

GENOMIC SELECTION FOR COMMERCIAL TRAITS IN NILE TILAPIA INCREASES PREDICTION ACCURACY AND GIVES UNBIASED ESTIMATES OF THE BREEDING VALUES

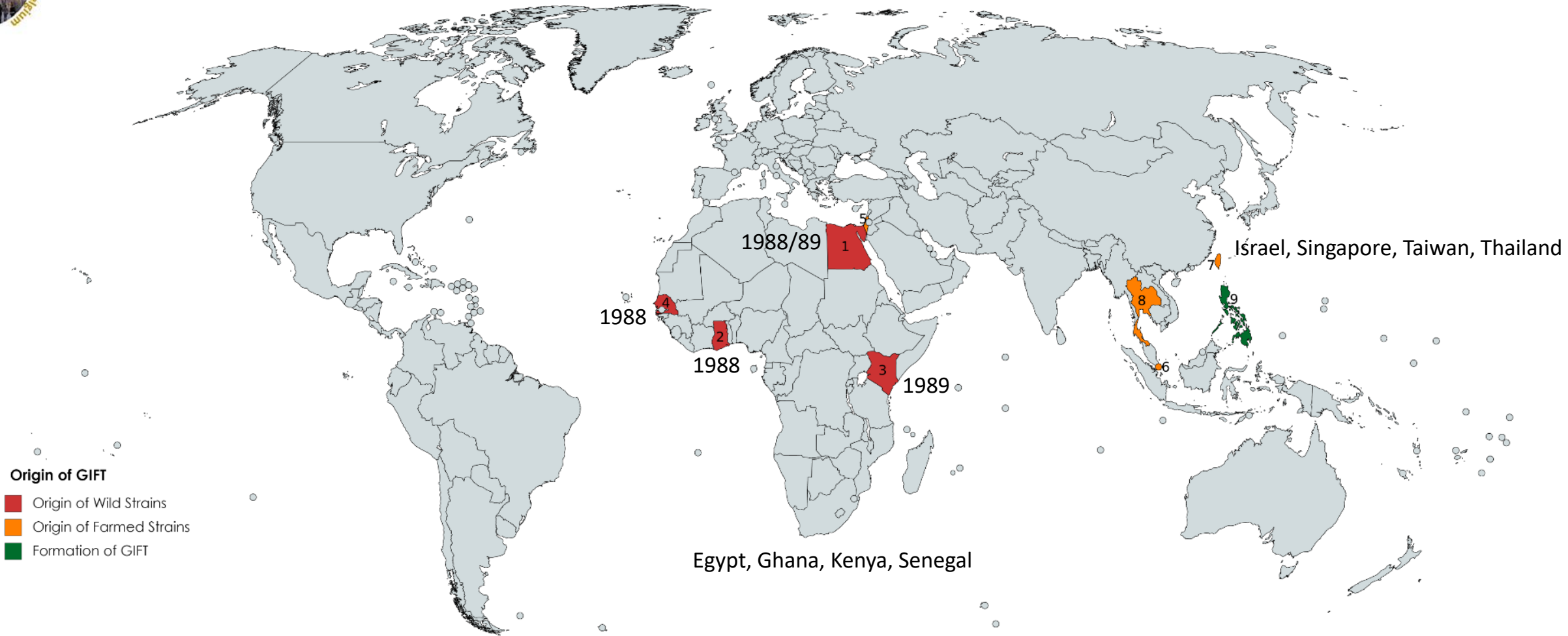
Rajesh Joshi, Anders Skaarud, Mayet de Vera, Alejandro Tola Alvarez, Jørgen Ødegård
GenoMar Genetics AS, Oslo, Norway and AquaGen AS, Norway



70th Annual Meeting of the European Federation of Animal Science (EAAP 2019)

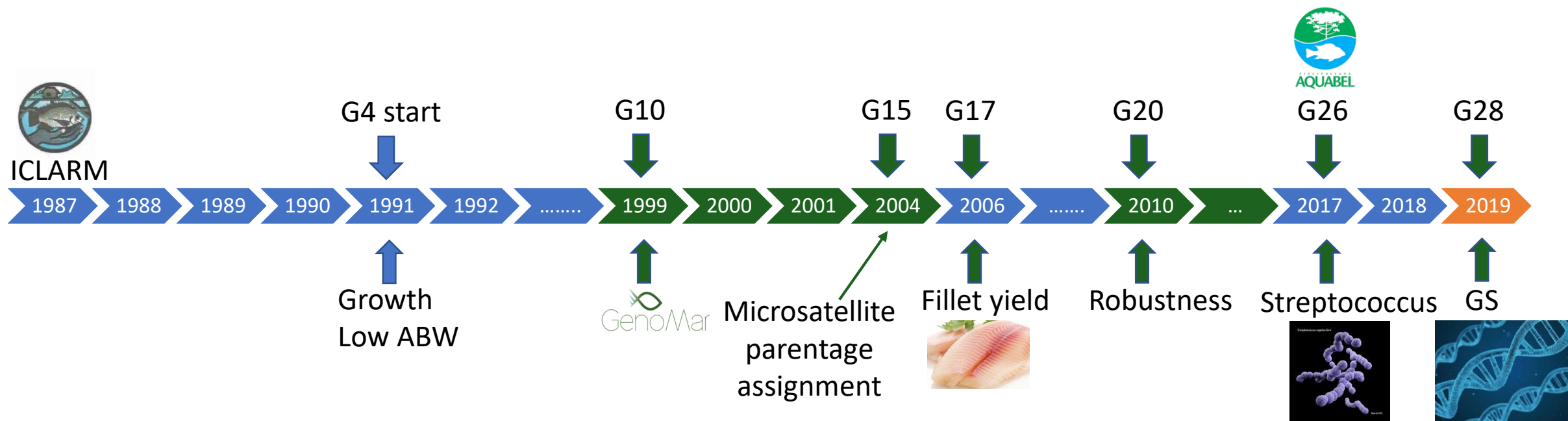
August 26-30, 2019

Formation of GIFT



Created with mapchart.net ©

GenoMar Supreme Tilapia (GST[®])



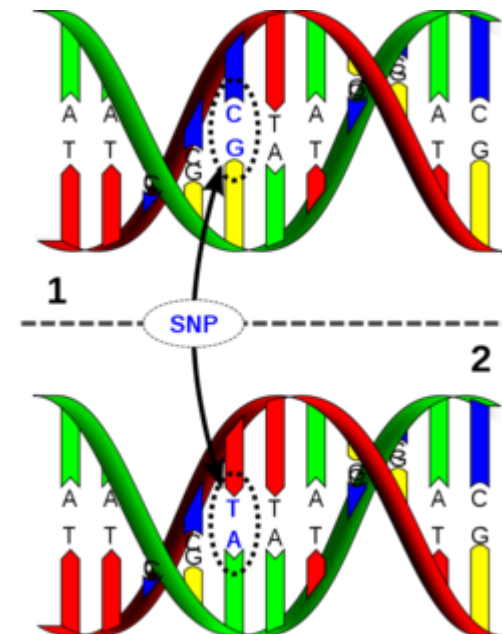


Breeding Nucleus



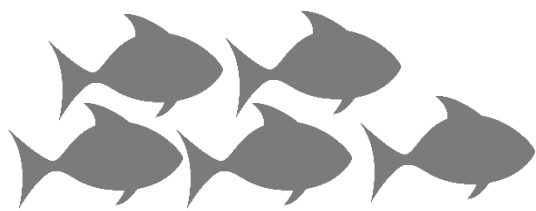
Genomic Selection

- Utilising abundant genome wide markers (SNPs)



Genomic Selection.....

- Use training set of population to obtain the prediction of SNP effects



Reference Population

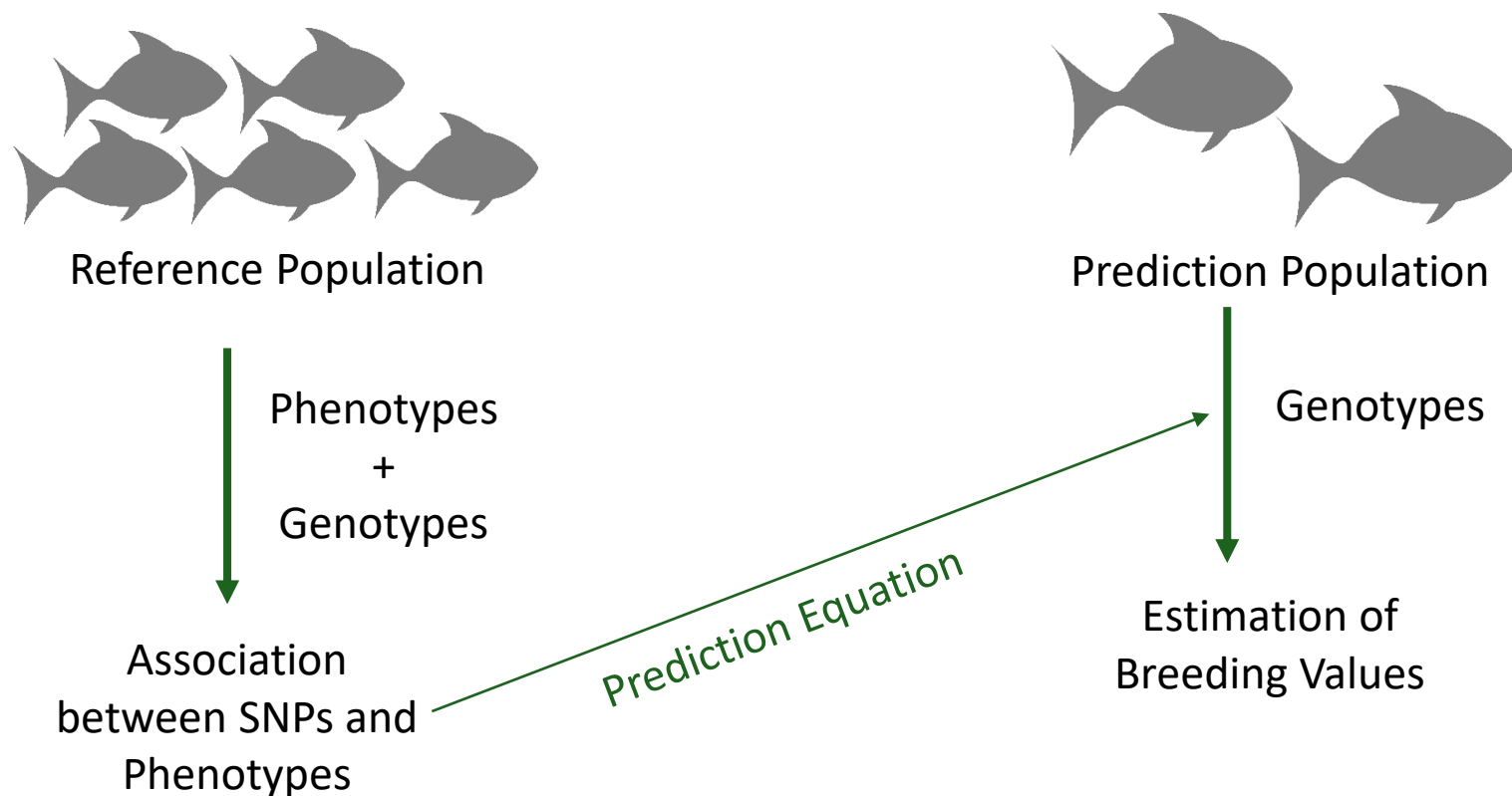


Phenotypes
+
Genotypes

Association
between SNPs and
Phenotypes

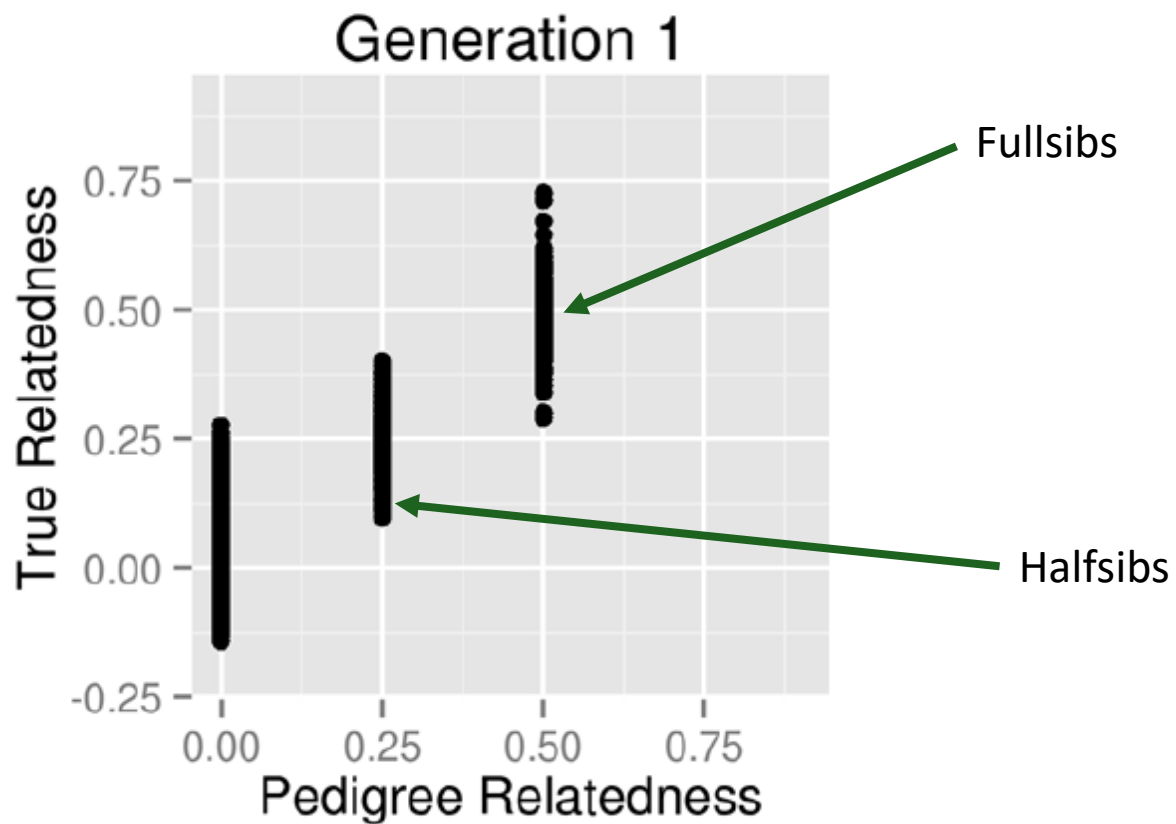
Genomic Selection.....

- Use training set of population to obtain the prediction of SNP effects



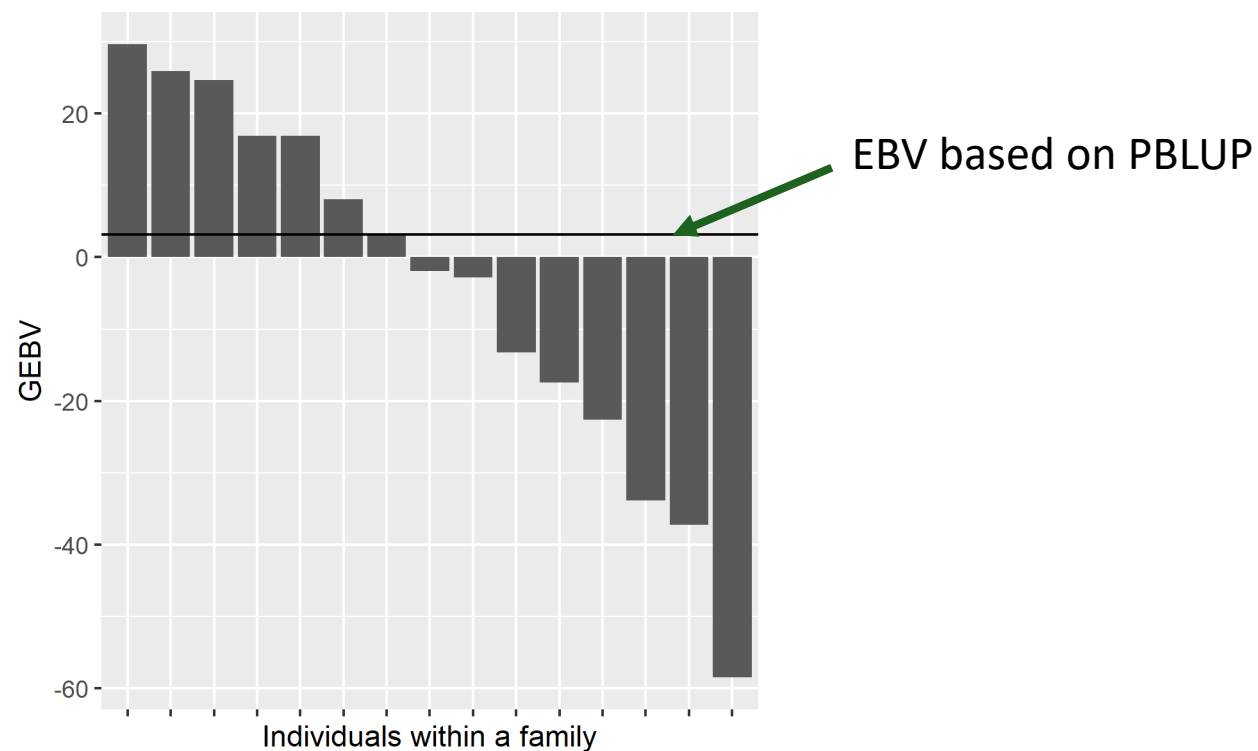
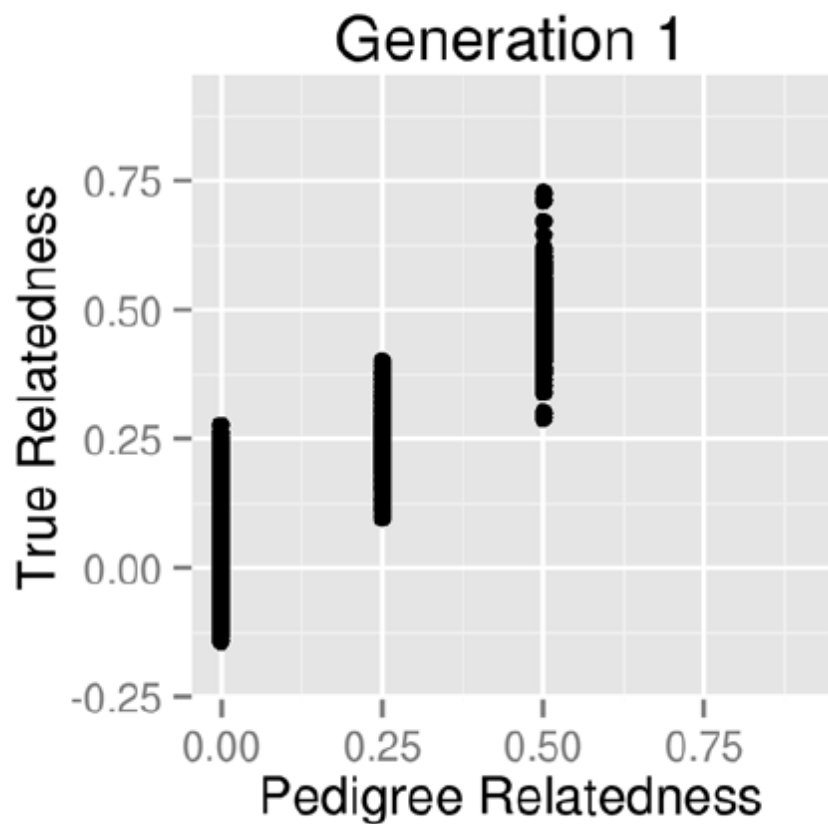
Advantages of Genomic Selection

SNP - more accurate genetic relationship



Advantages of Genomic Selection.....

Genomics- predicting different breeding values for full-sibs in sib-testing based on Mendelian segregation



Tilapia Reference Genome Assembly

Orenil 1.0 – released in 2011

- Updated to Orenil1.1 – at the end of 2012
- Based on short read sequencing

Tilapia Reference Genome Assembly....

Orenil 1.0 – released in 2011

- Updated to Orenil1.1 – at the end of 2012
- Based on short read sequencing

O_niloticus_UMD1 – 2017

- Novel long reads- using Pacific Bioscience Technology + publicly available illumine Short reads (Genome coverage: 44x)
- **O_niloticus_UMD_NMBU assembly:** re-anchoring using the high density linkage map
 - 90.2% (907.6 Mbp is now anchored to the genome assembly)



GigaScience, 8, 2019, 1–20

doi: 10.1093/gigascience/giz030
Research

RESEARCH

Chromosome-scale assemblies reveal the structural evolution of African cichlid genomes

Matthew A. Conte¹, Rajesh Joshi², Emily C. Moore³, Sri Pratima Nandamuri¹, William J. Gammerdinger¹, Reade B. Roberts³, Karen L. Carleton¹, Sigbjørn Lien² and Thomas D. Kocher^{1,*}

¹Department of Biology, University of Maryland, College Park, MD 20742, USA, ²Centre for Integrative Genetics (CIGENE), Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, PO Box 5003, Ås, Norway and ³Department of Biological Sciences and W. M. Keck Center for Behavioral Biology, North Carolina State University, Raleigh, NC 27695, USA

50K SNP arrays

Onil50 array (Affy) – 2016-17

- 58,466 SNPs
- WGS-32 fish
- GenoMar Genetics AS



ORIGINAL RESEARCH
published: 15 October 2018
doi: 10.3389/fgene.2018.00472



Development and Validation of 58K SNP-Array and High-Density Linkage Map in Nile Tilapia (*O. niloticus*)

Rajesh Joshi^{1*}, Mariann Árnýasi^{1†}, Sigbjørn Lien¹, Hans Magnus Gjøen¹, Alejandro Tola Alvarez² and Matthew Kent¹

¹ Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, Ås, Norway, ² GenoMar Genetics AS, Trondheim, Norway

50K SNP arrays....

Onil50 array (Affy) – 2016-17

- 58,466 SNPs
- WGS-32 fish
- GenoMar Genetics AS



ORIGINAL RESEARCH
published: 15 October 2018
doi: 10.3389/fgene.2018.00472



Development and Validation of 58K SNP-Array and High-Density Linkage Map in Nile Tilapia (*O. niloticus*)

Rajesh Joshi^{1*}, Mariann Árnýasi^{1†}, Sigbjørn Lien¹, Hans Magnus Gjøen¹, Alejandro Tola Alvarez² and Matthew Kent¹

¹ Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, Ås, Norway, ² GenoMar Genetics AS, Trondheim, Norway

50K (Illumina)- 2019

- 50,000 SNPs
- 3 different Latin American commercial tilapia
- WGS of 326 fish

High-throughput single nucleotide polymorphism (SNP) discovery and validation through whole-genome resequencing of hundreds of individuals in Nile tilapia (*Oreochromis niloticus*)

J.M. Yáñez, G. Yoshida, A. Barria, R. Palma-Véjares, D. Travisany, D. Díaz, G. Cáceres, M.I. Cádiz, M.E. López, J.P. Lhorente, A. Jedlicki, J. Soto, D. Salas, A. Maass

doi: <https://doi.org/10.1101/594671>

This article is a preprint and has not been peer-reviewed [what does this mean?].

Linkage maps

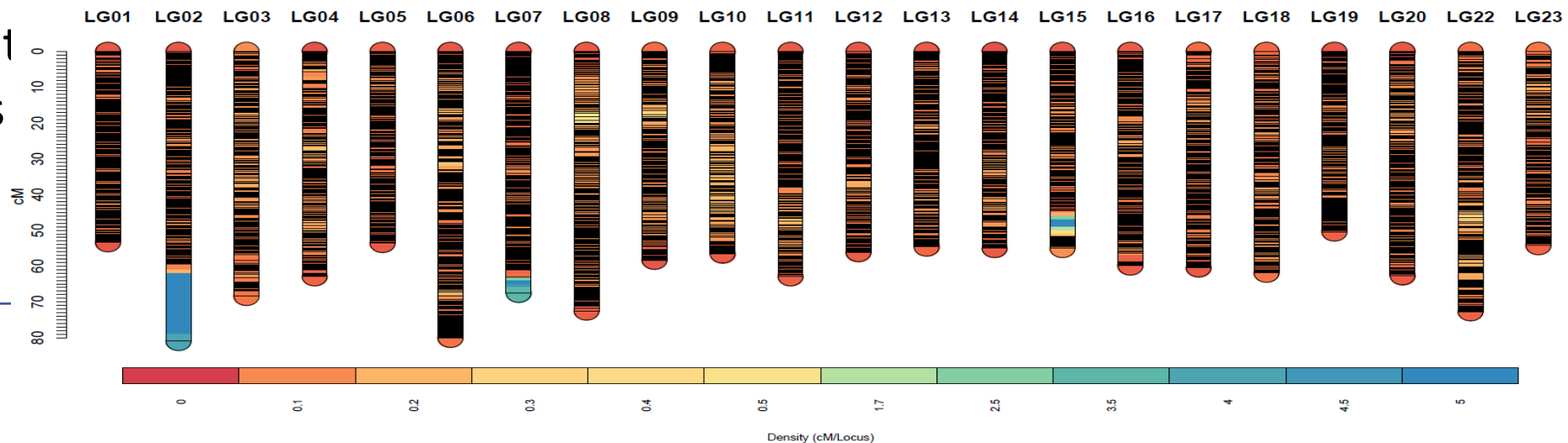
Low density Linkage Maps

- For linkage maps of varying resolution
- Markers found with
 - Restriction-site Associated DNA (RAD) sequencing (Palaikostas et al., 2013)
 - Microsatellites and/or AFLP markers (Guyon et al., 2012; Kocher et al., 1998; Lee et al., 2005).

Linkage maps....

Low density Linkage Maps

- For linkage maps of varying resolution
 - Markers found with
 - Restriction-site Associated DNA (RAD) se
 - Microsat markers
- al., 2005).



HD Linkage Map: 2017-18

- GST population
- 40,186 SNPs- 22 LGs
- Female (1632.9 cM): Male (1359.6 cM) = 1.2:1

Experimental design

- G26 fish reared in 8 different batches in 2017-18
- Mating design 1:1
- After hatching, all the fingerlings are reared together
- Treated with hormones to produce an all-male population
- Grown for entire 30 week period and was harvested
- Phenotypes available for: body weight at harvest (BW), fillet weight (FW) and fillet yield (FY).





Pedigree

- Pedigree constructed using microsatellites
- 14 generation deep
- 1:1 mating -> only full-sibs

Genotypes

- DNA extracted – fin clips and genotyped using Oni50[®] Affymetrix Array.
- Raw dataset: 58,466 SNPs
- “PolyHighResolution” & “NoMinorHom” : 50,275 SNPs (86.75%)
- MAF <0.05: 48,960 SNPs (83.74%)
- Individual call rate <0.9 : 3 animals filtered
- Phenotypes, pedigree and genotypes available for 1444 animals
- 188 full-sib families with an average of 7.68 offspring per full-sib family (range 1 to 15; standard deviation = 4.48).

Statistical models

- Univariate and Multivariate models in DMUv6

Vector of phenotypes

Vector of fixed effects

Vector of random genetic effects

Vector of residual errors

$$\begin{bmatrix} \mathbf{y}_{BW} \\ \mathbf{y}_{FW} \\ \mathbf{y}_{FY} \end{bmatrix} = \begin{bmatrix} \mathbf{X} & 0 & 0 \\ 0 & \mathbf{X} & 0 \\ 0 & 0 & \mathbf{X} \end{bmatrix} \begin{bmatrix} \mathbf{b}_{BW} \\ \mathbf{b}_{FW} \\ \mathbf{b}_{FY} \end{bmatrix} + \begin{bmatrix} \mathbf{Z} & 0 & 0 \\ 0 & \mathbf{Z} & 0 \\ 0 & 0 & \mathbf{Z} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{BW} \\ \mathbf{u}_{FW} \\ \mathbf{u}_{FY} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_{BW} \\ \mathbf{e}_{FW} \\ \mathbf{e}_{FY} \end{bmatrix}$$

Fixed effects
 Batch,
 Difference of age during harvesting

Filleter for the traits FW & FY

Distributional assumption of random effects

$$\begin{bmatrix} \mathbf{u}_{BW} \\ \mathbf{u}_{FW} \\ \mathbf{u}_{FY} \end{bmatrix} \sim N \left(0, \begin{bmatrix} \sigma_{a_{BW}}^2 & \sigma_{a_{BW,FW}} & \sigma_{a_{BW,FY}} \\ \sigma_{a_{BW,FW}} & \sigma_{a_{FW}}^2 & \sigma_{a_{FW,FY}} \\ \sigma_{a_{BW,FY}} & \sigma_{a_{FW,FY}} & \sigma_{a_{FY}}^2 \end{bmatrix} \otimes \mathbf{A} \right)$$

Genetic covariances
Genetic variances

$$\begin{bmatrix} \mathbf{e}_{BW} \\ \mathbf{e}_{FW} \\ \mathbf{e}_{FY} \end{bmatrix} \sim N \left(0, \begin{bmatrix} \sigma_{e_{BW}}^2 & \sigma_{e_{BW,FW}} & \sigma_{e_{BW,FY}} \\ \sigma_{e_{BW,FW}} & \sigma_{e_{FW}}^2 & \sigma_{e_{FW,FY}} \\ \sigma_{e_{BW,FY}} & \sigma_{e_{FW,FY}} & \sigma_{e_{FY}}^2 \end{bmatrix} \otimes \mathbf{I} \right)$$

$$\mathbf{G} = \frac{\mathbf{H}\mathbf{H}'}{\sum_1^i \sum 2p_i(1-p_i)} \quad \text{VanRaden 2008}$$

\mathbf{H} is a centered marker matrix, the sum in the denominator is over all loci and p_i is the allelic frequency at locus i

Prediction accuracy

- 5 replicates of 10-fold cross-validation
- Random, Within-family and Between-family cross-validation
- Univariate and Multivariate approaches
- PBLUP and GBLUP models
- Prediction accuracy=
 $\text{cor}[(G)EBVs, \text{fixed effect corrected phenotype}]/\text{sqrt}(h^2)$
- $SE = \frac{1 - \text{prediction accuracy}^2}{\sqrt{\text{No. of validation animals} - 1}}$

Prediction bias

- regression coefficient of phenotypes adjusted for the fixed effects on GEBVs or EBVs
- regression coefficient = 1 indicates unbiased prediction
 - <1 indicates inflation of GEBV or EBV and
 - >1 indicates deflation of GEBV or EBV

Results

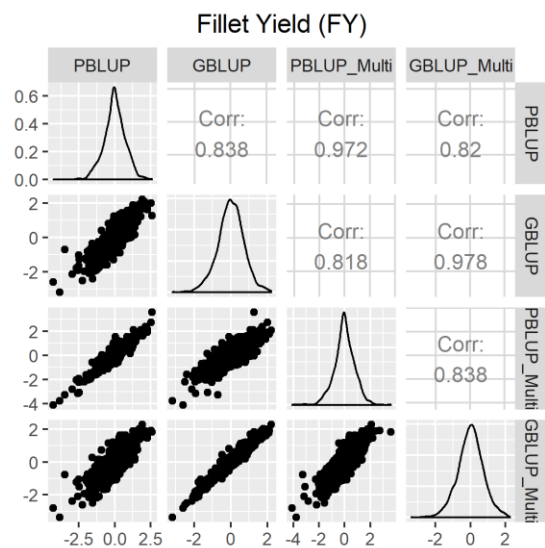
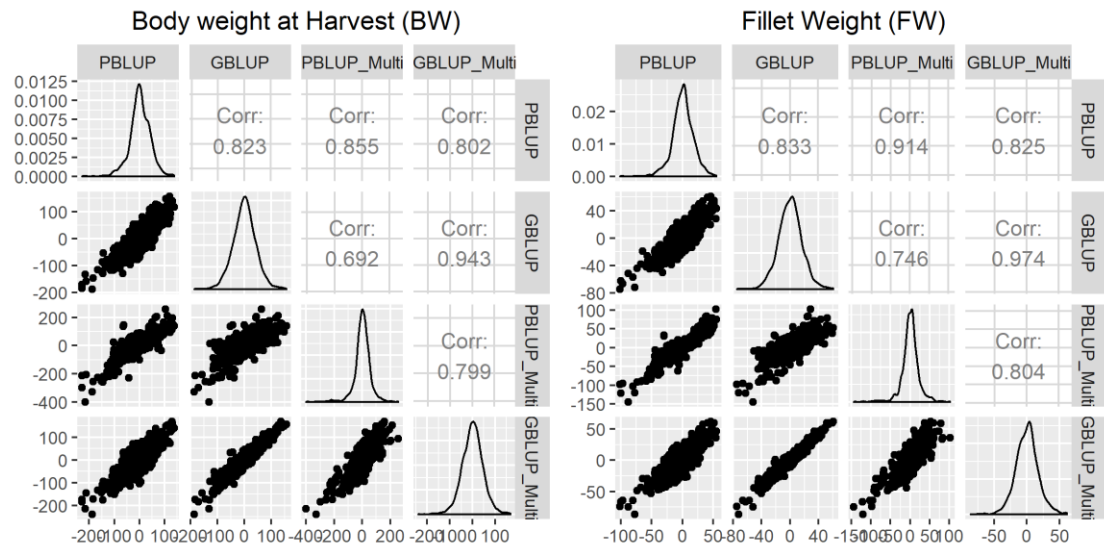
	Units	Min	Max	Median	Mean	Mean (SE)	SD	CV%
BW	g	138.70	1893.70	780.30	817.37	6.87	261.11	31.95
FW	g	39.10	754.60	284.25	300.01	2.82	107.34	35.78
FY	%	20.83	46.64	36.56	36.40	0.07	2.50	6.90

Heritability and correlations

Table: Heritability (diagonal), phenotypic correlation (above the diagonal) and genetic correlation (below the diagonal)

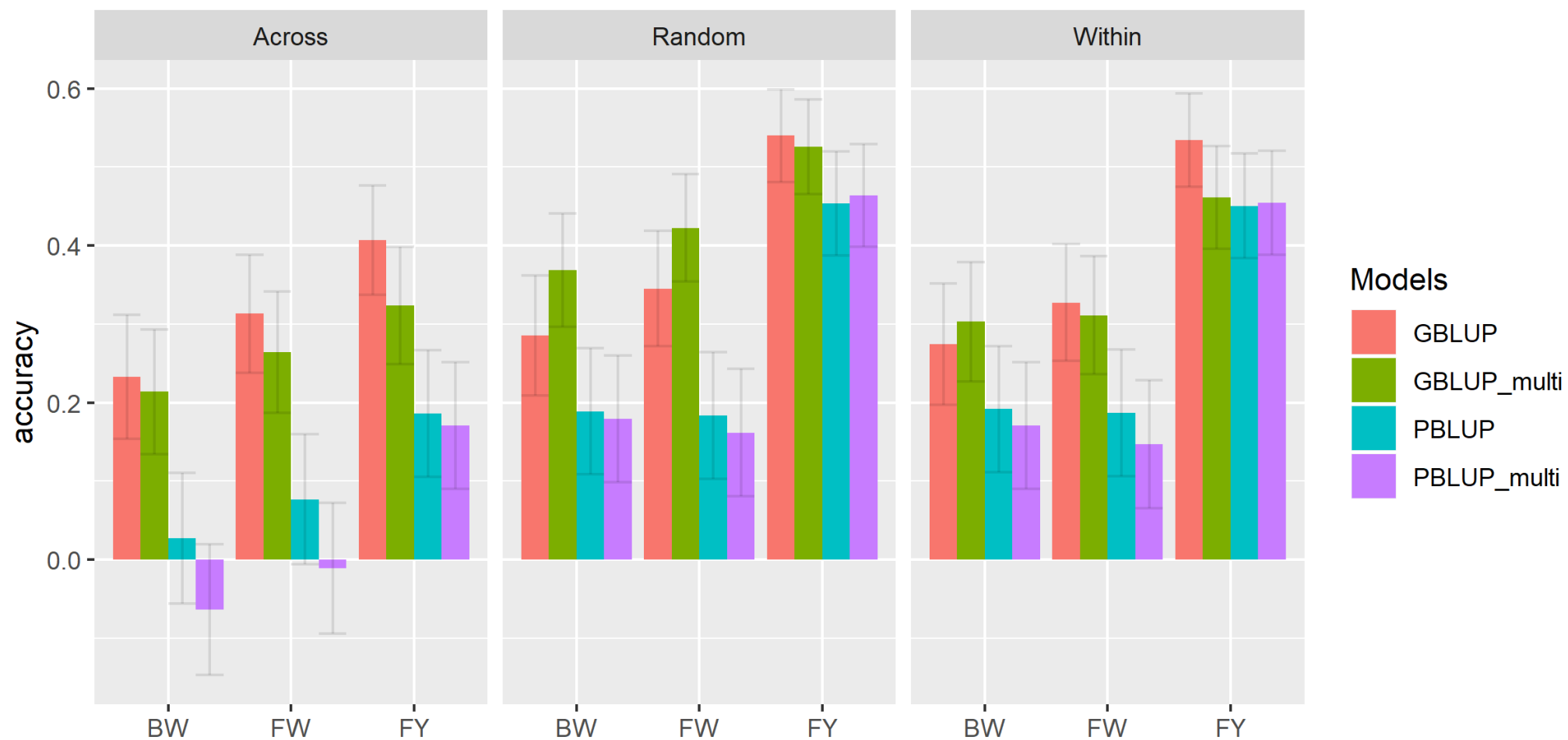
PBLUP	BW	FW	FY	GBLUP	BW	FW	FY
BW	0.28 ± 0.06	0.96 ± 0.01	0.23 ± 0.02	BW	0.19 ± 0.04	0.96 ± 0.01	0.23 ± 0.02
FW	0.96 ± 0.01	0.27 ± 0.07	0.47 ± 0.02	FW	0.96 ± 0.01	0.17 ± 0.04	0.47 ± 0.02
FY	-0.04 ± 0.17	0.23 ± 0.16	0.33 ± 0.07	FY	-0.11 ± 0.15	0.19 ± 0.15	0.23 ± 0.04

Impact on genetic evaluations

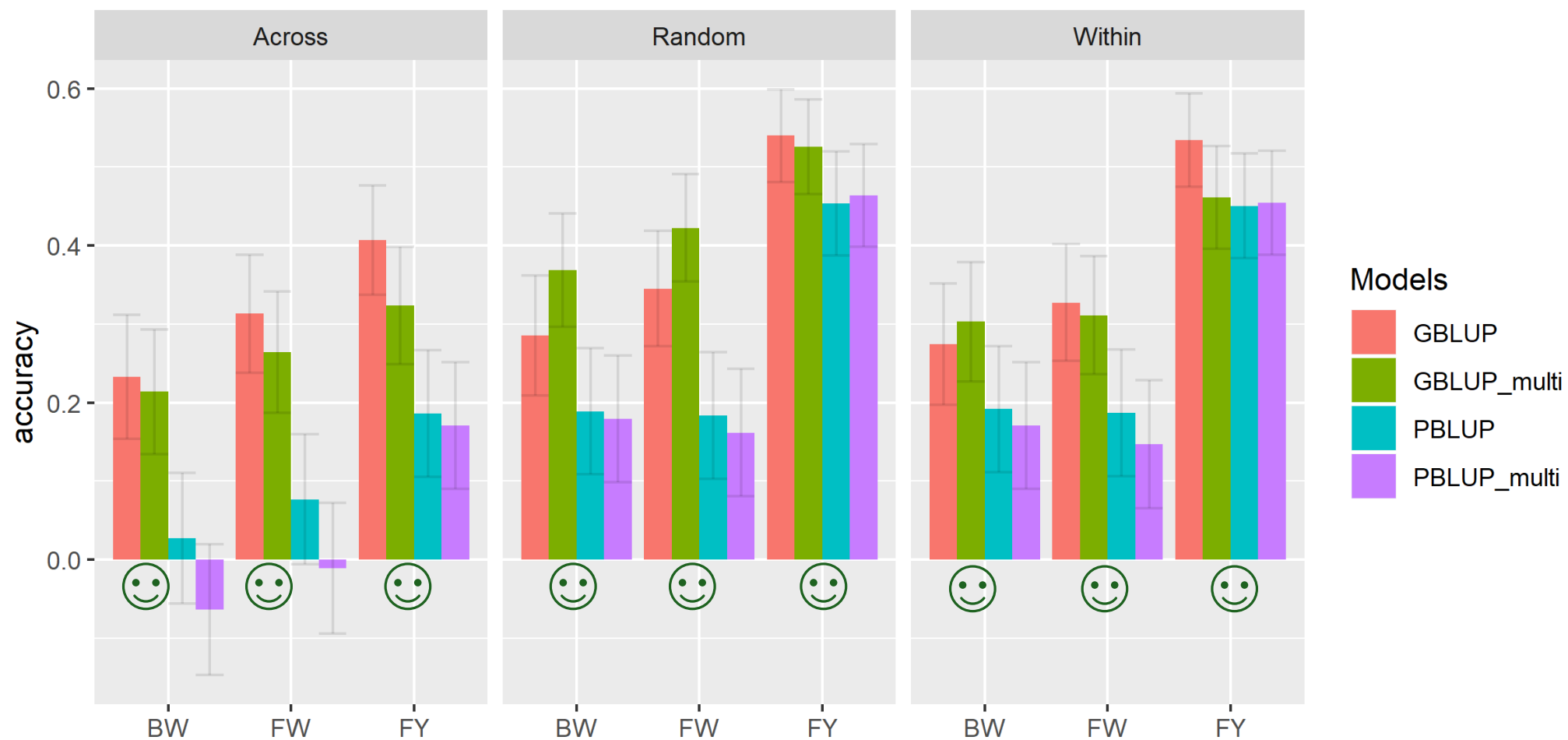


	BW	FW	FY
<i>PBLUP</i>	0	0	0
<i>GBLUP</i>	36	32	45
<i>PBLUP_Multi</i>	27	15	18
<i>GBLUP_Multi</i>	36	33	46

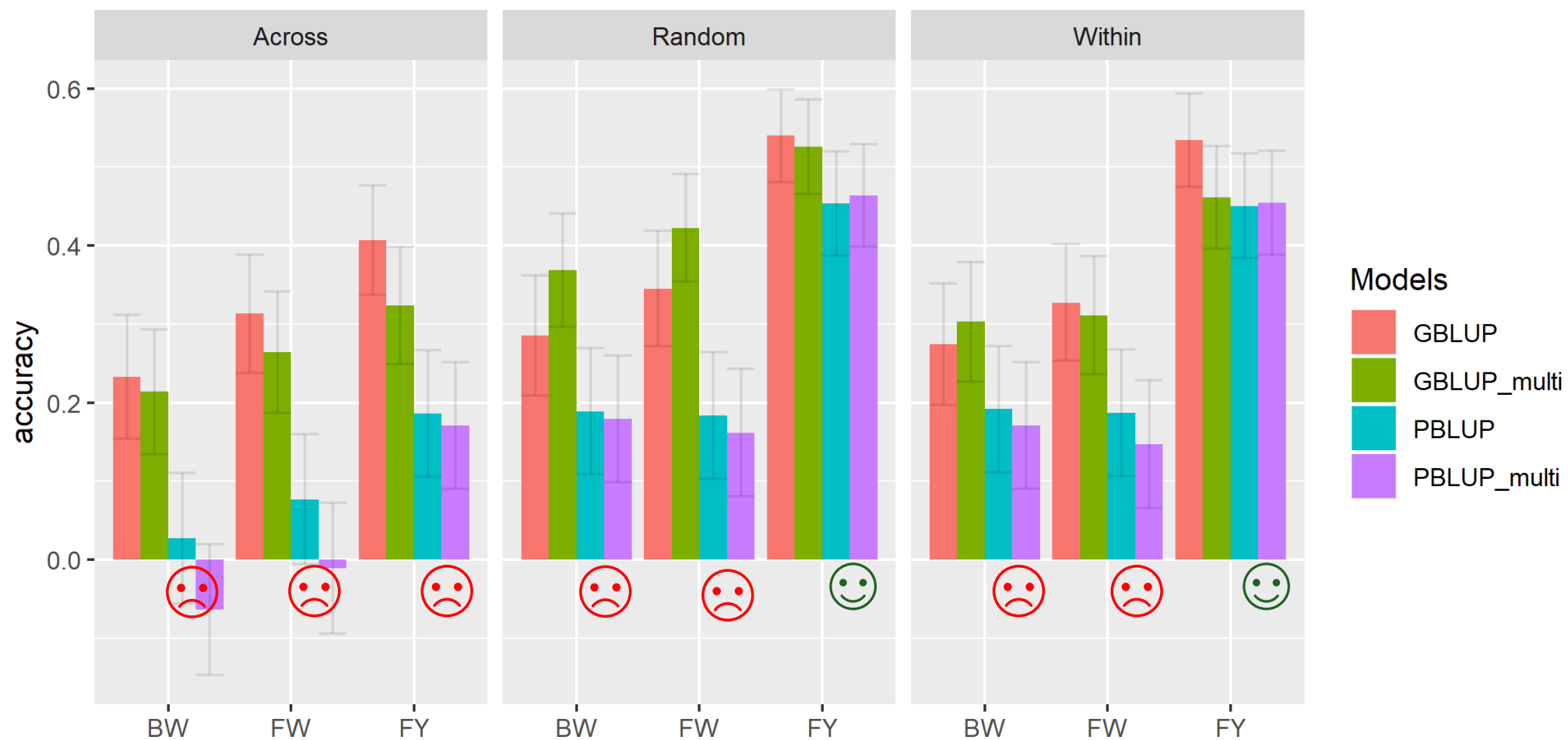
Prediction accuracy



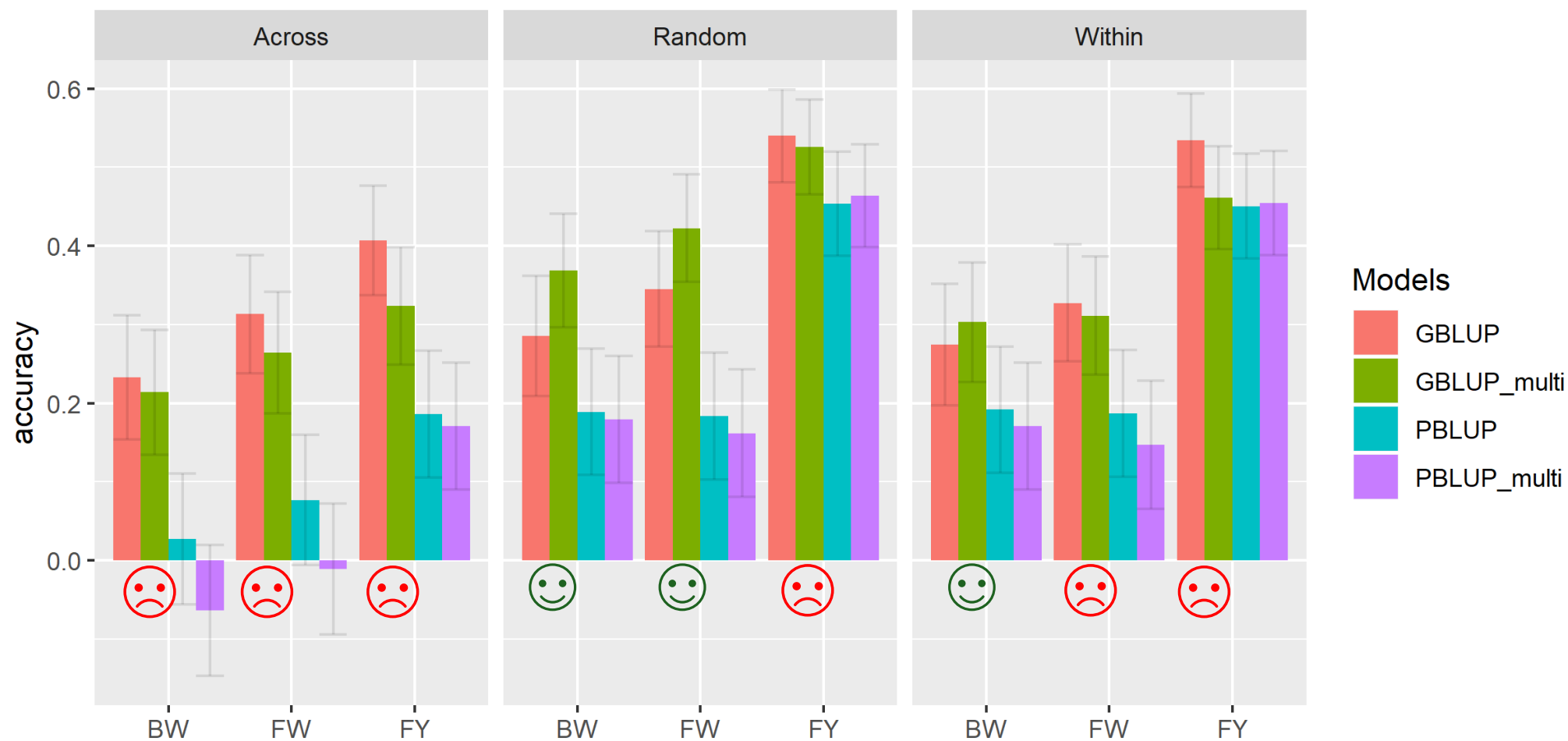
Prediction accuracy.. GBLUP vs PBLUP



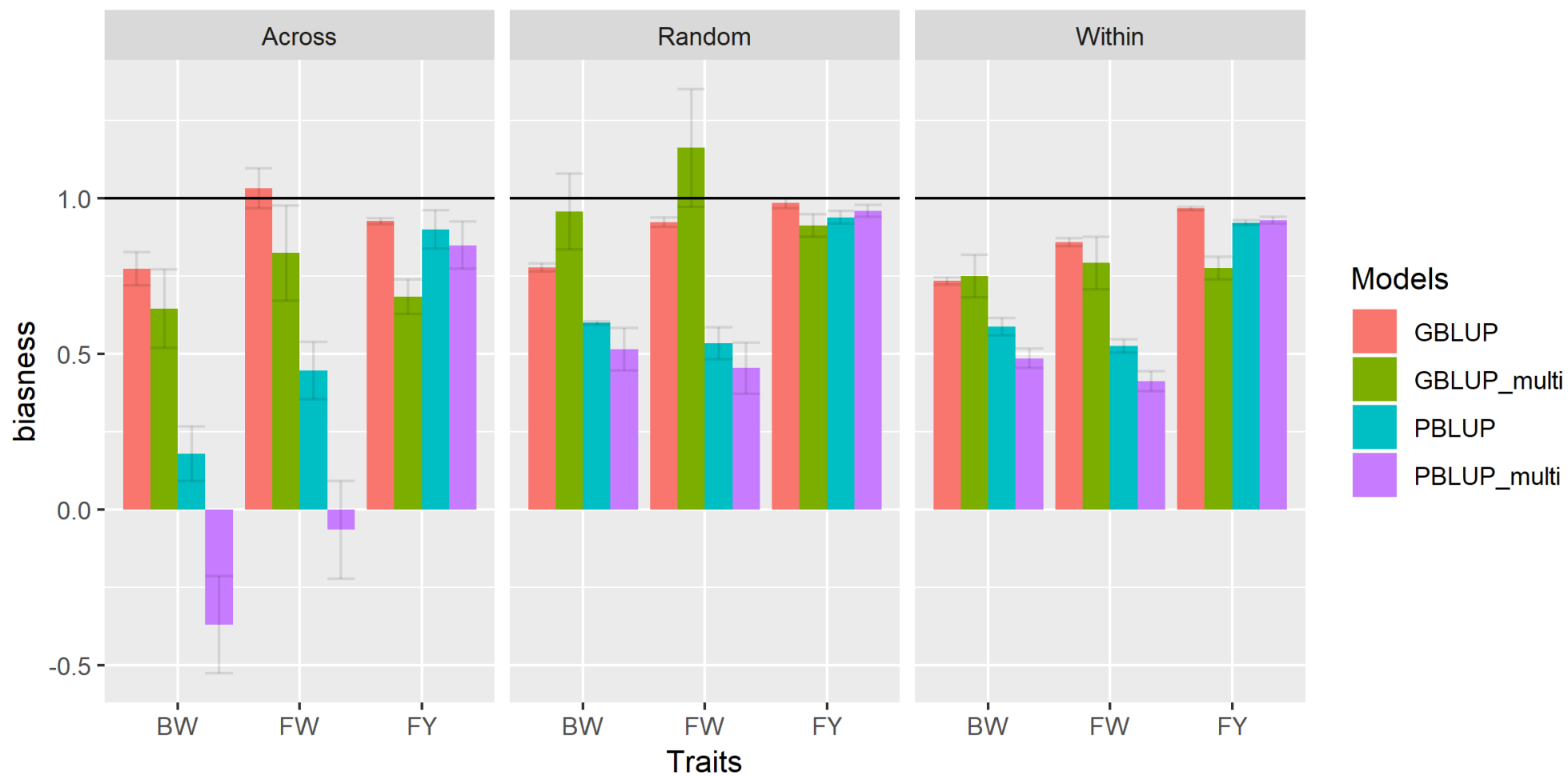
Prediction accuracy.. PBLUP vs PBLUP_multi



Prediction accuracy.. GBLUP vs GBLUP_multi



Prediction bias





Conclusions

- Genomic selection is beneficial to Nile tilapia breeding program
- It is recommended to use GBLUP univariate approach
- Selection index- both BW and FY