

Estimation of metafounder relationships for multibreed single-step genomic evaluation of Finnish beef breeds

Matti Taskinen, Timo J. Pitkänen, Anna-Maria Leino, Esa A.
Mäntysaari, Minna Koivula, Andrei Kudinov, Ismo Strandén
Natural Resources Institute Finland (Luke)

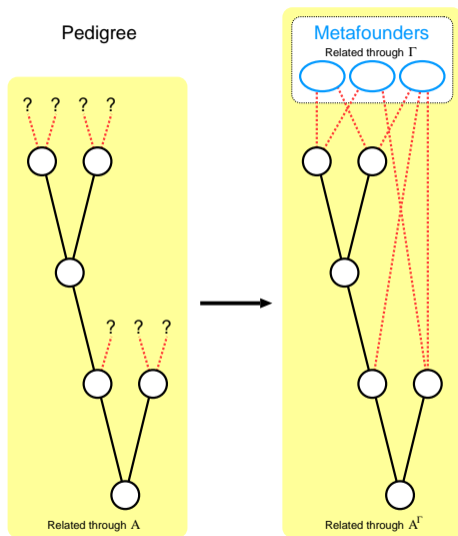
Interbull 2023, Aug. 26, 2023

Background

- Finnish multibreed beef evaluation.
- Applying **metafounder** concept.
- Estimating **metafounder relations** from genomic information.

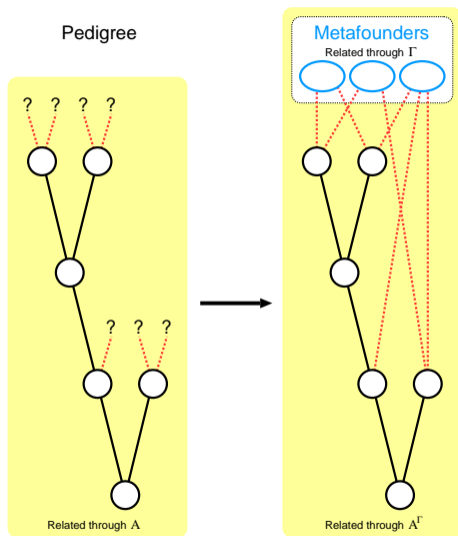
Metafounders: Supplementing pedigree with genomic data

- **Metafounders** concept: unknown parents in pedigree replaced with “metafounders”.
- Metafounders assumed to be related through Γ matrix that is estimated from **genomic data**.
- Goal: make pedigree relations closer to genomic relations.



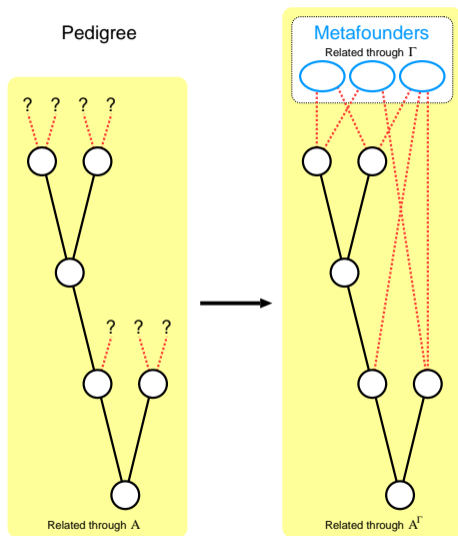
Metafounders: Supplementing pedigree with genomic data

- **Metafounders** concept: unknown parents in pedigree replaced with “metafounders”.
- Metafounders assumed to be related through Γ matrix that is estimated from **genomic data**.
- Goal: make pedigree relations closer to genomic relations.



Metafounders: Supplementing pedigree with genomic data

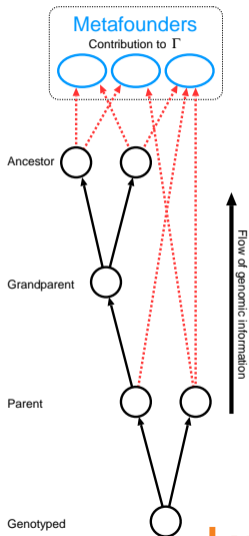
- **Metafounders** concept: unknown parents in pedigree replaced with “metafounders”.
- Metafounders assumed to be related through Γ matrix that is estimated from **genomic data**.
- Goal: make pedigree relations closer to genomic relations.



Metafounders: Estimating MF relations (Γ matrix)

Metafounder relations estimated from pedigree using base population allele frequencies of MFs:

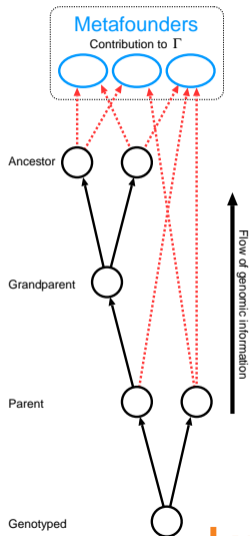
- Genomic information “inherited” from genotyped animals to their ancestors.
- MFs obtain contribution of genomic information through missing parents in pedigree.
- Allele frequencies estimated (Garcia-Baccino et al. 2017) using in-house `Bpop` program.
- Γ **matrix** calculated from allele frequencies (Legarra et al. 2014).



Metafounders: Estimating MF relations (Γ matrix)

Metafounder relations estimated from pedigree using base population allele frequencies of MFs:

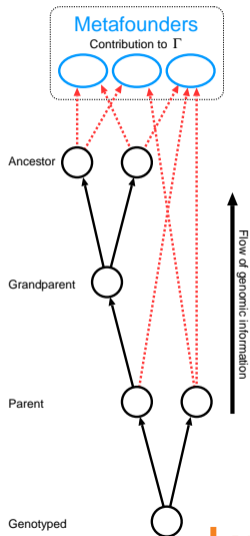
- Genomic information “inherited” from genotyped animals to their ancestors.
- MFs obtain contribution of genomic information through missing parents in pedigree.
- Allele frequencies estimated (Garcia-Baccino et al. 2017) using in-house `Bpop` program.
- Γ **matrix** calculated from allele frequencies (Legarra et al. 2014).



Metafounders: Estimating MF relations (Γ matrix)

Metafounder relations estimated from pedigree using base population allele frequencies of MFs:

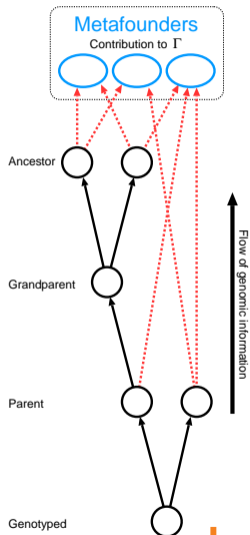
- Genomic information “inherited” from genotyped animals to their ancestors.
- MFs obtain contribution of genomic information through missing parents in pedigree.
- Allele frequencies estimated (Garcia-Baccino et al. 2017) using in-house Bpop program.
- Γ matrix calculated from allele frequencies (Legarra et al. 2014).



Metafounders: Estimating MF relations (Γ matrix)

Metafounder relations estimated from pedigree using base population allele frequencies of MFs:

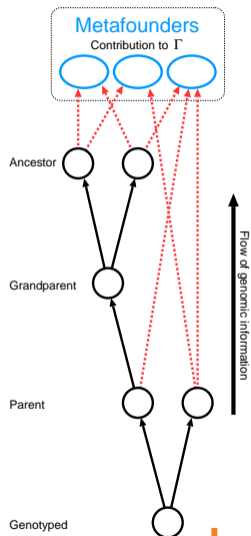
- Genomic information “inherited” from genotyped animals to their ancestors.
- MFs obtain contribution of genomic information through missing parents in pedigree.
- Allele frequencies estimated (Garcia-Baccino et al. 2017) using in-house B_{pop} program.
- Γ matrix calculated from allele frequencies (Legarra et al. 2014).



Metafounders: Estimating MF relations (Γ matrix)

Metafounder relations estimated from pedigree using base population allele frequencies of MFs:

- Genomic information “inherited” from genotyped animals to their ancestors.
- MFs obtain contribution of genomic information through missing parents in pedigree.
- Allele frequencies estimated (Garcia-Baccino et al. 2017) using in-house B_{pop} program.
- Γ **matrix** calculated from allele frequencies (Legarra et al. 2014).



Matching Unknown Parent Groups (UPG) with metafounders (MF)

In this project: MFs needed to match Unknown Parent Groups of old evaluations:

- **Pure breed** groups for 5 common breeds.
- Separate **domestic** and **foreign** groups.
- Dairy animals in beef pedigree (mk).
- Others (ot):
 - ▶ All crossbreds.
 - ▶ Purebreds of other breeds.
- Many year classes.
- Altogether 115 MFs needed.

Group	Breed	Country
1	hf	domestic
2	hf	foreign
3	aa	domestic
4	aa	foreign
5	ch	domestic
6	ch	foreign
7	li	domestic
8	li	foreign
9	si	domestic
10	si	foreign
11	mk	*
12	ot	*

Matching Unknown Parent Groups (UPG) with metafounders (MF)

In this project: MFs needed to match Unknown Parent Groups of old evaluations:

- **Pure breed** groups for 5 common breeds.
- Separate **domestic** and **foreign** groups.
- Dairy animals in beef pedigree (mk).
- Others (ot):
 - ▶ All crossbreds.
 - ▶ Purebreds of other breeds.
- Many year classes.
- Altogether 115 MFs needed.

Group	Breed	Country
1	hf	domestic
2	hf	foreign
3	aa	domestic
4	aa	foreign
5	ch	domestic
6	ch	foreign
7	li	domestic
8	li	foreign
9	si	domestic
10	si	foreign
11	mk	*
12	ot	*

Matching Unknown Parent Groups (UPG) with metafounders (MF)

In this project: MFs needed to match Unknown Parent Groups of old evaluations:

- **Pure breed** groups for 5 common breeds.
- Separate **domestic** and **foreign** groups.
- Dairy animals in beef pedigree (mk).
- Others (ot):
 - ▶ All crossbreds.
 - ▶ Purebreds of other breeds.
- Many year classes.
- Altogether 115 MFs needed.

Group	Breed	Country
1	hf	domestic
2	hf	foreign
3	aa	domestic
4	aa	foreign
5	ch	domestic
6	ch	foreign
7	li	domestic
8	li	foreign
9	si	domestic
10	si	foreign
11	mk	*
12	ot	*

Matching Unknown Parent Groups (UPG) with metafounders (MF)

In this project: MFs needed to match Unknown Parent Groups of old evaluations:

- **Pure breed** groups for 5 common breeds.
- Separate **domestic** and **foreign** groups.
- Dairy animals in beef pedigree (mk).
- Others (ot):
 - ▶ All crossbreds.
 - ▶ Purebreds of other breeds.
- Many year classes.
- Altogether 115 MFs needed.

Group	Breed	Country
1	hf	domestic
2	hf	foreign
3	aa	domestic
4	aa	foreign
5	ch	domestic
6	ch	foreign
7	li	domestic
8	li	foreign
9	si	domestic
10	si	foreign
11	mk	*
12	ot	*

Matching Unknown Parent Groups (UPG) with metafounders (MF)

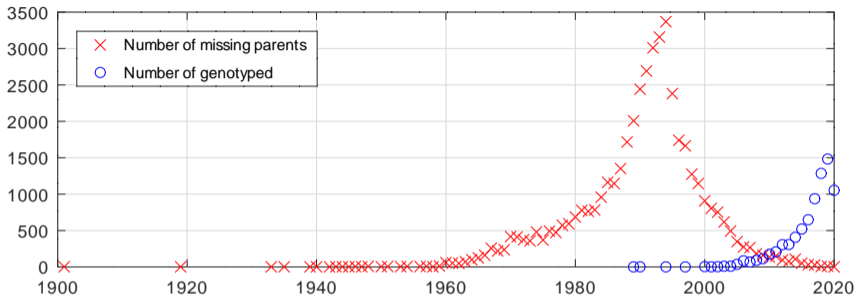
In this project: MFs needed to match Unknown Parent Groups of old evaluations:

- **Pure breed** groups for 5 common breeds.
- Separate **domestic** and **foreign** groups.
- Dairy animals in beef pedigree (mk).
- Others (ot):
 - ▶ All crossbreds.
 - ▶ Purebreds of other breeds.
- Many year classes.
- Altogether 115 MFs needed.

Unknown Parent Groups

hf	domestic	1980, 1985, 1990, 1995, 2000, 2005, 2010, 2015
hf	foreign	1980, 1984, 1989, 1994, 1999, 2004, 2009
aa	domestic	1980, 1986, 1991, 1996, 2001, 2006, 2011, 2016
aa	foreign	1980, 1984, 1989, 1994, 1999, 2004, 2009
ch	domestic	1980, 1985, 1990, 1995, 2000, 2005, 2010, 2015
ch	foreign	1980, 1981, 1986, 1991, 1996, 2001, 2006, 2011
li	domestic	1989, 1994, 1999, 2004, 2009, 2014
li	foreign	1980, 1983, 1988, 1993, 1998, 2003, 2008, 2013
si	domestic	1991, 1996, 2001, 2006, 2011, 2016
si	foreign	1980, 1981, 1986, 1991, 1996, 2001, 2006, 2011
rdc	*	1980, 1983, 1988, 1993, 1998, 2003, 2008
fic	*	1980, 1984, 1989, 1994, 1999, 2004
hol	*	1980, 1984, 1989, 1994, 1999, 2004, 2009
ot	*	1980, 1984, 1989, 1994, 1999, 2004, 2009

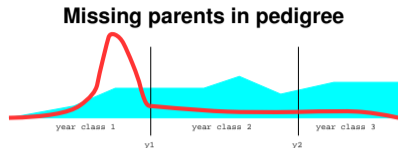
Metafounder challenge: Estimating Γ to the past



- Genotyped animals (blue) mostly from last few years (2018-2023).
- Metafounder relationships needed even before 1990.
- Base population relations need to be estimated 30 years back in time.

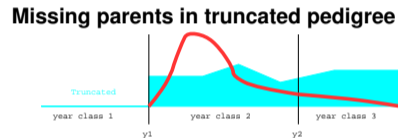
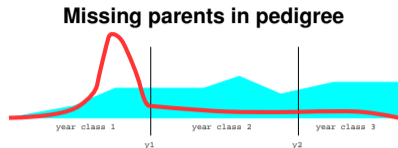
Challenge: Controlling distribution of missing parents

- Some year classes don't have enough missing parents to get genomic contributions
- When truncating pedigree, contribution from domestic offspring can “leak” to foreign MF.
- Similarly, crossbred offspring quite often “looses” genomic contribution to purebred MFs.
- Solution: different **pedigree truncation strategies** applied to control distribution of “missing” parents.



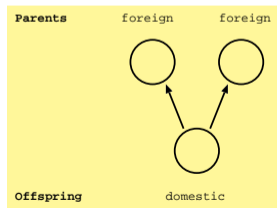
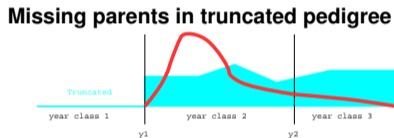
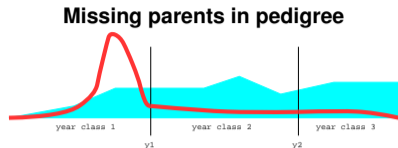
Challenge: Controlling distribution of missing parents

- Some year classes don't have enough missing parents to get genomic contributions \Rightarrow **pedigree truncation**.
- When truncating pedigree, contribution from domestic offspring can “leak” to foreign MF.
- Similarly, crossbred offspring quite often “looses” genomic contribution to purebred MFs.
- Solution: different **pedigree truncation strategies** applied to control distribution of “missing” parents.



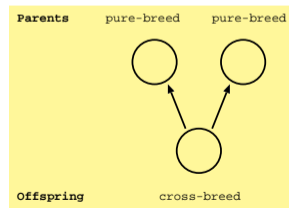
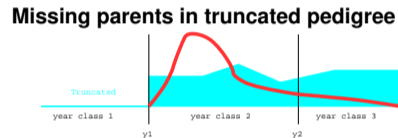
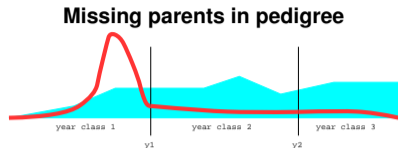
Challenge: Controlling distribution of missing parents

- Some year classes don't have enough missing parents to get genomic contributions \Rightarrow **pedigree truncation**.
- When truncating pedigree, contribution from domestic offspring can “leak” to foreign MF.
- Similarly, crossbred offspring quite often “looses” genomic contribution to purebred MFs.
- Solution: different **pedigree truncation strategies** applied to control distribution of “missing” parents.



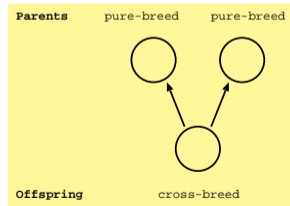
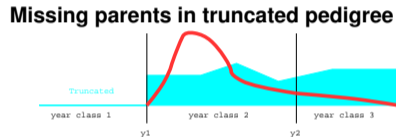
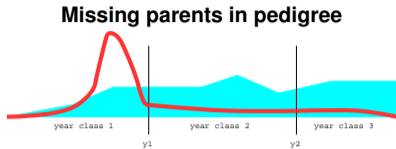
Challenge: Controlling distribution of missing parents

- Some year classes don't have enough missing parents to get genomic contributions \Rightarrow **pedigree truncation**.
- When truncating pedigree, contribution from domestic offspring can “leak” to foreign MF.
- Similarly, crossbred offspring quite often “looses” genomic contribution to purebred MFs.
- Solution: different **pedigree truncation strategies** applied to control distribution of “missing” parents.



Challenge: Controlling distribution of missing parents

- Some year classes don't have enough missing parents to get genomic contributions \Rightarrow **pedigree truncation**.
- When truncating pedigree, contribution from domestic offspring can “leak” to foreign MF.
- Similarly, crossbred offspring quite often “looses” genomic contribution to purebred MFs.
- Solution: different **pedigree truncation strategies** applied to control distribution of “missing” parents.



Estimating Γ matrix from genotypes

- Genomic data (currently):
 - ▶ Number of genotypes: 17494.
 - ▶ Most are domestic purebreds.
 - ▶ Foreign groups have far less genotypes.
- Not enough genotypes to match 115 UPG:
 - ▶ Solution: MF relations estimated first for fewer number of MF year classes.
 - ▶ Then interpolated to cover original UPG year classes.

Group	Genotypes	
	22.11	23.06
1 hf domestic	3730	3857
2 hf foreign	52	56
3 aa domestic	3426	3716
4 aa foreign	78	77
5 ch domestic	2894	3118
6 ch foreign	88	94
7 li domestic	1841	1987
8 li foreign	70	74
9 si domestic	2124	2275
10 si foreign	77	88
12 ot *	1362	1594
	15744	17494

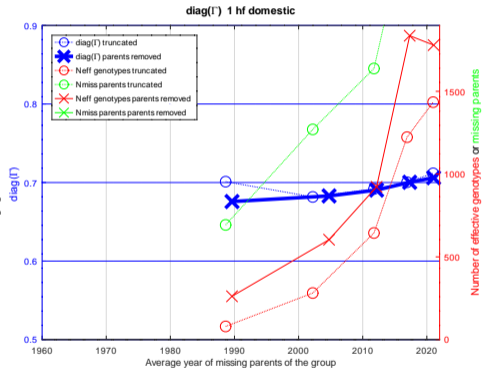
Estimating Γ matrix from genotypes

- Genomic data (currently):
 - ▶ Number of genotypes: 17494.
 - ▶ Most are domestic purebreds.
 - ▶ Foreign groups have far less genotypes.
- Not enough genotypes to match 115 UPG:
 - ▶ Solution: MF relations estimated first for fewer number of MF year classes.
 - ▶ Then **interpolated** to cover original UPG year classes.

Group	Genotypes	
	22.11	23.06
1 hf domestic	3730	3857
2 hf foreign	52	56
3 aa domestic	3426	3716
4 aa foreign	78	77
5 ch domestic	2894	3118
6 ch foreign	88	94
7 li domestic	1841	1987
8 li foreign	70	74
9 si domestic	2124	2275
10 si foreign	77	88
12 ot *	1362	1594
	15744	17494

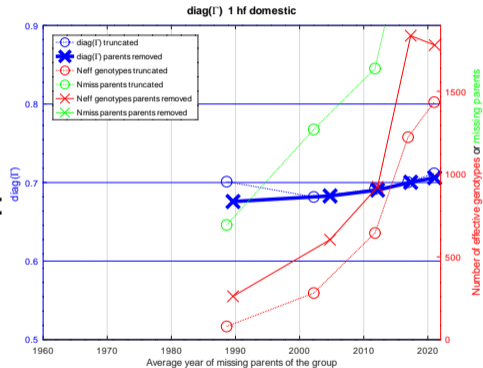
Estimating Γ matrix from genotypes

- Genomic data (currently):
 - ▶ Number of genotypes: 17494.
 - ▶ Most are domestic purebreds.
 - ▶ Foreign groups have far less genotypes.
- Not enough genotypes to match 115 UPG:
 - ▶ Solution: MF relations estimated first for fewer number of MF year classes.
 - ▶ Then interpolated to cover original UPG year classes.



Estimating Γ matrix from genotypes

- Genomic data (currently):
 - ▶ Number of genotypes: 17494.
 - ▶ Most are domestic purebreds.
 - ▶ Foreign groups have far less genotypes.
- Not enough genotypes to match 115 UPG:
 - ▶ Solution: MF relations estimated first for fewer number of MF year classes.
 - ▶ Then **interpolated** to cover original UPG year classes.

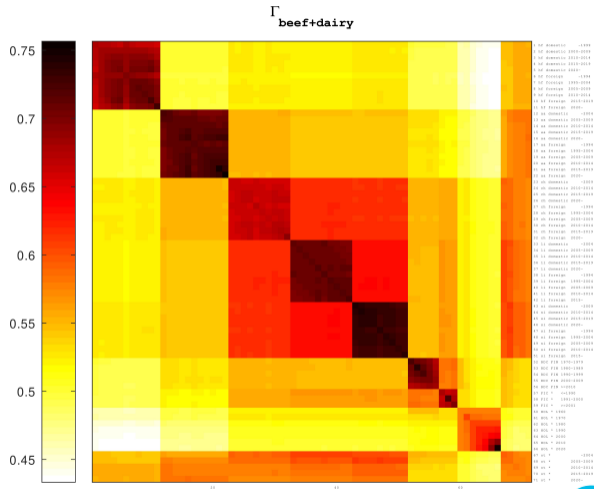


Combined beef+dairy Γ matrix

Allele frequencies of beef breeds were augmented with allele frequencies of **dairy breeds** from other in-house projects:

- Combined Γ matrix calculated from allele frequencies of **common markers**.
- Number of common markers:

	Original	Common	
		Beef	RDC+FIC
Beef	51809		
RDC+FIC	46914	43921	
Holstein	46342	43326	45823

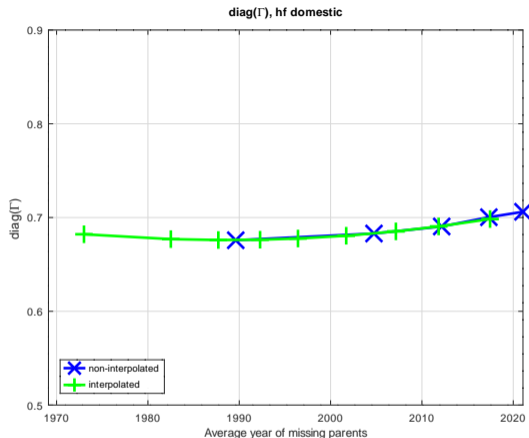


Interpolation of Γ matrix

- Γ matrix was further **interpolated** to cover original UPG year classes.
- Here using two dimensional bilinear fit over years x and y of each breed group pair i and j :

$$\hat{\Gamma}_{i,j}(x, y) = \begin{bmatrix} x & 1 \end{bmatrix} \begin{bmatrix} a_{i,j} & b_{i,j} \\ b_{j,i} & c_{i,j} \end{bmatrix} \begin{bmatrix} y \\ 1 \end{bmatrix}$$

- Example image (diagonal of Γ): Domestic Hereford had 5 MF year classes estimated that were further interpolated to match original 9 UPG.

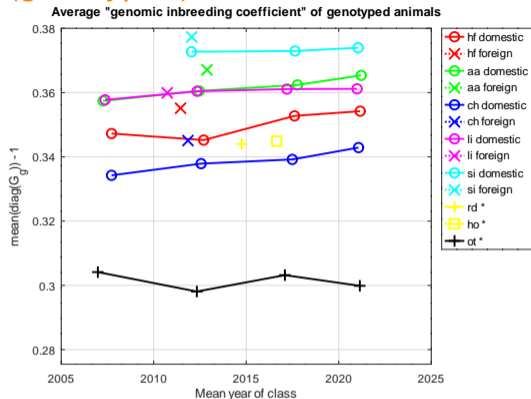
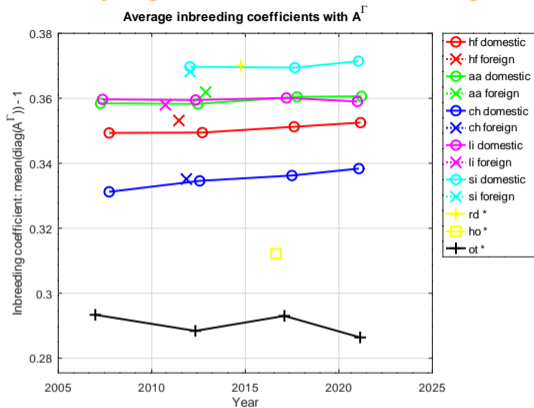


Verifying interpolated Γ matrix

Interpolated Γ matrix was verified by comparing Γ matrix adjusted pedigree relationships A^Γ with genomic relations for genotyped animals:

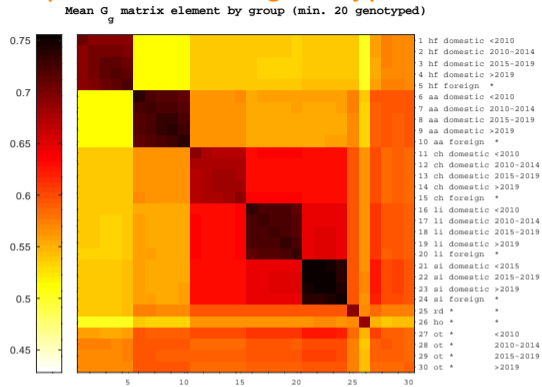
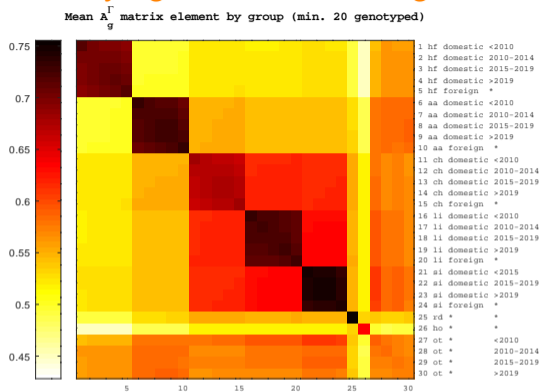
- Inbreeding levels.
- Average within and across breed group relations.

Verifying Γ matrix: Inbreeding levels (genotyped)



- Average inbreeding coefficients of A^Γ for genotyped animals (left).
- Average "genomic inbreeding" of genotyped animals (right).
- Breed groups in the same ("inbreeding") order and overall quite close match.

Verifying Γ matrix: Average breed group relations for genotyped



- Average within and across breed group A_g^Γ relations (left) for genotyped.
- Average **genomic relations** for genotyped animals (right).
- Pedigree relations with Γ are very close to genomic. Largest (avg.) diff: 0.020.

Conclusions

- Metafounder concept applied to Finnish multibreed beef evaluation.
- Metafounder relations estimated from genomic information:
 - ▶ First estimated for fewer number of metafounder year classes.
 - ▶ Then interpolated to cover original unknown parent groups.
- Results verified by comparing metafounder adjusted pedigree relations with genomic relations for genotyped animals.
 - ▶ On average very close.
- Validation (Timo Pitkänen's presentation later in this session):
Metafounder ssGBLUP slightly better than UPG ssGBLUP.

Thanks!

