Lessons from the application of genomic selection to salmonids aquaculture

Yniv Palti

USDA-ARS
National Center for Cool and Cold Water Aquaculture, Leetown, WV, USA

Some Historic Perspective

- **1976** - **Soller et al.** proposed a statistical approach for mapping quantitative trait loci (QTL) in segregating crosses. They predict that 1,000 individual animals with phenotypes and genotypes will be needed to detect a QTL explaining as little 1% of the genetic variation for the trait.
- **1983 – Beckman and Soller** discussed the idea of using RFLP markers for genetic improvement through marker assisted selection in agriculture animals: **"In most cases the anticipated costs appear to be commensurate with the scientific or economic value of the application"**.
- 2001 – Working draft of the Human Genome was published and **Meuwissen et al.** published a paper proposing the idea behind Genomic Selection.
- 2007 – Public release of the Bovine 50K SNP chip.
- 2009 – USDA and US dairy cattle breeders associations implement genomic evaluations in the selection of top bulls for breeding.

"It's tough to make **predictions**, especially about the **future**", Yogi Berra.

What is Marker Assisted Selection (MAS)?

- Marker-assisted selection can be used to select directly for favorable QTL alleles.
- It can be used for individual selection of genotyped animals even in absence of phenotyping.
- It can only be used effectively if the QTL effects are known in the population and carriers of the favorable QTL alleles can be identified through markers that are in very strong linkage disequilibrium (LD) with the QTL.
- For European Atlantic salmon, the method has been utilized with great success in selection for resistance to infectious pancreatic necrosis virus (IPNV), for which a single QTL explains most of the genetic variation. (Houston et al. 2008; Moen et al. 2009)

* Definition modified from Odegard et al. 2014, Frontiers in Genetics, Volume 5, Article 402.

What is Genomic Selection (GS)?

- Genomic Selection is an alternative for MAS when the genetic architecture of the trait is more complex and several to multiple loci have moderate to small effects.
- It is utilizing genotype information from a genome-wide scan with a large number of markers which are used jointly in the genetic analysis.
- It facilitates estimation of individual breeding values for breeding candidates using a “training” dataset from animals with phenotype and genotype data.
- It does not require prior knowledge of the QTL for the trait in the population.

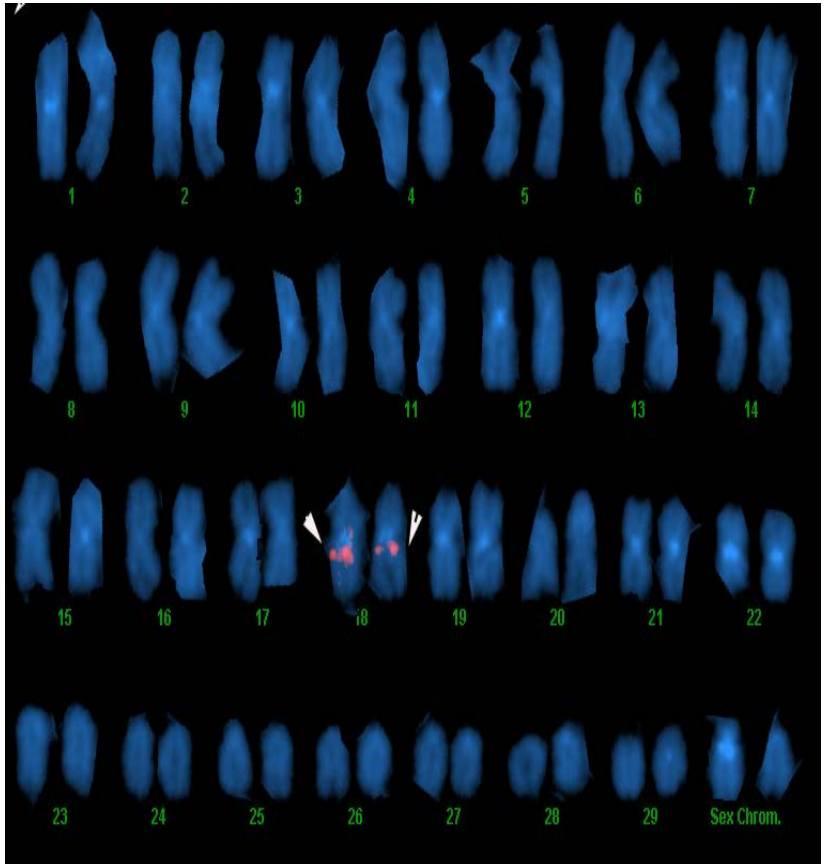
* Definition modified from Odegard et al. 2014, Frontiers in Genetics, Volume 5, Article 402.

What kind of traits can be improved faster using MAS or GS in aquaculture systems?

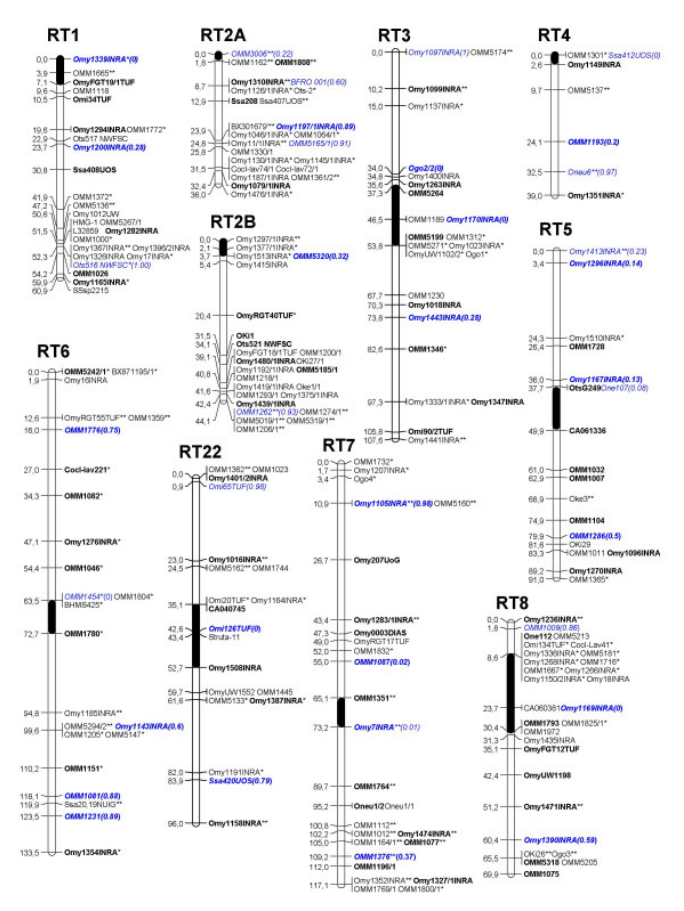
- Traits that cannot be measured directly on the breeding candidates, such us resistance to diseases and pathogens, fillet or muscle yield and flesh texture and color.
- MAS or GS allows for estimation of the genetic merit of the individual breeding candidates for those traits without phenotyping, using only their genotype data.
- Examples of traits from current commercial breeding programs in Norway and the USA include IPNV, Sea Lice, Pancreas Disease (PD) and Flesh Color for Atlantic salmon; and IPNV, Bacterial Cold Water Disease and Columnaris Disease in rainbow trout.

The Genetic Linkage Map Corresponds to the Actual Chromosomes

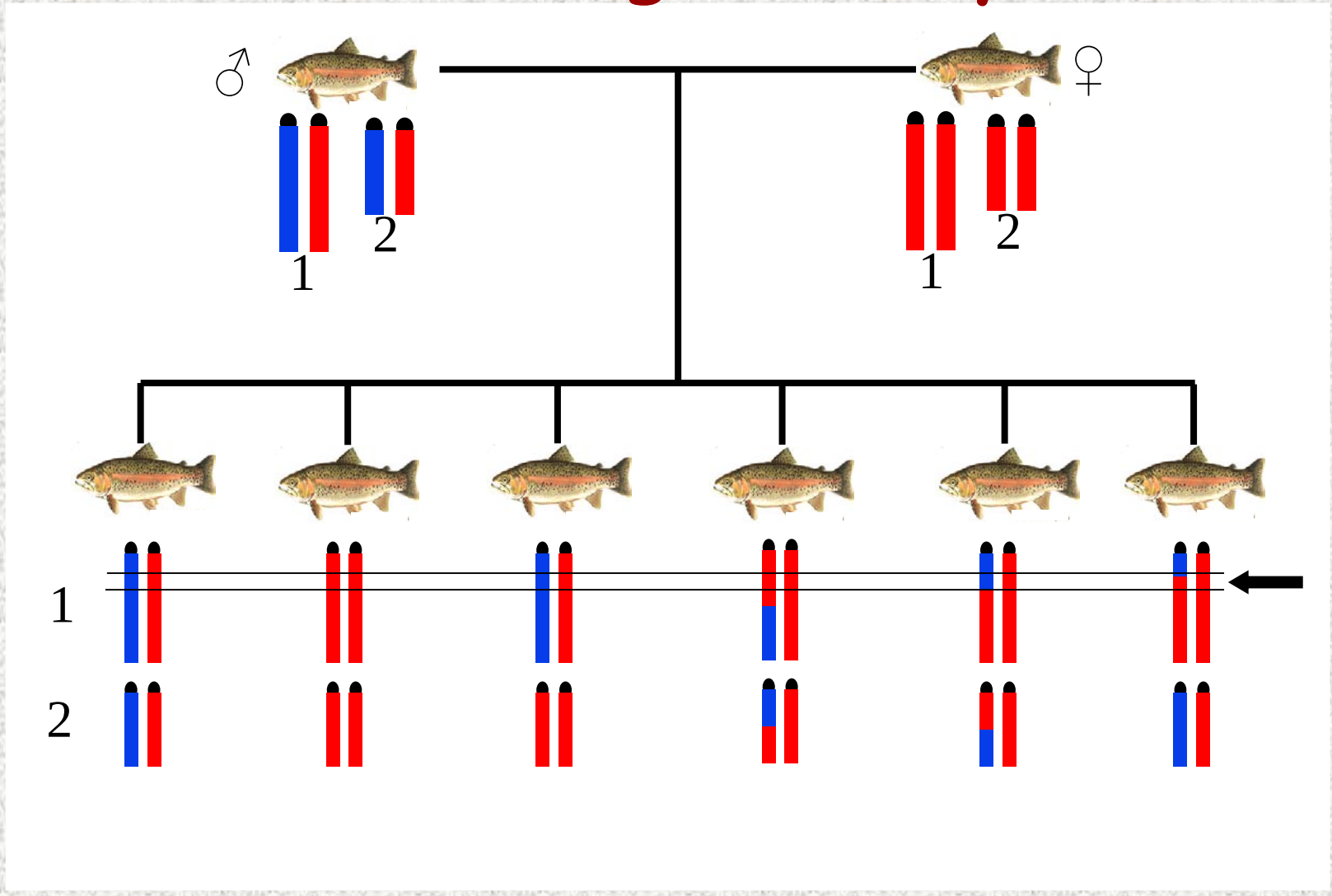
Chromosomes Karyotype



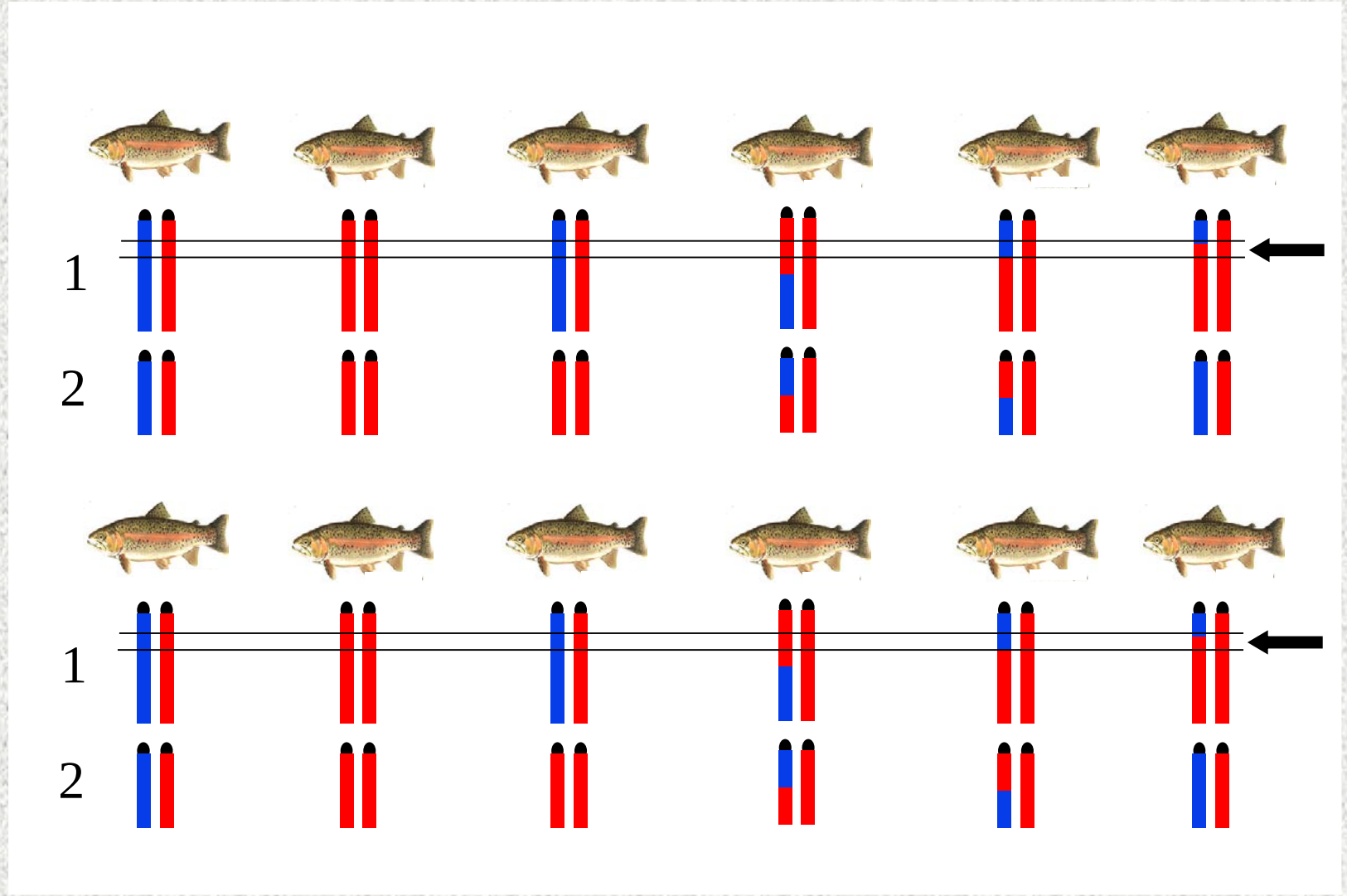
Linkage Groups



Mapping Disease Survival Trait In a Single Family



Linkage Disequilibrium Mapping



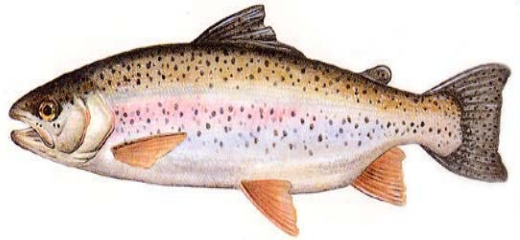
Lessons from Genome-Wide Association Analysis and Evaluation of Genomic Selection for Bacterial Cold Water Disease Resistance in Rainbow Trout

Yniv Palti, Roger Vallejo, Tim Leeds

USDA-ARS
National Center for Cool and Cold Water Aquaculture, Leetown, WV, USA

BACTERIAL COLD WATER DISEASE (BCWD)

- High priority disease problem in US trout aquaculture
- Caused by *Flavobacterium psychrophilum (Fp)*
- There is no licensed vaccine available
- Treatment is with antibiotics: Florfenicol & Oxytetracycline
- Antimicrobial resistance is a growing concern

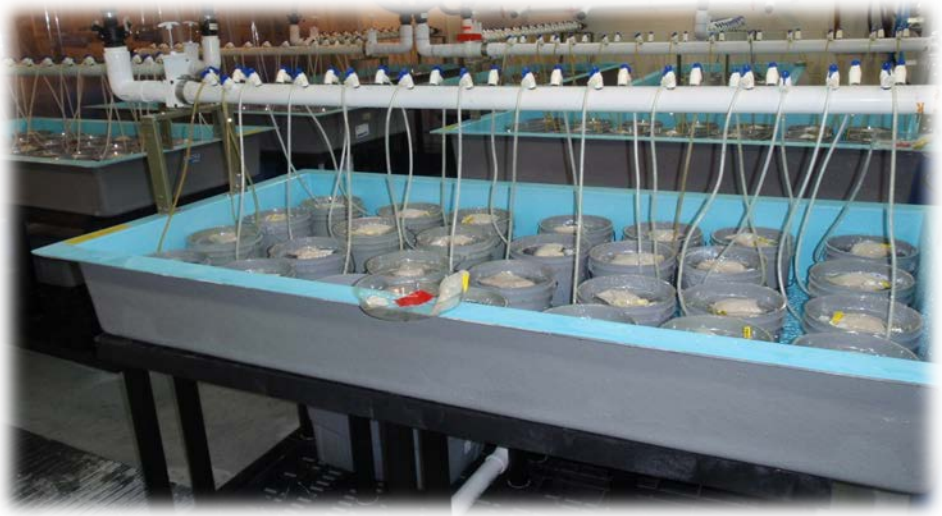


BCWD SYMPTOMS

- Fry-lethargy, lack of feeding, darkened skin, enlarged spleen, anemia & high mortality

BCWD RESISTANCE PHENOTYPES

- **Binary survival status (STATUS)**
  - 1= fish died during the 21 d post-challenge evaluation
  - 2= fish was alive on day 22 post-challenge



Genome Wide Association Study

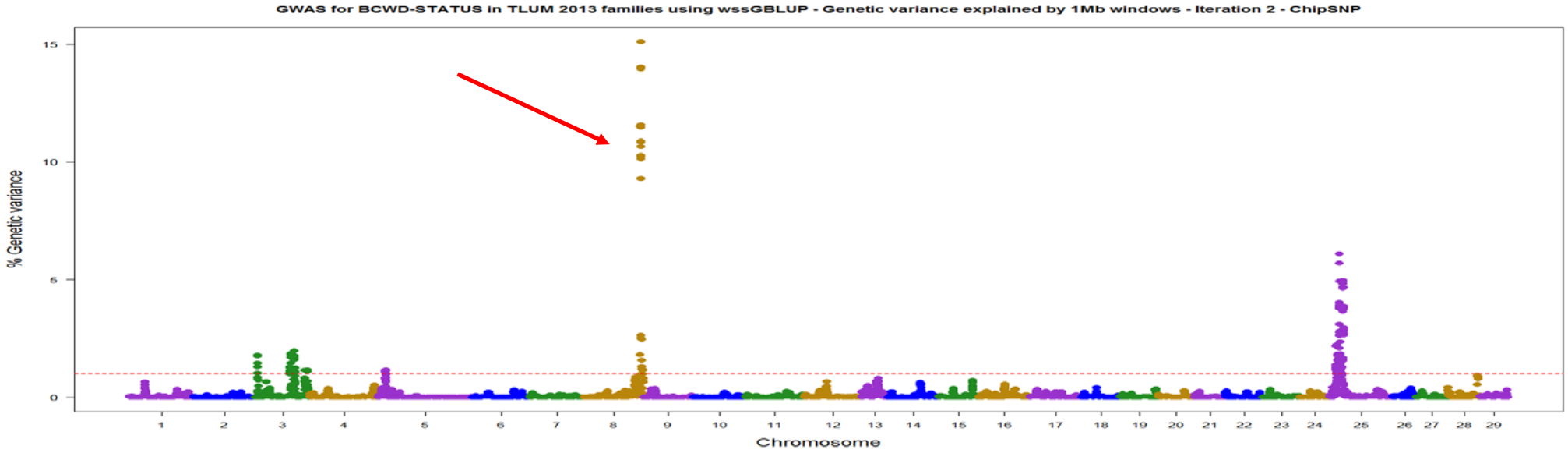
Vallejo et al., manuscript in preparation, to be submitted soon for publication in the peer reviewed journal G3.

Genome Wide Association Study - Design

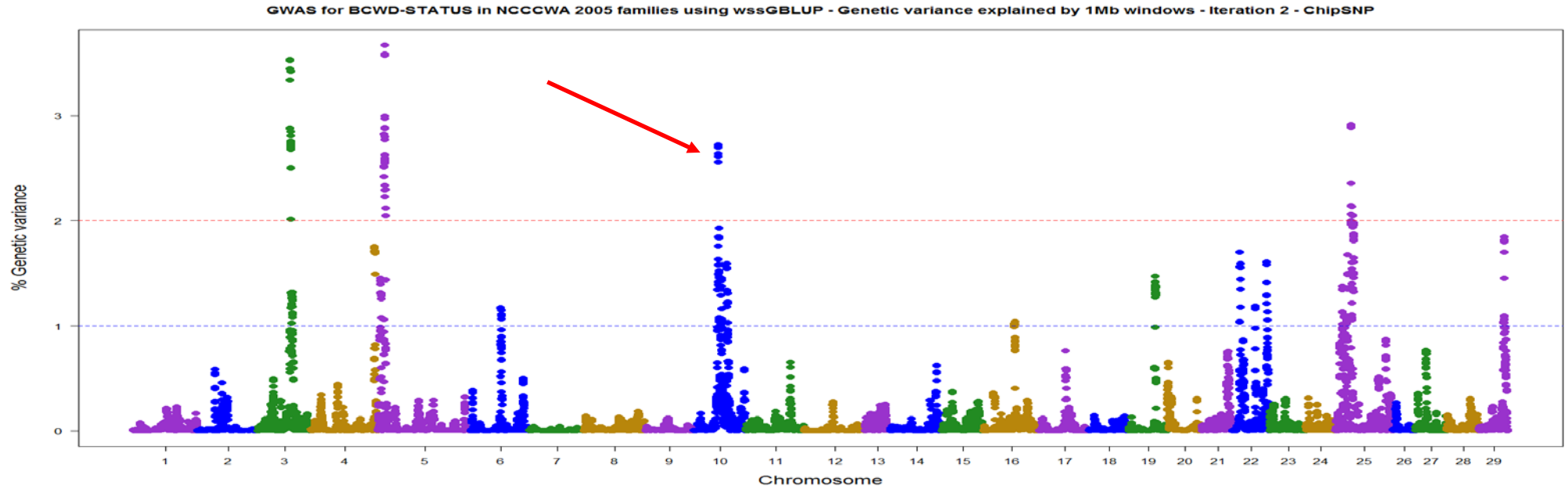
- TLUM2013 population: N=1,473 (57K SNP chip genotyping).
- NCCCWA2005 population: N=577 Genotyped with the SNP chip and with RAD SNPs (a genotyping by sequencing method)
- Data analyses with Single-step GBLUP (**ssGBLUP & WssGBLUP**) (**BLUPF90**; Aguilar et al., 2010) and the Bayesian method **BayesB** (**GENSEL**; Fernando & Garrick, 2009).
- The BayesB method used **1Mb** exclusive-consecutive windows and the wssGBLUP method used **1Mb** moving-sliding windows. (new trout genome)
- **Only showing results from WssGBLUP analyses.**

Comparing results between the two populations:
Similar genetic architecture, but not necessarily the same QTL or the same effect even if QTL location in the genome is similar.

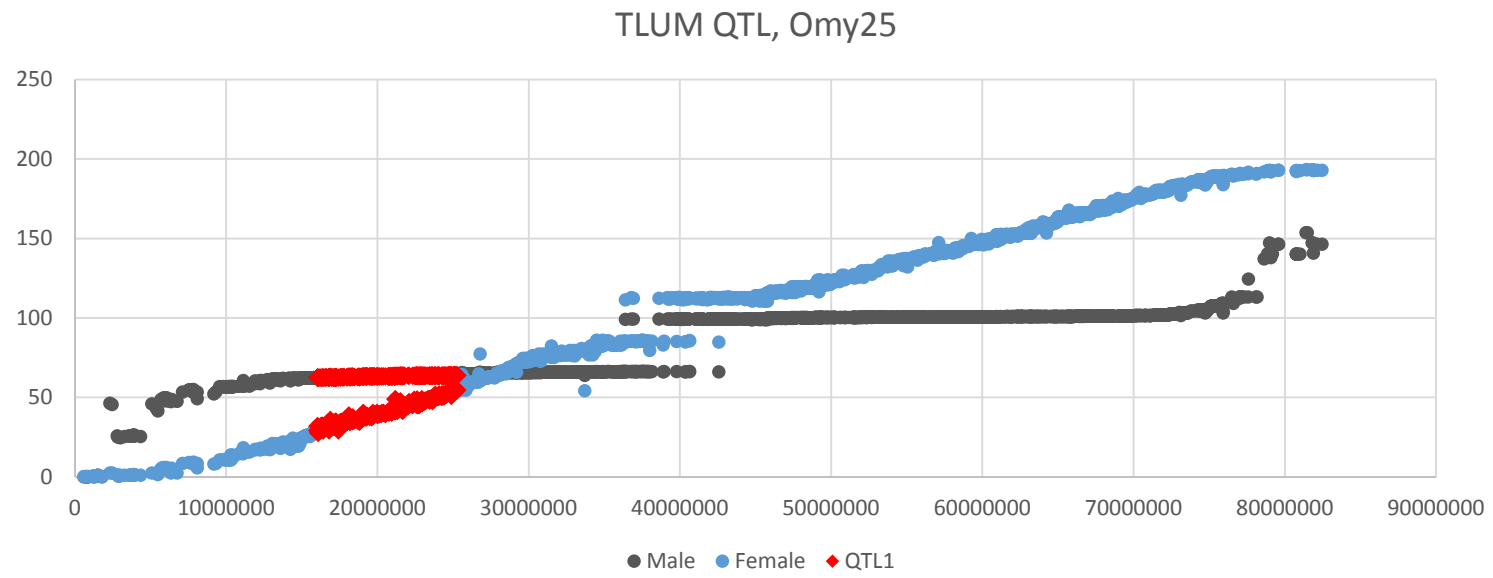
Scale Not The Same

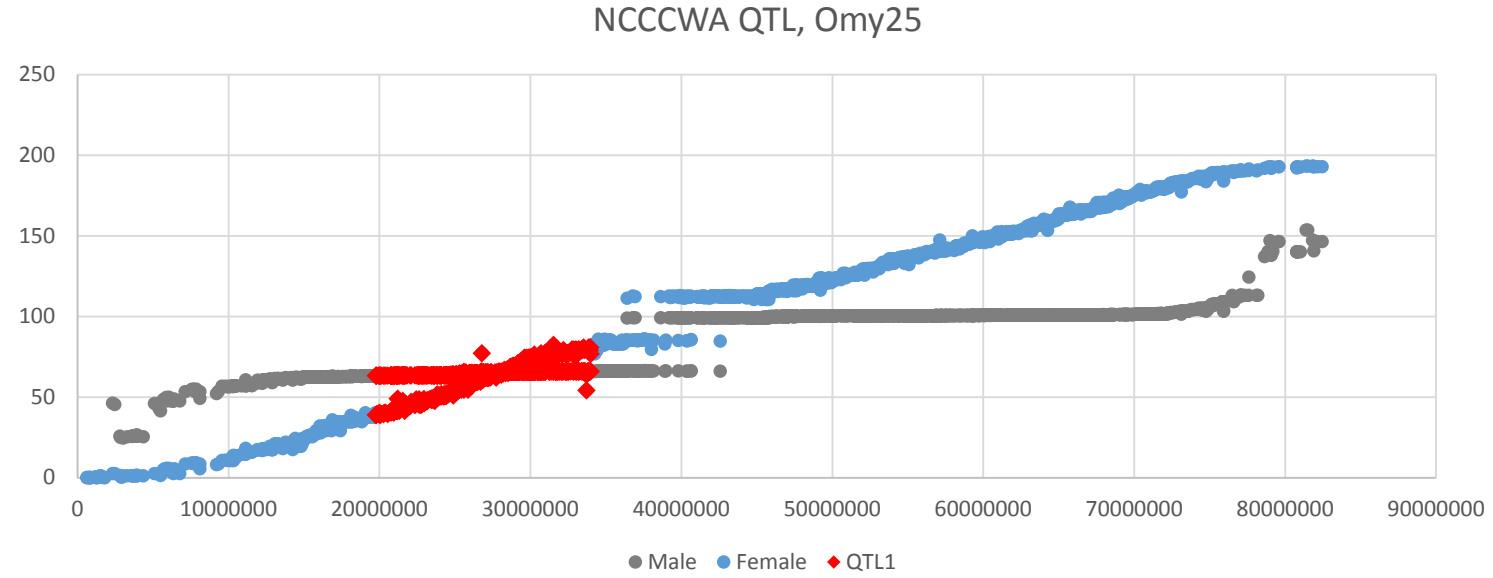


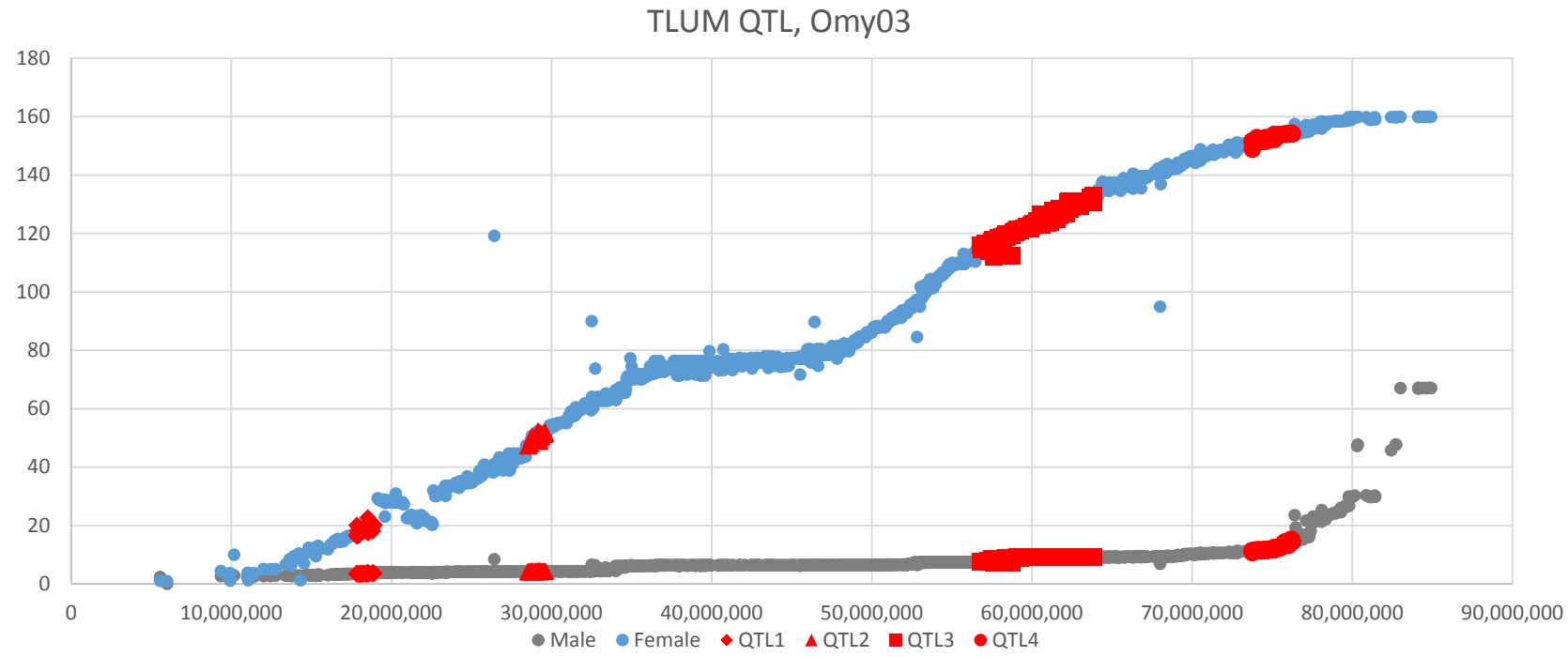
TLUM 2013

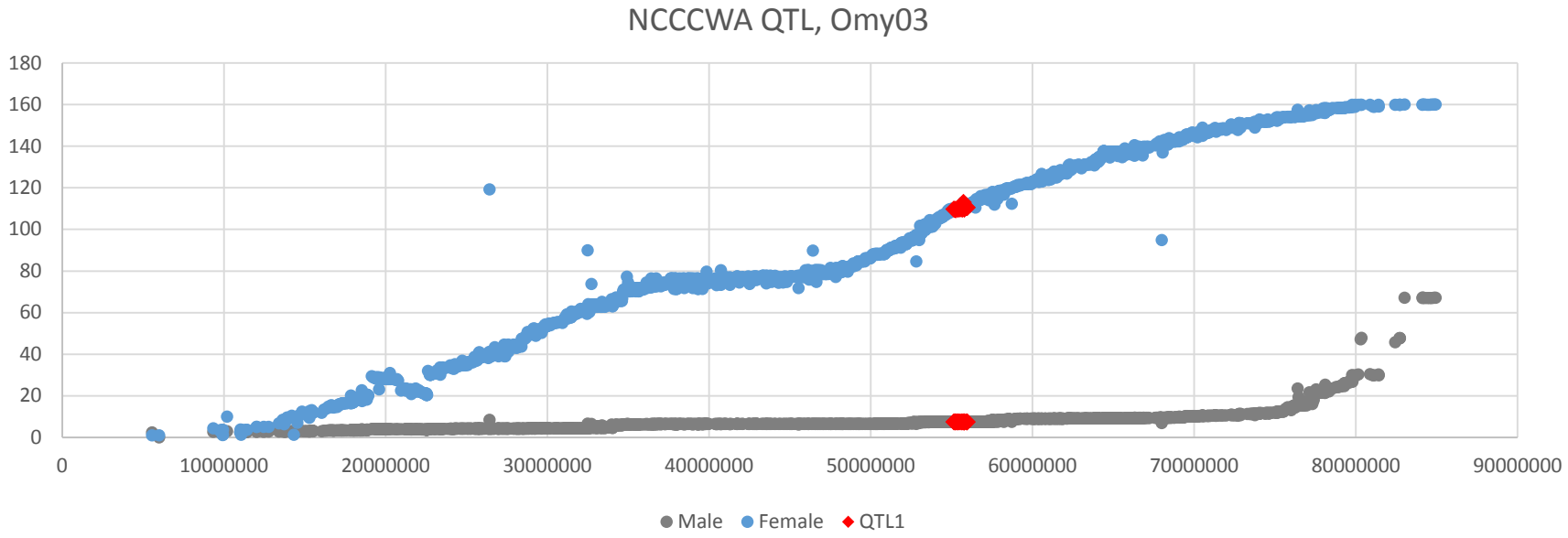


NCCCWA 2005


TLUM QTL, Omy25
250
200
150
100
50
0
0
10000000
20000000
30000000
40000000
50000000
60000000
70000000
80000000
90000000
Male
Female
QTL1


NCCCWA QTL, Omy25
250
200
150
100
50
0
0
10000000
20000000
30000000
40000000
50000000
60000000
70000000
80000000
90000000
Male
Female
QTL1

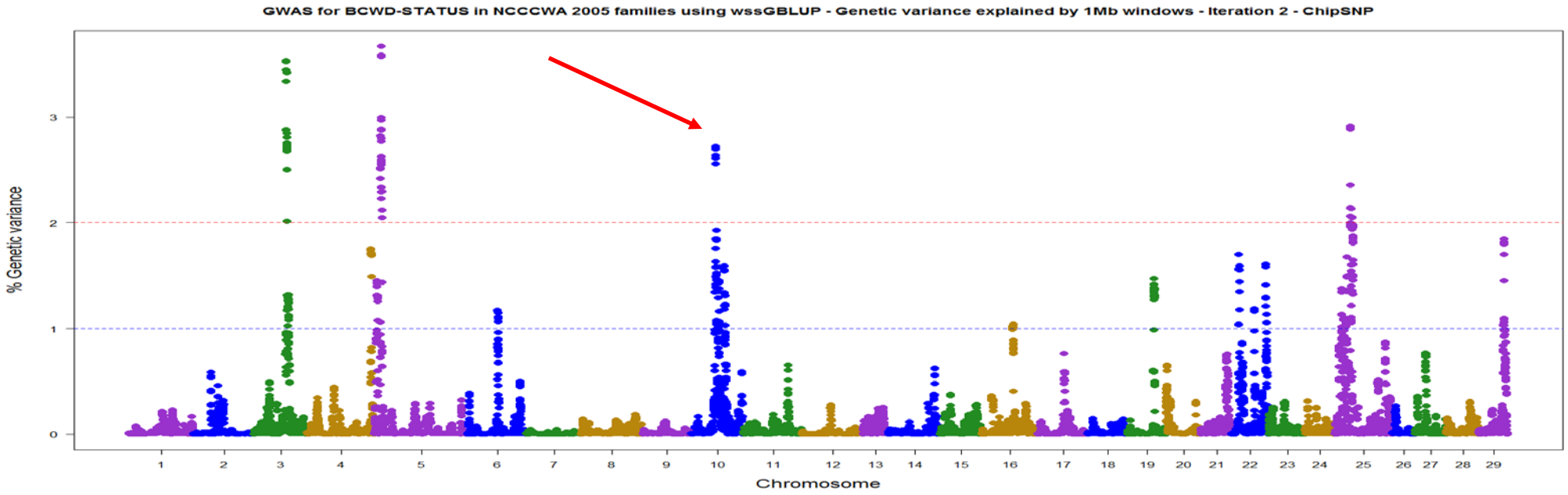

TLUM QTL, Omy03
180
160
140
120
100
80
60
40
20
0
0
10,000,000
20,000,000
30,000,000
40,000,000
50,000,000
60,000,000
70,000,000
80,000,000
90,000,000
Male
Female
QTL1
QTL2
QTL3
QTL4


NCCCWA QTL, Omy03
180
160
140
120
100
80
60
40
20
0
0
10000000
20000000
30000000
40000000
50000000
60000000
70000000
80000000
90000000
Male
Female
QTL1

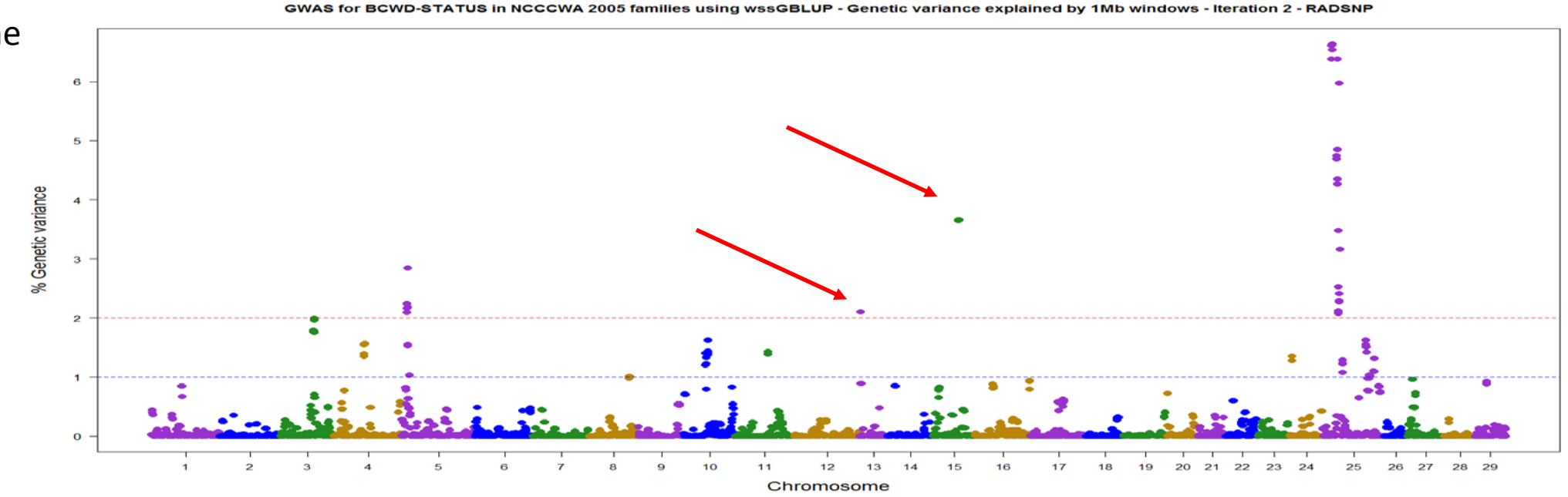
Comparing results between the two SNP genotyping platforms:

More QTL were detected using RAD SNP genotyping, but some were detected only with the SNP chip and genome coverage was better with the chip.

Scale Not The Same



SNP Chip



RAD SNPs

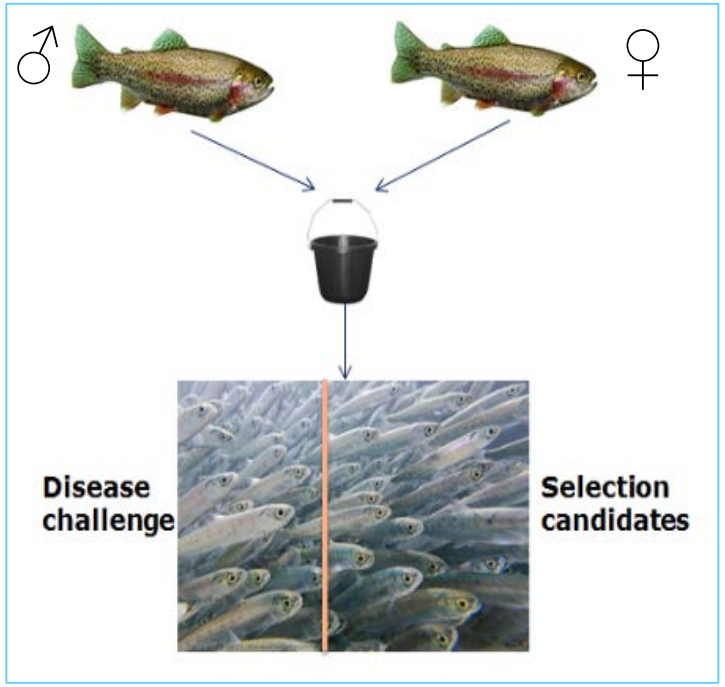
GWAS Results Summary

- Although the genetic architecture of the trait was similar, only six of 17 QTL were shared by the two populations.
- Overall, the WssGBLUP detected higher number of QTL than the BayesB and both GWAS models did not detect the same QTL which highlights the utility of using two different GWAS algorithms.
- The RAD genotyping platform detected higher number of QTL than the Chip technology and overall both genotyping platforms did not detect the same QTL in the NCCCWA population.
- **Sampling: N ≥ 1,000 is strongly recommended and balanced sampling from all families in the population is crucial.**

Evaluation of Genomic Selection for BCWD

Published: Vallejo *et al. Genet Sel Evol (2017) 49:17*
DOI 10.1186/s12711-017-0293-6

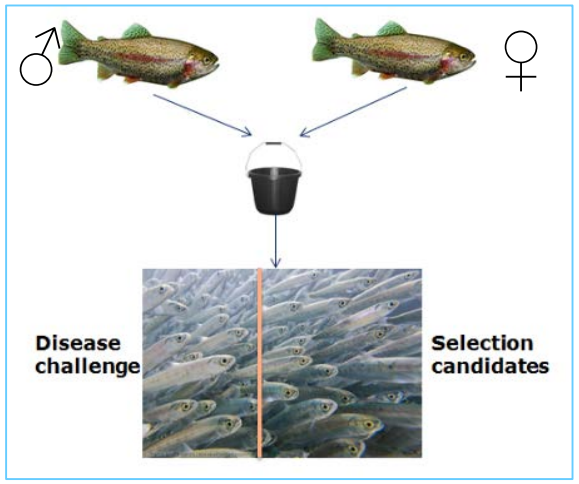
Generation 1



TRAINING
Genotyping and Phenotyping

Individual breeding values for Testing Fish based on Phenotype and Genotype data from the Training fish.

TESTING
Genotyping Only



Generation 2

Progeny Testing

- **Rainbow trout growth strain TLUM (Troutlodge, Inc.; USA; May)**

- **TRAINING sample:**
  - Offspring from 2013 year-class (YC)
  - Full-sib Families = 102
  - Phenotyped & Genotyped animals = 1,473 (from 50 families)

- **TESTING sample:**
  - Offspring from 2013 YC families (full-sibs of training animals)
  - Families = 25
  - Genotyped animals = 920 (Selection candidates with GEBVs)

- **Genotyping platform:**
  - SNP57K chip (Affymetrix: Axiom® Trout Genotyping Array)

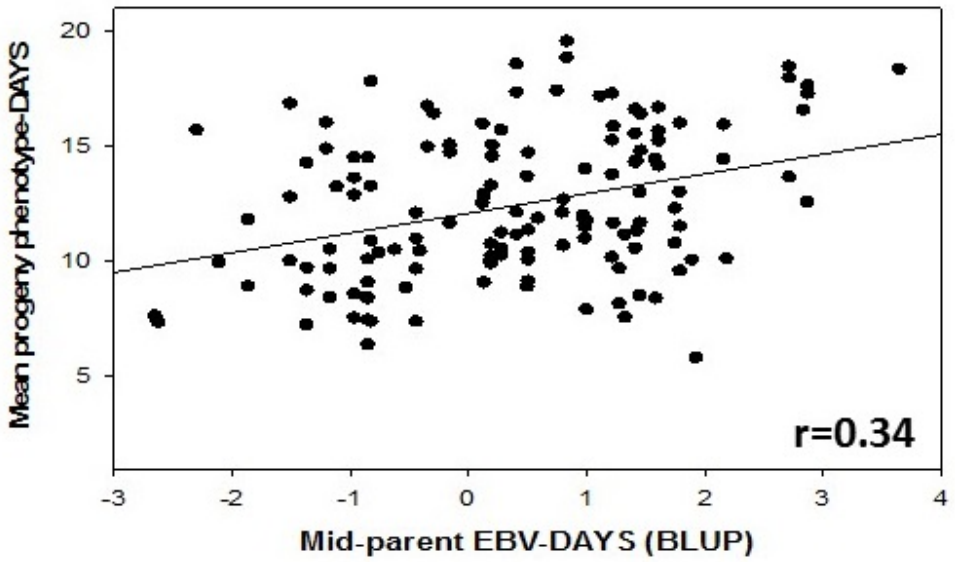
Pedigree & GS models

- Pedigree-based model (**BLUP**) (**BLUPF90**; Misztal et al., 2002, 2014)

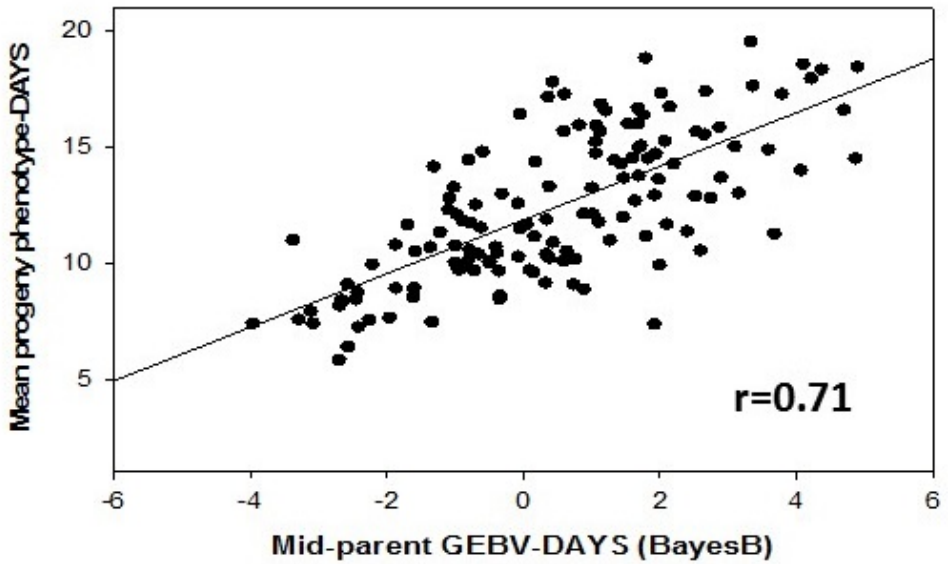
- GS Models:
- Single-step GBLUP (**ssGBLUP & WssGBLUP**) (**BLUPF90**; Aguilar et al., 2010)
- Or
- Bayesian method **BayesB** (**GENSEL**; Fernando & Garrick, 2009)

Correlation of Genetic predictions with disease survival performance of progeny

Pedigree EBVs

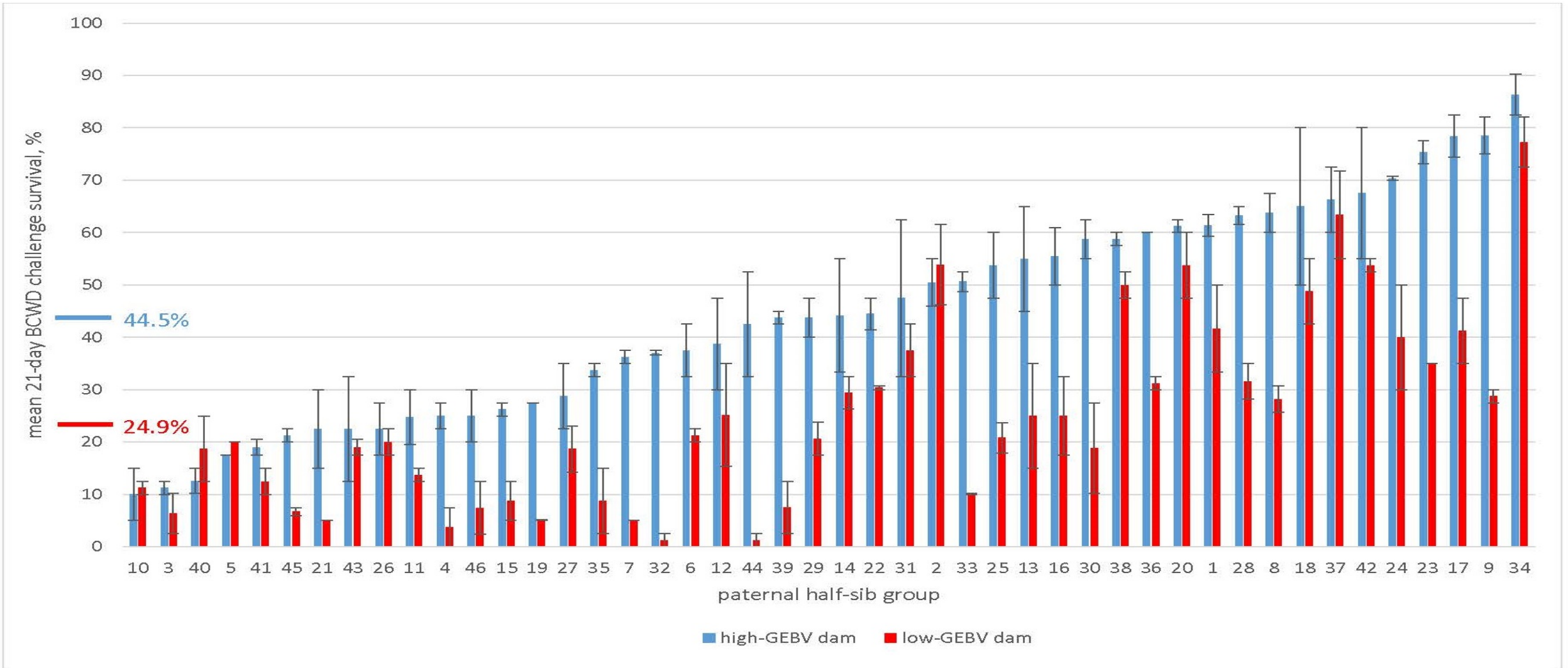


Genome Predictions



*** Genomic selection doubles the accuracy compared with traditional pedigree-based predictions!**

Survival of offspring from High-GEBV Dams is almost 80% better than their Low-GEBV sisters.



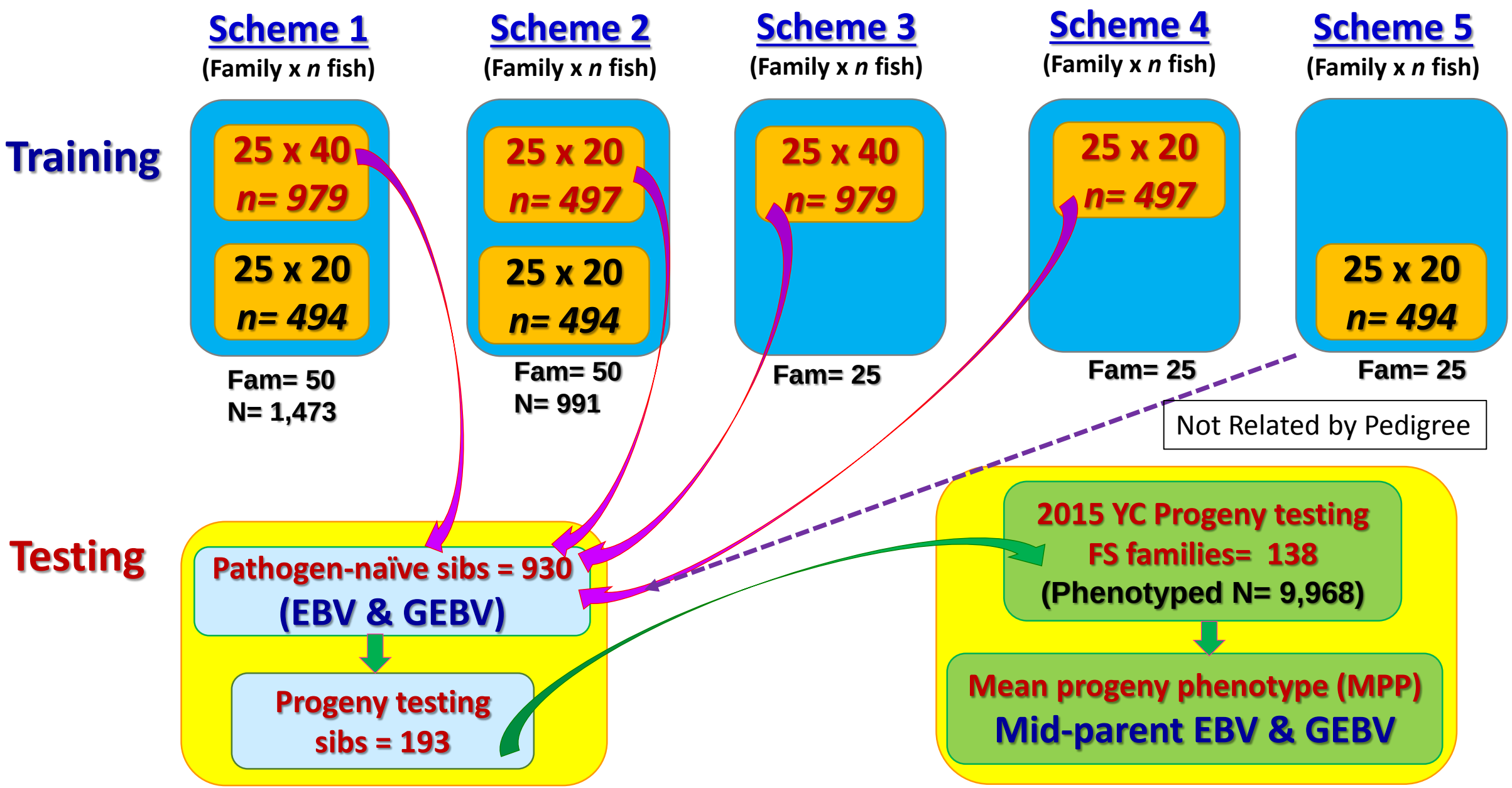
Impact of the experimental design on the accuracy of the genomic-based predictions

Now we are looking to reduce the cost of genotyping in two ways:

1. What will be the effect of reducing the samples size?

2. What will be the effect of using a smaller number of SNP markers?

GS schemes for BCWD resistance tested with BayesB



Scheme	Family		Training size	Training-testing fish relationship[1]	Accuracy
	#	Size			
1	50	20-40	**1,473**	0.66	**0.71**
2	50	20	991	0.50	0.67
3	25	40	979	1.00	**0.72**
4	25	20	497	1.00	0.61
5	25	20	494	0.00	**0.22**

[1]**Proportion of TRAINING fish that were full-sibs of TESTING fish**

The effect of the SNP density on the Accuracy of predicted GEBV for BCWD resistance in the TLUM2013 population

SNP Density	Effective SNPs	Accuracy (Status)
45195	41868	0.71
10000	9655	0.66
3000	2899	0.65
1000	964	0.57
500	485	0.50
300	292	0.49
200	194	0.48
70 QTL	70	0.66
BLUP (Pedigree Only)	N/A	0.36

Whole-Genome Selection Vs. MAS with QTL flanking Markers: Which should I use for getting the biggest "bang for the buck"?

1. Conduct GWAS for the specific trait of interest in the specific breeding population

2. Is large proportion of the genetic variation (>30%) explained by small number of major QTL (<10)?

3.

- Yes. (e.g. BCWD or IPNV resistance)
  - One may consider using markers flanking the QTL for genomic value predictions or for marker assisted selection.
- No. (e.g. Fillet Yield or Sea Lice resistance)
  - Marker assisted selection is not going to work. Whole-genome selection is the way to go.

ACKNOWLEDGEMENT

NCCCWA

- Roger Vallejo
- Tim Leeds
- Jason Evenhuis
- Guangtu Gao
- Greg Wiens
- Sixin Liu
- Tim Welch

Troutlodge, Inc.

- Jim Parsons
- Kyle Martin

Laboratory Mol. Biol.

- Roseanna Long
- Kristy Shewbridge
- Cassandra Parker
- Bryce Williams

Fish Caretakers

- Ryan Lipscomb
- Travis Moreland
- Clayton Birkett
- Josh Kretzer

University of Georgia

- Ignacy Misztal
- Breno Fragomeni

Iowa State University

- Dorian Garrick
- Rohan Fernando

USDA UNITED STATES DEPARTMENT OF AGRICULTURE

Agricultural Research Service