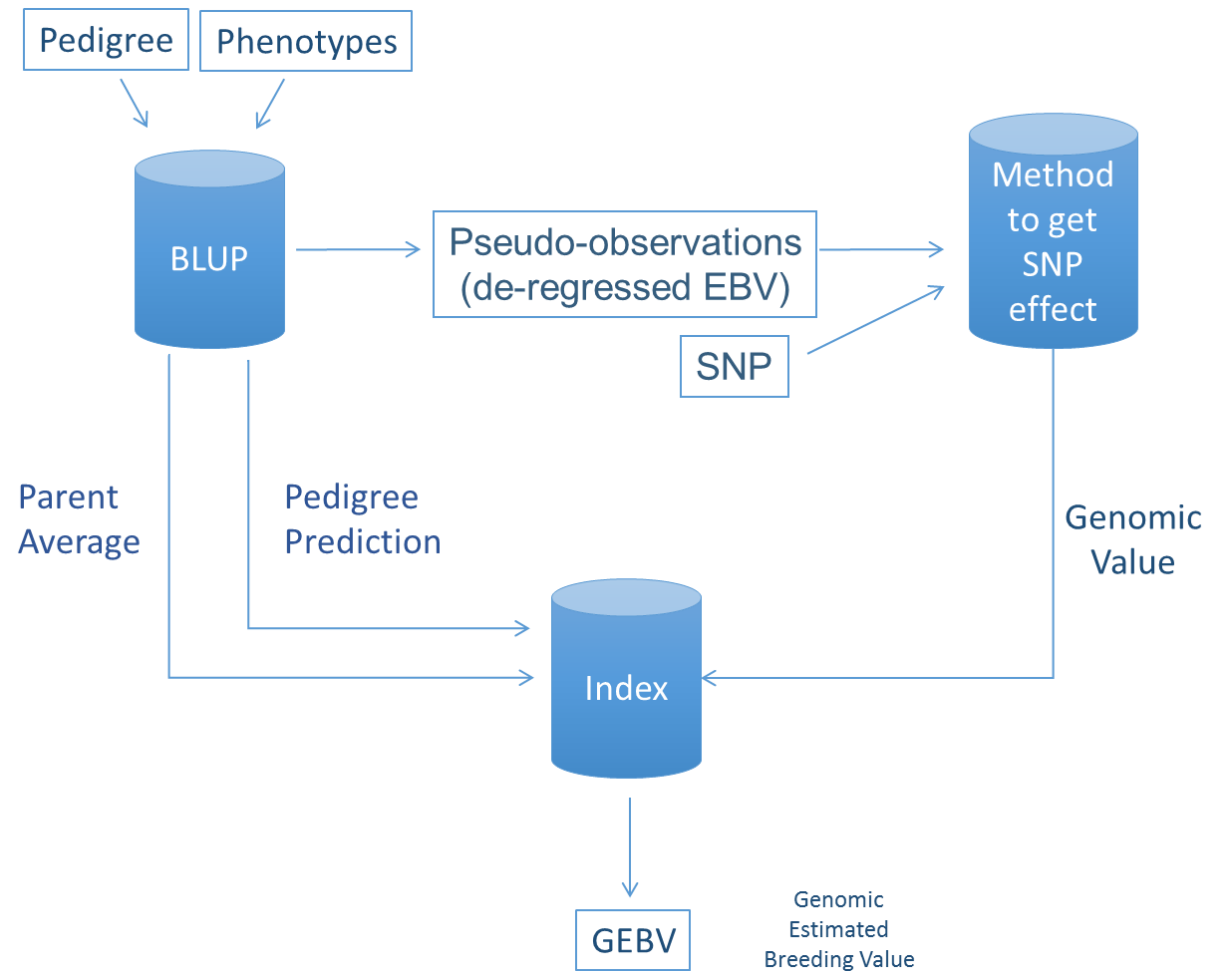
Biggest advance
since AI

Animal breeding timeline



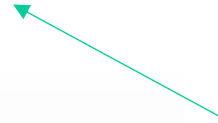
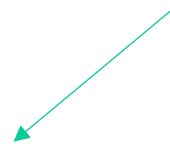
Methods for GS

- **Multistep**



Methods for GS

- **Single-Step GBLUP
(UGA)**



Methods for GS

- **Single-Step GBLUP**

BLUP

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{H}^{-1}\lambda \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

Henderson, 1963

ssGBLUP

single-step genomic BLUP

Data

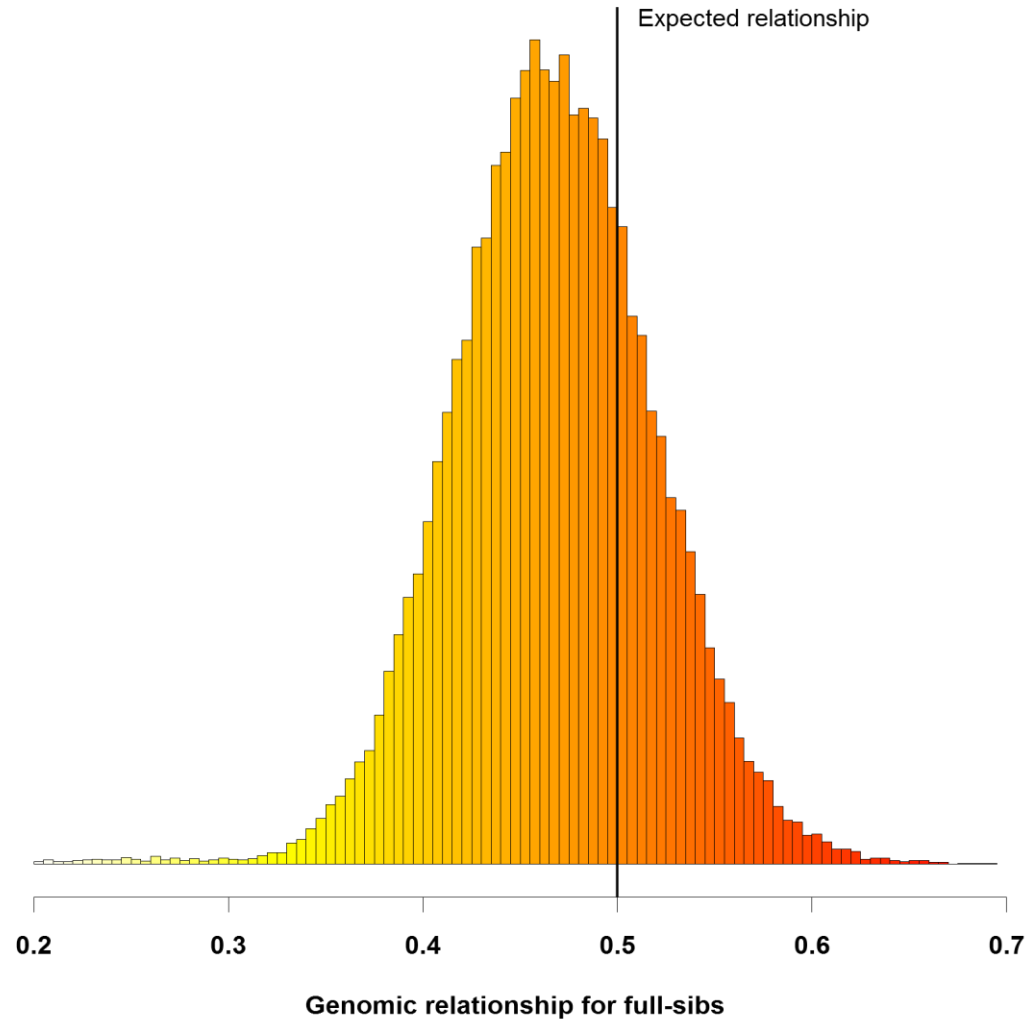
Pedigree

**Genomic
Info**

Realized relationships

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

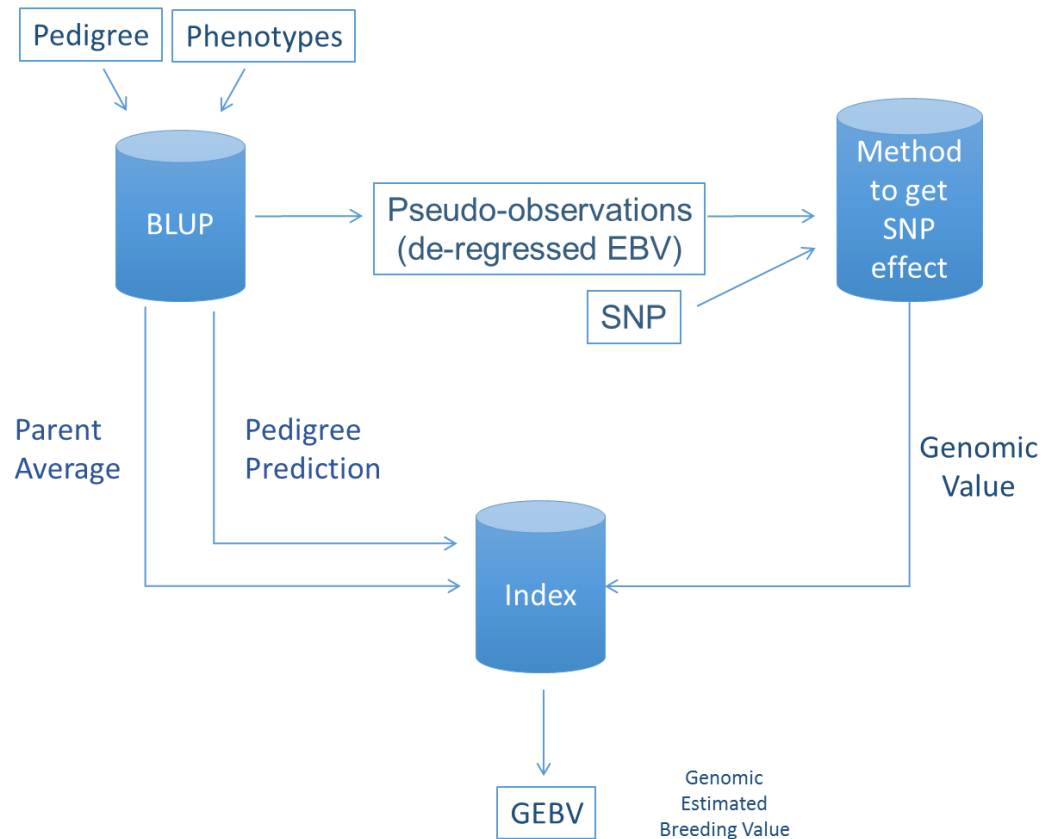
Aguilar et al., 2010



Lourenco et al., 2015

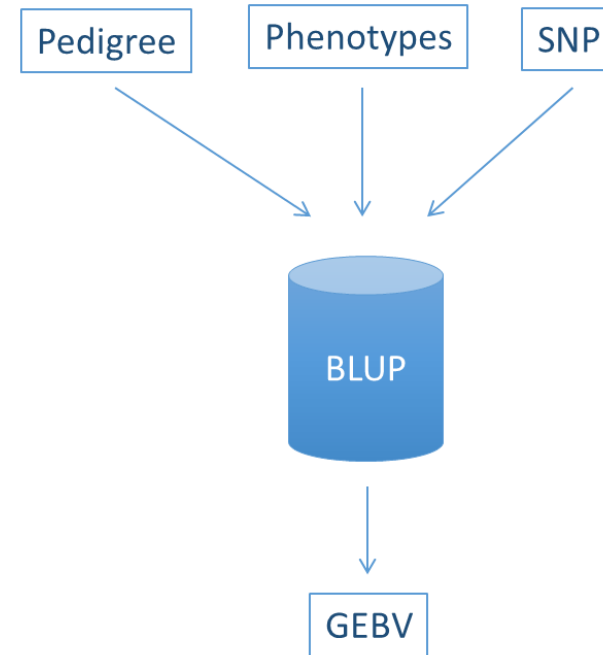
Methods for GS

- **Multistep**



- **Single-step**

- Simplicity
- Raw phenotypes
- Same models



Frequent questions - GS

- Is GS working?
- Which species are into GS?
 - Dairy cattle
 - Beef cattle
 - Chicken
 - Pigs
 - Sheep
 - Goats
 - Fish
 - Horse
 - Honey bee
 - ...

Frequent questions - GS

- Which species is getting more benefits?



<http://www.mbike.com/album-1001248/photo-4003750-original>

- \$ Genotyping < \$ Progeny testing
- High quality phenotypes (lots)
- AI highly used

Frequent questions - GS

- How much do we gain with GS?

Net Merit (NM\$)



<http://www.selectsires.com>

Reliability = 35%



<http://www.selectsires.com>

Reliability = 70%



<http://www.selectsires.com>



<http://www.wisholsteins.com>

Reliability = 85%

Frequent questions - GS

- Which animals should we genotype?
 - Should we genotype females?
- Does GS work for crossbred/multi-breed populations?
 - Are F1 worthy of genotyping?
- Will GS work when we have zillions of genotypes?
 - Are we able to handle that?

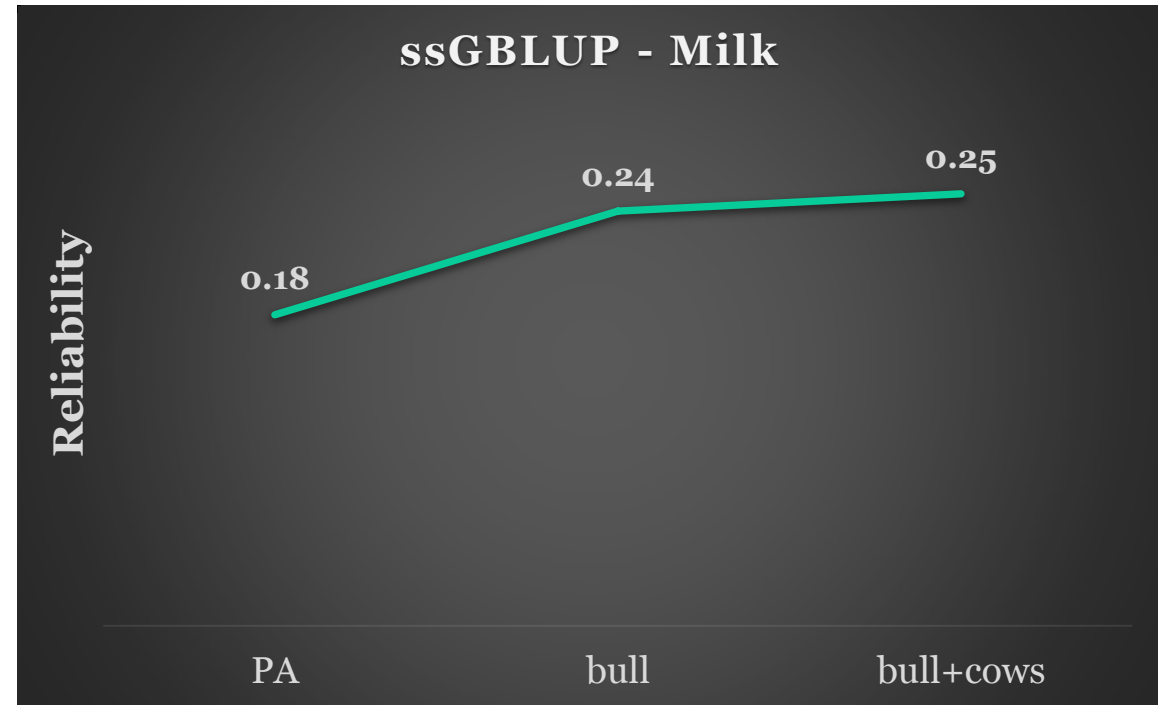
Should we genotype females?

- Females less impact
- Why to include females?
 - Increase the genotyped population
 - Increase accuracy / reliability of GEBV
 - Reduce error
 - Why not?

Should we genotype females?

- **Israeli Holsteins**
 - 1.5M MFP records
 - 830k animals in pedigree
 - 1.3k genotyped bulls
 - 350 genotyped cows

Small genotyped population
+6 points = 39%

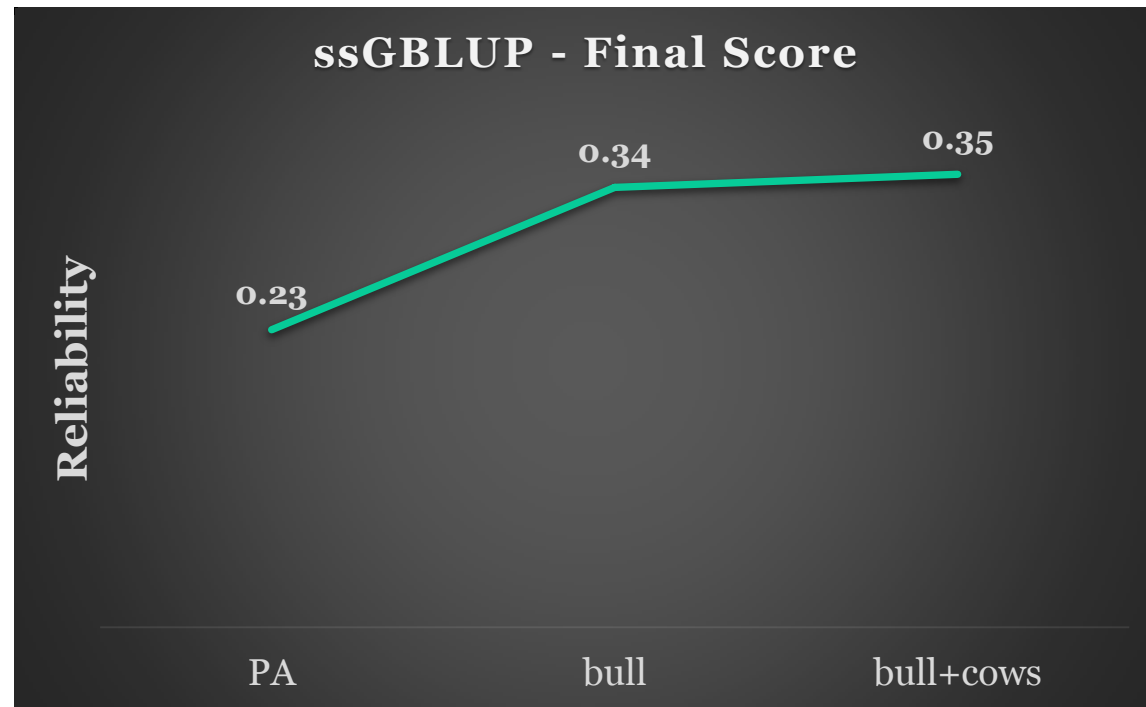


Lourenco et al., 2014

Should we genotype females?

- **US Holsteins**

- 11M FS records
- 10M animals in pedigree
- 34.5k genotyped bulls
- 5.2k genotyped cows

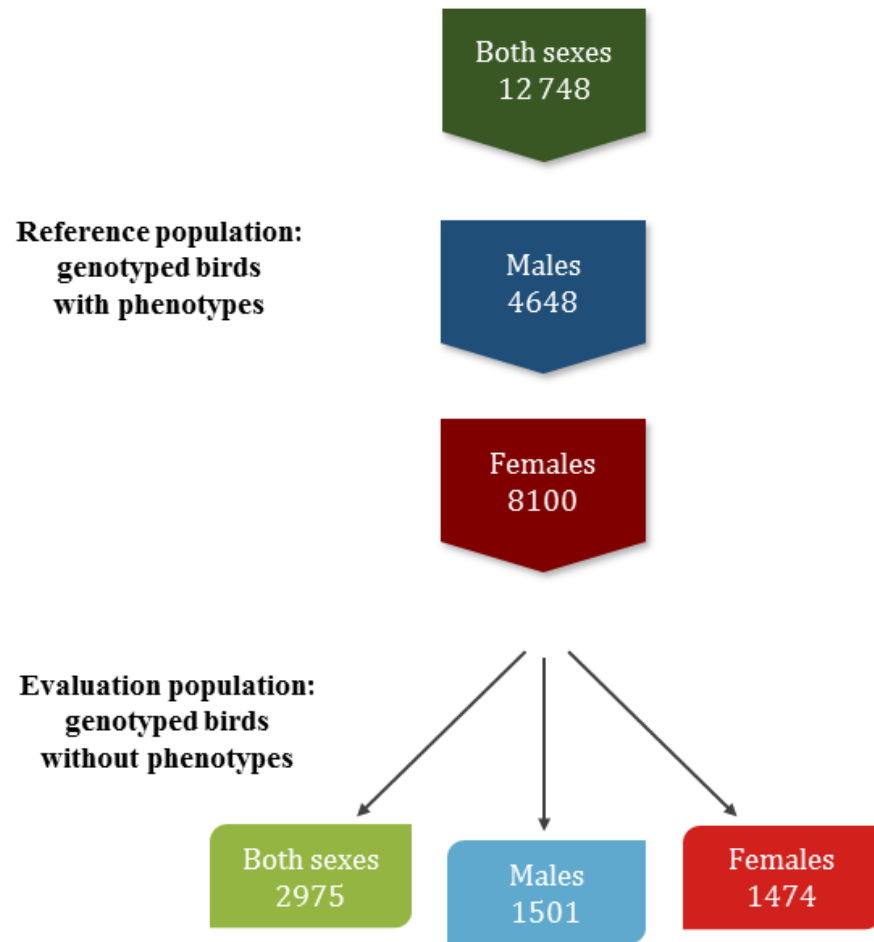


Tsuruta et al., 2014

Should we genotype females?

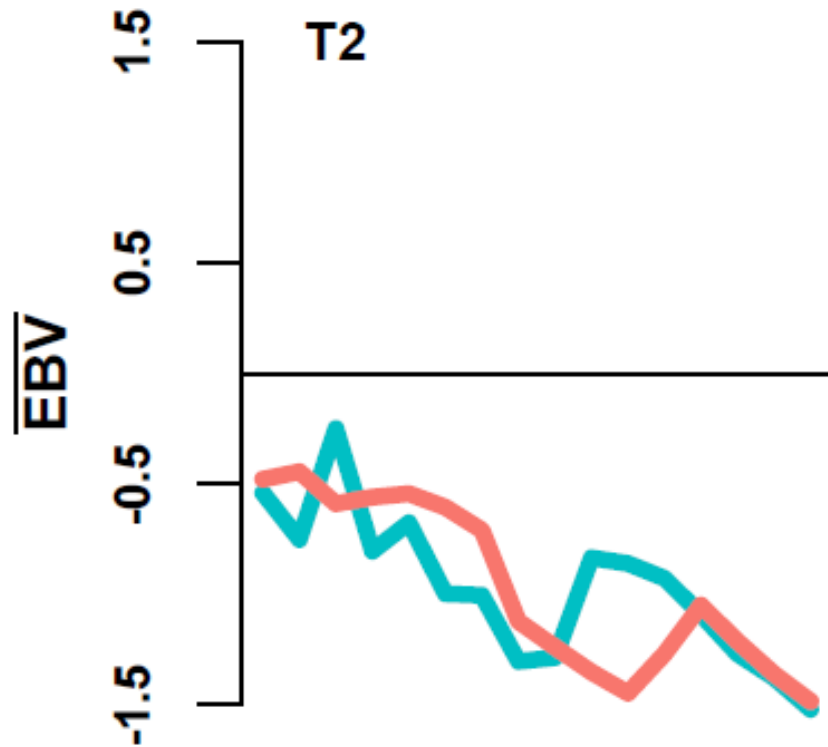
- Holsteins
 - Small increase in accuracy
 - Females have few progeny – less information
- Other species?
 - Females have higher impact – more progeny
 - Broiler chicken
 - Pigs
- Broiler chicken from Cobb-Vantress
 - 15k genotyped animals
 - 200k phenotypes / pedigree

Should we genotype females?



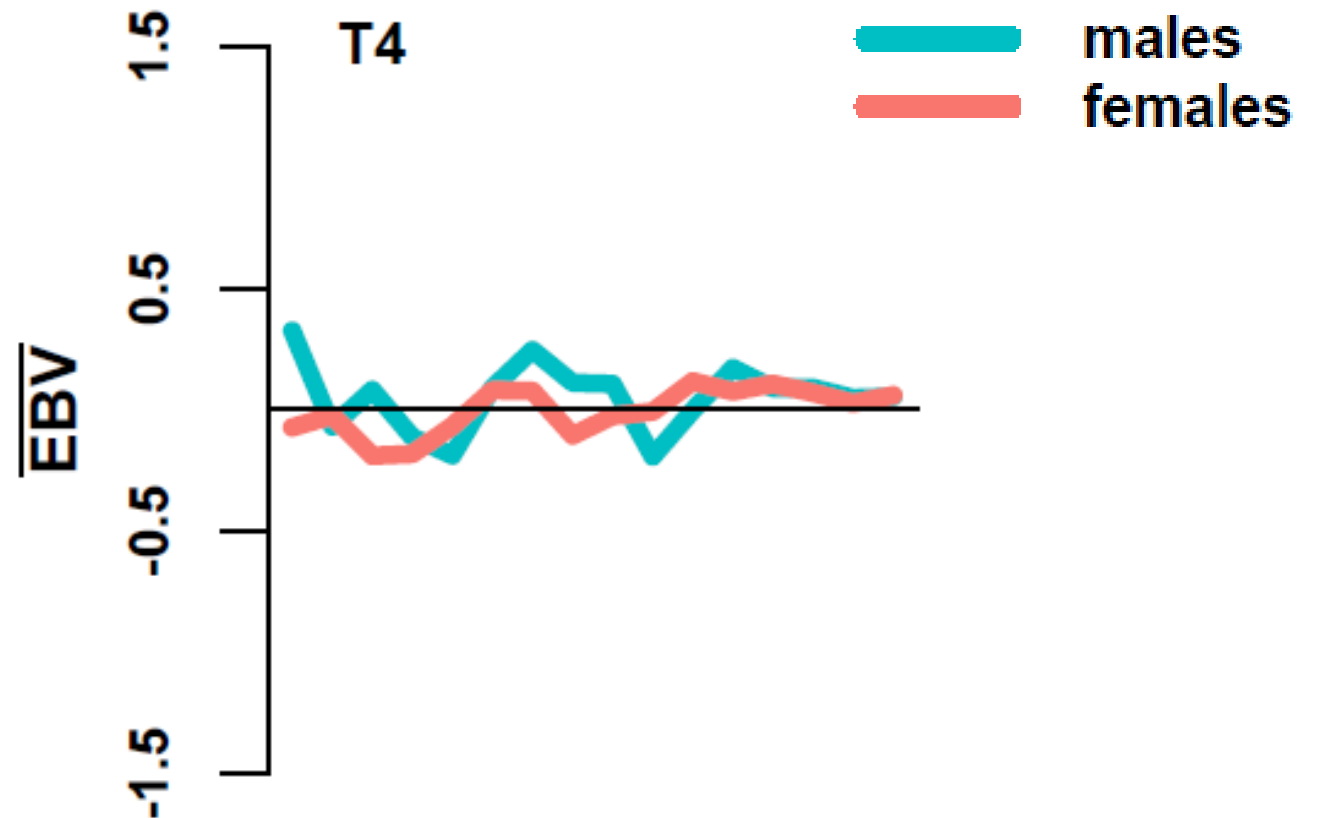
Should we genotype females?

$h^2 = 0.25$



$h^2 = 0.22$

= Amount of info



Should we genotype females?

TAKE HOME #1

- Small gain in accuracy in dairy
- Females already gave information through males - pedigree
- Big gain when females have higher reproductive impact

Does GS work for crossbred/multi-breed populations?

- Few genotyped animals
 - Increase the size of genotyped population
 - Higher accuracy
- \neq Genetic background
 - \neq Allele frequency



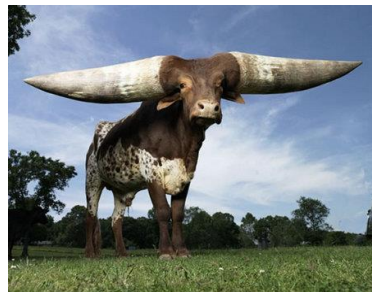
<http://myttingus.com/herd-sires.html>



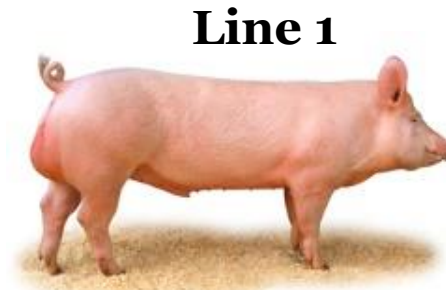
<http://www.thecattlesite.com/breeds/bee-f/40/highland/overview>



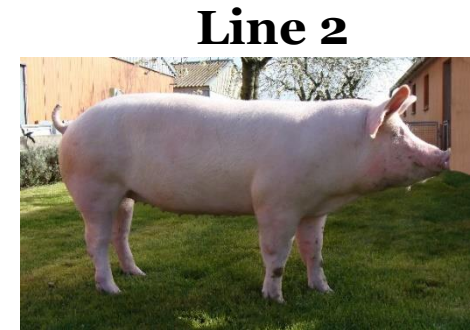
<http://www.richardbealblog.com/simmental-cattle/>



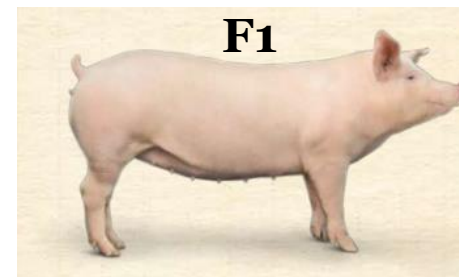
http://www.indiacatalog.com/photo_gallery/ic_images.php?id=40



<http://rattlerow.co.uk/gp-large-white-36/>

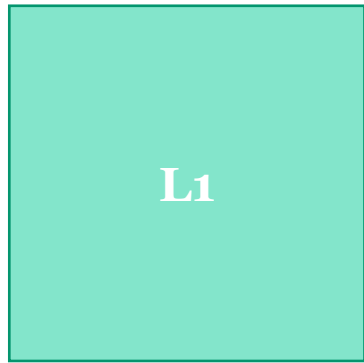


http://www.geneplus.com/GB/o2_selection/femelle.htm

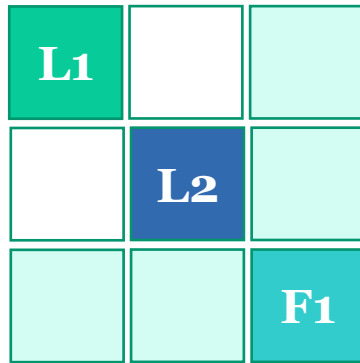


http://www.pic.com/Images/Users/1/usa/Gilts/PIC_Gilts.pdf

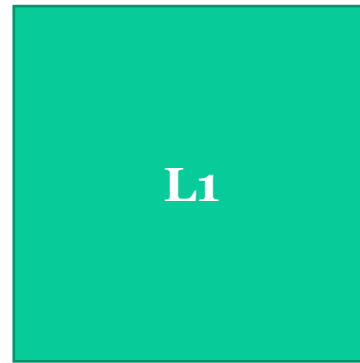
Does GS work for crossbred/multi-breed populations?



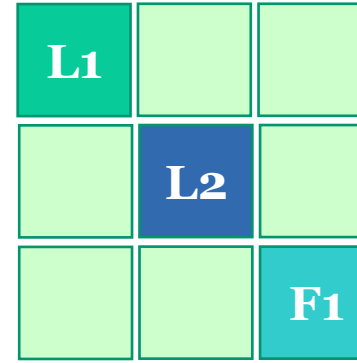
A
Single breed



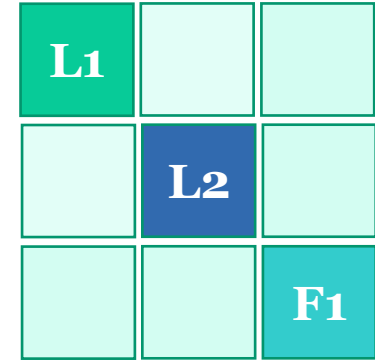
A
Multi-breed/
crossbred



G
Single breed



G
Across-breed
allele frequency



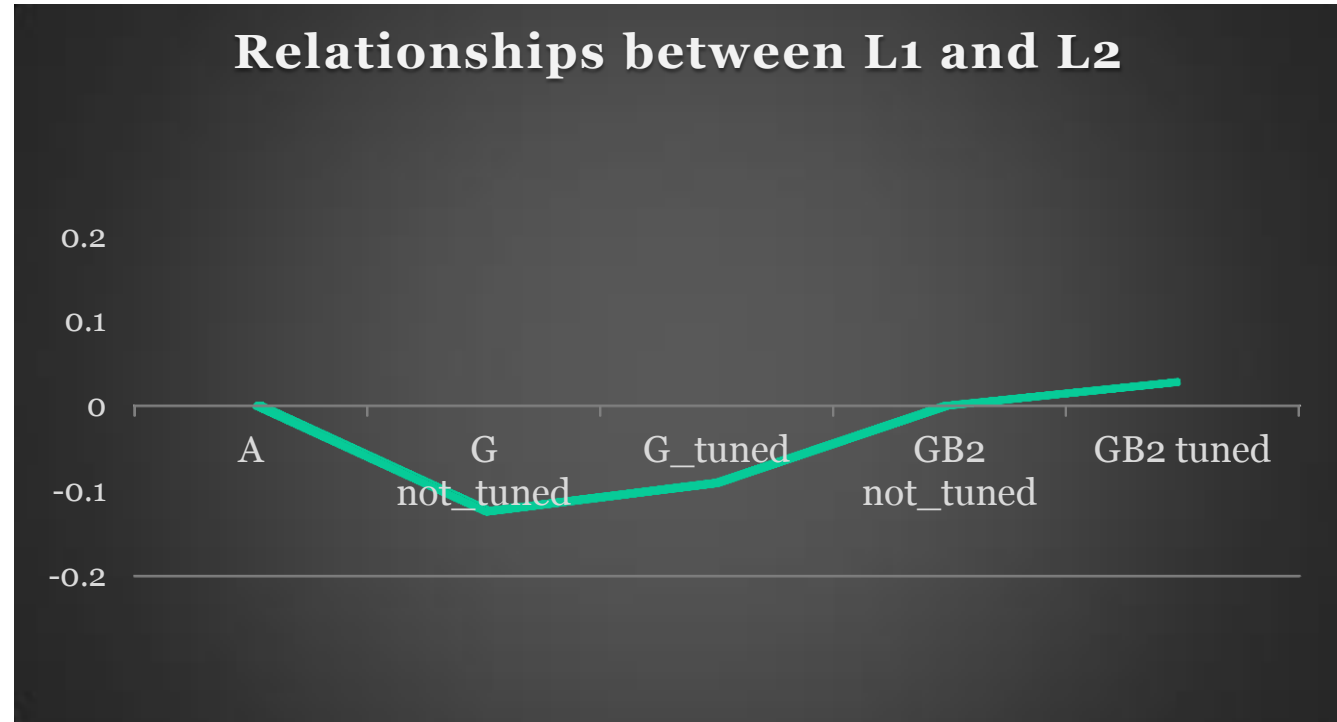
G
Allele frequency specific
for each breed

- G matrix accounting for breed-specific allele frequency

Does GS work for crossbred/multi-breed populations?

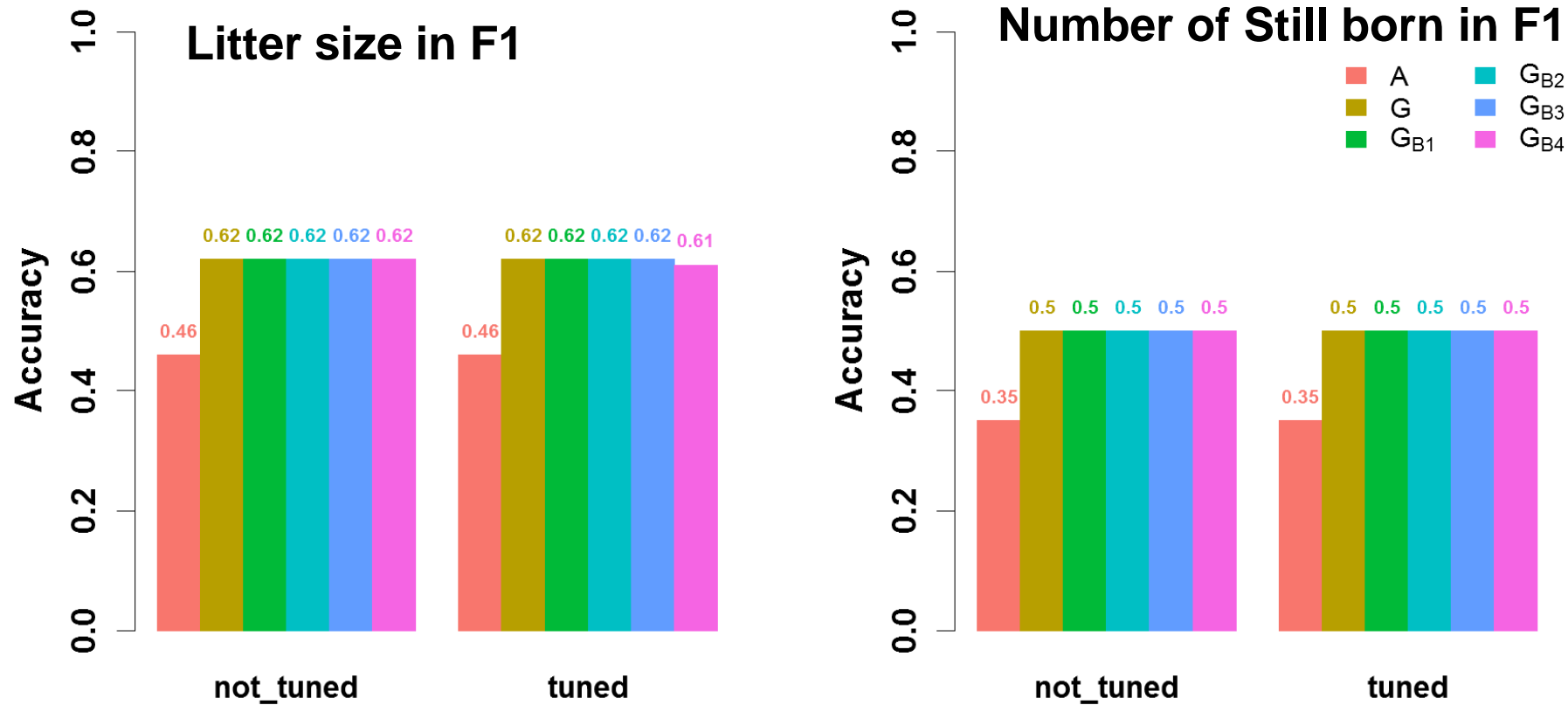
Lourenco et al., 2015

- Pigs from PIC
 - 2.3M records
 - 5.6M animals in pedigree
 - 15,800 genotyped animals
 - 6,000 Line 1
 - 6,700 Line 2
 - 3,100 F1



- Tuned **G** = correct **G** to match **A**₂₂
- Not_tuned **G** = no corrections

Does GS work for crossbred/multi-breed populations?



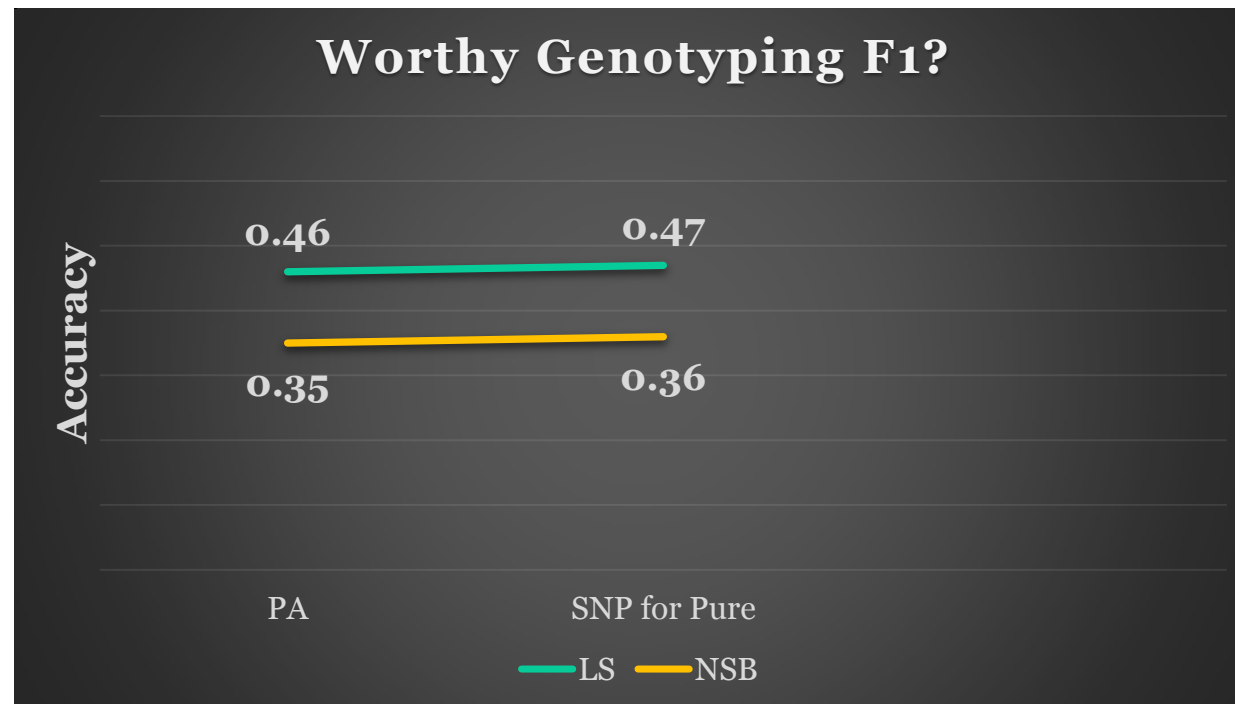
Lourenco et al., 2015

Accuracy of predicting future phenotypes for genotyped F1

Are F1's worthy of genotyping?

- Accuracy of predicting future phenotypes for genotyped F1
- Gains by genotyping F1

Lourenco et al., 2015



**Increase in
accuracy
35%**

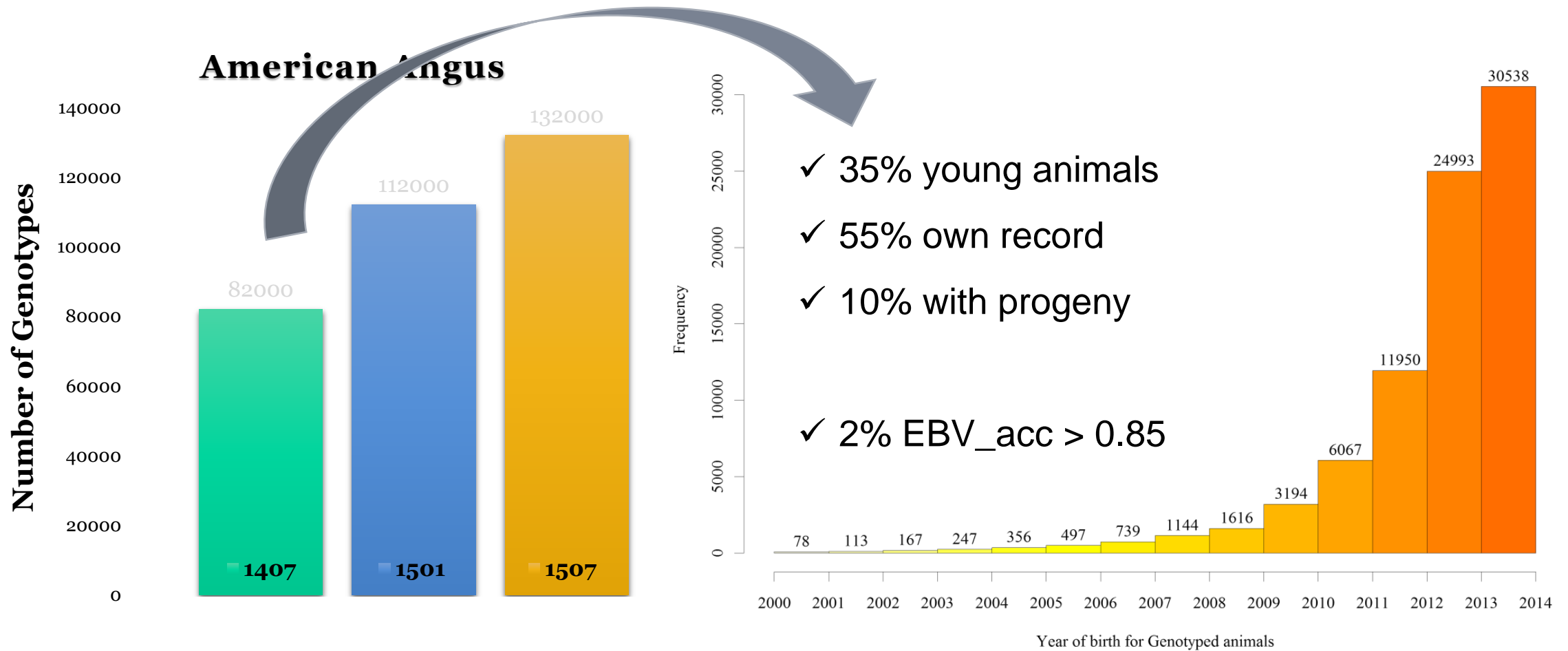
Does GS work for crossbred/multi-breed populations?

Are F1's worthy of genotyping?

TAKE HOME #2

- Do not need to use breed-specific allele frequency in ssGBLUP
- Scale regular **G** to match **A**₂₂
- Increase in accuracy by genotyping F1

Which animals should be genotyped?



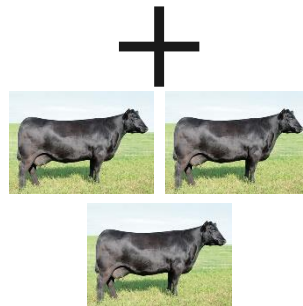
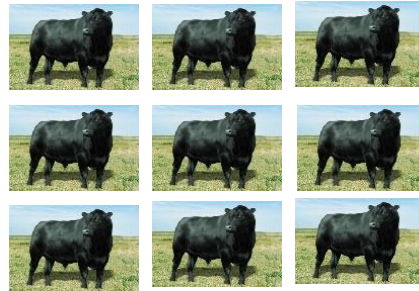
Which animals should be genotyped?

- Gains from GS in this population



<http://www.cattle.com/articles/title/Angus+Cattle.aspx>

Top bulls
Acc ≥ 0.85
N ~ 1,700



<http://www.newmanangus.com/donors/SAVBlackcapMay4568.html>

Top B+C
Acc ≥ 0.85
N ~ 2,000



<http://benfield-whiteridge.com/pl>



www.hansenangusranch.com/

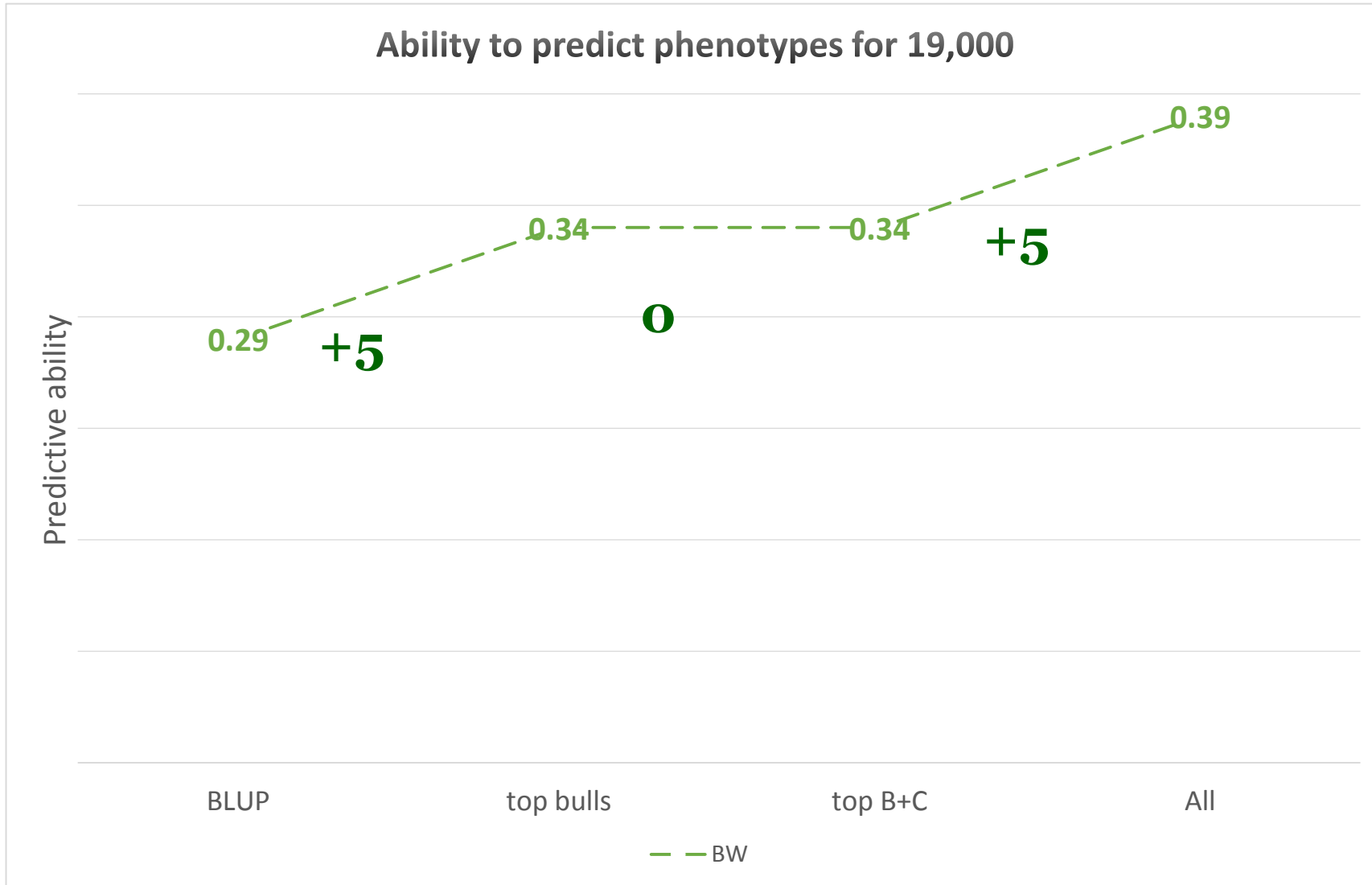


<https://farmcentrefrancais.wordpress.com/page/2/>

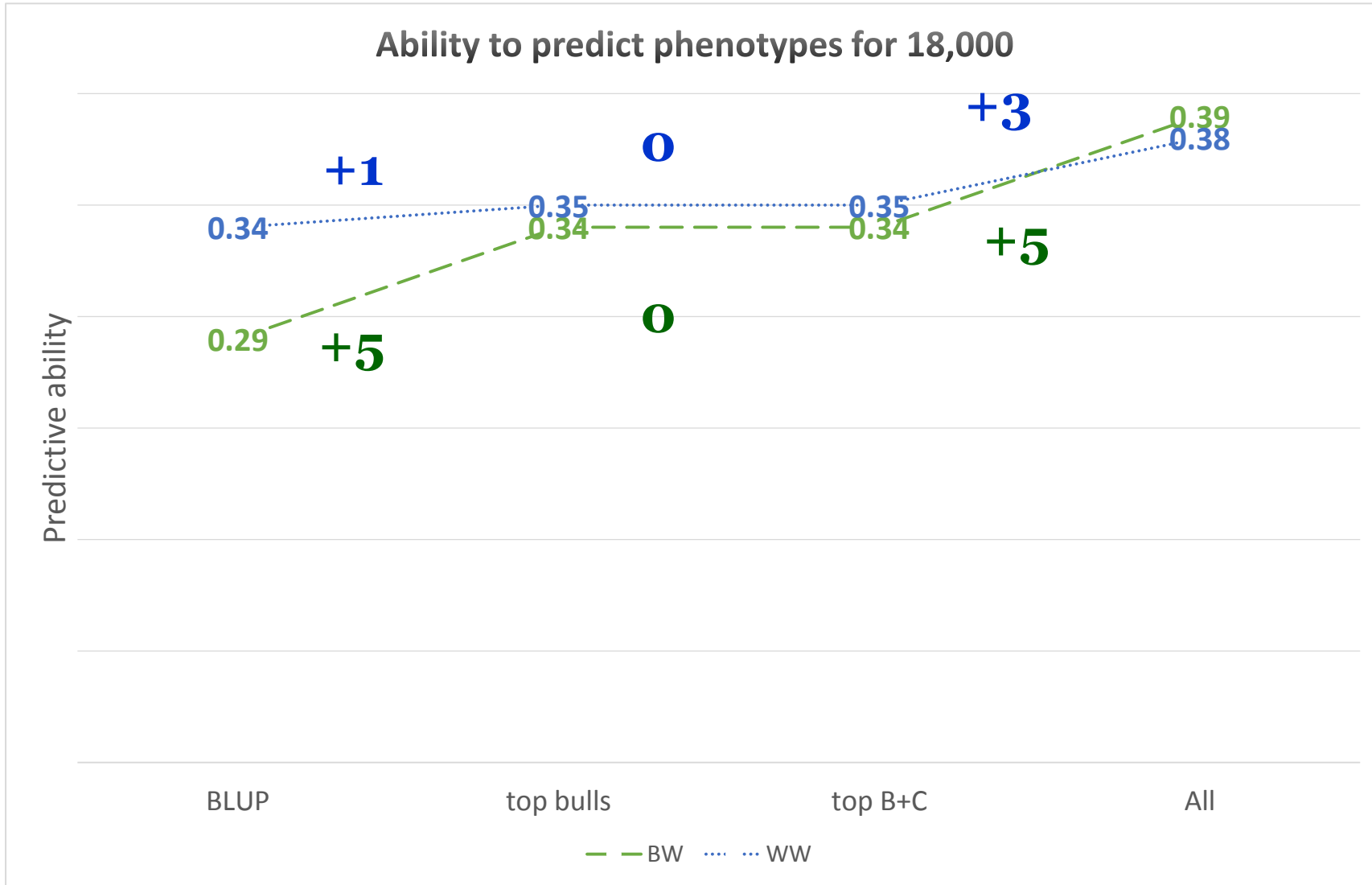
All
N ~ 33,000

Validation
N ~ 19,000

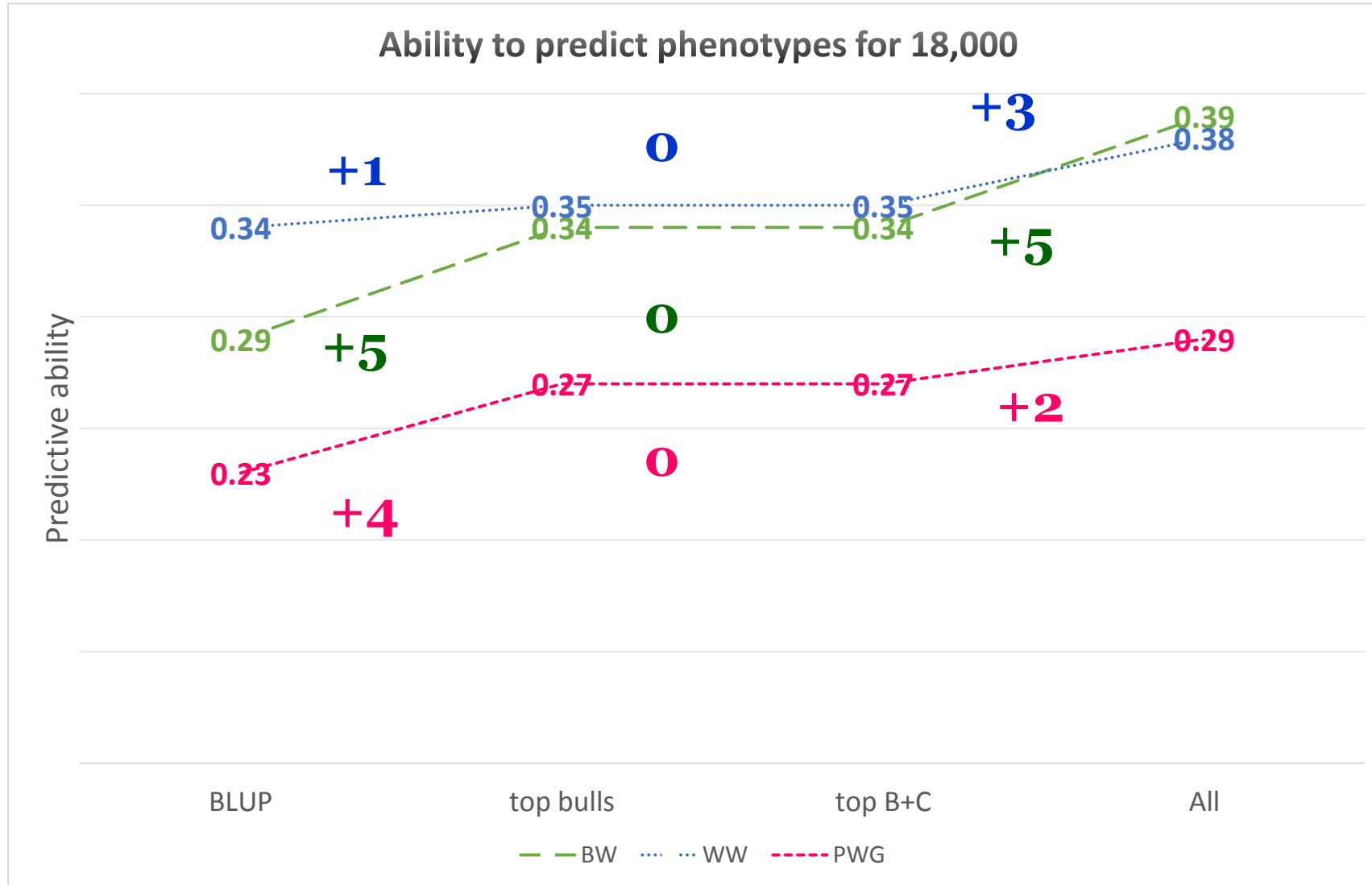
Which animals should be genotyped?



Which animals should be genotyped?



Which animals should be genotyped?

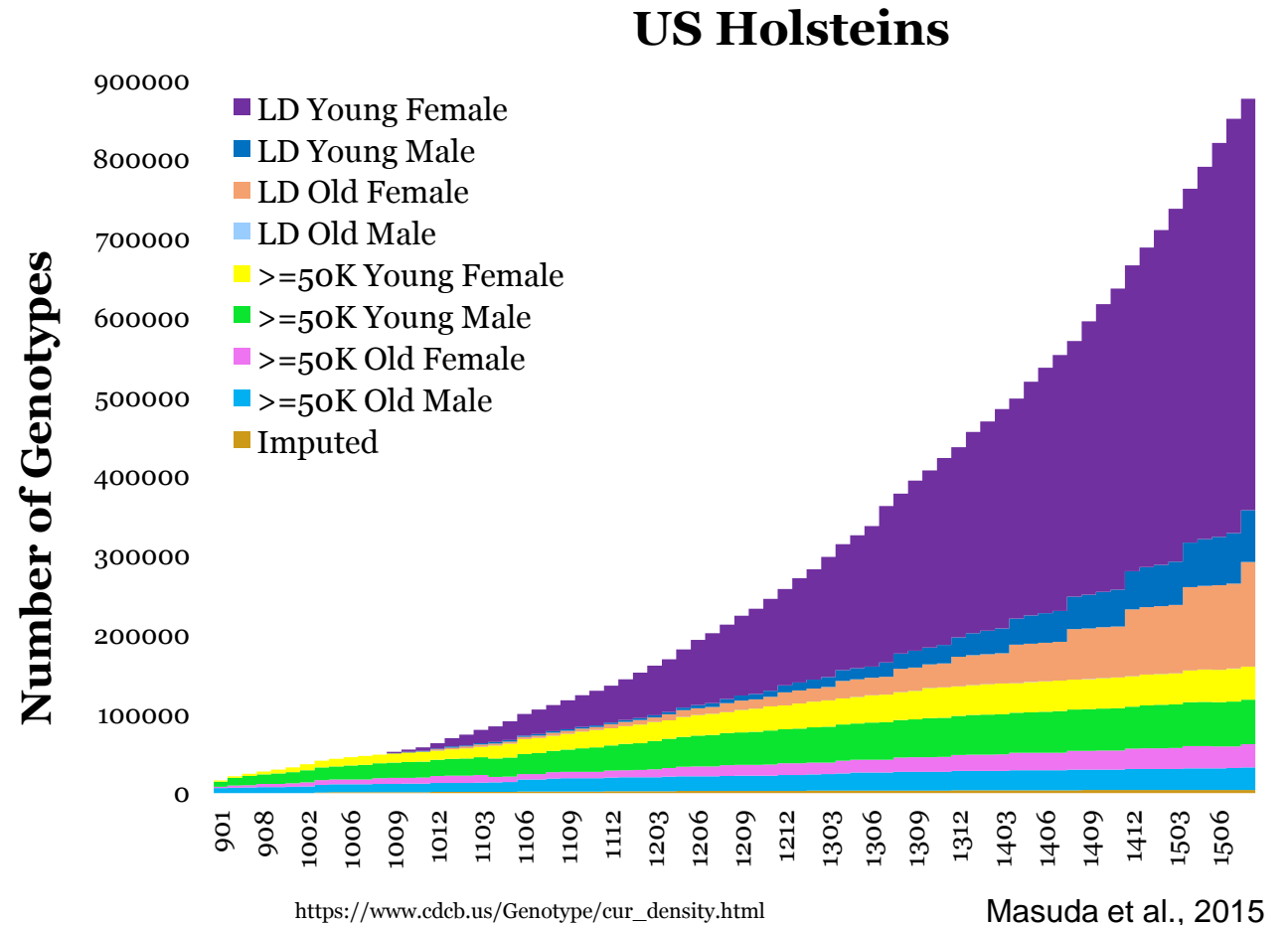
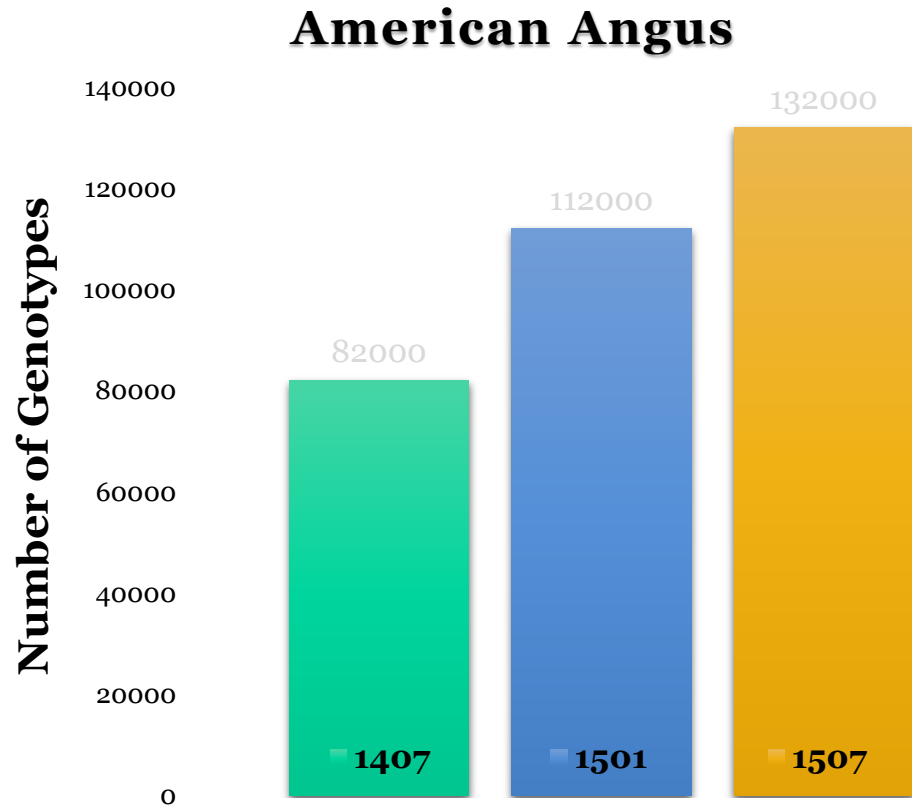


Which animals should be genotyped?

TAKE HOME #3

- Keep genotyping young animals
- Look for “very informative” animals – lots of links

GS for zillions of genotypes?



GS for zillions of genotypes?

- ssGBLUP
 - Create **G** and **A**₂₂
 - Direct inversion
 - Cubic cost
 - Limit ~ 150,000
 - 1.5h for 100,000

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

Aguilar et al., 2010

Association/ Species	# Genotyped	
US Charolais	2,000	✓
Fish	2,500	✓
Pigs	16,000	✓
Chicken	18,000	✓
Angus	132,000	✓
US Holstein	900,000	X

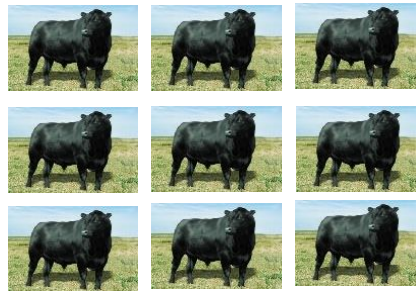
Limitation ~150,000

GS for zillions of genotypes?

- ssGBLUP - APY

Misztal et al., 2015

CORE



<http://www.cattle.com/articles/title/Angus+Cattle.aspx>

NON-CORE



<http://www.newmanangus.com/donors/SAVBlackcapMay4568.html>

<https://farmcentrefrancais.wordpress.com/page/2/>



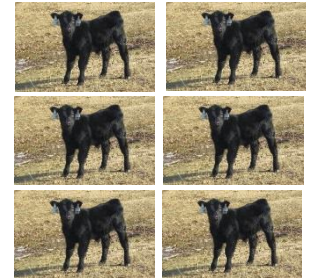
<http://www.streamcattle.com/images/2008/oct20/Lot6.jpg>



<http://benfield-whiteridge.com/phil.htm>



<http://www.hansenangusranch.com/>

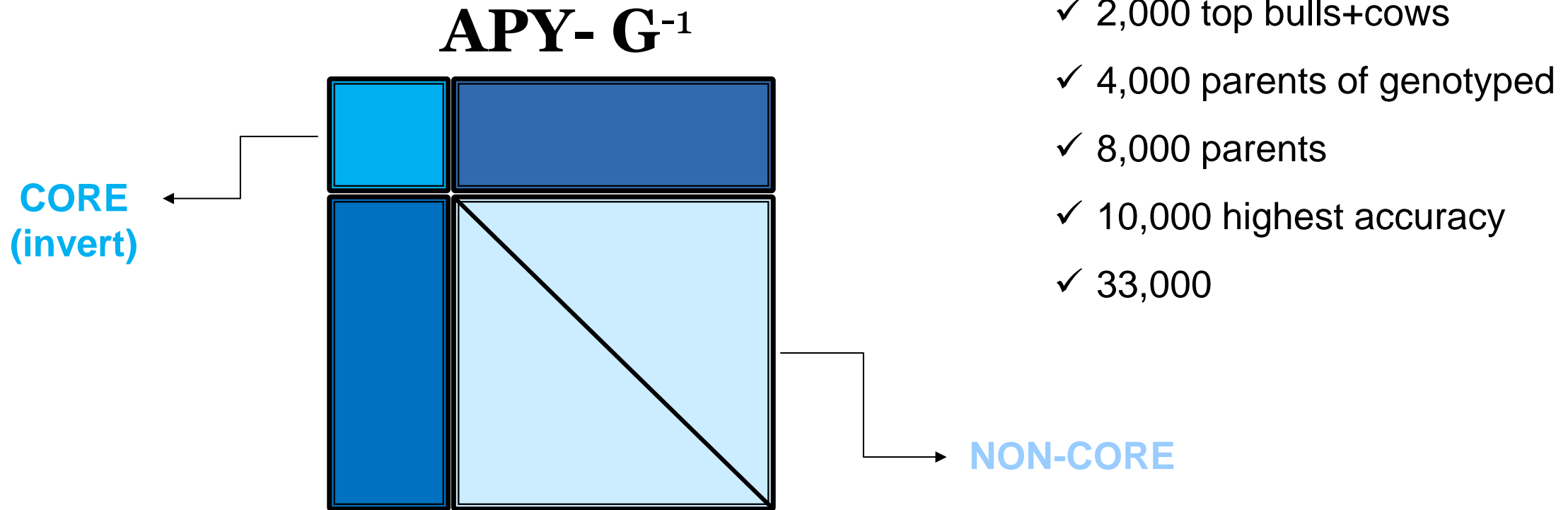


<http://www.rieksblackangus.com/>

GS for zillions of genotypes?

- **ssGBLUP - APY**

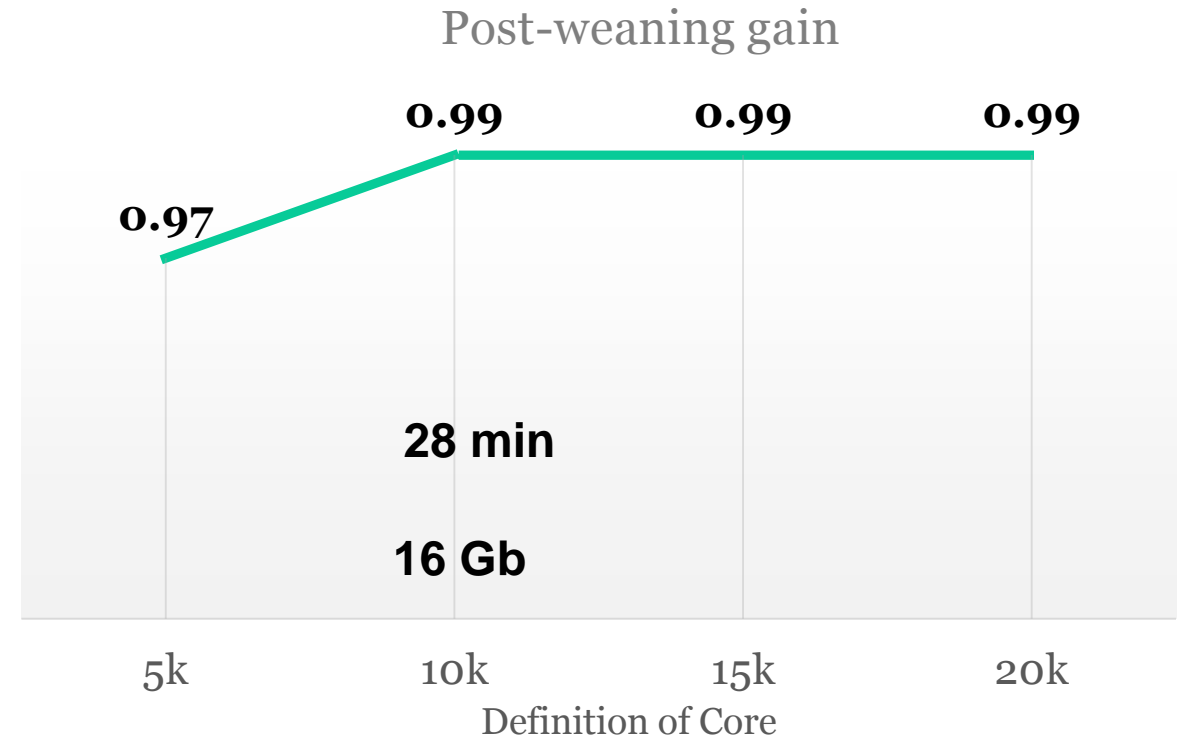
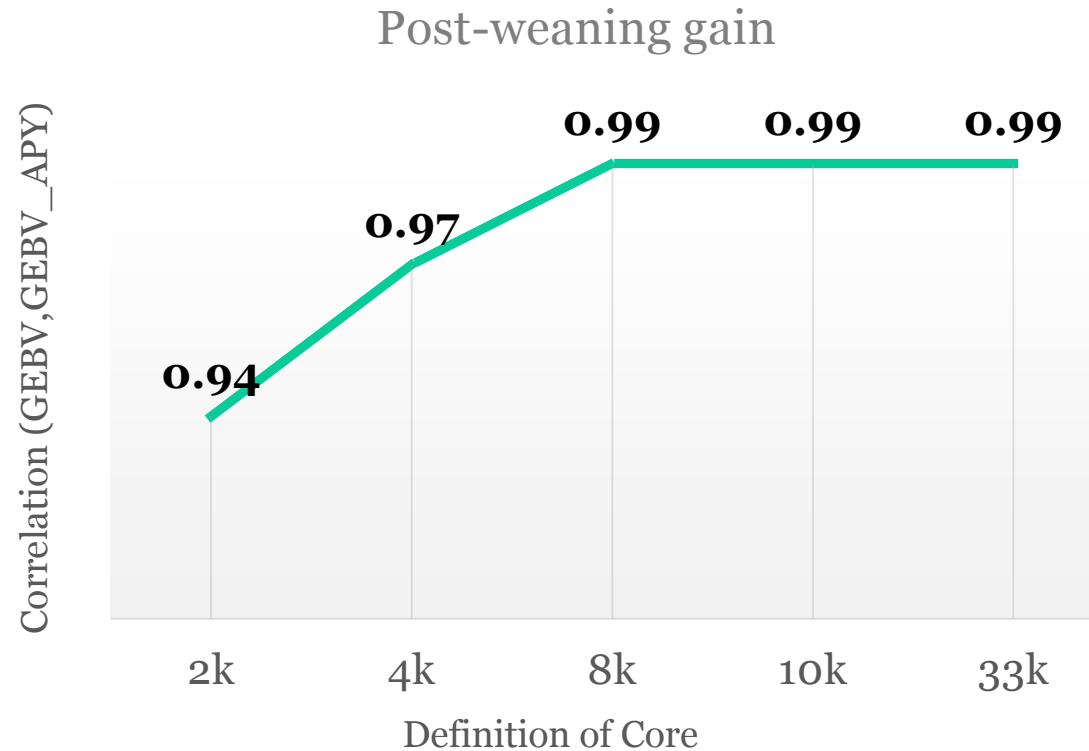
Total: 52k



GS for zillions of genotypes?

CORE animals defined by the amount of information

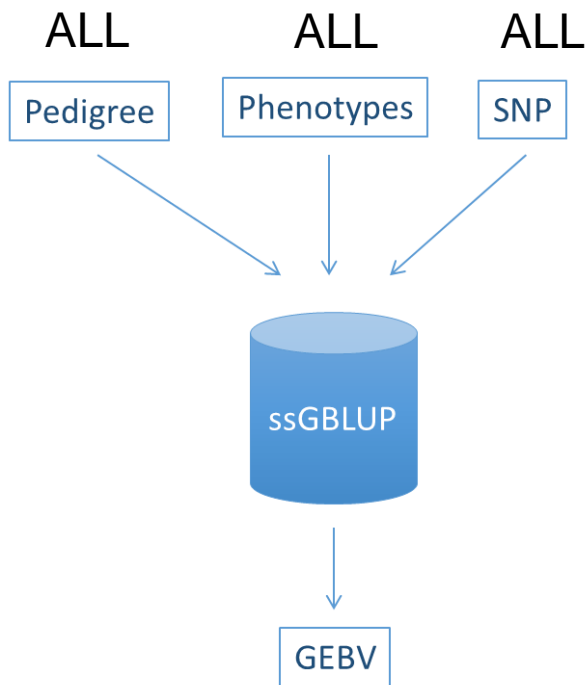
CORE animals randomly sampled from genotyped population



Regular G^{-1} = 213 min
230 Gb

GS for zillions of genotypes?

- **Direct Prediction**



✓ Genomic evaluations

✓ monthly

✓ weekly

✓ New genotypes

✓ daily

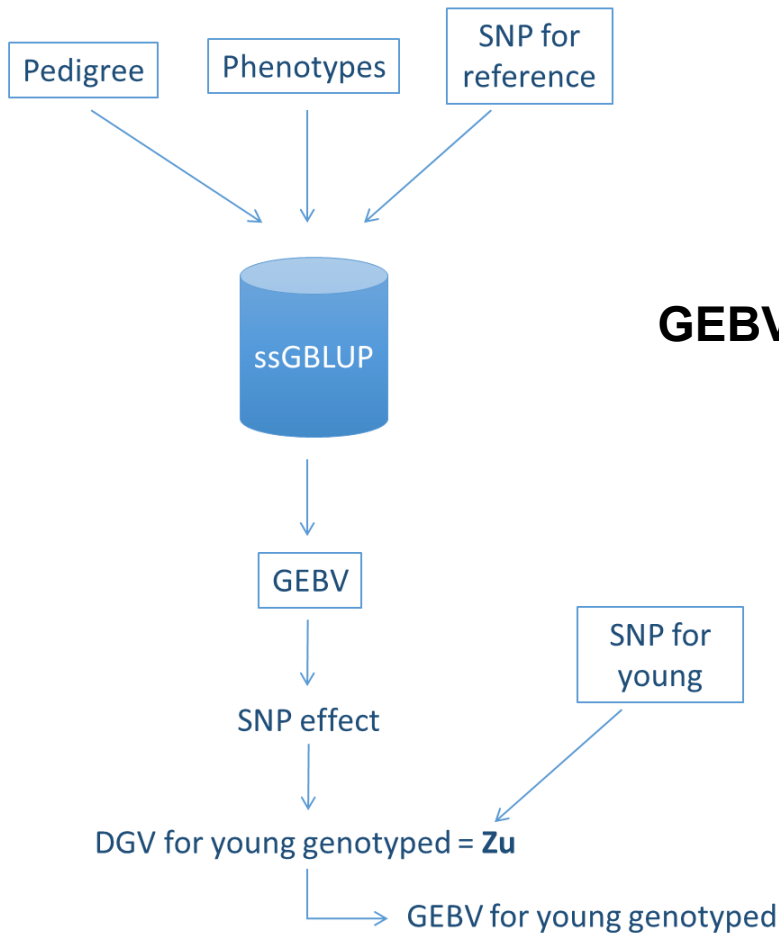
$$\text{GEBV} = \underbrace{w_1 \text{PA}}_{\text{parent average}} + \underbrace{w_2 \text{YD}}_{\text{yield deviation}} + \underbrace{w_3 \text{PC}}_{\text{progeny contribution}} + \underbrace{w_4 \text{DGV}}_{\text{direct genomic value}} - \underbrace{w_5 \text{PP}}_{\text{pedigree prediction}}$$

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

Arrows from the equation point to the matrix components: from 'parent average' to the top-left 0, from 'direct genomic value' to the bottom-right \mathbf{G}^{-1} , and from 'pedigree prediction' to the bottom-right $-\mathbf{A}_{22}^{-1}$.

GS for zillions of genotypes?

- **INDIRECT Prediction**



$$\text{GEBV} \approx w_1 \text{PA} + w_4 \text{DGV} - w_5 \text{PP}$$

=

$$\text{GEBV} \approx w_1 (\tfrac{1}{2} \text{EBV}_{\text{sire}} + \tfrac{1}{2} \text{EBV}_{\text{dam}}) + w_4 (\text{sum of all SNP effects})$$

Interim evaluations for young
genotyped animals

GS for zillions of genotypes?

- DIRECT Prediction**

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$$

VS

2,000
8,000
33,000

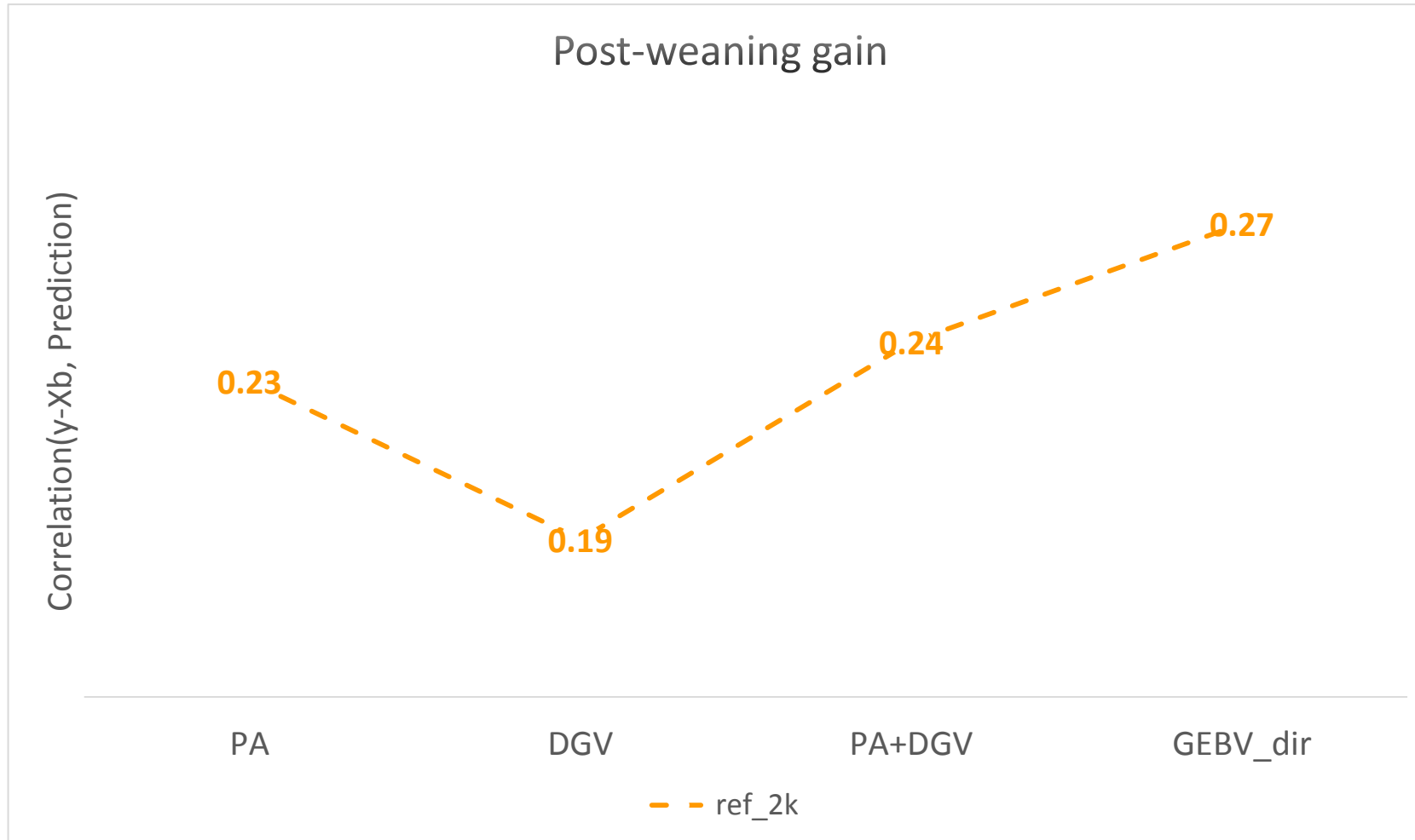
- INDIRECT Prediction**

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$$

2,000 + 18k
8,000 + 18k
33,000 + 18k

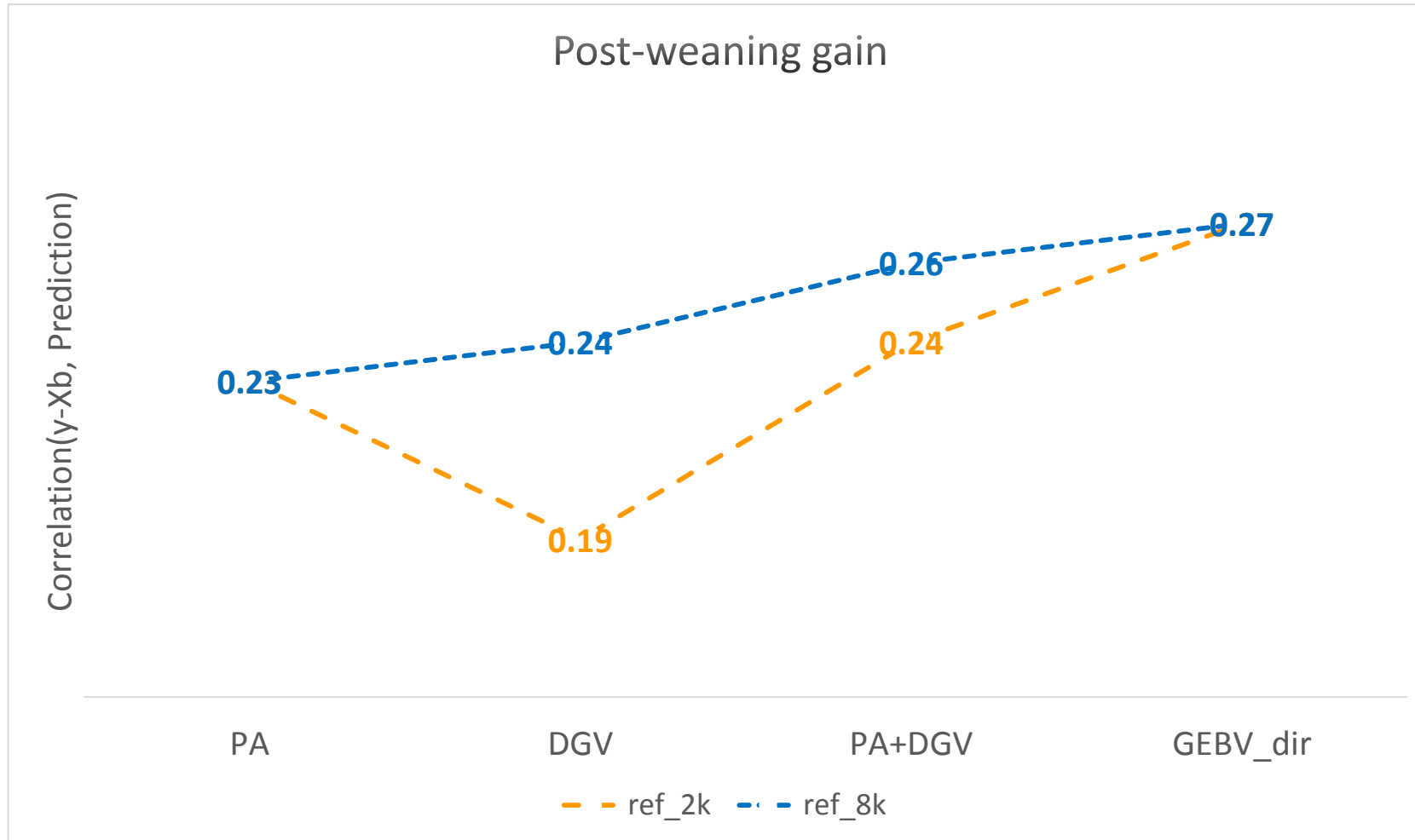
GS for zillions of genotypes?

- **INDIRECT Prediction**



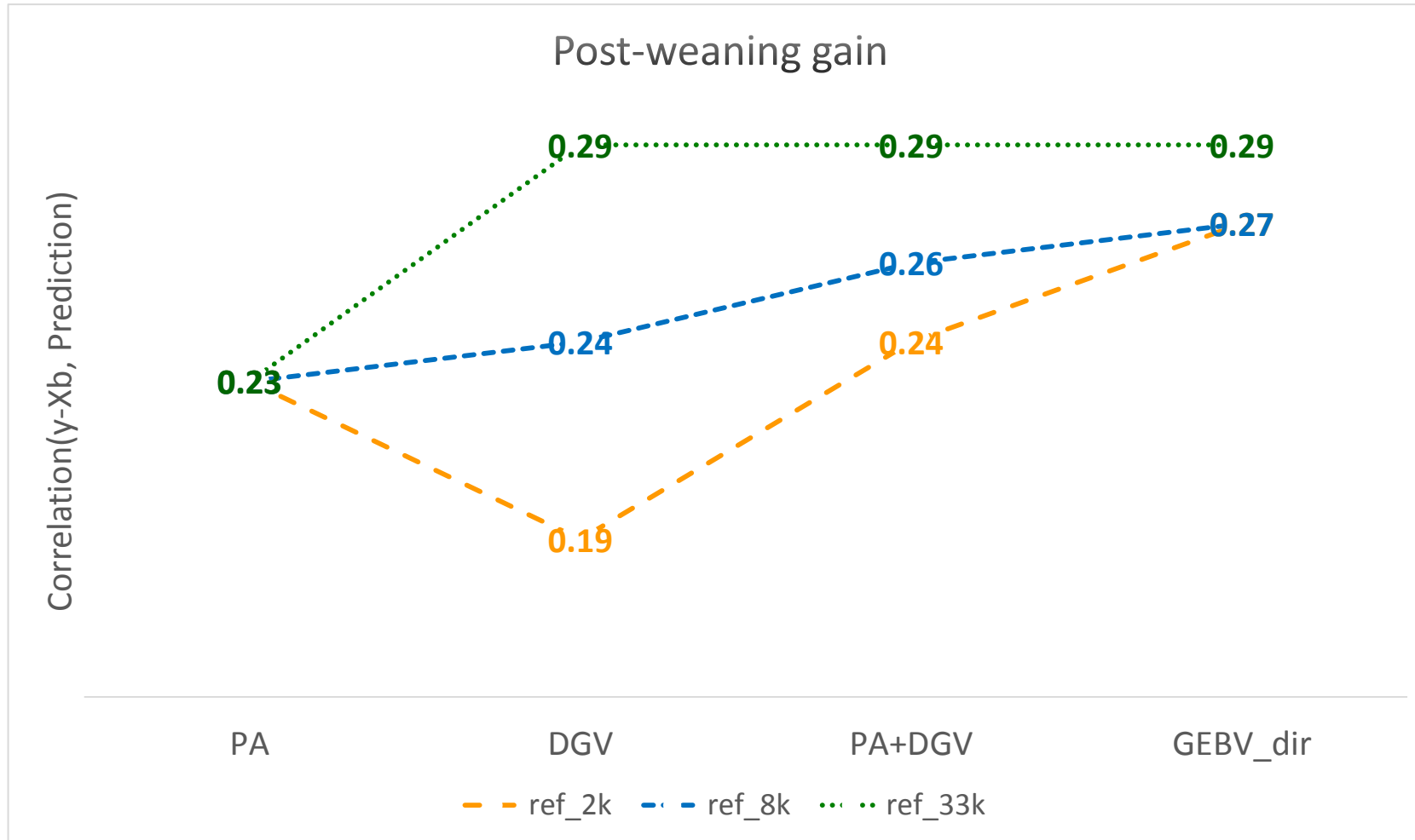
GS for zillions of genotypes?

- INDIRECT Prediction



GS for zillions of genotypes?

- INDIRECT Prediction



few genotyped
animals - Index

$$\text{GEBV} \approx w1 \text{ PA} + w4 \text{ DGV}$$

More genotyped
animals – only DGV

$$\text{GEBV} \approx w1 \text{ PA} + w4 \text{ DGV}$$

GS for zillions of genotypes?

TAKE HOME #4

- It is possible to run ssGBLUP for huge genotyped populations with the same accuracy as regular ssGBLUP
- APY with 10,000 core animals for cattle populations
- Young genotyped animals can have interim evaluations in ssGBLUP at a small computing cost

THANKS



- Shogo Tsuruta
- Breno Fragomeni
- Yutaka Masuda
- Ignacio Aguilar
- Andres Legarra
- Ivan Pocrnic