



Accuracy of Population-wide and Within-family Genomic Selection in Atlantic Salmon

Sonesson A.K., Meuwissen T.H.E., Baranski M., Moghadam H., Lillehammer M., Norris A., Bakke H. and Lund V.

Genomic selection (Meuwissen et al., 2001)

- Opportunities in aquaculture
 - increase within-family selection intensity
 - sib-traits (disease resistance traits, fillet traits)
- Genomic selection using population-wide marker effects
- Within-family genomic selection (Lillehammer et al., 2013)

Aim

Estimate accuracy of traditional and genomic breeding values for resistance to pancreas disease in Atlantic salmon

Variants:

- population-wide or within-family genomic selection
- marker densities
- family size

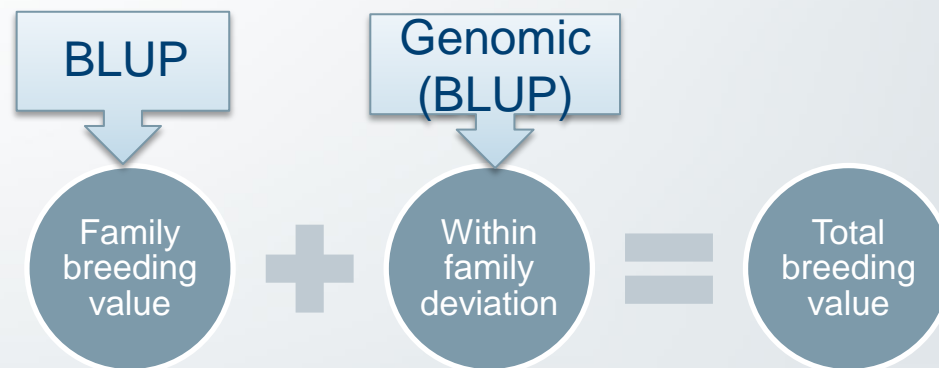
M&M- Breeding value estimation methods

G(EBV) by BLUP with alternative relationships matrices (ASReml)

1. Traditional BLUP breeding values (A-matrix)
 - Between family component
2. Across family GS based on pop-wide LD (GLD-matrix)
 - If marker alleles are Alike-in-State $\Rightarrow GLD_{(\text{position } i)} = 1$
 - $GLD_{(\text{position } i)} = 1$ can occur even for fish from different families
 - Averaged over all SNP positions
 - No pedigree used

3. Within family GS based on linkage analysis (GLA-matrix)

- Linkage analysis performed by LDMIP sets up a G-matrix
- Traces inheritance from parent to offspring using SNPs
 - if this shows that segments are identical => $GLA_{(\text{position } i)}=1$
- SNP-allele 1 in family A \neq SNP-allele 1 in family B
- Averaged over all SNP positions



M&M- fish material and SNP markers

- Marine Harvest 2014 fry test for PD resistance at VESO, Norway
- 1432 individuals from 59 full-sibling families+ parents genotyped
- Pedigree with 4198 individuals
- 3000~31000 SNP markers (Baranski et al.)

M&M- crossvalidation test

1. For every family the record of one animal was masked
2. (G)EBV calculation for masked animal
3. Correlation between masked/realised record and predicted (G)EBV
4. Accuracy= correlation scaled by $\sqrt{h^2}$

Results- Heritability estimate

- 0.47 with GLD-matrix
 - Overestimated because mortality was close to 50% per family

Results- accuracy A and GLD matrices

	N_SNPs	Accuracy((G)EBV)
A-matrix20perfam		0.311
GLD-matrix20perfam	31014	0.698
GLD-matrix10perfam	31014	0.551
GLD-matrix20perfam	10000	0.660
GLD-matrix10perfam	10000	0.483
GLD-matrix20perfam	3000	0.653
GLD-matrix10perfam	3000	0.478

Results-accuracy GLA-matrix

	N_SNPs	Accuracy((G)EBV)
GLA-matrix20perfam	31014	0.779
GLA-matrix10perfam	31014	0.757
GLA-matrix20perfam	3000	0.778

Conclusions

- GLA-matrix highest accuracy here
 - Overestimated by overestimated heritability and some very large half-sib families
- Family size very important for both GLA- and GLD-matrices
- Less sensitive to numbers of markers

One large family with 150 progeny in total for dam+sire (87+61)

- This couple of parents had 20 offspring
 - Cross-validation of these 20 offspring
 - Correlations- phenotype-GEBV
 - 30000 markers- 0.692
 - 3000 markers-0.679
- i.e. correlations not much affected by marker density
i.e. correlations close to maximum (0.69) ($=\sqrt{h^2}=h$)
i.e. accuracy is close to 100%

! Only one family was used- we wanted to avoid between fam info. to interfere

