7 December 2023

NAV blending and implementation

Aarhus University:	Ole Christensen, Emre Karaman, Ana Claudia Guillenea Golovanevsky
Luke:	Ismo Stranden, Esa Mäntysaari, Matti Taskinen, Hanni Kärkkäinen, Minna Kuivula, Andrei Kudinov, Hongding Gao, Timo Pitkänen; Martin Lidauer
Seges:	Ulrik Sander Nielsen, Kevin Byskov, Trine Andersen
Faba:	Jukka Pøsø
Vaxa:	Elisenda Rius-Vilarrasa

Invitation to NAV blending "TEAMS" meeting

Tuesday 12 December 09.30-12.00

1. Opening of the meeting and approval of the agenda
2. Approval of the minutes from last meeting
3. Single step
    a. Yield - handling bias – post processing (Jukka, Minna and Ulrik)
4. Dairyxdairy genomic prediction follow up from last meeting (Huiming)
5. Beefxdairy genomic prediction (AU, Seges)
6. Other matters

Best regards
Gert

Experiences from official genomic evaluations in DxD

Huiming Liu

2023-12-12

Projekt: DairyCross

STØTTET AF

Background

- In the case of crossbred animals, the same marker allele may have a different effect according to breed origin of the allele (BOA).
- Utilize BOM method and software developed by Jon (AU) in official NAV DxD genomic evaluation
- Official NAV genomic evaluation of DxD based on SNP solutions from genomic breeding value estimation of pure breeds are calculated monthly.
- It is possible to assess and rank crossbred animals within a herd using genomic values for the Nordic Total Merit (NTM) ranking system.

SEGES INNOVATION

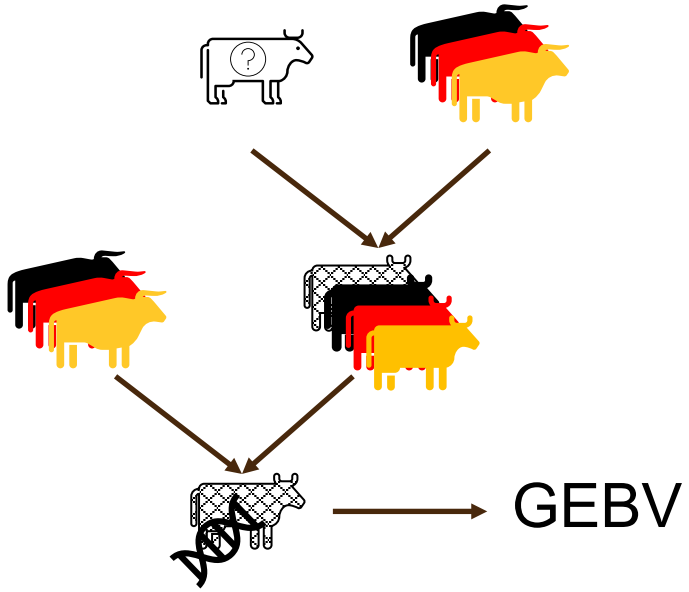
Steps

- Extraction of genotypes for crossbred animals and their ancestors
- Genotype imputation and phasing
- Assign BOA in crossbreds using AllOr
- Breed scaling
- Calculate genomic breeding values
- Postprocessing and standardization
- Calculation of NTM

SEGES INNOVATION

Extraction of the genotypes

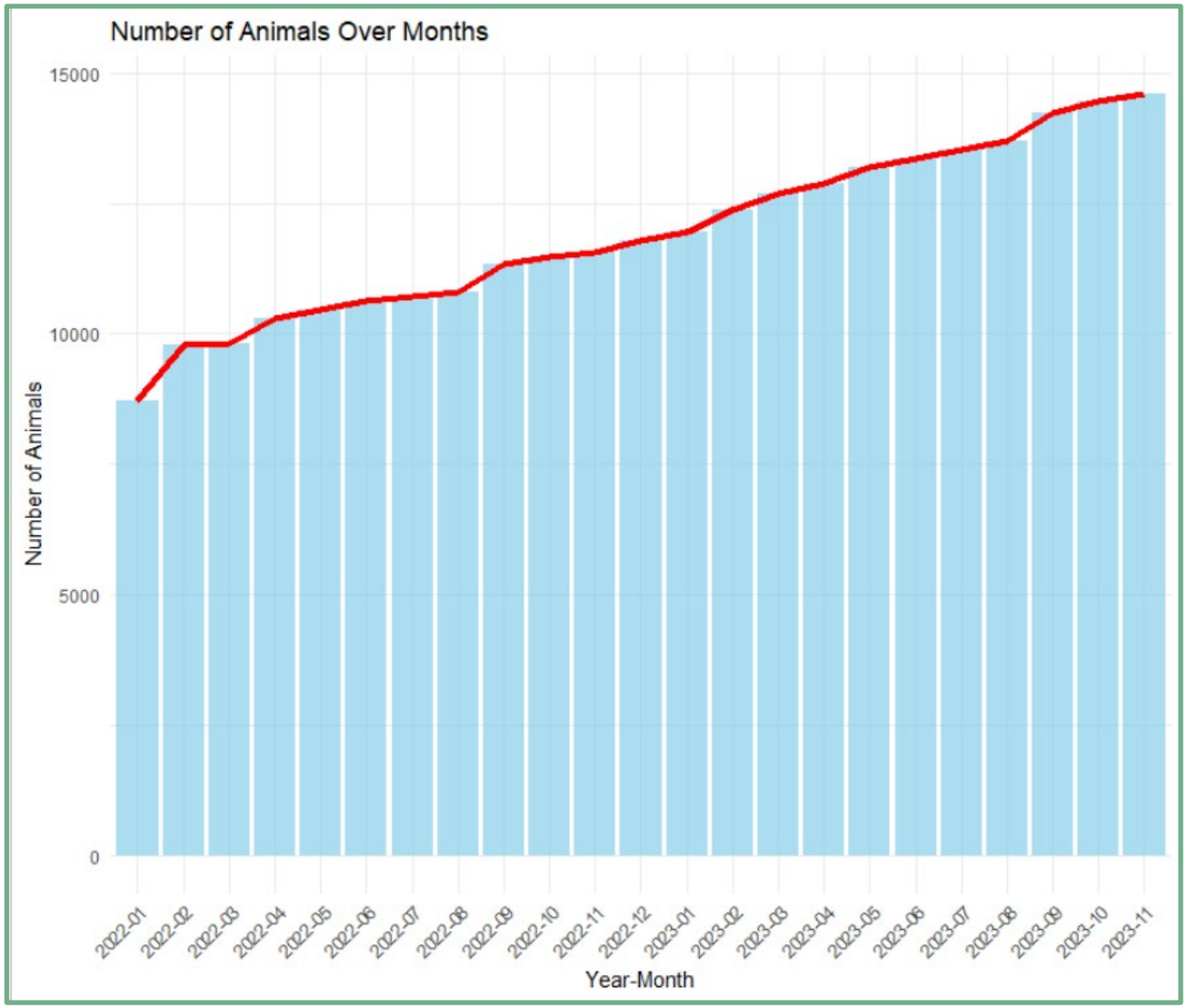
- XXX females with
  - HOL, JER or RDC sire and mgs
  - HOL, JER, RDC or XXX dam
  - excluding other breeds used for cross breeding (BSW, FLE, SIM, etc.)
  - Trace the pedigree for 5 generations



- Not possible to include MON crosses using the current method.

SEGES INNOVATION

Number of genotyped animals XXX over time



Month	Number of XXX in each country		
	DNK	SWE	FIN
Jan 22	7766	837	0
Nov 23	12443	2163	0

SEGES INNOVATION

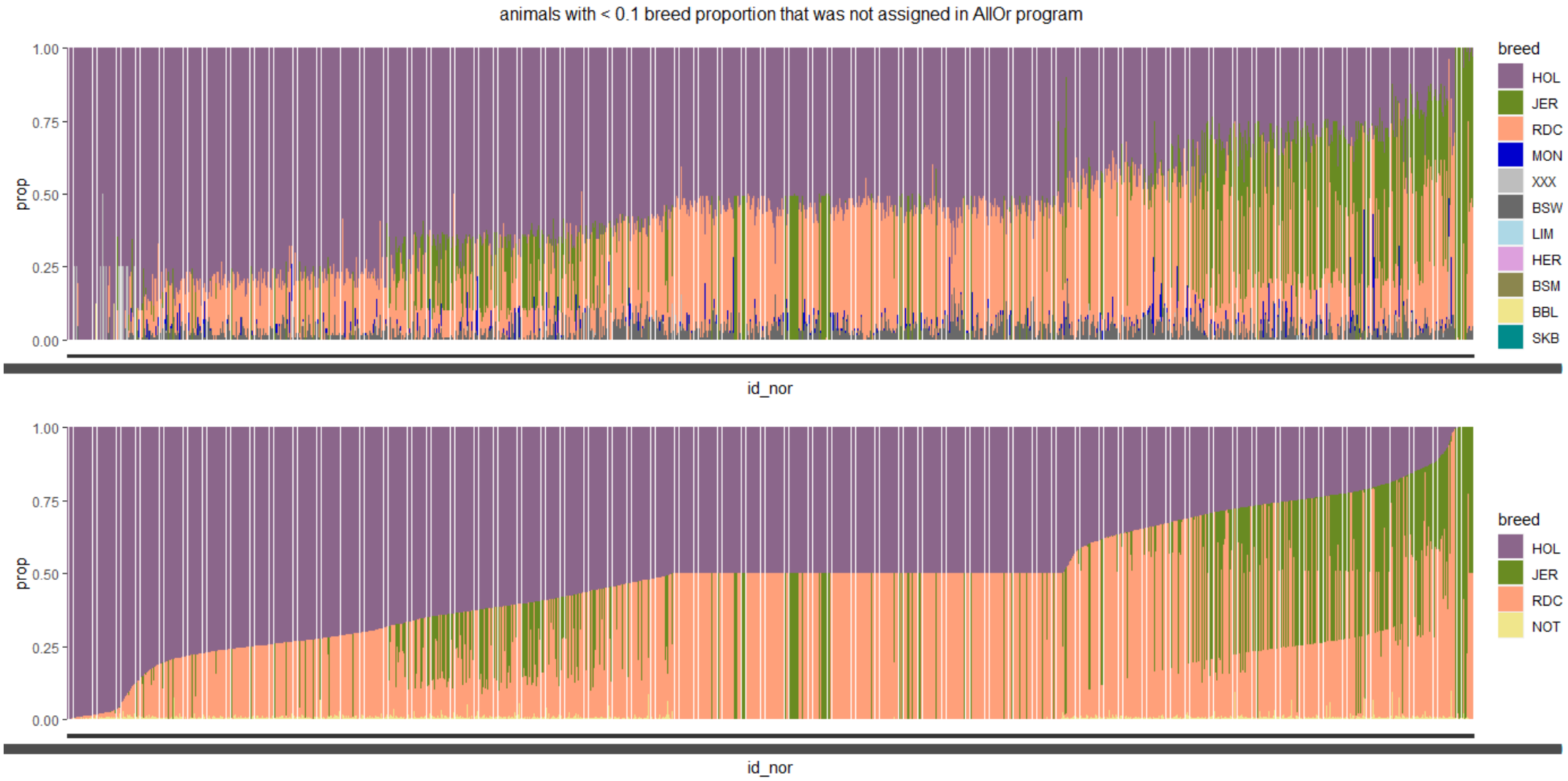
Imputation and breed origin of alleles detection

- Split of genotypes into 2 haplotypes – Fimpute v2.2
  - Input for the AllOr program for BOA detection
  - Super map – 47 586 markers
  - Requires complete imputation and phasing
  - 2 alleles were assigned to haplotypes randomly if Fimpute didn't identify the phase

- BOA detection
  - AllOr - designed to detect BOA in genotypes of crossbred animals from medium density SNP chips
  - Sire is known and of a purebred known breed, as in typical rotational crossbreeding
  - Genotypes of representative samples of all contributing pure breeds are required.
    - XXX's genotyped ancestors (tracing 5 generations)

SEGES INNOVATION

Breed proportions (pedigree vs AllOr)

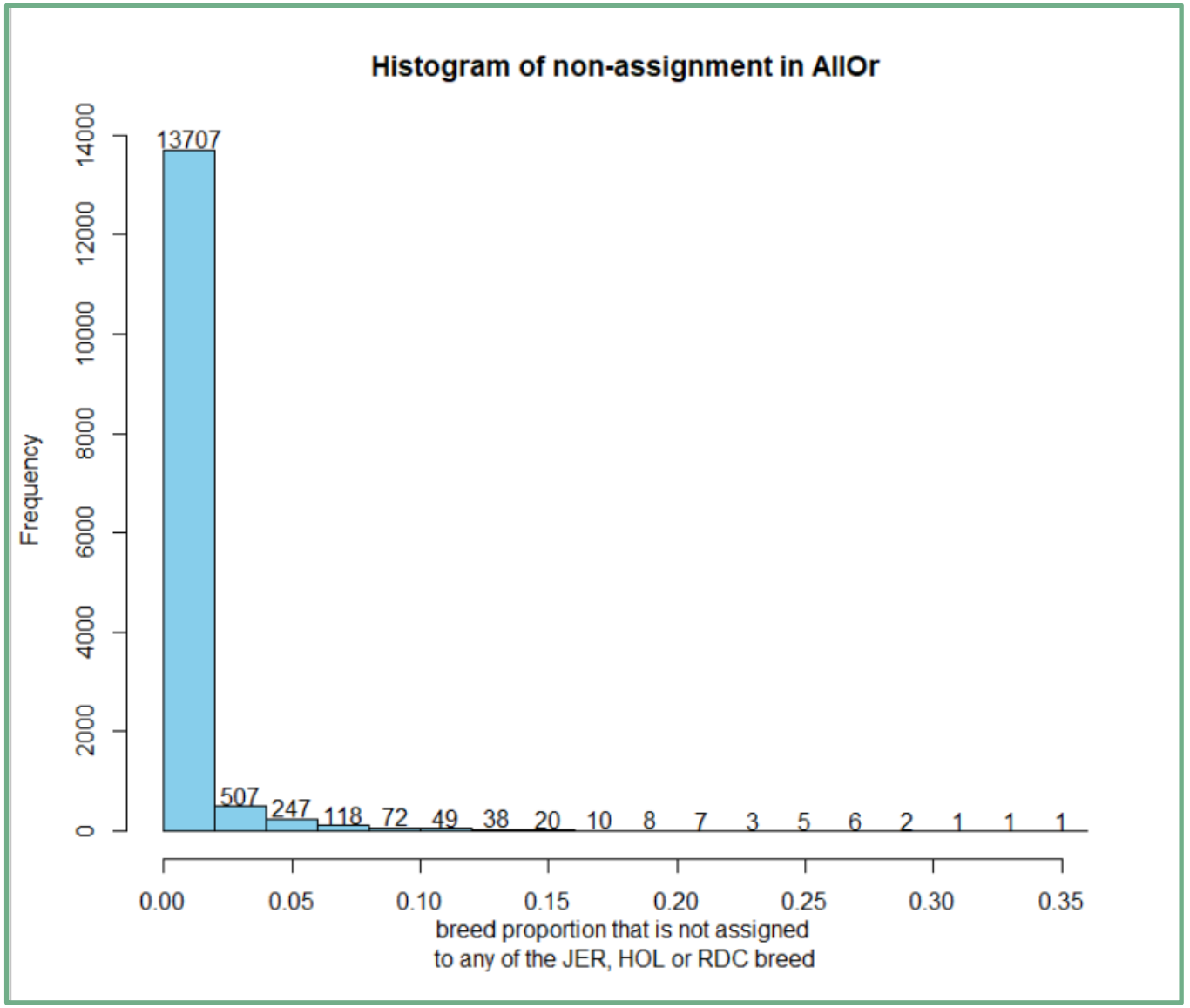


Pedigree

AllOr

SEGES INNOVATION

Histogram of non-assignment in AllOr (Nov 23)



SEGES INNOVATION

Jan 22 vs Nov 23 (AllOr)

Month	Breed proportion		
	HOL	JER	RDC
Jan 22	0.508	0.134	0.350
Nov 23	0.527	0.135	0.329

SEGES INNOVATION

Traits included

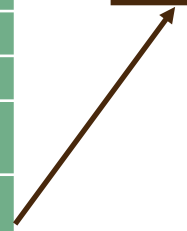
14 main traits

Yield (milk, fat and protein)
Growth
Fertility
Birth
Calving
Udder health
General health
Milkability
Temperament
Longevity
Claw health
Saved feed

23 type traits

1. Stature
2. Body depth
3. Chest width
4. Dairy form
5. Top line
6. Rump width
7. Rump angle
8. Rear legs, side view
9. Rear legs, back rear view
10. Hock quality
11. Bone quality
12. Foot angle
14. Fore udder attachment
15. Rear udder height
16. Rear udder width
17.Udder cleft/support
18. Udder depth
19. Teat length
20. Teat thickness
21. Teat placement (front)
22. Teat placement (back)
23. Udder balance

Frame
Feet & legs
Udder



SEGES INNOVATION

BOM model

Intercept

Snp solutions

polygenic

- $GEBV_{BOM,i} = \sum_{b=1}^{N_b} \mu_b \frac{\sum s_{1,i,b} + \sum s_{2,i,b}}{2m} + \sum_{b=1}^{N_b} (\boldsymbol{v}'_{\boldsymbol{b}}(\boldsymbol{w}_{\boldsymbol{i,1}} \circ \boldsymbol{s}_{\boldsymbol{1,i,b}}) + \boldsymbol{v}'_{\boldsymbol{b}}(\boldsymbol{w}_{\boldsymbol{i,2}} \circ \boldsymbol{s}_{\boldsymbol{2,i,b}})) + a_i$
  - μ_b is the intercept, accounting for difference in breed averages for breed *b* (phenotype average- average DGV)
  - $\boldsymbol{s}_{\boldsymbol{j,i,b}}$ is a vector of breed of origin indication for allele *j* of animal *i* to breed *b*, with 1 for alleles assigned to breed *b* and 0 for alleles assigned to other breeds and proportional values for alleles that could not be assigned
  - *m* is the number of markers
  - $\boldsymbol{v}_{\boldsymbol{b}}$ is a vector of marker effects for breed *b*
  - where $\boldsymbol{w}_{\boldsymbol{i,j}}$ contains haplotype *j* coded as 0 and 1 for the alternative alleles
  - $\circ$ is element vise multiplication
  - a_i is a residual polygenic effect

SEGES INNOVATION

BOM model

Intercept

Snp solutions

polygenic

- $GEBV_{BOM,i} = \sum_{b=1}^{N_b} \mu_b \frac{\sum s_{1,i,b} + \sum s_{2,i,b}}{2m} + \sum_{b=1}^{N_b} (\boldsymbol{v}'_b(\boldsymbol{w}_{i,1} \circ \boldsymbol{s}_{1,i,b}) + \boldsymbol{v}'_b(\boldsymbol{w}_{i,2} \circ \boldsymbol{s}_{2,i,b})) + a_i$

Scaled to phenotypic level

- μ_b is the intercept, accounting for difference in breed averages for breed *b* (phenotype average- average DGV)
- $\boldsymbol{s}_{j,i,b}$ is a vector of breed of origin indication for allele *j* of animal *i* to breed *b*, with 1 for alleles assigned to breed *b* and 0 for alleles assigned to other breeds and proportional values for alleles that could not be assigned
- *m* is the number of markers
- $\boldsymbol{v}_b$ is a vector of marker effects for breed *b*
- where $\boldsymbol{w}_{i,j}$ contains haplotype *j* coded as 0 and 1
- ∘ is element vise multiplication
- a_i is a residual polygenic effect

Official evaluation is on breed specific index scales and needed therefore to be converted to kg PY for example to make them comparable across breeds.

The SNP solutions were therefore * phenotypic effect (kg PY) of +1 index unit in each of the breeds (2.0, 1.7 and 2.1)

Polygenic effects –> phenotypic level

GEBVs at the phenotypic level

SEGES INNOVATION

Postprocessing and standardization of sub traits

- Rescaling GEBV back to index scale by dividing GEBVs by phenotypic effect of +1 index unit of HOL ($GEBV_{index}$).
- Rolling base for mean (MEAN).
  - XXX genotyped animals 1-7 years of age at the date of publication.
- Final XXX-GEBV = ($GEBV_{index}$ – MEAN)*HOL standardization factor+100
- HOL weight factors are used to calculate composite indices for Yield, Frame, F&L and Udder
- NTM

SEGES INNOVATION

Problem of inconsistency of DxD over months

Oct 22

The FREQ Procedure

diff_HOL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
-0.08	1	0.01	1	0.01
-0.06	3	0.03	4	0.03
-0.05	1	0.01	5	0.04
-0.04	8	0.07	13	0.11
-0.03	24	0.21	37	0.32
-0.02	60	0.52	97	0.85
-0.01	276	2.41	373	3.26
0	7916	69.10	8289	72.36
0.01	2713	23.68	11002	96.04
0.02	381	3.33	11383	99.36
0.03	54	0.47	11437	99.83
0.04	7	0.06	11444	99.90
0.05	6	0.05	11450	99.95
0.06	3	0.03	11453	99.97
0.07	2	0.02	11455	99.99
0.09	1	0.01	11456	100.00

Many animals have a different estimation of breed origin of alleles between different months

SEGES INNOVATION

Problem of inconsistency

Some animals have extreme changes in GEBV from month to month!

Oct 22

Obs	dif	dtm	dgebvredyld	dgebvredgwth
1	-11	1	1	.
2	-10	1	1	1
3	-9	1	.	.
4	-8	3	5	1
5	-7	7	6	.
6	-6	23	21	3
7	-5	60	51	9
8	-4	176	128	36
9	-3	444	404	130
10	-2	1254	1097	592
11	-1	2521	2488	2279
12	0	3239	3537	5031
13	1	2163	2237	2505
14	2	986	932	573
15	3	312	286	124
16	4	98	87	27
17	5	22	27	10
18	6	12	14	5
19	7	4	4	.
20	8	1	1	.
21	9	.	.	1
22	10	.	1	.
23	11	.	.	1
24	13	.	.	.

SEGES INNOVATION

Improve performance of imputation in DxD (Grum)

- Imputation: essential in official genomic evaluation
- In DxD, imputation jointly with purebred ancestors
- Inconsistencies across run (months) in GEBVs & NTM

Primarily due to performance of Imputation

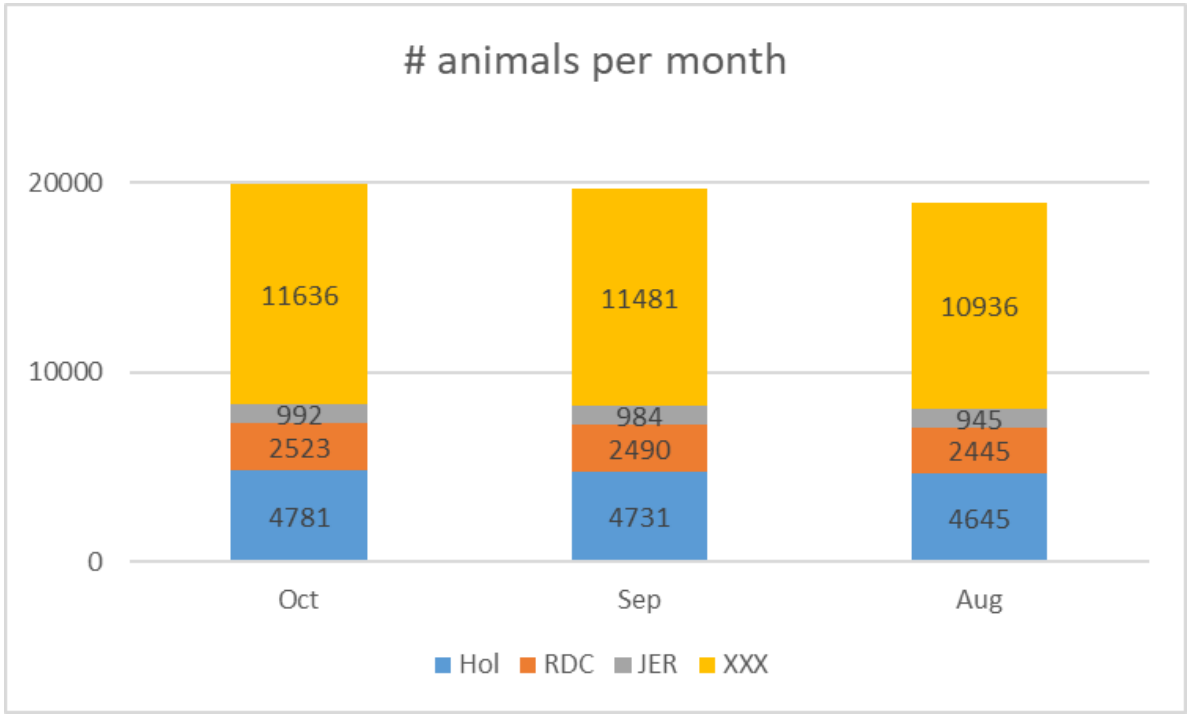
- Reference: Size & relatedness with target
- Marker: SNP panel (and target density), marker frequency,..
- Imputation tools: Methods and programs
- The extent and pattern of linkage disequlibrium differ in crossbred vs purebred animals

Purpose: Improve consistencies for DxD

SEGES INNOVATION

Initial checks

Three reference months to compare imputation and GEBV consistency



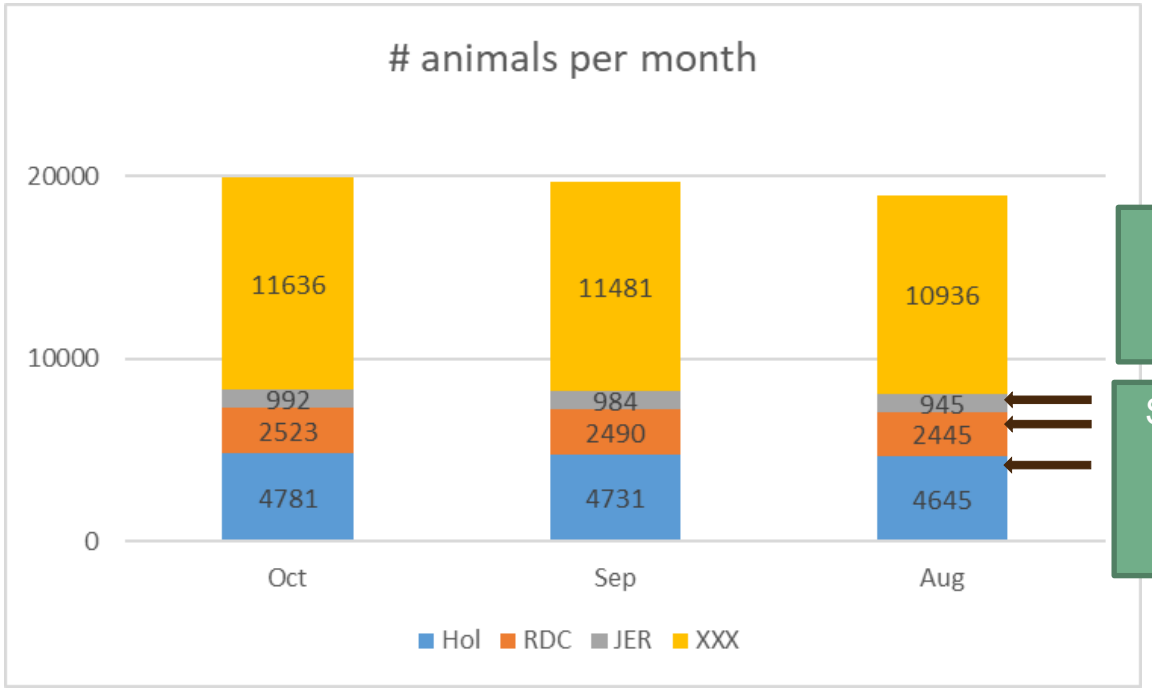
One-step joint Pure-Cross imputation

Fimpute v2.2

SEGES INNOVATION

New strategy

- Step 1: within-breed Purebred imputation +
  addition of 20K more individuals per breed
- Step 2: pure-cross imputation (limitation solved in FImpute3)
  - Purebred animals as reference
  - crossbreds animals as testing



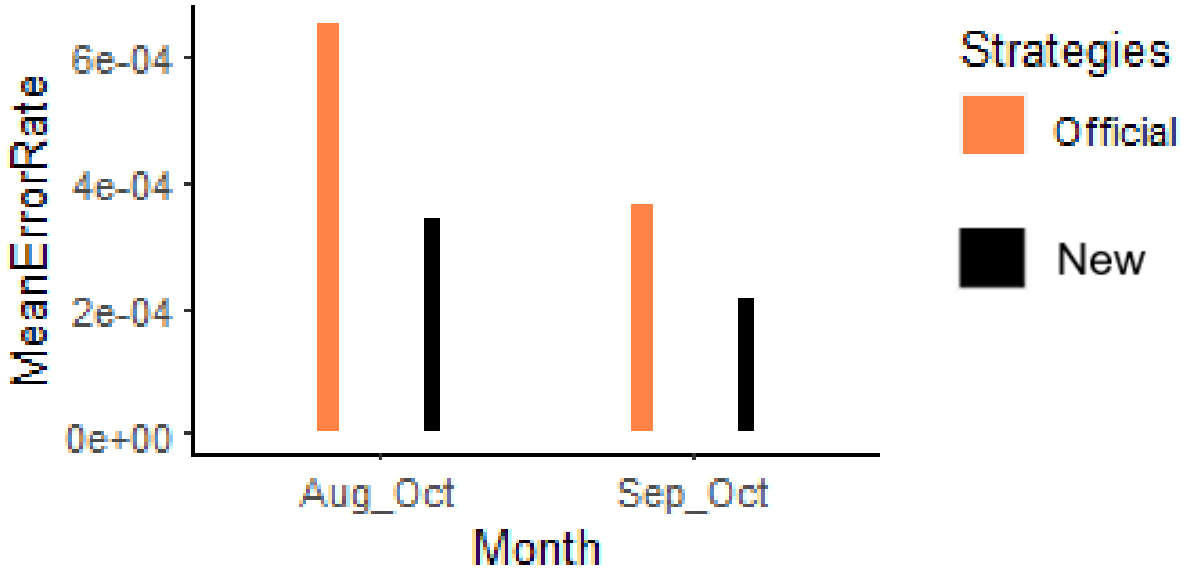
Step2: Impute cross with Imputed RDC+HOL+JER (No XXX in ref)

Step1: 20K more individuals per breed
+ Impute each pure breed separately

SEGES INNOVATION

Error – changes compared to Oct 22

- Step 1: within-breed Purebred imputation with addition of 20K more individuals per breed
- Step 2: pure-cross imputation
  - Purebred animals as reference
  - crossbreds animals as testing



SEGES INNOVATION

Stability of GEBV

- correlation between Aug22 and Sep22

	Official	New
NTM	0.980	0.989
Yield	0.978	0.989
Calving	0.971	0.986
Matitis	0.973	0.985
Health	0.978	0.985

New strategy:

No extreme changes in breed proportions
No extreme changes in GEBV

SEGES INNOVATION

Conclusions

- Two-step imputation where first purebreds are imputed alone in a within-breed setting
- Reference population composition and size matter: more purebred, avoid crossbreds
- New strategy was implemented from June 23

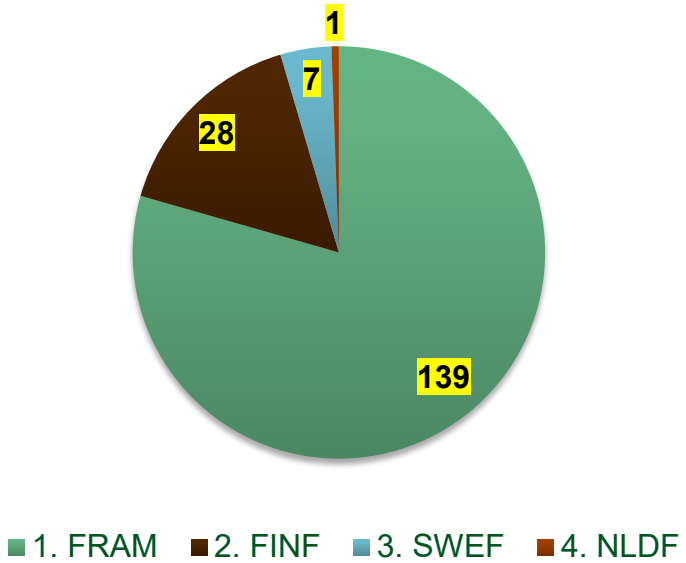
SEGES INNOVATION

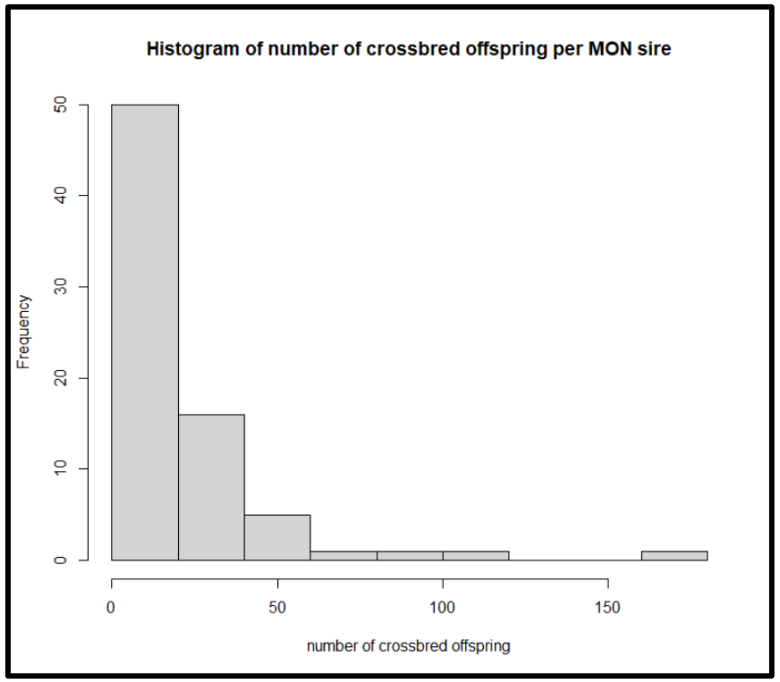
Current work

- Including crossbreds' own phenotypes in GEBV calculation (DONE)
- Test for the calculation of GEBV for MON crossbreds (DONE) - Emre's talk

- Implement genomic evaluation for MON crosses

Number of MON animals





There are 1558 genotyped XXX animals with a MON sire

There are 8 sires with more than 50 crossbred offspring

SEGES INNOVATION