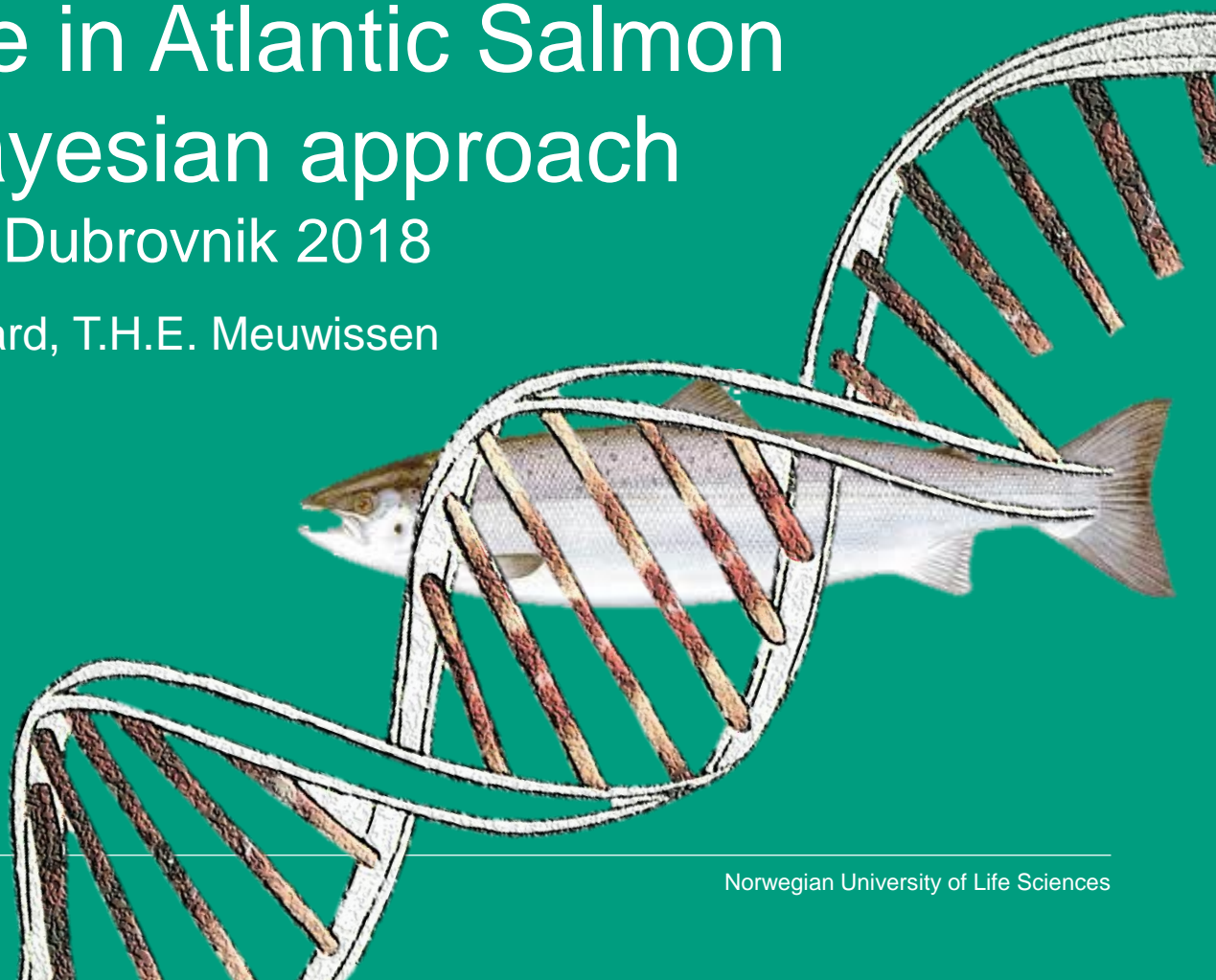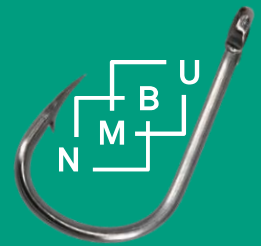# Accuracy of genomic prediction of salmon lice in Atlantic Salmon using a Bayesian approach

EAAP Dubrovnik 2018

Maria H. Kjetså, J. Ødegård, T.H.E. Meuwissen

30.08.2018 – session 66

Salmon Lice was estimated to cause US$436m in damages to the Norwegian fish farming industry in 2011.

- Abolofia et al. 2017

Jay Abolofia, Frank Asche, and James E. Wilen, "The Cost of Lice: Quantifying the Impacts of Parasitic Sea Lice on Farmed Salmon," Marine Resource Economics 32, no. 3 (July 2017): 329-349

Salmon Lice (*Lepeophtheirus Salmonis*)

Big problem both for animal welfare and economy in production of Atlantic Salmon and Rainbow Trout!

- Reduces fish growth and appetite

- Current treatments not very effective

  o problems with resistance

  o Treatments stressful and causes loss

  o "Cleaning fish" > high mortality

-Host resistance to salmon lice is heritable:

- Challenge test in tanks: **0.33**±0.05

  (Gjerde et al., 2011).

- Natural infection rates: **0.07-0.14**±0.02

  (Kolstad et al. 2005)
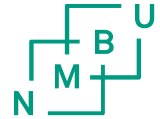
- Challenge test in sea cages: **0.14**±0.03

  (Ødegård et al., 2014).

Breeding schemes for disease traits in fish

- Disease traits:

  - Cannot challenge test elite breeding stock

  - Needs prediction of EBVs for un-phenotyped fish

- Sib-testing: Select Full-Sib families and split into

  - > Breeding stock

  - > Challenge test stock

Genomic selection in fish breeding schemes

- Genomic selection very important tool in fish breeding schemes

- In order to not loose any accuracy in the prediction;

    – challenge test needed every generation.

    –Just increasing the genotype density will not automatically increase the prediction accuracy

AIM:
Compare two methods of genomic prediction in terms of prediction accuracy for the trait Host resistance to sea lice in Atlantic Salmon

Test design

- We compared two methods of genomic prediction:

  - **GBLUP** and **Bayes GC**.

  - We also included pedigree **BLUP** as "control"

- Two different scenarios for the 5-fold-cross-validation:

  - **Random:** Random assignment of fish in the training and test data sets

  - **Across family:** Animals were sorted according to full-sib family, and an entire family were assigned to test- or training set.

Data

- 1385 Atlantic salmon with genotype and phenotype

- 99 full-sib families

- Genotyped with 220K SNP chip

- Phased and imputed to 750K SNP using FImpute*.

    – 60 parents with high density, no phenotype

- Phenotype: **logLC**:

    – logarithm of lice count from challenge test in sea cages.

    – host resistance to lice

- Heritability: 0.14

8   *Sargolzaei, M., J. P. Chesnais and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics, 15:478 (DOI: 10.1186/1471-2164-15-478).
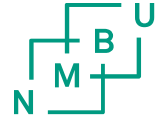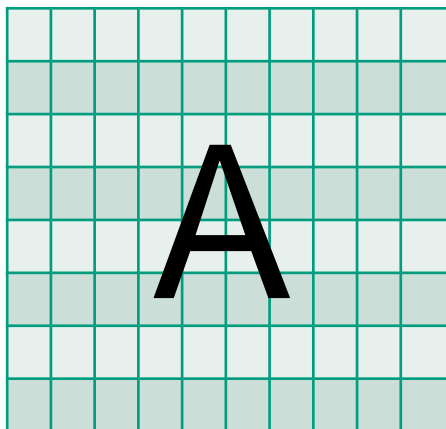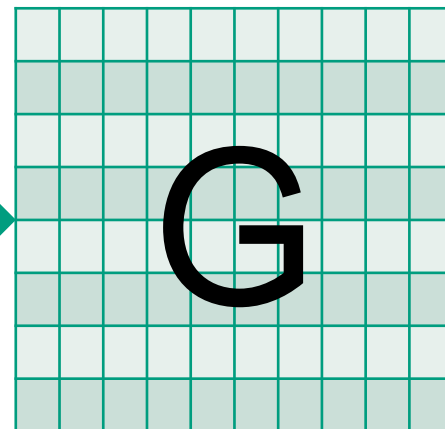
Linear Model

y = Xb + e

- LogLC = Person-TestDay + Fixed regression on weight + e

  – Run in DMU (Madsen & Jensen, 2007)

    • Produced Yield deviations (YD)

    • YD were analysed by BLUP, GBLUP and Bayesian model

      – overall mean was only fixed effect in these models
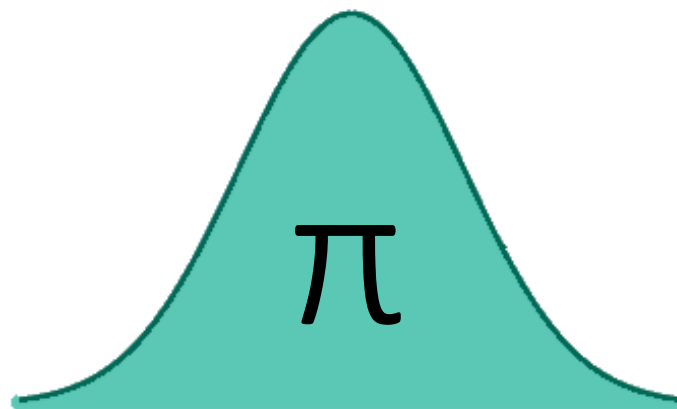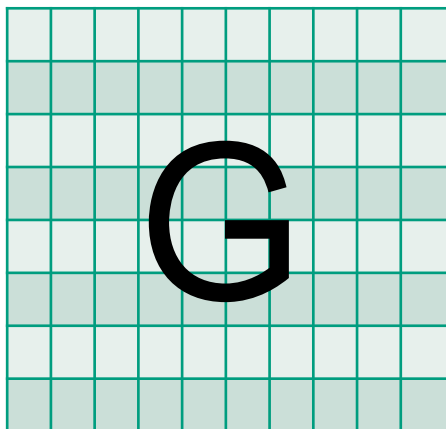
Pedigree to Genomic Prediction

BLUP

GBLUP

GBLUP simply replaces the pedigree relationship matrix in the mixed model equations by using a genomic relationship matrix.

The genomic relationships are derived from the SNPs.
- assumes all SNPs have equal variance
- finds the covariance between related animals based on their genome.

Bayes GC

Bayes GC

Bayes GC fits a polygenic effect through a G-matrix and a Bayes C term simultaneously

- Iheshiulor et al. 2017 (Iterative method of Bayes C)
- We use MCMC

11 Iheshiulor, O. O. M., Woolliams, J. A., Svendsen, M., Solberg, T., & Meuwissen, T. H. E. (2017). Simultaneous fitting of genomic-BLUP and Bayes-C components in a genomic prediction model. *Genetics, Selection, Evolution : GSE*, *49*, 63. http://doi.org/10.1186/s12711-017-0339-9
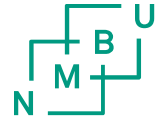
Bayes GC

- Bayes C fits SNPs with larger effects by using a prior probability (π) of a SNP having an effect or not.

- SNPs with large effects are assumed to be normally distributed.

- Runs multithreaded MCMC chains.

  – Tries to increase speed

Bayes GC
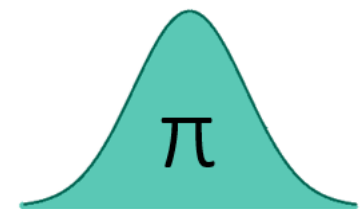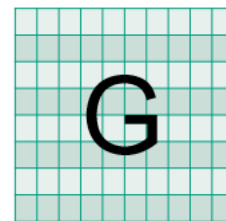
The complete Bayes GC model:

$$YD = \mu + u + \sum_i I_i X_i s_i + e$$

- (YD = Yield Deviation)

- u = polygenic effect $V(u) = G\sigma_u^2$

  - G matrix based on 750K SNP chip data

- I_i = indicator whether the SNP is in model in particular MCMC-cycle (0/1)

  - Prior probability I_i = 1 is π (here π = 0.003)

  - If SNP in model: $s_i \sim N(0, \sigma_u^2 /1000)$

Faster MCMC sampling

- Store genotype matrix once, run many threads in (relatively) short chains

- Sort (phased) genotypes into haplotypes

- Store haplotype IDs for each animal, instead of genotype IDs

    - (2 haplotype IDs/animal instead of 2000 SNP alleles/animal)

- Store which SNP alleles are in which haplotype

Results

Cross validation	Prediction Method	Accuracy of prediction	Standard Error
Random	BLUP	0.523	0.07
	GBLUP	0.665	0.07
	Bayes GC	0.670	0.07
Across family	BLUP	0.402	0.07
	GBLUP	0.568	0.07
	Bayes GC	0.595	0.07

Accuracy of prediction: 66,5% and 67% for GBLUP and Bayes GC respectively

Results

Cross validation	Prediction Method	Accuracy of prediction			Standard Error
Random	BLUP			0.523	**0.07**
	GBLUP	-23.1%		0.665	**0.07**
	Bayes GC		-14.8%	0.670	**0.07**
Across family	BLUP			-11.2% 0.402	0.07
	GBLUP			0.568	0.07
	Bayes GC			0.595	0.07

The Full-sib-family scenario performs worse overall for all the methods

- The reduction between scenarios is less for Bayes GC.

Results

Cross validation	Prediction Method	Accuracy of prediction	Standard Error
Random	BLUP	0.523	0.07
	GBLUP	0.665	0.07
	Bayes GC	0.670	0.07
Across family	BLUP	0.402	0.07
	GBLUP	0.568	0.07
	Bayes GC	0.595	0.07

+0.7%

+4.9%

For the Random five-fold predictions, The Increase in % from GBLUP to BGC is 0.76%.
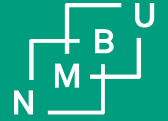For the non-full-sib, the increase in % from GBLUP to BGC is 4.9%.

Conclusions

- Sea lice resistance GEBV were quite accurate (67%)

- Accuracy difference between GBLUP and Bayes GC is not very large

  - But increased when predicting across families

    - i.e. predicting over larger genetic distances

Thank you for your attention

Questions?