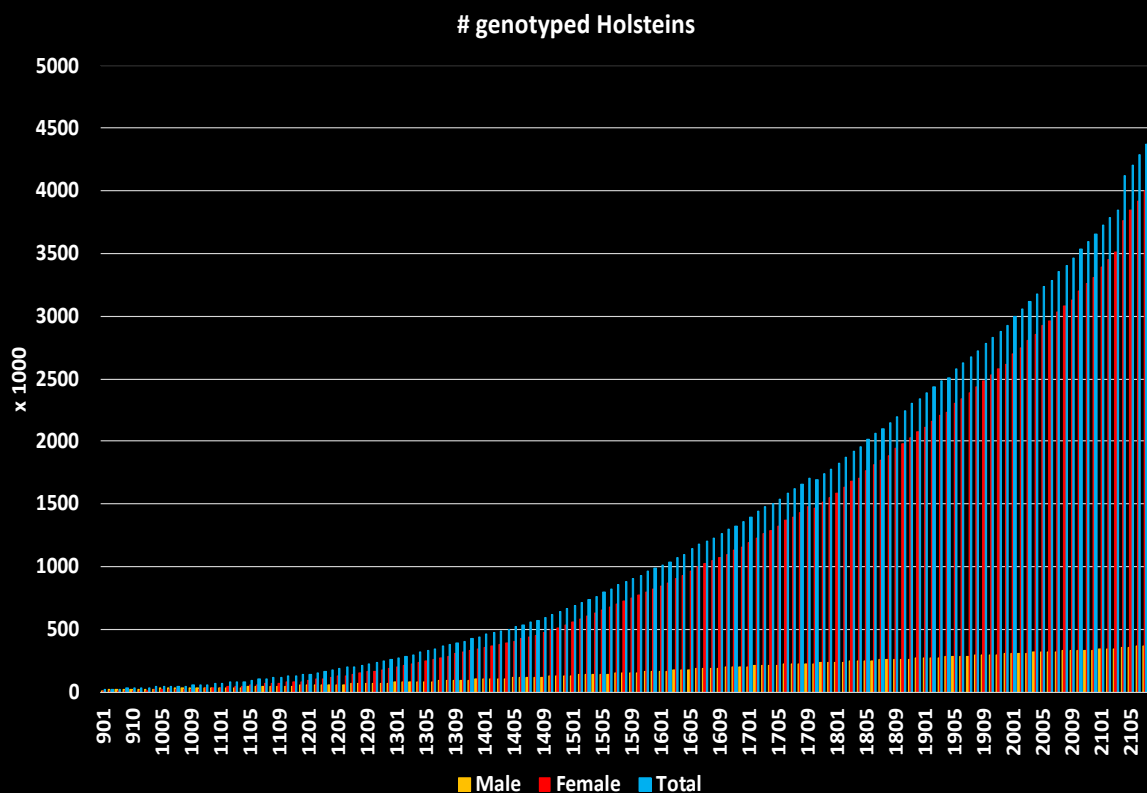


Are indirect genomic predictions a good option as the number of genotypes continues to rise?



Shogo Tsuruta, Daniela Lourenco, Yutaka Masuda, Ignacy Misztal - University of Georgia
Tom Lawlor - Holstein Association USA Inc.



Indirect **G**enomic **P**rediction
↓
Reducing Computational Cost
by
GEBV and SNP
for randomly selected genotyped
animals

Introduction

Computing cost ↑ # genotyped animals ↑

How to minimize the cost?

1. Get a high-speed computer or develop a faster computing program
2. Use APY with a small number of core animals
3. Remove old data
4. Calculate genomic predictions separately and indirectly
 - which animals?

Genotyped animals with no or less influence on other animals

- **target animals**

i.e., no progeny, no phenotypes

Data set

Phenotypes, pedigrees, and genotypes up to 2018
– full dataset for genomic prediction (benchmark)
10.3M records for 18 linear type traits for 13.6M animals
2.3M genotyped animals with 79K SNPs



Genomic prediction (GEBV)
for up to 13.6M animals using ssGBLUP with APY
except for **target genotyped animals**
(males: no progeny + females: no phenotypes)



Indirect **Genomic Prediction (IGP)**

Target animals for IGP - no progeny / no phenotypes

YOB	# genotyped animals for IGP (x 1000)			# genotyped animals for GEBV (x 1000)	# IGP in %
	Male	Female	Total		
2014 - 18	145	1304	1449	886	62
2015 - 18	118	1171	1289	1046	55
2016 - 18	89	1000	1089	1246	47
2017 - 18	58	707	765	1570	33
2018	23	257	280	2055	12
All	-	-	-	2335	-

Method

1. Genomic prediction (**GEBV**) for animals with progeny or phenotypes



2. **SNP effects**

$$\hat{\mathbf{u}} = \lambda \mathbf{DZ}'\mathbf{G}^{-1} \hat{\mathbf{a}}$$

($\hat{\mathbf{u}}$ = SNP effects, $\lambda = \frac{\sigma_u^2}{\sigma_a^2}$, $\mathbf{D} = \mathbf{I}$, $\mathbf{Z} = \text{\#animals} \times \text{\#SNP matrix}$, $\hat{\mathbf{a}} = \text{GEBV}$)

$\hat{\mathbf{u}}$ from $\hat{\mathbf{a}}$ for all genotyped animals up to each year group

or

$\hat{\mathbf{u}}$ from $\hat{\mathbf{a}}$ for randomly sampled genotyped animals



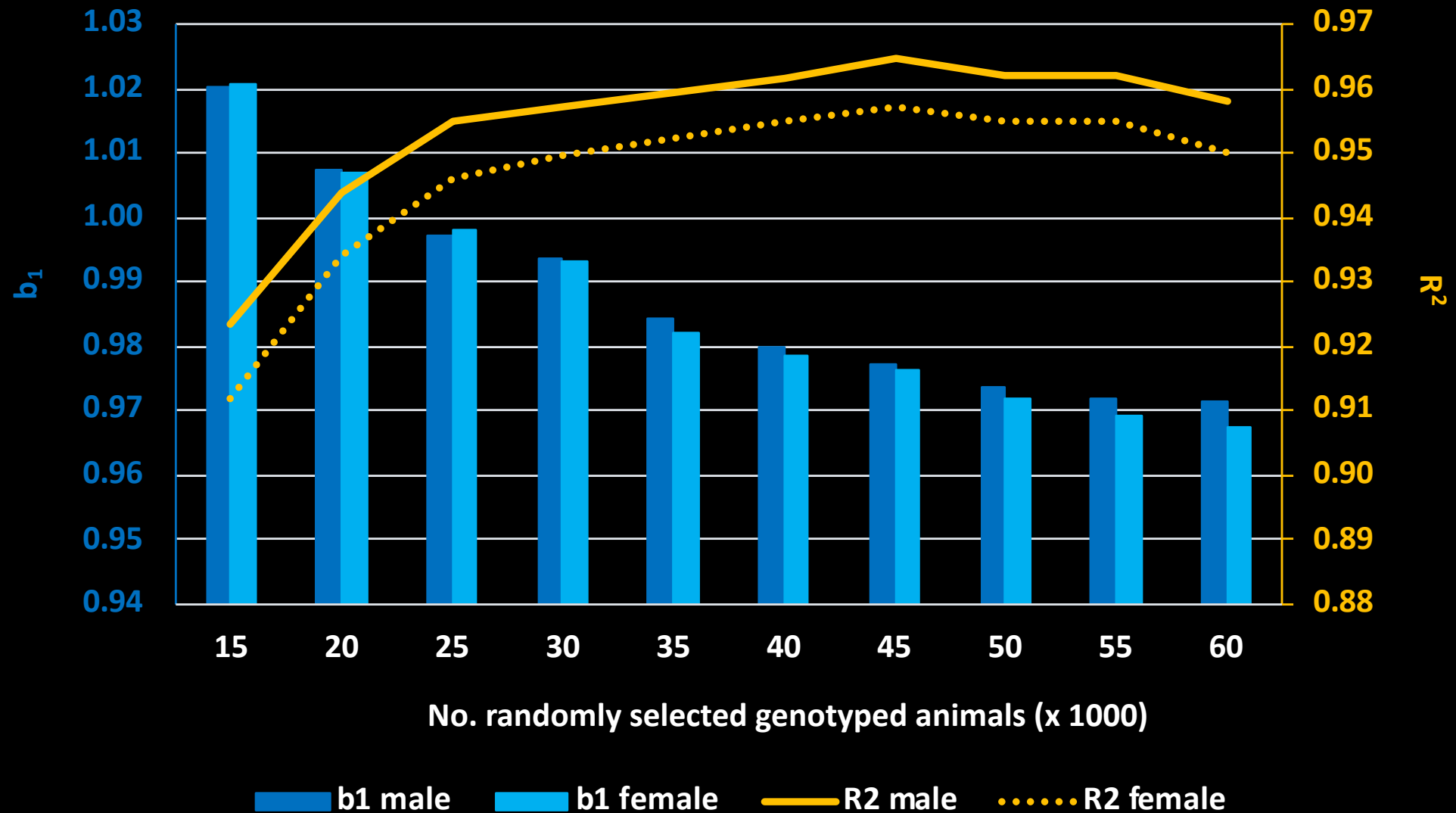
3. **Indirect Genomic Prediction (IGP)** = $\mathbf{Z}\hat{\mathbf{u}}$

$$\text{GEBV} = b_0 + b_1 \times \text{IGP using 2014 – 2018 IGP group}$$

	ΔG	b_0		b_1		R^2	
		Male	Female	Male	Female	Male	Female
Mean	1.11	1.29	1.21	0.99	0.99	0.950	0.949
SD	0.68	0.75	0.72	0.03	0.03	0.016	0.017
Correlation ($\Delta G, *$)	-	0.98	0.98	0.84	0.84	-0.57	-0.55

ΔG = Genetic progress / year

Randomly selected animals to obtain SNPs for 2014 – 2018 IGP



YOB	Computing time (hrs: wall clock)			No. iterations for GEBV
	GEBV	IGP (selected)	Total (selected)	
2014 - 18	72	27-5 (1)	99-77 (73)	1003
2015 - 18	106	35-8 (1)	141-114 (107)	1052
2016 - 18	118	45-10 (1)	163-128 (119)	1101
2017 - 18	140	60-12 (1)	200-152 (141)	1177
2018	153	70-13 (1)	223-166 (151)	1332
All	-	-	177	1433

Conclusions

- Indirect genomic prediction feasible in large-scale genomic evaluation
- Reducing the computing cost by Indirect Genomic Evaluation
- Significant reduction in computing cost by using randomly selected genotyped animals for IGP
- Indirect Genomic Predictions as accurate as GEBV



Acknowledgements

Holstein Association USA Inc.
and
Council on Dairy Cattle Breeding

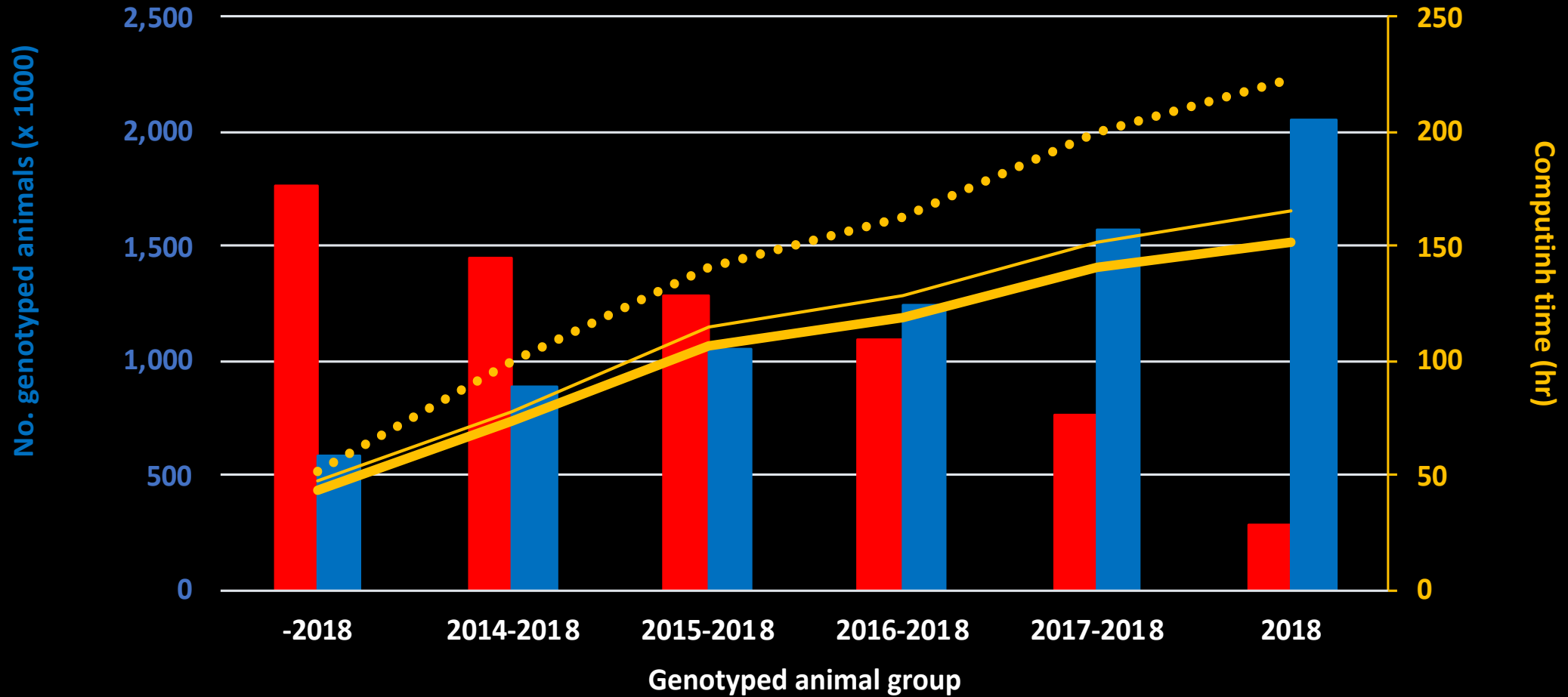
Anonymous reviewers in JDS Communications

Thank you



Graphical Summary

Computing Cost (hours)



■ No. genotyped animals for IGP

■ No. genotyped animals for GEBV

●●●● Cost with all genotyped animals

— Cost with selected genotyped animals

— Without constructing G-1 twice