Received: 21 May 2022 | Revised: 3 October 2022 | Accepted: 3 October 2022

DOI: 10.1111/raq.12750

REVIEW

REVIEWS IN Aquaculture

Genome-wide association and genomic selection in aquaculture

José M. Yáñez[1,2] | Agustín Barría[3] | María E. López[4] | Thomas Moen[5] | Baltasar F. Garcia[1] | Grazyella M. Yoshida[1] | Peng Xu[6,7]

[1]Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Santiago, Chile

[2]Millennium Nucleus of Austral Invasive Salmonids (INVASAL), Concepción, Chile

[3]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh Easter Bush, Midlothian, UK

[4]Department of Aquatic Resources, Institute of Freshwater Research, Swedish University of Agricultural Sciences, Drottningholm, Sweden

[5]AquaGen, Trondheim, Norway

[6]Fujian Key Laboratory of Genetics and Breeding of Marine Organisms, College of Ocean and Earth Sciences, Xiamen University, Xiamen, China

[7]State Key Laboratory of Large Yellow Croaker Breeding, Ningde Fufa Fisheries, Ningde, China

Correspondence
José M. Yáñez, Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Santiago, Chile.
Email: jmayanez@uchile.cl

Funding information
Alliance of internacional Scince Organization - ANSO, Grant/Award Number: ANSO-CR-PP-2021-03; Fondo Nacional de Desarrollo Científico y Tecnológico, Grant/Award Number: 1211761; National Natural Science Foundation of China, Grant/Award Number: U21A20264

Abstract

Recent advancements in genomic technologies have led to the discovery and application of DNA-markers [e.g. single nucleotide polymorphisms (SNPs)] for the genetic improvement of several aquaculture species. The identification of specific genomic regions associated with economically important traits, using, for example, genome-wide association studies (GWAS), has allowed the discovery and incorporation of markers linked to quantitative trait loci (QTL) into aquaculture breeding programs through marker-assisted selection (MAS). However, most of the traits of economic relevance are expected to be controlled by many QTLs, each one explaining only a small proportion of the genetic variation. For traits under polygenic control, prediction of the genetic merit of animals based on the sum of effects at positions across the entire genome (i.e. genomic estimated breeding values, GEBV, which are used for what has become known as genomic selection), has been demonstrated to speed the rate of genetic gain for several traits in aquaculture breeding. The aim of this review was to provide an overview of the development and application of genomic technologies in uncovering the genetic basis of complex traits and accelerating the genetic progress in aquaculture species, as well as providing future perspectives about the deployment of novel molecular technologies for selective breeding in coming years.

KEYWORDS
aquaculture, genomic selection, GWAS, quantitative trait loci, selection signature, single-nucleotide polymorphism

1 | INTRODUCTION

Rapidly expanding aquaculture production represents the potential for a fast increase in fish supply.[1] The efficiency and sustainability of fish farming will be crucial to meet the need for protein for human consumption in the near future. Selective breeding of fish and shellfish species represents an efficient way of sustainably increasing production by means of improving traits such as rapid growth, product quality, disease resistance and tolerance to diverse environmental stressors.[2,3] Encouragingly, annual genetic gains reported for aquatic species are in general substantially higher than those for terrestrial farm animals, and there is considerable scope for achieving significant positive economic impact via selective breeding schemes.[2,4,5]

In aquaculture, as in terrestrial livestock species, there is increasing interest in applying genomic information to accelerate genetic progress for traits that are difficult to measure on the selection candidates.[6] Quantitative trait loci (QTLs; refer to Table 1 for full list of abbreviations) with a large effect on economically important traits were initially mapped using sparse molecular markers (e.g. microsatellites) and linkage analyses in several fish and shellfish species; some examples were reviewed in Yue et al.[7] The most successful case of implementation of marker-assisted selection (MAS) in aquaculture species is related to a major QTL controlling nearly 80% of the genetic variance for resistance to infectious pancreatic necrosis virus (IPNV) in Atlantic salmon (*Salmo salar*),[8-10] which has been successfully implemented in commercial stocks.[8] Moen et al.[8,9] showed how, in the case of IPNV, haplotyping could be used to determine quantitative trait nucleotide (QTN) genotypes accurately at the population level (as opposed to within individual family groups), even with sparse microsatellite markers. However, this case is not representative, and in most cases, the practical application of QTL for selective breeding has been limited by the low proportion of the heritability explained by marked QTLs and the inconsistency of the linkage phase among families from the same population (i.e. which marker allele is associated with the performance-increasing QTL allele).

TABLE 1 List of abbreviations

Abbreviation	Meaning
DNA	Deoxyribonucleic acid
SNP	Single-nucleotide polymorphism
GWAS	Genome-wide association study
QTL	Quantitative trait loci
MAS	Marker-assisted selection
GEBV	Genomic estimated breeding values
IPNV	Infectious pancreatic necrosis virus
QTN	Quantitative trait nucleotide
LD	Linkage disequilibrium
EBV	Estimated breeding value
GBS	Genotyping-by-sequencing
RAD-seq	Restriction site-associated DNA sequencing
SSA	Salmon alphavirus
PMCV	Piscine myocarditis virus
qPCR	Quantitative polymerase chain reaction
BCWD	Bacterial cold-water disease
IHNV	Infectious haematopoietic necrosis virus
PD	Pancreas disease
CMS	Cardiomyopathy syndrome
VNN	Viral nervous necrosis
LG	Linkage group
ddRAD-seq	Double-digest restriction site-associated DNA sequencing
WSSV	White spot syndrome virus
IBD	Identity-by-descent
GO	Gene ontology
H_P	Pooled heterozygosity
XP-EHH	Cross Population Extended haplotype homozygosity
SRS	Salmon rickettsial syndrome

With the emergence of genome-wide single-nucleotide polymorphism (SNP) genotyping panels, the genetic architecture of quantitative traits and loci controlling them have been studied with a deeper level of resolution by applying genome-wide association studies (GWAS). This approach exploits the population linkage disequilibrium (LD) to identify association between particular genetic variants and phenotypic variation for relevant traits.[6] Furthermore, dense SNP panels can be used to implement genomic selection schemes, which can yield accurate individual estimated breeding values (EBVs), by assuming that the genetic variance for a particular trait is explained by the additive effects of thousands of loci with very small effects uniformly distributed throughout the genome.[11] In a common aquaculture genomic selection scheme (Figure 1), individuals genetically connected to the selection candidates (e.g. full- and half-siblings) are phenotyped and genotyped (i.e. as the 'training' population), while the selection candidates are only genotyped (i.e. as the 'testing' population). Data from training population is used to estimate the individual SNP effects, and genomic EBVs are predicted for the selection candidates using all genotyped SNPs.[12-14] Thus, genomic selection has the potential to increase accuracy of selection and accelerate the genetic progress for traits which cannot be directly measured on the selection candidates (e.g. carcass quality and disease resistance traits). As an example of the impact of genomic selection in animal breeding, it is estimated that in the dairy cattle industry, more than 3 million animals have been genotyped since 2008 and genomic selection is considered to have doubled the rate of genetic progress.[15] It is expected that genomic selection will be increasingly adopted by the aquaculture industry.

In this review, we cover the status of genomic resources available in fish and shellfish species with relevance for aquaculture (e.g. salmon, trout, carp, tilapia, catfish, shrimp, scallops and oysters) and how these tools are being used to (i) leverage the discovery of loci controlling economically important traits, (ii) accelerate genetic progress using genomic selection and (iii) detect the genetic variants involved in domestication and selection in aquaculture species.

2 | GENOMIC RESOURCES FOR AQUACULTURE SPECIES

Genomic resources for the most important aquaculture species globally have become more widely available in recent years. These resources include high-quality reference genome assemblies and genome-wide SNP panels for several species. The current status of reference genome sequences available for important aquaculture species is shown in Table 2. The genome biology of aquaculture species can prove highly complex. For instance, the duplication of the complete set of chromosomes has occurred in an extensive and independent manner in many groups of fish,[16] such as pseudotetraploidy in

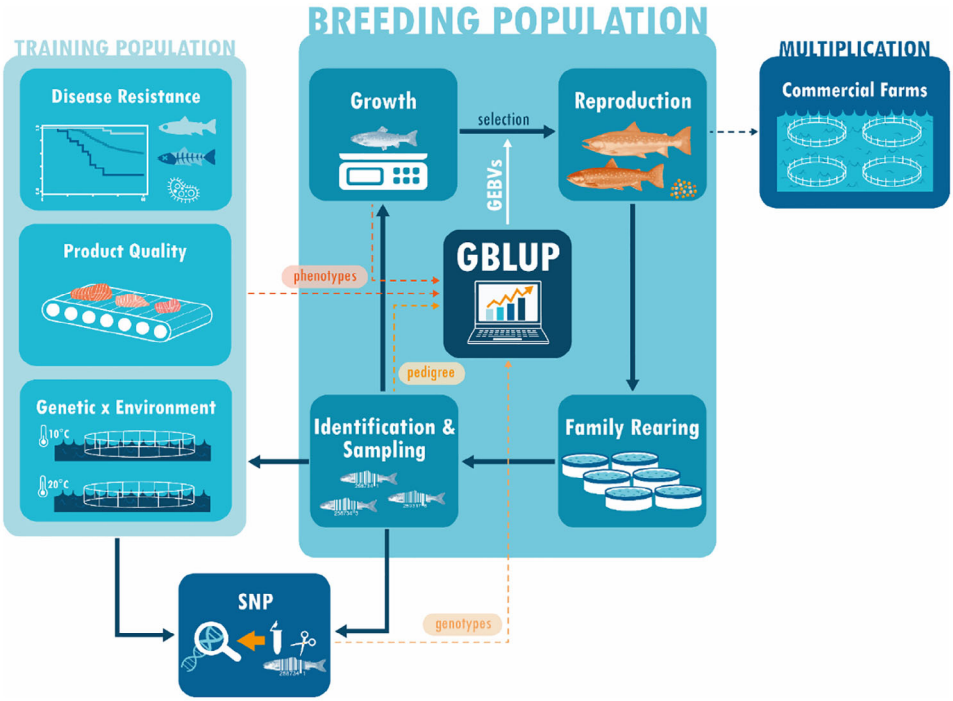

FIGURE 1 The breeding population is comprised broodstock that are spawned in the reproduction season to generate several (hundreds) of full- and half-sib families reared in separate tanks. When the individuals are large enough to be physically tagged, they are individually identified to maintain pedigree traceability during the whole cycle. Some of the tagged animals are kept as selection candidates in the breeding nucleus until they are measured at commercial weight. Other groups of tagged animals (full- and half-sibs of the selection candidates), representing all families from the breeding nucleus, are sent to genotype-by-environment, product-quality and disease-resistance testing to evaluate training genotypes with phenotypes (i.e. to become training populations). All physically tagged animals are sampled and genotyped using SNP panels. The information from pedigree, phenotypes and genotypes from the training populations and selection candidates are jointly analysed through genomic evaluation methods (e.g. GBLUP) to predict GEBVs, which in turn are used to make selection decisions for the next reproduction season and to accelerate the transfer of superior genetics to commercial farms via multiplication

TABLE 2 Current status of the reference genomes for important aquaculture species

Species	Assembly	Release date	Coverage	Number of contigs	Contig N50 (bp)	Total (GB)
Atlantic salmon (*Salmo salar*)	GCA_ 905237065.2	21 April 2021	70×	4222	28,058,890	2.8
Rainbow trout (*Oncorhynchus mykiss*)	GCA_002163495.1	3 May 2017	244.0×	559,855	13,827	2.2
Coho salmon (*O. kisutch*)	GCA_002021735.1	3 May 2017	213.0×	97,074	58,118	2.4
German mirror carp (*Cyprinus carpio*)	GCA_004011555.1	10 January 2019	221.74×	53,446	94,545	1.4
Yellow River carp (*Cyprinus carpio haematoperus*)	GCA_004011575.1	10 January 2019	211.94×	184,435	59,678	1.42
Hebao red carp (*Cyprinus carpio*)	GCA_004011595.1	10 January 2019	186.66×	316,365	36,306	1.46
Grass carp (*Ctenopharyngodon idella*)	http://www.ncgr.ac.cn/grasscarp	29 July 2015	132.1×	5701 scaffolds	40,781	0.87
Goldfish (*Carassius auratus*)	GCA_003368295.1	9 August 2018	71.0×	8463	821,153	1.8
Nile tilapia (*Oreochromis niloticus*)	GCA_001858045.3	31 October 2016	44.0×	3010	2,923,640	1
Channel catfish (*Ictalurus punctatus*)	GCA_001660625.1	9 June 2016	54.0×	34,544	77,201	0.8
Pangasius (*Pangasianodon hypophthalmus*)	GCA_003671635.1	22 October 2018	130.0×	23,339	62,522	0.7
Eastern oyster (*Crassostrea virginica*)	GCA_002022765.4	1 September 2017	87×	669	1,971,208	0.7
Large yellow croaker (*Larimichthhys crocea*)	GCA_003845795.1	27 November 2018	430.50×	1576	2,833,482	0.72
Pacific white shrimp (*Litopenaeus vannamei*)	GCA_003730335.1	15 November 2018	1.0×	19,584	6600	1

TABLE 3 Dense SNP arrays available for fish and shellfish species reported in the literature

Species	Technology	Number of SNPS	References
S. salar	Affymetrix Axiom (ThermoFisher)	132K	247
	Affymetrix Axiom (ThermoFisher)	220K	150
	Affymetrix Axiom (ThermoFisher)	200K and 50K	25,246
	Affymetrix Axiom (ThermoFisher)	930K	245
O. mykiss	Affymetrix Axiom (ThermoFisher)	57K	250
C. carpio	Affymetrix Axiom (ThermoFisher)	250K	251
I. punctatus	Affymetrix Axiom (ThermoFisher)	250K	252
	Affymetrix Axiom (ThermoFisher)	690K	253
O. niloticus	Affymetrix Axiom (ThermoFisher)	58K	254
	Illumina iSelect	50K	255
Tambaqui (*Colossoma macropomum*) and Pacu (*Piaractus mesopotamicus*)	Affymetrix (ThermoFisher)	30K	256
L. crocea	Affymetrix Axiom	600K	257
Japanese flounder (*Paralichthys olivaceus*)	Affymetrix Axiom	50K	258
Pacific oyster (*Crassostrea gigas*)	Affymetrix Axiom (ThermoFisher)	190K	259
	Affymetrix Axiom (ThermoFisher)	23K	174
	Affymetrix Axiom (ThermoFisher)	41K	260
L. vannamei	Illumina iSelect	9K	261

salmonids[17] and allotetraploidy in carp.[18] Repetitive DNA elements can be highly abundant and variable in aquatic species.[19] Moreover, some species from the phylum Mollusca have a highly polymorphic genome.[20] Thus, the application of complementary sequencing technologies (e.g. long-read sequencing),[21] and alternative methodologies (e.g. optical mapping, which consists of a microscopy visualization on the characteristics of DNA)[22]—accounting for polyploidy, highly repetitive and polymorphic genome content—is key when generating and applying genomic tools in aquaculture species.

Reference genomes are very important resources for the design and generation of genome-wide SNP panels, which will have a significant impact on disentangling the genetic architecture of complex traits and accelerating the response to selection in aquaculture species.[23] Genotyping-by-sequencing (GBS) approaches, for example, restriction site associated DNA sequencing (RAD-seq) and double-digestion RAD-seq, have been applied in aquatic species with scarce or even null development of a reference sequence (see Robledo et al.[24] for a review). Although these efforts have been very important to take the first steps towards introducing genomic information into selective breeding of aquaculture species, SNP arrays are the most widespread tools for routine genomic evaluations in major farmed species. This is most likely due to comparative advantages of the SNP arrays over GBS techniques, including the simplicity of sample processing and downstream bioinformatics analysis, lower proportion of missing data, faster turn-around time for genotyping and reproducibility of results.

Next-generation DNA sequencing technologies have facilitated assessing the genetic variation of several species of interest at a genomic scale. The discovery of abundant SNP markers has been performed together with the validation of informativeness. Thus, various SNP panels have been developed for important aquaculture species (Table 3) and their applications are increasingly being reported. As will mentioned below, these tools are being currently used to identify QTLs controlling relevant traits and to estimate the genetic merit of animals with increased accuracy in several leading breeding programs.

3 | GENETIC ARCHITECTURE OF DESIRABLE TRAITS

The aim of GWAS is to identify the genomic regions involved in determining the phenotypic variation for a particular trait. In aquaculture, the traits of interest vary across species and even within species, being defined based on the breeding objective. However, there are several traits of interest for selection across different species, including growth-related traits, host disease resistance, time to sexual maturation, carcass quality traits, and tolerance to different environmental factors.

GWAS results can be visualized by (i) plotting the additive genetic variance explained by each SNP and (ii) plotting the $-\log_{10}$ (*p*-value) for the association of each SNP with a specific trait. In the first case, the proportion of additive genetic variance captured by the SNPs is presumed to be a function of the LD between the SNPs and the causal variants affecting the traits. The attention is focused on the size of the SNP effect, irrespective of its statistical significance. In the second case, if the *p*-value surpasses the significance threshold (commonly using an $= 0.05$, corrected using the false discovery rate or Bonferroni method) the SNP is considered associated with the trait at

a genome-wide level. In some cases, suggestive or chromosome-wide associations are also reported (e.g. for 50K SNPs, the suggestive *p*-value threshold is 10^{-6}).

3.1 | GWAS in salmonid species

The vast majority of GWAS studies in aquaculture has been performed on salmonid species, in particular, Atlantic salmon and rainbow trout (*Oncorhynchus mykiss*). For Atlantic salmon, most of the studies are related to host disease resistance. All the studies were conducted using dense SNP arrays and include resistance to diseases caused by bacteria,[25,26] viruses[27–30] and parasites.[31–38]

Resistance against bacterial infections in Atlantic salmon was assessed by experimental challenges conducted by intraperitoneal injection and host resistance was defined as time-to-death and binary survival/death.[25,26] Two genome-wide significant QTLs were found for resistance to *Piscirickettsia salmonis* on chromosomes *Ssa 01* and *Ssa 17*,[25] and two chromosome-wide significant SNPs were identified for host resistance to *Renibacterium salmoninarum* on chromosomes *Ssa 04* and *08*.[26] Nevertheless, the proportion of the genetic variance explained by each QTL was relatively low for both traits.[25,26]

Genomic regions linked to resistance to viral disease have been identified in Atlantic salmon. For instance, two studies have shown the presence of one genome-wide significant QTL on *Ssa 03* for resistance against salmon alphavirus (SAV), the causative agent of pancreas disease, when the trait is defined as either survival or viral load.[27,28] In addition, a QTL for survival was detected in *Ssa 07*, which together with the QTL in *Ssa 03* explained about 60% of the additive genetic variance.[28] These results suggest the presence of two major genes controlling SAV resistance. Several members of immunoglobulin-heavy-chain locus B (*igh-B*) and grass carp reovirus-induced gene 1 (*gig1*), both located on *Ssa 03*, and immunoglobulin-light-chain (*iglc*), located on *Ssa 07*, were proposed as candidate genes involved in SAV resistance based on GWAS results and global gene expression analyses.[27,28] In addition, a major QTL explaining 57% of the genetic variance for viral load after an experimental infection against piscine myocarditis virus (PMCV), the causative agent of cardiomyopathy syndrome, has been found.[29,30] The QTL is located on *Ssa 27*, and a cluster of major histocompatibility complex class I (*mhc-I*) genes is suggested to include the causative gene.[30]

GWAS for resistance against two sea lice species, *Caligus rogercresseyi*[31,35] and *Lepeophtheirus salmonis*,[32,36,38] have been reported in Atlantic salmon. All the studies defined resistance as the number of lice attached to fins and body or lice density after an experimental challenge. Regarding *C. rogercresseyi* resistance, three QTL with small to moderate effects were identified and characterized on *Ssa 03*, *Ssa 08* and *Ssa 21*[31,35]; however, only the QTL on *Ssa 03* reached genome-wide statistical significance.[35] In the case of resistance to *L. salmonis*, evidence from three studies indicated that this trait is polygenic in nature, with only suggestive QTLs found on *Ssa 01*, *Ssa 04*, *Ssa 14* and *Ssa 20*.[32,36,38] Identification of genomic regions associated with resistance to amoebic gill disease, caused by *Paramoeba perurans* and *Neoparamoeba perurans*, was assessed in three independent studies.[33,34,37] The authors defined resistance as amoebic load (measured using qPCR) or gill injury score (ranging from 0 = healthy red gills to 5 = extensive lesions)[39] after either experimental challenges[33,37] or a field test.[34] One genome-wide significant QTL was identified for gill injury score on *Ssa* 02,[34] and several suggestive QTLs were found for both amoebic gill disease resistance traits, suggesting that several genomic regions are involved in trait variation.

Genomic regions associated with growth rate have been evaluated at different time-points in *S. salar*.[40–42] Polygenic architecture was suggested for body weight for 1-year-old fish, with no SNPs surpassing the genome-wide significance threshold. This architecture was expected given the results observed in other species. At the same age, a single SNP located on *Ssa 17* was associated with body length at a chromosome-wide significance level.[40] Using Bayesian and frequentist GWAS approaches, polygenic architecture for body weight at tagging (~13 g) and at 25 months of age (300 g),[41] as well as for days to reach 5 kg,[42] respectively, has been suggested.

Using more than 4000 animals and 45K SNPs, five different loci were identified for sex determination in a Tasmanian Atlantic salmon population.[43] Similar results were found in a wild Spanish Atlantic salmon population, with *Ssa 02* explaining the highest proportion of genetic variance,[44] although these authors also suggested *Ssa 21* as a candidate chromosome for sex determination. In the case of sexual maturation, two studies showed that a major QTL on *Ssa 25*, linked to the *vestigial-like protein 3* (*vgll3*) gene, is controlling age-at-maturity in European Atlantic salmon.[45,46] In addition, a single significant genomic region associated with grilling (precocious male maturation) was identified in *Ssa 21* (*Ssa 25* homeolog) in North American Atlantic salmon.[47] However, several other genomic regions have been associated with age-at-maturity in other farmed Atlantic salmon populations, including QTLs on almost all chromosomes.[42,48,49] For instance, Sinclair-Waters et al.,[49] analysed 11K Atlantic salmon males with 512K imputed SNP genotypes (i.e. genotypes predicted using a reference population genotyped at higher SNP density). The authors suggested a mixed architecture for time-at-maturation, with several loci of small effect and a few loci of large effect, with the latter located on *Ssa 09* and *Ssa 25* and most likely linked to the sine *oculis-related* homeobox *6* (*six6*) and *vestigial-like protein 3* (*vgll3*) genes, respectively.

Regarding carcass quality traits, Sodeland et al.[50] evaluated fillet fat content and firmness at 2 and 5 days post-mortem. These authors suggested that most of the genetic variation affecting fillet fat and firmness are located on *Ssa 09* and *Ssa 10*, and *Ssa 03* and *Ssa 11*, respectively, although no major-effect QTL was found. In addition, an important QTL on *Ssa 26* showed strong association to flesh colour in Atlantic salmon, with beta-carotene oxygenase 1 like (*bco1l*) being the most likely causative gene.[51] More recently, significant QTLs on *Ssa 19* and *Ssa 21* were found to be associated with omega-3 fatty acid composition in fillets of Atlantic salmon.[52]

For rainbow trout, most of the studies ($n = 21$) were focused on identifying genomic regions associated with resistance to infectious

diseases, similar to Atlantic salmon. Among them, four were related to the bacterium *Flavobacterium psychrophilum*, the etiological agent for bacterial cold water disease (BCWD), either using SNP arrays or RAD-seq.[53–56] Several QTLs were found for resistance, defined as survival status or time-to-death after either experimentally or naturally induced infection. However, only one genomic region located on *Omy 08* seems to be consistently associated when using different genomic and statistical approaches, suggesting the importance of this chromosome on resistance to BCWD.[53–55] Recently, two novel genome-wide significant QTLs for resistance to BCWD were identified in a French rainbow trout population. These QTLs were identified on *Omy 07* and *Omy 17* using 30K SNPs.[56] Because the data were collected after a natural *F. psychrophilum* field outbreak, these QTLs may be related to natural host response against infection, which may explain the differences from results of previous studies based on experimentally induced infections by inoculum injection. Additionally, several genomic regions have been associated with resistance to other important bacterial and parasitic disease agents in rainbow trout, including *Piscirickettsia salmonis*,[57] *Yersinia ruckeri*[58] and sea lice.[59] Major QTLs have been mapped to *Omy 21* for *Vibrio anguillarum*,[60] and in *Omy 16* for *Ichthyophthirius multifiliis*[61] and *Aeromonas salmonicida*[62] resistance, respectively. In addition, growth under chronic thermal stress and cortisol response to crowding seem to be under polygenic control in this species.[63,64]

Dissecting the genetic basis of resistance to viral diseases has been addressed in rainbow trout. Two studies identified genomic regions associated with resistance to IPNV and infectious haematopoietic necrosis virus (IHNV) using Bayesian approaches, although the IHNV sutdy also used single and a weighted single-step GBLUP approach. Brief explanations about the most-cited statistical

TABLE 4 Summary of the main BLUP and Bayesian methods employed in GWAS and genomic selection

Approach	Acronym	Definition	Method's application	References
Best linear unbiased prediction (BLUP)	PBLUP	Pedigree best linear unbiased prediction	Breeding values are predicted using only pedigree information and mixed model equations with the inverse of the numerator relationship matrix (A^{-1})	262
	rrBLUP	Ridge-regression best linear unbiased prediction	A 'penalized' BLUP method in which the estimation suffers shrinkage towards zero to avoid over-fitting and multicollinearity problems caused by the large number of markers	263
	GBLUP	Genomic best linear unbiased prediction	Genomic breeding values are estimated using genomic relationships (pedigree is unnecessary) which are based on genome-wide dense marker data to compose the *G* matrix	192
	ssGBLUP	Single-step genomic best linear unbiased prediction	A GBLUP method that uses all pedigree, phenotypic and genotypic available information, simultaneously combining them into the *H* matrix	264
	wssGBLUP	Weighted single-step genomic best linear unbiased prediction	A ssGBLUP method which allows weighting each marker according to the variance they explain using an iterative method and fitting better to oligogenic traits	66
Bayesian	BayesA	–	The prior distribution of marker effects is assumed to be identical with independent univariate-*t* distribution and a null mean	11
	BayesB	–	Each marker effect has a distribution with point mass at zero with probability π and a univariate-*t* distribution with probability $1-\pi$ and a null mean	11
	BayesC	–	Each marker effect has a distribution with point mass at zero with probability π and a univariate-normal distribution with probability $1-\pi$ and a null mean	265
	BLasso	Bayesian least absolute shrinkage and selection operator	A 'penalized' regression method based on the sum of absolute regression coefficients assuming an exponential prior distribution for variances of marker effects	265

approaches covered in this review for performing GWAS and genomic selection are available in Table 4. The main difference between these statistical methods is the prior distribution of SNP effects in which single-step GBLUP (ssGBLUP) fits better to polygenic traits and Bayesian methods to oligogenic (i.e. few QTLs of moderate effects) or major-effects traits.[65] However, the weighted single-step GBLUP (wssGBLUP) allows weighting each marker based on the variance they explain also fitting properly for oligogenic traits.[66] . Three genomic regions were suggested a associated with IPNV resistance, although the proportion of the genetic variance explained by each QTL was smaller than 0.2%, suggesting polygenic architecture of this trait.[67] Based on evidence of 10 genomic regions with moderate effect and several loci with small effect associated with resistance to IHNV, an oligogenic architecture was suggested for this trait.[68]

Regarding growth-related traits, several QTL for body weight-at-tagging and at age 18 months, explaining up to 3% of the genetic variance, were found using 20-SNP windows (i.e. the percentage of genetic variance explained by 20 adjacent SNPs was accumulated in non-overlapping windows) and the wssGBLUP approach.[69] A similar approach was used to identify a marker located on *Omy 05* associated with body weight at 10 months[70] and other regions on seven different chromosomes explaining from 2% to 6% of the genetic variance for body weight gain.[71] Regarding carcass quality-related traits, two QTLs for muscle yield, located on *Omy 14* and *Omy 16*, were found using a window of 50 adjacent SNPs, together explaining up to 23% of the genetic variance for the trait.[72] Further, the authors also found 10 QTLs of smaller effect (explaining between 1% and 3% of the genetic variance) for muscle yield, including a region located on *Omy 09*, confirming QTLs for fillet yield, fillet weight and carcass weight found in a previous study.[70] QTLs of small effect have also been associated with protein content, firmness, flesh colour, intramuscular fat and moisture content in fillet on different chromosomes of rainbow trout.[72-74] Low-to moderate-effect QTLs have been identified for female reproductive traits (spawning date, body weight, egg size and number) and masculinization in XX females,[75,76] respectively, in rainbow trout.

In the case of coho salmon (*Oncorhynchus kisutch*), there have been fewer genomic studies to date compared to Atlantic salmon and rainbow trout. The recent release of a publicly available SNP chip[77] might increase the number GWAS for this species in the near future. Based on 9K SNP markers from a double-digest RAD approach, moderate-effect QTLs associated with resistance to the facultative intracellular bacterium *P. salmonis* were found.[78]

A recent study in another cultured species, Chinook salmon (*Oncorhynchus tshawytscha*), identified a major QTL for flesh colour on *Chr 30*, explaining 66% of the variation in colour, located close to the beta-carotene oxygenase 2-like (*bco2l*) gene.[79]

In the case of the most widely cultured salmonids, a handful of commercial breeding companies produce eggs of genetically improved strains used by producers worldwide. These companies are active in genomic research, with or without academic partners, and the results from their studies are not always published in the public domain. These companies perform marker-assisted selection for a number of traits for which major QTLs have been found, for example, resistance to viral diseases (IPN; pancreas disease, PD and cardiomyopathy syndrome, CMS), resistance to bacterial diseases (BCWD, flavobacteriosis and piscirickettsiosis), fillet colour and age at maturity.

3.2 | GWAS in non-salmonid species

Studies for the detection of QTLs have been conducted in a variety of non-salmonid aquaculture species. Eight studies were found performing GWAS for ictalurid catfish (notably, channel catfish *Ictalurus punctatus*). Using 14K SNPs, only one marker, located on linkage group (*LG*) *07*, was significantly associated with *Flavobacterium columnare* resistance at a genome-wide level, whereas three different QTLs on *LG 07*, *12* and *14* were suggestively associated with resistance.[80] In addition, a QTL for enteric septicemia (*Edwardsiella ictaluri*) resistance was identified in *LG 1* in two independent studies using 13 and 6K SNP panels, respectively; however, several other LGs were also found to be suggestively associated, indicating polygenic architecture for the trait.[81,82] Several genome- and chromosome-wide significant QTLs with small effects have been identified for body conformation traits in ictalurid catfish, including body weight, deheaded body length, body length, body depth, body breadth and head size.[83-85] Three QTLs, explaining more than 10% of the genetic variance each, were found to be associated with heat stress in hybrid catfish (channel catfish × blue catfish).[86] In the same hybrid population, two significant and two suggestive small-effect QTLs for low oxygen tolerance were found using a 250K SNP chip.[87]

Studies aiming to identify genomic regions associated with resistance to viral nervous necrosis (VNN) have been conducted in Asian and European sea bass, *Lates calcarifer* and *Dicentrarchus labrax*, respectively.[88,89] Two QTLs, on *LG 19* and *20*, were significantly associated with resistance to VNN in *L. calcarifer*, when the trait was defined as survival/mortality and time-to-death after an experimental challenge.[89] Several QTLs with small effect were associated with resistance to VNN in different populations of *D. labrax* using both RAD sequencing[90] and SNP chip genotyping technologies,[88] with the most important QTL found in *LG 12;* explaining about 10% of the genetic variance for the trait.[88]

QTL mapping and association analysis in common carp (*Cyprinus carpio*) identified 22 QTLs for growth-related traits and seven QTLs for sexual dimorphism using a 250K SNP array. Candidate genes underlying growth-related traits, included important regulators, such as kisspeptin 2 (*kiss2*), insulin-like growth factor 1 (*igf1*), somatolactin beta (*smtlb*), neuropeptide FF receptor 1 (*npffr1*) and carboxypeptidase E (*cpe*), and sexual dimorphism-related genes, such as 3-keto-steroid reductase (*3ksr*) and mab-3-related transcription factor 2b (*dmrt2b*).[91] GWAS analysis identified 12 significant SNPs for head-size traits of common carp, based on 433 individuals from multiple families.[92] Several loci were associated, at the suggestive level, with fat content in dorsal and abdominal muscle and abdominal fat weight in common carp, with only one genome-wide significant SNP associated with the latter, suggesting polygenic architecture for these traits.[93]

Later, several markers associated with polyunsaturated fatty acid content, at both genome- and chromosome-wide significance thresholds, were identified.[94] An association study of abnormal scale pattern in 82 indigenous Yellow River carp individuals identified the causative mutation in fibroblast growth factor receptor 1 a1 (*fgfr1a1*).[95] A QTL located on *LG 44* and explaining 7% of the genetic variance was found for resistance to koi herpesvirus in common carp.[96] Similarly, three significant QTLs and one suggestive QTL were identified for host resistance to koi herpesvirus in mirror carp.[97] For rohu carp (*Labeo rohita*), 10 LGs exhibited 21 SNPs significantly associated to *Aeromonas hydrophila* resistance, and several of them were homologous to genes related to immune function, such as heat shock proteins, mucins and lectins.[98]

In the case of Nile tilapia (*Oreochromis niloticus*), several loci were found to be associated with acute salinity tolerance on *Oni 05* and *Oni 18* using a ddRAD-seq genotyping approach.[99] A large-effect QTL associated with sex determination, and mapped within the anti-Mullerian hormone gene (*amh*) on *Oni 23*, was identified using whole-genome resequencing data from fish belonging to three breeding populations.[100] In contrast, polygenic architecture was suggested for fillet yield and harvest weight, with markers explaining 1.5%–2.0% of the genetic variance for each trait, respectively.[101]

In large yellow croaker (*Larimichthys crocea*), the sex determination region was localized within a 2.4 Mb region on *Chr 22*, based on GWAS using ddRAD sequencing in 905 individuals from a breeding population.[102] Significant SNP loci related to growth and body-shape traits were identified on multiple chromosomes using the GWAS approach based on ddRAD-seq genotyping data.[103,104] A significance peak associated with the high plant protein utilization trait was identified on *Chr 18*.[105] Five significant SNPs related to heat tolerance were identified on *Chr 04* based on the use of a 600K SNP array.[106] Similarly, seven significant QTLs, located on four chromosomes, were found for resistance against the ciliate parasite *Cryptocaryon irritans*[107]; subsequently, a more sophisticated GWAS study revealed 15 QTLs associated with that resistance trait.[108]

In the case of resistance of gilthead sea bream (*Sparus aurata*) to pasteurellosis (*Photobacterium damselae* subsp. *piscicida*), no genome-wide significant markers were found when using a 12K SNP panel identified through 2b-RAD.[109]

Using whole-genome resequencing data for 505 individuals, a total of 33 SNPs located on six different chromosomes were found to be significantly associated with host resistance of half-smooth tongue sole (*Cynoglossus semilaevis*) to *Vibrio harveyi* infection.[110]

3.3 | GWAS in shellfish

Compared to salmonids and non-salmonid fish, genome-wide association studies in shellfish are considerably fewer. In Pacific oyster (*Crassostrea gigas*), a number of studies have been carried out to map genomic regions associated with nutritional traits, including glycogen, amino acid and fatty acid content, and Zn, Cu and Se accumulation, among others.[111-113] A strong QTL associated with glycogen content[111] was functionally characterized, and the protein phosphatase 1 regulatory subunit 3B (*ppp1r3b*) gene was suggested to be responsible for promoting glycogen synthesis.[114] Several candidate genomic regions with small effect have been associated with heat tolerance in Pacific abalone (*Haliotis discus hannai*),[115] and ostreid herpesvirus[116] and *Vibrio alginolyticus*[117] resistance in Pacific oysters. Polygenic architecture was found for resistance to herpes virus in *C. gigas*.[116] The authors found several markers showing genome-wide association with binary survival and viral load. However, *LG 06* had the most promising QTL, explaining the larger proportion of genetic variance. Similarly, no major QTLs were found for shell growth and shape.[118] Recently, an association study was performed for survival under low salinity in eastern oysters (*Crassostrea virginica*) using RAD-seq technologies. Significant QTLs were found on *LG 01* and *07*.[119]

In scallops, all studies used RAD-seqtechnology to genotype animals. This approach is common for species that do not have extensive genomic information, in particular, a reference genome assembly.[24] Most of these studies were performed for the Yesso scallop (*Patinopecten yessoensis*) considering different traits, such as growth,[120] pH tolerance,[121] and colour of shell[122] and muscle.[123,124] In the case of shell colour, genes were identified with carotenoid-related functions flanking two SNPs on *LG 11*.[122] Li et al.[124] found a genomic region strongly associated with carotenoid coloration on *LG 08*. Through transcriptomic analysis in that region, the authors confirmed down-regulation of the beta-carotene oxygenase-like *1* (*pybcol1) gene* encoding carotenoid oxygenase, the enzyme known to catalyse carotenoid cleavage in a variety of organisms, which was *differentially expressed* between white and orange muscles. These results suggest that marker-assisted selection may be a feasible alternative to obtain scallops with reddish-orange shell and muscle.

A limited number of GWASs have been published for shrimp. Several genomic regions and candidate genes associated with growth have been found in Pacific white shrimp (*Litopenaeus vannamei*)[125-127] using RAD-seq or similar technologies; however, none found QTLs with large effects. For sex determination, a major QTL was reported using a 6.4K SNP panel for discovering and genotyping SNPs.[128] For the black tiger shrimp (*Penaeus monodon*), association analysis was performed for white spot syndrome virus (WSSV) resistance and sex determination.[129] However, for WSSV resistance, only weak associations were detected; for sex, a strong association was identified on *LG 30*. This major association was confirmed later, in which a locus explaining 77.4% of the phenotype variance for sex determination, but in *LG 23* using linkage mapping.[130] Recently, a new 50K SNP panel was described for *L. vannamei*,[131] which may result in more GWAS studies for this species in the near future.

3.4 | Validation of GWAS

A recurrent issue in the application and validation of GWAS is inconsistent results across different populations. The main reasons for such inconsistency are the complex genetic architectures that drive most economically relevant traits, and the genetic background of the

TABLE 5 GWAS available for fish and shellfish species reported in the literature

Species	Trait	Number of SNPS (K)	Statistical Method	Associated chromosomes	Candidate genes	References
S. salar	Fillet texture	6	No mentioned	3, 9, 10 and 11	Not assessed	50
S. salar	Growth traits	112	GRAMMAR[a]	17	Myh9, gucy2f and gapdhs	40
S. salar	Growth rate and sexual maturation	4	GRAMMAR	1, 2, 10, 12, 13, 21, 25 and 28	E2f4, npm1 and magi-1	42
S. salar	Body weight	116	BayesC	21 and 27	Vtn, wnt1 and mrap2	41
S. salar	Flesh pigment content	6	LMM-GWAS[b]	26	Bco1 and bco1l	51
S. salar	Fatty acid composition	57	EMMAX[c]	21	Elovl2	52
S. salar	*P. salmonis* resistance	48	GRAMMAR	1 and 15	Fut10, il31ra and il6st	25
S. salar	*C. rogercresseyi* resistance	39	FASTA[d]	21	Col1a1	31
S. salar	*L. salmonis* resistance	132 and 35	GRAMMAR	1, 3, 9 and 23	Not assessed	267
S. salar	*R. salmoninarum* resistance	44	FASTA	4 and 8	Not assessed	26
S. salar	*N. perurans* resistance	7	FASTA	17 and 18	Not assessed	33
S. salar	*L. salmonis* resistance	7	GRAMMAR	1 and 23	Not assessed	36
S. salar	*L. salmonis* resistance	37	QFAM[e]	4, 14 and 20	Cptp, chrng and neu4	32
S. salar	*Piscine myocarditis* resistance	55	EMMAX	27 and 12	Eb-like gene, i-e beta chain and h2-ab1	29
S. salar	*Amoebic Gill Disease* Resistance	55	MLMA-LOCO[f]	4, 9 and 13	*Fat4 and* il-*18 bp*	37
S. salar	*C. rogercresseyi* resistance	50[g]	sSNP[h]	3, 8 and 21	Tob1 and stk17b	35
S. salar	Pancreas disease resistance	54	MLMA-LOCO	3	Gig1-like, mbl2-like, tlr13, glyr1 and abat	27
S. salar	Pancreas disease resistance	55	EMMAX	3 and 7	Igh-b and immunoglobulin-light-chain	28
S. salar	Sea-age variation	4	No mentioned	9, 12, 16, 27	Otub2, itpk1 and llp60	268
S. salar	Sexual maturation	4600	Chi-square[i]	25	Vgll3	46
S. salar	Age at maturity	220	ACPOM[j]	25	Vgll3, akap11 and six6	45
S. salar	Sex determination	46	Chi-square	2, 3, 5, 6, 12 and 25	Sdy	43
S. salar	Sex determination	36	EMMAX	2 and 21	Sdy, sox and dnd	44
S. salar	Age at maturity	0.08 and 50	GLMM[k]	21	Ropn1	47
S. salar	Sexual maturation	50	MLMA-LOCO	10 and 11	Picalm, magi2 and maguk	48
S. salar	Age of maturity	512	BOLT-LMM[l]	9, 25, 28 and 29	Vgll3, six6 and ndufs4	49
O. mykiss	Fillet yield, carcass and body weight	38	wssGBLUP[m]	5, 8, 9, 10, 13, 14, 17, 22, 23 and 27	Sox2, capn2 and ksr1	70
O. mykiss	Muscle yield	35	wssGBLUP	14 and 16	Slc26a9, cd34a and mcts1	72
O. mykiss	Filet firmness and protein content	35	wssGBLUP	1, 3, 4, 5, 7, 8, 10, 11, 13, 21 and 28	Nucb1, kcna1 and actc1	73
O. mykiss	Body weight	35	wssGBLUP	15 and 24	Fam60a, wnt16 and bpm2	69

(Continues)

TABLE 5 (Continued)

Species	Trait	Number of SNPS (K)	Statistical Method	Associated chromosomes	Candidate genes	References
O. mykiss	Body weight	50	wssGBLUP	2, 4, 8, 9, 13, 14 and 18	Cav-1, tes, eif4g2, slc6a15, Ap-1, prrc2c, myoc, prdx6, plpp6, acss2 and pecr	71
O. mykiss	Growth under chronic thermal stress	1400[g]	wssGBLUP	3 and 7	Stat5b, stat3 and cish	64
O. mykiss	Body weight, carcass traits, total fat, percentage of fat in flesh, flesh colour	57	BayesC	8, 13, 17 and 22	Htr1, gnpat, ephx1, bcmo1, and cyp2x	74
O. mykiss	*F. psychrophilum* resistance	5	GRAMMAR	8, 11, 13, 17 and 21	Socs6, jak1 and tlr-13	54
O. mykiss	*F. psychrophilum* resistance	5	GRAMMAR	2, 8, 15, 19, 22, 26 and 28	Not assessed	53
O. mykiss	*F. psychrophilum* resistance	24, 41 and 48	wssGBLUP and BayesB	3, 5, 8, 10, 13, 15 and 25	Il1r-like-1 and tnfrsf1a-like-a	55
O. mykiss	*F. psychrophilum* resistance	30	MLMA-LOCO, ssGBLUP[n] and wssGBLUP	3, 7, 10 and 17	Not assessed	56
O. mykiss	Infectious pancreatic necrosis virus resistance	38	BayesC	5, 13 and 23	Senp5, ccr7 and itga11	67
O. mykiss	Infectious haematopoietic necrosis virus resistance	42	ssGBLUP and BayesB	1, 2, 4, 5, 6, 8, 16, 17, 21, 25, 26 and 28	Drd4, cadps2 and plekha7	68
O. mykiss	*P. salmonis* resistance	26	ssGBLUP	3, 5, 14, 24, 27 and 29	Usp2, nlrc3 and tap	57
O. mykiss	*Yersinia ruckeri* resistance	50	MLMA-LOCO	3, 12, 20 and 22	Not assessed	58
O. mykiss	*Vibrio anguillarum* resistance	57	MLMA-LOCO	21	Ppp1c and ig	60
O. mykiss	*Ichthyophthirius multifiliis* resistance	57	MLMA-LOCO	16 and 17	Igm, igt, t-cell receptor β, complement factor c3, lysozyme, cathelicidins 1 and 2 and saa	61
O. mykiss	*C. rogercresseyi* resistance	57	wssGBLUP	15 and 26	Tbx21, fgf11, fgf13, ifng2, frrs1 and tr16l	59
O. mykiss	*Aeromonas salmonicida* resistance	57	MLMA-LOCO	16	Pra1 family protein 3	62
O. mykiss	Female reproduction traits	57	GBLUP[o] and BayesC	1, 2, 6, 8, 11, 12, 15 and 27	Ipo11, arhgef4, nr2e1, mprd, phc1, prkg2 and wapla	76
O. mykiss	Sex determination	57 and 275[g]	MLMA-LOCO and BayesC	1, 12 and 20	Fgfa8, cyp17a1 and (LOC11052793)	75
O. kisutch	*P. salmonis* resistance	9	wssGBLUP	11 and 29	Pik3ap1, phpt1 and tial1	78
O. tshawytscha	Flesh colour	10	MLM[p]	4, 12 and 30	Bco2-l	79
I. punctatus	Head length, head width and head depth	218	EMMAX and QFAM	5, 7, 9 and 16	Cad22l, cyp24a1 and snx25	83
I. punctatus	Deheaded body length, body length, body depth and body breadth	250	EMMAX	2, 5, 6, 7, 13, 14, 24 and 25	Snx8, col17a1, anln, macf1, bmp8a, col21a1, mmp16b, rab3a and fgf12a	84
I. punctatus	Deheaded body length, body length, body depth and body breadth	218	EMMAX and QFAM	2, 5, 7, 13, 14, 18, 24, 26	Rab8a, ddr2l and arhgap22	85
I. punctatus	*F. columnnare* resistance	14	EMMAX	7, 12 and 14	Pik3r3b, cyld-like and adcyap1r1	80
I. punctatus	Heat stress	5	EMMAX	14 and 16	Traf2, fbxw5 and anapc2	86
I. punctatus	Resistance to enteric septicemia	6	EMMAX and QFAM	1, 12 and 16	Nlrc3, nck1 and agtr1b	82

TABLE 5 (Continued)

Species	Trait	Number of SNPS (K)	Statistical Method	Associated chromosomes	Candidate genes	References
I. punctatus	Low Oxygen tolerance	25	EMMAX	2, 4, 23 and 29	Dmbx1a, cyp1a1 and klhl5	87
I. punctatus	Resistance to enteric septicemia	13	EMMAX and QFAM	1, 3, 21 and 26	Nlrc3, nlrp12 and uba5	81
C. carpio	Muscle fat content and abdominal fat-related traits	85	GLMM	9, 17, 21, 23, 33, 34, 35 and 40	Ankrd10a, tanc2 and zfand5a	93
C. carpio	Growth-related and sex dimorphism	250	Chi-square	1, 7, 8, 10, 11, 16, 20, 27, 37, 43	Kiss2, igf1, smtlb, npffr1, cpe3ksr and dmrt2b	91
C. carpio	Polyunsaturated fatty acid content	108	GLMM	20, 21, 22, 35 and 50	Duox2 and trh	94
C. carpio	Head-size related traits	250	GLMM	6, 24, 27, 35, 46, 50	Srpk2, fsrp5, igf1, igf3, grb10, igf1r, notch2 and sfrp2	92
C. carpio	Resistance to Koi herpesvirus	12	wssGBLUP	34, 42 and 44	Trim25	96
C. carpio	Koi herpesvirus resistance	250	Chi-square, GLMM	17, 33, 39, 41, 43	Tnfa, hif1a, galectin-8, rootletin and palladin	97
C. carpio	abnormal scattered scale patterns	250	GLMM	15, 16	Fgfr1a1	95
L. rohita	Resistance to *A. hydrophila*	6	QFAM, FASTA and GRAMMAR	4, 7, 14, 15, 18–21, 23, 24	Hsp60, dse, cd22, tcr, shsps, tbt, muc5b and dpp7	98
L. crocea	Growth and body-shape-related traits	28	GLMM	2, 7, 16, 19, 21, 22, 23	Fgf18, fgf1, nr3c1, cyp8b1, fabp2, cyp2r1, ppara and ccm21	103
L. crocea	Body shape-related traits	23	GLMM	2, 3, 4, 6, 7, 11, 14, 16, 18, 19	Fabp1, acrv1, bcor, mstn, bambi and neo1	104
L. crocea	Resistance to *Cryptocaryon irritans*	5	GLMM	5, 6, 7, 19	Ifnar1, ifngr2, ikbke and cd112	107
L. crocea	Sex determination and gonadosomatic index	23	MLM	18, 22	Dmrt1, dmrt3, piwil2, fam102a and odf2	102
O. niloticus	Harvest weight and fillet yield	50	wssGBLUP	4, 16, 18 and 23	Utp6, rab31 and npr1	101
O. niloticus	Salt tolerance	2	MLM	5 and 18	Epc1, dennd4c and plaa	99
O. niloticus	Sex determination	2400	MLM	23	Amh	100
Cynoglossus semilaevis	Resistance to vibrio harveyi	1000	Chi-square	5, 7, 12, 14, 17	Fblx19, plekha7, nucb2 and fgfr2	110
S. aurata	*Pasteurellosis* resistance	12	GBLUP	1, 2, 3, 10, 7, 20 and 21	Not assessed	109
D. labrax	Viral nervous necrosis resistance	9	GBLUP and wssGBLUP	3, 20 and 25 (unassigned scaffold)	Not assessed	90
D. labrax	Viral nervous necrosis resistance	57	GBLUP and BayesC	12 and 8	Not assessed	88
L. calcarifer	Viral nervous necrosis resistance	45	MLM	1, 8, 14, 15, 16, 19, 20 and 21	Ube2g2, kdm2a and map3k11	89
C. gigas	Glycogen, protein and amino acid contents	52,143	MLM	6, 8 and 9	Yp17a, pepck and g6pase	111
C. gigas	Fatty acids composition, glycogen, Zn and Se	8	MLM	2, 3, 4, 5, 6, 8 and 9	Osbpl11, mc5r, klf3, adamts, insig2, pecr, acsl1, lipe and agpat6	113

(Continues)

TABLE 5 (Continued)

Species	Trait	Number of SNPS (K)	Statistical Method	Associated chromosomes	Candidate genes	References
C. gigas	Shell growth and shape-related traits	52,143	MLM	unassigned scaffolds	Rpsa, trim3, tnn, ky, fndc2, slc7a9 and ankrd44	118
C. gigas	Ostreid Herpesvirus resistance	16	FASTA and EMMAX	6 and 8	Fbne, ranbpm and b3galnt2	116
C. gigas	Zn and Cu accumulation	190	MLM	1, 3, 4, 6, 7, 8 and 11	Cnr3, cav1, donson and t4 dna ligase	112
C. gigas	*Vibrio alginolyticus* resistance	48	BLINK[q]	3, 5, 6, 9 and 10	Tlr6, scarf1, mdm2, hace1, cyld, laccase-19, spidr, pgg1b, gpcr161 and rbbp6	117
C. virginica	Acute low salinity tolerance	29	Linkage QTL mapping	1 and 7	Ubr5, *slco4a1* and ncoa-2	119
Patinopecten yessoensis	Shell colour	62	Chi-square	11	Ldlr, fris and friy	122
P. yessoensis	shell length, shell width, shell height, body weight, soft tissue weight and adductor muscle weight	9	MLM	17 and 18	E2f3	120
P. yessoensis	Muscle coloration	30	Chi-square	8	Pybco-like 1	124
P. yessoensis	Muscle coloration	9	Chi-square	7	Parp and cyp2j6	123
Haliotis discus hannai	Heat tolerance	1431	EMMAX	1, 7, 8, 11, 12, 14, 15,16 17 and 18	Mdhc, achβ3, ach92, cah2 and cah7	269
L. vannamei	Body weight	23	GRAMMAR	11, 15, 18, 21, 28 and 44	Lvsrc	126
L. vannamei	Body weight	4	GRAMMAR	7, 27, 33, and 38	Prkcd and rap2a	125
L. vannamei	Body weight	94	sSNP	19 and 39	Dcmpd and nptk	127
L. vannamei	Sex determination	5	MLMA-LOCO	42 and 44	Pvfem-1	128
Penaeus monodon	White spot syndrome virus resistance and sex determination	6	sSNP	2, 3, 5, 6, 17, 18, 19, 22, 27, 30 and 43	Pai-rbp1, tap26, abra, flik and ubfd1	129

[a]Genome-wide rapid association using mixed model and regression.
[b]Mixed linear model.
[c]Efficient mixed-model association eXpedited.
[d]Fast score test for association.
[e]Family-based association test for quantitative traits.
[f]Mixed linear model-based association analysis leaving-one-chromosome-out.
[g]Imputed genotypes.
[h]Single-SNP Genome-Wide Association Study.
[i]Chi-square allelic test.
[j]Additive cumulative proportional odds model.
[k]General linear mixed model.
[l]Bayesian mixture of normal prior with linear mixed model association.
[m]Weighted single-step genomic best linear unbiased prediction.
[n]Single-step genomic best linear unbiased prediction.
[o]Genomic best linear unbiased prediction.
[p]Mixed linear model.
[q]Bayesian-information and linkage-disequilibrium iteratively nested keyway.

studied populations. Ideally, the target population should present a homogenous background (i.e. lack of population stratification); otherwise, the discrepant allele frequencies among subpopulations may cause confounding effects.[132] This is advantageous for aquaculture species due to the family structure in which homogeneity may be obtained by mating few animals to generate large progenies.[133] Nevertheless, the number of animals in founder populations may limit the power for detecting causal variants. Thus, the recommendation is to have a target population with an homogenous genetic background and enough phenotypic variation to increase mapping resolution.

Another issue related to GWAS results is related to sample size. False discovery rate is inversely proportional to the sample size, but given costs associated with phenotyping and genotyping, it is difficult to have a very large sample size (i.e. several thousand animals). Alternative strategies, such as genotype imputation, genotyping of individuals with extreme phenotypes and meta-analysis of GWAS results of the same trait may be adopted to amend the low sample size problem, but they may also lead to overestimation of the markers effect.[134] We found a diverse number of methods being adopted for GWAS in aquaculture species (Table 5), some of them based on fitting each marker at a time (e.g. GRAMMAR and EMMAX methods) and others assuming a prior distribution of SNP effects fitting them simultaneously (e.g. GBLUP and BayesB). There is no consensus on which method is the most accurate and statistically powerful However, previous knowledge on the genetic architecture of traits of interest might be relevant, as some methods may fit better to traits controlled by major effect QTL (e.g. BayesC and wssGBLUP) while others may be more suitable for polygenic traits (e.g. GBLUP and ssGBLUP).[65] The adoption of multiple statistical methods and overlapping of results may be a viable option to decrease the number of false positives as well.[135]

4 | GENOMIC SELECTION

The genomic selection approach refers to making selection decisions based on GEBV, first proposed by Meuwissen et al.[11] Genomic selection uses dense marker genotypes and phenotypic data to predict genomic breeding values and has been proven to improve accuracy of selection when compared to pedigree-based EBV. The implementation of genomic selection depends on the LD between markers and QTLs affecting the trait. If the level of LD is enough to capture the effects of these QTLs, a prediction equation to estimate GEBV can be fitted using genotypes and phenotypes of animals from a reference population. Finally, the GEBVs of selection candidates can be estimated using only genotypic data based on the prediction equation calculated in the previous step.[136,137] However, in some species, including aquaculture species, recorded phenotypes in animals related to the selection is continuously needed for updating the prediction equation and avoiding the decrease of GEBV accuracy.[14,138]

The main benefits of applying genomic selection is the aquaculture sector are related to the evaluation of difficult-to-measure traits, higher accuracy in the estimation of additive genetic variance and GEBVs, and reduction in the generation interval.[139] The strategy of genomic selection may be relevant for traits that are difficult or impossible to be measured directly on selection candidates, such as resistance to diseases, carcass traits, sex-limited traits and when a genotype-by-environment interaction is present.[12,14,140,141] Another advantage of genomic selection is the more accurate estimate of relationship among animals than using pedigree records only.[142] It o The pedigree matrix expresses only the expected relationship between individuals, while the genomic matrix captures the actual genomic relatedness shared between relatives. For example, Forneris et al.[143] found from 5% to 15% variation in the relationship among relatives when comparing marker and pedigree kinship matrices. Meuwissen et al.[144] suggested that, in principle, the pedigree information must be omitted, considering that the markers may capture the genetic relationship. Further, in genomic selection, the genomic information may be used to estimate the effect of several loci at same time (e.g. Bayesian methods) or by estimating the average genomic relationship (e.g. GBLUP methods). However, for MAS, only a small number of loci is considered as significant loci linked to the trait of interest. In addition, most economic traits are influenced by several genes with several SNPs explaining a small proportion of genetic variance, highlighting that in this situation, MAS could not be successfully implemented.[136]

4.1 | Genomic selection in salmonid species

The highest number of genomic selection studies has been reported for salmonids, followed by non-salmonid fish and shellfish species (Table 6). We found 22 studies covering three salmonid species (*Salmo salar*, *Oncorhynchus mykiss* and *O. kisutch)* applying different SNP densities and statistical methods to evaluate the use of genomic information for genomic selection. Further, the accuracy of genomic selection was most often reported for growth and disease resistance traits.

Values of relative increase in accuracy compared to pedigree-based BLUP ranged from 9% to 52% for resistance to *L. salmonis*, *C. rogercresseyi*, *P. salmonis* and *N. perurans* in *S. salar*.[33,34,145–150] For growth traits (body weight and length), fillet colour and uniformity of growth, higher variation in the increase in accuracy was reported, with values ranging from 0% to 78%.[40,145,150–152]

Few studies were found evaluating the realized benefits of genomic selection in aquaculture species. For Tasmanian Atlantic salmon, a cumulative effect in the rate of genetic gain was observed after 2-years of genomic selection was implemented for improvement of amoebic gill disease resistance, fillet colour, maturation and growth traits.[153] This realized prediction accuracy was estimated by comparing the GEBVs at two different timepoints: before (at spawning) and after progeny information was available, which produced more robust and reliable validation of the prediction equation. Likewise, Vallejo et al.[154] showed that, using genomic selection, the accuracy of GEBVs for BCWD resistance in rainbow trout was 50% higher than pedigree-based methods. The authors used a progeny-test validation scheme which estimates more accurately the impact of genomic selection on commercial breeding. More modest figures (0%–15%) were also

TABLE 6 Genomic prediction assessment for fish and shellfish species reported in the literature

Species	Genotype technique	Number of markers (K)	Trait	Statistical Method	Range accuracy increase (%)[a]	References
S. salar	SNP array	0.5, 1, 5,10, 20, 33 and 112	Weight and length at 1 year	GBLUP[b]	0–20	40
S. salar	SNP array	31	Body weight and uniformity of growth	ssGBLUP[c]	0.19–0.78[d]	151
S. salar	SNP array	50[e]	Body weight	GBLUP	11	152
S. salar	SNP array[f]	1, 2, 4, 22, 55 and 220	*L. salmonis* resistance and fillet colour	IBD-GS[g] and GBLUP	15–52	150
S. salar	SNP array	0.25 and 25[e]	*C. rogercresseyi* resistance and body weight	GBLUP	19–21	145
S. salar	SNP array	0.5, 1, 3, 10, 20 and 50	*P. salmonis* resistance	GBLUP, BayesC and BLasso[h]	8–30	146
S. salar	SNP array	0.5, 1, 5, 10, 25 and 37	*C. rogercresseyi* resistance	GBLUP, BayesC and Blasso	19–22	147
S. salar	SNP array	0.5, 0.9, 1.8, 2.7, 3.5, 4.4, 5.3, 6.1, 6.8 and 7.1	*N. perurans* resistance	GBLUP	17–22	33
S. salar	SNP array	57	*N. perurans* resistance	GBLUP and ssGBLUP	9–17	34
S. salar	SNP array	215 and 750	*L. salmonis* resistance	GBLUP, BayesC and BayesGC[i]	38–50	148
S. salar	SNP array	3[e]	Amoebic gill disease resistance, fillet colour, sexual maturation and harvest weight	ssGBLUP	6–30[d]	153
S. salar	SNP array	0.1, 0.5, 1, 5, 10 and 50[e]	*C. rogercresseyi* resistance	GBLUP	0.49–0.78[j]	149
O. mykiss	SNP array	50	Fillet yield and firmness	ssGBLUP	35–42	270
O. mykiss	SNP array	57	Fatty acid composition	GBLUP	12–120	271
O. mykiss	RAD-Seq[k] and SNP array	24 and 57	*F. psychrophilum* resistance	ssGBLUP and wssGBLUP[l]	26–50	272
O. mykiss	SNP array	57	*F. psychrophilum* resistance	ssGBLUP, wssGBLUP and BayesB	83–109	154
O. mykiss	SNP array	0.5, 3, 10, 20 and 27	*P. salmonis* resistance	GBLUP, ssGBLUP, BLasso and BayesC	28–41	156
O. mykiss	SNP array	57	Infectious pancreatic necrosis virus resistance	ssGBLUP	7–11	186
O. mykiss	SNP array	57	Infectious haematopoietic necrosis resistance	ssGBLUP and wssGBLUP	42–154	68
O. mykiss	SNP array	57	*F. columnare* resistance	ssGBLUP and wssGBLUP	37–46	273
O. mykiss	SNP array	32 and 7	Bacterial cold water disease resistance	ssGBLUP, wssGBLUP, ssBMR-BayesB[m]	10–35	274
O. mykiss	SNP array	57	Female reproduction	GBLUP	10–35	76
O. kisutch	ddRAD-Seq[n]	9	*P. salmonis* resistance	GBLUP, ssGBLUP, wssGBLUP and BayesC	6–155	78
D. labrax	SNP array	57	Body weight, cortisol level, glucose level, lactate level and lysozyme level	GBLUP	0–28	166

TABLE 6 (Continued)

Species	Genotype technique	Number of markers (K)	Trait	Statistical Method	Range accuracy increase (%)[a]	References
D. labrax	RAD-Seq	9	Viral nervous necrosis resistance	rrBLUP[o], BayesA, BayesB and BayesC	8–13	90
L. crocea	ddRAD-Seq	35	Body weight gain rate, body length and survival rate	GBLUP and BayesB	0.10–0.11[d]	105
L. crocea	ddRAD-Seq	17	*Cryptocaryon irritans* resistance	GBLUP	0.31[d]	158
Paralichthys olivaceus	Genome-resequencing	5000	*Edwardsiella tarda* resistance	BayesB, ssGBLUP and wssGBLUP	8–22	159
O. niloticus	SNP array	0.5, 1, 3 and 32[e]	Harvest weight and fillet yield	ssGBLUP	4–27	101
O. niloticus	SNP array	50	Body weight, fillet weight and fillet yield	GBLUP	20–75	167
O. niloticus	SNP array	65	Feed conversion ratio, body weight gain, residual feed intake and feed intake	GBLUP and ssGBLUP	5–34	171
O. niloticus	Genome-resequencing	50 and 1000	*Streptococcus agalactiae* resistance	GBLUP, wGBLUP and BayesCπ	129–155	160
O. niloticus	SNP array	50	*Francisella orientalis* resistance and body weight	GBLUP, BayesB, BayesC and BayesS	0–217	274
C. carpio	GBS[p]	22	Body length	GBLUP	18	165
C. carpio haematoperus	SNP array	0.1, 0.5, 1, 2.5, 5, 10, 25, 50 and 75	Body weight, body length, body thick, gonad weight, carcass weight and morphological traits	GBLUP	0–38	168
C. carpio	RAD-Seq	15	Koi herpesvirus resistance	rrBLUP, BayesA, BayesB and BayesC	8–18	157
I. punctatus	SNP array	57	Harvest and carcass weight	ssGBLUP	28–36	164
I. punctatus	SNP array	250	*Flavobacterium columnare* resistance	GBLUP, ENGBLUP[q]	0.49–0.73[d]	162
Lates calcarifer	GBS	25	Body weight, body length, fillet weight and fillet yield	GBLUP	5–76	170
L. calcarifer	SNP array	70	Body weight, body length, body depth, body shape index, Fulton's condition factor and fillet weight	GBLUP	10–49	169
Epinephelus coioides	Genome-resequencing	0.005, 0.05, 0.1, 0.5, 1, 5, 10, 50, 200,500, 800 and 2000[e]	Ammonia tolerance	rrBLUP, BayesA, BayesB and BayesC	0.72–0.74[d]	163
Oplegnathus fasciatus	ddRAD-Seq	16	Body weight, body length and body depth	GBLUP, BLasso, BayesRR[r], BayesA, BayesB and BayesC	0.23–0.31[d]	172
C. farreri	2b-RAD-Seq[s]	31	Shell length, shell height, shell width, and whole weight	GBLUP, BayesB, RKHS[t] and SNN[u]	41–60	173
C. gigas	SNP array	23	Shell height, shell length, and wet weight	GBLUP	25–30	174
C. gigas	SNP array	23	Ostreid herpesvirus resistance	GBLUP	6–19	175

(Continues)

TABLE 6 (Continued)

Species	Genotype technique	Number of markers (K)	Trait	Statistical Method	Range accuracy increase (%)[a]	References
C. angulata	GBS	19	Shell length, shell width, shell depth, tenderness, taste, polychaete infestation and presence of *Marteilioides chungmuensis*	GBLUP, BayesA, BayesC,	15–200	177
C. angulata	GBS	19	Whole weight, body shape, meat yield and colour	GBLUP	12–171	176
C. virginica	ddRAD-Seq	29	Acute low salinity tolerance	BayesRR, BayesA, RKHS	0.48–0.57[d]	119
L. vannamei	2b-RAD-Seq	0.05, 0.1, 0.2, 0.4, 0.8, 1.6, 3.2, 6.4, 12.8, 20, 23	Body weight and body length	rrBLUP, BayesA and BLasso	0.50–0.62[j]	178
L. vannamei	2b-RAD-Seq	23	*Vibrio parahaemolyticus* resistance	GBLUP	3–7	179
L. vannamei	SNP array	18	White spot syndrome virus resistance	GBLUP	0.64–0.69[d]	180
Fenneropenaeus merguiensis	GBS	10	*Hepatopancreatic parvovirus* resistance, growth traits	GBLUP	11–567	181

[a]Compared to pedigree-based best linear unbiased prediction.
[b]Genomic best linear unbiased prediction.
[c]Single-step genomic best linear unbiased prediction.
[d]In terms of ability of prediction.
[e]Genotype imputed.
[f]Single-nucleotide polymorphism array.
[g]Identity-by-descent genomic selection.
[h]Bayesian least absolute shrinkage and selection operator.
[i]Model combines a polygenic term and a BayesC term.
[j]In terms of accuracy of GEBVs.
[k]Restriction site associated DNA sequencing.
[l]Weight single-step genomic best linear unbiased prediction.
[m]Single-step Bayesian multiple regression BayesB.
[n]Double digest restriction associated DNA sequencing.
[o]Ridge regression genomic best linear unbiased prediction.
[p]Genotyping by sequencing.
[q]Elastic net genomic best linear unbiased prediction.
[r]Bayesian ridge-regression.
[s]2b-restriction site-associated DNA sequencing.
[t]Kernel Hilbert spaces regression.
[u]Sparse neural networks.

obtained for rainbow trout using the same progeny evaluation scheme for IHNV resistance.[155]

A common strategy to evaluate the possibility of cost-effective use of genomic information in genomic selection is to estimate GEBVs using different SNP densities (e.g. from 0.5 to 220K[31,33,38,145,146,150,154,156]). This strategy may help to determine the SNP-density threshold needed to obtain GEBVs without reduction in accuracy. Different statistical methods have been tested for salmonid species; among them, the most used were the genomic best linear unbiased prediction (GBLUP) and single-step GBLUP (ssGBLUP), reported in 64% and 50% of the genomic selection studies for salmonids, respectively.

4.2 | Genomic selection in non-salmonid fish

Reports of genomic selection in non-salmonid fish species include 19 studies in nine different species: European sea bass (*D. labrax*), Nile tilapia (*O. niloticus*), common carp (*C. carpio*), large yellow croaker (*L. crocea*), Japanese flounder (*Paralichthys olivaceus*), channel catfish (*Ictalurus punctatus*), Asian sea bass (*L. calcariifer*), grouper (*Epinephelus coioides*) and striped knifejaw (*Oplegnathus fasciatus*). Palaiokostas et al.[90] found relative increases in genomic selection accuracy ranging from 8% to 13% for resistance against viral nervous necrosis in *D. labrax*. For resistance to koi herpes virus disease in *C. carpio*, the values ranged from 8% to 18%.[157] Zhao et al.[158] reported successful genomic selection in large yellow croaker (*Larimichthys crocea*) for resistance against *Cryptocaryon irritans*, a parasitic ciliate which causes huge economic losses to the aquaculture industry. The survival rate was significantly improved after one generation of genomic selection.[158] Lu et al.[159] found that genomic prediction of resistance against *Edwardsiella tarda* in Japanese flounder with 50K SNPs was better than pedigree-based prediction, and the accuracy of genomic breeding values were similar using ssGBLUP, wssGBLUP and BayesB (refer to Table 4 for acronyms) methods for prediction. In Nile tilapia, genomic selection for resistance against *Streptococcus agalactiae* and *Francisella orientalis*, resulted in both more than twice the ability of prediction over pedigree-based methods.[160,161] For *I. punctatus*, genomic selection demonstrated satisfactory results for *Flavobacterium columnare* resistance, but no pedigree was available, precluding comparison with genomic methods.[162] For ammonia tolerance in *E. coioides*, in which different SNP densities (5–2000K) were evaluated using different methods (rrBLUP, BayesA, BayesB and BayesC), significant improvement in prediction ability was revealed when using SNPs found significant in previous GWAS (i.e. the use of informative SNPs offered higher prediction accuracy in comparison to randomly selected SNPs).[163] Increase in growth (body weight and body length) and carcass quality (carcass weight and fillet yield) traits increased from 4% to 76% comparing the accuracy of GEBVs using genomic information and pedigree-based BLUP (EBVs) for *D. labrax*, *O. niloticus*, *C. carpio*, *I. punctatus* and *L. calcarifer*.[101,164–170] Joshi et al.[167] found the highest values (20%, 43% and 75%) for fillet yield, body weight and fillet weight, respectively, for *O. niloticus* using a high-density chip (50K), whereas Yoshida et al.[101] found values ranging from 4% to 27% using different SNP densities (0.5, 1, 3, and 32K) with true and imputed genotypes for harvest weight and fillet yield in the same species. Only one study was found for feed efficiency traits, with substantial improvement in accuracy of genomic prediction compared to pedigree breeding (up to 34%) in Nile tilapia.[171]

A number of genomic selection studies for *D. labrax*,[90] *L. crocea*,[105] *P. olivaceus*,[159] *O. niloticus*,[160,161] *C. carpio*,[157] *E. coioides*[163] and *O. fasciatus*[172] compared the performance of different genomic selection methods to estimate the GEBVs, including Bayesian (rrBLUP, BayesA, BayesB, BayesC and BLasso) and GBLUP (GBLUP, ssGBLUP and wssGBLUP) approaches. Although some methods outperformed others in terms of genomic prediction ability (e.g. higher prediction accuracy of Bayesian methods over GBLUP for *Streptococcus agalactiae* resistance in Nile tilapia[161]), the performance of each method may vary according to the architecture of the trait, heritability, reference population size and SNP density, among others factors, and no ideal method may be indicated without further investigation about breeding program design and determination of quantitative traits.

4.3 | Genomic selection in shellfish

Studies of genomic selection are scarce for shellfish. Until recently there have been studies only for Zhikong scallop (*Chlamys farreri*),[173] Pacific oyster (*C. gigas*),[174,175] Portuguese oyster (*C. angulata*),[176,177] eastern oyster (*C. virginica*),[119] Pacific white shrimp (*Litopenaeus vannamei*),[178-180] and banana shrimp (*Fenneropenaeus merguiensis*)[181] in the literature.

For scallops (*C. farreri*) and oysters (*C. gigas*, *C. angulata* and *C. virginica*), most of the studies are related to shell traits. Vu et al.[176,177] found an advantage of including genomic information in breeding programs evaluating colour of shell and mantle in *C. angulata*, doubling the prediction accuracy relative to only pedigree information. For *C. gigas*, genomic prediction increase was also positive for growth (up to 30%) and ostreid herpesvirus resistance (up to 19%) traits.[174,175]

In shrimp, the GEBVs were estimated mostly for growth and disease resistance traits. Significant benefit of genomic selection was found for body weight in *L. vannamei* (up to 62%) and more modest results for *Vibrio parahaemolyticus* resistance (up to 7%) using 2b-RAD-Seq technology to genotype animals.[178,179] For the same species, the application of a low-density SNP array was used for genomic selection purposes and, in spite of pedigree not being available to perform a direct comparison, there was a high realized genetic gain (up to 58%) observed for white spot syndrome virus resistance.[180] In *Fenneropenaeus merguiensis*, 12%–38% increase in accuracy using genomic selection was reported for growth traits. Accuracy was six times higher for *Hepatopancreatic parvovirus* resistance compared to applying only pedigree information.[181]

Most of the studies in shellfish used GBLUP. Four studies also applied Bayesian methods to estimate the GEBVs, with no significant difference respective approaches.[119,173,177,178]

4.4 | General discussion for genomic selection

Genomic selection has been successfully applied in some livestock species and for different economic traits.[15,182,183] In aquaculture, most of the genomic selection applications were evaluated for salmonids (43% of studies) and more limited numbers of studies for other freshwater (26%) and saltwater (11%) fish species, molluscs (11%), and crustaceans (6%; Table 6). The targeted traits were mostly growth-related (e.g. body weight and body length), carcass qualities (e.g. fillet yield and fillet colour) and resistance against diseases (e.g. *P. salmonis*, *C. rogercresseyi* and *F. psychrophilum*). The accuracies and advantages over conventional selection or MAS were dependent upon the trait involved, as well as the selection schemes and phenotyping strategies. Results showed that particular attention to the population structure and the traits included for genetic evaluation should be paid if multiple-traits are considered. It is important to mention that genomic selection including more than one trait simultaneously may be computationally challenging if the effect of each locus is considered in the evaluation.[184]

The first genomic selection study for aquaculture species was reported by Ødegård et al.[150] in an admixed *S. salar* population evaluated for fillet colour and resistance against *L. salmonis*. The authors evaluated different subsets of low-density markers (1, 2, 4, 22 and 55K) from a 220K SNP panel and used identity-by-descent genomic selection (IBS-GS) and GBLUP models to compare the accuracy of genomic selection to that of a pedigree-based model. The results of using genomic information outperformed the pedigree-based model independent of panel density and the model used for both traits, with relative increase in accuracy ranging from 15% to 52%. Further, increasing the number of markers from 22 to 220K only marginally improved the reliability (~1%). Similar studies for Atlantic salmon,[33,40,147] rainbow-trout[156] and Nile tilapia[101] also reported that low-density panels ranging from 0.5 to 20K were sufficient to obtain an accuracy close to the maximum value compared to a high-density panel. The population background (e.g. effective population size and admixture) could result in long-range linkage disequilibrium that may increase the accuracy of genomic selection even when using low-density panels.[24,77,185]

To estimate the accuracy of genomic prediction, it is common to use cross-validation strategies. In practical terms, the individuals in both validation and training data are related to some degree, resulting in accuracy of GEBVs higher than zero. In addition, family-based population structures are commonly used in cross-validation designs for aquaculture species, meaning that all animals can be related to some degree.[38,146,147,152,186] This close relationship between validation and training animals could increase the accuracy of genomic selection,[187] and eventually, overestimates GEBVs (e.g. when the markers are not in linkage disequilibrium with QTL[142]). It is important to highlight that the true potential of implementing genomic selection should be assessed using full evaluation methodologies comprising both predicted and realized genetic gains.[13] Further, the accuracy of genomic selection is affected by other factors, such as the number of individuals used in the reference population, number of markers, effective population size, degree of LD and architecture and heritability of traits.[188,189] Additionally, the accuracy of GEBV may be influenced using statistical methods, which mainly differ with respect to the prior assumption of marker distribution effects (Table 4) and the calculation of the genetic relationship matrix (e.g. ssGBLUP proposed by Aguilar et al.[190] as an alternative approach to GBLUP).

The genomic selection model generally outperformed the PBLUP models for all traits studied in aquaculture species (Table 6). For *S. salar*, for example, the improvement in accuracies ranged from 0% to 20% and from 8% to 52% for growth and resistance against disease traits, respectively, whereas the reliabilities for *O. kisutch* and *O. mykiss* for resistance against BCWD and *P. salmonis*, respectively, were up to 100% compared to PBLUP models.[78,155] The reliability of GEBV was marginally different between genomic selection models,[146,147,156,157,178] whereas Barría et al.[78] suggested superior accuracy for GBLUP and BayesC models compared to ssGBLUP and wssGBLUP. In agreement, Vallejo et al.[154] reported higher predictability for a Bayesian model (Bayes B) than ssGBLUP or wssGBLUP models.

The performance of genomic selection models is strongly dependent on the genetic architecture of the trait.[191] The GBLUP model, for example, assumes that the effect at all markers has a normal distribution,[11,192] which fits better for traits with polygenic inheritance, as for example, resistance against *P. salmonis* for Atlantic salmon,[25] and is more efficient using closely related individuals.[146,150,156] In contrast, the Bayesian models, such as BayesC and BayesB, presented better results for moderate- to large-effect QTLs controlling the trait,[193] similarly to marker-assisted selection methods, for instance, for fillet colour in Atlantic salmon.[150] Additionally, statistical models, such as ssGBLUP and wsGBLUP, have been commonly used for genomic selection in aquaculture species[78,151,186] (Table 6). Despite ssGBLUP having the same normal distribution of marker effects as GBLUP, the advantages are due to the use of additional phenotypes from related animals that have no genomic information but may be connected with genotyped animals via pedigree information.[190] In wssGBLUP, the marker variances are updated on each iteration and used as weights for the next iteration round,[194] emulating shrinkage Bayesian models, such as BayesC. More accurate results are expected for ssGBLUP compared to the GBLUP model as in Yoshida et al.,[156] whereas for oligogenic traits, better results are expected with wssGBLUP over ssGBLUP, as observed by Vallejo et al.[154] Some authors reported that the GEBVs estimated for different models were highly correlated, resulting in values up to 0.80 and some correlations close or equal to one as BayesB versus BayesC (correlation = 0.99, Vallejo et al.[154]) and BayesC versus ssGBLUP (correlation = 1.00, Bangera et al.[146]). However, both studies reported correlation between PBLUP and genomic selection models from 0.60 to 0.81, suggesting higher differences of predicted genetic merit between EBVs and GEBVs than between GEBVs estimated using different models. Despite several statistical methods being available for estimation of SNP effects, most BLUP studies are related to the simple replacement of the *A* matrix (pedigree) by the *G* matrix (genomic). Instead of modelling each SNP individually, the animal's effect is fitted

based on a genomic relationship matrix, a more efficient strategy in terms of computation time and demand.[195]

Previously, it was suggested that the high number of markers and the genotyping cost were the major limitations to the implementation of genomic selection.[136] However, the advances in molecular techniques and the continuous reduction in the cost per unit to genotype the individuals has changed the cost-effectiveness of implementing genomic selection.[196] Further, alternative strategies have been tested for reducing the cost to genotype the individuals, such as the use of low-density panels[24,40,146,147,150,156] and genotype imputation.[145,152,197] Genotype imputation consists of genotyping at high density (or sequencing) a few animals as reference to infer and impute non-genotyped markers of many target animals that were genotyped using lower-density SNP panels.[198]

The choice of panel density may represent a cost-effective strategy for genomic selection when a high number of selection candidates must be genotyped, as in the case of aquaculture species. Therefore, the combination of genotyping the selection candidates using low-density panels and a small proportion of highly related individuals (e.g. parents) using a high-density panel followed by genotypic imputation[199] could be a cost-effective strategy with little or no loss in accuracy of GEBV, in comparison to using all animals genotyped with high-density SNP panels.[101,145,152,197] Tsai et al.[145] imputed genotypes of *S. salar* from 0.25 to ~25K markers and found marginally lower values (~3% for both resistance to sea lice and body weight) of prediction accuracy than using true genotypes. Also in Atlantic salmon, Yoshida et al.[152] found identical values (0.73) of prediction accuracy for body weight using genotype imputation (from 0.5K to 50K) and true 50K genotypes. In Nile tilapia, the increase in GEBV value ranged from 8% to 25% using genotype imputation, and it was suggested that directed savings of 69% could be achieved by genotyping fewer animals using a high-density panel and proceeding with genotyping imputation.[101]

5 | SIGNATURES OF DOMESTICATION AND SELECTION

Domestication and selective breeding have resulted in important phenotypic changes in aquaculture species.[200] Intense selective breeding and adaptation to local environments have given raise to different strains of Atlantic salmon, Nile tilapia, rainbow trout and other species. Further, when positive selection pressures occur in a population undergoing domestication and artificial breeding, changes in allele frequencies, linkage disequilibrium and haplotype patterns will arise.[201] Characterizing genomic regions that are affected by selection may allow inferences about genomic regions, functionality and genes underlying the expression of specific traits[202]; thus, selection signature studies have been carried out in several species, in both wild and domestic populations. Selection signature studies performed on aquaculture species will be described below.

5.1 | Selection signatures in salmonids

In aquaculture species, the study of selection signatures has focused mostly on salmonids, especially Atlantic salmon. Some studies have concentrated on describing genetic differences between hatchery strains and wild populations to evaluate the effect of farmed salmon escaping into nature.[203,204] Posterior studies have been applied to investigate genomic regions underlying traits of importance to production and the effect of domestication in farmed Atlantic salmon. Gutierrez et al.[205] performed a study using a 6.5K SNP array and three methods based on genetic differentiation (F_{ST}) between populations to detect selection signatures in the Cermaq population, a Mowi strain, originated from wild Norwegian populations. They found 44 loci showing evidence of selection signatures, associated with molecular functions that could be related to traits such as growth, response to pathogens and environmental stressors. Interestingly, they found evidence of markers previously associated with early sexual maturation in Atlantic salmon.[42] Liu et al.[206] performed a study comparing a domestic population with its presumed wild founder population, finding several F_{ST} outlier loci putatively under selection near genes and quantitative trait loci (QTLs) for growth (somatostatin receptor 5, *sstr5*), appetite (melanocortin 4 receptor-like, *mc4r*), maturity (*vgll3* and thyrotropin receptor, *tshr*) and disease resistance (major histocompatibility complex, *mhc class2*). Likewise, López et al.[207] performed a study with two independent pairs of domesticated/wild populations to evaluate the effect of domestication and artificial selection. They found evidence of both nonparallel and parallel signatures of selection upon genes with molecular functions that might be associated with traits under domestication, such as growth (e.g. supervillin, *svil* and plexin-b2, *plxnb2*), behaviour (e.g. autism susceptibility candidate 2, *auts2* and bromodomain- and wd repeat-containing protein 3, *brwd3*), immune response (e.g. collagen alpha-1XIII chain, *coda1* and ubiquitin-conjugating enzyme E2F putative, *ube2f*), response to environmental stimuli (e.g. melanopsins, *opn4x1b2*) and reproduction (e.g. *myopalladin*, *mypn* and zona pellucida sperm-binding protein 3-like, *zp3*). In a different study, four farmed populations of Atlantic salmon with a common geographical origin, all of them derived from the Mowi strain, were used to assess how selective pressures have affected populations cultivated in different production environments.[208] They found potential candidate genes for traits with both biological and economic importance for Atlantic salmon, such as growth (e.g. kinase non-catalytic c-lobe domain containing 1, *kind1* and calcineurin like EF-hand protein 2, *chp2*), immune system function (e.g. potassium voltage-gated channel subfamily b member 2, *kcnb2*; zinc finger protein, *rlf*; Synergin gamma, *synrg*; sorting nexin 14, *snx14*; f-box and leucine-rich repeat protein 5, *fbxl5*; e2f transcription factor 4, *e2f4* and Bloom syndrome gene, *blm*) and behaviour (e.g. gamma-aminobutyric acid type a receptor subunit beta1, *gabrb1*; s-phase cyclin-a associated protein in the er, *scaper*; calsyntenin 3 *clstn3* and peroxisomal biogenesis factor 5, *pex5*). Additionally, using whole-genome sequencing, Bertolotti et al.[209] investigated changes in structural variant allele frequencies between wild and farmed Atlantic salmon, finding evidence of

polygenic selection upon behaviour, immunity, circadian control of metabolism and other traits.

Coho salmon (*O. kisutch*) were studied to identify genomic signatures of domestication; by genotyping 137K SNPs in two different lines selectively bred to improve growth rate for approximately eight generations in Chile. Several genomic regions that contain genes potentially involved in growth, immune system, behaviour and maturation traits showed evidence of selection.[210] Finally, in rainbow trout (*O. mykiss*), using a commercial population of 749 individuals genotyped with 36K SNPs, Cadiz et al.[211] reported approximately 100 SNPs under selection, including markers within autosomal inversions on *Omy 05* and *Omy 20*.

5.2 | Selection signatures in other species

Tilapia are a group of cichlid fish native to the Middle East and Africa[212] and include some of the most important warmwater fish species used in aquaculture.[213] The main species of tilapia used for cultivation are Nile tilapia, blue tilapia (*O. aureus*), Mozambique tilapia (*O. mossambicus*) and several species of *Sarotherodon*.[213] Many breeding programs for these species have been established at universities, small businesses, large multinational companies and consortia of combinations of these entities since the 1980s, which have led to the existence of many strains and phenotypic differences.[214] In comparison to salmonid species, fewer studies have been done on tilapia; to date, only two studies of signatures of selection have been carried out. Hong Xia et al.[214] sequenced the genome of 47 tilapia individuals, belonging to Mozambique, Nile and red tilapia species and strains; they detected over a hundred regions harbouring selection signatures in each evaluated tilapia strain. Candidate genes in these regions were mapped into five gene ontology (GO) categories of which the *wnt* signalling, *gnrh* (gonadotropin-releasing hormone) and *integrin* signalling pathways overlapped all populations evaluated. These pathways have important roles in animal growth, development and disease resistance.[215] Cádiz et al.[216] performed whole-genome re-sequencing of 326 individuals belonging to three strains of farmed Nile tilapia cultivated in Brazil and Costa Rica. They applied two haplotype-based tests (integrated haplotype score, iHS; and extended haplotype homozygosity between pairs of populations, Rsb) to detect selection signatures within the genomes of these populations and detected 16, 174, and 96 candidate genes subjected to selection, in the three evaluated strains, respectively. Enrichment analysis of these genes revealed associations with growth, immune system, reproduction, behaviour, adaptation to environmental conditions and nervous system. In a more recent study performed using whole-genome sequences from 20 individuals from the Sukamandi Indonesian strain, several selection signatures were found, revealing eight potential genes related to salinity tolerance (cell cycle-associated protein 1a, *caprin1a*; nucleobindin 2; ATP binding cassette subfamily; solute carrier family 12 member 1; calcium channel; unc-51 like autophagy activating kinase; solute carrier and cadherin-1, *cdh1*).[217] The authors also found that the Sukamandi strain is approximately 10% derived from blue tilapia (*O. aureus*), indicating a past hybridization event with Nile tilapia (*O. niloticus*). These results may be important to better understand the effect of artificial selection and domestication within the genome of Nile tilapia as well as informing future selective breeding.

Channel catfish (*I. punctatus*), native to North America, is an important species for freshwater cultivation.[218] Changes in morphological, behavioural and growth traits have been found in channel catfish during domestication[219,220]; however, the molecular bases of such changes are unknown. Sun et al.,[221] by sequencing 150 individuals belonging to four domestic and one wild populations, identified genomic regions harbouring selection signatures using the pooled heterozygosity (H_P) test. They detected 23 genomic regions with putative selective sweeps, spanning 11 genes. Some of these genes play roles related to aquaculture performance traits, such as hypoxia-inducible factor 1 ß (*hif1ß*), which is involved in response to hypoxia and tolerance of low dissolved oxygen levels.

Grass carp (*Ctenopharyngodon idella*) is one of the four major Chinese carp of important economic value, used as both food fish and for aquatic vegetation control, and has been cultured for over 1300 years in China.[222] Using 43K SNPs, scored by GBS, Shen et al.[223] conducted a study to infer population structure and evidence of local adaptation. They found evidence of both positive and balancing selection. Genes associated with loci under selection were involved in many biological functions, such as anatomical structure and function, developmental process, metabolic process, reproduction and immune system, among others.

Red sea bream (*Pagrus major*) is a species in the family Sparidae. Due to its rapid growth and easy adaptation to environmental conditions, red sea bream is one of the most important species cultivated in Japan, Korea and China.[224] In Japan and Korea, it has been cultivated since the 1960s and 1980s, respectively,[225,226] where breeding programs have been established to improve growth rate. Nam et al.[227] conducted a study by whole-genome re-sequencing in one wild and three farmed populations from Japan and Korea to study the effect of artificial selection. They applied the cross-population extended haplotype homozygosity (XP-EHH) and relative nucleotide diversity tests to perform comparisons between wild and farmed populations. They detected 420, 549 and 325 genes in each of three farmed populations with significantly enriched GO terms related to metabolic processes, such as fatty acid and monocarboxylic acid metabolic processes. Enrichment of metabolic processes has also been observed in other farmed fish species selected for faster growth rate.[228] Apart from metabolic processes, they found GO terms related to developmental processes, among them neuron development and positive regulation of neuron projection development that are related to central nervous system development.[229] The authors related these GO terms to behaviour traits, since change of behavioural traits has been observed in other domesticated fish species.[230,231] Nam et al.[227] discovered different genes under selection between these farmed populations, although all the breeding programs were aimed at improving growth, showing that artificial selection acted on different genomic regions, many independent breeding programs.

Finally, tambaqui or cachama (*Colossoma macropomum*), one of the most important neotropical freshwater fish used for aquaculture in South America, was investigated using the ddRAD sequencing approach in farmed populations from Brazil, Colombia and Peru to study the effects of recent artificial selection and domestication. Agudelo et al.[232] identified several genomic regions potentially associated with stress tolerance and immunity, suggesting local adaptation to the culture environment.

For non-fish species, only two studies were found evaluating selection signatures, and both for the mollusc phylum. A total of 371 greenlip abalones (*Haliotis laevigata*) from 13 different costal sites were genotyped with approximately 9K SNPs.[233] The sample sites were located close to commercial farms, in order to investigate the effect of the seascape upon selection signatures. Geographical mapping of the sample sites and oceanographic variables were also included in the analysis making it possible to associate several genes, to environmental heterogeneity in oxygen concentrations and minimum temperatures such as cytochrome c oxidase (*cox*), heme α-synthase and Pumilio homologue (*pum1*). Another study performed using whole-genome sequences from 30 individuals from two different strains and natural populations of Yesso scallop (*M. yessoensis*) found several genes potentially associated with growth and disease resistance traits, among them the lysophosphatidylcholine acyltransferase 1 (*lpcat1)* and tumour necrosis factor receptor-associated factor (*traf)*, related to carotenoid accumulation in the muscle and immune response in molluscs, respectively.[234]

6 | INCORPORATING WHOLE-GENOME RESEQUENCING

The availability of genome sequences for many individuals would be useful for searching for rare genetic variants associated with economic traits for aquaculture. Recent advances in next-generation sequencing technologies are contributing to reduced cost of sequencing and have made it possible to apply genomic selection using WGS data. The potential advantage of genomic selection using WGS was first suggested because WGS provides a large proportion of genome coverage, thus including most of the causal mutations. Therefore, the accuracy attributed to LD between SNPs and causative mutations is not as relevant for WGS as it is for screening on SNP chips, given that the causal mutations are mostly present in the WGS data.[15,235] However, previous genomic selection studies in *Drosophila melanogaster*, cattle, poultry or simulated data suggested little or no increase in prediction accuracy comparing the use of WGS over dense SNP chips.[235–241] The main reason for these results is the high number of rare SNPs present in WGS data. These rare SNPs most likely will not be represented in both validation and reference populations, decreasing the percentage of variance captured by them.[238] So far, studies to evaluate the accuracy of genomic selection with WGS data in aquaculture species are few, as was recently reported by Liu et al.[242] for resistance to *E. tarda* in Japanese flounder (*Paralichthys olivaceus*). The authors used information from whole-genome resequencing of 1052 Japanese flounder and reported predicted accuracy values of 0.60 and 0.61 through cross-validation strategies for GBLUP and BayesC, respectively, with high Pearson correlation between methods (0.95). Unfortunately, in this study pedigree information was not available to compare the accuracies of GEBV and EBV, and neither did they test different chip densities to evaluate the benefits in terms of accuracy of WGS using a high-density chip. Yoshida and Yánez[64] performed a GWAS using WGS-imputed genotypes and selected a subset of 50K SNPs which were more important for growth under chronic thermal stress in rainbow trout. The authors found that prioritizing significant SNPs selected from GWAS produced better prediction accuracy (1.2%–13.3% higher accuracy in comparison to the pedigree-based scenario).

7 | OUTSIDE THE PUBLIC DOMAIN

Scientific papers published in the public domain do not tell the full story of applying genomics in aquaculture. Genomic- and marker-assisted selection are also employed by commercial breeding companies, providing the aquaculture industries with genetically selected broodstock. The work done by these companies is sometimes reported in the public domain, other times not.

The Atlantic salmon breeding sector is a good example of such 'private' use of genomic selection and GWAS, as Atlantic salmon breeding companies were the first to employ these techniques in commercial breeding. The above-mentioned identification of a major QTL for resistance to the viral disease IPNV[9,10] raised the salmonid breeding sector's awareness of what marker-assisted selection can offer on top of established methods of selective breeding. In addition to the obvious benefit of faster genetic gain,[8] marker-assisted selection offered new opportunities for product diversification; genetic material with increased IPNV resistance could be sold at a premium, so that the costs of performing marker-assisted selection could be recovered.[8] In the following years, breeding companies working on Atlantic salmon added an increasing number of value-added products to their portfolio, all of which were based on one or more major or moderate QTLs. In particular, eggs could be sold which were particularly resistant to the viral diseases cardiomyopathy syndrome or pancreas disease,[28,243] resistant to the bacterial disease salmon rickettsial syndrome, or had a particularly red fillet colour.[244]

Many traits, however, turned out to be highly polygenic, controlled by many genes with small individual effects. Genomic selection became a realistic option with the advent of affordable high-throughput SNP genotyping systems, and from 2013 onwards, genomic selection was implemented in several breeding programs. Breeding companies often chose a policy of limited collaboration, meaning that SNP-chips were produced that were often tailor-made to suit one or a few populations, rather than being generic and based on wide collaborative consortia. Thus, the five published SNP-arrays for Atlantic salmon[25,150,245-247] are mirrored by a handful of arrays developed (and continuously updated) by breeding companies for private use. The most extensive collection of validated SNPs originated from

two ~450K Affymetrix Axiom arrays which were put into the public domain. A condensed version of these two arrays was later used for identifying a major QTL for age-at-maturity in Atlantic salmon, implying *vestigial-like protein 3* (*vgll3*) as the causative gene.[46]

Currently, commercial Atlantic salmon breeding programs typically employ genomic selection on polygenic traits such as growth rate and resistance to the ectoparasites salmon louse (*L. salmonis*) and *C. rogercresseyi*, as well as to disease resistance traits. Genomic selection has boosted genetic gain in aquaculture breeding, not least because most traits of importance cannot be measured directly upon the breeding candidates. Prior to the introduction of genomic selection, disease resistance traits, fillet quality traits, and so forth could be selected for only using family selection, which exploits only the between-family component of genetic variation (~50% of the total genetic variance). Genomic selection facilitates exploitation also of the within-family component.

8 | CONCLUSIONS AND FUTURE DIRECTIONS

In only two decades, the application of genomics in aquaculture breeding has gone from 'none' to 'common'. Today, many breeding programs in aquaculture employ marker-assisted or genomic selection in some form for propagating their stocks or for producing eggs. However, it is the lowest-hanging fruits which have been gathered: some major genes with very large effects on specific traits have been selected for and brought close to fixation. Genomic selection has been applied, but mostly in situations where training data has been available from close relatives of the breeding candidates. The exploitation of genomics in aquaculture breeding has thus just begun. Further focus is needed and expected for the future, particularly within these areas:

- Fine mapping and identification of causative variation: In aquaculture as in other species, a typical GWAS currently is performed using a SNP-chip harbouring ~50K SNPs. Likewise, genomic selection is done by genotyping breeding candidates and reference animals on a chip of similar density. This choice of marker density is often motivated more by practical and cost-related issues than by scientific needs, and one could argue whether 50K SNPs is 'high-density' considering that a genome is likely to harbour millions of SNPs in addition to more cryptic variations. If marker densities were higher, or if causative mutations could be assayed rather than (more or less) random SNPs, genetic testing would become more precise and more generic, that is, less reliant on continuous updating of reference data sets. Using causative mutations instead of using high-density SNP panels (≥50K SNPs) for genomic prediction is desirable, given that it has been shown that accuracy of selection can be maximized when accounting for all loci controlling a particular trait, which in turn may relax the need for updating prediction equations every generation.[238]
- Gene editing to introduce 'novel' traits or to fast-track genetic improvement: More so than land-based species, aquaculture species may potentially benefit from extant wild-living relatives as well as a plethora of related species, from which they often differ phenotypically.[248] For example, while Atlantic salmon and rainbow trout are susceptible to sea lice infestations, coho salmon and pink salmon are not. Future and ongoing projects will aim to identify the genetic factors which create resistance in some species and not in others, by comparing the genomes, transcriptomes and so forth, of the different species. If technology and policy allow,[249] these variants could be applied from one species to another using gene editing. Gene editing could be used to increase the frequency of beneficial, but minor, alleles without risking inbreeding.

These steps will be crucial for the utilization of more advanced genomic technologies in aquaculture species, allowing more effective utilization of genetic variation in production traits via precision breeding, which can be considered paramount to the continued successful development, efficiency and sustainability of aquaculture.

AUTHOR CONTRIBUTIONS

José Manuel Yáñez: Conceptualization; data curation; project administration; supervision; writing – original draft; writing – review and editing. **Agustín Barría:** Data curation; writing – original draft; writing – review and editing. **Maria Eugenia López:** Data curation; writing – original draft; writing – review and editing. **Thomas Moen:** Data curation; writing – original draft; writing – review and editing. **Baltasar Fernandes Garcia:** Data curation; writing – original draft; writing – review and editing. **Grazyella Yoshida:** Data curation; writing – original draft; writing – review and editing. **Peng Xu:** Conceptualization; data curation; writing – original draft; writing – review and editing.

ACKNOWLEDGEMENTS

José M. Yáñez acknowledges funding from FONDECYT regular grant (No. 1211761). José M. Yáñez and Peng Xu acknowledge funding from the ANSO grant (No. ANSO-CR-PP-2O21-03). Peng Xu acknowledges funding from the National Natural Science Foundation of China (U21A20264).

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

José M. Yáñez https://orcid.org/0000-0002-6612-4087
María E. López https://orcid.org/0000-0001-9272-0694
Peng Xu https://orcid.org/0000-0002-1531-5078

REFERENCES

1. Food and Agriculture Organization of the UN (FAO). The State of World Aquaculture Production: Sustainability in Action; 2020. https://www.fao.org/3/ca9229en/ca9229en.pdf

2. Gjedrem T. Genetic improvement for the development of efficient global aquaculture: a personal opinion review. *Aquaculture*. 2012; 344–349(44):12-22. doi:10.1016/j.aquaculture.2012.03.003
3. Rye M, Gjerde B, Gjedrem T. Genetic improvement programs for aquaculture species in developed countries. In *Proceedings of the 9th world congress on genetics applied to livestock production*, 2010;8:1-8.
4. Lhorente JP, Araneda M, Neira R, Yáñez JM. Advances in genetic improvement for salmon and trout aquaculture: the Chilean situation and prospects. *Rev Aquac*. 2019;11(2):340-353. doi:10.1111/raq.12335
5. Yáñez JM, Joshi R, Yoshida GM. Genomics to accelerate genetic improvement in tilapia. *Anim Genet*. 2020;51(5):658-674. doi:10.1111/AGE.12989
6. Georges M, Charlier C, Hayes B. Harnessing genomic information for livestock improvement. *Nat Rev Genet*. 2019;20(3):135-156. doi:10.1038/s41576-018-0082-2
7. Yue GH. Recent advances of genome mapping and marker-assisted selection in aquaculture. *Fish Fish*. 2014;15(3):376-396. doi:10.1111/faf.12020
8. Moen T, Torgersen J, Santi N, et al. Epithelial cadherin determines resistance to infectious pancreatic necrosis virus in Atlantic Salmon. *Genetics*. 2015;200(4):1313-1326.
9. Moen T, Baranski M, Sonesson AK, Kjøglum S. Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (Salmo salar): population-level associations between markers and trait. *BMC Genomics*. 2009;10:1-14. doi:10.1186/1471-2164-10-368
10. Houston RD, Haley CS, Hamilton A, et al. Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic Salmon (Salmo salar). *Genetics*. 2008;178(2):1109-1115. doi:10.1534/GENETICS.107.082974
11. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of Total genetic value using genome-wide dense marker maps. *Genetics*. 2001; 157(4):1819-1829. Accessed May 9, 2018. http://www.genetics.org/content/genetics/157/4/1819.full.pdf
12. Villanueva B, Fernández J, García-Cortés LA, Varona L, Daetwyler HD, Toro MA. Accuracy of genome-wide evaluation for disease resistance in aquaculture breeding programs. *J Anim Sci*. 2011;89(11):3433-3442. doi:10.2527/jas.2010-3814
13. Taylor JF. Implementation and accuracy of genomic selection. *Aquaculture*. 2014;420–421:S8-S14. doi:10.1016/j.aquaculture.2013.02.017
14. Sonesson AK, Meuwissen TH. Testing strategies for genomic selection in aquaculture breeding programs. *Genet Sel Evol*. 2009;41(1): 37. doi:10.1186/1297-9686-41-37
15. Wiggans GR, Cole JB, Hubbard SM, Sonstegard TS. Genomic selection in dairy cattle: the USDA experience. *Annu Rev Anim Biosci*. 2017;5(1):309-327. doi:10.1146/annurev-animal-021815-111422
16. Braasch I, Postlethwait JH. Polyploidy in fish and the teleost genome duplication. In: Soltis PS, Soltis DE, eds. *Polyploidy and Genome Evolution*. Vol 1. 1st ed. Springer-Verlag; 2012:341-383. doi:10.1007/978-3-642-31442-1_17/FIGURES/7
17. Danzmann RG, Davidson EA, Ferguson MM, et al. Distribution of ancestral proto-Actinopterygian chromosome arms within the genomes of 4R-derivative salmonid fishes (rainbow trout and Atlantic salmon). *BMC Genomics*. 2008;9(1):1-16. doi:10.1186/1471-2164-9-557/FIGURES/6
18. Xu P, Xu J, Liu G, et al. The allotetraploid origin and asymmetrical genome evolution of the common carp Cyprinus carpio. *Nat Commun*. 2019;10(1):1-11. doi:10.1038/s41467-019-12644-1
19. Yuan Z, Liu S, Zhou T, et al. Comparative genome analysis of 52 fish species suggests differential associations of repetitive elements with their living aquatic environments. *BMC Genomics*. 2018;19(1):1-10. doi:10.1186/s12864-018-4516-1
20. Zhang G, Fang X, Guo X, et al. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*. 2012; 490(7418):49-54. doi:10.1038/nature11413
21. Vij S, Kuhl H, Kuznetsova IS, et al. Chromosomal-level assembly of the Asian seabass genome using long sequence reads and multi-layered scaffolding. *PLoS Genet*. 2016;12(4):e1005954. doi:10.1371/JOURNAL.PGEN.1005954
22. Gupta A, Kounovsky-Shafer KL, Ravindran P, Schwartz DC. Optical mapping and nanocoding approaches to whole-genome analysis. *Microfluid Nanofluid*. 2016;20(3):1-14. doi:10.1007/S10404-015-1685-Y/FIGURES/7
23. Yáñez JM, Newman S, Houston RD. Genomics in aquaculture to better understand species biology and accelerate genetic progress. *Front Genet*. 2015;6(APR):1-3. doi:10.3389/fgene.2015.00128
24. Robledo D, Palaiokostas C, Bargelloni L, Martínez P, Houston R. Applications of genotyping by sequencing in aquaculture breeding and genetics. *Rev Aquac*. 2018;10(3):670-682. doi:10.1111/raq.12193
25. Correa K, Lhorente JP, López ME, et al. Genome-wide association analysis reveals loci associated with resistance against *Piscirickettsia salmonis* in two Atlantic salmon (*Salmo salar* L.) chromosomes. *BMC Genomics*. 2015;16(1):1-9. doi:10.1186/s12864-015-2038-7
26. Holborn MK, Ang KP, Elliott JAK, Powell F, Boulding EG. Genome wide association analysis for bacterial kidney disease resistance in a commercial north American Atlantic salmon (*Salmo salar*) population using a 50K SNP panel. *Aquaculture*. 2018;495:465-471. doi:10.1016/j.aquaculture.2018.06.014
27. Aslam ML, Robledo D, Krasnov A, et al. Quantitative trait loci and genes associated with salmonid alphavirus load in Atlantic salmon: implications for pancreas disease resistance and tolerance. *Sci Rep*. 2020;10(1):1-15. doi:10.1038/s41598-020-67405-8
28. Hillestad B, Makvandi-Nejad S, Krasnov A, Moghadam HK. Identification of genetic loci associated with higher resistance to pancreas disease (PD) in Atlantic salmon (*Salmo salar* L.). *BMC Genomics*. 2020; 21(1):1-13. doi:10.1186/s12864-020-06788-4
29. Hillestad B, Moghadam HK. Genome-wide association study of piscine myocarditis virus (PMCV) resistance in Atlantic Salmon (*Salmo salar*). *J Hered*. 2019;110(6):720-726. doi:10.1093/jhered/esz040
30. Hillestad B, Kristjánsson ÓH, Makvandi-Nejad S, Moghadam HK. Genome-wide association study confirms previous findings of major loci affecting resistance to piscine myocarditis virus in Atlantic salmon (*Salmo salar* L.). *Genes (Basel)*. 2020;11(6):608. doi:10.3390/genes11060608
31. Correa K, Lhorente JP, Bassini L, et al. Genome wide association study for resistance to Caligus rogercresseyi in Atlantic salmon (*Salmo salar* L.) using a 50K SNP genotyping array. *Aquaculture*. 2017;472:61-65. doi:10.1016/j.aquaculture.2016.04.008
32. Holborn MK, Rochus CM, Ang KP, et al. Family-based genome wide association analysis for salmon lice (*Lepeophtheirus salmonis*) resistance in north American Atlantic salmon using a 50K SNP array. *Aquaculture*. 2019;511:734215. doi:10.1016/j.aquaculture.2019.734215
33. Robledo D, Matika O, Hamilton A, Houston RD. Genome-wide association and genomic selection for resistance to amoebic gill disease in Atlantic Salmon. *G3 Genes Genomes Genet*. 2018;8(4):1195-1203. doi:10.1534/G3.118.200075
34. Aslam ML, Boison SA, Lillehammer M, Norris A, Gjerde B. Genome-wide association mapping and accuracy of predictions for amoebic gill disease in Atlantic salmon (*Salmo salar*). *Sci Rep*. 2020;10(1):1-9. doi:10.1038/s41598-020-63423-8
35. Robledo D, Gutiérrez AP, Barría A, Lhorente JP, Houston RD, Yáñez JM. Discovery and functional annotation of quantitative trait loci affecting resistance to sea lice in Atlantic salmon. *Front Genet*. 2019;10:1-10. doi:10.3389/fgene.2019.00056
36. Rochus CM, Holborn MK, Ang KP, et al. Genome-wide association analysis of salmon lice (*Lepeophtheirus salmonis*) resistance in a north American Atlantic salmon population. *Aquacult Res*. 2018;49(3): 1329-1338. doi:10.1111/are.13592

37. Boison SA, Gjerde B, Hillestad B, Makvandi-Nejad S, Moghadam HK. Genomic and transcriptomic analysis of amoebic gill disease resistance in Atlantic Salmon (*Salmo salar* L.). *Front Genet*. 2019;10:68. doi:10.3389/fgene.2019.00068
38. Tsai HY, Hamilton A, Tinch AE, et al. Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. *Genet Sel Evol*. 2016;48(1):1-11. doi:10.1186/s12711-016-0226-9
39. Taylor RS, Muller WJ, Cook MT, Kube PD, Elliott NG. Gill observations in Atlantic salmon (Salmo salar, L.) during repeated amoebic gill disease (AGD) field exposure and survival challenge. *Aquaculture*. 2009;290(1–2):1-8. doi:10.1016/j.aquaculture.2009.01.030
40. Tsai HY, Hamilton A, Tinch AE, et al. Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. *BMC Genomics*. 2015;16(1):969. doi:10.1186/s12864-015-2117-9
41. Yoshida GM, Lhorente JP, Carvalheiro R, Yáñez JM. Bayesian genome-wide association analysis for body weight in farmed Atlantic salmon (Salmo salar L.). *Anim Genet*. 2017;48(6):698-703. doi:10.1111/age.12621
42. Gutierrez AP, Yáñez JM, Fukui S, Swift B, Davidson WS. Genome-wide association study (GWAS) for growth rate and age at sexual maturation in Atlantic Salmon (*Salmo salar*). *PLoS One*. 2015;10(3):e0119730. doi:10.1371/journal.pone.0119730
43. Kijas J, McWilliam S, Naval Sanchez M, et al. Evolution of sex determination loci in Atlantic Salmon. *Sci Rep*. 2018;8(1):1-11. doi:10.1038/s41598-018-23984-1
44. Gabián M, Morán P, Fernández AI, et al. Identification of genomic regions regulating sex determination in Atlantic salmon using high density SNP data. *BMC Genomics*. 2019;20(1):764. doi:10.1186/s12864-019-6104-4
45. Barson NJ, Aykanat T, Hindar K, et al. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature*. 2015;528(7582):405-408. doi:10.1038/nature16062
46. Ayllon F, Kjærner-Semb E, Furmanek T, et al. The vgll3 locus controls age at maturity in wild and domesticated Atlantic Salmon (*Salmo salar* L.) males. *PLoS Genet*. 2015;11(11):e1005628. doi:10.1371/journal.pgen.1005628
47. Boulding EG, Ang KP, Elliott JAK, Powell F, Schaeffer LR. Differences in genetic architecture between continents at a major locus previously associated with sea age at sexual maturity in European Atlantic salmon. *Aquaculture*. 2019;500:670-678. doi:10.1016/j.aquaculture.2018.09.025
48. Mohamed AR, Verbyla KL, Al-Mamun HA, et al. Polygenic and sex specific architecture for two maturation traits in farmed Atlantic salmon. *BMC Genomics*. 2019;20(1):1-12. doi:10.1186/s12864-019-5525-4
49. Sinclair-Waters M, Ødegård J, Korsvoll SA, et al. Beyond large-effect loci: large-scale GWAS reveals a mixed large-effect and polygenic architecture for age at maturity of Atlantic salmon. *Genet Sel Evol*. 2020;52(1):1-11. doi:10.1186/s12711-020-0529-8
50. Sodeland M, Gaarder M, Moen T, et al. Genome-wide association testing reveals quantitative trait loci for fillet texture and fat content in Atlantic salmon. *Aquaculture*. 2013;408-409:169-174. doi:10.1016/j.aquaculture.2013.05.029
51. Helgeland H, Sodeland M, Zoric N, et al. Genomic and functional gene studies suggest a key role of beta-carotene oxygenase 1 like (bco1l) gene in salmon flesh color. *Sci Rep*. 2019;9(1):1-12. doi:10.1038/s41598-019-56438-3
52. Horn SS, Ruyter B, Meuwissen THE, Moghadam H, Hillestad B, Sonesson AK. GWAS identifies genetic variants associated with omega-3 fatty acid composition of Atlantic salmon fillets. *Aquaculture*. 2020;514:734494. doi:10.1016/j.aquaculture.2019.734494
53. Liu S, Vallejo RL, Palti Y, et al. Identification of single nucleotide polymorphism markers associated with bacterial cold water disease resistance and spleen size in rainbow trout. *Front Genet*. 2015;6(SEP):298. doi:10.3389/fgene.2015.00298
54. Palti Y, Vallejo RL, Gao G, et al. Detection and validation of QTL affecting bacterial cold water disease resistance in rainbow trout using restriction-site associated DNA sequencing. *PLoS One*. 2015;10(9):e0138435. doi:10.1371/JOURNAL.PONE.0138435
55. Vallejo RL, Liu S, Gao G, et al. Similar genetic architecture with shared and unique quantitative trait loci for bacterial cold water disease resistance in two rainbow trout breeding populations. *Front Genet*. 2017;8:1-15. doi:10.3389/fgene.2017.00156
56. Fraslin C, Brard-Fudulea S, D'Ambrosio J, et al. Rainbow trout resistance to bacterial cold water disease: two new quantitative trait loci identified after a natural disease outbreak on a French farm. *Anim Genet*. 2019;50(3):293-297. doi:10.1111/age.12777
57. Barria A, Marín-Nahuelpi R, Cáceres P, et al. Single-step genome-wide association study for resistance to *Piscirickettsia salmonis* in rainbow trout (*Oncorhynchus mykiss*). *G3 Genes Genomes Genet*. 2019;9(11):3833-3841. doi:10.1534/g3.119.400204
58. Zuo S, Karami AM, Ødegård J, et al. Immune gene expression and genome-wide association analysis in rainbow trout with different resistance to Yersinia ruckeri infection. *Fish Shellfish Immunol*. 2020;106(April):441-450. doi:10.1016/j.fsi.2020.07.023
59. Cáceres P, Barría A, Christensen K, et al. Genome-scale comparative analysis for host resistance against sea lice between Atlantic salmon and rainbow trout. *Sci Rep*. 2021;11(1):1-11. doi:10.1101/624031
60. Karami AM, Ødegård J, Marana MH, et al. A major QTL for resistance to vibrio anguillarum in rainbow trout. *Front Genet*. 2020;11(December):607558. doi:10.3389/fgene.2020.607558
61. Jaafar R, Ødegård J, Mathiessen H, et al. Quantitative trait loci (QTL) associated with resistance of rainbow trout Oncorhynchus mykiss against the parasitic ciliate Ichthyophthirius multifiliis. *J Fish Dis*. 2020;43(12):1591-1602. doi:10.1111/jfd.13264
62. Marana MH, Karami AM, Ødegård J, et al. Whole-genome association study searching for QTL for Aeromonas salmonicida resistance in rainbow trout. *Sci Rep*. 2021;11(1):1-12. doi:10.1038/s41598-021-97437-7
63. Liu S, Vallejo RL, Gao G, et al. Identification of single-nucleotide polymorphism markers associated with cortisol response to crowding in rainbow trout. *Marine Biotechnol*. 2015;17(3):328-337. doi:10.1007/s10126-015-9621-4
64. Yoshida GM, Yáñez JM. Increased accuracy of genomic predictions for growth under chronic thermal stress in rainbow trout by prioritizing variants from GWAS using imputed sequence data. *Evol Appl*. 2021;15(4):537-552. doi:10.1111/eva.13240
65. Legarra A, Croiseau P, Sanchez MP, et al. A comparison of methods for whole-genome QTL mapping using dense markers in four livestock species. *Genet Sel Evol*. 2015;47(1):1-10. doi:10.1186/s12711-015-0087-7
66. Wang H, Misztal I, Aguilar I, et al. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. *Front Genet*. 2014;5:134. doi:10.3389/fgene.2014.00134
67. Rodríguez FH, Flores-Mara R, Yoshida GM, et al. Genome-wide association analysis for resistance to infectious pancreatic necrosis virus identifies candidate genes involved in viral replication and immune response in rainbow trout (*Oncorhynchus mykiss*). *G3 Genes Genomes Genet*. 2019;9(9):2897-2904. doi:10.1534/g3.119.400463
68. Vallejo RL, Cheng H, Fragomeni BO, et al. Genome-wide association analysis and accuracy of genome-enabled breeding value predictions for resistance to infectious hematopoietic necrosis virus in a commercial rainbow trout breeding population. *Genet Sel Evol*. 2019;51(1):1-14. doi:10.1186/S12711-019-0489-Z/FIGURES/4
69. Reis Neto RV, Yoshida GM, Lhorente JP, Yáñez JM. Genome-wide association analysis for body weight identifies candidate genes related to development and metabolism in rainbow trout

(*Oncorhynchus mykiss*). *Mol Genet Genomics*. 2019;294(3):563-571. doi:10.1007/S00438-018-1518-2/TABLES/2
70. Gonzalez-Pena D, Gao G, Baranski M, et al. Genome-wide association study for identifying loci that affect fillet yield, carcass, and body weight traits in rainbow trout (*Oncorhynchus mykiss*). *Front Genet*. 2016;7:1-14. doi:10.3389/fgene.2016.00203
71. Ali A, Al-Tobasei R, Lourenco D, Leeds T, Kenney B, Salem M. Genome-wide identification of loci associated with growth in rainbow trout. *BMC Genomics*. 2020;21(1):1-16. doi:10.1186/s12864-020-6617-x
72. Salem M, Al-Tobasei R, Ali A, et al. Genome-wide association analysis with a 50K transcribed gene SNP-Chip identifies QTL affecting muscle yield in rainbow trout. *Front Genet*. 2018;9(SEP):387. doi:10.3389/fgene.2018.00387
73. Ali A, Al-Tobasei R, Lourenco D, Leed T, Kenney B, Salem M. Genome-wide scan for common variants associated with intramuscular fat and moisture content in rainbow trout. *BMC Genomics*. 2020;21(1):529. doi:10.21203/rs.3.rs-18231/v1
74. Blay C, Haffray P, Bugeon J, et al. Genetic parameters and genome-wide association studies of quality traits characterised using imaging Technologies in Rainbow Trout, Oncorhynchus mykiss. *Front Genet*. 2021;12:1-17. doi:10.3389/fgene.2021.639223
75. Fraslin C, Phocas F, Bestin A, et al. Genetic determinism of spontaneous masculinisation in XX female rainbow trout: new insights using medium throughput genotyping and whole-genome sequencing. *Sci Rep*. 2020;10(1):1-13. doi:10.1038/s41598-020-74757-8
76. D'Ambrosio J, Morvezen R, Brard-Fudulea S, et al. Genetic architecture and genomic selection of female reproduction traits in rainbow trout. *BMC Genomics*. 2020;21(1):1-14. doi:10.1186/s12864-020-06955-7
77. Barría A, Christensen KA, Yoshida G, et al. Whole genome linkage disequilibrium and effective population size in a Coho Salmon (Oncorhynchus kisutch) breeding population using a high-density SNP Array. *Front Genet*. 2019;10(MAY):498. doi:10.3389/fgene.2019.00498
78. Barría A, Christensen KAKA, Yoshida GM, et al. Genomic predictions and genome-wide association study of resistance against *Piscirickettsia salmonis* in Coho Salmon (*Oncorhynchus kisutch*) using ddRAD sequencing. *G3*. 2018;8(4):1183-1194. doi:10.1534/g3.118.200053
79. Lehnert SJ, Christensen KA, Vandersteen WE, et al. Carotenoid pigmentation in Salmon: variation in expression at BCO2-l locus controls a key fitness trait affecting red coloration. *Proc R Soc B Biol Sci*. 2019;286:20191588. doi:10.1098/rspb.2019.1588
80. Geng X, Sha J, Liu S, et al. A genome-wide association study in catfish reveals the presence of functional hubs of related genes within QTLs for columnaris disease resistance. *BMC Genomics*. 2015;16(1):196. doi:10.1186/s12864-015-1409-4
81. Shi H, Zhou T, Wang X, et al. Genome-wide association analysis of intra-specific QTL associated with the resistance for enteric septicemia of catfish. *Mol Genet Genomics*. 2018;293(6):1365-1378. doi:10.1007/s00438-018-1463-0
82. Zhou T, Liu S, Geng X, et al. GWAS analysis of QTL for enteric septicemia of catfish and their involved genes suggest evolutionary conservation of a molecular mechanism of disease resistance. *Mol Genet Genomics*. 2017;292(1):231-242. doi:10.1007/s00438-016-1269-x
83. Geng X, Liu S, Yao J, et al. A genome-wide association study identifies multiple regions associated with head size in catfish. *G3 Genes Genomes Genet*. 2016;6(10):3389-3398. doi:10.1534/g3.116.032201
84. Geng X, Liu S, Yuan Z, Jiang Y, Zhi D, Liu Z. A genome-wide association study reveals that genes with functions for bone development are associated with body conformation in catfish. *Marine Biotechnol*. 2017;19(6):570-578. doi:10.1007/s10126-017-9775-3
85. Li N, Zhou T, Geng X, et al. Identification of novel genes significantly affecting growth in catfish through GWAS analysis. *Mol Genet Genomics*. 2018;293(3):587-599. doi:10.1007/s00438-017-1406-1
86. Jin Y, Zhou T, Geng X, et al. A genome-wide association study of heat stress-associated SNPs in catfish. *Anim Genet*. 2017;48(2):233-236. doi:10.1111/age.12482
87. Zhong X, Wang X, Zhou T, et al. Genome-wide association study reveals multiple novel QTL associated with low oxygen tolerance in hybrid catfish. *Marine Biotechnol*. 2017;19(4):379-390. doi:10.1007/s10126-017-9757-5
88. Griot R, Allal F, Phocas F, et al. Genome-wide association studies for resistance to viral nervous necrosis in three populations of European sea bass (Dicentrarchus labrax) using a novel 57k SNP array DlabChip. *Aquaculture*. 2021;530:735930. doi:10.1016/J.AQUACULTURE.2020.735930
89. Wang L, Liu P, Huang S, et al. Genome-wide association study identifies loci associated with resistance to viral nervous necrosis disease in Asian seabass. *Marine Biotechnol*. 2017;19(3):255-265. doi:10.1007/s10126-017-9747-7
90. Palaiokostas C, Cariou S, Bestin A, et al. Genome-wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (Dicentrarchus labrax) using RAD sequencing. *Genet Sel Evol*. 2018;50(1):30. doi:10.1186/s12711-018-0401-2
91. Peng W, Xu J, Zhang Y, et al. An ultra-high density linkage map and QTL mapping for sex and growth-related traits of common carp (Cyprinus carpio). *Sci Rep*. 2016;6(1):1-16. doi:10.1038/srep26693
92. Chen L, Peng W, Kong S, et al. Genetic mapping of head size related traits in common carp (*Cyprinus carpio*). *Front Genet*. 2018;9:448. doi:10.3389/fgene.2018.00448
93. Zheng X, Kuang Y, Lv W, Cao D, Sun Z, Sun X. Genome-wide association study for muscle fat content and abdominal fat traits in common carp (*Cyprinus carpio*). *PLoS One*. 2016;11(12):e0169127. doi:10.1371/journal.pone.0169127
94. Zhang H, Xu P, Jiang Y, et al. Genomic, transcriptomic, and epigenomic features differentiate genes that are relevant for muscular polyunsaturated fatty acids in the common carp. *Front Genet*. 2019;10:217. doi:10.3389/fgene.2019.00217
95. Zhou Z, Chen L, Dong C, et al. Genome-scale association study of abnormal scale pattern in Yellow River carp identified previously known causative gene in European Mirror carp. *Marine Biotechnol*. 2018;20(5):573-583. doi:10.1007/s10126-018-9827-3
96. Palaiokostas C, Robledo D, Vesely T, et al. Mapping and sequencing of a significant quantitative trait locus affecting resistance to koi herpesvirus in common carp. *G3 Genes Genomes Genet*. 2018;8(11):3507-3513. doi:10.1534/g3.118.200593
97. Jia Z, Chen L, Ge Y, et al. Genetic mapping of koi herpesvirus resistance (KHVR) in Mirror carp (Cyprinus carpio) revealed genes and molecular mechanisms of disease resistance. *Aquaculture*. 2020;519:734850. doi:10.1016/j.aquaculture.2019.734850
98. Robinson N, Baranski M, Mahapatra KD, et al. A linkage map of transcribed single nucleotide polymorphisms in rohu (Labeo rohita) and QTL associated with resistance to *Aeromonas hydrophila*. *BMC Genomics*. 2014;15(1):1-23. doi:10.1186/1471-2164-15-541/FIGURES/5
99. Jiang DL, Gu XH, Li BJ, et al. Identifying a long QTL cluster across chrLG18 associated with salt tolerance in tilapia using GWAS and QTL-seq. *Marine Biotechnol*. 2019;21(2):250-261. doi:10.1007/s10126-019-09877-y
100. Cáceres G, López ME, Cádiz MI, et al. Fine mapping using whole-genome sequencing confirms anti-Müllerian hormone as a major gene for sex determination in farmed *Nile tilapia* (*Oreochromis niloticus* L.). *G3 Genes Genomes Genet*. 2019;9(10):3213-3223. doi:10.1534/g3.119.400297
101. Yoshida GM, Lhorente JP, Correa K, Soto J, Salas D, Yáñez JM. Genome-wide association study and cost-efficient genomic predictions for growth and fillet yield in *Nile tilapia* (*Oreochromis niloticus*). *G3 Genes Genomes Genet*. 2019;9(8):2597-2607. doi:10.1534/g3.119.400116
102. Lin H, Zhou Z, Zhao J, et al. Genome-wide association study identifies genomic loci of sex determination and Gonadosomatic index

traits in large yellow croaker (*Larimichthys crocea*). *Marine Biotechnol*. 2021;23(1):127-139. doi:10.1007/s10126-020-10007-2
103. Zhou Z, Han K, Wu Y, et al. Genome-wide association study of growth and body-shape-related traits in large yellow croaker (*Larimichthys crocea*) using ddRAD sequencing. *Marine Biotechnol*. 2019;21(5):655-670. doi:10.1007/s10126-019-09910-0
104. Kong S, Zhou Z, Zhou T, et al. Genome-wide association study of body shape-related traits in large yellow croaker (*Larimichthys crocea*). *Marine Biotechnol*. 2020;22(5):631-643. doi:10.1007/s10126-020-09983-2
105. Ke Q, Wang J, Bai Y, et al. GWAS and genomic prediction revealed potential for genetic improvement of large yellow croaker adapting to high plant protein diet. *Aquaculture*. 2022;553:738090. doi:10.1016/J.AQUACULTURE.2022.738090
106. Wu Y, Zhou Z, Pan Y, et al. GWAS identified candidate variants and genes associated with acute heat tolerance of large yellow croaker. *Aquaculture*. 2021;540:736696. doi:10.1016/J.AQUACULTURE.2021.736696
107. Kong S, Ke Q, Chen L, et al. Constructing a high-density genetic linkage map for large yellow croaker (*Larimichthys crocea*) and mapping resistance trait against ciliate parasite *Cryptocaryon irritans*. *Marine Biotechnol*. 2019;21(2):262-275. doi:10.1007/s10126-019-09878-x
108. Zhao J, Zhou T, Bai H, et al. Genome-wide association analysis reveals the genetic architecture of parasite (*Cryptocaryon irritans*) resistance in large yellow croaker (*Larimichthys crocea*). *Marine Biotechnol*. 2021;23(2):242-254. doi:10.1007/S10126-021-10019-6/TABLES/4
109. Palaiokostas C, Ferarreso S, Franch R, Houston RD, Bargelloni L. Genomic prediction of resistance to pasteurellosis in gilthead sea bream (*Sparus aurata*) using 2b-RAD sequencing. *G3 Genes Genomes Genet*. 2016;6:3693-3700. doi:10.1534/g3.116.035220
110. Zhou Q, Su Z, Li Y, et al. Genome-wide association mapping and gene expression analyses reveal genetic mechanisms of disease resistance variations in *Cynoglossus semilaevis*. *Front Genet*. 2019;10:1167. doi:10.3389/fgene.2019.01167
111. Meng J, Song K, Li C, et al. Genome-wide association analysis of nutrient traits in the oyster *Crassostrea gigas*: genetic effect and interaction network. *BMC Genomics*. 2019;20(1):1-14. doi:10.1186/s12864-019-5971-z
112. Meng J, Wang W, Shi R, et al. Identification of SNPs involved in Zn and Cu accumulation in the Pacific oyster (*Crassostrea gigas*) by genome-wide association analysis. *Ecotoxicol Environ Saf*. 2020;192:110208. doi:10.1016/j.ecoenv.2020.110208
113. Shi R, Li C, Qi H, et al. Construction of a high-resolution genetic map of *Crassostrea gigas*: QTL mapping and GWAS applications revealed candidate genes controlling nutritional traits. *Aquaculture*. 2020;527:735427. doi:10.1016/j.aquaculture.2020.735427
114. Liu S, Li L, Meng J, et al. Association and functional analyses revealed that PPP1R3B plays an important role in the regulation of glycogen content in the Pacific oyster *Crassostrea gigas*. *Front Genet*. 2019;10:106. doi:10.3389/fgene.2019.00106
115. Yu F, Peng W, Tang B, et al. A genome-wide association study of heat tolerance in Pacific abalone based on genome resequencing. *Aquaculture*. 2021;536:736436. doi:10.1016/j.aquaculture.2021.736436
116. Gutierrez AP, Bean TP, Hooper C, et al. A genome-wide association study for host resistance to ostreid herpesvirus in Pacific oysters (*Crassostrea gigas*). *G3 Genes Genomes Genet*. 2018;8(4):1273-1280. doi:10.1534/g3.118.200113
117. Yang B, Zhai S, Zhang F, et al. Genome-wide association study toward efficient selection breeding of resistance to vibrio alginolyticus in Pacific oyster, *Crassostrea gigas*. *Aquaculture*. 2022;548:737592. doi:10.1016/J.AQUACULTURE.2021.737592
118. He X, Li C, Qi H, et al. A genome-wide association study to identify the genes associated with shell growth and shape-related traits in *Crassostrea gigas*. *Aquaculture*. 2021;543:736926. doi:10.1016/j.aquaculture.2021.736926
119. McCarty AJ, Allen SK, Plough LV. Genome-wide analysis of acute low salinity tolerance in the eastern oyster *Crassostrea virginica* and potential of genomic selection for trait improvement. *G3 Genes Genomes Genet*. 2022;12(1):1-14. doi:10.1093/G3JOURNAL/JKAB368
120. Ning X, Li X, Wang J, et al. Genome-wide association study reveals E2F3 as the candidate gene for scallop growth. *Aquaculture*. 2019;511:734216. doi:10.1016/J.AQUACULTURE.2019.734216
121. Yang Z, Sun F, Liao H, et al. Genome-wide association study reveals genetic variations associated with ocean acidification resilience in yesso scallop *Patinopecten yessoensis*. *Aquat Toxicol*. 2021;240:105963. doi:10.1016/J.AQUATOX.2021.105963
122. Zhao L, Li Y, Li Y, et al. A genome-wide association study identifies the genomic region associated with Shell color in yesso scallop, *Patinopecten yessoensis*. *Marine Biotechnol*. 2017;19(3):301-309. doi:10.1007/S10126-017-9751-Y/FIGURES/5
123. Wang S, Wang H, Zhao L, et al. Identification of genes associated with carotenoids accumulation in scallop (*Patinopecten yessoensis*). *Aquaculture*. 2022;550:737850. doi:10.1016/J.AQUACULTURE.2021.737850
124. Li X, Wang S, Xun X, et al. A carotenoid oxygenase is responsible for muscle coloration in scallop. *Biochim Biophys Acta Mol Cell Biol Lipids*. 2019;1864(7):966-975. doi:10.1016/J.BBALIP.2019.03.003
125. Yu Y, Wang Q, Zhang Q, et al. Genome scan for genomic regions and genes associated with growth trait in Pacific white shrimp *Litopeneaus vannamei*. *Marine Biotechnol*. 2019;21(3):374-383. doi:10.1007/s10126-019-09887-w
126. Wang Q, Yu Y, Zhang Q, et al. A novel candidate gene associated with body weight in the Pacific white shrimp *Litopenaeus vannamei*. *Front Genet*. 2019;10(MAY):520. doi:10.3389/fgene.2019.00520
127. Lyu D, Yu Y, Wang Q, et al. Identification of growth-associated genes by genome-wide association study and their potential application in the breeding of Pacific white shrimp (*Litopenaeus vannamei*). *Front Genet*. 2021;12:465. doi:10.3389/FGENE.2021.611570/BIBTEX
128. Jones DB, Nguyen HT, Khatkar MS, et al. The identification of a major sex QTL in the white-leg shrimp, *Litopenaeus vannamei*. *Aquaculture*. 2020;529:735673. doi:10.1016/J.AQUACULTURE.2020.735673
129. Robinson NA, Gopikrishna G, Baranski M, et al. QTL for white spot syndrome virus resistance and the sex-determining locus in the Indian black tiger shrimp (*Penaeus monodon*). *BMC Genomics*. 2014;15(1):1-21. doi:10.1186/1471-2164-15-731/FIGURES/5
130. Guo L, Xu YH, Zhang N, et al. A high-density genetic linkage map and QTL mapping for sex in black tiger shrimp (*Penaeus monodon*). *Front Genet*. 2019;10:326. doi:10.3389/FGENE.2019.00326/BIBTEX
131. Garcia BF, Bonaguro Á, Araya C, Carvalheiro R, Yáñez JM. Application of a novel 50K SNP genotyping array to assess the genetic diversity and linkage disequilibrium in a farmed Pacific white shrimp (*Litopenaeus vannamei*) population. *Aquac Rep*. 2021;20:100691. doi:10.1016/J.AQREP.2021.100691
132. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. 2019;20(8):467-484. doi:10.1038/s41576-019-0127-1
133. You X, Shan X, Shi Q. Research advances in the genomics and applications for molecular breeding of aquaculture animals. *Aquaculture*. 2020;526:735357. doi:10.1016/J.AQUACULTURE.2020.735357
134. Geng X, Zhi D, Liu Z. Genome-wide association studies of performance traits. In: Liu Z, ed. *Bioinformatics in Aquaculture: Principles and Methods*. Vol 1. 1st ed. John Wiley & Sons; 2017:415-433.
135. Garcia BF, de Melo TP, de Neves HH, Carvalheiro R. Comparison of GWA statistical methods for traits under different genetic structures: a simulation study. *Livest Sci*. 2020;241:104213. doi:10.1016/j.livsci.2020.104213

136. Goddard ME, Hayes BJ. Genomic selection. *J Anim Breed Genet*. 2007;124(6):323-330. doi:10.1111/J.1439-0388.2007.00702.X
137. Heffner EL, Sorrells ME, Jannink JL. Genomic selection for crop improvement. *Crop Sci*. 2009;49(1):1-12. doi:10.2135/CROPSCI2008.08.0512
138. Muir WM. Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. *J Anim Breed Genet*. 2007;124(6):342-355. doi:10.1111/J.1439-0388.2007.00700.X
139. Zenger KR, Khatkar MS, Jones DB, Khalilisamani N, Jerry DR, Raadsma HW. Genomic selection in aquaculture: application, limitations and opportunities with special reference to marine shrimp and pearl oysters. *Front Genet*. 2019;10(JAN):693. doi:10.3389/FGENE.2018.00693/BIBTEX
140. Vela-Avitúa S, Meuwissen TH, Luan T, Ødegård J. Accuracy of genomic selection for a sib-evaluated trait using identity-by-state and identity-by-descent relationships. *Genet Sel Evol*. 2015;47(1):1-6. doi:10.1186/S12711-014-0084-2/FIGURES/2
141. Cao L, Liu H, Mulder HA, et al. Genomic breeding programs realize larger benefits by cooperation in the presence of genotype × environment interaction than conventional breeding programs. *Front Genet*. 2020;11:251. doi:10.3389/FGENE.2020.00251/BIBTEX
142. Habier D, Fernando RL, Dekkers JCM. The impact of genetic relationship information on genome-assisted breeding values. *Genetics*. 2007;177(4):2389-2397. doi:10.1534/GENETICS.107.081190
143. Forneris NS, Steibel JP, Legarra A, et al. A comparison of methods to estimate genomic relationships using pedigree and markers in livestock populations. *J Anim Breed Genet*. 2016;133(6):452-462. doi:10.1111/JBG.12217
144. Meuwissen T, Hayes B, Goddard M. Accelerating improvement of livestock with genomic selection. *Annu Rev Anim Biosci*. 2013;1:221-237. doi:10.1146/ANNUREV-ANIMAL-031412-103705
145. Tsai HY, Matika O, Edwards SMK, et al. Genotype imputation to improve the cost-efficiency of genomic selection in farmed Atlantic salmon. *G3 Genes Genomes Genet*. 2017;7(4):1377-1383. doi:10.1534/g3.117.040717
146. Bangera R, Correa K, Lhorente JP, Figueroa R, Yáñez JM. Genomic predictions can accelerate selection for resistance against *Piscirickettsia salmonis* in Atlantic salmon (*Salmo salar*). *BMC Genomics*. 2017;18(1):1-12. doi:10.1186/S12864-017-3487-Y/FIGURES/2
147. Correa K, Bangera R, Figueroa R, Lhorente JP, Yáñez JM. The use of genomic information increases the accuracy of breeding value predictions for sea louse (*Caligus rogercresseyi*) resistance in Atlantic salmon (*Salmo salar*). *Genet Sel Evol*. 2017;49(1):1-5. doi:10.1186/s12711-017-0291-8
148. Kjetså MH, Ødegård J, Meuwissen THE. Accuracy of genomic prediction of host resistance to salmon lice in Atlantic salmon (*Salmo salar*) using imputed high-density genotypes. *Aquaculture*. 2020;526:735415. doi:10.1016/J.AQUACULTURE.2020.735415
149. Fraslin C, Yáñez JM, Robledo D, Houston RD. The impact of genetic relationship between training and validation populations on genomic prediction accuracy in Atlantic salmon. *Aquac Rep*. 2022;23:101033. doi:10.1016/J.AQREP.2022.101033
150. Ødegård J, Moen T, Santi N, Korsvoll SA, Kjøglum S, Meuwisse THE. Genomic prediction in an admixed population of Atlantic salmon (Salmo salar). *Front Genet*. 2014;5:402. doi:10.3389/FGENE.2014.00402/BIBTEX
151. Sae-Lim P, Kause A, Lillehammer M, Mulder HA. Estimation of breeding values for uniformity of growth in Atlantic salmon (*Salmo salar*) using pedigree relationships or single-step genomic evaluation. *Genet Sel Evol*. 2017;49(1):1-12. doi:10.1186/s12711-017-0308-3
152. Yoshida GM, Carvalheiro R, Lhorente JP, et al. Accuracy of genotype imputation and genomic predictions in a two-generation farmed Atlantic salmon population using high-density and low-density SNP panels. *Aquaculture*. 2018;491:147-154. doi:10.1016/j.aquaculture.2018.03.004
153. Verbyla KL, Kube PD, Evans BS. Commercial implementation of genomic selection in Tasmanian Atlantic salmon: scheme evolution and validation. *Evol Appl*. 2022;15(4):631-644. doi:10.1111/EVA.13304
154. Vallejo RL, Leeds TD, Gao G, et al. Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture. *Genet Sel Evol*. 2017;49(1):1-13. doi:10.1186/s12711-017-0293-6
155. Vallejo RL, Fragomeni BO, Cheng H, et al. Assessing accuracy of genomic predictions for resistance to infectious hematopoietic necrosis virus with progeny testing of selection candidates in a commercial rainbow trout breeding population. *Front Vet Sci*. 2020;7:939. doi:10.3389/FVETS.2020.590048/BIBTEX
156. Yoshida GM, Bangera R, Carvalheiro R, et al. Genomic prediction accuracy for resistance against *Piscirickettsia salmonis* in farmed rainbow trout. *G3 Genes Genomes Genet*. 2018;8(2):719-726. doi:10.1534/g3.117.300499
157. Palaiokostas C, Vesely T, Kocour M, et al. Optimizing genomic prediction of host resistance to koi herpesvirus disease in carp. *Front Genet*. 2019;10:543. doi:10.3389/FGENE.2019.00543/BIBTEX
158. Zhao J, Bai H, Ke Q, et al. Genomic selection for parasitic ciliate *Cryptocaryon irritans* resistance in large yellow croaker. *Aquaculture*. 2021;531:735786. doi:10.1016/J.AQUACULTURE.2020.735786
159. Lu S, Liu Y, Yu X, et al. Prediction of genomic breeding values based on pre-selected SNPs using ssGBLUP, WssGBLUP and BayesB for Edwardsiellosis resistance in Japanese flounder. *Genet Sel Evol*. 2020;52(1):1-10. doi:10.1186/S12711-020-00566-2/TABLES/2
160. Lu S, Zhu J, Du X, et al. Genomic selection for resistance to *Streptococcus agalactiae* in GIFT strain of *Oreochromis niloticus* by GBLUP, wGBLUP, and BayesCπ. *Aquaculture*. 2020;523:735212. doi:10.1016/j.aquaculture.2020.735212
161. Joshi R, Skaarud A, Alvarez AT, Moen T, Ødegård J. Bayesian genomic models boost prediction accuracy for survival to *Streptococcus agalactiae* infection in *Nile tilapia* (*Oreochromus nilioticus*). *Genet Sel Evol*. 2021;53(1):37. doi:10.1186/s12711-021-00629-y
162. Zhang Y, Liu Z, Li H. Genomic prediction of Columnaris disease resistance in catfish. *Marine Biotechnol*. 2020;22(1):145-151. doi:10.1007/S10126-019-09941-7/TABLES/4
163. Shan X, Xu T, Ma Z, et al. Genome-wide association improves genomic selection for ammonia tolerance in the orange-spotted grouper (*Epinephelus coioides*). *Aquaculture*. 2021;533:736214. doi:10.1016/J.AQUACULTURE.2020.736214
164. Garcia ALS, Bosworth B, Waldbieser G, Misztal I, Tsuruta S, Lourenco DAL. Development of genomic predictions for harvest and carcass weight in channel catfish. *Genet Sel Evol*. 2018;50(1):1-12. doi:10.1186/S12711-018-0435-5/FIGURES/4
165. Palaiokostas C, Kocour M, Prchal M, Houston RD. Accuracy of genomic evaluations of juvenile growth rate in common carp (*Cyprinus carpio*) using genotyping by sequencing. *Front Genet*. 2018;9:82. doi:10.3389/FGENE.2018.00082/BIBTEX
166. Oikonomou S, Samaras A, Tekeoglou M, et al. Genomic selection and genome-wide association analysis for stress response, disease resistance and body weight in European seabass. *Animals*. 2022;12(3):277. doi:10.3390/ANI12030277
167. Joshi R, Skaarud A, de Vera M, Alvarez AT, Ødegård J. Genomic prediction for commercial traits using univariate and multivariate approaches in *Nile tilapia* (*Oreochromis niloticus*). *Aquaculture*. 2020;516:734641. doi:10.1016/j.aquaculture.2019.734641
168. Wang J, Chen L, Li B, et al. Performance of genome prediction for morphological and growth-related traits in Yellow River carp. *Aquaculture*. 2021;536:736463. doi:10.1016/J.AQUACULTURE.2021.736463

169. Jerry DR, Jones DB, Lillehammer M, et al. Predicted strong genetic gains from the application of genomic selection to improve growth related traits in barramundi (*Lates calcarifer*). *Aquaculture*. 2022;549:737761. doi:10.1016/J.AQUACULTURE.2021.737761
170. Sukhavachana S, Senanan W, Pattarapanyawong N, et al. Multiple-trait genomic prediction of harvest and fillet traits in Asian seabass (*Lates calcarifer*, Bloch 1790). *Aquaculture*. 2021;544:737069. doi:10.1016/J.AQUACULTURE.2021.737069
171. Barría A, Benzie JAH, Houston RD, De Koning DJ, de Verdal H. Genomic selection and genome-wide association study for feed-efficiency traits in a farmed *Nile tilapia* (*Oreochromis niloticus*) population. *Front Genet*. 2021;12:1796. doi:10.3389/FGENE.2021.737906/BIBTEX
172. Gong J, Zhao J, Ke Q, et al. First genomic prediction and genome-wide association for complex growth-related traits in rock bream (*Oplegnathus fasciatus*). *Evol Appl*. 2021;15(4):523-536. doi:10.1111/EVA.13218
173. Wang Y, Sun G, Zeng Q, et al. Predicting growth traits with genomic selection methods in Zhikong scallop (*Chlamys farreri*). *Marine Biotechnol*. 2018;20(6):769-779. doi:10.1007/S10126-018-9847-Z/FIGURES/4
174. Gutierrez AP, Matika O, Bean TP, Houston RD. Genomic selection for growth traits in Pacific oyster (*Crassostrea gigas*): potential of low-density marker panels for breeding value prediction. *Front Genet*. 2018;9:391. doi:10.3389/FGENE.2018.00391
175. Gutierrez AP, Symonds J, King N, Steiner K, Bean TP, Houston RD. Potential of genomic selection for improvement of resistance to ostreid herpesvirus in Pacific oyster (*Crassostrea gigas*). *Anim Genet*. 2020;51(2):249-257. doi:10.1111/AGE.12909
176. Vu SV, Knibb W, Gondro C, et al. Genomic prediction for whole weight, body shape, meat yield, and color traits in the Portuguese oyster *Crassostrea angulata*. *Front Genet*. 2021;12:1140. doi:10.3389/FGENE.2021.661276/BIBTEX
177. Vu SV, Gondro C, Nguyen NTH, et al. Prediction accuracies of genomic selection for nine commercially important traits in the Portuguese oyster (*Crassostrea angulata*) using DArT-Seq technology. *Gene*. 2021;12(2):210. doi:10.3390/GENES12020210
178. Wang Q, Yu Y, Yuan J, et al. Effects of marker density and population structure on the genomic prediction accuracy for growth trait in Pacific white shrimp *Litopenaeus vannamei*. *BMC Genet*. 2017;18(1):45. doi:10.1186/s12863-017-0507-5
179. Wang Q, Yu Y, Zhang Q, et al. Evaluation on the genomic selection in *Litopenaeus vannamei* for the resistance against *Vibrio parahaemolyticus*. *Aquaculture*. 2019;505:212-216. doi:10.1016/J.AQUACULTURE.2019.02.055
180. Lillehammer M, Bangera R, Salazar M, et al. Genomic selection for white spot syndrome virus resistance in whiteleg shrimp boosts survival under an experimental challenge test. *Sci Rep*. 2020;10(1):20571. doi:10.1038/s41598-020-77580-3
181. Nguyen NH, Phuthaworn C, Knibb W. Genomic prediction for disease resistance to Hepatopancreatic parvovirus and growth, carcass and quality traits in Banana shrimp *Fenneropenaeus merguiensis*. *Genomics*. 2020;112(2):2021-2027. doi:10.1016/J.YGENO.2019.11.014
182. VanRaden PM, Van Tassell CP, Wiggans GR, et al. Invited review: reliability of genomic predictions for north American Holstein bulls. *J Dairy Sci*. 2009;92(1):16-24. doi:10.3168/JDS.2008-1514
183. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. Invited review: genomic selection in dairy cattle: progress and challenges. *J Dairy Sci*. 2009;92(2):433-443. doi:10.3168/jds.2008-1646
184. Karaman E, Lund MS, Su G. Multi-trait single-step genomic prediction accounting for heterogeneous (co)variances over the genome. *Heredity*. 2019;124(2):274-287. doi:10.1038/s41437-019-0273-4
185. Yoshida GM, Barria A, Correa K, et al. Genome-wide patterns of population structure and linkage disequilibrium in farmed *Nile tilapia* (*Oreochromis niloticus*). *Front Genet*. 2019;10:745. doi:10.3389/fgene.2019.00745
186. Yoshida GM, Carvalheiro R, Rodríguez FH, Lhorente JP, Yáñez JM. Single-step genomic evaluation improves accuracy of breeding value predictions for resistance to infectious pancreatic necrosis virus in rainbow trout. *Genomics*. 2019;111(2):127-132. doi:10.1016/j.ygeno.2018.01.008
187. Nirea KG, Sonesson AK, Woolliams JA, Meuwissen TH. Strategies for implementing genomic selection in family-based aquaculture breeding schemes: double haploid sib test populations. *Genet Sel Evol*. 2012;44(1):1-9. doi:10.1186/1297-9686-44-30/TABLES/4
188. Daetwyler HD, Villanueva B, Woolliams JA. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS One*. 2008;3(10):e3395. doi:10.1371/JOURNAL.PONE.0003395
189. Goddard M. Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica*. 2009;136:245-257. doi:10.1007/s10709-008-9308-0
190. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J Dairy Sci*. 2010;93(2):743-752. doi:10.3168/JDS.2009-2730
191. Daetwyler HD, Pong-Wong R, Villanueva B, Woolliams JA. The impact of genetic architecture on genome-wide evaluation methods. *Genetics*. 2010;185(3):1021-1031. doi:10.1534/GENETICS.110.116855
192. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci*. 2008;91(11):4414-4423. doi:10.3168/JDS.2007-0980
193. Garrick DJ, Fernando RL. Implementing a QTL detection study (GWAS) using genomic prediction methodology. In: Gondro C, van der Werf J, Hayes B, eds. *Genome-Wide Association Studies and Genomic Prediction*. Vol 1019. Humana Press; 2013:275-298. doi:10.1007/978-1-62703-447-0_11
194. Wang H, Misztal I, Aguilar I, Legarra A, Muir WM. Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet Res*. 2012;94(2):73-83. doi:10.1017/S0016672312000274
195. Misztal I, Legarra A. Invited review: efficient computation strategies in genomic selection. *Animal*. 2017;11(5):731-736. doi:10.1017/S1751731116002366
196. Houston RD. Future directions in breeding for disease resistance in aquaculture species. *Rev Bras Zootec*. 2017;46(6):545-551. doi:10.1590/S1806-92902017000600010
197. Dufflocq P, Pérez-Enciso M, Lhorente JP, Yáñez JM. Accuracy of genomic predictions using different imputation error rates in aquaculture breeding programs: a simulation study. *Aquaculture*. 2019;503:225-230. doi:10.1016/j.aquaculture.2018.12.061
198. Hickey JM, Kinghorn BP, Tier B, Wilson JF, Dunstan N, van der Werf JH. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. *Genet Sel Evol*. 2011;43(1):1-13. doi:10.1186/1297-9686-43-12
199. Sargolzaei M, Chesnais JP, Schenkel FS. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*. 2014;15(1):1-12. doi:10.1186/1471-2164-15-478
200. López ME, Neira R, Yáñez JM. Applications in the search for genomic selection signatures in fish. *Front Genet*. 2015;5:458. doi:10.3389/FGENE.2014.00458/BIBTEX
201. Ma Y, Ding X, Qanbari S, Weigend S, Zhang Q, Simianer H. Properties of different selection signature statistics and a new strategy for combining them. *Heredity*. 2015;115(5):426-436. doi:10.1038/hdy.2015.42
202. Pérez O'Brien AM, Utsunomiya YT, Mészáros G, et al. Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. *Genet Sel Evol*. 2014;46(1):1-14. doi:10.1186/1297-9686-46-19/TABLES/4
203. Karlsson S, Moen T, Lien S, Glover KA, Hindar K. Generic genetic differences between farmed and wild Atlantic salmon identified

from a 7K SNP-chip. *Mol Ecol Resour*. 2011;11(Suppl 1):247-253. doi:10.1111/J.1755-0998.2010.02959.X
204. Mäkinen H, Vasemägi A, Mcginnity P, Cross TF, Primmer CR. Population genomic analyses of early-phase Atlantic Salmon (*Salmo salar*) domestication/captive breeding. *Evol Appl*. 2015;8(1):93-107. doi:10.1111/EVA.12230
205. Gutierrez AP, Yáñez JM, Davidson WS. Evidence of recent signatures of selection during domestication in an Atlantic salmon population. *Mar Genomics*. 2016;26:41-50. doi:10.1016/J.MARGEN.2015.12.007
206. Liu L, Ang KP, Elliott JAK, et al. A genome scan for selection signatures comparing farmed Atlantic salmon with two wild populations: testing colocalization among outlier markers, candidate genes, and quantitative trait loci for production traits. *Evol Appl*. 2017;10(3):276-296. doi:10.1111/EVA.12450
207. López ME, Benestan L, Moore JS, et al. Comparing genomic signatures of domestication in two Atlantic salmon (*Salmo salar* L.) populations with different geographical origins. *Evol Appl*. 2019;12(1):137-156. doi:10.1111/EVA.12689
208. López ME, Linderoth T, Norris A, Lhorente JP, Neira R, Yáñez JM. Multiple selection signatures in farmed Atlantic Salmon adapted to different environments across hemispheres. *Front Genet*. 2019;10:901. doi:10.3389/FGENE.2019.00901
209. Bertolotti AC, Layer RM, Gundappa MK, et al. The structural variation landscape in 492 Atlantic salmon genomes. *Nat Commun*. 2020;11(1):1-16. doi:10.1038/s41467-020-18972-x
210. López ME, Cádiz MI, Rondeau EB, Koop BF, Yáñez JM. Detection of selection signatures in farmed coho salmon (*Oncorhynchus kisutch*) using dense genome-wide information. *Sci Rep*. 2021;11(1):1-13. doi:10.1038/s41598-021-86154-w
211. Cádiz MI, López ME, Díaz-Domínguez D, et al. Detection of selection signatures in the genome of a farmed population of anadromous rainbow trout (*Oncorhynchus mykiss*). *Genomics*. 2021;113(5):3395-3404. doi:10.1016/J.YGENO.2021.07.027
212. Lim C, Webster CD. *Tilapia: Biology, Culture, and Nutrition*. Food Products Press; 2006.
213. Fitzsimmons K. Tilapia: the most important aquaculture species of the 21st century. Paper presented at Proceedings from the Fifth International Symposium on Tilapia Aquaculture; 1999:8.
214. Hong Xia J, Bai Z, Meng Z, et al. Signatures of selection in tilapia revealed by whole genome resequencing. *Sci Rep*. 2015;5(1):1-10. doi:10.1038/srep14168
215. Schomburg D, Michal G. *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*. 2nd ed. John Wiley & Sons; 2012.
216. Cádiz MI, López ME, Díaz-Domínguez D, et al. Whole genome resequencing reveals recent signatures of selection in three strains of farmed *Nile tilapia* (*Oreochromis niloticus*). *Sci Rep*. 2020;10(1):1-14. doi:10.1038/s41598-020-68064-5
217. Yu X, Setyawan P, Bastiaansen JWM, et al. Genomic analysis of a *Nile tilapia* strain selected for salinity tolerance shows signatures of selection and hybridization with blue tilapia (*Oreochromis aureus*). *Aquaculture*. 2022;560:738527. doi:10.1016/J.AQUACULTURE.2022.738527
218. Chen X, Zhong L, Bian C, et al. High-quality genome assembly of channel catfish, *Ictalurus punctatus*. *Gigascience*. 2016;5(1):1-4. doi:10.1186/S13742-016-0142-5/2737418
219. Fine ML, Lahiri S, Sullivan ADH, Mayo M, Newton SH, Sismour EN. Reduction of the pectoral spine and girdle IN domesticated channel catfish is likely caused by changes IN selection pressure. *Evolution*. 2014;68(7):2102-2107. doi:10.1111/EVO.12379
220. Dunham R. *Aquaculture and Fisheries Biotechnology: Genetic Approaches*. 1st ed. CABI; 2011.
221. Sun L, Liu S, Wang R, et al. Identification and analysis of genome-wide SNPs provide insight into signatures of selection and domestication in channel catfish (*Ictalurus punctatus*). *PLoS One*. 2014;9(10):e109666. doi:10.1371/JOURNAL.PONE.0109666
222. Gui J, Tang Q, Li Z, Liu J, De Silva SS. *Aquaculture in China: Success Stories and Modern Trends*. Vol 1. 1st ed. Wiley; 2018.
223. Shen Y, Wang L, Fu J, Xu X, Yue GH, Li J. Population structure, demographic history and local adaptation of the grass carp. *BMC Genomics*. 2019;20(1):1-16. doi:10.1186/S12864-019-5872-1/TABLES/4
224. Dawood MAO, Koshio S, Ishikawa M, et al. Dietary supplementation of β-glucan improves growth performance, the innate immune response and stress resistance of red sea bream, *Pagrus major*. *Aquacult Nutr*. 2017;23(1):148-159. doi:10.1111/ANU.12376
225. Murata O, Harada T, Miyashita S, et al. Selective breeding for growth in Red Sea bream. *Fish Sci*. 1996;62(6):845-849. doi:10.2331/FISHSCI.62.845
226. Noh CH, Hong KP, Oh S-Y, et al. Comparative growth performance of the selected and the non-selected Red Sea bream (*Pagrus major*) lines. *Korean J Fish Aquat Sci*. 2004;37(5):400-404. doi:10.5657/KFAS.2004.37.5.400
227. Nam BH, Yoo DA, Kim YO, et al. Whole genome sequencing reveals the impact of recent artificial selection on red sea bream reared in fish farms. *Sci Rep*. 2019;9(1):1-9. doi:10.1038/s41598-019-42988-z
228. Fleming IA, Agustsson T, Finstad B, Johnsson JI, Björnsson BT. Effects of domestication on growth physiology and endocrinology of Atlantic salmon (*Salmo salar*). *Can J Fish Aquat Sci*. 2002;59(8):1323-1330. doi:10.1139/F02-082
229. Frantz LAF, Schraiber JG, Madsen O, et al. Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nat Genet*. 2015;47(10):1141-1148. doi:10.1038/ng.3394
230. Johnsson JI, Petersson E, Jönsson E, Björnsson BT, Järvi T. Domestication and growth hormone alter antipredator behaviour and growth patterns in juvenile brown trout, *Salmo trutta*. *Can J Fish Aquat Sci*. 1996;53(7):1546-1554. doi:10.1139/F96-090/ASSET/F96-090.FP.PNG_V03
231. Robison BD, Rowland W. A potential model system for studying the genetics of domestication: behavioral variation among wild and domesticated strains of zebra danio (*Danio rerio*). *Can J Fish Aquat Sci*. 2005;62(9):2046-2054. doi:10.1139/F05-118
232. Agudelo JFG, Mastrochirico-Filho VA, Borges CHS, et al. Genomic selection signatures in farmed *Colossoma macropomum* from tropical and subtropical regions in South America. *Evol Appl*. 2022;15(4):679-693. doi:10.1111/EVA.13351
233. Sandoval-Castillo J, Robinson NA, Hart AM, Strain LWS, Beheregaray LB. Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (*Haliotis laevigata*), along a longitudinal environmental gradient. *Mol Ecol*. 2018;27(7):1603-1620. doi:10.1111/MEC.14526
234. Lv J, Cai Y, Liu P, et al. Genomic differentiation and selection signatures of two elite varieties of yesso scallop *Mizuhopecten yessoensis*. *Aquaculture*. 2022;550:737842. doi:10.1016/J.AQUACULTURE.2021.737842
235. Ni G, Cavero D, Fangmann A, Erbe M, Simianer H. Whole-genome sequence-based genomic prediction in laying chickens with different genomic relationship matrices to account for genetic architecture. *Genet Sel Evol*. 2017;49(1):1-14. doi:10.1186/S12711-016-0277-Y/FIGURES/7
236. Ober U, Ayroles JF, Stone EA, et al. Using whole-genome sequence data to predict quantitative trait phenotypes in *Drosophila melanogaster*. *PLoS Genet*. 2012;8(5):e1002685. doi:10.1371/JOURNAL.PGEN.1002685
237. van Binsbergen R, Bink MC, Calus MP, et al. Accuracy of imputation to whole-genome sequence data in Holstein Friesian cattle. *Genet Sel Evol*. 2014;46(1):1-13. doi:10.1186/1297-9686-46-41

238. Pérez-Enciso M, Rincón JC, Legarra A. Sequence- vs. chip-assisted genomic selection: accurate biological information is advised. *Genet Sel Evol*. 2015;47(1):1-14. doi:10.1186/s12711-015-0117-5
239. Heidaritabar M, Calus MPL, Megens HJ, Vereijken A, Groenen MAM, Bastiaansen JWM. Accuracy of genomic prediction using imputed whole-genome sequence data in white layers. *J Anim Breed Genet*. 2016;133(3):167-179. doi:10.1111/JBG.12199
240. Brøndum RF, Su G, Janss L, et al. Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. *J Dairy Sci*. 2015;98(6):4107-4116. doi:10.3168/JDS.2014-9005
241. Hayes B, MacLeod IM, Daetwyler HD, et al. Genomic prediction from whole genome sequence in livestock: 1000 bull genomics project. Paper presented at 10th World Congress on Genetics Applied to Livestock Production (WCGALP); 2014. Accessed April 27, 2022. https://library.wur.nl/WebQuery/wurpubs/479811
242. Liu Y, Lu S, Liu F, et al. Genomic selection using BayesCπ and GBLUP for resistance against *Edwardsiella tarda* in Japanese flounder (*Paralichthys olivaceus*). *Marine Biotechnol*. 2018;20(5):559-565. doi:10.1007/S10126-018-9839-Z/TABLES/4
243. Boison S, Ding J, Leder E, et al. QTLs associated with resistance to cardiomyopathy syndrome in Atlantic Salmon. *J Hered*. 2019;110(6):727-737. doi:10.1093/JHERED/ESZ042
244. Baranski M, Moen T, Våge D. Mapping of quantitative trait loci for flesh colour and growth traits in Atlantic salmon (*Salmo salar*). *Genet Sel Evol*. 2010;42(1):1-14. doi:10.1186/1297-9686-42-17/TABLES/7
245. Lien S, Koop BF, Sandve SR, et al. The Atlantic salmon genome provides insights into rediploidization. *Nature*. 2016;533(7602):200-205. doi:10.1038/nature17164
246. Yáñez JM, Naswa S, López ME, et al. Genomewide single nucleotide polymorphism discovery in Atlantic salmon (*Salmo salar*): validation in wild and farmed American and European populations. *Mol Ecol Resour*. 2016;16(4):1002-1011. doi:10.1111/1755-0998.12503
247. Houston RD, Taggart JB, Cézard T, et al. Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). *BMC Genomics*. 2014;15(1):1-13. doi:10.1186/1471-2164-15-90/FIGURES/5
248. Blix TB, Dalmo RA, Wargelius A, Myhr AI. Genome editing on finfish: current status and implications for sustainability. *Rev Aquac*. 2021;13(4):2344-2363. doi:10.1111/RAQ.12571
249. Hallerman E. Genome editing in cultured fishes. *CABI Agric Biosci*. 2021;2(1):1-19. doi:10.1186/S43170-021-00066-3
250. Palti Y, Gao G, Liu S, et al. The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout. *Mol Ecol Resour*. 2015;15(3):662-672. doi:10.1111/1755-0998.12337
251. Xu J, Zhao Z, Zhang X, et al. Development and evaluation of the first high-throughput SNP array for common carp (*Cyprinus carpio*). *BMC Genomics*. 2014;15(1):1-10. doi:10.1186/1471-2164-15-307
252. Liu S, Sun L, Li Y, et al. Development of the catfish 250K SNP array for genome-wide association studies. *BMC Res Notes*. 2014;7(1):1-12. doi:10.1186/1756-0500-7-135
253. Zeng Q. Development of 690K SNP arrays for whole genome mapping and genetic studies in catfish; 2016. Accessed April 28, 2022. http://etd.auburn.edu/handle/10415/5374
254. Joshi R, Árnyasi M, Lien S, Gjøen HM, Alvarez AT, Kent M. Development and validation of 58K SNP-array and high-density linkage map in Nile tilapia (*O. niloticus*). *Front Genet*. 2018;9(OCT):472. doi:10.3389/FGENE.2018.00472
255. Yáñez JM, Yoshida G, Barria A, et al. High-throughput single nucleotide polymorphism (SNP) discovery and validation through whole-genome resequencing in Nile tilapia (*Oreochromis niloticus*). *Mar Biotechnol*. 2020;22(1):109-117. doi:10.1007/s10126-019-09935-5
256. Mastrochirico-Filho VA, Ariede RB, Freitas MV, et al. Development of a multi-species SNP array for serrasalmid fish *Colossoma macropomum* and *Piaractus mesopotamicus*. *Sci Rep*. 2021;11(1):1-11. doi:10.1038/s41598-021-98885-x
257. Zhou T, Chen B, Ke Q, et al. Development and evaluation of a high-throughput single-nucleotide polymorphism array for large yellow croaker (*Larimichthys crocea*). *Front Genet*. 2020;11:1297. doi:10.3389/FGENE.2020.571751
258. Zhou Q, Chen YD, Lu S, et al. Development of a 50k SNP array for japanese flounder and its application in genomic selection for disease resistance. *Engineering*. 2021;7(3):406-411. doi:10.1016/J.ENG.2020.06.017
259. Qi H, Song K, Li C, et al. Construction and evaluation of a high-density SNP array for the Pacific oyster (*Crassostrea gigas*). *PLoS One*. 2017;12(3):e0174007. doi:10.1371/JOURNAL.PONE.0174007
260. Vendrami DLJ, Houston RD, Gharbi K, et al. Detailed insights into pan-European population structure and inbreeding in wild and hatchery Pacific oysters (*Crassostrea gigas*) revealed by genome-wide SNP data. *Evol Appl*. 2019;12(3):519-534. doi:10.1111/EVA.12736
261. Jones DB, Jerry DR, Khatkar MS, et al. A comparative integrated gene-based linkage and locus ordering by linkage disequilibrium map for the Pacific white shrimp, *Litopenaeus vannamei*. *Sci Rep*. 2017;7(1):1-16. doi:10.1038/s41598-017-10515-7
262. Henderson CR. Best linear unbiased estimation and prediction under a selection model. *Biometrics*. 1975;31(2):423. doi:10.2307/2529430
263. Whittaker JC, Thompson R, Denham MC. Marker-assisted selection using ridge regression. *Genet Res*. 2000;75(2):249-252. doi:10.1017/S0016672399004462
264. Legarra A, Aguilar I, Misztal I. A relationship matrix including full pedigree and genomic information. *J Dairy Sci*. 2009;92(9):4656-4663. doi:10.3168/jds.2009-2061
265. Habier D, Fernando RL, Kizilkaya K, Garrick DJ. Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics*. 2011;12:186. doi:10.1186/1471-2105-12-186
266. De Los CG, Naya H, Gianola D, et al. Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics*. 2009;182(1):375-385. doi:10.1534/genetics.109.101501
267. Tsai HY, Hamilton A, Tinch AEA, et al. Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. *Genet Sel Evol*. 2016;48(1):47. doi:10.1186/s12711-016-0226-9
268. Johnston SE, Orell P, Pritchard VL, et al. Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a wild population of Atlantic salmon (*Salmo salar*). *Mol Ecol*. 2014;23(14):3452-3468. doi:10.1111/MEC.12832
269. Al-Tobasei R, Ali A, Garcia ALS, Lourenco D, Leeds T, Salem M. Genomic predictions for fillet yield and firmness in rainbow trout using reduced-density SNP panels. *BMC Genomics*. 2021;22(1):1-11. doi:10.1186/S12864-021-07404-9
270. Blay C, Haffray P, D'Ambrosio J, et al. Genetic architecture and genomic selection of fatty acid composition predicted by Raman spectroscopy in rainbow trout. *BMC Genomics* 2021;22(1):1-19. doi:10.1186/S12864-021-08062-7
271. Vallejo RL, Leeds TD, Fragomeni BO, et al. Evaluation of genome-enabled selection for bacterial cold water disease resistance using progeny performance data in rainbow trout: Insights on genotyping methods and genomic prediction models. *Front Genet*. 2016;7(MAY):96. doi:10.3389/FGENE.2016.00096
272. Silva RMO, Evenhuis JP, Vallejo RL, et al. Whole-genome mapping of quantitative trait loci and accuracy of genomic predictions for resistance to columnaris disease in two rainbow trout breeding populations. *Genet Sel Evol*. 2019;51(1):1-13. doi:10.1186/S12711-019-0484-4
273. Vallejo RL, Cheng H, Fragomeni BO, et al. The accuracy of genomic predictions for bacterial cold water disease resistance remains higher than the pedigree-based model one generation after model training in

a commercial rainbow trout breeding population. *Aquaculture*. 2021; 545:737164. doi:10.1016/J.AQUACULTURE.2021.737164

274. Joshi R, Almeida DB, da Costa AR, et al. Genomic selection for resistance to Francisellosis in commercial Nile tilapia population: genetic and genomic parameters, correlation with growth rate and predictive ability. *Aquaculture*. 2021;537:736515. doi:10.1016/J.AQUACULTURE.2021.736515

How to cite this article: Yáñez JM, Barría A, López ME, et al. Genome-wide association and genomic selection in aquaculture. *Rev Aquac*. 2023;15(2):645-675. doi:10.1111/raq.12750