Imputation for cost-effective genomic selection for disease resistance in Atlantic salmon

Binyam Dagnachew[1]

S.A. Boison[2], M. Baranski[2]

[1]Nofima AS, Norway, [2]Mowi Genetics ASA, Norway

Nofima

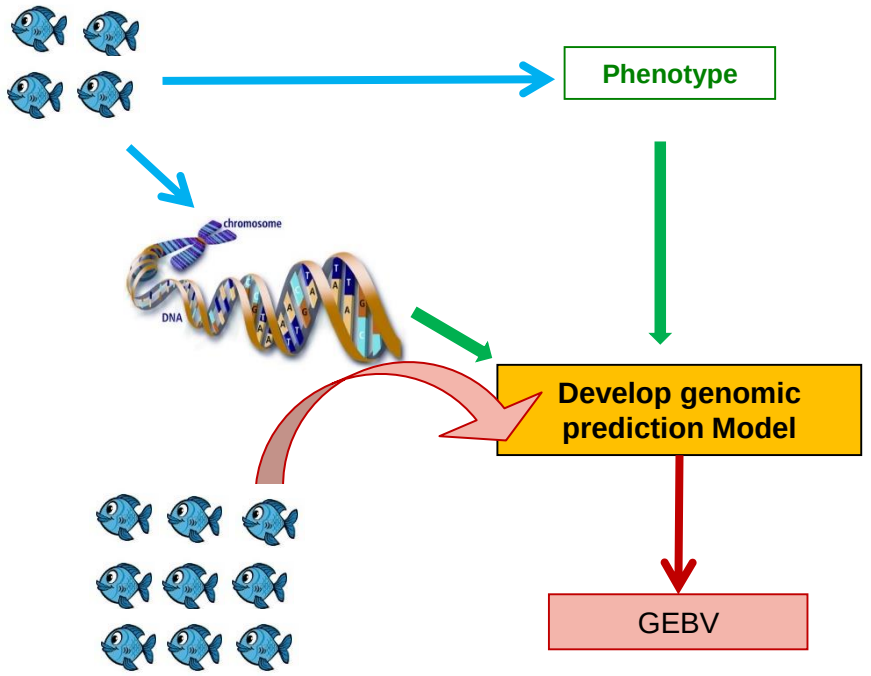
EAAP2019

28/08/2019
GHENT, BELGIUM

Genomic selection (GS)

➢ GS in aquaculture could increase accuracy and genetic gain

- ✓ Increases within-family selection intensity
- ✓ Traits measured only in sibs
  - ✓ disease traits
- ✓ Traits difficult to improve by traditional selection

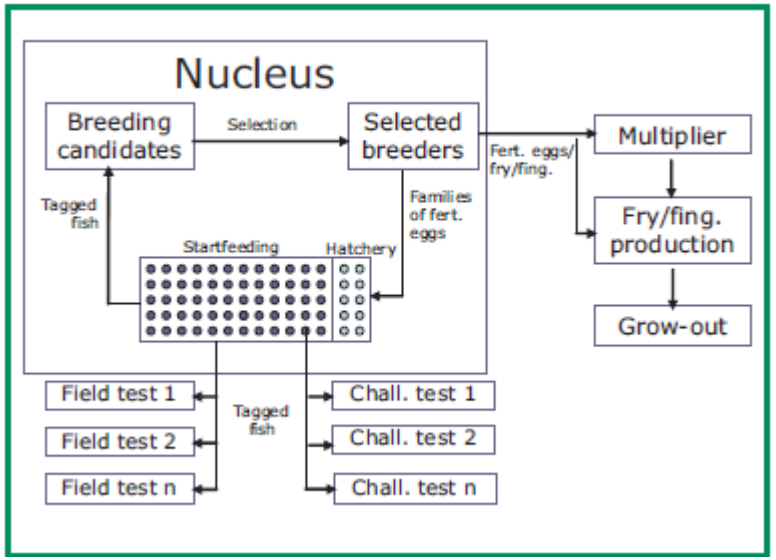


Nofima

GS: the challenge in aquaculture

➢ Conventional GS is expensive in aquaculture

- ❖ Many individuals needed to be genotyped
- ❖ Most of the genotypes are no/limited re-usable



➢ Approaches to reduce cost

- ❖ Reduce # of candidates
- ❖ Reduce # of test individuals
- ❖ Use of DNA pools for the reference population
- ❖ Reduce # of markers combined with imputation

Nofima

Why imputation?

- Decrease genotyping cost
- Increase selection intensity by genotyping more selection candidates with low-density markers
- Increase prediction accuracy by genotyping more training individuals with fewer markers

$$\text{Response } R_t = h^2 S = \frac{i\, r\, \sigma_A}{L}$$

Questions

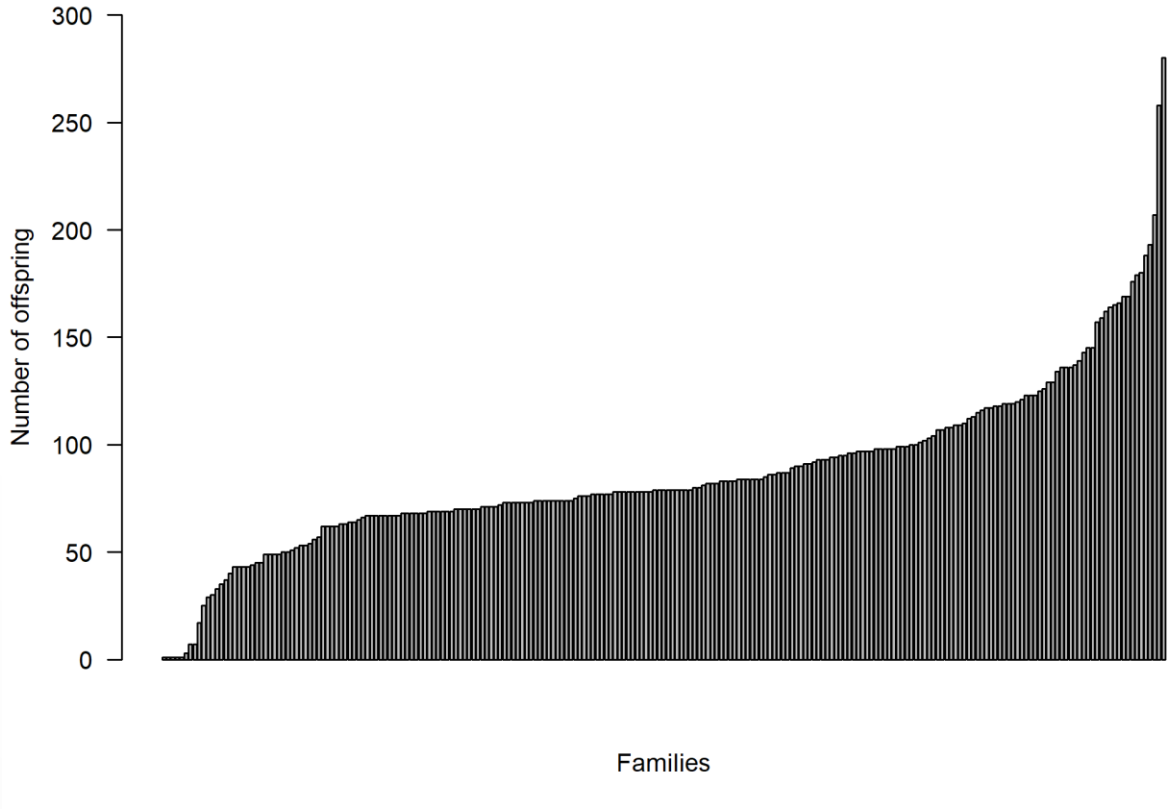
- **Given 54k SNP array, can we select 0.5k (~500 SNPs) and maximize imputation accuracy with different marker selection approach?**

- **How big is the loss in accuracy?**

The population

	Number
Sires	121
Dams	348

Traits	# Parents	# offspring
AGD	451	3,511
CMS	450	4,312
GWT	428	3,428
PD	468	7,645
PIGM	435	3,425
SL	469	4,564
SWT	242	4,592



Nofima

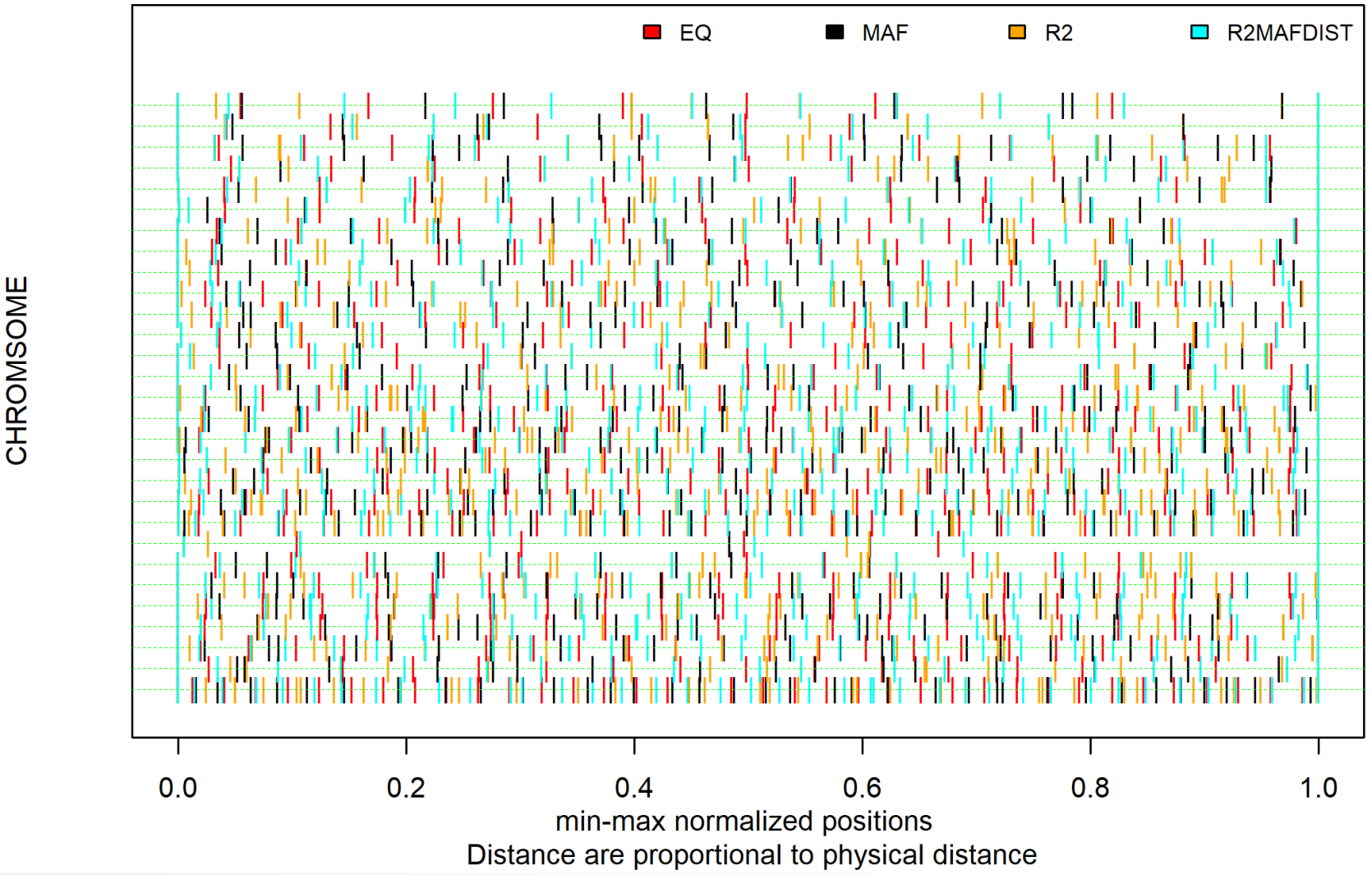
Marker Selection approaches

- Given 54K SNP, 0.5K were selected using different selection approaches
  - *Markers are selected based on the statistics from the parental population*

- **Approaches**
  - **Equidistance (EQ) –** proximity to the center of SNP windows
  - **MAF –** Highest allele frequency in a SNP window
  - **R2 –** based LD of a marker with all markers in the SNP window
  - Combination of the above approaches
    - **EQ:MAF**
    - **EQ:R2**
    - **R2:MAF**
    - **EQ:MAF:R2**

When using these combinations, weight are given to each. Weights sums up to 1

Marker selection by method



Different markers are selected

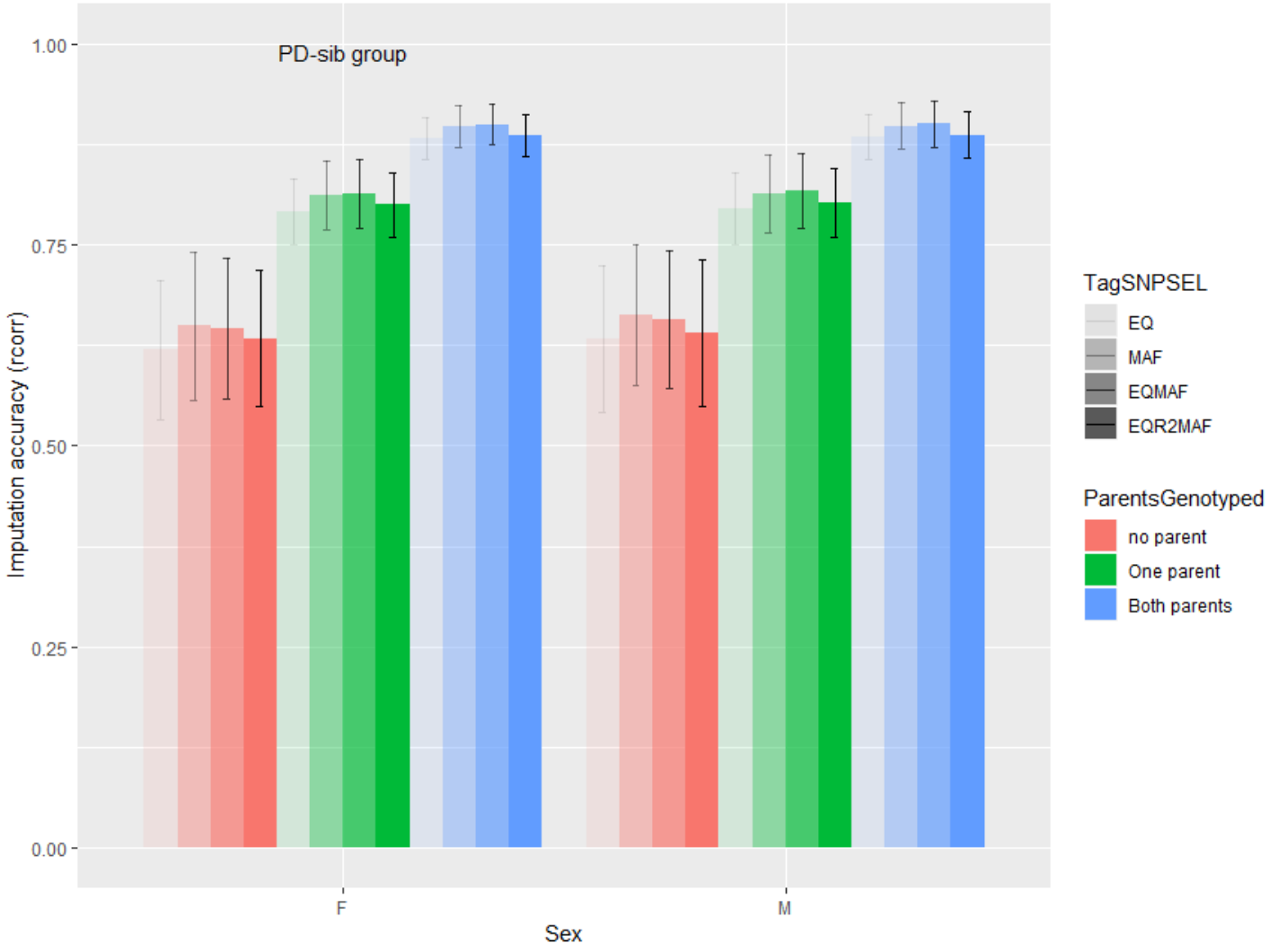
Nofima

Imputation

- Parents were genotyped in ~53k SNP (chip 1)
- Offspring were genotyped ~0.5k SNP (chip 2)
- FIMPUTE software

Chr	Length(MB)	No. SNP	
		Chip 1	Chip 2
1	158.93	3919	38
2	74.81	1759	19
3	92.29	2578	26
4	86.74	2160	21
5	83.59	2195	20
6	87.01	2167	22
7	88.98	1524	15
8	36.4	482	5
9	141.7	2918	28
10	116.09	2862	28
11	93.84	2042	20
12	91.86	2175	21
13	107.73	2678	27
14	93.89	2415	22
15	103.9	2175	22
16	87.49	1775	19
17	57.55	1302	15
18	70.6	1521	16
19	82.83	1680	18
20	86.73	2130	22
21	57.93	1302	15
22	63.06	1575	16
23	55.84	1495	16
24	59.89	1281	14
25	53.38	1249	14
26	47.7	1086	13
27	45.5	1299	15
28	57.77	1090	12
29	50.69	957	11
Overall	2334.72	53.791	550

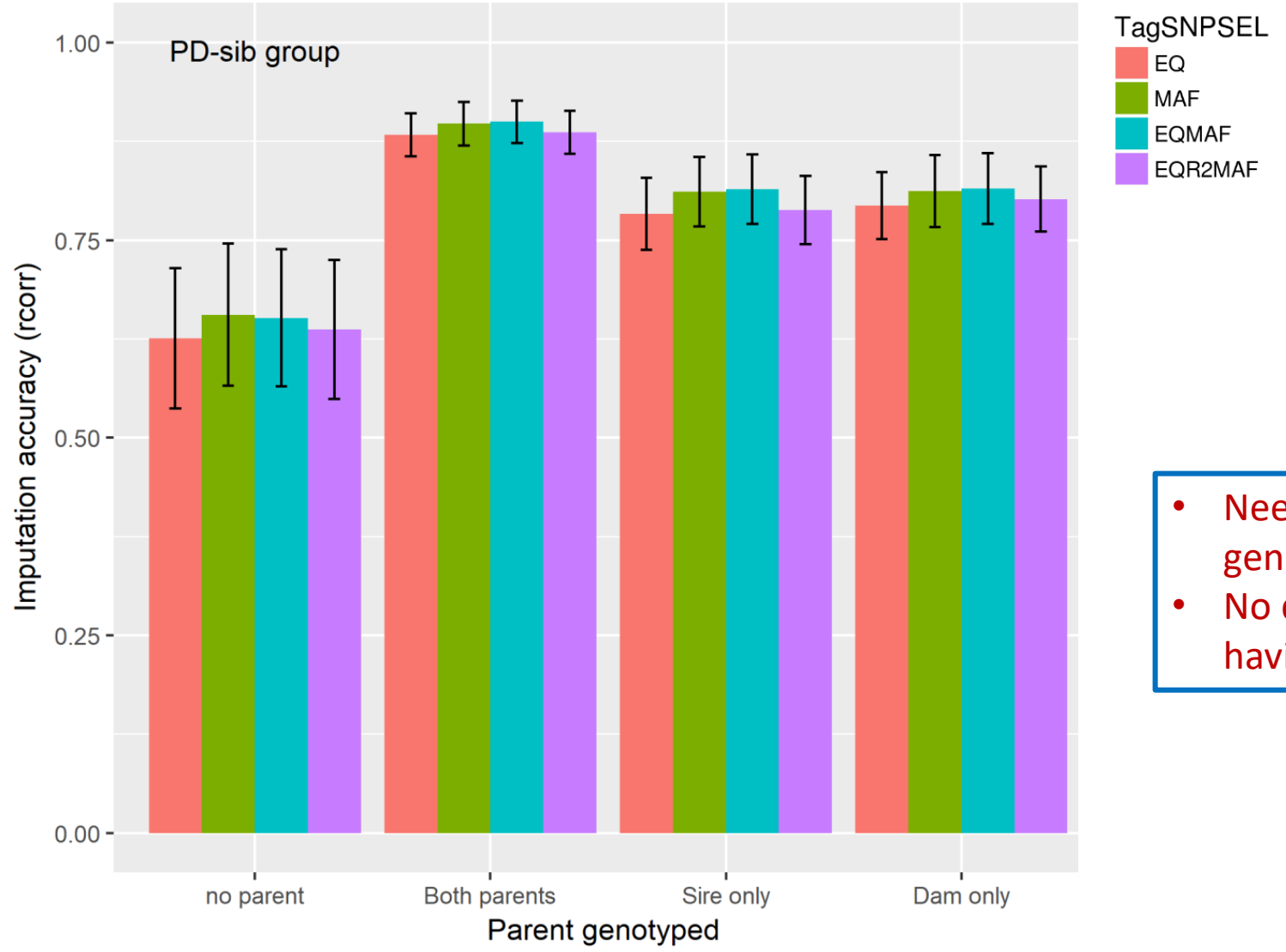
Sex-wise imputation accuracy



No difference in accuracy for male and female offspring

Nofima

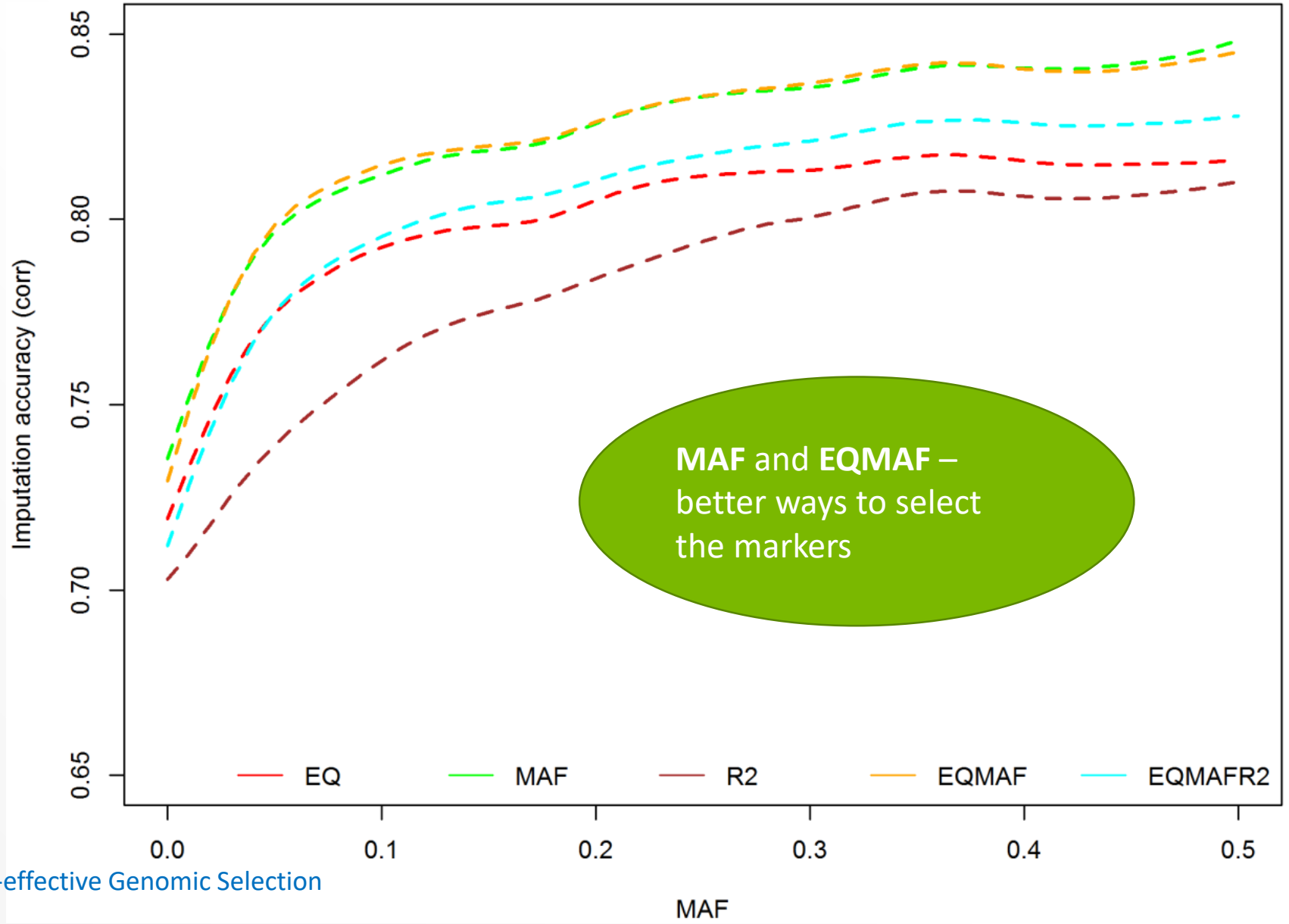
Which parent to genotype?



- Need to have both parents genotyped
- No difference between having dam only or sire only

Nofima

Imputation accuracy per marker selection method



Nofima

Genomic prediction

- **Models**
  - SNP-BLUP model
    - Non-imputed genotypes
    - Imputed
  - Pedigree based analysis

- **Training and validation sets**
  - Family-wise splitting of data

Trait	Nsamples	NTrain (85% of N)	NTest (15% of N)
AGD	3511	2986	525
CMS	4312	3662	650
Growth (Bodyweight)	3428	2913	515
Pancreas Disease (PD)	7645	6495	1150
Pigmentation (Colour)	3425	2910	515
Sea-lice (LC)	4564	3879	685
Smolt weight (SWT)	4592	3902	690

Accuracy of selection

$$accuracy\ (r) = \frac{\rho(G[P]EBV, y_{adj})}{\sqrt{h^2}}$$

$GEBV$- Genomic breeding values computed from **true** or **imputed** genotypes

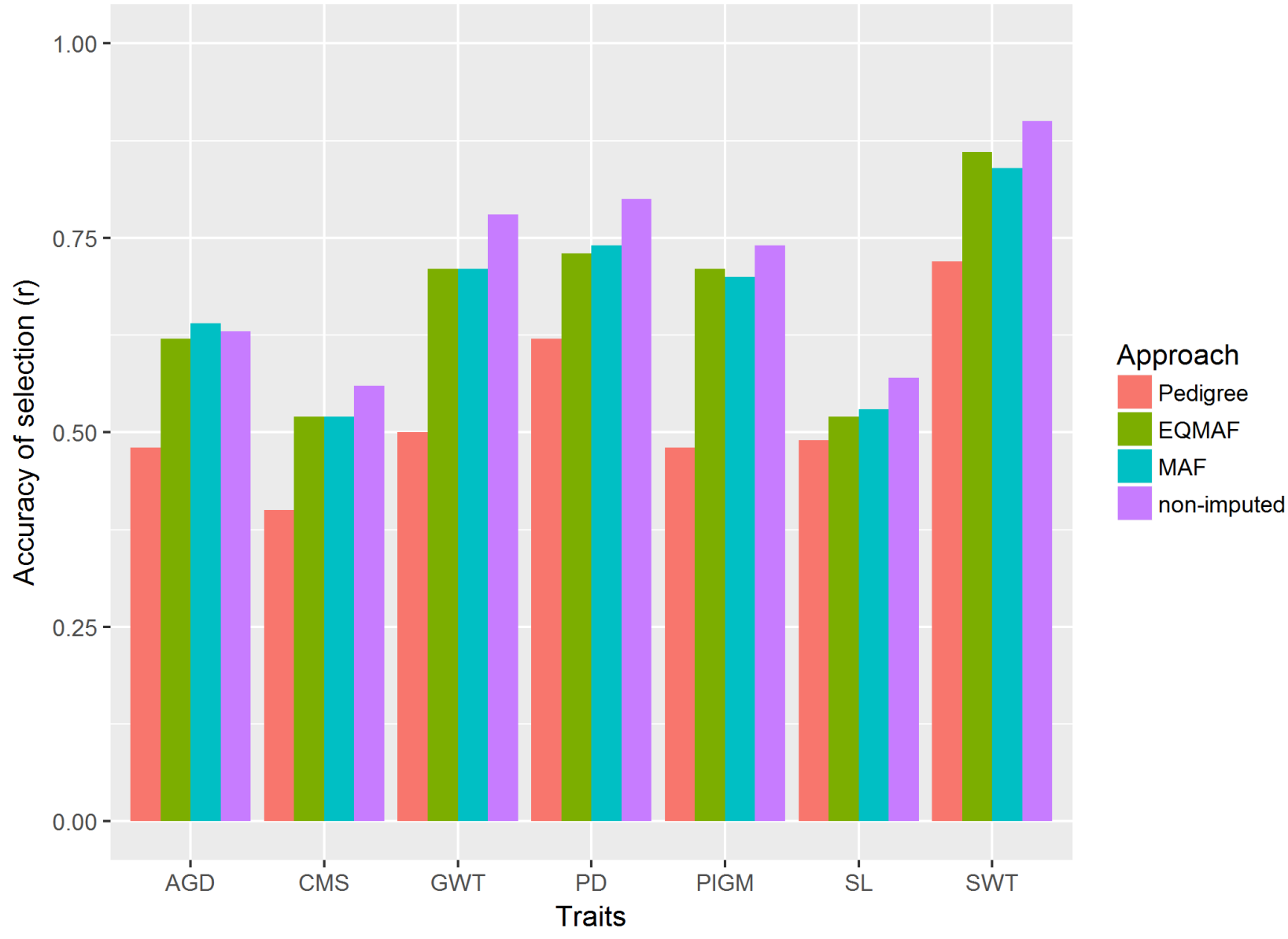
$PEBV$ – Pedigree based breeding values

h^2 - heritability based on pedigree estimate

y_{adj} - adjusted phenotype

GEBV prediction accuracies

Little loss in accuracy of prediction



Nofima

Summary

- EQMAF and MAF were better approaches for selection of the markers
- No difference in imputation accuracies for male and female offspring
- Genotyping of both parents is necessary
  - No or limited difference in imputation accuracies for genotyping sire or dam only
- Limited loss in selection accuracy

MOWI®

FutureFish

Thank you!

Questions and comments

Nofima