Exploiting genomics to fast-track genetic improvement in aquaculture

Ross D. Houston[1], Tim P. Bean[1], Daniel J. Macqueen[1], Manu Kumar Gundappa[1], Ye Hwa Jin[1], Tom L. Jenkins[2], Sarah-Louise C. Selly[3], Samuel A.M. Martin[4], Jamie R. Stevens[2], Eduarda M. Santos[2], Andrew Davie[3] and Diego Robledo[1]*

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

[2] Sustainable Aquaculture Futures, Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter, UK

[3] Institute of Aquaculture, University of Stirling, Stirling, UK

[4] School of Biological Sciences, University of Aberdeen, Aberdeen, UK

* e-mail: ross.houston@roslin.ed.ac.uk

Abstract | Aquaculture is the fastest growing farmed food sector and will soon become the primary source of fish and shellfish for human diets. However, in contrast to crops and livestock, production is derived from a numerous and exceptionally diverse set of species, typically in the early stages of domestication. Genetic improvement via well-designed and managed breeding programs has major potential to help meet the rising seafood demand driven by human population growth. Genomics and biotechnology are increasingly applied from an early stage, enabled by continuous advances in sequencing and bioinformatics. This review highlights the application of these technologies across the broad range of aquaculture species and stages of domestication, and explores the potential to combine genomics with the amenable reproductive biology of most aquatic species in order to expedite domestication and genetic improvement.

Introduction

Aquaculture plays a crucial and rapidly increasing role in food security and economic stability worldwide. Over 90 % of global aquaculture takes place in low and middle income countries, where it makes major contributions to the Sustainable Development Goals of the United Nations, either directly through human consumption or indirectly through economic growth[1]. Global production of finfish and shellfish reached 172.6 million tons in 2017, approximately half of which is currently derived from aquaculture[2]. Capture fisheries are placing serious pressures on wild stocks, with minimal scope for sustainable expansion[3]. In contrast, aquaculture is the fastest growing food production sector globally, and with major limitations on wild capture and terrestrial farmland exploitation[4], its future importance as a source of affordable and nutritious animal protein for human diets is clear. However, intensification of aquaculture production poses environmental concerns such as habitat destruction[5] and infectious disease outbreaks, which negatively impact the health and welfare of farmed, and potentially wild populations[6], and may be exacerbated by climate change[7].

Genetic improvement has major potential to improve the efficiency and reduce the environmental footprint of aquaculture. However, in contrast to the terrestrial livestock and crop sectors, aquaculture comprises a hugely diverse group of finfish and shellfish species (Fig. 1A), comprising an estimated 543 different animal species (362 finfish, 104 molluscs, 62 crustaceans, 9 other aquatic invertebrates, and 6 frogs and reptiles; plants and algae being outside of the scope of this review). Farming of approximately 70 of these species underpins 80 % of the world's aquaculture production volume, compared to just three livestock species (pig, chicken, cow) underpinning 80 % of global meat production (Fig. 1B and Supplementary Tables 5-6), and four species (rice, wheat, maize, potatoes) underlying two thirds of crop production[8]. Despite their diversity, aquaculture species tend to share two key features which enhance their potential for genetic improvement. Firstly they remain in the early stages of the domestication process[9] (Fig. 1B), and secondly they are highly fecund with typically external fertilization. These features of their reproductive biology enable flexibility in

breeding program design and widespread dissemination of selectively bred strains to producers often without the need for several multiplication layers. Therefore, there is a pressing opportunity to utilize domestication and selective breeding programs to harness the as-yet largely untapped genetic potential of farmed aquatic species[10], as highlighted in a recent landmark report by the FAO[10]. This potential for cumulative and permanent improvement of production traits is evident from the typically high genetic gains in aquaculture breeding programs, for example an average of 13 % growth increase per generation in Atlantic salmon (*Salmo salar*)[11]; substantially higher than observed in terrestrial livestock species breeding programs[12,13].

Genomic tools are hugely valuable to inform sustainable genetic improvement[14], and their affordability and accessibility now means they can be applied at all stages of the domestication and genetic improvement continuum, from informing the choice of base populations through to advanced genomic selection in closed commercial breeding nuclei (Box 1). Furthermore, they can be applied to characterize, utilize, and conserve wild aquatic genetic resources, and inform the management of interaction between farmed and wild aquatic animals throughout this continuum. This review provides an overview of the status of domestication and selective breeding in aquaculture species, highlights how tailored application of genomic tools can expedite sustainable genetic improvement in diverse species and environments, and explores the potential of emerging genomic and biotechnology techniques such as genome editing or surrogate broodstock technologies to make further step-improvements in aquaculture breeding and production.

Genomics-enabled domestication in aquaculture

The recent and rapid domestication of aquaculture species: Domestication in the context of this review is considered as the process of moving from an exclusive reliance on wild broodstock to the use of closed breeding populations together with selective breeding programs for genetic improvement of production traits. For certain major aquaculture species [e.g. carp (*Cyprinidae* spp.) and tilapia (*Cichlidae* spp.)], aquaculture and domestication has been ongoing in some form for millennia[15], but selective breeding

programs to enable genetic improvement are much more recent ([16]; Fig. 1B). Currently only a minority of aquaculture production is derived from selectively-bred stocks; estimated at approximately 10 % eight years ago[17]. However, this is increasing rapidly, particularly for species with high production volume and value, with approximately 75 % of top 10 finfish, crustacean, and mollusc species (by production volume) benefitting from some form of modern selective breeding program (Supplementary table 3). The availability and exploitation of selective breeding depends on the local environmental, social, political and economic landscape, all of which can present major hurdles, especially in low and middle income countries[18]. As such, the use of genetic technologies varies dramatically by continent with > 80 % of European aquaculture production derived from selective breeding programs[19]. These programs enable cumulative, permanent, and sustainable genetic gain for target production traits[12,20], and are fundamental to scale-up production in the context of finite resources[10].

Moving towards genetic improvement via selective breeding requires progression along the 'levels of domestication' scale[21], which reflects the degree of control humans have over the lifecycle of the farmed species. While the number and diversity of aquaculture species presents a challenge to this process, new husbandry techniques linked to improved understanding of reproductive biology and larval rearing will help overcome these challenges. Historically, the selection of species amenable to reproduction in farmed environments has been key in defining which livestock and aquaculture species were farmed. For example, domesticated species tend to display behavioral plasticity that enables them to adapt to a range of captive environments[22,23]. A key difference between livestock and aquaculture species is that domestication of terrestrial livestock occurred in tandem with global human migration several millennia prior to informed management of breeding populations, and modern livestock lines have typically undergone multiple major genetic bottlenecks[9]. In contrast, the time lag between domestication and selective breeding is considerably shorter in aquaculture species, with both occurring in tandem in many cases. Consequently, genomic tools

94 can be used from the outset to inform, optimize, and expedite the two processes (Box 1), and provide a
95 more detailed understanding of their impact on the species' genomes and physiology.

96 ***The burgeoning genomic toolbox:*** Genomic resources for aquaculture generally lag behind terrestrial
97 livestock, in particular for sequencing and assembly of reference genomes (Box 2). In part, this reflects the
98 tradtionally challenging nature of genome assembly in non-mammalian and non-avian species, particularly
99 for aquatic species with complex genomic features (Box 2), such as recent whole genome duplication in
100 salmonid species and the exceptionally high heterozygosity observed in bivalve species[33]. However, these
101 genomic features may underlie adaptive capacity and phenotypic plasticity in production environments, and
102 may contribute to the genetic regulation of production-relevant traits. Fortunately, long-read sequencing
103 (such as Nanopore and PacBio) and improved scaffolding technologies (such as Hi-C and Optical mapping)
104 now provide the toolbox to rapidly and cost-effectively generate contiguous reference genomes in
105 established and new aquaculture species (Box 2). Therefore, sequencing a target species' genome has
106 become within the reach of individual laboratories, and no longer requires the degree of coordinated effort
107 and funding that led to the first farmed animal species' reference genome assemblies, including Atlantic
108 salmon[34]. However, standardisation and coordination of multiple assemblies, including population- or
109 'breed'-specific assemblies, and their functional annotation remains a challenge for which international
110 coordination and community-led initiatives are required to meet (Box 2). A key component of the genomic
111 toolbox to inform domestication and selective breeding is genotyping technologies. Single nucleotide
112 polymorphism (SNP) array platforms have been created for many high value aquaculture species (Box 2),
113 and genotyping by sequencing (GBS) techniques including Restriction site-Associated DNA sequencing (RAD-
114 Seq[30]) and derivatives have been applied in many species to obtain population-level SNP data without
115 significant prior investment, and without the immediate need for a reference genome[31,32].

116 ***Genomics to inform domestication and formative breeding programs:*** GBS is a mainstay of studying the
117 genetics of new aquaculture species, and has been applied to support domestication and formation of

breeding programs. The establishment and management of genetically diverse base populations is essential to this process, as it underlies the future genetic potential to be exploited via selective breeding[24]. Poor broodstock management and hatchery practices leading to inbreeding depression has been hypothesised to result in reduced population fitness, increased susceptibility to stress and disease, and ultimately 'boom-and-bust' production cycles[25,26]. Tailored use of genomic tools can be applied at each stage of the domestication and selective breeding continuum to inform and optimise the process (Box 1). However, it should be noted that the reliability of genomic data alone to predict adaptive potential of populations is questionable[27], and they should be used as a complement to phenotypic evaluations of stocks. These evaluations may include diallelic crosses between strains in multiple environments, which can inform on the importance of heterosis and genotype by environment interaction (discussed in more detail below[28]). However, while hybrid vigour resulting from strain crosses can result in significant one off gains in production, and genomic tools can be used to investigate the underlying molecular mechanisms of this heterosis[29], exploiting additive genetic variation via within-strain breeding programs is likely to result in superior performance after a small number of generations of selection[28].

An example of genomics-enabled domestication of a new target species is the Australasian snapper (*Pagrus auratus*) in New Zealand. Rapid generation of *de novo* genome maps[36], transcriptomes[37], genotyping by sequencing methods[36,38], and estimation of genetic diversity and genetic parameters[38] were applied to inform selection of base populations, retention of genetic diversity during domestication, and investigations into the biology of production traits. Likewise, the recent widespread use of cleanerfish (e.g. Ballan wrasse, *Labrus bergylta*, and lumpfish, *Cyclopterus lumpus*) for co-culture with Atlantic salmon farming to help tackle sea lice (*Lepeophtheirus salmonis* and *Caligus rogercresseyi*) has led to expedited, genomics-enabled domestication and breeding of lumpfish (Box 1). These cases are early examples of how genomics technology has rapidly become accessible and should be applied from the outset to inform domestication

and subsequent genetic improvement, with a roadmap for appropriate genomics tools matched to the different stages given in Box 1.

Genomics tools are valuable to tackle species-specific breeding and production issues related to the highly diverse biology of aquaculture species. For example, a key component of the domestication-genetic improvement continuum in aquaculture species is an early understanding of sex determination, where a diverse array of genetic and non-genetic systems have been described[39]. These can vary within genus and even within species, and sequential hermaphroditism presents an additional challenge in several commercially important aquaculture species[40]. The GBS techniques described above and in Box 1 have been widely applied to assess the genetic basis of sex determination[41], for example in Nile tilapia (*Oreochromis niloticus*)[42], Atlantic halibut (*Hippoglossus hippoglossus*)[43], European seabass (*Dicentrarchus labrax*)[44], and mud crabs (Scylla sp.)[45]. The genetic markers identified in these studies can be applied to predict sex of juveniles and control sex ratio in both broodstock and production animals, thereby playing a key role in domestication and genetic improvement. An additional species-specific reproductive challenge is mass spawning, which is a feature of several marine aquaculture species, such as gilthead sea bream and barramundi. This causes practical challenges such as uneven parental contribution and difficulty in tracking individual pedigrees, which can result in inbreeding[46]. While multiple interventions are possible to enable pedigree tracking (e.g. pair spawning or stripping using hormonal induction)[47] genetic markers are frequently applied to track stock relatedness is to minimize loss of genetic diversity within a closed breeding nucleus[46]. These examples highlight the advantages of harnessing the increased availability and affordability of genomic technologies to optimize each stage of the domestication process, as species transition from wild-sourced broodstock towards closed breeding programs for targeted genetic improvement.

Genomics to accelerate genetic gain in aquaculture breeding

The establishment of well-managed selective breeding programs for aquaculture based on recording of pedigree and routine measurements of traits has been successful in increasing production of several

165 species[11]. Just as genomic tools are applied to inform and optimize domestication, they can improve
166 selective breeding in several ways, including maximizing genetic gain and minimizing inbreeding[14].

167 ***Major effect loci in recently domesticated populations***. A key factor in defining the optimal use of genomic
168 tools is the genetic architecture of production traits in the breeding goal; i.e. whether genetic variation in
169 target traits is underpinned by few major-effect loci or (as is typically the case in farmed animal
170 populations[13]) many loci of minor effect. Farmed aquatic populations face selection pressures that are vastly
171 different to their wild counterparts. Due to the recent and ongoing domestication processs, previously
172 neutral alleles in wild populations may be beneficial for production phenotypes, and these will remain
173 amongst the standing genetic variation in aquaculture populations. During the millenia of domestication of
174 terrestrial livestock, such loci are likely to already be fixed via soft sweeps, but in aquaculture species they
175 may present a one-off opportunity for rapid genetic improvement via marker-assisted selection (MAS). A
176 well-known example is the major quantitative trait locus (QTL) affecting resistance to Infectious Pancreatic
177 Necrosis Virus (IPNV) in Atlantic salmon, for which rapid uptake of MAS by the industry had a major role in
178 preventing outbreaks of this disease (Box 3). Other applications of QTL for disease resistance include
179 breeding of a Japanese flounder (*Paralichthys olivaceus*) strain with resistance to lymphocystis disease
180 (LD)[48], based on a major QTL for LD resistance[49], and use of MAS based on QTL affecting resistance to
181 bacterial cold water disease in rainbow trout (*Oncorhynchus mykiss*)[50]. Other noteworty examples of major
182 effect loci in salmon include *vgll3*, which controls the timing of sexual maturation and explains 30 – 40 % of
183 the phenotypic variation in age at maturity[51,52], as well as loci for resistance to pancreas disease[53], and
184 cardiomyopathy syndrome[54,55]. Similarly, in Nile tilapia, a locus explaining 79 % of the phenotypic variation
185 in salinity tolerance was detected[56], although validation of the size of effect in independent populations is
186 required to make generalised conclusions about this trait. The list of loci of major effect will presumably
187 increase in the near future as genomics is increasingly used to study traits of interest to aquaculture in
188 additional species and populations. While MAS has had limited success in terrestrial livestock, its use within

189 aquaculture populations at the early stages of domestication can provide rare but striking examples which
190 help to highlight the value of genetic improvement to the industry (Box 3).

191 ***Genomic selection is transformative for aquaculture breeding:*** Genome-wide association studies in
192 aquaculture species highlight that most traits of relevance to production are polygenic in nature[57,58]. For
193 genetic improvement of such traits, routine trait measurement and tracking of relationships between
194 individual animals in a breeding population is required[59]. The availability of large full-sibling families gives
195 both power and flexibility to a breeding program design, for example allowing the routine testing of full-
196 siblings of the selection candidates (sib-testing) for traits that are practically challenging or impossible to
197 measure on the selection candidates themselves, such as disease resistance (Fig. 2). However, in sib-testing,
198 selection candidates from a given family have the same estimated breeding value, placing limitations on
199 genetic gain that can be achieved while maintaining genetic diversity. Genetic marker data are required to
200 accurately capture the within-family (or Mendelian sampling) component of genetic variation for such traits.
201 Genomic selection[60] was first tested in Atlantic salmon breeding, enabled by development of the first high
202 density SNP arrays[61,62], and demonstration of its utility to accurately predict breeding values in a typical
203 salmon breeding program setting[62,63]. Genomic selection in aquaculture breeding is based on the same
204 concept as for terrestrial livestock, with genotype and phenotype measurements taken on a reference
205 population used to train a prediction model which is then applied to genotyped selection candidates to
206 predict genomic estimated breeding values (gEBVs,[13,60].) Importantly, the high fecundity and large family
207 sizes in aquaculture species offers two major advantages. Firstly, the close relationship between the
208 reference population and the selection candidates enables high selection accuracy, even at low marker
209 density, due to long shared genomic segments between close relatives. Secondly, routine phenotyping can
210 be performed on these close relatives for different traits and in diverse environments, including 'field'
211 testing in commercial farm settings (Fig 2). In the past five years, the majority of advanced breeding
212 programs for major aquaculture species have routinely employed genomic selection[58,64], and the

developments in low cost genotyping technologies are enabling technology transfer to the smaller and more fragmented sectors.

The availability of large full-sibling families can be exploited using within-family genomic selection, with very low density markers used to estimated gEBVs within families with known pedigree-based EBVs[65]. The increased accuracy of genomic prediction compared to pedigree prediction is evident in a range of aquaculture species, with a median increase in prediction accuracy of 24 % for growth-related traits and 22 % for disease resistance traits (Table 1). These increases in prediction accuracy are relatively consistent across species and genotyping platforms, with SNP arrays primarily used in the high value species, but GBS giving equivalent findings in several other finfish, crustacean and shellfish species (Table 1). The majority of studies of genomic selection in aquaculture species use GBLUP approaches which harness genomic relationships to estimate genetic merit of individuals. A range of Bayesian models have been tested in several species, but without consistent differences in prediction accuracy compared to the simpler GBLUP approach. Adequate sample size for the genotyped and phenotyped population is key to fully assess the efficacy of genomic selection (e.g. > 1000 individuals), but the population structure is equally important as prediction accuracy is very dependent on the proximity of relationships between animals in the training and validation sets[66]. While several thousand genome-wide markers are also required, it is noteworthy that a reduction in SNP density down to only one or two thousand SNPs tends to be sufficient to achieve the asymptote of prediction accuracy where these close relationships exist[58]. However, the accuracy drops drastically as the relationship between the reference and test populations becomes more distant, as demonstrated in Atlantic salmon[67] and common carp (*Cyprinus carpio*)[68], and therefore routine trait measurement and genotyping is required each generation to retrain the genomic prediction models.

Low cost solutions for democratizing genomic selection. Capitalizing on the advantages offered by high fecundity in aquaculture breeding programs requires genotyping of thousands of animals per generation which can be prohibitively expensive. While genomic selection has become commonplace in a few highly

developed aquaculture sectors (e.g. salmonids, tilapia, shrimp), genomic tools are yet to be routinely incorporated into breeding programs for many species (Box 2; Supplementary Table 7). Hence, to translate the benefits of genomic selection to most aquaculture species, there is a clear need to develop cost-effective and species-specific tools, together with effective knowledge transfer to help democratize the technologies. Lower density SNP panels, potentially typed using targeted GBS techniques (e.g. GT-Seq,[69]) or fluorescence-based assays, tend to be cheaper than SNP arrays. Low density genotyping can be integrated with genotype imputation to increase the accuracy of genomic selection to levels approaching those obtained with high-density genotyping[70,71]. Imputation relies on genotyping only a subset of the animals at high density (in an aquaculture breeding scheme, typically the parents of the reference population and selection candidates), defining the set of haplotypes in this subset, followed by genotyping offspring at low density and imputing to high density based on those haplotypes[70]. Considering that breeding programs for many aquaculture species routinely use low density SNP panels for parentage assignment[46], combined purpose low density panels can offer the benefit of genomic selection at little added cost (and may reduce the need for physical tagging). The addition of selected functional markers linked to major QTL would add further value to combined purpose panels to enable concurrent parentage assignment, MAS and imputation-based genomic selection. Further research to develop cost-effective and pragmatic genomic selection approaches is essential to translate its benefits to aquaculture sectors with smaller margins, including in many low and middle income countries.

From sequence to consequence: identifying causative variants for target traits. Mapping and understanding the causative or functional variants impacting complex traits is a fundamental goal of biology, but also has potential additional benefits for improving rates of genetic gain in breeding either via improved selection accuracy or as targets for genome editing (Fig 3). The reduction in prediction accuracy with more distant relationships between reference and validation sets[72] is partly due to the fact that QTL are captured via linked markers rather than causative genetic variants. Research from terrestrial livestock

261 breeding hints at the potential of harnessing whole genome sequencing data[73], and incorporating weighting
262 on putative functional genomic variants (e.g. Bayes RC[74]) into genomic selection models to improve
263 accuracy, albeit improvements in prediction accuracy have been rather minor in most cases. Movement
264 towards use of whole genome sequencing of key selected individuals (e.g. parents) combined with
265 imputation to WGS based on genome-wide SNP genotypes will result in population-scale sequence data for
266 aquaculture species to allow testing of such approaches in the near future. However, the cost of WGS and
267 the effectiveness of low-density SNP panels described above means that significant improvements in
268 selection accuracy would be necessary to justify its routine use in breeding programmes.

269 The high fecundity harnessed for sib-testing is also advantageous for high-resolution genetic mapping
270 experiments, and GWAS are used to highlight genomic regions associated with traits of interest. However
271 such regions often contain hundreds to thousands of candidate causative variants and dozens of genes, and
272 most of these variants are in non-coding regions potentially impacting on transcriptional regulation.
273 Shortlisting these variants and genes to those which are more likely to be causal can be facilitated by
274 employing a pipeline of functional genomics techniques, together with knowledge of the biology of the trait
275 in question (Fig. 3). The genomic toolbox to bridge this gap is growing rapidly (Box 2), with RNA sequencing
276 routinely used for gene expression profiling, and emerging technologies being increasingly employed to
277 elucidate patterns of cytosine methylation, chromatin accessibility, histone modifications, transcriptional
278 start sites and transcript variants[75]. These tools enhance the scope to identify putative causative variants
279 within regulatory sequences (e.g. enhancers) active under specific environmental conditions (e.g. during
280 disease outbreaks). In addition, aquaculture species also benefit from existence of extant and recently
281 diverged wild counterparts, and use of comparative genomics and orthology analysis can help predict
282 functional variants based on sequence conservation[76]. Ultimately, the identification of functional variants
283 will require functional studies such as genome editing of a specific allele to assess consequences for the trait

284 of interest in cell culture and / or whole animal systems (see section '*Genome editing to accelerate genetic*
285 *improvement*' below).

286 ***Genetic improvement of disease resistance in aquaculture populations.*** Infectious disease outbreaks
287 present a major and ongoing threat to economic and environmental sustainability of aquaculture[77]. Most
288 farming occurs in open water environments, providing frequent contact with pathogens (including wild
289 reservoirs of infection), and at high stocking densities conducive to the rapid spread of infection. Outbreaks
290 of single pathogens can destroy national aquaculture industries, highlighted by outbreaks of Infectious
291 Salmon Anaemia Virus in Chile in 2007-2010[78], and annual losses of shrimp due to White Spot Syndrome
292 Virus equating to ~10 % of the global industry[79]. Options to fully mitigate such diseases via vaccination (in
293 finfish only), biosecurity, and pharmaceutical interventions are limited in aquaculture systems for several
294 reasons. Firstly, physical handling is logistically and financially challenging; secondly, the open-water nature
295 of many farming systems makes outbreaks difficult to contain; and thirdly, the early-stage of research in
296 many species means there is a paucity of vaccination and / or treatment options for diseases. The power of
297 genetic and breeding technologies to prevent or mitigate infectious diseases is increasingly recognized (e.g.
298 Box 3). Encouragingly, host resistance to most aquaculture diseases is heritable[80–82], and sibling testing
299 schemes together with genomic selection provide an effective route to breeding more resistant stocks
300 without compromising the biosecurity of the breeding nucleus (Fig. 2). Indeed, disease resistance has
301 become a major component of advanced aquaculture breeding programs[19], whereas in terrestrial livestock
302 this is limited by logistical and financial challenges relating to routine measurement of disease resistance
303 traits[83]. However, refining and optimizing collection of disease resistance data in both experimental and
304 production environments is an important goal. Disease resistance is typically measured using laboratory-
305 based pathogen challenges of pedigreed populations of animals, using outcomes such as survival or
306 pathogen burden to quantify the resistance traits[80]. However, disease outcomes in an outbreak depend on
307 several epidemiological factors, and new traits such as the propensity of an infected individual to transmit

disease have been suggested to have a genetic basis in farmed fish[84]. Benchmarking disease resistance traits measured in experimental settings with respect to outcomes in production environments is key to achieving disease prevention and control via improved genetics.

Towards accurate high-throughput phenotyping. Obtaining accurate phenotypes *en masse* is critical for any breeding program since the accuracy of trait measurement directly impacts genetic gain per generation. Phenotype measurements can be particularly challenging for aquaculture species, because manual measurements prior to harvest typically require handling large numbers of animals outside the water, presenting a logistical and financial challenge. Therefore, the ability to collect such data both directly on the selection candidates in the breeding nucleus, and on relatives of those candidate in test or production environments, can present a limitation to genetic progress in breeding programs. Computer vision technologies are being widely applied to automate plant and terrestrial livestock phenotyping, and its utility to accurately predict traits of interest has been demonstrated in several aquaculture species[58,85]. Optical sensors and machine vision systems can be used to monitor behavioural and health traits in tank or cage environments, while hyperspectral imaging approaches can inform on fillet content and characteristics[85]. For instance, the use of underwater cameras for real-time *in situ* data collection is being exploited for tasks such as sea lice monitoring in Atlantic salmon farms[86], and their use is likely to expand for more widespread data collection and phenotyping[85]. Connected mobile devices for affordable on-farm monitoring and automation of aquaculture environments (i.e. sensor technologies and the 'internet of things') has major potential for monitoring individual traits such as behavior and feed intake, in parallel to collection of huge volumes of environmental data. Transforming such data into meaningful phenotypes for breeding is a substantial challenge, and consequently data interpretation and descision tools such as machine learning and artificial intelligence will assume greater importance in aquaculture[87]. The effective combination of increasingly high resolution and high volume phenotyping in breeding nuclei, production environments, and

post-harvest together with routine genomic evaluations will lead to increasingly precise and more effective genetic improvement of aquaculture species.

The interaction between genetics and environment in aquaculture

Tackling genotype by environment interactions in aquaculture breeding. The performance and robustness of a farmed animal is dependent on the interaction between its genotype and the environment, which can vary substantially in aquaculture both within and across farms. For example, water quality presents a key challenge with limited environmental control, resulting in substantial within and across farm variation in partial pressure of CO_2, temperature, and other parameters. The transition to on-land recirculating aquaculture systems or floating closed containment systems with close control of environmental conditions is plausible for certain species[88], but the level of investment required to establish and maintain these systems is substantial and it is unlikely to be feasible for the majority of situations. As such, genetic improvement in a breeding program is intrinsically linked to the environment where traits are recorded, and genotype by environment interactions (G x E) commonly result in genotype re-ranking such that the best-performing genotypes in one environment are not the best in another, placing a limitation to realizing genetic gain in breeding programs[89,90]. The extent and nature of the G x E depends on the trait in question, and can be quantified by measuring the genetic correlation between the trait in different environments. Studies across multiple aquaculture species have highlighted that such correlations tend to be positive, but only moderate in magnitude for growth and survival traits[89], highlighting the need to account for G x E in aquaculture breeding programs.

The domestication and genetic improvement of local strains and species, which may be better adapted to the local environment, is one route to reducing the impact of G x E. However, well-managed breeding programs are expensive, and as such the current trend is consolidation into large and high-tech programs that harness high fecundity (often including multiplication layers) to disseminate single lines into production facilities worldwide. In this scenario, breeding programs need to account for G x E to maximize the benefits

355 of genetic improvement[91]. The possibility of disseminating many closely related animals to diverse
356 geographical locations and environmental conditions (Fig. 2) can be coupled with phenotyping technologies
357 for routine data collection to feedback information on performance under diverse settings. This may
358 facilitate production of differentiated strains tailored for specific environments, or inclusion of robustness
359 as a target trait such that a single strain has phenotypic plasticity within and across diverse environments[92].
360 An example of a robust strain that performs well in multiple environments is the Genetically Improved
361 Farmed Tilapia (GIFT) strain. In the late 1970s, inadequate tilapia stocks were hampering the development
362 of aquaculture in Asia. To develop a strain with robust performance in high and low input systems across
363 diverse environments, a base population including wild and farmed strains from eight African and Asian
364 countries was established. The breeding program focused primarily on improving growth rate, but involved
365 multiple farmers in different countries in evaluations to account for G X E. The GIFT strain is now farmed in
366 16 countries across Asia, Africa and Latin America, and grows 85 % faster than the base population[93].

367 Genomic selection can facilitate the breeding of more robust strains in aquaculture species where reference
368 populations (including close relatives of selection candidates) are tested in diverse environments[89,94]. The
369 performance of a genotype along an environmental gradient for any measurable trait can be used to
370 calculate the response curve, or reaction norm, of that genotype[89]. This reaction norm can be used as a
371 target trait for genomic selection to reduce sensitivity to environmental variation, with notably superior
372 results to sibling testing schemes alone[94]. The variation within and between production environments is
373 typically larger for aquaculture in low and middle incoming countries, and as breeding programs in such
374 settings increase in sophistication, the low-cost genomic selection methods described above should be
375 applied to help improve resilience of stock performance within and across environments to maximize the
376 benefits of genetic gain for producers.

377 ***Epigenetic programming to improve performance and environmental adaptation:*** Epigenetic mechanisms
378 or 'marks' (e.g. cytosine methylation, histone modifications, chromatin accessibility state) can be influenced

by the environment to result in substantial phenotypic variation from the same genomic DNA blueprint[103]. Recent domestication can profoundly alter the epigenome of hatchery-reared animals[104] via alterations to methylation profile, highlighting the potential for rapid epigenetic reprogramming. This potential can be harnessed by intentional environmental manipulation during crucial life stages (in particular larvae and broodstock) to improve production traits later in life and / or in subsequent generations[103,105,106]. For example, early-life use of plant-based diets improved the acceptance and utilization of these diets in later life in rainbow trout[107], and early-life stress can modulate future stress or immune responses in Atlantic salmon, which may have implications for robustness in adult stages[99,108]. Multigenerational epigenetic effects are of most relevance to selective breeding, and have been proposed to play a role in the fitness of the Manila clam (*Ruditapes philippinarum*), where adults exposed to low pH during gonadal maturation had faster-growing offspring compared to controls[109], and in the Sydney rock oyster (*Saccostrea glomerata*), where larvae of parents incubated under low-pH conditions grew and developed faster in low-pH conditions and had higher fitness as adults[110]. The development of assays to assess genome-wide cytosine modification, chromatin structure and accessibility across multiple aquaculture species (Box 2) will provide the toolbox to understand more about the mechanisms underpinning these epigenetic phenomena, and the availability of isogenic finfish lines is a useful resource to help distinguish genetic and epigenetic effects[113].

For heritable epigenetic marks that impact on production traits, it is highly likely that their impact will be directly captured and utilized by conventional sib testing and genomic selection, which are both based on phenotypic similarity between relatives[114]. However, distinguishing additive genetic and epigenetic components of phenotypic variation may facilitate improvement in genetic parameter estimation and prediction of response to selection[106]. Furthermore, an interesting intersection between epigenetic programming and genetic improvement via selective breeding may be related to optimizing of robust performance of improved stocks in multiple environments. The use of genomics to support breeding of 'robust' strains for multiple environments described above can be augmented with tailored epigenetic

programming to improve the performance of these strains in specific farmed environments. Furthermore, there is likely to be genetic variation in the response to targeted environmental manipulation, and genomic prediction using large full sibling families each split into groups tested with targeted environmental treatments can be used to assess this. Therefore, selection for improved response to epigenetic programming could be a route to realizing genetic improvement for impact across diverse production environments.

The microbiome as a predictor of performance. The microbiome is a critical component of the interaction between animals and their environment, and contributes to the health and performance of farmed animals[95,96]. Colonization and development of bacterial communities are essential to immune function and influenced by host physiology and immune response. Host microbial composition is heritable to some extent in marine species[9798], and differences have also been observed between farmed and wild strains of Atlantic salmon[99] and Pacific whiteleg shrimp (*Litopenaeus vannamei*)[100]. Microbiome research in aquaculture species is currently primarily focused on gaining understanding of its composition in various species[96,101]. Developments in DNA sequencing technologies have provided drastic improvements in microbiome analyses, in particular metagenomics approaches to sequencing all genomes within a sample. Microbiome sequencing may have potential when paired with host genotyping for prediction of production traits, with a potential example trait being ability of salmonids to tolerate increasingly vegetarian diets[102]. In terrestrial livestock, microbiome similarity matrices have been used to replace or complement the host genomic relationship matrix, with an improved predictive ability for feed conversion efficiency in Holstein Friesian dairy cattle[13]. In this context, microbiome composition can be considered as an 'intermediate phenotype' resulting from both host genetic and environmental influences, and has potential value in prediction of trait performance in later life, rather than prediction of offspring performance. The latter may depend in part on the heritable component of the microbiome, but is likely to be captured within additive genetic variation and breeding values for production traits.

Interaction between farmed and wild animals. The recent domestication of aquaculture species means that farmed species often co-exist in close proximity to wild counterparts, and there can be frequent interaction and interbreeding between the two groups. As species move along the domestication scale towards closed selective breeding populations, the genetic divergence between these farmed and wild populations widens. The genomes of farmed species are significantly altered by domestication and genetic improvement programs, which exert intense selection pressures[116]. Genomic tools can be applied to gauge these impacts as domestication progresses via high density genotyping or sequencing of multiple populations of farmed and wild populations, and comparison of genetic diversity across the genome to identify common signatures of selection[117,118]. The divergence between wild and farmed populations results in notable differences in growth, morphology, life history, behaviour, and physiology[119]. The impact of domestication on physiology of the animals has been demonstrated via studies of gene expression and genome methylation, which show marked differences after few generations of hatchery breeding in salmonids[120]. Introgression of potentially maladapted alleles into wild populations can lead to undesirable changes in life history traits, reduced population productivity, and decreased resilience[121]. Many species of marine fish and invertebrates are characterized by high connectivity, with associated high gene-flow, and high effective population size[122], such that the effects of introgression from farm-reared animals is rapidly diluted. Such introgression may even be beneficial in some species, e.g. bivalve shellfish, by contributing to natural recruitment and adding genetic variation to wild populations[123,124]. In contrast, freshwater and anadromous species are characterized by relatively small effective population sizes[125], and gene flow can be heavily modified (or blocked)[126,127]. Consequently, inflow of genes from farmed animals can result in rapid and substantial alterations to the genepool in populations of these species[125]. Therefore, methods of preventing escapees and interbreeding of farmed and wild animals are important for the sustainability of aquaculture and its long-term coexistence with extant wild populations[125,128,129]. Engineering and management solutions are unlikely to completely prevent escapees, and genetic technologies to prevent such introgression include

triploidy, currently used in a range of species including salmonids and oysters[130,131], or other means of inducing sterility in production stocks such as germ cell ablation via genome editing[132] (see section '*Genome editing to accelerate genetic improvement'* below).

In addition to protecting wild stocks, it is important to maintain genetic resources for farmed strains as they undergo domestication. Biobanking is applied for conservation of germplasm of aquatic animals, both for vulnerable wild species and farmed strains to avoid losing genetic diversity. There are established repositories and gene banks for finfish and shellfish, and technologies for preservation of gametes, tissues, and cell lines are developing rapidly, with detailed reviews available [133,134]. However, the field remains at a relatively early stage compared to equivalent efforts in crops and terrestrial livestock. While cryopreservation of sperm is routine for several fish and shellfish species, cryopreservation of oocytes is much more challenging to achieve. Cryopreservation of ovarian tissues is a promising alternative, but would require research into *in vitro* culture of these tissues[134], and surrogate broodstock (discussed below) hold promise to preserve genetic resources via transplant of primordial germ cells[135]. As these methods develop, preservation of aquatic genetic resources will also benefit from more centralized efforts, akin to the seedbanks for crops together with associated FAO standards and procedures for biobanking[136].

Biotechnology in the future of aquaculture breeding

While there is much potential to be realized in the optimal use of genomic tools to support domestication and genetic improvement, innovation in the application of biotechnology to aquaculture genetics also holds promise to tackle production barriers. This includes use of genome editing technologies to make targeted changes to the genomes of aquaculture species', resulting in improved health and performance, use of reproductive biotechnologies such as surrogate broodstock to expedite genetic gain, and combinations of both approaches.

Genome editing to accelerate genetic gain. Genome editing tools such as engineered CRISPR/Cas9 systems[137,138] are invaluable to understanding genetic regulation of economically-important traits, and have

potential to accelerate genetic gain in aquaculture breeding programs (Fig. 3). The Cas9 enzyme makes a double-stranded cut at a genomic site corresponding to a guide RNA, which results in either small insertions or deletions that can lead to loss-of-function mutations (non-homologous end joining), or in user-defined edits to the genome based on a provided DNA template (homology directed repair). Since the first demonstration of effective genome editing in Atlantic salmon [139], CRISPR/Cas9 has been successfully applied in various farmed finfish and mollusc species, primarily for gene knockout and as proof of principle[140]. Microinjection into early stage embryos is the most commonly used delivery method, but can be inefficient and alternative delivery methods, such as electroporation of sperm, hold promise[141]. Genome editing can be used as a component of pipelines to identify putative causative genes and variants, for example by assessment of gene knockout on traits of interest. Exploitation of genome-wide loss of function CRISPR screens such as GeCKO (Genome-scale CRISPR Knock-Out)[142] in aquaculture species offers a powerful tool to explore the genetic basis for resistance to certain pathogens, and the successful editing of a salmonid fish cell line using a lentivirus delivery system suggests that this approach is technically viable[143]. However, cell line resources for many aquaculture species are limited, in particular invertebrate species, and targeted development of suitable cell lines for important aquaculture species is required. As an alternative approach, *in vivo* GeCKO may be plausible in some species, due to their external fertilization, abundance of embryos, and feasibility of early life screening[140]. This is likely to require the development of Cas9-stable broodstock and a method of delivering guide RNA libraries *en masse* to early-stage embryos. Combining such genome-wide screening approaches with mapping and shortlisting causative functional variants in QTL regions, will create opportunities for targeted experiments testing candidate causative alleles, followed by assessment of the consequences on the trait (Fig. 3).

There are several potential applications of genome editing to expedite genetic improvement and tackle production barriers in aquaculture. Firstly, it could allow the rapid fixation of favorable alleles at QTL segregating within breeding populations[144]. Secondly, since most modern aquaculture breeding programs

499 are closed systems, introgression of favorable alleles from other populations, potentially including wild
500 stocks, is logistically and biologically challenging. Genome editing can facilitate introgression-by-editing of
501 such alleles from other populations, strains or species into a breeding population[140]. Finally, it is possible to
502 create *de novo* alleles based on knowledge of the biology of the trait in question, or utilizing targets from
503 GeCKO screens. For example, removal of an exon of the CD163 gene in pigs (*Sus scrofa*) resulted in
504 complete resistance to the porcine reproductive and respiratory syndrome virus[145]. While disease resistance
505 is also likely to be the primary focus for genome editing in aquaculture, other traits such as adaptation of
506 stocks to plant-based diets, and sterility to prevent introgression and unwanted effects of precocious
507 maturity[146,147] are other key objectives. Knockout of germline-specific genes such as dead end (*dnd*), *nanos2*
508 or *nanos3* resulted in sterility in Atlantic salmon[132] and Nile tilapia[148], respectively. For practical applications,
509 genome editing needs to be integrated into well-managed breeding programs to ensure maintenance of
510 genetic diversity. Genome editing *en masse* in production animals is unlikely to be feasible, and therefore
511 editing of the germline of broodstock animals is highly likely to be the most effective approach. Sterility
512 requires special consideration because it is by definition non-heritable, and inducible transgenic targets may
513 be required. However, sterility may be a useful trait to include with other genome editing targets to negate
514 the risk of edited alleles being transferred to wild stocks (e.g. via escapees). Refinement of genome editing
515 methods are occurring constantly, and use of modified CRISPR/Cas systems such as CRISPR activation
516 (CRISPRa) or CRISPR interference (CRISPRi) can induce differences in expression levels of target genes
517 instead of complete knock-out[149–151]. Such tools will be valuable in elucidating the functional genetic basis of
518 production traits, for fundamental understanding of genome function and for future application in
519 aquaculture breeding programs. However, it is critical that edited stocks are carefully studied to detect and
520 avoid off-target editing, and rigorously monitored to discount deleterious pleiotropic effects, and
521 aquaculture can follow procedures used in terrestrial livestock to achieve this[152]. Furthermore, any practical
522 application for aquaculture depends entirely on an acceptable regulatory and public approval landscape[153],

523 and the approval of the genetically-modified AquaAdvantage salmon (Aquabounty) as fit for human
524 consumption by the US Food and Drug Administration and the Canadian Food Inspection Agency was a
525 recent landmark[154]. Target traits that have concurrent production and animal welfare / environmental
526 benefits should be a focus for genome editing in aquaculture, and public and policy-maker engagement on
527 the technology, its benefits, and its risks, is absolutely vital.

528 ***Surrogate broodstock to reduce generation intervals.*** A key factor in the rate of genetic gain in a breeding
529 program is the length of the generation interval. Consider the breeder's equation;

$$\Delta G = \frac{i\, r\, \sigma_A}{y}$$

530

531 Where ΔG is genetic gain over time, i is selection intensity, r is selection accuracy, σ_A is additive genetic
532 variance, and y is generation time. Genomic selection has resulted in a step increase in selection accuracy,
533 and much research is now devoted to achieving more incremental increases. However, decreasing
534 generation time has potential for more drastic changes to genetic gain, especially considering that many of
535 the major aquaculture species have relatively long generation intervals (e.g. up to 20 years in sturgeon,
536 family *Acipenseridae*). Surrogate broodstock technologies are based on the concept of isolation of the
537 primordial germ cells (PGC) of selected broodstock animals at an early life stage, and transplantation of
538 these cells into the surrogate [a germ cell-ablated specimen of a species with shorter generation time (Fig.
539 4). When combined with genomic selection using samples from embryos or juveniles, surrogate broodstock
540 technology could potentially reduce generation interval without significant loss of selection accuracy (Fig.
541 4). Germ cell isolation, transplantation and successful gamete production in surrogate broodstock has been
542 demonstrated across species within a genus, and even across genera[155], for example rainbow trout offspring
543 were produced when spermatogonia from rainbow trout were injected into newly-hatched sterile masu
544 salmon (*Oncorhynchus masou*)[156]. The same technology has other potential applications, for example to
545 produce offspring from a species which is challenging to rear in captivity using surrogates, such as Atlantic

546 bluefin tuna (*Thunnus thynnus*) gametes from chub mackerel (*Scomber japonicus*)[155]. In addition, surrogate
547 technology can be coupled with genome editing of PGC to create germline-edited animals, as successfully
548 demonstrated in chickens[157]. This approach is a route to genome editing for aquaculture species where
549 access to the newly fertilized embryos is challenging, like certain crustaceans[158] or ovoviviparous species
550 such as rockfish (*Sebastes* spp.)[159]. While clearly a long-term and high-risk research goal, the combination of
551 surrogate broodstock, genome editing, and genomic selection has potential to drastically increase the rate
552 of genetic gain in breeding programs via the reduction of generation interval. While extensive effort and
553 resources have been put into the use of functional genomic data to improve selection accuracy in breeding,
554 such reproductive technologies require equivalent attention.

555 **Conclusions**

556 In contrast to terrestrial livestock and crop production, most aquaculture production derives from species
557 for which domestication and breeding is at an early stage. Genetic improvement and dissemination of
558 germplasm originating from a well-managed breeding program enables cumulative increases in production
559 traits, and facilitates adaptation to emerging challenges, such as climate change or infectious disease
560 outbreaks. With the recent growth and accessibility of the genomic toolbox, genomics should be utilized
561 from the outset of domestication and breeding program design to inform base population composition,
562 maintain genetic diversity, and understand sex determination and differentiation. Genomic selection has
563 revolutionized terrestrial livestock breeding and is commonplace in advanced aquaculture sectors like
564 salmon, but judicious application of multi-purpose cost-effective marker panels may be necessary to
565 translate those benefits to most aquaculture species where the industries are smaller and more
566 fragmented. The ability to disseminate closely-related individuals to diverse testing and production
567 environments, combined with genomic selection, should be applied to tackle genotype by environment
568 interactions and improve robustness. Genomic tools can also inform on the potential of the microbiome and
569 epigenome as useful intermediate phenotypes, and as conduits to improve capacity for adaptation of stocks

to environmental challenges. For the more advanced aquaculture sectors, the immediate future will include mapping and understanding functional genomic variants, harnessing the species' high fecundity to perform high-resolution genetics and genomics experiments paired with highly contiguous and well-annotated genome assemblies. Genome editing is key to this process, and as such requires species-specific optimization both *in vivo* and in cell culture, with the development of suitable cell lines for aquaculture species also being an important focus, e.g. to assist with genome-wide CRISPR screens for disease resistance. The widespread commercial application of genome editing in aquaculture appears several years away, but it has clear potential for step-changes in trait improvement to help address production barriers. In the longer term, developments in surrogate broodstock technology combined with genomic selection has the potential for shortening of generation interval to expedite genetic gain. Underpinning many of these advances is an improved knowledge of the genetics and biology of key production traits, which is particularly pertinent for the many aquaculture species from understudied taxa with major knowledge gaps relating to fundamental inheritance and genome biology. Overall, there is now an unprecedented opportunity to harness genomics to fast-track the domestication and genetic improvement of farmed aquatic species, which will be necessary to secure the sustainable growth of aquaculture as one of the most promising solutions to the current global food security challenge.

Box 1 | **Genomic tools to optimise the domestication process**

601 Culture of new aquatic species is established each year, and there is potential to use genomic tools to
602 optimize and characterize the domestication process from the outset. Historically, the mismanagement of
603 genetic resources and diversity during this process has led to reduced genetic resilience[25] and the
604 subsequent emergence of "crowd" diseases in farmed populations[168] which can be catastrophic for
605 emerging industries. Targeted use of appropriate genomic tools throughout the domestication process,
606 could delay, mitigate, or even remove the potential for this to happen by retaining genetic resilience in both
607 wild and farmed populations. Genomic tools have already made significant contributions to optimize
608 scientific breeding programs, and to proactive species conservation strategies for both farmed and wild
609 populations of target species[169]. However, the recent and rapid development of genomic tools, together
610 with their accessibility and cost-efficiency, means that optimal tools can be applied at each stage of the
611 progression along the domestication and selective breeding continuum (see Figure).

612 Cleaner fish such as Ballan wrasse (*Labrus bergylta*) and lumpfish (*Cyclopterus lumpus*) are used in
613 commercial salmon production to eat sea lice from the skin of the fish, and are a key aspect of integrated
614 pest management. Wrasse and lumpfish[170] production began in 2009 and 2011 respectively[171], with life
615 cycles in captivity closed in 2018 and 2016[172] and reference genomes released by 2016[173] and 2018[174]. Both
616 domestication processes have combined animal biology, health management and nutritional requirements
617 together with development of genomic tools for genetic management and enhancement[172]. Trial crosses
618 are crucial when selecting broodstock, as this allows for the greatest gains in early stage production and the
619 resolution of incompatibilities. For example, in lumpfish initial gains in production from trial crosses
620 included 100–fold differences in growth rates between age-matched individuals. Where trials are executed
621 in combination with relatively cheap GBS, these crosses can capture both phenotypes of interest and
622 maintain the broad genetic diversity which makes the significant phenotypic differences evident at these
623 early stages. In addition, broodstock selection should run concurrently with evaluation of wild stock

624 population structure, using genomic tools to inform management strategies for species conservation and
625 rapid diagnostics of genetic introgression[169] (see Figure).

626 When moving towards more advanced selective breeding programs, bespoke tools such as SNP-arrays can
627 be applied, but their cost-effectiveness needs considered and compared to GBS. These tools can then be
628 applied to understand the genetic architecture of production traits, and to support genomic selection to
629 maximize genetic gain and minimizing inbreeding. This SNP discovery and high density genotyping also
630 paves the way for generation of targeted low density SNP panels, which can have concurrent uses to
631 support parentage assignment, stock management, traceability, and low-cost genomic selection. Finally,
632 due to the relative ease of generating reference genome assemblies, they should be undertaken from the
633 outset of the domestication of a new species for aquaculture, as they inform the choice of marker panels for
634 genotyping and subsequent studies to understand the biology of production traits.

635 Figure: Progression of domestication and the potential input of molecular genetics for optimization at each
636 stage of this process.

637

Box 1 | **Genomic tools and resources for aquaculture species**

Genomic resources for farmed finfish, crustaceans, molluscs and echinoderms of significant global value are summarized in the table below (full data provided for the top 20 species per each taxonomic group in Supplementary Table 7). The development and uptake of genomic resources has lagged behind livestock by several years. While genomic resources have been accumulating rapidly for aquaculture species, several high value species remain without a publicly available high-quality reference genome and have limited genomic resources (see Table). Genomic complexities largely absent in livestock present a pervasive challenge to the generation of high-quality reference genomes in aquaculture species, and as a result many of the existing assemblies are currently very fragemented (see Table). These include the widespead presence of duplicated loci due to genome duplication events, e.g. in salmonids[34], cyprinids[160], and sturgeons[161], and extreme rates of heterozygosity, e.g. in bivalves[162] and crustaceans[163]. Such features seriously hinder assembly algorithms using short-read sequence data. Nonetheless, more recent sequencing technologies, including platforms generating long reads (e.g. PacBio and Oxford Nanopore) and linked short reads (10X Genomics) are increasingly applied to generate improved quality assemblies. When combined with long-range scaffolding technologies such as high-throughput chromatin conformation capture approaches (Hi-C; e.g. Dovetail Genomics) and / or optical mapping (e.g. Bionano Genomics), high quality contiguous assemblies are possible even for challenging genomes[164], and such strategies are currently being applied to aquaculture species (Supplementary Table 7). For example, a recent genome assembly of the yellow perch (*Perca flavescens*) resulted in 24 (2n = 24) chromosome-size scaffolds covering 99 % of the complete assembly, with an N50 of 37.4 Mb[165]. All major aquaculture species are likely to benefit from such high-quality assemblies in the near future. A crucial next step is to improve genome annotation to help reveal the functional basis of phenotypic variation and facilitate identification of causative genetic variants. RNA sequencing combined with advances in software for read alignment and quantification has facilitated genome-wide prediction of coding and non-coding genes in many aquaculture species, replacing

microarrays as the standard for global quantification of gene expression. Single cell RNA sequencing is yet to be applied to aquaculture species, but offers opportunities to understand complex and rare cell populations, uncover regulatory relationships between genes, and study trajectories of distinct cell lineages in development[166]. Discovery and exploitation of epigenetic marks, including DNA and histone modifications impacting gene regulation in aquaculture species represents a crucial step to help bridge the genotype-phenotype gap[103]. The Functional Annotation of Animal Genomes (FAANG) initiative[144] is a concerted effort to map such features in livestock, with the Functional Annotation of All Salmonid Genomes (FAASG) being an equivalent community initiative for salmonid fish[35], and equivalent initiatives are likely to follow for other major aquaculture species.

Table: The genomic resources available for the highest value finfish, crustacean, mollusc and other species.

Species	**Production value** ($Bn)	**Genome size** (Gbp)	**Scaffold N50** (Mbp)	**Number of coding genes**	**Published SNP arrays**	**Total number of re-sequenced genomes**
Atlantic salmon (*Salmo salar*)	16.69	2.96	1.36	48,775	7 (15k-286K)	165
Grass carp (Ctenopharyngodon idella)	12.64	0.90	6.45	27,263	-	1
Silver carp (*Hypophthalmichthys molitrix*)	10.26	1.10	0.31	-	-	-
Nile tilapia (*Oreochromis niloticus*)	7.61	1.00	38.8	29,550	2 (50K-58K)	65
Bighead carp (*Hypophthalmichthys nobilis*)	7.31	1.01	0.08	-	-	-
Whiteleg shrimp (*Litopenaeus vannamei*)	26.74	1.63	0.6	24,987	1 (6K)	-
Red swamp crawfish (*Procambarus clarkii*)	10.00	2.07	0.001	136,962	-	-
Chinese mitten crab (*Eriocheir sinensis*)	9.54	1.54	0.49	-	-	-
Giant tiger prawn (*Penaeus monodon*)	5.59	1.44	0.007	18,115	1 (6K)	2
Oriental river prawn (*Macrobrachium nipponense*)	2.09	-	-	-	-	-
Japanese carpet shell (*Ruditapes philippinarum*)	6.95	2.56	0.048	108,034	-	15

Chilean mussel (*Mytilus platensis*)	2.50	-	-	-	-	-
Constricted tagelus (*Sinonovacula constricta*)	1.41	-	-	-	-	-
Pacific cupped oyster (*Crassostrea gigas*)	1.24	0.55	0.4	28,398	2 (27K -190K)	516
Blood cockle (*Tegillarca granosa*)	1.02	-	-	-	-	-
Japanese sea cucumber (*Apostichopus japonicus*)	1.40	0.8	0.48	30,350	-	1

673

674

675

676

677

678

679

680

681

682 Box 3 | **Genetic solutions to major diseases in aquaculture: The example of IPN in salmon**

683 Infectious pancreatic necrosis (IPN) is a viral disease that was one of the primary concerns for salmon
684 farming around the turn of the century, with frequent outbreaks causing high levels of mortality (up to 90
685 %) to stocks both in freshwater hatcheries and following transfer to sea cages. Resistance to IPNV was
686 shown to be moderately to highly heritable[175], and breeding companies began to implement family-based
687 selection. In parallel, teams from the UK and Norway undertook projects to investigate the molecular
688 genetic basis of resistance, using microsatellite markers and sparse linkage maps available at the time. Both
689 groups identified a single major QTL on Chromosome 26, and showed that it could explain 80 – 100 % of
690 genetic variation in resistance to IPNV in sea water field trials[176], and also in experimental freshwater
691 trials[177–179]. High-throughput sequencing approaches subsequently allowed development of SNP-based
692 genetic tests to predict IPN resistance of salmon without the need for regular disease challenge
693 experiments[180,181]. The practical outcome of these experiments was extensive use of MAS for the favourable
694 allele in all major salmon breeding programs, assisted by the fact that the resistance allele is dominant[178,181].
695 The results were striking, with a sustained decrease in the incidence of IPN outbreaks to near zero (see
696 Figure below,[64]). Follow up functional studies highlighted the marked differences in gene expression
697 response to infection between resistant and susceptible salmon fry[182] and suggested that epithelial cadherin
698 may be part of the mechanism underlying the QTL[181]. However, the exact causative mutation(s) and nature
699 of their effect remain at least partly elusive. The use of MAS to tackle IPN resulted in widespread
700 recognition of the potential of (molecular) genetics in selective breeding to tackle infectious diseases of
701 aquaculture. While other examples of major QTL for production relevant traits exist (see main text), the IPN
702 case resulted in a Gartner Hype Cycle whereby the 'Peak of Inflated Expectations' was that other diseases
703 could be solved by mapping a single QTL and using MAS, which was followed by a 'Trough of
704 Disillusionment' when studies began to show the typically oligogenic nature of disease resistance traits.

However, the field has now reached a 'Plateau of Productivity' where genomic selection has become state of the art for most traits, and MAS is applied for selected major QTL such as IPN.

Figure: The marked reduction in mortalities due to the IPN virus in Marine Harvest (now Mowi) production farms in Norway (adapted from [64]).

737 **Table 1** | Summary of studies testing genomic prediction for production traits in aquaculture species, including comparing prediction accuracy using
738 genomic and pedigree-based approaches.
739

Species	Trait	Measurement	Heritability (pedigree)	Accuracy (pedigree)	Relative increase	Genotyping technology	Ref
Atlantic salmon (Salmo salar)	Growth	Weight	0.60 (0.48)	0.70 (0.58)	21%	SNP array (132K, 112K post-filtering)	183
		Length	0.61 (0.51)	0.66 (0.56)	18%		183
	Resistance to sea lice	Lice count	0.33 (0.27)	0.60 (0.48)	25%	SNP array (132K, 33K post-filtering)	184
		Lice count	0.22(0.27)	0.46 (0.43)	7%		184
		Lice count	0.11 (0.10)	0.50 (0.41)	22%	SNP array (50K, 37K post-filtering)	185
		Log lice density	(0.14)	0.52 (0.34)	52%	SNP array (220K)	62
	Resistance to amoebic gill disease	Gill score	0.24 (0.25)	0.62 (0.51)	22%	Two species SNP array (17K, 7K post-filtering)	186
		Amoebic load	0.25 (0.36)	0.70 (0.60)	17%		186
		Gill score	0.28 (0.32)	0.72 (0.61)	18%	SNP array (55K, 53K post-filtering)	187
	Resistance to salmon rickettsial syndrome	Time to death	0.27 (0.18)	0.41* (0.34)	21%	SNP array (50K, 50K post-filtering)	188
		Binary survival	0.39 (0.26)	0.26 (0.20)	30%		188
	Fillet pigmentation	-	(0.43)	0.44 (0.36)	22%	SNP array (220K)	62
	Muscle fat	-	0.25 (0.36)	0.56 (0.60)	-7%	SNP array (57K, 50K post-filtering)	189
	Omega-3 fatty acid content	DHA	0.20 (0.21)	0.41 (0.33)	24%		189
		EPA	0.04 (0.06)	0.32 (0.37)	-14%		189
Rainbow trout (Oncorhynchus mykiss)	Resistance to bacterial cold water disease	Binary survival	–	0.68* (0.36)	89%	SNP array (57K, 45K post-filtering)	190
		Time to death	0.33 (0.37)	0.67* (0.34)	97%	SNP array (57K, 36K post-filtering)	191
		Binary survival	0.35 (0.35)	0.70* (0.36)	94%		191
		Time to death	0.29 (0.31)	0.49 (0.50)	-2%	SNP array (57K, 41K post-filtering)	192
		Binary survival	0.45 (0.48)	0.46 (0.41)	12%		192
	Resistance to infectious pancreatic necrosis virus	Time to death	0.25 (0.40)	0.53 (0.49)	8%	SNP array (57K, 38K post-filtering)	193
		Binary survival	0.24 (0.35)	0.56 (0.50)	12%		193
	Resistance to salmon rickettsial syndrome	Time to death	0.45 (0.38)	0.78* (0.61)	28%	SNP array (57K, 27K post-filtering)	194
		Binary survival	0.55 (0.54)	0.60* (0.47)	28%		194
	Resistance to Infectious hematopoietic necrosis virus	Time to death	0.23 (0.33)	0.33 (0.13)	154%	SNP array (57K, 35K post-filtering)	195
		Binary survival	0.25 (0.28)	0.39 (0.24)	63%		195
	Resistance to columnaris disease	Binary survival	0.32 (–)	0.11 (-0.02)	-650%	SNP array (57K, 36K post-filtering)	196
		Binary survival	0.51 (–)	0.22 (0.06)	267%	SNP array (57K, 34K post-filtering)	196
Coho salmon	Resistance to salmon	Time to death	– (0.14)	0.52 (0.27)	93%	ddRAD (9K)	197

(Oncorhynchus kisutch)	rickettsial syndrome	Binary survival	– (0.27)	0.81 (0.31)	161%		197
Carp (Cyprinus carpio)	Growth	Length	0.33 (0.33)	0.71 (0.60)	18%	RAD-seq (20K)	198
	Resistance to koi herpesvirus	Binary survival	0.50 (0.61)	0.53* (0.49)	8%	RAD-seq (16K)	68
Nile tilapia (Oreochromis niloticus)	Growth	Harvest weight	0.36 (0.31)	0.60 (0.48)	25%	SNP array (43K, 32K post-filtering)	199
		Fillet yield	0.21 (0.21)	0.62 (0.54)	15%		199
		Harvest weight	0.17 (0.22)	0.29 (0.19)	53%	SNP array (59K, 48K post-filtering)	200
		Fillet weight	0.16 (0.24)	0.34 (0.18)	89%		200
		Fillet yield	0.23 (0.33)	0.54 (0.46)	17%		200
European sea bass (Dicentrarchus labrax)	Resistance to viral nervous necrosis	Binary survival	0.43 (0.27)	0.62* (0.67)	-7%	RAD-seq (9K)	201
Gilthead sea bream (Sparus aurata)	Resistance to pasteurellosis	Time to death	0.28 (0.22)	0.44* (0.30)	47%	2b-RAD (22K)	202
	Resistance to pasteurellosis	Time to death	0.32 (0.32)	0.54* (0.45)	20%	2b-RAD (28K)	203
		Binary survival	0.33 (0.31)	0.56* (0.46)	22%		203
Turbot (Scophthalmus maximus)	Resistance to Scuticociliatosis	Resilience	0.15 (–)	0.46 (0.41)	12%	2b-RAD (18K)	204
		Resistance	0.26 (–)	–	–		204
		Endurance	0.12 (–)	–	–		204
Japanese Flounder (Paralichthys olivaceus)	Resistance to Edwardsiella tarda	Binary survival	– (–)	0.603 (–)	–	WGS (1.9M)	205
Channel catfish (Ictalurus punctatus)	Growth	Harvest weight	0.27 (–)	0.37 (0.29)	28%	SNP array (660K, 55K post-filtering)	206
		Residual carcass weight	0.34 (–)	0.31 (0.24)	29%		206
Large yellow croaker (Larimichthys crocea)	Growth	Body weight	0.60 (–)	0.41 (–)	–	ddRAD (30K)	207
		Body length	0.59 (–)	0.40 (–)	–		207
	n-3HUFA	–	0.44 (–)	0.30 (–)	–	ddRAD (32K)	207
Yellowtail kingfish (Seriola lalandi)	Growth	Weight	0.47 (0.42)	0.69 (–)	–	DArT-Seq (14K)	208
		Length	0.43 (0.42)	0.67 (–)	–		208
		Condition index	0.21 (0.11)	0.44 (–)	–		208
Yellow drum (Nibea albiflora)	Growth	Body length	– (–)	0.38* (–)	–	GBS (54K)	209
		Swimming bladder index	– (–)	0.17* (–)	–		209
		Swimming	– (–)	0.22* (–)	–		209

		bladder weight					
		Body thickness	– (–)	0.24* (–)	–		209
		Body height	– (–)	0.30* (–)	–		209
		Body length / body height ratio	– (–)	0.36* (–)	–		209
		Gonad weight index	– (–)	0.37* (–)	–		209
Oyster (Crassostrea gigas)	Growth	Shell length	0.26 (0.23)	0.54 (0.44)	23%	Two species SNP array (38K, 23K post-filtering)	210
		Shell height	0.23 (0.20)	0.60 (0.47)	28%		210
		Wet weight	0.35 (0.31)	0.67 (0.54)	24%		210
	Resistance to Osterid Herpesvirus	Binary survival	0.37 (0.25)	0.76 (0.64)	19%		211
Yesso scallop (Patinopecten yessoensis)	Growth	Shell height	0.48 (–)	0.53 (–)	–	2b-RAD (2K)	212
		Shell length	0.48 (–)	0.46 (–)	–		212
		Shell width	0.36 (–)	0.55 (–)	–		212
Zhikong scallop (Chlamys farreri)	Growth	Shell length	0.42 (–)	0.65* (–)	–	2b-RAD (31K)	213
		Shell height	0.47 (–)	0.70* (–)	–		213
		Shell width	0.54 (–)	0.63* (–)	–		213
		Whole weight	0.28 (–)	0.64* (–)	–		213
Whiteleg shrimp (Litopenaeus vannamei)	Growth	Body weight	0.32 (–)	0.62 (–)	–	2b-RAD (23K)	214
		Body length	0.45 (–)	0.61 (–)	–		214
		Body length	– (–)	0.30* (–)	–	SLAF-seq (6K)	215
		Body weight	– (–)	0.41* (–)	–		215
	Resistance to AHPND	Time to death	0.26 (0.24)	0.50 (0.47)	6%	2b-RAD (23K)	216
		Binary survival	0.16 (0.15)	0.21 (0.20)	5%		216
Banana shrimp (Fenneropenaeus merguiensis)	Growth	Body weight	0.55	0.76 (0.65)	17%	DArT-Seq (9K)	217
		Body length	0.49	0.73 (0.60)	22%		217
		Head length	0.39	0.42 (0.32)	31%		217
		Body width	0.61	0.72 (0.60)	20%		217
		Tail weight	0.45	0.77 (0.66)	17%		217
		Meat yield	0.10	–	–		217
	Colour	Dark (raw shrimp)	0.18	0.59 (0.53)	11%		217
		Red (cooked	0	NA	–		217

	shrimp)					
	‘Flesh streaks’	–	0	NA	–	[217]
	Yellow hepatopancreas	–	0.03	NA	–	[217]
	Resistance to HPV	Viral load	0.35	0.60 (0.09)	567%	[217]

* Alternative statistical models to GBLUP were used, e.g. Bayesian models or RRBLUP

Fig. 1 | **A summary of global aquaculture diversity and production. A**). Phylogenetic tree showing farmed species with an annual production value higher than $1,000M per annum (see Supplementary Table 1). Estimated divergence times are from refs [218–224]. **B**). The time at which species were first farmed or domesticated including species which account for 80 % of all farmed seafood production and 95 % of all meat globally. Arrow in the bar denotes the point at which the first scientifically-driven selective breeding studies were undertaken for each species (note this could not be identified precisely for chickens or goats). Fading of timelines denotes uncertainty (see Supplementary Table 3.) **C**). Seafood production globally by sector and continent[2] (see Supplementary Table 4).

Fig. 2 | **Genomic selection within an aquaculture breeding program.** Full-siblings from a number of families are split into selection candidates and animals for phenotypic evaluation. These full-siblings of the selection candidates can be grown in different environmental conditions and phenotyped for different traits, for example using pathogen challenges to estimate resistance to different diseases or measuring performance traits in diverse production environments. The selection candidates and their phenotyped full-siblings are all genotyped, and a genomic relationship matrix reflecting the genetic similarity between each pair of animals is built. This relationship matrix and the collected phenotypes enable the estimation of breeding values for the selection candidates through the use of genomic selection models such as GBLUP or Bayesian models [13].

Fig. 3 | **Discovering functional variants using genomics and genome editing.** Three complementary strategies to discover causative variants affecting traits of interest for aquaculture breeding are represented. The first is 'Mapping and understanding QTL' which harnesses GWAS and within-family QTL mapping approaches to detect genomic regions associated with these traits, followed by functional genomic comparison of animals carrying alternate genotypes at the identified QTL. Identified SNPs within the region of candidate genes are then annotated according to their position in the genome to prioritise them as targets for validation using CRISPR/Cas9 genome editing. The second is 'Comparative genomics' where two closely related species that differ for a high priority trait (e.g. resistance to sea lice) are compared using comparative and functional genomics, again leading to potential genome editing targets for validation. The third is 'Reverse genetics' where pooled, genome-wide CRISPR screens can be applied in cell culture, followed by screening based on markers of infection or resistance to infection to identify key genes involved in disease resistance. The high fecundity of aquaculture species may allow analogous approaches *in vivo* using Cas9 transgenic broodstock followed by screening of embryos or juveniles. The three categories of functional variants identified in the inner circle all have potential for genetic improvement, either via marker-assisted or functionally-enriched genomic selection, or directly via genome editing of broodstock after a further testing and validation phase of research.

Fig. 4 | **Potential application of surrogate broodstock technology to accelerate genetic gain.** This approach involves the transplantation of germ cells from a donor species ('Target') to a recipient species ('Surrogate'), which then produces gametes of the donor. The main interest for aquaculture is to transfer the germ cells of the selected breeders of the farmed species to a surrogate which is easier to maintain in captivity and has a shorter generation time, reducing the time between two successive rounds of selection, which will assure the success of production and accelerate the rate of genetic gain of the breeding program. The germ cells of the surrogate must be ablated before transplantation. In this respect, germ cell free animals can be obtained through chromosome set manipulation (i.e. triploidy [156]) or the functional manipulation of genes fundamental for germ cell survival (e.g. through genome editing[132]).

Glossary

BROODSTOCK

A group of sexually mature individuals used in aquaculture for breeding purposes.

HYBRIDIZATION

Crossing between genetically divergent strains or species, usually aiming to achieve improved performance in offspring.

POLYPLOIDIZATION

Induction of increased ploidy levels to achieve improved production performance.

GENETIC GAIN

Improvement in average genetic value, and therefore improved phenotypes, in a population due to selection over cycles of selective breeding.

OVIPAROUS

Producing offspring by means of eggs which are hatched after they have been laid by the parent.

OVOVIVIPAROUS

Producing offspring by means of eggs which are hatched within the body of the parent.

BROADCAST SPAWNING / MASS SPAWNING

Release of high numbers of eggs and sperm into the water, where fertilization occurs externally.

GONOCHORISM

The state of having just one of at least two distinct sexes in any one individual organism.

SYNCHRONOUS HEMAPHRODISM

Where an adult organism in a species has both male and female sexual organs at the same time.

SEQUENTIAL HERMAPHRODITISM

Where an individual in a species is born as one sex, but can later change into the opposite sex.

BEHAVIORAL PLASTICITY

The ability of an organism to change its behavior following exposure to stimuli, such as changing environmental conditions.

GENETIC BOTTLENECK

A sharp reduction in genetic diversity, typically due to a large reduction in population size caused by environmental events or human activities.

BASE POPULATION

The population of animals used to start a selective breeding program.

INBREEDING DEPRESSION

The reduced biological fitness in a given population as a result of inbreeding, typically due to deleterious recessive alleles.

SNP ARRAY

A type of DNA microarray which is used to genotype genome-wide polymorphisms within a population.

GENOTYPING BY SEQUENCING

A method using high-throughput sequencing to discover and genotype genome-wide SNPs within a population.

SCAFFOLDING

An approach during genome assembly where contigs (i.e. continious assembled sequences) are linked into larger continguous sequences including gaps of known length.

INTROGRESSION

The deliberate movement of a target locus from one species or strain (donor) into another (recipient) by the creation and repeated backcrossing of a hybrid with one of the donor species or strains.

EFFECTIVE POPULATION SIZE

The size of an idealised population which would give rise to the rate of inbreeding and the rate of change in variance of allele frequencies actually observed in the population under consideration. It is approximate to the number of individuals that contribute gametes to the next generation.

SOFT SWEEP

The increase in frequency and / or fixation of a favourable allele at an existing polymorphic locus due to strong positive selection pressure.

MARKER-ASSISTED SELECTION

The selection of breeding individuals for genetic improvement of a trait of interest based on genetic markers linked to a quantitative trait locus affecting that trait.

QUANTITATIVE TRAIT LOCUS

A region of the genome which explains a significant component of variation in a trait of interest.

MENDELIAN SAMPLING

The chance factor in the process of distributing half the genetic material from each parent to their offspring, which is the source of within-family genetic variation.

POLYGENIC

A heritable trait under the control of many loci, typically of small effect.

GENOMIC SELECTION

The selection of breeding individuals for genetic improvement of a trait of interest based on the use of genome-wide genetic markers to estimate genomic breeding values. Genetic marker genotypes and phenotypes are measured in a reference population to predict breeding values of selection candidates that have genotypes only.

REFERENCE POPULATION

In genomic selection, the population of animals which have both genotypes and phenotypes. These data are used to estimate genetic marker effects, which are then applied to predict breeding values for genotyped selection candidates.

PHENOTYPING

Collection of measurements relating to traits of interest to the goals of a breeding program.

ACCURACY

In the context of genomic selection, accuracy is the correlation between the estimated genomic breeding values and the true breeding values.

GENOTYPE IMPUTATION

The statistical inference of unobserved genotypes based on knowledge of haplotypes in a population, typically used to predict high density marker genotypes when most individuals are genotyped for low density marker genotypes.

CAUSATIVE VARIANT

A polymorphism within the genome of a population that has a direct effect on a trait of interest, as opposed to simply being a genetic marker associated with the trait.

INTERNET OF THINGS

A network of physical objects that use sensors and application program interfaces to connect and exchange data over the Internet.

GENOMIC RELATIONSHIP MATRIX

A matrix containing the estimation of the proportion of total genomic DNA shared by any two individuals based on genome-wide genetic marker data.

BREEDING NUCLEUS

The elite broodstock animals that are maintained only for breeding, which is followed by multiplication and disemination of the genetically improved animals for production.

SURROGATE BROODSTOCK

Sterile animals used for the production of gametes of another individual, strain, or species.

INTROGRESSION-BY-EDITING

Using genome editing to change a target allele in the recipient population to correspond to the sequence of the allele in the donor population.

GECKO SCREEN

The process of transducing a genome-scale CRISPR-Cas9 knockout guide RNA library into a cell culture, then performing a screening experiment (e.g. survival after pathogen challenge), followed by sequencing to detect enrichment of guide RNAs suggestive of genes with a role in the trait of interest.

PLEIOTROPIC EFFECTS

In the context of genome editing, the unintended impacts on traits other than the target trait due to a specific edit.

PRIMORDIAL GERM CELLS

The stem cells specified during early development that will differentiate to form male and female gametes, therefore representing the precursors of the germline.

GERMPLASM

In the context of animal breeding, the genetic material of a breeding program.

PARTHENOGENESIS

A form of asexual reproduction in which formation of embryos occurs without fertilization.

HETEROGAMETIC

The existence of sex chromosomes, typically with a difference in sequence and / or morphology. The heterogametic sex is the one containing one copy of each sex chromosome.

MONOSEX PRODUCTION

The production of a single sex population of a farmed animal species, typically due to improved performance of that sex or to avoid precocious maturation.

LINKED READS

Linking together of short sequence reads to provide long range orientation, based on the addition of a unique DNA barcode to each read generated from an individual molecule.

CHROMATIN CONFORMATION CAPTURE

Methods used to analyze the spatial organization of chromatin in a cell by measuring the number of interactions between genomic regions that are nearby in 3-D space, but may be separated by many nucleotides in the linear genome.

OPTICAL MAPPING

A technique for constructing ordered, genome-wide, high-resolution restriction enzyme-based maps from single, stained molecules of DNA.

SINGLE CELL RNA SEQUENCING

Obtaining the gene expression profiles of individual cells, typically via the encapsulation of cells into droplets where each droplet carries a unique DNA barcode that labels the transcripts derived from single cells.

GENOTYPE-PHENOTYPE GAP

The gap in knowledge of how variation at the level of the genome causes an effect on a phenotype of interest.

DOUBLY UNIPARENTAL INHERITANCE

The existence of two distinct mitochondrial DNA lineages within an individual of a species, one of which is inherited through females and the other through males.

SEGREGATION DISTORTION

A phenomenon where the observed genotype frequencies at a locus fall outside the expected Mendelian segregation ratio.

VIABILITY SELECTION

The selection of individual organisms who can survive until a certain lifecycle stage, typically to reproduction.

SELECTION INTENSITY

The number of phenotypic standard deviation units that selected parents are superior to the mean of a population.

GBLUP

Genomic best linear unbiased prediction is a modification of the pedigree-based best linear unbiased prediction method, and incorporates SNP information in the form of a genomic relationship matrix which defines the additive genetic covariance among individuals to predict breeding values.

BAYESIAN MODELS

In the context of genomic selection, the use of multiple-regression methods incorporating prior information on marker effects which are used widely for genomic prediction of breeding values.

Competing interests

The authors have no competing interests.

Acknowledgements

The authors acknowledge funding from the Biotechnology and Biological Sciences Research Council (BBSRC), the Natural Environment Research Council, and the Scottish Aquaculture Innovation Centre via the consortium project 'AquaLeap' (BB/S004343/1, BB/S004181/1, BB/S004416/1, BB/S004300/1), and BBSRC Institute Strategic Programme grants (BBS/E/D/20241866 & BBS/E/D/20002172 & BBS/E/D/20002174).

References

1. Anderson, J. L., Asche, F., Garlock, T. & Chu, J. Aquaculture: Its role in the future of food. *Front. Econ. Glob.* **17**, 159–173 (2017).
2. FAO. *FAO yearbook. Fishery and Aquaculture Statistics 2017*. (FAO, 2019).
3. Longo, S. B., Clark, B., York, R. & Jorgenson, A. K. Aquaculture and the displacement of fisheries captures. *Conserv. Biol.* **33**, cobi.13295 (2019).
4. Froehlich, H. E., Runge, C. A., Gentry, R. R., Gaines, S. D. & Halpern, B. S. Comparative terrestrial feed and land use of an aquaculture-dominant world. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 5295–5300 (2018).
5. Ahmed, N., Thompson, S. & Glaser, M. Global Aquaculture Productivity, Environmental Sustainability, and Climate Change Adaptability. *Environ. Manage.* **63**, 159–172 (2019).
6. Jennings, S. *et al.* Aquatic food security: insights into challenges and solutions from an analysis of interactions between fisheries, aquaculture, food safety, human health, fish and human welfare, economy and environment. *Fish Fish.* **17**, 893–938 (2016).

7. Handisyde, N., Telfer, T. C. & Ross, L. G. Vulnerability of aquaculture-related livelihoods to changing climate at the global scale. *Fish Fish.* **18**, 466–488 (2017).

8. Troell, M. *et al.* Does aquaculture add resilience to the global food system? *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13257–63 (2014).

9. Teletchea, F. Animal Domestication: A Brief Overview. in *Animal Domestication* (IntechOpen, 2019). doi:10.5772/intechopen.86783

10. FAO. *The State of the World's Aquatic Genetic Resources for Food and Agriculture.* (FAO, 2019).

11. Gjedrem, T. & Rye, M. Selection response in fish and shellfish: a review. *Rev. Aquac.* **10**, 168–179 (2018).

12. Hill, W. G. Is continued genetic improvement of livestock sustainable? *Genetics* **202**, 877–81 (2016).

13. Georges, M., Charlier, C. & Hayes, B. Harnessing genomic information for livestock improvement. *Nat. Rev. Genet.* **20**, 135–156 (2019).

14. Abdelrahman, H. *et al.* Aquaculture genomics, genetics and breeding in the United States: current status, challenges, and priorities for future research. *BMC Genomics* **18**, 191 (2017).

15. Harland, J. The origins of aquaculture. *Nat. Ecol. Evol.* 1–2 (2019). doi:10.1038/s41559-019-0966-3

16. Nguyen, N. H. Genetic improvement for important farmed aquaculture species with a reference to carp, tilapia and prawns in Asia: achievements, lessons and challenges. *Fish Fish.* **17**, 483–506 (2016).

17. Gjedrem, T., Robinson, N. & Rye, M. The importance of selective breeding in aquaculture to meet future demands for animal protein: A review. *Aquaculture* **350–353**, 117–129 (2012).

18. Kumar, G., Engle, C. & Tucker, C. Factors driving aquaculture technology adoption. *J. World Aquac. Soc.* **49**, 447–476 (2018).

19. Janssen, K., Chavanne, H., Berentsen, P. & Komen, H. Impact of selective breeding on European aquaculture. *Aquaculture* **472**, 8–16 (2017).

20. Janssen, K., Saatkamp, H. & Komen, H. Cost-benefit analysis of aquaculture breeding programs. *Genet. Sel. Evol.* **50**, 2 (2018).

21. Teletchea, F. & Fontaine, P. Levels of domestication in fish: implications for the sustainable future of aquaculture. *Fish Fish.* **15**, 181–195 (2014).

22. Mignon-Grasteau, S. *et al.* Genetics of adaptation and domestication in livestock. *Livest. Prod. Sci.* **93**, 3–14 (2005).

23. Driscoll, C. A., Macdonald, D. W. & O'Brien, S. J. From wild animals to domestic pets, an evolutionary view of domestication. *Proc. Natl. Acad. Sci. U. S. A.* **106 Suppl**, 9971–8 (2009).

24. Fernández, J., Toro, M. Á., Sonesson, A. K. & Villanueva, B. Optimizing the creation of base populations for aquaculture breeding programs using phenotypic and genomic data and its consequences on genetic progress. *Front. Genet.* **5**, 414 (2014).

25. You, W. & Hedgecock, D. Boom-and-bust production cycles in animal seafood aquaculture. *Rev. Aquac.* 1–16 (2018). doi:10.1111/raq.12278

26. Doyle, R. W. Inbreeding and disease in tropical shrimp aquaculture: a reappraisal and caution. *Aquac. Res.* **47**, 21–35 (2016).

27. Reed, D. H. & Frankham, R. How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution (N. Y).* **55**, 1095–1103 (2001).

28. Bentsen, H. B. *et al.* Genetic improvement of farmed tilapias: growth performance in a complete diallel cross experiment with eight strains of Oreochromis niloticus. *Aquaculture* **160**, 145–173 (1998).

29. Govindaraju, D. R. An elucidation of over a century old enigma in genetics—Heterosis. *PLOS Biol.* **17**, e3000215 (2019).

30. Baird, N. A. *et al.* Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* **3**, e3376 (2008).

31. Campbell, C. R., Poelstra, J. W. & Yoder, A. D. What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biol. J. Linn. Soc.* **124**, 561–583 (2018).

32. Robledo, D., Palaiokostas, C., Bargelloni, L., Martínez, P. & Houston, R. Applications of genotyping by sequencing in aquaculture breeding and genetics. *Rev. Aquac.* **10**, 670–682 (2018).

33. Plough, L. V. Genetic load in marine animals: a review. *Curr. Zool.* **62**, 567–579 (2016).

34. Lien, S. *et al.* The Atlantic salmon genome provides insights into rediploidization. *Nature* **533**, 200–205 (2016).

35. Macqueen, D. J. *et al.* Functional Annotation of All Salmonid Genomes (FAASG): an international initiative supporting future salmonid research, conservation and aquaculture. *BMC Genomics* **18**, 484 (2017).

36. Ashton, D. T., Ritchie, P. A. & Wellenreuther, M. High-density linkage map and QTLs for growth in snapper (Chrysophrys auratus). *G3 (Bethesda).* **9**, 1027–1035 (2019).

37. Wellenreuther, M., Le Luyer, J., Cook, D., Ritchie, P. A. & Bernatchez, L. Domestication and temperature modulate gene expression signatures and growth in the Australasian snapper Chrysophrys auratus. *G3 (Bethesda).* **9**, 105–116 (2019).

38. Ashton, D. T., Hilario, E., Jaksons, P., Ritchie, P. A. & Wellenreuther, M. Genetic diversity and heritability of economically important traits in captive Australasian snapper (Chrysophrys auratus). *Aquaculture* **505**, 190–198 (2019).

39. Devlin, R. H. & Nagahama, Y. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture* **208**, 191–364 (2002).

40. Kobayashi, Y., Nagahama, Y. & Nakamura, M. Diversity and plasticity of sex determination and differentiation in fishes. *Sex Dev.* **7**, 115–25 (2013).

41. Martínez, P. *et al.* Genetic architecture of sex determination in fish: applications to sex ratio control in aquaculture. *Front. Genet.* **5**, 340 (2014).

42. Palaiokostas, C. *et al.* Mapping and validation of the major sex-determining region in Nile tilapia (Oreochromis niloticus L.) Using RAD sequencing. *PLoS One* **8**, e68389 (2013).

43. Palaiokostas, C. *et al.* Mapping the sex determination locus in the Atlantic halibut (Hippoglossus hippoglossus) using RAD sequencing. *BMC Genomics* **14**, 566 (2013).

44. Palaiokostas, C. *et al.* A new SNP-based vision of the genetics of sex determination in European sea bass (Dicentrarchus labrax). *Genet. Sel. Evol.* **47**, 68 (2015).

45. Shi, X. *et al.* Female-specific SNP markers provide insights into a WZ/ZZ sex determination system for mud crabs Scylla paramamosain, S. tranquebarica and S. serrata with a rapid method for genetic sex identification. *BMC Genomics* **19**, 981 (2018).

46. Vandeputte, M. & Haffray, P. Parentage assignment with genomic markers: a major advance for understanding and exploiting genetic variation of quantitative traits in farmed aquatic animals. *Front. Genet.* **5**, 432 (2014).

47. Khang, P. Van, Phuong, T. H., Dat, N. K., Knibb, W. & Nguyen, N. H. An 8-Year Breeding Program for Asian Seabass Lates calcarifer: Genetic Evaluation, Experiences, and Challenges. *Front. Genet.* **9**, 191 (2018).

48. Fuji, K. *et al.* Marker-assisted breeding of a lymphocystis disease-resistant Japanese flounder (Paralichthys olivaceus). *Aquaculture* **272**, 291–295 (2007).

49. Fuji, K. *et al.* Identification of a single major genetic locus controlling the resistance to lymphocystis disease in Japanese flounder (Paralichthys olivaceus). *Aquaculture* **254**, 203–210 (2006).

50. Liu, S. *et al.* Retrospective evaluation of marker-assisted selection for resistance to bacterial cold water disease in three generations of a commercial rainbow trout breeding population. *Front. Genet.* **9**, 286 (2018).

51. Ayllon, F. *et al.* The vgll3 locus controls age at maturity in wild and domesticated Atlantic salmon (Salmo salar L.) males. *PLoS Genet.* **11**, e1005628 (2015).

52. Barson, N. J. *et al.* Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* **528**, 405–408 (2015).

53. Gonen, S. *et al.* Mapping and validation of a major QTL affecting resistance to pancreas disease (salmonid alphavirus) in Atlantic salmon (Salmo salar). *Heredity (Edinb).* **115**, 405–414 (2015).

54. Boison, S. *et al.* QTLs associated with resistance to cardiomyopathy syndrome in Atlantic salmon. *J. Hered.* (2019). doi:10.1093/jhered/esz042

55. Hillestad, B. & Moghadam, H. K. Genome-wide association study of piscine myocarditis virus (PMCV) resistance in Atlantic salmon (Salmo salar). *J. Hered.* (2019). doi:10.1093/jhered/esz040

56. Gu, X. H. *et al.* Identifying a major QTL associated with salinity tolerance in Nile Tilapia using QTL-Seq. *Mar. Biotechnol.* **20**, 98–107 (2018).

57. Houston, R. D. Future directions in breeding for disease resistance in aquaculture species. *Rev. Bras. Zootec.* **46**, 545–551 (2017).

58. Zenger, K. R. *et al.* Genomic selection in aquaculture: application, limitations and opportunities with special reference to marine shrimp and pearl oysters. *Front. Genet.* **9**, 693 (2019).

59. Gjedrem, T. & Baranski, M. *Selective breeding in aquaculture : an introduction*. (Springer Science & Business Media, 2010).

60. Meuwissen, T. H., Hayes, B. J. & Goddard, M. E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–29 (2001).

61. Houston, R. D. *et al.* Development and validation of a high density SNP genotyping array for Atlantic salmon (Salmo salar). *BMC Genomics* **15**, 90 (2014).

62. Odegård, J. *et al.* Genomic prediction in an admixed population of Atlantic salmon (Salmo salar). *Front. Genet.* **5**, 402 (2014).

63. Tsai, H. Y. *et al.* The genetic architecture of growth and fillet traits in farmed Atlantic salmon (Salmo salar). *BMC Genet.* **16**, 51 (2015).

64. Norris, A. Application of genomics in salmon aquaculture breeding programs by Ashie Norris: Who

knows where the genomic revolution will lead us? *Mar. Genomics* **36**, 13–15 (2017).

65. Lillehammer, M., Meuwissen, T. H. E. & Sonesson, A. K. A low-marker density implementation of genomic selection in aquaculture using within-family genomic breeding values. *Genet. Sel. Evol.* **45**, 39 (2013).

66. Daetwyler, H. D., Calus, M. P. L., Pong-Wong, R., de los Campos, G. & Hickey, J. M. Genomic prediction in animals and plants: Simulation of data, validation, reporting, and benchmarking. *Genetics* **193**, 347–365 (2013).

67. Tsai, H. Y. *et al.* Construction and annotation of a high density SNP linkage map of the Atlantic salmon (Salmo salar) genome. *G3 GENES, GENOMES, Genet.* **6**, 2173–2179 (2016).

68. Palaiokostas, C. *et al.* Optimizing genomic prediction of host resistance to koi herpesvirus disease in carp. *Front. Genet.* **10**, 543 (2019).

69. Campbell, N. R., Harmon, S. A. & Narum, S. R. Genotyping-in-Thousands by sequencing (GT-seq): a cost effective SNP genotyping method based on custom amplicon sequencing. *Mol. Ecol. Resour.* **15**, 855–867 (2015).

70. Tsai, H.-Y. *et al.* Genotype imputation to improve the cost-efficiency of genomic selection in farmed Atlantic salmon. *G3 (Bethesda).* **7**, 1377–1383 (2017).

71. Yoshida, G. M. *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* **491**, 147–154 (2018).

72. Daetwyler, H. D., Calus, M. P. L., Pong-Wong, R., de los Campos, G. & Hickey, J. M. Genomic prediction in animals and plants: simulation of data, validation, reporting, and benchmarking. *Genetics* **193**, 347–365 (2013).

73. 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.* **49**, 8 (2017).

74. MacLeod, I. M. *et al.* Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics* **17**, 144 (2016).

75. Giuffra, E., Tuggle, C. K. & FAANG Consortium. Functional Annotation of Animal Genomes (FAANG): current achievements and roadmap. *Annu. Rev. Anim. Biosci.* **7**, 65–88 (2019).

76. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157 (2015).

77. Lafferty, K. D. *et al.* Infectious diseases affect marine fisheries and aquaculture economics. *Ann. Rev. Mar. Sci.* **7**, 471–496 (2015).

78. Asche, F., Hansen, H., Tveteras, R. & Tveteras, S. The salmon disease crisis in Chile. *Mar. Resour. Econ.* **24**, 405–411 (2009).

79. Verbruggen, B. *et al.* Molecular mechanisms of white spot syndrome virus infection and perspectives on treatments. *Viruses* **8**, 23 (2016).

80. Ødegård, J., Baranski, M., Gjerde, B. & Gjedrem, T. Methodology for genetic evaluation of disease resistance in aquaculture species: challenges and future prospects. *Aquac. Res.* **42**, 103–114 (2011).

81. Yáñez, J. M. *et al.* Genetic co-variation between resistance against both Caligus rogercresseyi and Piscirickettsia salmonis, and body weight in Atlantic salmon (Salmo salar). *Aquaculture* **433**, 295–298

(2014).

82. Gjedrem, T. Disease resistant fish and shellfish are within reach: a review. *J. Mar. Sci. Eng.* **3**, 146–153 (2015).

83. Bishop, S. C. & Woolliams, J. A. Genomics and disease resistance studies in livestock. *Livest. Sci.* **166**, 190–198 (2014).

84. Anacleto, O. *et al.* Genetic differences in host infectivity affect disease spread and survival in epidemics. *Sci. Rep.* **9**, 4924 (2019).

85. Saberioon, M., Gholizadeh, A., Cisar, P., Pautsina, A. & Urban, J. Application of machine vision systems in aquaculture with emphasis on fish: state-of-the-art and key issues. *Rev. Aquac.* **9**, 369–387 (2017).

86. Føre, M. *et al.* Precision fish farming: a new framework to improve production in aquaculture. *Biosyst. Eng.* **173**, 176–193 (2018).

87. Liakos, K. *et al.* Machine learning in agriculture: a Review. *Sensors* **18**, 2674 (2018).

88. Badiola, M., Basurko, O. C., Piedrahita, R., Hundley, P. & Mendiola, D. Energy use in Recirculating Aquaculture Systems (RAS): a review. *Aquac. Eng.* **81**, 57–70 (2018).

89. Sae-Lim, P., Gjerde, B., Nielsen, H. M., Mulder, H. & Kause, A. A review of genotype-by-environment interaction and micro-environmental sensitivity in aquaculture species. *Rev. Aquac.* **8**, 369–393 (2016).

90. Nguyen, N. H. Genetic improvement for important farmed aquaculture species with a reference to carp, tilapia and prawns in Asia: achievements, lessons and challenges. *Fish Fish.* **17**, 483–506 (2016).

91. Sae-Lim, P. *et al.* Genetic (co)variance of rainbow trout (Oncorhynchus mykiss) body weight and its uniformity across production environments. *Genet. Sel. Evol.* **47**, 46 (2015).

92. Saltz, J. B. *et al.* Why does the magnitude of genotype-by-environment interaction vary? *Ecol. Evol.* **8**, 6342–6353 (2018).

93. Eknath, A. E. & Acosta, B. O. *Genetic Improvement of Farmed Tilapias (GIFT) project: final report, March 1988–December 1997*. (1998).

94. Mulder, H. A. Genomic Selection Improves Response to Selection in Resilience by Exploiting Genotype by Environment Interactions. *Front. Genet.* **7**, 178 (2016).

95. Brugman, S. *et al.* A comparative review on microbiota manipulation: lessons from fish, plants, livestock, and human research. *Front. Nutr.* **5**, 80 (2018).

96. Derome, N. *Microbial communities in aquaculture ecosystems: improving productivity and sustainability*. (Springer International Publishing, 2019). doi:10.1007/978-3-030-16190-3

97. Smith, C. C. R., Snowberg, L. K., Gregory Caporaso, J., Knight, R. & Bolnick, D. I. Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *ISME J.* **9**, 2515–2526 (2015).

98. Li, W. *et al.* Genetic effects on the gut microbiota assemblages of hybrid fish from parents with different feeding habits. *Front. Microbiol.* **9**, 2972 (2018).

99. Uren Webster, T. M., Consuegra, S., Hitchings, M. & Garcia de Leaniz, C. Interpopulation variation in the Atlantic salmon microbiome reflects environmental and genetic diversity. *Appl. Environ. Microbiol.* **84**, e00691-18 (2018).

100. Cornejo-Granados, F. *et al.* Microbiome of Pacific Whiteleg shrimp reveals differential bacterial community composition between wild, aquacultured and AHPND/EMS outbreak conditions. *Sci. Rep.* **7**, 11783 (2017).

101. Llewellyn, M. S., Boutin, S., Hoseinifar, S. H. & Derome, N. Teleost microbiomes: the state of the art in their characterization, manipulation and importance in aquaculture and fisheries. *Front. Microbiol.* **5**, 207 (2014).

102. Naylor, R. L. *et al.* Feeding aquaculture in an era of finite resources. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15103–10 (2009).

103. Gavery, M. R. & Roberts, S. B. Epigenetic considerations in aquaculture. *PeerJ* **5**, e4147 (2017).

104. Luyer, J. Le *et al.* Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. *Proc. Natl. Acad. Sci.* **114**, 12964–12969 (2017).

105. Jonsson, B. & Jonsson, N. Early environment influences later performance in fishes. *J. Fish Biol.* **85**, 151–188 (2014).

106. Moghadam, H., Mørkøre, T. & Robinson, N. Epigenetics—potential for programming fish for aquaculture? *J. Mar. Sci. Eng.* **3**, 175–192 (2015).

107. Geurden, I. *et al.* Early-feeding exposure to a plant-based diet improves its future acceptance and utilization in rainbow trout. *Commun. Agric. Appl. Biol. Sci.* **78**, 157–60 (2013).

108. Robinson, N. A., Johnsen, H., Moghadam, H., Andersen, Ø. & Tveiten, H. Early developmental stress affects subsequent gene expression response to an acute stress in Atlantic salmon: an approach for creating robust fish for aquaculture? *G3 (Bethesda).* **9**, 1597–1611 (2019).

109. Zhao, L. *et al.* Transgenerational acclimation to seawater acidification in the Manila clam Ruditapes philippinarum: Preferential uptake of metabolic carbon. *Sci. Total Environ.* **627**, 95–103 (2018).

110. Parker, L. M., O'Connor, W. A., Raftos, D. A., Pörtner, H.-O. & Ross, P. M. Persistence of positive carryover effects in the oyster, Saccostrea glomerata, following transgenerational exposure to ocean acidification. *PLoS One* **10**, e0132276 (2015).

111. Lubzens, E., Bobe, J., Young, G. & Sullivan, C. V. Maternal investment in fish oocytes and eggs: The molecular cargo and its contributions to fertility and early development. *Aquaculture* **472**, 107–143 (2017).

112. Jiang, L. *et al.* Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell* **153**, 773–784 (2013).

113. Franěk, R. *et al.* Isogenic lines in fish – a critical review. *Rev. Aquac.* raq.12389 (2019). doi:10.1111/raq.12389

114. Goddard, M. E. & Whitelaw, E. The use of epigenetic phenomena for the improvement of sheep and cattle. *Front. Genet.* **5**, 247 (2014).

115. González-Recio, O., Toro, M. A. & Bach, A. Past, present, and future of epigenetics applied to livestock breeding. *Front. Genet.* **6**, 305 (2015).

116. Wang, G.-D., Xie, H.-B., Peng, M.-S., Irwin, D. & Zhang, Y.-P. Domestication genomics: evidence from animals. *Annu. Rev. Anim. Biosci.* **2**, 65–84 (2014).

117. Lopez Dinamarca, M. E. *et al.* Multiple selection signatures in farmed Atlantic salmon adapted to different environments across Hemispheres. *Front. Genet.* **10**, 901 (2019).

118. López, M. E. *et al.* Comparing genomic signatures of domestication in two Atlantic salmon (Salmo salar L.) populations with different geographical origins. *Evol. Appl.* **12**, 137–156 (2019).

119. Glover, K. A. *et al.* Half a century of genetic interaction between farmed and wild Atlantic salmon: status of knowledge and unanswered questions. *Fish Fish.* **18**, 890–927 (2017).

120. Roberge, C., Einum, S., Guderley, H. & Bernatchez, L. Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. *Mol. Ecol.* **15**, 9–20 (2005).

121. Skaala, Ø. *et al.* An extensive common-garden study with domesticated and wild Atlantic salmon in the wild reveals impact on smolt production and shifts in fitness traits. *Evol. Appl.* **12**, 1001–1016 (2019).

122. Cowen, R. K. & Sponaugle, S. Larval dispersal and marine population connectivity. *Ann. Rev. Mar. Sci.* **1**, 443–466 (2009).

123. Varney, R. L., Watts, J. C. & Wilbur, A. E. Genetic impacts of a commercial aquaculture lease on adjacent oyster populations. *Aquaculture* **491**, 310–320 (2018).

124. Hornick, K. M. & Plough, L. V. Tracking genetic diversity in a large-scale oyster restoration program: effects of hatchery propagation and initial characterization of diversity on restored vs. wild reefs. *Heredity (Edinb).* **123**, 92–105 (2019).

125. Hindar, K., Fleming, I. A., McGinnity, P. & Diserud, O. Genetic and ecological effects of salmon farming on wild salmon: modelling from experimental results. *ICES J. Mar. Sci.* **63**, 1234–1247 (2006).

126. Hansen, M. M., Limborg, M. T., Ferchaud, A.-L. & Pujolar, J.-M. The effects of Medieval dams on genetic divergence and demographic history in brown trout populations. *BMC Evol. Biol.* **14**, 122 (2014).

127. Horreo, J. L. *et al.* Long-term effects of stock transfers: synergistic introgression of allochthonous genomes in salmonids. *J. Fish Biol.* **85**, 292–306 (2014).

128. Heggberget, T. G. *et al.* Interactions between wild and cultured Atlantic salmon: a review of the Norwegian experience. *Fish. Res.* **18**, 123–146 (1993).

129. Naylor, R. *et al.* Fugitive salmon: assessing the risks of escaped fish from net-pen aquaculture. *Bioscience* **55**, 427–437 (2005).

130. O'Flynn, F., McGeachy, S. A., Friars, G. W., Benfey, T. J. & Bailey, J. K. Comparisons of cultured triploid and diploid Atlantic salmon (Salmo salar L.). *ICES J. Mar. Sci.* **54**, 1160–1165 (1997).

131. Piferrer, F. *et al.* Polyploid fish and shellfish: production, biology and applications to aquaculture for performance improvement and genetic containment. *Aquaculture* **293**, 125–156 (2009).

132. Wargelius, A. *et al.* Dnd knockout ablates germ cells and demonstrates germ cell independent sex differentiation in Atlantic salmon. *Sci. Rep.* **6**, 21284 (2016).

133. Adams, S. L., Smith, J. F., Taylor, J., McGowan, L. T. & Tervit, H. R. Cryopreservation of Greenshell™ Mussel (Perna canaliculus) Sperm. in 329–336 (Springer, New York, NY, 2015). doi:10.1007/978-1-4939-2193-5_14

134. Goswami, M., Mishra, A., Ninawe, A., Trudeau, V. & Lakra, W. Bio-banking: an emerging approach for conservation of fish germplasm. *Poultry, Fish. Wildl. Sci.* **4**, (2016).

135. Robles, V. *et al.* Biology of teleost primordial germ cells (PGCs) and spermatogonia: Biotechnological applications. *Aquaculture* **472**, 4–20 (2017).

136. FAO. *Genebank Standards for Plant Genetic Resources for Food and Agriculture*. (2014).

137. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–23 (2013).

138. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–6 (2013).

139. Edvardsen, R. B., Leininger, S., Kleppe, L., Skaftnesmo, K. O. & Wargelius, A. Targeted mutagenesis in Atlantic salmon (Salmo salar L.) using the CRISPR/Cas9 system induces complete knockout individuals in the F0 generation. *PLoS One* **9**, e108622 (2014).

140. Gratacap, R. L., Wargelius, A., Edvardsen, R. B. & Houston, R. D. Potential of genome editing to improve aquaculture breeding and production. *Trends Genet.* **35**, 672–684 (2019).

141. Sato, M. & Nakamura, S. Possible production of genome-edited animals using gene-engineered sperm. in *Gene Editing - Technologies and Applications* (IntechOpen, 2019). doi:10.5772/intechopen.84859

142. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).

143. Gratacap, R. L. *et al.* Efficient CRISPR/Cas9 genome editing in a salmonid fish cell line using a lentivirus delivery system. *bioRxiv* 734442 (2019). doi:10.1101/734442

144. Jenko, J. *et al.* Potential of promotion of alleles by genome editing to improve quantitative traits in livestock breeding programs. *Genet. Sel. Evol.* **47**, 55 (2015).

145. Burkard, C. *et al.* Pigs lacking the scavenger receptor cysteine-rich domain 5 of CD163 are resistant to porcine reproductive and respiratory syndrome virus 1 infection. *J. Virol.* **92**, (2018).

146. Zohar, Y. Endocrinology and fish farming: Aspects in reproduction, growth, and smoltification. *Fish Physiol. Biochem.* **7**, 395–405 (1989).

147. Wong, T.-T. & Zohar, Y. Production of reproductively sterile fish: a mini-review of germ cell elimination technologies. *Gen. Comp. Endocrinol.* **221**, 3–8 (2015).

148. Li, M. *et al.* Efficient and heritable gene targeting in tilapia by CRISPR/Cas9. *Genetics* **197**, 591–599 (2014).

149. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–51 (2013).

150. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–83 (2013).

151. Sanson, K. R. *et al.* Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, (2018).

152. Tait-Burkard, C. *et al.* Livestock 2.0 – genome editing for fitter, healthier, and more productive farmed animals. *Genome Biol.* **19**, 204 (2018).

153. Bruce, A. Genome edited animals: Learning from GM crops? *Transgenic Res.* **26**, 385–398 (2017).

154. Waltz, E. First genetically engineered salmon sold in Canada. *Nature* (2017). doi:10.1038/nature.2017.22116

155. Yoshizaki, G. & Yazawa, R. Application of surrogate broodstock technology in aquaculture. *Fish. Sci.* **85**, 429–437 (2019).

156. Okutsu, T., Shikina, S., Kanno, M., Takeuchi, Y. & Yoshizaki, G. Production of trout offspring from triploid salmon parents. *Science* **317**, 1517 (2007).

157. Sid, H. & Schusser, B. Applications of gene editing in chickens: a new era is on the horizon. *Front. Genet.* **9**, 456 (2018).

158. Shelley, C. & Lovatelli, A. *Mud crab aquaculture –A practical manual*. *FAO Fisheries and Aquaculture Technical Paper No. 567* (FAO, 2011).

159. Xu, S., Zhao, L., Xiao, S. & Gao, T. Whole genome resequencing data for three rockfish species of Sebastes. *Sci. Data* **6**, 97 (2019).

160. Xu, P. *et al.* Genome sequence and genetic diversity of the common carp, Cyprinus carpio. *Nat. Genet.* **46**, 1212–1219 (2014).

161. Ludwig, A., Belfiore, N. M., Pitra, C., Svirsky, V. & Jenneckens, I. Genome duplication events and functional reduction of ploidy levels in sturgeon (Acipenser, Huso and Scaphirhynchus). *Genetics* **158**, 1203–15 (2001).

162. Hollenbeck, C. M. & Johnston, I. A. Genomic tools and selective breeding in molluscs. *Front. Genet.* **9**, 253 (2018).

163. Zhang, X. *et al.* Penaeid shrimp genome provides insights into benthic adaptation and frequent molting. *Nat. Commun.* **10**, 356 (2019).

164. van Dijk, E. L., Jaszczyszyn, Y., Naquin, D. & Thermes, C. The third revolution in sequencing technology. *Trends Genet.* **34**, 666–681 (2018).

165. Feron, R. *et al.* Characterization of a Y-specific duplication/insertion of the anti-Mullerian hormone type II receptor gene based on a chromosome-scale genome assembly of yellow perch, Perca flavescens. *bioRxiv* 717397 (2019). doi:10.1101/717397

166. Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp. Mol. Med.* **50**, 96 (2018).

167. Andersson, L. *et al.* Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project. *Genome Biol.* **16**, 57 (2015).

168. Diamond, J. Evolution, consequences and future of plant and animal domestication. *Nature* **418**, 700–707 (2002).

169. Bernatchez, L. *et al.* Harnessing the Power of Genomics to Secure the Future of Seafood. *Trends in Ecology and Evolution* **32**, 665–680 (2017).

170. Brooker, A. J. *et al.* Sustainable production and use of cleaner fish for the biological control of sea lice: recent advances and current challenges. *Vet. Rec.* **183**, 383–383 (2018).

171. Sveier, H. & Breck, O. Cleaner fish application in Norway. in *Cleaner fish Biology and Aquaculture applications* (ed. Treasurer, J.) 370–385 (5m Publishing Ltd, 2018).

172. Treasurer, J. *Cleaner Fish Biology and Aquaculture Applications*. (5M Publishing Ltd, 2018).

173. Lie, K. K. *et al.* Loss of stomach, loss of appetite? Sequencing of the ballan wrasse (Labrus bergylta) genome and intestinal transcriptomic profiling illuminate the evolution of loss of stomach function in fish. *BMC Genomics* **19**, 186 (2018).

174. Knutsen, T. M. Lumpfish (Cyclopterus lumpus) draft genome assembly. *figshare* (2018). doi:10.6084/m9.figshare.7301546.v1

175. Storset, A., Strand, C., Wetten, M., Kjøglum, S. & Ramstad, A. Response to selection for resistance against infectious pancreatic necrosis in Atlantic salmon (Salmo salar L.). *Aquaculture* **272**, (2007).

176. Houston, R. D. *et al.* Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic salmon (Salmo salar). *Genetics* **178**, 1109–15 (2008).

177. Moen, T., Baranski, M., Sonesson, A. K. & 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* **10**, 368 (2009).

178. Houston, R. D. *et al.* The susceptibility of Atlantic salmon fry to freshwater infectious pancreatic necrosis is largely explained by a major QTL. *Heredity (Edinb).* **105**, 318–327 (2010).

179. Gheyas, A. A. *et al.* Segregation of infectious pancreatic necrosis resistance QTL in the early life cycle of Atlantic Salmon (Salmo salar). *Anim. Genet.* **41**, 531–536 (2010).

180. Houston, R. D. *et al.* Characterisation of QTL-linked and genome-wide restriction site-associated DNA (RAD) markers in farmed Atlantic salmon. *BMC Genomics* **13**, 244 (2012).

181. Moen, T. *et al.* Epithelial cadherin determines resistance to infectious pancreatic necrosis virus in Atlantic salmon. *Genetics* **200**, 1313–26 (2015).

182. Robledo, D. *et al.* Gene expression comparison of resistant and susceptible Atlantic salmon fry challenged with Infectious Pancreatic Necrosis virus reveals a marked contrast in immune response. *BMC Genomics* **17**, 279 (2016).

183. Tsai, H.-Y. *et al.* Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. *BMC Genomics* **16**, 969 (2015).

184. Tsai, H.-Y. *et al.* Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. *Genet. Sel. Evol.* **48**, 47 (2016).

185. Correa, K., Bangera, R., Figueroa, R., Lhorente, J. P. & Yáñez, J. M. 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.* **49**, 15 (2017).

186. Robledo, D., Matika, O., Hamilton, A. & Houston, R. D. Genome-Wide Association and Genomic Selection for Resistance to Amoebic Gill Disease in Atlantic Salmon. *G3 (Bethesda).* **8**, 1195–1203 (2018).

187. Boison, S. A., Gjerde, B., Hillestad, B., Makvandi-Nejad, S. & Moghadam, H. K. Genomic and Transcriptomic Analysis of Amoebic Gill Disease Resistance in Atlantic Salmon (Salmo salar L.). *Front. Genet.* **10**, 68 (2019).

188. Bangera, R., Correa, K., Lhorente, J. P., Figueroa, R. & Yáñez, J. M. Genomic predictions can accelerate selection for resistance against Piscirickettsia salmonis in Atlantic salmon (Salmo salar). *BMC Genomics* **18**, 121 (2017).

189. Horn, S. S., Meuwissen, T. H. E., Moghadam, H., Hillestad, B. & Sonesson, A. K. Accuracy of selection for omega-3 fatty acid content in Atlantic salmon fillets. *Aquaculture* 734767 (2019). doi:10.1016/j.aquaculture.2019.734767

190. Vallejo, R. L. *et al.* Accurate genomic predictions for BCWD resistance in rainbow trout are achieved using low-density SNP panels: Evidence that long-range LD is a major contributing factor. *J. Anim. Breed. Genet.* **135**, 263–274 (2018).

191. Vallejo, R. L. *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.* **49**, 17 (2017).

192. Vallejo, R. L. *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.* **7**, 96 (2016).

193. Yoshida, G. M., Carvalheiro, R., Rodríguez, F. H., Lhorente, J. P. & Yáñez, J. M. Single-step genomic evaluation improves accuracy of breeding value predictions for resistance to infectious pancreatic necrosis virus in rainbow trout. *Genomics* **111**, 127–132 (2019).

194. Yoshida, G. M. *et al.* Genomic Prediction Accuracy for Resistance Against Piscirickettsia salmonis in Farmed Rainbow Trout. *G3 (Bethesda).* **8**, 719–726 (2018).

195. Vallejo, R. L. *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.* **51**, 47 (2019).

196. Silva, R. M. O. *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.* **51**, 42 (2019).

197. Barría, A. *et al.* Genomic predictions and genome-wide association study of resistance against piscirickettsia salmonis in coho salmon (Oncorhynchus kisutch) using ddRAD sequencing. *G3 (Bethesda).* **8**, 1183–1194 (2018).

198. Palaiokostas, C., Kocour, M., Prchal, M. & Houston, R. D. Accuracy of genomic evaluations of juvenile growth rate in common carp (Cyprinus carpio) using genotyping by sequencing. *Front. Genet.* **9**, 82 (2018).

199. Yoshida, G. M. *et al.* Genome-wide association study and cost-efficient genomic predictions for growth and fillet yield in Nile tilapia (Oreochromis niloticus). *G3 Genes, Genomes, Genet.* **9**, 2597–2607 (2019).

200. Joshi, R., Skaarud, A., de Vera, M., Alvarez, A. T. & Ødegård, J. Genomic prediction for commercial traits using univariate and multivariate approaches in Nile tilapia (Oreochromis niloticus). *Aquaculture* **516**, 734641 (2020).

201. Palaiokostas, C. *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.* **50**, 30 (2018).

202. Palaiokostas, C., Ferraresso, S., Franch, R., Houston, R. D. & Bargelloni, L. Genomic prediction of resistance to pasteurellosis in gilthead sea bream (Sparus aurata) using 2b-RAD sequencing. *G3: Genes|Genomes|Genetics* **6**, 3693–3700 (2016).

203. Aslam, M. L. *et al.* Genetics of resistance to photobacteriosis in gilthead sea bream (Sparus aurata) using 2b-RAD sequencing. *BMC Genet.* **19**, 43 (2018).

204. Saura, M. *et al.* Disentangling genetic variation for resistance and endurance to scuticociliatosis in turbot using pedigree and genomic information. *Front. Genet.* **10**, 539 (2019).

205. Liu, Y. *et al.* Genomic Selection Using BayesCπ and GBLUP for Resistance Against Edwardsiella tarda in Japanese Flounder (Paralichthys olivaceus). *Mar. Biotechnol.* **20**, 559–565 (2018).

206. Garcia, A. L. S. *et al.* Development of genomic predictions for harvest and carcass weight in channel catfish. *Genet. Sel. Evol.* **50**, 66 (2018).

207. Dong, L., Xiao, S., Wang, Q. & Wang, Z. Comparative analysis of the GBLUP, emBayesB, and GWAS algorithms to predict genetic values in large yellow croaker (Larimichthys crocea). *BMC Genomics* **17**,

460 (2016).

208. Nguyen, N. H., Premachandra, H. K. A., Kilian, A. & Knibb, W. Genomic prediction using DArT-Seq technology for yellowtail kingfish Seriola lalandi. *BMC Genomics* **19**, 107 (2018).

209. Liu, G. *et al.* Evaluation of genomic selection for seven economic traits in yellow drum (Nibea albiflora). *Mar. Biotechnol.* **21**, 806–812 (2019).

210. Gutierrez, A. P., Matika, O., Bean, T. P. & Houston, R. D. Genomic Selection for Growth Traits in Pacific Oyster (Crassostrea gigas): Potential of Low-Density Marker Panels for Breeding Value Prediction. *Front. Genet.* **9**, (2018).

211. Gutierrez, A. P. *et al.* Potential of genomic selection for improvement of resistance to Ostreid Herpes virus in Pacific oyster (Crassostrea gigas). *bioRxiv* 754473 (2019). doi:10.1101/754473

212. Dou, J. *et al.* Evaluation of the 2b-RAD method for genomic selection in scallop breeding. *Sci. Rep.* **6**, 19244 (2016).

213. Wang, Y. *et al.* Predicting growth traits with genomic selection methods in Zhikong scallop (Chlamys farreri). *Mar. Biotechnol.* **20**, 769–779 (2018).

214. Wang, Q. *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.* **18**, 45 (2017).

215. Wang, Q., Yu, Y., Li, F., Zhang, X. & Xiang, J. Predictive ability of genomic selection models for breeding value estimation on growth traits of Pacific white shrimp Litopenaeus vannamei. *Chinese J. Oceanol. Limnol.* **35**, 1221–1229 (2017).

216. Wang, Q. *et al.* Evaluation on the genomic selection in Litopenaeus vannamei for the resistance against Vibrio parahaemolyticus. *Aquaculture* **505**, 212–216 (2019).

217. Nguyen, N. H., Phuthaworn, C. & Knibb, W. Genomic prediction for disease resistance to Hepatopancreatic parvovirus and growth, carcass and quality traits in Banana shrimp Fenneropenaeus merguiensis. *Genomics* (2019). doi:10.1016/J.YGENO.2019.11.014

218. Zhang, X. *et al.* The sea cucumber genome provides insights into morphological evolution and visceral regeneration. *PLOS Biol.* **15**, e2003790 (2017).

219. Hughes, L. C. *et al.* Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proc. Natl. Acad. Sci.* **115**, 6249–6254 (2018).

220. Wanninger, A. & Wollesen, T. The evolution of molluscs. *Biol. Rev. Camb. Philos. Soc.* **94**, 102 (2018).

221. Wolfe, J. M. *et al.* A phylogenomic framework, evolutionary timeline and genomic resources for comparative studies of decapod crustaceans. *Proc. R. Soc. B Biol. Sci.* **286**, 20190079 (2019).

222. Dohrmann, M. & Wörheide, G. Dating early animal evolution using phylogenomic data. *Sci. Rep.* **7**, 3599 (2017).

223. Plazzi, F. & Passamonti, M. Towards a molecular phylogeny of mollusks: bivalves' early evolution as revealed by mitochondrial genes. *Mol. Phylogenet. Evol.* **57**, 641–657 (2010).

224. Kumar, S., Stecher, G., Suleski, M. & Hedges, S. B. TimeTree: a resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017).

A

Livestock
Duck and geese
Chicken and turkey
Cattle, sheep, goat
Pig
Camels
Horses
Rabbits and hares
Finfish
Mandarin fish
Largemouth black bass
Tilapia
Asian swamp eel
Salmonids
Carps
Catfishes
Milkfish
Japanese eel
Echinoderms
Sea cucumbers
Crustaceans
Penaeid shrimps
Palaemonid shrimps
Chinese mitton crab
Mud crabs
Red swamp crawfish
Molluscans
Blood cockle
Oysters
Mussels
Venus clams
Constricted tagelus
800
600
400
200
0
Million years ago

B

% global meat production
Pig
Chicken
Cattle
Sheep
Turkey
Goat
% global live seafood weight
Cupped oysters
Grass carp
Silver carp
Whiteleg shrimp
Japanese carpet shell
Nile tilapia
Common carp
Bighead carp
Pangasius catfish
Catla
Atlantic salmon
Scallops
Roho labeo
Mytilid mussels
Milkfish
Red swamp crawfish
Constricted tagelus
Wuchang bream
Rainbow trout
Chinese mitten crab
Giant tiger prawn
Black carp
Mrigal carp
Snakehead
Yellow catfish
0 - 50 %
10,000
1,000
100
0
Time of first farming - Years Before Present

C

Oceania
Africa
Americas
Europe
Asia
Crustaceans
Echinoderms and others
Molluscans
Finfish
0
40,000
80,000
120,000
160,000
200,000
Value in million USD

Family 1 x

Family 2 x

Family 3 x

... Family n x

Selection candidates

Full-sibs of the selection candidates

Genotyping

Phenotyping

Selected breeders

Prediction equation

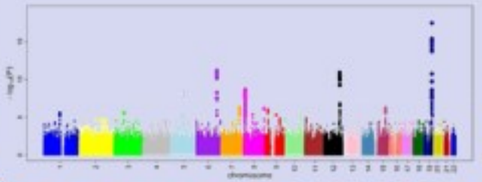
$gEBV = w_1x_1 + w_2x_2 + ... + w_nx_n$

Similarity matrix

Mapping and understanding QTL

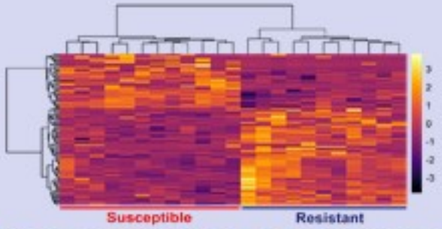
Population-scale GWAS

Association analyses to detect QTL in the population



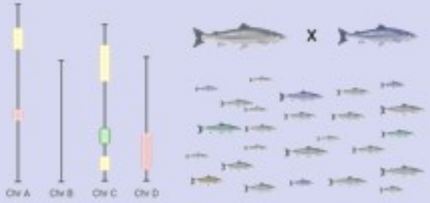
Functional genomics of alternate genotypes

I.e. gene expression comparison using RNA-Seq



Family QTL mapping

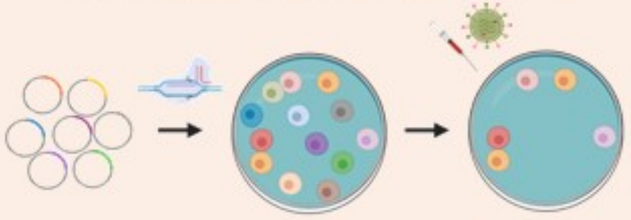
Detect low frequency or small effect QTL



Pooled CRISPR screens

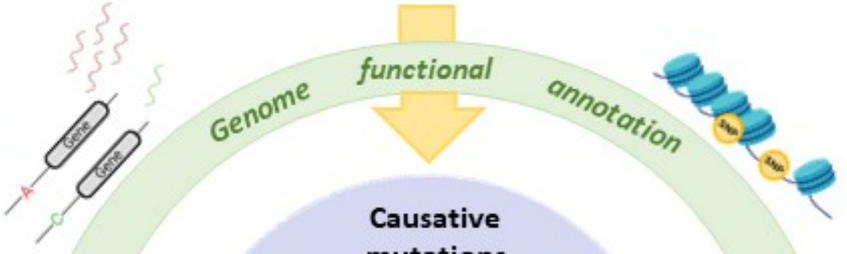
CRISPR/Cas9 in vitro genetic screens

Genome-wide editing in cell cultures to study gene function



In vivo adaptation of CRISPR screens

Genome-wide editing in embryos to study gene function



Genome *functional* annotation

Causative mutations

In vivo validation using genome editing

De novo variation

Cross-species variation

Enhance aquaculture stocks

Functionally-prioritised genomic selection

Genome editing in broodstock

Comparative genomics

Cross-species comparisons

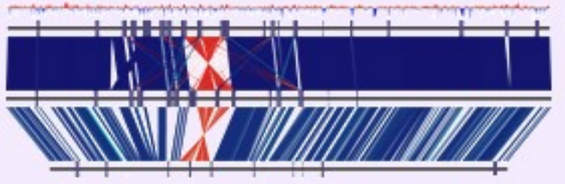
Closely-related species with divergent phenotypes

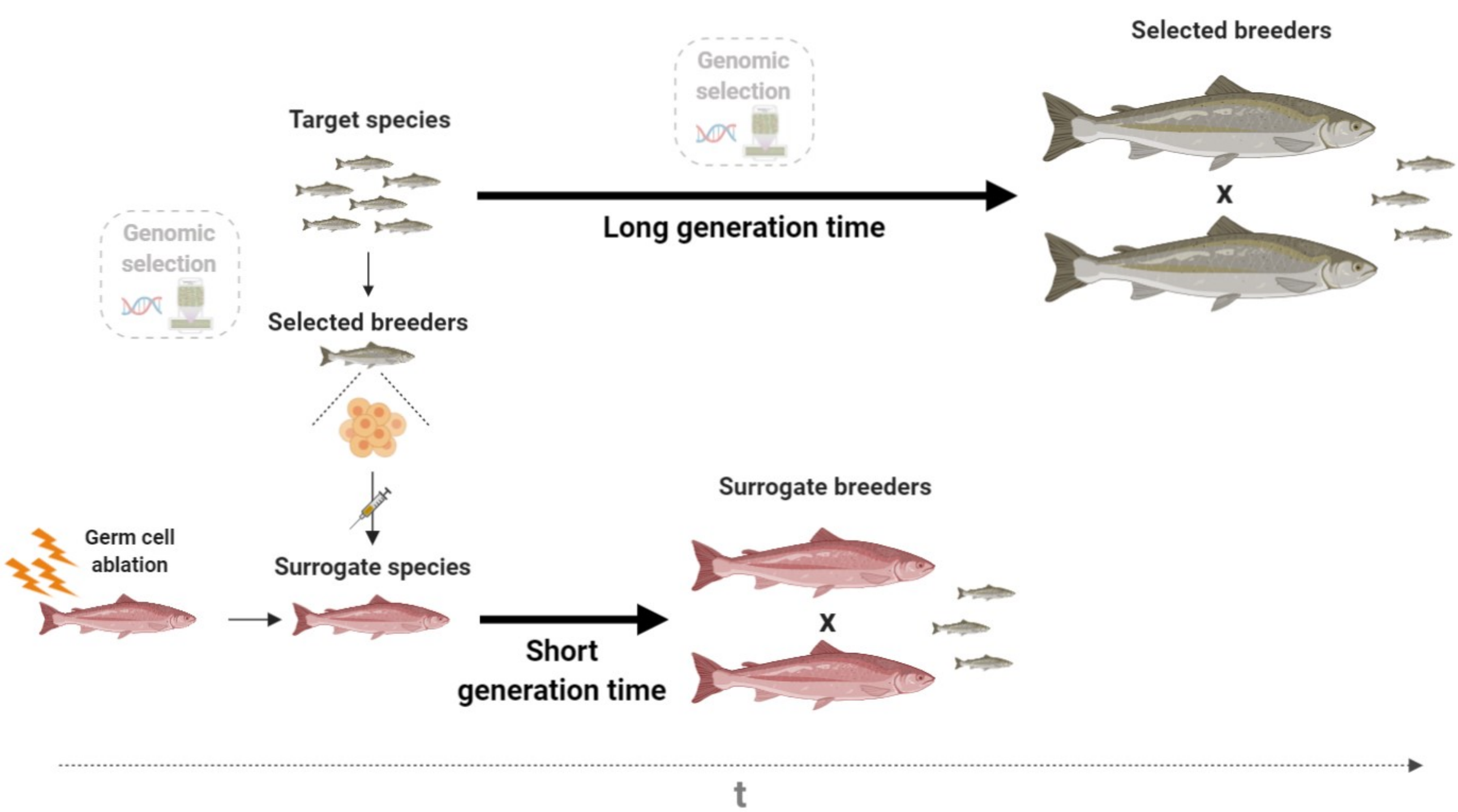
Atlantic salmon
Sea lice susceptible

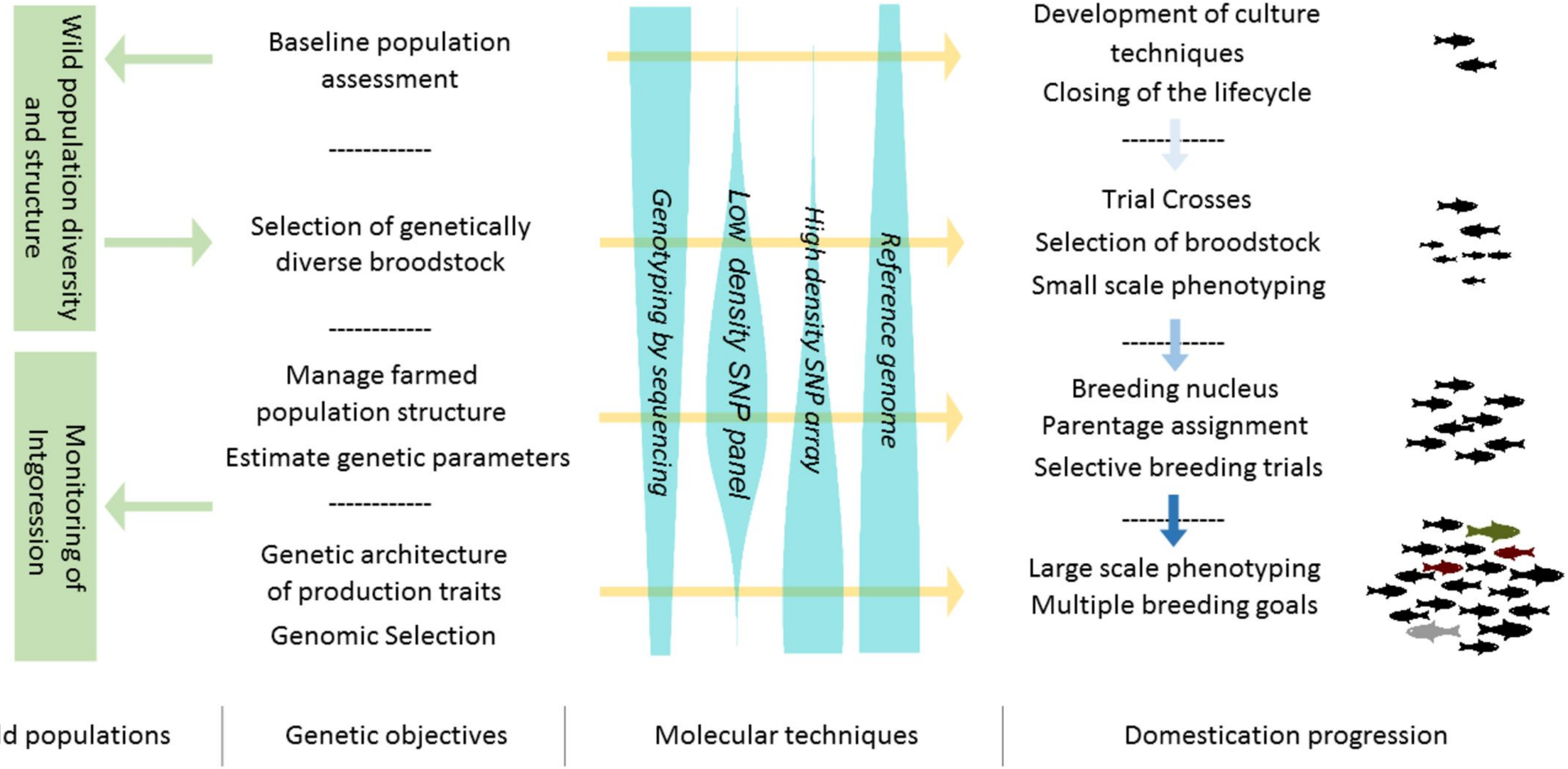
Coho salmon
Sea lice resistant

Distinct genomic features

Associate phenotypic differences to specific genomic differences




Selected breeders
Genomic selection
Target species
Genomic selection
Long generation time
X
Selected breeders
Germ cell ablation
Surrogate species
Surrogate breeders
X
Short generation time
t


Wild population diversity and structure
Monitoring of Intgoression
Baseline population assessment
Selection of genetically diverse broodstock
Manage farmed population structure
Estimate genetic parameters
Genetic architecture of production traits
Genomic Selection
Genotyping by sequencing
Low density SNP panel
High density SNP array
Reference genome
Development of culture techniques
Closing of the lifecycle
Trial Crosses
Selection of broodstock
Small scale phenotyping
Breeding nucleus
Parentage assignment
Selective breeding trials
Large scale phenotyping
Multiple breeding goals
Wild populations
Genetic objectives
Molecular techniques
Domestication progression

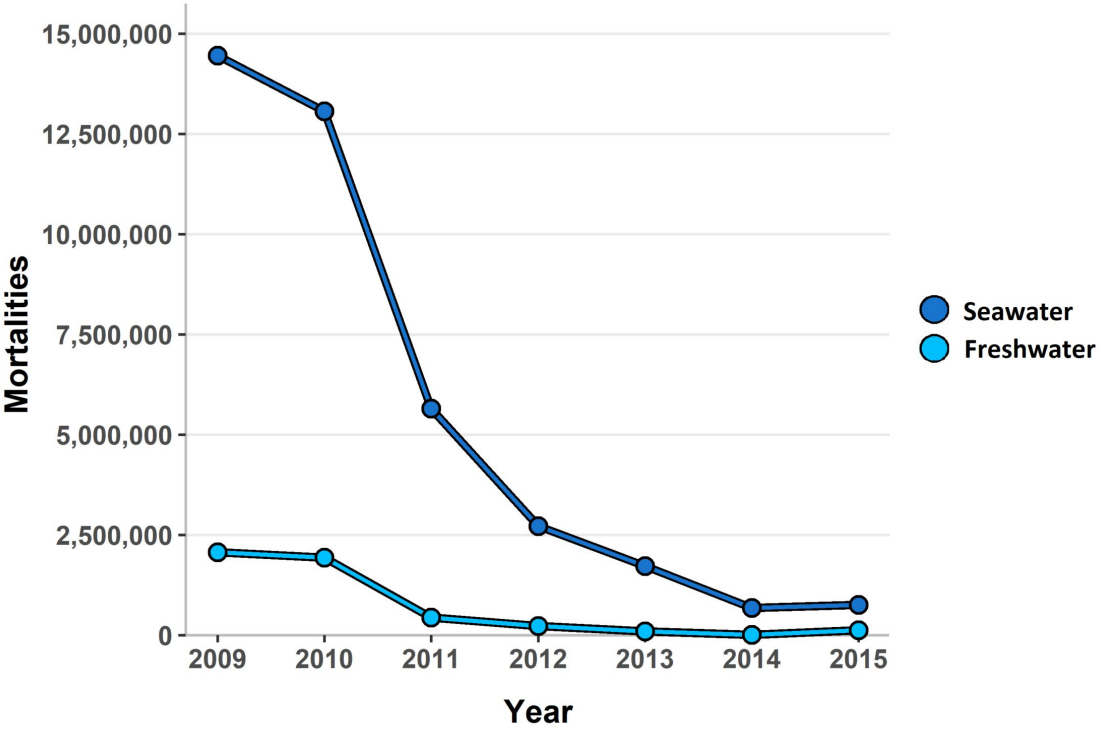

15,000,000
12,500,000
10,000,000
7,500,000
5,000,000
2,500,000
0
Mortalities
2009
2010
2011
2012
2013
2014
2015
Year
Seawater
Freshwater

Table 1 | Summary of studies testing genomic prediction for production traits in aquaculture species, including comparing prediction accuracy using genomic and pedigree-based approaches.

Species	Trait	Measurement	Heritability (pedigree)	Accuracy (pedigree)	Relative increase	Genotyping technology	Ref
Atlantic salmon (Salmo salar)	Growth	Weight	0.60 (0.48)	0.70 (0.58)	21%	SNP array (132K, 112K post-filtering)	183
		Length	0.61 (0.51)	0.66 (0.56)	18%		183
	Resistance to sea lice	Lice count	0.33 (0.27)	0.60 (0.48)	25%	SNP array (132K, 33K post-filtering)	184
		Lice count	0.22(0.27)	0.46 (0.43)	7%		184
		Lice count	0.11 (0.10)	0.50 (0.41)	22%	SNP array (50K, 37K post-filtering)	185
		Log lice density	(0.14)	0.52 (0.34)	52%	SNP array (220K)	62
	Resistance to amoebic gill disease	Gill score	0.24 (0.25)	0.62 (0.51)	22%	Two species SNP array (17K, 7K post-filtering)	186
		Amoebic load	0.25 (0.36)	0.70 (0.60)	17%		186
		Gill score	0.28 (0.32)	0.72 (0.61)	18%	SNP array (55K, 53K post-filtering)	187
	Resistance to salmon rickettsial syndrome	Time to death	0.27 (0.18)	0.41* (0.34)	21%	SNP array (50K, 50K post-filtering)	188
		Binary survival	0.39 (0.26)	0.26 (0.20)	30%		188
	Fillet pigmentation	-	(0.43)	0.44 (0.36)	22%	SNP array (220K)	62
	Muscle fat	-	0.25 (0.36)	0.56 (0.60)	-7%	SNP array (57K, 50K post-filtering)	189
	Omega-3 fatty acid content	DHA	0.20 (0.21)	0.41 (0.33)	24%		189
		EPA	0.04 (0.06)	0.32 (0.37)	-14%		189
Rainbow trout (Oncorhynchus mykiss)	Resistance to bacterial cold water disease	Binary survival	–	0.68* (0.36)	89%	SNP array (57K, 45K post-filtering)	190
		Time to death	0.33 (0.37)	0.67* (0.34)	97%	SNP array (57K, 36K post-filtering)	191
		Binary survival	0.35 (0.35)	0.70* (0.36)	94%		191
		Time to death	0.29 (0.31)	0.49 (0.50)	-2%	SNP array (57K, 41K post-filtering)	192
		Binary survival	0.45 (0.48)	0.46 (0.41)	12%		192
	Resistance to infectious pancreatic necrosis virus	Time to death	0.25 (0.40)	0.53 (0.49)	8%	SNP array (57K, 38K post-filtering)	193
		Binary survival	0.24 (0.35)	0.56 (0.50)	12%		193
	Resistance to salmon rickettsial syndrome	Time to death	0.45 (0.38)	0.78* (0.61)	28%	SNP array (57K, 27K post-filtering)	194
		Binary survival	0.55 (0.54)	0.60* (0.47)	28%		194
	Resistance to Infectious hematopoietic necrosis virus	Time to death	0.23 (0.33)	0.33 (0.13)	154%	SNP array (57K, 35K post-filtering)	195
		Binary survival	0.25 (0.28)	0.39 (0.24)	63%		195
	Resistance to columnaris	Binary survival	0.32 (–)	0.11 (-0.02)	-650%	SNP array (57K, 36K post-filtering)	196

	disease	Binary survival	0.51 (–)	0.22 (0.06)	267%	SNP array (57K, 34K post-filtering)	[196]
Coho salmon (Oncorhynchus kisutch)	Resistance to salmon rickettsial syndrome	Time to death	– (0.14)	0.52 (0.27)	93%	ddRAD (9K)	[197]
		Binary survival	– (0.27)	0.81 (0.31)	161%		[197]
Carp (Cyprinus carpio)	Growth	Length	0.33 (0.33)	0.71 (0.60)	18%	RAD-seq (20K)	[198]
	Resistance to koi herpesvirus	Binary survival	0.50 (0.61)	0.53* (0.49)	8%	RAD-seq (16K)	[68]
Nile tilapia (Oreochromis niloticus)	Growth	Harvest weight	0.36 (0.31)	0.60 (0.48)	25%	SNP array (43K, 32K post-filtering)	[199]
		Fillet yield	0.21 (0.21)	0.62 (0.54)	15%		[199]
		Harvest weight	0.17 (0.22)	0.29 (0.19)	53%	SNP array (59K, 48K post-filtering)	[200]
		Fillet weight	0.16 (0.24)	0.34 (0.18)	89%		[200]
		Fillet yield	0.23 (0.33)	0.54 (0.46)	17%		[200]
European sea bass (Dicentrarchus labrax)	Resistance to viral nervous necrosis	Binary survival	0.43 (0.27)	0.62* (0.67)	-7%	RAD-seq (9K)	[201]
Gilthead sea bream (Sparus aurata)	Resistance to pasteurellosis	Time to death	0.28 (0.22)	0.44* (0.30)	47%	2b-RAD (22K)	[202]
	Resistance to pasteurellosis	Time to death	0.32 (0.32)	0.54* (0.45)	20%	2b-RAD (28K)	[203]
		Binary survival	0.33 (0.31)	0.56* (0.46)	22%		[203]
Turbot (Scophthalmus maximus)	Resistance to Scuticociliatosis	Resilience	0.15 (–)	0.46 (0.41)	12%	2b-RAD (18K)	[204]
		Resistance	0.26 (–)	–	–		[204]
		Endurance	0.12 (–)	–	–		[204]
Japanese Flounder (Paralichthys olivaceus)	Resistance to Edwardsiella tarda	Binary survival	– (–)	0.603 (–)	–	WGS (1.9M)	[205]
Channel catfish (Ictalurus punctatus)	Growth	Harvest weight	0.27 (–)	0.37 (0.29)	28%	SNP array (660K, 55K post-filtering)	[206]
		Residual carcass weight	0.34 (–)	0.31 (0.24)	29%		[206]
Large yellow croaker (Larimichthys crocea)	Growth	Body weight	0.60 (–)	0.41 (–)	–	ddRAD (30K)	[207]
		Body length	0.59 (–)	0.40 (–)	–		[207]
	n-3HUFA	–	0.44 (–)	0.30 (–)	–	ddRAD (32K)	[207]
Yellowtail kingfish (Seriola lalandi)	Growth	Weight	0.47 (0.42)	0.69 (–)	–	DArT-Seq (14K)	[208]
		Length	0.43 (0.42)	0.67 (–)	–		[208]
		Condition index	0.21 (0.11)	0.44 (–)	–		[208]

Yellow drum (Nibea albiflora)	Growth	Body length	– (–)	0.38* (–)	–	GBS (54K)	209
		Swimming bladder index	– (–)	0.17* (–)	–		209
		Swimming bladder weight	– (–)	0.22* (–)	–		209
		Body thickness	– (–)	0.24* (–)	–		209
		Body height	– (–)	0.30* (–)	–		209
		Body length / body height ratio	– (–)	0.36* (–)	–		209
		Gonad weight index	– (–)	0.37* (–)	–		209
Oyster (Crassostrea gigas)	Growth	Shell length	0.26 (0.23)	0.54 (0.44)	23%	Two species SNP array (38K, 23K post-filtering)	210
		Shell height	0.23 (0.20)	0.60 (0.47)	28%		210
		Wet weight	0.35 (0.31)	0.67 (0.54)	24%		210
	Resistance to Osterid Herpesvirus	Binary survival	0.37 (0.25)	0.76 (0.64)	19%		211
Yesso scallop (Patinopecten yessoensis)	Growth	Shell height	0.48 (–)	0.53 (–)	–	2b-RAD (2K)	212
		Shell length	0.48 (–)	0.46 (–)	–		212
		Shell width	0.36 (–)	0.55 (–)	–		212
Zhikong scallop (Chlamys farreri)	Growth	Shell length	0.42 (–)	0.65* (–)	–	2b-RAD (31K)	213
		Shell height	0.47 (–)	0.70* (–)	–		213
		Shell width	0.54 (–)	0.63* (–)	–		213
		Whole weight	0.28 (–)	0.64* (–)	–		213
Whiteleg shrimp (Litopenaeus vannamei)	Growth	Body weight	0.32 (–)	0.62 (–)	–	2b-RAD (23K)	214
		Body length	0.45 (–)	0.61 (–)	–		214
		Body length	– (–)	0.30* (–)	–	SLAF-seq (6K)	215
		Body weight	– (–)	0.41* (–)	–		215
	Resistance to AHPND	Time to death	0.26 (0.24)	0.50 (0.47)	6%	2b-RAD (23K)	216
		Binary survival	0.16 (0.15)	0.21 (0.20)	5%		216
Banana shrimp (Fenneropenaeus merguiensis)	Growth	Body weight	0.55	0.76 (0.65)	17%	DArT-Seq (9K)	217
		Body length	0.49	0.73 (0.60)	22%		217
		Head length	0.39	0.42 (0.32)	31%		217

	Body width	0.61	0.72 (0.60)	20%	[217]
	Tail weight	0.45	0.77 (0.66)	17%	[217]
	Meat yield	0.10	–	–	[217]
Colour	Dark (raw shrimp)	0.18	0.59 (0.53)	11%	[217]
	Red (cooked shrimp)	0	NA	–	[217]
'Flesh streaks'	–	0	NA	–	[217]
Yellow hepatopancreas	–	0.03	NA	–	[217]
Resistance to HPV	Viral load	0.35	0.60 (0.09)	567%	[217]

* Alternative statistical models to GBLUP were used, e.g. Bayesian models or RRBLUP