Aquaculture
November 2023, Volume 576 Pages 739878 (13p.)
https://doi.org/10.1016/j.aquaculture.2023.739878
https://archimer.ifremer.fr/doc/00846/95839/

Potential of genomic selection for growth, meat content and colour traits in mixed-family breeding designs for the Pacific oyster *Crassostrea gigas*

Jourdan Antoine [1, 2], Morvezen Romain [1, *], Enez Florian [1], Haffray Pierrick [1], Lange Adeline [3], Vétois Emilie [4], Allal Francois [5], Phocas Florence [6], Bugeon Jérôme [7], Dégremont Lionel [2], Boudry Pierre [8]

[1] SYSAAF, Station LPGP/INRAE, Campus de Beaulieu, Rennes 35042, France
[2] Ifremer, RBE, ASIM, Avenue lede Mus de Loup – Ronce-les-Bains, La Tremblade 17390, France
[3] France Naissain, Bouin 85230, France
[4] SATMAR, Gatteville 50760, France
[5] MARBEC, Université Montpellier, CNRS, Ifremer, IRD, Palavas-les-Flots, France
[6] Université Paris-Saclay, INRAE, AgroParisTech, GABI, Jouy-en-Josas 78350, France
[7] INRAE, LPGP, Rennes 35000, France
[8] Ifremer, Département Ressources Biologiques et Environnement, Plouzané 29280, France

* Corresponding author : Romain Morvezen, email address : romain.morvezen@inrae.fr

Abstract :

Selective breeding programs have been initiated worldwide for the Pacific oyster Crassostrea gigas to improve economically important traits such as growth and disease resistance. The emergence of genomic tools has allowed novel perspectives for using genomic selection (GS) in mixed-family breeding designs, which are cheaper and easier to develop than classical breeding schemes. In this study, we evaluated the potential of GS for different growth-related and shell colour traits in two independent commercially selected populations (P1 and P2), based on mixed-family designs. From ≈14.5K informative SNPs genotyped with the bi-species Axiom Affymetrix 57K oyster array, ≈12.5K were mapped on the reference genome. A strong heterogeneity of marker density between and within chromosomes was reported, with a low linkage disequilibrium (below 0.1 at 0.1 Mb) between pairs of SNPs. The within-population structure was homogenous for each population, with effective sizes of 107 for P1 and 76 for P2. Heritability was estimated for each trait and population and ranged from 0.08 ± 0.04 (for mean darkness intensity in P1) to 0.56 ± 0.08 (for the mean upper valve b* value in P2) for a pedigree-based model and from 0.04 ± 0.02 (for mean darkness intensity in P1) to 0.69 ± 0.04 (for the mean darkness intensity in P2) for a genomic-based model. Growth-related traits were generally highly genetically and positively correlated with each other, but weakly correlated with colour traits. Accuracy of prediction was generally higher with the genomic model (GBLUP) than with the classical BLUP model, with a maximum gain of accuracy (from 0.38 to 0.66) for flesh weight adjusted by total weight in P2. Accuracy of breeding values was slightly higher for colour traits for P2, with higher heritability estimates. Overall, our results indicate that GS has a good potential to be implemented in mixed-family breeding programs in a shellfish such as C. gigas.

Highlights

► Genomic selection is suitable for commercial mixed-family oyster breeding programs. ► Although the SNP array quality was limited it was sufficient for genomic selection. ► Growth-related traits were moderately heritable. ► There were weak correlations between growth-related traits and colour traits. ► Prediction accuracy was higher with the genomic model than with the pedigree model.

Keywords : Mollusc, Aquaculture, Genomic selection, Prediction accuracy, Breeding program, Linkage disequilibrium

Introduction

Oyster farming is a major aquaculture activity worldwide. The Pacific oyster (*Crassostrea gigas,* Thunberg 1793) has been introduced into numerous countries and is the most widely cultivated oyster species, with a global annual production estimated at 610 k tons in 2020 (FAO, 2022). Its production can be based either on the collection of juveniles (known as seed or spat) from the wild or on hatchery-produced individuals. Reproduction in hatcheries makes it possible to produce sterile triploids and improved seed using selective breeding. Considering its socioeconomic added value, Pacific oyster selective breeding programs are being encouraged internationally (Nascimento-Schulze et al., 2021; Robert and Gérard, 1999) to improve commercially important traits, such as body weight, survival and meat yield, based on morphometric predictors. Selective breeding programs were first successfully applied on traits easy to measure on individuals that are candidates for selection or on families for group traits (de Melo et al., 2016; Evans and Langdon, 2006; Langdon et al., 2003; Li et al., 2011).

Mass selection has been carried out since the mid-1990s in different oyster species, including European flat oyster *O. edulis* (Naciri-Graven et al., 1998), Sydney rock oyster *Saccostrea glomerata* (Nell et al., 2000), Pacific oyster (Li et al., 2011), American oyster *Crassostrea virginica* (Allen et al., 2021) and Iwagaki oyster *C. nippona* (Y. Hu et al., 2022). It has been widely used to mitigate the mortality caused by Ostreid Herpesvirus 1 (OsHV-1) on *C. gigas* (Dégremont et al., 2015), notably in France. Although effective, mass selection may quickly lead to inbreeding if genetic diversity is not properly monitored, however (Y. Hu et al., 2022; Launey et al., 2001).

As an alternative strategy to individual selection, pedigree-based selective breeding programs have been initiated in several oyster species to estimate breeding values by combining phenotypic information and pedigree records. These programs can be conducted using separate family designs (Evans et al., 2009; Evans and Langdon, 2006; Langdon et al., 2003) or, more recently, mixed-family designs (Vandeputte and Haffray, 2014). The first approach requires a large financial investment in infrastructure and logistic organisation to breed and rear a sufficient number of families before any economic benefit is seen. Common environmental effects (e.g. a tank effect) may also reduce the reliability of estimated breeding values (EBV) of selection candidates, and lead to overestimations of heritability values and expected genetic gains (Gjerde et al., 1996; Hollenbeck and Johnston, 2018; Kong et al., 2015). The second approach involves grouping all families in a single rearing batch and then assigning offspring to their parents using molecular markers. This has already been tested and implemented in *C. gigas* using microsatellites or SNPs (Kong et al., 2015; Lapègue et al., 2014; Matson, 2011; Wan et al., 2020, 2017). This alternative does not require specific facilities and should limit environmental variance and allow a more accurate estimate to be made of their breeding values. The mixed-family approach is more advanced in fish and crustaceans than in bivalves, where only the genetic parameters have so far been estimated (Enez et al., 2018; Kong et al., 2015; Smits et al., 2020). To our knowledge, the suitability of this approach for commercial-scale Pacific oyster breeding programs has not yet been evaluated.

In recent years, a revolution in genetic selection in livestock and crop species has taken place based on the availability of high-density genomic information (Meuwissen et al., 2016). Next-generation sequencing (NGS) and genotyping-by-sequencing (GBS) tools have been developed for different oyster species (McCarty et al., 2022; Vu et al., 2021; Yang et al., 2022). However, they do not provide the same set of markers from one population to another (e.g. between training population and breeding population) and are dependent on DNA quality, which limits their potential to develop repeatable genomic analyses. Alternatively, SNP arrays have been developed for some bivalve species such as the silver-lip pearl oyster *Pinctada maxima*, with an Illumina 3 k iSelect custom array (Jones et al., 2017),

or in the Pacific oyster, with a high-density SNP 190K SNP array (Qi et al., 2017) or the medium bi-species (Pacific oyster and European flat oyster) 57K SNP array (Gutierrez et al., 2017). SNP arrays have been used for various applications in aquaculture species, including genome-wide association studies (GWAS), characterisation of genetic resources, genome and QTL mapping, and genomic selection (GS) (Yáñez et al., 2022) but not as widely in molluscs.

Genomic selection is particularly pertinent for traits that are expensive or difficult to measure (*e.g.* resistance to diseases, meat quantity) because less phenotypic data is needed to obtain similar accuracies from EBVs resulting from GS and pedigree-based selection. GS can improve the selection response by capturing both within and between family genetic variation components. In aquaculture, the salmon industry has been leading the way in GS for several years (Correa et al., 2017; Ødegård et al., 2014; Robledo et al., 2018; Tsai et al., 2016, 2015). To date, more and more aquaculture species are following this trend (see for review (Allal and Nguyen, 2022; Boudry et al., 2021; Houston et al., 2020; Song et al., 2023; Yáñez et al., 2022; Zenger et al., 2019), such as rainbow trout, European sea bass, sea bream, Nile tilapia, Channel catfish or whiteleg shrimp. The recent development of genotyping tools in molluscs has so far resulted in a relatively limited number of studies on the potential of genomic selection. GS has been investigated in the Portuguese oyster for morphometric traits, edibility traits and disease traits (Vu et al., 2021), in the American oyster for low salinity tolerance (McCarty et al., 2022) and in the silver-lip oyster for pearl quality traits (Jones et al., 2017). Genetic selection for growth traits has been studied in the triangle sail mussel (Wang et al., 2022), European flat oyster (Penaloza et al., 2022) and Pacific oyster (Gutierrez et al., 2018). For the latter, GS to improve resistance to OsHV-1 has also been considered (Gutierrez et al., 2020).

The aim of our study was to assess the potential of genomic selection for growth, meat content and color traits in two independent mixed-family breeding designs at commercial scale in *C. gigas* selected lines. Using a bi-species Axiom Affymetrix 57K SNP array (Gutierrez et al., 2017), we first characterised the genetic structure and diversity of these two populations. Then, we estimated genetic parameters for commercial traits (growth, yield, shell colour, etc.). Finally, we compared the accuracies of genomic selection and pedigree-based selection to provide recommendations to optimise selection in two mixed-family selective breeding programs.

Materials and Methods

Biological material

Broodstock from two French breeding companies (Vendée Naissain and SATMAR) was used. This came from two populations that had undergone six to eight generations of mass selection, mostly for resistance to OsHV-1, growth and morphology, and whose genetic variability had been monitored with genetic markers (Gutierrez et al., 2017; Lapègue et al., 2014) in the two last generations. The first population (P1), resulted from seven full-factorial crosses (factorials) of ten males and ten females each, as proposed by Dupont-Nivet et al. (2006) to minimise inbreeding and increase genetic links between families. Three of these males were used in two of the factorial crosses and one in three of them, leading to a total of 65 males and 70 females potentially generating up to 700 full-sib families. Gill tissue from each parent was sampled and stored in 95% ethanol. Sperm of each male and eggs of each female were individually collected by scarification of the gonads. Sperm mobility was checked visually under the microscope and egg numbers for all females were estimated by cell counting. Sperm from each male was subdivided into 10 beakers, each containing approximately 400 000 oocytes from a single different female. The crosses were performed over two consecutive days. In each case, fertilisation was assessed visually after 20 minutes, and the embryos were rinsed. Embryos were mixed

per factorial in a 200-L tank. The hatching rates were estimated at day 1 for factorials 1, 2, 3, 4 and at day 2 for factorials 5, 6 and 7. After counting, D-larvae from all factorials were grouped together, with in equal proportions from each factorial, into a single batch constituting population P1. Larvae were reared following the company standard larval and nursery protocols of the hatchery following the FAO recommendations (Helm et al., 2004) . Seed was transferred into mesh bags for inshore rearing at an average size of 6 mm and transported to a growing site (Baie des Veys, Normandy).

The second population (P2) was produced following a similar protocol with six full-factorial crosses of ten males and eight females each, potentially generating up to 480 full-sib families. The fertilisation process was similar to P1. Then, larvae were mixed by factorial in 10 000-L tanks and reared separately following the company standard larval and nursery protocols of the hatchery following the FAO recommendations (Helm et al., 2004). At 3 months old and an average size of 6 mm, seed were transferred into mesh bags for inshore rearing in three growing sites (Morsalines, Normandy; Saint Vaast, Normandy; Landéda, Brittany). When 7 months old, the oysters were collected, pooled between sites and reared on a single site (Morsalines), still in six separate factorial batches.

Phenotyping

Populations P1 and P2 were phenotyped when 36 and 31 months old, respectively, which is a common age for oysters to reach commercial size (≈ 60 g) in France. For P2, individuals were randomly and equally sampled within each six-factorial batch. All individuals were phenotyped for the following growth-related traits: Total Individual Weight Before Opening (TW), Shell Length (L), Shell Width (W), Shell Height (H), Upper (UVW) and Lower Valve Weight (LVW), and Wet Flesh Weight (FW). Additionally, traits related to meat content and shell colour of the oyster were also recorded. Flesh Weight Adjusted for Total Weight (FWA) was estimated using the residual of the linear regression between flesh weight and total weight independently in each population. A digital photo of the external right valve (upper/flat) was taken in a tent for ambient light homogenisation between individuals, using a Canon EOS 2000D camera. Images were analysed with an automatic pipeline, using FIJI software (Jourdan et al., 2021) to measure the surfaces of the two valves and the mean external colour of the upper valve, estimated in the CIE LAB space (Robertson, 1977). L*a*b* is a device-independent colour space, providing colour that is consistent and very close to human perception and is most suitable for representation of surfaces or materials illuminated by a light source as less affected by the degree of curvature, shadows and gloss of surfaces (Mendoza et al., 2006). After setting a threshold for the dimension L* starting at a value of 30 to determine which pixels were dark, we could calculate the Percentage of Dark Area (PDA) and the Mean Darkness Intensity (MDI) of all pixels over this threshold. For all traits, outlying values deviating by more than four standard deviations for all phenotypes were excluded from the analyses.

Genotyping

All parents and offspring were genotyped on the bi-species Axiom Affymetrix 57K oyster array, Axiom_Oyster02 (Thermo Fisher Scientific, Waltham, Massachusetts, USA) comprising 40 625 markers for *C. gigas* (Gutierrez et al., 2017). Genotyping was performed at the Gentyane INRAE Platform (Clermont-Ferrand, France). Quality control analyses were carried out using the Axiom Analysis Suite software (AxAS) with the default best practice workflow suggested by manufacturer. Firstly, individuals with a data quality control (DQC) greater than 0.82 and a quality control (QC) call rate greater than 97% were kept. Secondly, this quality analysis allowed us to filter markers with a call rate greater than 98.5%, a minor allele frequency (MAF) greater than 0.01, a Hardy-Weinberg deviation test p-value greater than 0.0001 and a Fisher's linear discriminant (FLD) that measured the clustering quality of

SNPs greater than 5.3 rather than 3.6. Fisher's linear discriminant is essentially the smallest distance between the heterozygous cluster centre and the two homozygous cluster centres in the X dimension (Axiom Genotyping Solution Data Analysis Guide, 2011). The filters applied to the SNPs were more stringent than in the best practice workflow, which explains the limited number of markers, but the quality of the genotyping was improved for the subsequent analyses. For the number of SNPs after each filtration step in P1 and P2, see Supplementary Table 1.

Parentage assignment and genetic characterisation of the population

Parentage assignment was performed using APIS (Griot et al., 2020), with an error rate of 1%. The 1000 best SNPs, selected with a call rate greater than 99.5% and ranked on the MAF value were used. The minimum MAF values were 0.45 and 0.42 for P1 and P2, respectively. Parents without any offspring were excluded from the subsequent analysis.

The number of parents required to maintain genetic diversity, or allele frequencies, within a population was estimated using estimates of effective population size ($\mathrm{N_e}$). The two metrics used to estimate $\mathrm{N_e}$: ($\mathrm{N_{ev}}$) were based on the number of parents and the variance of reproductive success (Chevassus, 1989), whereas ($\mathrm{N_{et}}$) was based on linkage disequilibrium (LD) (D'Ambrosio et al., 2019).

As the individuals were produced by factorials, $\mathrm{N_{ev}}$ could be estimated within each factorial assuming random mating and equal parent representation among the offspring according to Falconer and Mackay (1996): $Ne = 4N_m N_f/(N_m + N_f)$, where N_m and N_f are the numbers of sires and dams, respectively, with offspring in each factorial. To take into account the variability in family representation, $\mathrm{N_e}$ was also calculated according to equation 7 from Chevassus (1989). The $\mathrm{N_{ev}}$ for P1 and P2 populations were obtained by summing the $\mathrm{N_{ev}}$ per factorial.

For the calculation of the $\mathrm{N_{et}}$ based on the expected LD, we followed the approach proposed by D'Ambrosio et al. (2019) using estimates of genetic distances between SNPs (i.e. expressed in Morgans). These distances were derived from the physical distance between SNPs and the recombination rate estimated as the average across all 10 chromosomes. Three different recombination rates (1.8, 2.0 and 2.2) were considered based on published estimates (Gagnaire et al., 2018; Yin et al., 2020).

The rate of inbreeding ($\Delta\mathrm{F}$) can be calculated for a population based on the $\mathrm{N_e}$ (Falconer and Mackay, 1996), expressed as:

$$\Delta\mathrm{F} = \frac{1}{2\mathrm{N_e}}$$

To evaluate potential genetic sub-structuration of each population, and then potential bias, a principal component analysis (PCA) was carried out with the filtered markers on the ade4 package (Dray and Dufour, 2007) developed under R (R Core Team, 2020).

SNP mapping, genome coverage and linkage disequilibrium estimation

All markers of the array with their flanking regions were blasted on the reference genome (cgigas_uk_roslin_v1, GenBank accession number: GCA_902806645.1) (Peñaloza et al., 2021).
To map SNPs and considering the very high polymorphism in the oyster genome, four mismatches were accepted over a length of about 70 bases. Only the markers located at a unique position on the reference genome were kept for further analysis (Supplementary table 3). Linkage disequilibrium (LD), was estimated using the squared correlation based on genotypic allele counts (number of non-reference alleles at each locus) using the PLINK v1.9 software (Chang et al., 2015). Pairwise LD between

all SNPs and adjacent SNPs in a 75-Mb window were computed for each linkage group and population. All pairs of SNPs were sorted based on the distance between the two markers. The mean r^2 values were then determined using the average distances between SNPs: 0 kb for markers less than 15 kb apart; 30, 60, 90, etc. up to 1020 kb with a 30 kb-window (5–15 kb; 15–25 kb; 25–35 kb; etc.).

Estimation of genetic parameters

Estimated breeding values (EBV) and heritabilities were estimated with the BLUPF90 software package (Misztal et al., 2002) under univariate BLUP linear mixed models:

$$\mathbf{y_i} = \mathbf{X_i b_i} + \mathbf{Z_i u_i} + \mathbf{e_i} \quad (1)$$

where $\mathbf{y_i}$ is the performance vector for trait **i** explained by $\mathbf{u_i}$, $\mathbf{b_i}$ and $\mathbf{e_i}$, which are the vectors of additive genetic effects, the vector of the fixed effect of mesh bags for P1 and the vector of residual errors, respectively. For population 2 a group effect equivalent to the factorial was considered but degraded the model with a higher AIC. A mean effect was taken into account for both populations. $\mathbf{X_i}$ is the incidence matrix for $\mathbf{b_i}$, $\mathbf{Z_i}$ is the incidence matrix for $\mathbf{u_i}$. $\mathbf{e_i}$~N(0, Iσ^2e) is a normally distributed vector containing the residuals with mean zero, I is the identity matrix and $\boldsymbol{\sigma^2 e}$ is the residual variance.

Under the BLUP model, the (co)variance structure for the breeding values was calculated using the pedigree relationship matrix (A), where $\mathbf{u_i}$ follows a normal distribution N(0, Aσ^2a) with mean zero, and $\boldsymbol{\sigma^2 a}$ is the additive genetic variance.

Genomic BLUP (GBLUP) uses the same approach as the BLUP model described in eq. (1) except that the pedigree relationship matrix (A) is replaced by a genomic relationship matrix (G) (VanRaden, 2008). Variance components were estimated using the average information restricted maximum likelihood algorithm (Gilmour et al., 1995).

Heritability for BLUP and GBLUP models was estimated by the ratio of the additive genetic variance $\boldsymbol{\sigma^2 a}$ and total phenotypic variance $\boldsymbol{\sigma^2 p = \sigma^2 a + \sigma^2 e}$.

Genetic correlations were estimated using bivariate GBLUP linear mixed models for each pair of traits.

Accuracy

For both BLUP and GBLUP models, the prediction accuracy, often named selection accuracy, was estimated using 40 replicates of standard five-fold cross-validation sets (training population 80%, validation population 20%). Prediction accuracy was calculated as the correlation between the EBVs and the corrected phenotypes of individuals from the validation population divided by the square root of the heritability estimated using the genomic-based relationship matrix (Legarra et al., 2008). Then, accuracy values obtained for both pedigree and genomic EBVs were compared.

Results

1. Genotyping

1.1. Quality

For P1, 130 of the 135 parents and 1131 of the 1135 offspring passed the QC filtering. In P2, 100 of the 104 parents and 1032 of the 1046 offspring passed the QC filtering. Four parents are missing from P2 because there was no biological material for genotyping. Quality analysis led us to retain 14,469 SNPs (35.6%) and 14 452 SNPs (35.6%) from the 40 625 markers available for P1 and P2, respectively.

Between the two populations, 8524 (59 %) SNPs were shared, 5945 and 5928 were unique to P1 and P2, respectively. Among the 5945 SNPs unique to P1, some were excluded from P2 because of a call rate lower than 98.5 (14%), monomorphism (10%), being below the MAF (> 0.01) or Hardy-Weinberg equilibrium deviation p-value (> 0.0001) thresholds (62%), or not passing the best practice workflow recommended by the manufacturer (5%), nor the modified FLD Axiom genotyping filter (9%). Among the 5928 SNPs unique to P2, some were excluded from P1 because of a call rate lower than 98.5 (37%), monomorphism (14%), being below the MAF and Hardy-Weinberg equilibrium deviation p-value thresholds (24%), or not passing the best practice workflow (3%), nor the modified FLD Axiom genotyping filter 22%.

1.2. SNP mapping and genome coverage

In all, 12 454 and 12 267 SNPs matching our mapping requirement were located on the reference genome, representing a loss of 2015 and 2185 markers for P1 and P2, respectively. The positions of markers on the chromosomes are illustrated in Figure 1. The average number of SNPs per Mb range between 2.33 and 36.44, depending on the chromosomes (Table 1) and 30% of all 1 Mb segments have less than 5 SNPs. In both populations, SNP density is not homogeneous across the genome. SNPs are numerous for chromosomes 1, 3, 6, 7 and 10, with denser regions with more than 70 SNPs per Mb (see Supplementary Figures 3 and 4). Conversely, chromosomes 4, 5, 8, and 9 contain fewer SNPs per Mb. Despite its greater length (73 Mb), chromosome 2 has the least SNP coverage, with about 200 markers in each population and large areas (>1 Mb) without any SNPs. The low density of markers implies high average distances between adjacent SNPs that range between 27 kb and 432 kb, depending on the chromosome. Although a large number of markers were not identical between populations, they generally covered the same regions in similar proportions. By slicing the genome by 1-Mb windows, 381 Mb out of a total of 578 Mb (66%) were shown to have a difference in the number of SNPs between P1 and P2 that was less than 5.

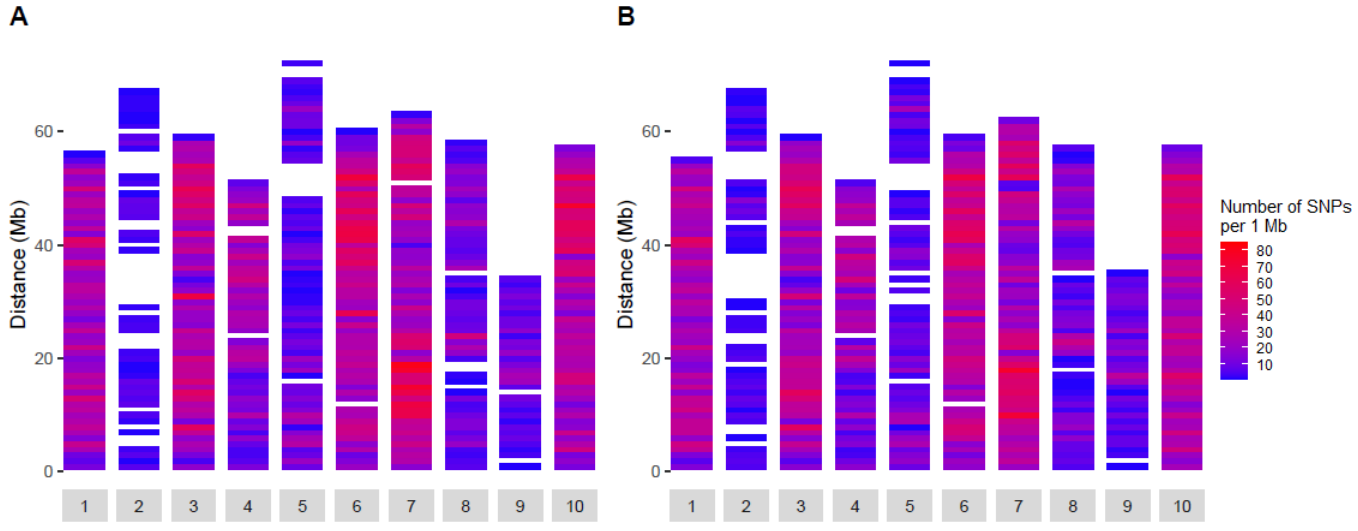


Figure 1. Density map of valid SNPs on each chromosome in population 1 (A) and population 2 (B)

Table 1. Number of SNPs, mean distance and mean r^2 between adjacent SNPs in each linkage group in population 1 (P1) and population 2 (P2).

Chromosome	Number of SNPs	Average number of SNPs per Mb	Average distance between adjacent SNPs (kb)	Average r^2 between adjacent SNPs

	P1	P2	P1	P2	P1	P2	P1	P2
1	1510	1571	26.98	28.07	36.45	34.86	0.107	0.090
2	155	195	2.33	2.93	432.05	343.55	0.066	0.056
3	1886	1791	31.98	30.37	30.75	32.39	0.104	0.084
4	1004	983	19.71	19.29	50.00	50.90	0.110	0.082
5	590	533	8.21	7.42	121.14	133.85	0.091	0.081
6	2098	2067	34.98	34.47	28.18	28.52	0.108	0.084
7	2295	2124	36.44	33.73	27.03	29.06	0.120	0.093
8	617	672	10.66	11.60	92.48	84.69	0.114	0.096
9	299	328	8.57	9.40	112.43	105.41	0.118	0.074
10	2000	2003	35.11	35.16	28.13	28.09	0.094	0.085

1.3. Linkage disequilibrium

Figure 2 shows that the linkage disequilibrium (LD) strongly decreases with distance between pairs of SNPs for both populations up to a distance of 0.1 Mb. At greater distances, LD continues to decline and stabilises. P1 has a higher LD than P2 throughout the genome. On average over the genome, the r^2 values under 15 kb distance between two SNPs are 0.12 and 0.10 for P1 and P2, respectively. Linkage disequilibrium values are generally low (e.g. Table 1: < 0.120 for P1; < 0.96 for P2), even between adjacent SNPs. This is particularly the case for chromosomes 2, 5, and 9 in P2, for which the distances between adjacent SNPs were higher.

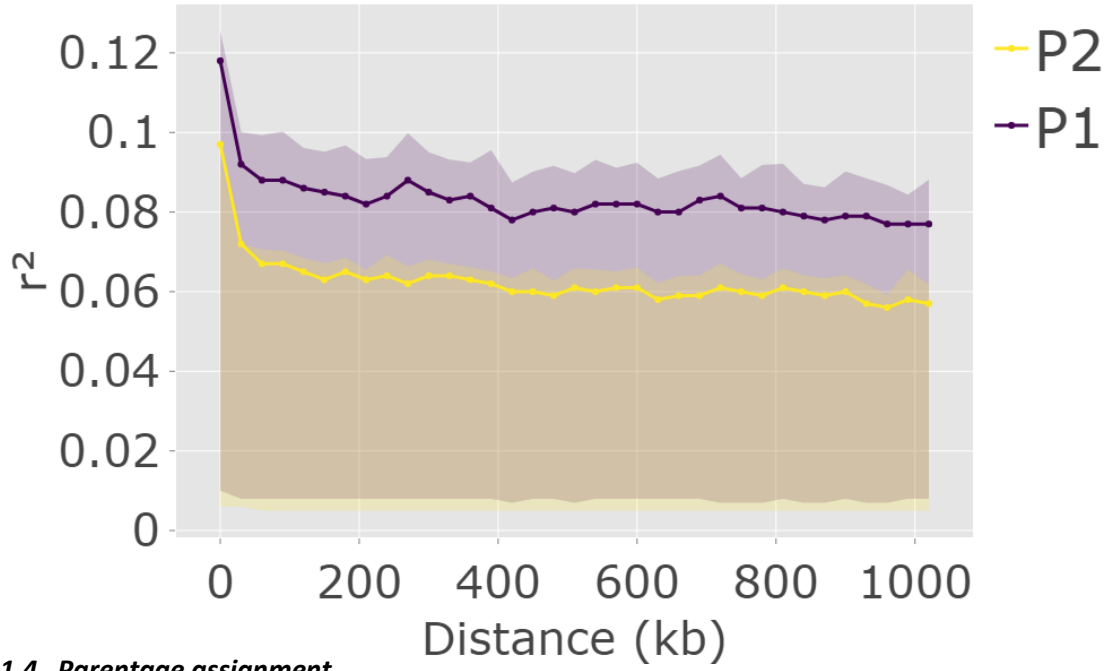


Figure 2. Linkage disequilibrium (r^2, lines) and the interquartile intervals (shaded areas) as a function of genomic distance; purple for population 1 (P1) and yellow for population 2 (P2).

1.4. Parentage assignment

Parentage assignment allowed the reconstruction of the pedigree of 1116 and 915 offspring for P1 and P2, with assignment rates reaching 98.7% and 88.1%, respectively. These offspring were assigned to a single couple of parents, corresponding to the expected crossing designs (see Supplementary Figures 1 & 2). In P2, missing genotypes for four parents (three females in three factorials and one male) is likely to explain the non-assignment of 10.7% of the progeny. Table 2 presents the remaining observed number of full-sib families as well as the minimum and maximum numbers of assigned offspring per

full-sib family, male and female parents, as well as the variances of contribution of sire (V_m) and dam (V_f) reproductive success.

Table 2. Total, minimum, maximum and mean numbers of assigned offspring per full-sib family, sire and dam as well as variance of reproductive success in sires (V_m) and dams (V_f) in population 1 (P1) and population 2 (P2).

	n		Minimum number of offspring		Maximum number of offspring		Mean number of offspring		Variance of reproductive success	
	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
Full-sibs	454	305	1	1	12	20	2.70	2.99	-	-
Sires	64	59	2	3	59	56	19.13	15.46	148.05	152.60
Dams	69	45	1	3	53	66	17.74	20.27	117.40	163.70

1.5. Effective size (Ne) of populations

The expected effective sizes calculated according to Falconer and Mackay (1996) from the initial crossing schemes were 140 and 107 for P1 and P2, respectively. As some individuals were not assigned to their parents (see section 1.4), these expected sizes were adjusted to 135 and 102. The estimated values of N_e with both metrics and for both populations are shown in Table 3. Taking into account the variance in reproductive success by sex (V_m and V_f, Chevassus (1989)), the derived N_{ev} values were 107 for P1 and 76 for P2. In both populations, variance in reproductive success had limited effects on N_{ev}. For the effective size of populations based on the expected LD (N_{et}), the values ranged between 66 and 84 for P1 and between 59 and 75 for P2, depending on the recombination rate considered. The changes in inbreeding coefficient with a minimal theoretical increase in inbreeding was between 0.47% and 0.76% for P1 and between 0.66% and 0.85% for P2 at each generation.

Table 3. Estimation of the effective population size in P1 and P2 based, in the first row, on the variance of reproductive success (N_{ev}) or based on the linkage disequilibrium considering three different recombination rates (N_{et}); estimations of the rate of inbreeding (ΔF) are also given.

Recombination rate	N_{ev} or N_{et}		ΔF (%)	
	P1	P2	P1	P2
-	107	76	0.47	0.66
1.8	84	75	0.60	0.67
2.0	74	66	0.68	0.76
2.2	66	59	0.76	0.85

1.6. Structure of populations under selection

Figure 3 illustrates the results of the principal component analysis, showing the structure of P1 and P2 by factorial cross. The two principal components expressed 4.9% and 3.1% of the variance for P1 and 3.4% and 2.7% for P2. The two populations were homogeneous overall, except for factorial eight in P2, for which offspring were more isolated from the others. As the variance expressed on the X axis for

the P2 population is limited (Figure B), this result suggests that P1 and P2 could both be considered as homogeneous populations. This allowed us to pursue the analysis of GS under good conditions.

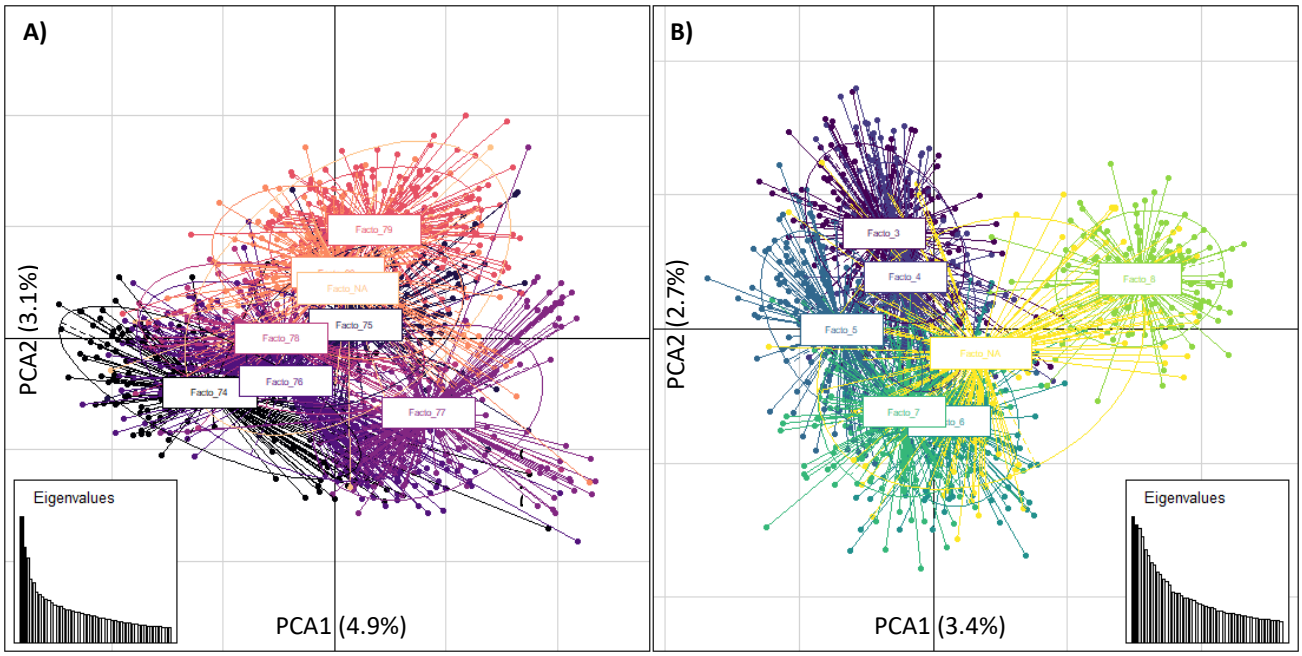


Figure 3. First two axes and associated variances resulting from the principal component analysis (PCA) of genotypic data from population 1 (A) and population 2 (B). The ellipses are constructed with axes defined as 1.5 times the standard deviation of the projections of individual coordinates on the axes.

2. Genomic selection

2.1. Phenotyping

Table 4 presents descriptive statistics on the recorded traits. Depending on the traits considered, between 1232 and 1249 individuals were analysed for P1 and between 1067 and 1077 individuals for P2. The coefficients of variation were mostly different between the populations for Lower (cupped) Valve Weight, Upper (flat) Valve Weight, Total Weight and Mean Upper Valve a* Value and were similar for the other traits.

Table 4. Summary statistics of studied oyster traits in populations P1 (population 1) and P2 (population 2).

Traits		N		Mean		Standard error		Minimum		Maximum		CV (%)	
		P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
Height	(H, mm)	1236	1076	29.04	27.38	4.19	3.91	16.51	15.20	49.57	42.23	14	14
Width	(W, mm)	1236	1076	50.23	47.11	6.79	5.39	28.97	31.14	76.26	64.78	14	11
Length	(L, mm)	1236	1077	88.22	84.22	11.88	9.96	55.19	54.98	129.73	117.55	13	12
Wet Flesh Weight	(FW, g)	1236	1076	7.49	4.55	2.66	1.45	0.64	1.36	17.81	10.98	36	32
Flesh Weight Adjusted	(FWA)	1236	1075	0	0.04	1.17	0.73	-4.17	-2.77	5.05	3.17	-	-
Lower Valve Weight	(LVW, g)	1232	1076	28.36	24.78	9.46	5.92	6.25	11.60	63.79	47.05	33	24
Upper Valve Weight	(UVW, g)	1232	1075	16.93	11.89	5.83	2.90	3.72	5.07	38.13	22.66	34	24
Total Weight	(TW, g)	1236	1076	67.65	57.45	21.86	12.90	15.88	27.40	144.16	108.30	32	22
Lower Valve Area	(LVA, mm^2)	1249	1073	3540.83	3314.26	845.91	576.97	1182.15	1738.51	6423.52	5546.72	24	17
Upper Valve Area	(UVA, mm^2)	1249	1072	2862.13	2491.6	685.87	452.47	1031.78	1268.48	5409.84	4198.98	24	18
L* Value	(LV)	1249	1074	42.70	45.90	5.65	5.46	24.93	25.81	59.61	59.43	13	12
a* Value	(AV)	1249	1074	3.00	1.40	1.58	1.33	-2.12	-2.16	8.63	6.70	53	95
b* Value	(BV)	1249	1074	19.79	19.17	2.58	2.60	10.53	9.14	28.65	28.11	13	14
Mean Darkness Intensity	(MDI)	1249	1074	22.97	21.32	1.51	2.36	16.78	12.57	26.53	26.59	7	11
Percent of Dark Area	(PDA, %)	1249	1067	21.40	16.28	12.07	9.38	1.16	0.85	67.93	60.63	56	58

2.2. Estimation of genetic parameters

Heritability estimates of Height, Weight and Length showed low to moderate values, ranging from 0.15 (± 0.05) to 0.31 (± 0.07) (Table 5). Heritability estimates for these traits were nearly always equivalent whether they were based solely on pedigree information or if genomic information was considered. The only exceptions were Width, which had a higher heritability with only pedigree information in P2, and Flesh Weight Adjusted for TW, which had a higher heritability with genomic information in both populations.

For growth traits, there were no large differences between the two populations except for Upper Valve Weight, which had a pedigree-based estimated heritability of 0.20 (± 0.05) for P1 and 0.31 (± 0.07) for P2 (this was also found using genomic-based estimation with 0.22 (± 0.04) and 0.29 (± 0.04), respectively).

For colour traits, greater differences between the two populations were observed. The heritability of the Mean Upper Valve b* (BV) for P1 was 0.35 (± 0.06) to 0.34 (± 0.05) while for P2 it was 0.56 (± 0.08) to 0.54(± 0.05). For the mean value of L* (LV), heritability ranged between 0.23 (± 0.06) to 0.18 (± 0.04) for P1 to a heritability of 0.45 (± 0.08) to 0.46 (± 0.05) for P2. The mean value of a* (AV) was very limited for the two populations (≈ 0.11) and even lower with the GBLUP model in P1 (0.04 ± 0.02). The MDI had a very low estimated heritability in P1 (<0.1), while in P2 there were high estimates of 0.54 (± 0.08) in BLUP to 0.69 (± 0.04) in GBLUP. The same was observed for the Percent of Dark Area, a heritability lower than 0.2 for P1 and higher than 0.5 for P2.

Table 5. Heritability estimates and their standard errors using BLUP and GBLUP models. P1: population 1, P2: population 2.

		P1 h^2 (SE)		P2 h^2 (SE)	
		BLUP	GBLUP	BLUP	GBLUP
Height	H	0.20 (0.05)	0.21 (0.04)	0.15 (0.05)	0.16 (0.04)
Width	W	0.17 (0.05)	0.21 (0.04)	0.31 (0.07)	0.23 (0.04)
Length	L	0.28 (0.06)	0.26 (0.04)	0.22 (0.06)	0.19 (0.04)
Wet Flesh Weight	FW	0.22 (0.06)	0.25 (0.04)	0.26 (0.06)	0.24 (0.04)
Lower Valve Weight	LVW	0.22 (0.05)	0.24 (0.04)	0.20 (0.05)	0.18 (0.04)
Upper Valve Weight	UVW	0.20 (0.05)	0.22 (0.04)	0.31 (0.07)	0.29 (0.04)
Total Weight	TW	0.19 (0.05)	0.23 (0.04)	0.23 (0.06)	0.21 (0.04)
Lower Valve Area	LVA	0.22 (0.05)	0.24 (0.04)	0.30 (0.06)	0.25 (0.04)
Upper Valve Area	UVA	0.22 (0.05)	0.25 (0.04)	0.30 (0.06)	0.26 (0.04)
Flesh Weight Adjusted	FWA	0.21 (0.05)	0.27 (0.04)	0.19 (0.06)	0.29 (0.05)
L* Value	LV	0.23 (0.06)	0.18 (0.04)	0.45 (0.08)	0.46 (0.05)
a* Value	AV	0.11 (0.04)	0.04 (0.02)	0.11 (0.04)	0.11 (0.03)
b* Value	BV	0.35 (0.06)	0.34 (0.05)	0.56 (0.08)	0.54 (0.05)
Mean Darkness Intensity	MDI	0.08 (0.04)	0.04 (0.02)	0.54 (0.08)	0.69 (0.04)
Percent of Dark Area	PDA	0.18 (0.05)	0.13 (0.04)	0.52 (0.08)	0.54 (0.05)

The genetic correlations between traits were similar in the two populations (Table 6). First, correlations between growth traits were high and positive (greater than 0.5) with slightly lower correlations for Width and Height, especially in P2 (W: 0.16–0.84; H: 0.16–0.70). Genetic correlations of Flesh Weight Adjusted with other growth traits were very low (less than 0.16). Then, regarding genetic correlations between colour traits, models for Mean Upper Valve a* Value (AV) and Mean Darkness Intensity (MDI) failed to converge for P1. For the remaining correlations between colour traits, the same trends were observed in both populations, with strong correlations for Mean Upper Valve L* Value (LV), Mean Upper Valve b* Value (BV), MDI and Percent of Dark Area (PDA). The least correlated colour trait was AV, with a range between 0.27 and 0.54 in P2. Colour traits were positively correlated except AV and PDA, which were negatively correlated with the other colour traits in both populations. Finally, for correlations between colour and growth traits, BV in P1 was the trait that was the most correlated with growth traits (ranging from 0.33 to 0.53, excluding FWA). All the other traits were weakly correlated (less than 0.32). In P2, the colour and growth traits were weakly correlated, with the highest correlation of 0.33 for LV with L and with LVA.

Table 6. Genetic correlations and standard errors for growth and colour traits in population 1 (above the diagonal) and population 2 (below the diagonal). TW: total weight, H: height, W: width, L: length, FW: wet flesh weight, LVW: lower valve weight, UVW: upper valve weight, LVA: lower valve area, UVA: upper valve area, FWA: flesh weight adjusted by TW, LV: mean upper valve L value, AV: mean upper valve a* value, BV: mean upper valve b* value, MDI: mean darkness intensity, PDA: percent of dark area*

	TW	H	W	L	FW	LVW	UVW	LVA	UVA	FWA	LV	AV	BV	MDI	PDA
TW		0.81 (0.06)	0.89 (0.04)	0.85 (0.05)	0.9 (0.03)	0.96 (0.01)	0.88 (0.04)	0.95 (0.02)	0.92 (0.02)	0.06 (0.15)	0.21 (0.16)	-0.22 (0.34)	0.53 (0.11)	-0.15 (0.39)	-0.25 (0.18
H	0.70 (0.10)		0.49 (0.13)	0.56 (0.11)	0.78 (0.07)	0.8 (0.06)	0.62 (0.10)	0.61 (0.10)	0.56 (0.11)	0.16 (0.15)	0.26 (0.17)	-0.17 (0.36)	0.43 (0.13)	0.02 (0.45)	-0.29 (0.25)
W	0.50 (0.12)	0.16 (0.17)		0.66 (0.10)	0.78 (0.07)	0.79 (0.07)	0.84 (0.06)	0.9 (0.04)	0.91 (0.04)	0.01 (0.16)	0.15 (0.18)	-0.32 (0.44)	0.33 (0.13)	-0.16 (0.41)	-0.18 (0.20)
L	0.83 (0.07)	0.39 (0.17)	0.43 (0.15)		0.78 (0.06)	0.8 (0.06)	0.73 (0.07)	0.91 (0.03)	0.87 (0.04)	0.1 (0.14)	0.2 (0.16	-0.16 (0.70)	0.5 (0.11)	-0.29 (1.58)	-0.20 (0.18)
FW	0.83 (0.05)	0.65 (0.12)	0.47 (0.12)	0.7 (0.09)		0.83 (0.05)	0.71 (0.08)	0.88 (0.04)	0.84 (0.05)		0.15 (0.16)		0.4 (0.12)		-0.17 (0.18)
LVW	0.87 (0.04)	0.68 (0.11)	0.46 (0.14)	0.7 (0.10)	0.67 (0.09)		0.75 (0.07)	0.87 (0.04)	0.79 (0.05)	0.01 (0.14)	0.15 (0.16)	-0.16 (0.34)	0.52 (0.11)	-0.23 (0.35	-0.19 (0.27)
UVW	0.90 (0.03)	0.57 (0.12)	0.51 (0.11)	0.71 (0.08)	0.68 (0.08)	0.72 (0.08)		0.85 (0.04)	0.89 (0.03)	-0.13 (0.15)	0.2 (0.17)	-0.17 (0.36)	0.47 (0.12)	-0.19 (0.7)	-0.23 (0.19)
LVA	0.83 (0.05)	0.35 (0.15)	0.84 (0.06)	0.84 (0.06)	0.69 (0.08)	0.71 (0.08)	0.77 (0.06)		0.98 (0.01)	0.1 (0.14)	0.17 (0.16)	-0.15 (0.35)	0.44 (0.12)	-0.28 (0.66)	-0.16 (0.19)
UVA	0.81 (0.05)	0.29 (0.16)	0.73 (0.08)	0.85 (0.05)	0.67 (0.08)	0.59 (0.10)	0.82 (0.04)	0.94 (0.02)		0.06 (0.14)	0.17 (0.16)	-0.17 (0.34)	0.42 (0.12)	-0.25 (1.16)	-0.18 (0.19)
FWA	0.03 (0.15)		0.11 (0.15)	0.05 (0.16)		-0.07 (0.16)		0.03 (0.14)	0.05 (0.14)		-0.06 (0.16)		-0.14 (0.13)		0.11 (0.18)
LV	0.3 (0.13)	0.09 (0.15)	0.09 (0.13)	0.33 (0.13)	0.21 (0.13)	0.21 (0.14)	0.18 (0.12)	0.33 (0.12)	0.3 (0.12)	-0.08 (0.13)		-0.7 (0.30)	0.67 (0.10)	0.4 (1.04)	-0.99 (0.02)
AV	0.04 (0.22)		0.06 (0.22)	-0.09 (0.23)		0.19 (0.24)	0.03 (0.46)	-0.08 (0.20)	-0.18 (0.23)		-0.51 (0.15)		-0.27 (0.31)		0.74 (0.42)
BV	0.19 (0.13)	0.24 (0.15)	0.16 (0.13)	0.13 (0.14)	0.05 (0.13)	0.24 (0.14)	0.18 (0.12)	0.17 (0.13)	0.09 (0.12)		0.76 (0.05)	-0.27 (0.17)			-0.68 (0.11)
MDI	0.15 (0.13)	-0.04 (0.15)	0.13 (0.13)	0.22 (0.13)	0.11 (0.12)	0.12 (0.14)	0.11 (0.11)	0.18 (0.12)	0.2 (0.12)		0.75 (0.05)	-0.4 (0.17)	0.64 (0.06)		-0.56 (0.45)
PDA	-0.21 (0.13)	-0.05 (0.15)	-0.08 (0.13)	-0.27 (0.13)	-0.14 (0.13)	-0.17 (0.14)	-0.09 (0.12)	-0.27 (0.12)	-0.23 (0.12)	0.1 (0.12)	-0.97 (0.01)	0.54 (0.14)	-0.8 (0.04)	-0.85 (0.03)	

2.3. Accuracy of EBVs

When comparing the two selection models, the prediction accuracy was always higher with the GBLUP model except for AV and MDI in P1 (Figure 4, see all results in Supplementary Table 2). However, the genomic information provided better prediction accuracy for all traits in P2. In both populations, the greatest difference in prediction accuracy between the two models was for FWA, which ranged from 0.42 to 0.68 for P1 and from 0.38 to 0.66 for P2.

LV, AV and MDI were among traits with lowest accuracy using a GBLUP model in both populations. The smallest gain in accuracy from BLUP to GBLUP model was for L, LV and PDA in P1 with a difference below 0.07 and for W, LVW, TW and LVA in P2 with a difference below 0.10 (Supplementary Table 2).

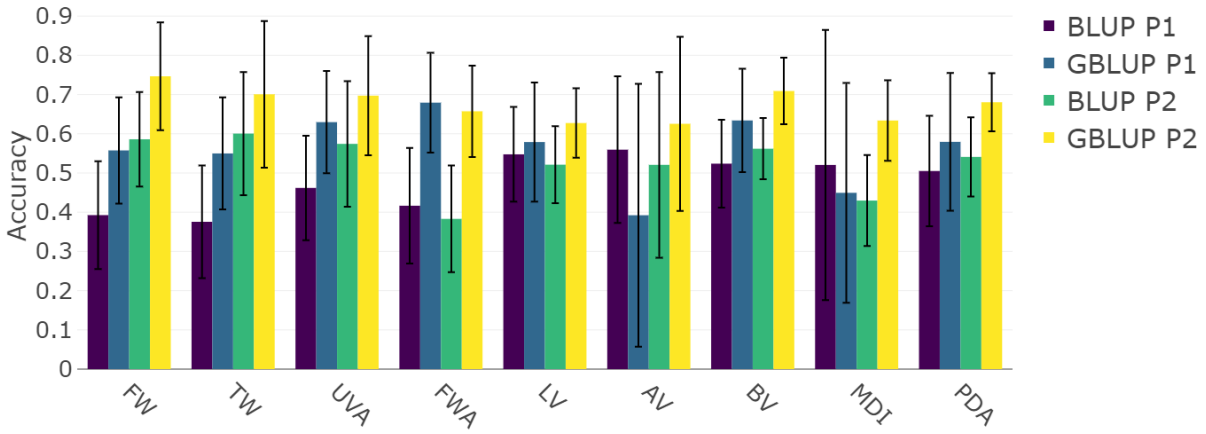


Figure 4. Prediction accuracy and standard deviations using BLUP or GBLUP in two studied population for growth and colour traits. P1: Population 1, P2: Population 2, FW: Wet Flesh Weight, TW: Total Weight, UVA: Upper Valve Area, FWA: Flesh Weight Adjusted by TW, LV: Mean Upper Valve L value, AV: Mean Upper Valve a* value, BV: Mean Upper Valve b* value, MDI: Mean Darkness Intensity, PDA: Percent of Dark Area*

Discussion

Several recent articles have highlighted the potential of genomic selection in Pacific oyster breeding (Gutierrez et al., 2018, 2020; Kriaridou et al., 2020) and started to demonstrate its feasibility in commercial programs (Delomas et al., 2023). Here, we aimed to further demonstrate the feasibility of GS at the commercial scale in mixed family designs, using two independent populations under selection in France. This was made possible thanks to the development of the bi-species Axiom Affymetrix 57K oyster array (Gutierrez et al., 2017), which allowed us to study genetic structure and diversity in our two populations. The genetic parameter estimations of the commercial traits revealed a potential interest in the application of selective breeding. Additionally, comparing the accuracy of EBVs estimated by the BLUP or GBLUP models provides further evidence for the switch to GS for these two mixed-family selective breeding programs.

1. Genotyping

1.1. Quality

The bi-species Axiom Affymetrix 57K oyster array (Gutierrez et al., 2017) has already been used in other studies. After the filters proposed by the best practice workflow of the AxAS software, Gutierrez et al. (2018) obtained 23 400 informative SNPs on 820 individuals and Vendrami et al. (2019) obtained 21 500 on 232 individuals. Using the same thresholds, approximately 24 000 SNPs would have been obtained for P1 and 27 200 SNPs for P2. However, those filtering thresholds were not sufficient to clearly discriminate genotype clusters and be confident in genotype calls in our datasets. To overcome this issue, we were more stringent on the FLD threshold than recommended by the best practice workflow of AxAS, leading to a much lower final number of SNPs retained than other studies using the same array. The choice of quality over quantity was also motivated by the fact that in many species, including Pacific oyster, reducing the number of markers up to a certain number has little impact on accuracy results. In many studies a number of markers between 1000 and 3000 is sufficient to obtain good prediction accuracy (Delomas et al., 2023; Kriaridou et al., 2020; Robledo et al., 2018; Song and Hu, 2022). These previous results, obtained on different species and using different SNP selection methods, consolidated the choice to keep only very high quality SNPs. In the case of the oyster array, many markers are of poor quality, and stringent filtering of genotyping data is to be recommended.

1.2. Genome covering by selected SNPs

The Pacific oyster genome is a complex one that has benefited from a succession of assemblies since 2012 (Peñaloza et al., 2021; Qi et al., 2021; Wang et al., 2019; Yin et al., 2020). The array used in the present study was developed based on the first assembly of a first reference genome (Zhang et al., 2012). Various GWAS have been performed with this assembly (Gutierrez et al., 2018, 2020). However, Hedgecock et al. (2015) revealed assembly errors on this reference genome related to several factors such as extremely high levels of genetic polymorphism, non-Mendelian segregation of marker loci in pair crosses, and high incidence of null alleles for genetic markers. A recent assembly of better quality for a new reference genome has been released (Peñaloza et al., 2021). A large proportion of our selected markers mapped well onto this latest reference genome (86% and 85% for P1 and P2 respectively). However, large disparities in the distribution of SNPs across the genome were observed in the marker density maps. This could lead to missing QTLs in some regions but this might be overcome either by using an existing high density array (like the 190K array, Qi et al., 2017) or by developing a new, more optimised one.

1.3. Linkage disequilibrium

Genomic selection makes use of the linkage disequilibrium (LD) between SNPs and QTLs or causal mutations that contribute to the variance of the trait (Goddard and Hayes, 2009). Linkage disequilibrium was previously reported in both wild and selected *C. gigas* populations (Gutierrez et al., 2017; Zhong et al., 2017). Values for r^2 ranged between 0.15 and 0.25 for SNPs within a distance of 0 to 1 kb and between 0.09 and 0.15 within 100 kb for the studied populations, although in one of their populations, the LD level still decreased to 0.05 at 500 kb. The strong decay of LD was also found in both wild and selected populations, with a r^2 below 0.2 at ~0.13 kb or at ~0.35 kb (Hu et al., 2022). In the other mollusc species common cockle (*Cerastoderma edule*) and silver lip oyster (*Pinctada maxima*), the maximum r^2 reported were 0.05 and 0.082, respectively, even for very short distances (Jones et al., 2013; Vera et al., 2022). Those estimated LD values were low overall compared with fish species such as salmonids, cyprinids or perciforms (D'Ambrosio et al., 2019; Hong Xia et al., 2015; Xu et al., 2014). In our study, the LD estimates were calculated on two large commercial-scale populations, which contrasts with the work of Gutierrez et al. (2017), who used a small sample of individuals. The levels of LD we report in P1 and P2 were calculated with high variance of family sizes, potentially leading to an over estimation of LD in our populations. However, our results confirm the low level of LD in *C. gigas* populations as being in the range obtained in previous studies on molluscs. Low LD results can be explained by several factors, such as the relatively high effective population sizes reported in our study and the limited number of generations for which these populations had been bred in isolation.

2. Diversity and structure of breeding populations

Implementation of a genomic selection program implies carefully considering several factors. One of these is the mating design. Comparing different mating designs (factorial, partial factorial and nested) by simulation, it was demonstrated that a factorial mating system showed the lowest inbreeding rate and the highest selection response (Dupont-Nivet et al., 2006; D'Agaro et al., 2007; Sonesson and Ødegård, 2016). Mass spawning mating could be a way to make a full factorial design, but high variance in reproductive success is to be expected (Boudry et al., 2002; Smits et al., 2020). However, full-factorials with individual, couple-by-couple fertilisations can be difficult to implement. So, partial factorial seemed to be a good compromise to achieve high genetic responses while preserving genetic variability (Dupont-Nivet et al., 2006). The Pacific oyster has a well-controlled reproductive cycle, with the possibility of generating synchronised spawning of males and females, making this type of mating design feasible. This meant that the partial factorial mating design could be carried out with our populations, as represented in Supplementary Figures 1 and 2 (showing results obtained after parentage assignment).

Another factor to consider is how to rear all the families. On the one hand, mixing all the families together is the least demanding way to do this in terms of infrastructure and workload. It also ensures a common environment for comparing candidates for selection. On the other hand, there is a potential for unequal contribution of parents (Boudry et al., 2002) and this approach raises the risk of inbreeding in later generations (García-Ballesteros et al., 2021). In our data, we observe some variance in reproductive success, but this has little impact on the N_{ev} of the two populations. Our estimates are higher than the N_e estimated by Zhong et al. (2017): between 48 and 58 for populations under selection, and Smits et al. (2020): 39 or 87 considering variance in reproductive success or not, respectively. The N_{et} estimated based on linkage disequilibrium gives similar results to N_{ev} when considering a recombination rate of 1.8, but differs a little from it when increasing the recombination rate. However, all the values of N_e are consistent and a limited increase in inbreeding for the next generations is expected (< 1%). Therefore, mixed-family breeding programs should allow for an inbreeding increase rate lower than suggested by the FAO for fish breeders in order to assure long term management of genetic diversity (Fisheries and Aquaculture Management Division, 2008).

Finally, another point requiring vigilance is to keep genetic links between families and factorials. Factorial designs generate a very high number of families, which can result in a very low number of individuals phenotyped and genotyped in each family (3.5–4) (Haffray et al., 2018). Small and variable family sizes could be considered unfavourable for the accuracy of the estimated breeding values (EBVs). Nevertheless, an intermediate to high accuracy of EBVs can be estimated even with a low number of full-sibs per family (Haffray et al., 2018), thanks to the very high number of genetic connections within each factorial (70 to 80 families in our two studied populations) and between all factorials when previous generations are known. The relative genomic homogeneity of P1 and P2 populations is also a good indicator that no high variance of reproductive success appeared in recent past generations. Mixed-family designs with egg equalisation thus appear to be a relevant way to limit costs without sacrificing control of parental representation. These two populations from mixed families and *a posteriori* pedigree were, therefore, a suitable manner to test the feasibility of genomic selection.

3. Estimation of genetic parameters

3.1. Heritability

Heritability estimates depend on parameters of the studied population such as genetic diversity, life stage and environmental conditions (Gjedrem and Thodesen, 2005). As a result, they may vary between populations and/or rearing environments. Hollenbeck and Johnston (2018), reviewing selective breeding in molluscs, reported low to moderate heritability for most traits of commercial interest. On diverse molluscan species, like bay scallop, blue mussel or manila clam, they reported a range of heritability for total weight between 0.22 and 0.39. For morphometric traits (H, W, L), heritability estimates ranged between 0.11 and 0.73 (Smits et al., 2020; Vu et al., 2021; Wang et al., 2014; Zheng et al., 2004). Focusing on *C. gigas,* heritability estimates were reported in the range of 0.07 to 0.58 for total weight and 0.1 to 0.49 for shell growth traits (Dégremont et al., 2007; Gomes et al., 2018; Gutierrez et al., 2018; Xu et al., 2017). Our results are in accordance with these previous studies, with a low estimated heritability for Total Weight between 0.19 and 0.23 and low to moderate heritability values for shell growth traits (H, W, L) ranging from 0.15 to 0.31 in our two populations. For Wet Flesh Weight, we obtained heritability estimates between 0.22 and 0.26. This is lower than the 0.35 reported by Gutierrez et al. (2018) or the 0.53 reported in the pearl oyster by Wang et al. (2010). These results indicate genetic variation in growth traits, providing a significant opportunity to improve these phenotypes through genomic selection.

A major difference between all these studies is the standard error (SE) associated with the heritability estimates. The low SE values obtained in our study (between 0.02 and 0.08) are slightly lower than those of Wang et al. (2022), with an SE of about 0.11 using genomics on 764 individuals from random group matings of *Hyriopsis cumingii,* and in agreement with those obtained by Gutierrez et al. (2018), with an SE of 0.06. Conversely, we should not ignore the numbers of individuals and families (Houston et al., 2020). In several studies, heritability values had high standard errors (0.09–0.25) (Kong et al., 2015; Smits et al., 2020; Xu et al., 2017), probably due to a small number of families (< 50 full-sib) and/or offspring (≤ 400). All these results show the possibility of accurately estimating heritability with a mixed-family breeding program, provided that there is a sufficient number of genotyped individuals.

It has been well demonstrated that differences in rearing methods (suspended vs bag) and practices (immersed vs on shore) may have a significant impact on the expression of additive genetic components of oyster growth (Sheridan, 1997). Oyster production in mesh bags with twice daily tide emersion, as used along the European Atlantic coast, may increase environmental variance, impacting the estimation of the additive genetic variance. As an example, Melo et al. (2018) found moderate to high GxE correlations between marine and brackish environments for growth and survival traits, indicating that creating independent breeding programs, instead of a single breeding program for generalist families, would be more effective at improving those traits in different geographical locations. A GxE study on the impact of novel rearing conditions (lantern nets, ropes and rotative systems, etc.) could help find the best conditions for evaluating additive genetic components and EBV for oyster growth when compared with standard French rearing conditions on the Atlantic coast.

Colour is an often studied trait in molluscs species because a large phenotypic variability is available and market opportunities may exist for different phenotypes (Williams, 2017). Many means are available to obtain colour phenotype data; for example, to establish categories with threshold parameters or to identify individuals with higher shell contrast. Other methods of acquiring colour data on molluscs were used in other studies, such as Raman spectroscopy (Stenger et al., 2021) or diffuse reflectance UV-Vis spectra (Yan et al., 2021). The main advantage of our method is that it is fast, inexpensive, close to the human visual perception and allows the measurement of other traits at the same time. It is often reported that shell colour in molluscs is associated with QTLs (see for review Saenko and Schilthuizen, 2021; Williams, 2017). Here, we considered colour as a quantitative trait under polygenic control. For shell pigmentation equivalent to L* (Xu et al., 2017), a large range of estimated heritabilities have been reported, ranging between 0.59 (Evans et al., 2009) and 0.156 (± 0.078) (Xing et al., 2018). Strong realised heritabilities (between 0.52 and 0.69) for the parameters L*, a* and b* were described in a black shell lineage (Xu et al., 2017). In our study, the same measurement procedure was used for both populations, and we observed differences between them, mainly for the traits related to the parameters L* (luminance) and b* (blue-yellow axis). The heritability estimates for these traits in the P1 population were lower than 0.35, with a very low estimate for the Mean Darkness Intensity (< 0.1). In P2, however, the values range from 0.44 to 0.65, which is in accordance with previous studies. These differences may be due to the difference in phenotypic variance of these colour parameters between the two populations, with P1 being lighter and more uniform than P2. Studying lineages of different shell background colours may explain the difference in heritability of the a* parameter obtained in our study (0.1) compared with Xu et al. in 2017 (0.65). In addition, De et al. (2019) showed that different families performed differently across two different rearing sites revealing the genotype by environment (G × E) interactions on growth and shell colour traits. In our study, site effects were confounded with our populations, which could explain the observed differences in heritabilities. In the future, in order to better identify the selection objectives for colour, a classification should be established according to the perception of the consumer's eye.

Nevertheless, our results of low to high shell colour heritabilities confirm that oyster shell colour is under genetic control and could be improved through a mixed-family selective breeding program.

3.2. Genetic correlation estimates

Genetic correlation is notably used to evaluate whether the genetic basis of a desired trait could be antagonistic with another trait. In the literature, genetic correlation estimates between growth traits were reported between 0.55 and 0.95 (Gomes et al., 2018; Alejandro P. Gutierrez et al., 2018; Kong et al., 2015). The same characteristics have been found in other oyster species such as *C. virginica*, with moderate to high positive genetic correlations between total weight and other morphometric traits (0.58–0.90) (Vu et al., 2021). Most growth traits here have moderate to strong genetic correlations (0.38–0.98 except for the low genetic correlation between height and width in P2: 0.16) in both populations, which is consistent with previous studies. Length traits, mainly width and height, are very likely prone to measurement errors. Considering that oysters have a great morphological plasticity, it is difficult to get an accurate, repeatable measure of these traits, which may explain their lower genetic correlation in our two populations. The high genetic correlations between some traits could indicate that common genes are involved in the expression of several traits. In a selection scheme, evaluating and selecting on one of these traits would simultaneously improve all the others. The Flesh Weight Adjusted trait can also be a breeding objective. If the flesh weight is corrected by the total weight, then this trait can be selected independently of the other traits.

Conflicting results about correlation of shell colour parameters are found in the literature, with genetic correlations ranging from -0.11 to 0.98 depending on genetic background – especially when comparing different lines selected for different colours – (Wan et al., 2017; Xu et al., 2017). In our study, the genetic correlations of colour traits vary, ranging from -0.27 to 0.76 considering the three main colour parameters (L*, a*, b*). It is noteworthy that between the colour parameters a* and b* in the two populations we find a medium negative correlation of -0.27. It is important to know the relationships of colour and growth traits well in order to choose breeding strategies that best suit the improvement objectives. To date, these relationships seem very dependent on the lines studied. These correlations vary depending on studies, sometimes showing traits to be uncorrelated (Han and Li, 2021; Wan et al., 2017) and sometimes highly correlated (Xu et al., 2017). In our two populations, it seems that the colour traits are either uncorrelated or only slightly genetically correlated with growth traits, except for the P1 where the colour parameter b* is moderately correlated with the total weight, length and weight of the valves (≈ 0.5). In our case, given the weak genetic correlations, different selection plans aiming to improve colour or growth traits can be developed without negative interaction between these traits.

4. Genomic prediction accuracy

Accurate estimations of breeding values are essential for developing a breeding program and predicting the responses of traits of interest to selection. An overview made of the application of genomic prediction strategies to a variety of cultured aquatic species reported a 22% improvement in prediction accuracy for growth-related traits over pedigree-based strategies (Houston et al., 2020), indicating that substantial improvements in the rate of genetic gain can be achieved through genomics-based selection techniques. In 2018, Gutierrez et al. published results on three traits going in the same direction, with expected improvements of ∼25% for wet weight and shell length to ∼30% for shell height. On *C. virginica*, a 15% increase in shell length was found by Vu et al. (2021). Regardless of the population, our results comparing the accuracy of EBV from genomic models with the accuracy from pedigree models show that genomic models are better for all the traits studied, with only the exceptions of the Mean Darkness Intensity and the Mean Upper Valve a* Value in P1, where the

heritability was close to zero. The gain in accuracy for growth traits is between 15 % and 60 % in P1 and between 10% and 30 % in P2. For colour traits with consistent heritability (L*, b*), the gain in accuracy with genomic information is between 6% and 26% and even 48% for MDI in P2. This is the first time to our knowledge that the accuracy of selection for shell colour parameters has been evaluated for *C. gigas*. These results are in agreement with previous studies in aquaculture showing improved accuracy in predicting breeding values from genomic rather than pedigree-based models in Atlantic salmon (Ødegård et al., 2014; Robledo et al., 2018; Tsai et al., 2015; Yoshida et al., 2018), sea bass (Palaiokostas *et al*., 2016, 2018), rainbow trout (D'Ambrosio et al., 2020; Vallejo et al., 2017) and whiteleg shrimp (Wang et al., 2017). Our results confirm the relevance of using genomic tools in the breeding programs of the Pacific oyster.

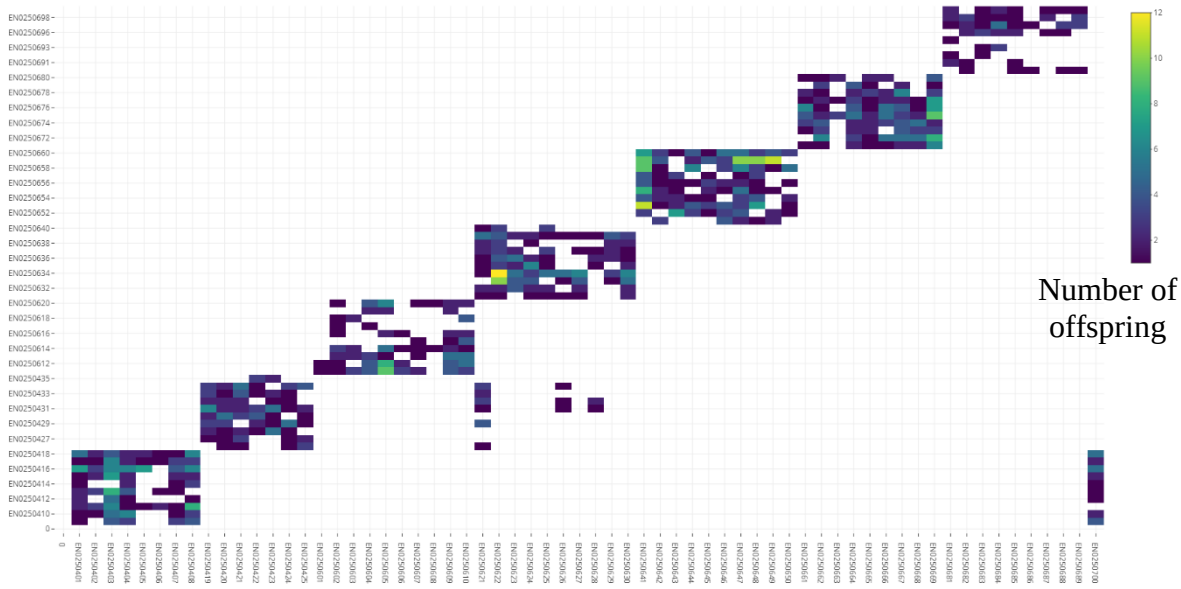
Conclusion

Our results confirm the possibility of improving growth and colour traits in commercial populations of Pacific oyster and the relevance of genomic selection and mixed-family designs for doing this. The two breeding programs evaluated both manage to show substantial genetic variation for genetic improvement and good genetic diversity, as shown by effective size estimates. However, better genomic tools need to be developed because the number of SNPs of good quality was low and provided poor genome coverage and density. Finally, interactions between genotype and environment, including rearing practices, should be evaluated by further studies to optimise breeding programs for hatcheries that provide seed able to grow in highly contrasted environments.

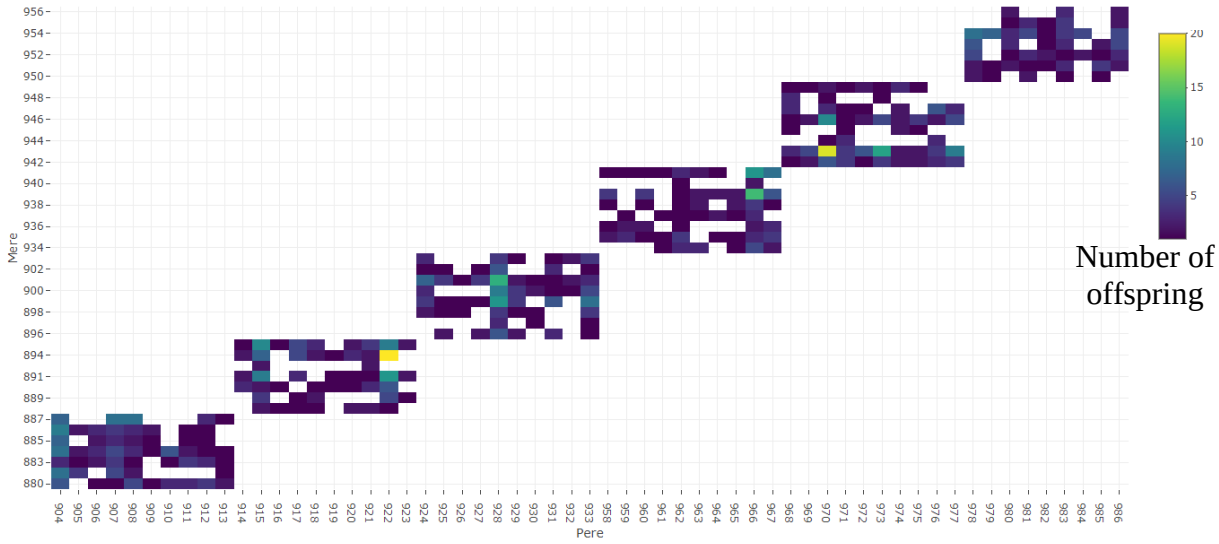
Acknowledgements

This work was supported by the European Maritime and Fisheries Fund (EMFF) and FranceAgriMer as part of 'Quality Huitre' under grant number PFEA470018FA1000011. We acknowledge the staff of the INRAE platform GENTYANE for their genotyping expertise and those of the Pôle de Calcul et de Données Marines (PCDM) for providing DATARMOR data access and computational resources. We thank Helen McCombie from Brest University translation bureau for her professional English editing services. Finally, we thank Frédéric Chenier and Stéphane Angeri from Vendée Naissain, and Laura Gribouval and Bertrand du Mesnildot from SATMAR, for their participation in the project.

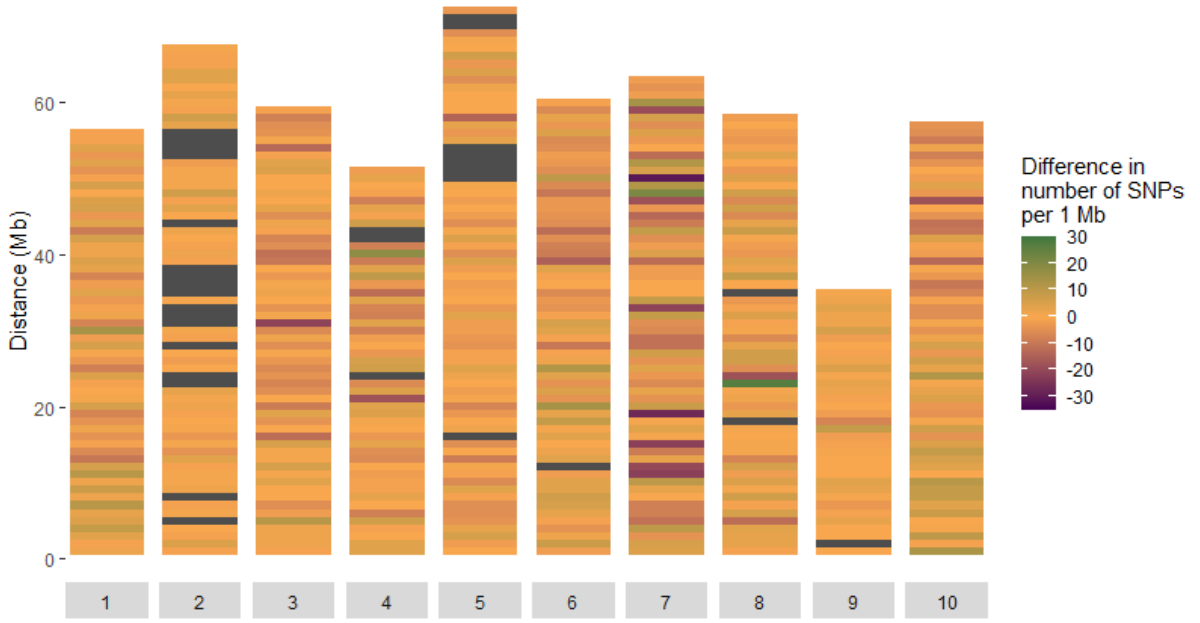
Appendices



Supplementary Figure 1. Breeding design of population 1 with sires on the x-axis and dams on the y-axis, showing number of individuals in each full-sib family assessed by parentage assignment.



Supplementary Figure 2. Breeding design of population 2 with sires on the x-axis and dams on the y-axis, showing number of individuals in each full-sib family assessed by parentage assignment.



Supplementary Figure 3. Density map of the difference in the numbers of SNPs per megabase between population 1 (redish) and population 2 (greenish) for each linkage group. Megabases with a dark grey colour did not include any SNPs in both populations.

Supplementary Table 1. Number of SNPs after each filtration step in population 1 and population 2. CR: Call rate, MAF: Minor allele frequency, H.W. p-value: Hardy-Weinberg p-value, FLD: Fisher's linear discriminant.

Filters	**SNP number in P1**	**SNP number in P2**
None	40 625	40 625
PolyHighRes/ NoMinorHom	24 002	27 252
CR (98.5)	21 012	25 057
MAF (>0.01)	19 933	23 519
H.W. p-value (>0.0001)	17 051	16 237
FLD (5.3)	14 469	14 452

Supplementary Table 2. Accuracy of EBVs in population 1 (P1) and population 2 (P2) estimated with BLUP and GBLUP models.

Traits		P1		P2	
		BLUP	GBLUP	BLUP	GBLUP
Height	(H, mm)	0.29 (0.18)	0.40 (0.19)	0.51 (0.20)	0.65 (0.21)
Width	(W, mm)	0.23 (0.18)	0.37 (0.18)	0.68 (0.15)	0.78 (0.17)
Length	(L, mm)	0.35 (0.13)	0.43 (0.15)	0.54 (0.17)	0.68 (0.13)
Wet Flesh Weight	(FW, g)	0.39 (0.14)	0.56 (0.14)	0.59 (0.12)	0.75 (0.14)
Lower Valve Weight	(LVW, g)	0.39 (0.13)	0.55 (0.14)	0.60 (0.16)	0.66 (0.18)
Upper Valve Weight	(UVW, g)	0.40 (0.15)	0.53 (0.15)	0.60 (0.11)	0.73 (0.11)
Total Weight	(TW, g)	0.38 (0.14)	0.55 (0.14)	0.60 (0.16)	0.70 (0.19)
Lower Valve Area	(LVA, mm^2)	0.47 (0.13)	0.63 (0.14)	0.62 (0.14)	0.72 (0.14)
Upper Valve Area	(UVA, mm^2)	0.46 (0.13)	0.63 (0.13)	0.57 (0.16)	0.70 (0.15)
Flesh Weight Adjusted	(FWA)	0.42 (0.15)	0.68 (0.13)	0.38 (0.14)	0.66 (0.12)
Mean Upper Valve L* Value	(LV)	0.55 (0.12)	0.58 (0.15)	0.52 (0.10)	0.63 (0.09)
Mean Upper Valve a* Value	(AV)	0.56 (0.19)	0.39 (0.34)	0.52 (0.24)	0.63 (0.22)
Mean Upper Valve b* Value	(BV)	0.52 (0.11)	0.63 (0.13)	0.56 (0.08)	0.71 (0.08)
Mean Darkness Intensity	(MDI)	0.52 (0.34)	0.45 (0.28)	0.43 (0.12)	0.63 (0.10)
Percent of Dark Area	(PDA, %)	0.51 (0.14)	0.58 (0.18)	0.54 (0.10)	0.68 (0.07)

Supplementary Table 3. Location of the SNP site from the bi-species Axiom Affymetrix 57K oyster array, Axiom_Oyster02, on the latest assembly (cgigas_uk_roslin_v1, GenBank accession number: GCA_902806645.1).

References

Allal, F., Nguyen, N.H., 2022. Genomic Selection in Aquaculture Species. In: Ahmadi, N., Bartholomé, J. (Eds.), Genomic Prediction of Complex Traits, Methods in Molecular Biology. Springer US, New York, NY, pp. 469–491. https://doi.org/10.1007/978-1-0716-2205-6_17

Axiom Genotyping Solution Data Analysis Guide [WWW Document], 2011. URL https://www.affymetrix.com/support/downloads/manuals/axiom_genotyping_solution_analysis_guide.pdf

Boudry, P., Allal, F., Aslam, M.L., Bargelloni, L., Bean, T.P., Brard-Fudulea, S., Brieuc, M.S.O., Calboli, F.C.F., Gilbey, J., Haffray, P., Lamy, J.-B., Morvezen, R., Purcell, C., Prodöhl, P.A., Vandeputte, M., Waldbieser, G.C., Sonesson, A.K., Houston, R.D., 2021. Current status and potential of genomic selection to improve selective breeding in the main aquaculture species of International Council for the Exploration of the Sea (ICES) member countries. Aquac. Rep. 20, 100700. https://doi.org/10.1016/j.aqrep.2021.100700

Boudry, P., Collet, B., Cornette, F., Hervouet, V., Bonhomme, F., 2002. High variance in reproductive success of the Pacific oyster (*Crassostrea gigas*, Thunberg) revealed by microsatellite-based parentage analysis of multifactorial crosses. Aquaculture, 204(3-4), 283-296. https://doi.org/10.1016/S0044-8486(01)00841-9

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4, s13742-015-0047–8. https://doi.org/10.1186/s13742-015-0047-8

Chevassus, B., 1989. Constitution of aquacultural stocks: Genetic aspects. Presented at the Advances in Tropical Aquaculture, Workshop at Tahiti, French Polynesia, 20 Feb - 4 Mar 1989. https://archimer.ifremer.fr/doc/00000/1442/

Correa, K., Bangera, R., Figueroa, R., Lhorente, J.P., Yáñez, J.M., 2017. 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. https://doi.org/10.1186/s12711-017-0291-8

D'Agaro, E., Woolliams, J.A., Haley, C.S., Lanari, D., 2007. Optimizing mating schemes in fish breeding. Ital. J. Anim. Sci. 6, 795–796. https://doi.org/10.4081/ijas.2007.1s.795

D'Ambrosio, J., Morvezen, R., Brard-Fudulea, S., Bestin, A., Acin Perez, A., Guéméné, D., Poncet, C., Haffray, P., Dupont-Nivet, M., Phocas, F., 2020. Genetic architecture and genomic selection of female reproduction traits in rainbow trout. BMC Genom. 21, 558. https://doi.org/10.1186/s12864-020-06955-7

D'Ambrosio, J., Phocas, F., Haffray, P., Bestin, A., Brard-Fudulea, S., Poncet, C., Quillet, E., Dechamp, N., Fraslin, C., Charles, M., Dupont-Nivet, M., 2019. Genome-wide estimates of genetic diversity, inbreeding and effective size of experimental and commercial rainbow trout lines undergoing selective breeding. Genet. Sel. Evol. 51, 26. https://doi.org/10.1186/s12711-019-0468-4

de Melo, C.M.R., Durland, E., Langdon, C., 2016. Improvements in desirable traits of the Pacific oyster, *Crassostrea gigas*, as a result of five generations of selection on the West Coast, USA. Aquaculture 460, 105–115. https://doi.org/10.1016/j.aquaculture.2016.04.017

De, X., Qi, L., JingXiao, Z., 2019. Genotype by environment (G × E) interaction for growth and shell color traits in the white-shell strain of Pacific oyster (*Crassostrea gigas*). J. Fish. China. 43, 474–482.

Dégremont, L., Ernande, B., Bédier, E., Boudry, P., 2007. Summer mortality of hatchery-produced Pacific oyster spat (Crassostrea gigas). I. Estimation of genetic parameters for survival and growth. Aquaculture 262, 41–53. https://doi.org/10.1016/j.aquaculture.2006.10.025

Dégremont, L., Garcia, C., Allen, S.K., 2015. Genetic improvement for disease resistance in oysters: A review. J. Invertebr. Pathol. 131, 226–241. https://doi.org/10.1016/j.jip.2015.05.010

Delomas, T.A., Hollenbeck, C.M., Matt, J.L., Thompson, N.F., 2023. Evaluating cost-effective genotyping strategies for genomic selection in oysters. Aquaculture 562, 738844. https://doi.org/10.1016/j.aquaculture.2022.738844

Dray, S., Dufour, A.-B., 2007. The ade4 Package: Implementing the duality diagram for ecologists. J. Stat. Softw. 22, 1–20. https://doi.org/10.18637/jss.v022.i04

Dupont-Nivet, M., Vandeputte, M., Haffray, P., Chevassus, B., 2006. Effect of different mating designs on inbreeding, genetic variance and response to selection when applying individual selection in fish breeding programs. Aquaculture 252, 161–170. https://doi.org/10.1016/j.aquaculture.2005.07.005

Enez, F., Lorgeoux, B., H, M., Bugeon, J., Vandeputte, M., Gagnaire, P.-A., Bierne, N., Blanc, P.P., Haffray, P., 2018. Genetic parameters for growth and colour traits in Pacific blue shrimp *Litopenaeus stylirostris* in a mixed family design with SNP parentage assignement in New-Caledonia. In 11. World Congress on Genetics Applied to Livestock Production. https://hal.inrae.fr/hal-03156790

Evans, S., Camara, M.D., Langdon, C.J., 2009. Heritability of shell pigmentation in the Pacific oyster, *Crassostrea gigas*. Aquaculture 286, 211–216. https://doi.org/10.1016/j.aquaculture.2008.09.022

Evans, S., Langdon, C., 2006. Direct and indirect responses to selection on individual body weight in the Pacific oyster (*Crassostrea gigas*). Aquaculture 261, 546–555. https://doi.org/10.1016/j.aquaculture.2006.07.037

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to Quantitative Genetics (Pearson, ed. 4).

FAO, 2022. The State of World Fisheries and Aquaculture 2022: Towards Blue Transformation, The State of World Fisheries and Aquaculture (SOFIA). FAO, Rome, Italy. https://doi.org/10.4060/cc0461en

Fisheries and Aquaculture Management Division, 2008. Aquaculture Development 3. Genetic Resource Management, FAO Technical Guidelines for Responsible Fisheries. FAO, Rome, Italy.

Gagnaire, P.-A., Lamy, J.-B., Cornette, F., Heurtebise, S., Dégremont, L., Flahauw, E., Boudry, P., Bierne, N., Lapègue, S., 2018. Analysis of genome-wide differentiation between native and introduced populations of the cupped oysters *Crassostrea gigas* and *Crassostrea angulata*. Genome Biol. Evol. 10, 2518–2534. https://doi.org/10.1093/gbe/evy194

García-Ballesteros, S., Fernández, J., Toro, M.Á., Villanueva, B., 2021. Benefits of genomic evaluation in aquaculture breeding programs with separate rearing of families. Aquaculture 543, 737004. https://doi.org/10.1016/j.aquaculture.2021.737004

Gilmour, A.R., Thompson, R., Cullis, B.R., 1995. Average information REML: An efficient algorithm for variance parameter estimation in linear mixed models. Biometrics 51, 1440–1450. https://doi.org/10.2307/2533274

Gjedrem, T., Thodesen, J., 2005. Selection, in: Gjedrem, T. (Ed.), Selection and breeding programs in aquaculture. Springer Netherlands, Dordrecht, pp. 89–111. https://doi.org/10.1007/1-4020-3342-7_7

Gjerde, B., Gjøen, H.M., Villanueva, B., 1996. Optimum designs for fish breeding programmes with constrained inbreeding Mass selection for a normally distributed trait. Livest. Prod. Sci. 47, 59–72. https://doi.org/10.1016/S0301-6226(96)01000-7

Goddard, M.E., Hayes, B.J., 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat. Rev. Genet. 10, 381–391. https://doi.org/10.1038/nrg2575

Gomes, R.B., Dybas, P.R., Silva, F.C. da, Sühnel, S., Albuquerque, M.C.P. de, Melo, C.M.R. de, 2018. Estimates of genetic parameters for growth and survival in Pacific oyster (*Crassostrea gigas*). Bol. Inst. Pesca 44. https://doi.org/10.20950/1678-2305.2018.337

Griot, R., Allal, F., Brard-Fudulea, S., Morvezen, R., Haffray, P., Phocas, F., Vandeputte, M., 2020. APIS: An auto-adaptive parentage inference software that tolerates missing parents. Mol. Ecol. Resour. 20, 579–590. https://doi.org/10.1111/1755-0998.13103

Gutierrez, Alejandro P, Bean, T.P., Hooper, C., Stenton, C.A., Sanders, M.B., Paley, R.K., Rastas, P., Bryrom, M., Matika, O., Houston, R.D., 2018. A Genome-wide association study for host resistance to *Ostreid Herpesvirus* in Pacific oysters (*Crassostrea gigas*). G3: Genes Genomes Genet. 8, 1273–1280. https://doi.org/10.1534/g3.118.200113

Gutierrez, Alejandro P., Matika, O., Bean, T.P., Houston, R.D., 2018. Genomic selection for growth traits in Pacific oyster (*Crassostrea gigas*): Potential of low-density marker panels for breeding value prediction. Front. Genet. 9, 391. https://doi.org/10.3389/fgene.2018.00391

Gutierrez, A.P., Symonds, J., King, N., Steiner, K., Bean, T.P., Houston, R.D., 2020. Potential of genomic selection for improvement of resistance to ostreid herpesvirus in Pacific oyster (*Crassostrea gigas*). Anim. Genet. 51, 249–257. https://doi.org/10.1111/age.12909

Gutierrez, A.P., Turner, F., Gharbi, K., Talbot, R., Lowe, N.R., Peñaloza, C., McCullough, M., Prodöhl, P.A., Bean, T.P., Houston, R.D., 2017. Development of a medium density

combined-species SNP array for Pacific and European oysters (*Crassostrea gigas* and *Ostrea edulis*). G3: Genes Genomes Genet. 7, 2209–2218. https://doi.org/10.1534/g3.117.041780

Haffray, P., Enez, F., Bugeon, J., Chapuis, H., Dupont-Nivet, M., Chatain, B., Vandeputte, M., 2018. Accuracy of BLUP breeding values in a factorial mating design with mixed families and marker-based parentage assignment in rainbow trout *Oncorhynchus mykiss*. Aquaculture 490, 350–354. https://doi.org/10.1016/j.aquaculture.2018.03.003

Han, Z., Li, Q., 2021. Relationship between shell color and growth and survival traits in the Pacific oyster *Crassostrea gigas*. J. Ocean Univ. China. 20, 985–991. https://doi.org/10.1007/s11802-021-4676-5

Hedgecock, D., Shin, G., Gracey, A.Y., Den Berg, D.V., Samanta, M.P., 2015. Second-generation linkage maps for the Pacific oyster *Crassostrea gigas* reveal errors in assembly of genome scaffolds. G3: Genes Genomes Genet. 5, 2007–2019. https://doi.org/10.1534/g3.115.019570

Helm, M.M., Bourne, N., Lovatelli, A., 2004. Hatchery culture of bivalves: A practical manual. FAO Fisheries Technical Paper, FAO.

Hollenbeck, C.M., Johnston, I.A., 2018. Genomic tools and selective breeding in molluscs. Front. Genet. 9. https://doi.org/10.3389/fgene.2018.00253

Hong Xia, J., Bai, Z., Meng, Z., Zhang, Y., Wang, L., Liu, F., Jing, W., Yi Wan, Z., Li, J., Lin, H., Hua Yue, G., 2015. Signatures of selection in tilapia revealed by whole genome resequencing. Sci. Rep. 5, 14168. https://doi.org/10.1038/srep14168

Houston, R.D., Bean, T.P., Macqueen, D.J., Gundappa, M.K., Jin, Y.H., Jenkins, T.L., Selly, S.L.C., Martin, S.A.M., Stevens, J.R., Santos, E.M., Davie, A., Robledo, D., 2020. Harnessing genomics to fast-track genetic improvement in aquaculture. Nat. Rev. Genet. 21, 389–409. https://doi.org/10.1038/s41576-020-0227-y

Hu, B., Tian, Y., Li, Q., Liu, S., 2022. Genomic signatures of artificial selection in the Pacific oyster, *Crassostrea gigas*. Evol. Appl. 15, 618–630. https://doi.org/10.1111/eva.13286

Hu, Y., Li, Q., Xu, C., Liu, S., Kong, L., Yu, H., 2022. Genetic variability of mass-selected and wild populations of Iwagaki oyster (*Crassostrea nippona*) revealed by microsatellites and mitochondrial COI sequences. Aquaculture 738737. https://doi.org/10.1016/j.aquaculture.2022.738737

Jones, D., Toole, P., Khatkar, M., Raadsma, H., Jerry, D., Zenger, K., 2017. Developing a genomic selection breeding program for complex pearl colour traits within the silver-lipped pearl oyster. Proc. Assoc. Advmt. Anim. Breed. Genet. 22, 537–540. http://www.aaabg.org/aaabghome/AAABG22papers/125Jones22537.pdf.

Jones, D.B., Jerry, D.R., Khatkar, M.S., Raadsma, H.W., Zenger, K.R., 2013. A high-density SNP genetic linkage map for the silver-lipped pearl oyster, *Pinctada maxima*: a valuable resource for gene localisation and marker-assisted selection. BMC Genom. 14, 810. https://doi.org/10.1186/1471-2164-14-810

Jourdan, A., Bugeon, J., Lamy, J.-B., Enez, F., Dégremont, L., Boudry, P., Haffray, P., Chesnier, F., Vetois, E., Morvezen, R., 2021. An image analysis pipeline for molluscs shells

phenotyping: Development and validation on the pacific oyster Crassostrea gigas. Presented at the Aquaculture Europe 2021.

Kong, N., Li, Q., Yu, H., Kong, L.-F., 2015. Heritability estimates for growth-related traits in the Pacific oyster (*Crassostrea gigas*) using a molecular pedigree. Aquac. Res. 46, 499–508. https://doi.org/10.1111/are.12205

Kriaridou, C., Tsairidou, S., Houston, R.D., Robledo, D., 2020. Genomic prediction using low density marker panels in aquaculture: Performance across species, traits, and genotyping platforms. Front. Genet. 11. https://doi.org/10.3389/fgene.2020.00124

Langdon, C., Evans, F., Jacobson, D., Blouin, M., 2003. Yields of cultured Pacific oysters *Crassostrea gigas* Thunberg improved after one generation of selection. Aquaculture 220, 227–244. https://doi.org/10.1016/S0044-8486(02)00621-X

Lapègue, S., Harrang, E., Heurtebise, S., Flahauw, E., Donnadieu, C., Gayral, P., Ballenghien, M., Genestout, L., Barbotte, L., Mahla, R., Haffray, P., Klopp, C., 2014. Development of SNP-genotyping arrays in two shellfish species. Mol. Ecol. Resour. 14, 820–830. https://doi.org/10.1111/1755-0998.12230

Launey, S., Barre, M., Gerard, A., Naciri-Graven, Y., 2001. Population bottleneck and effective size in Bonamia ostreae-resistant populations of *Ostrea edulis* as inferred by microsatellite markers. Genet. Res. 78, 259–270. https://doi.org/10.1017/S0016672301005353

Legarra, A., Robert-Granié, C., Manfredi, E., Elsen, J.-M., 2008. Performance of genomic selection in mice. Genet. 180, 611–618. https://doi.org/10.1534/genetics.108.088575

Li, Q., Wang, Q., Liu, S., Kong, L., 2011. Selection response and realized heritability for growth in three stocks of the Pacific oyster *Crassostrea gigas*. Fish. Sci. 77, 643–648. https://doi.org/10.1007/s12562-011-0369-0

Matson, S.E., 2011. Development, evaluation and application of a mixed-family selective breeding method for the Pacific oyster (*Crassostrea gigas*). Dissertation work for the Oregon state university for the doctor of philosophy in animal science [WWW Document]. URL https://seagrant.oregonstate.edu/sgpubs/development-evaluation-and-application-mixed-family-selective-breeding-method-pacific-oyster.

McCarty, A.J., Allen, S.K., Jr,, Plough, L.V., 2022. Genome-wide analysis of acute low salinity tolerance in the eastern oyster *Crassostrea virginica* and potential of genomic selection for trait improvement. G3: Genes Genomes Genet. 12, jkab368. https://doi.org/10.1093/g3journal/jkab368

Melo, C.M.R.D., Morvezen, R., Durland, E., Langdon, C., 2018. Genetic by Environment Interactions for Harvest Traits of the Pacific Oyster *Crassostrea gigas* (Thunberg) across Different Environments on the West Coast, USA. J. Shellfish Res. 37, 49–61. https://doi.org/10.2983/035.037.0104

Mendoza, F., Dejmek, P., Aguilera, J.M., 2006. Calibrated color measurements of agricultural foods using image analysis. Postharvest Biol. Technol. 41, 285–295. https://doi.org/10.1016/j.postharvbio.2006.04.004

Meuwissen, T., Hayes, B., Goddard, M., 2016. Genomic selection: A paradigm shift in animal breeding. Anim. Front. 6, 6–14. https://doi.org/10.2527/af.2016-0002

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., Lee, D.H., 2002. BLUPF90 and related programs (BGF90). Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, August, 2002. Session 28 1–2.

Naciri-Graven, Y., Martin, A.-G., Baud, J.-P., Renault, T., Gérard, A., 1998. Selecting the flat oyster Ostrea edulis (L.) for survival when infected with the parasite *Bonamia ostreae*. J. Exp. Mar. Biol. Ecol.224, 91–107. https://doi.org/10.1016/S0022-0981(97)00171-8

Nascimento-Schulze, J.C., Bean, T.P., Houston, R.D., Santos, E.M., Sanders, M.B., Lewis, C., Ellis, R.P., 2021. Optimizing hatchery practices for genetic improvement of marine bivalves. Rev. Aquac. 13, 2289–2304. https://doi.org/10.1111/raq.12568

Nell, J.A., Smith, I.R., McPhee, C.C., 2000. The Sydney rock oyster *Saccostrea glomerata* (Gould 1850) breeding programme: progress and goals. Aquac. Res. 31, 45–49. https://doi.org/10.1046/j.1365-2109.2000.00387.x

Ødegård, J., Moen, T., Santi, N., Korsvoll, S.A., Kjøglum, S., Meuwissen, T.H.E., 2014. Genomic prediction in an admixed population of Atlantic salmon (*Salmo salar*). Front. Genet. 5. https://doi.org/10.3389/fgene.2014.00402

Penaloza, C., Barria, A., Papadopoulou, A., Hooper, C., Preston, J., Green, M., Helmer, L., Hammerson, J.K., Schulze, J.N., Minardi, D., Gundappa, M.K., Macqueen, D., Hamilton, J., Houston, R.D., Bean, T.P., 2022. Genome-wide association and genomic prediction of growth traits in the European flat oyster (*Ostrea edulis*) (preprint). Genet. https://doi.org/10.1101/2022.06.10.495672

Peñaloza, C., Gutierrez, A.P., Eöry, L., Wang, S., Guo, X., Archibald, A.L., Bean, T.P., Houston, R.D., 2021. A chromosome-level genome assembly for the Pacific oyster *Crassostrea gigas*. GigaScience 10, giab020. https://doi.org/10.1093/gigascience/giab020

Qi, H., Li, L., Zhang, G., 2021. Construction of a chromosome-level genome and variation map for the Pacific oyster *Crassostrea gigas*. Mol. Ecol. Resour. 21, 1670–1685. https://doi.org/10.1111/1755-0998.13368

Qi, H., Song, K., Li, C., Wang, W., Li, B., Li, L., Zhang, G., 2017. Construction and evaluation of a high-density SNP array for the Pacific oyster (*Crassostrea gigas*). Plos One 12, e0174007. https://doi.org/10.1371/journal.pone.0174007

R Core Team, 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria.

Robert, R., Gérard, A., 1999. Bivalve hatchery technology: The current situation for the Pacific oyster *Crassostrea gigas* and the scallop *Pecten maximus* in France. Aquat. Living Resour. 12, 121–130. https://doi.org/10.1016/S0990-7440(99)80021-7

Robertson, A.R., 1977. The CIE 1976 Color-Difference Formulae. Color Res. Appl. 2, 7–11. https://doi.org/10.1002/j.1520-6378.1977.tb00104.x

Robledo, D., Matika, O., Hamilton, A., Houston, R.D., 2018. Genome-wide association and genomic selection for resistance to amoebic gill disease in Atlantic salmon. G3: Genes Genomes Genet. 8, 1195–1203. https://doi.org/10.1534/g3.118.200075

Saenko, S.V., Schilthuizen, M., 2021. Evo-devo of shell colour in gastropods and bivalves. Curr. Opin. Genet. Dev. 69, 1–5. https://doi.org/10.1016/j.gde.2020.11.009

Sheridan, A.K., 1997. Genetic improvement of oyster production—a critique. Aquaculture 153, 165–179. https://doi.org/10.1016/S0044-8486(97)00024-0

Smits, M., Enez, F., Ferraresso, S., Dalla Rovere, G., Vetois, E., Auvray, J.-F., Genestout, L., Mahla, R., Arcangeli, G., Paillard, C., Haffray, P., Bargelloni, L., 2020. Potential for genetic improvement of resistance to *Perkinsus olseni* in the Manila clam, *Ruditapes philippinarum*, using DNA parentage assignment and mass spawning. Front. Vet. Sci. 7. https://doi.org/10.3389/fvets.2020.579840

Sonesson, A.K., Ødegård, J., 2016. Mating structures for genomic selection breeding programs in aquaculture. Genet. Sel. Evol. 48, 46. https://doi.org/10.1186/s12711-016-0224-y

Song, H., Dong, T., Yan, X., Wang, W., Tian, Z., Sun, A., Dong, Y., Zhu, H., Hu, H., 2023. Genomic selection and its research progress in aquaculture breeding. Rev. Aquac. 15, 274–291. https://doi.org/10.1111/raq.12716

Song, H., Hu, H., 2022. Strategies to improve the accuracy and reduce costs of genomic prediction in aquaculture species. Evol. Appl. 15, 578–590. https://doi.org/10.1111/eva.13262

Stenger, P.-L., Ky, C.-L., Reisser, C., Duboisset, J., Dicko, H., Durand, P., Quintric, L., Planes, S., Vidal-Dupiol, J., 2021. Molecular pathways and pigments underlying the colors of the pearl oyster *Pinctada margaritifera* var. *cumingii* (Linnaeus 1758). Genes 12, 421. https://doi.org/10.3390/genes12030421

Tsai, H.-Y., Hamilton, A., Tinch, A.E., Guy, D.R., Bron, J.E., Taggart, J.B., Gharbi, K., Stear, M., Matika, O., Pong-Wong, R., Bishop, S.C., Houston, R.D., 2016. Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. Genet. Sel. Evol. 48, 47. https://doi.org/10.1186/s12711-016-0226-9

Tsai, H.-Y., Hamilton, A., Tinch, A.E., Guy, D.R., Gharbi, K., Stear, M.J., Matika, O., Bishop, S.C., Houston, R.D., 2015. Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. BMC Genom. 16, 969. https://doi.org/10.1186/s12864-015-2117-9

Vallejo, R.L., Leeds, T.D., Gao, G., Parsons, J.E., Martin, K.E., Evenhuis, J.P., Fragomeni, B.O., Wiens, G.D., Palti, Y., 2017. 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. https://doi.org/10.1186/s12711-017-0293-6

Vandeputte, M., Haffray, P., 2014. Parentage assignment with genomic markers: a major advance for understanding and exploiting genetic variation of quantitative traits in farmed aquatic animals. Front. Genet. 5. https://doi.org/10.3389/fgene.2014.00432

VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91, 4414–4423. https://doi.org/10.3168/jds.2007-0980

Vendrami, D.L.J., Houston, R.D., Gharbi, K., Telesca, L., Gutierrez, A.P., Gurney-Smith, H., Hasegawa, N., Boudry, P., Hoffman, J.I., 2019. Detailed insights into pan-European population structure and inbreeding in wild and hatchery Pacific oysters (*Crassostrea gigas*) revealed by genome-wide SNP data. Evol. Appl. 12, 519–534. https://doi.org/10.1111/eva.12736

Vera, M., Maroso, F., Wilmes, S.B., Hermida, M., Blanco, A., Fernández, C., Groves, E., Malham, S.K., Bouza, C., Consortium, T.C., Robins, P.E., Martínez, P., 2022. Genomic survey of edible cockle (*Cerastoderma edule*) in the Northeast Atlantic: A baseline for sustainable management of its wild resources. Evol. Appl. 15, 262–285. https://doi.org/10.1111/eva.13340

Vu, S.V., Gondro, C., Nguyen, N.T.H., Gilmour, A.R., Tearle, R., Knibb, W., Dove, M., Vu, I.V., Khuong, L.D., O'Connor, W., 2021. Prediction accuracies of genomic selection for nine commercially important traits in the Portuguese oyster (*Crassostrea angulata*) Using DArT-Seq Technology. Genes 12, 210. https://doi.org/10.3390/genes12020210

Wan, S., Li, Q., Liu, T., Yu, H., Kong, L., 2017. Heritability estimates for shell color-related traits in the golden shell strain of Pacific oyster (*Crassostrea gigas*) using a molecular pedigree. Aquaculture 476, 65–71. https://doi.org/10.1016/j.aquaculture.2017.04.012

Wan, S., Li, Q., Yu, H., Liu, S., Kong, L., 2020. Estimating heritability for meat composition traits in the golden shell strain of Pacific oyster (*Crassostrea gigas*). Aquaculture 516, 734532. https://doi.org/10.1016/j.aquaculture.2019.734532

Wang, H., Du, X., Lü, W., Liu, Z., 2010. Estimating the heritability for growth-related traits in the pearl oyster, *Pinctada fucata martensii* (Dunker). Aquac. Res. 42, 57–64. https://doi.org/10.1111/j.1365-2109.2010.02552.x

Wang, H., Liu, J., Li, Y., Zhu, X., Liu, Z., 2014. Responses to two-way selection on growth in mass-spawned F1 progeny of *Argopecten irradians concentricus* (Say). J. Oceanol. Limnol 32, 349–357. https://doi.org/10.1007/s00343-014-3153-z

Wang, Q., Yu, Y., Yuan, J., Zhang, X., Huang, H., Li, F., Xiang, J., 2017. 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. https://doi.org/10.1186/s12863-017-0507-5

Wang, X., Xu, W., Wei, L., Zhu, C., He, C., Song, H., Cai, Z., Yu, W., Jiang, Q., Li, L., Wang, K., Feng, C., 2019. Nanopore sequencing and *de novo* assembly of a black-shelled Pacific oyster (*Crassostrea gigas*) Genome. Front. Genet. 10. https://doi.org/10.3389/fgene.2019.01211

Wang, Z., Hu, H., Sun, T., Li, X., Lv, G., Bai, Z., Li, J., 2022. Genomic selection for improvement of growth traits in triangle sail mussel (*Hyriopsis cumingii*). Aquaculture 561, 738692. https://doi.org/10.1016/j.aquaculture.2022.738692

Williams, S.T., 2017. Molluscan shell colour. Biol. Rev. 92, 1039–1058. https://doi.org/10.1111/brv.12268

Xing, D., Li, Q., Kong, L., Yu, H., 2018. Heritability estimate for mantle edge pigmentation and correlation with shell pigmentation in the white-shell strain of Pacific oyster, *Crassostrea gigas*. Aquaculture 482, 73–77. https://doi.org/10.1016/j.aquaculture.2017.09.026

Xu, J., Zhao, Z., Zhang, X., Zheng, X., Li, J., Jiang, Y., Kuang, Y., Zhang, Y., Feng, J., Li, C., Yu, J., Li, Q., Zhu, Y., Liu, Y., Xu, P., Sun, X., 2014. Development and evaluation of the first high-throughput SNP array for common carp (*Cyprinus carpio*). BMC Genom. 15, 307. https://doi.org/10.1186/1471-2164-15-307

Xu, L., Li, Q., Yu, H., Kong, L., 2017. Estimates of heritability for growth and shell color traits and their genetic correlations in the black shell strain of Pacific oyster *Crassostrea gigas*. Mar. Biotechnol. 19, 421–429. https://doi.org/10.1007/s10126-017-9772-6

Yan, J., Zhang, J., Sun, Q., Yan, X., Sheng, J., 2021. The unique UV–Vis reflection features of the nacre of Hyriopsis cumingii shells, and its formation mechanisms. Dyes Pigm.184, 108753. https://doi.org/10.1016/j.dyepig.2020.108753

Yáñez, J.M., Barría, A., López, M.E., Moen, T., Garcia, B.F., Yoshida, G.M., Xu, P., 2022. Genome-wide association and genomic selection in aquaculture. Rev. Aquac. https://doi.org/10.1111/raq.12750

Yang, B., Zhai, S., Zhang, F., Wang, H., Ren, L., Li, Y., Li, Q., Liu, S., 2022. Genome-wide association study toward efficient selection breeding of resistance to *Vibrio alginolyticus* in Pacific oyster, *Crassostrea gigas*. Aquaculture 548, 737592. https://doi.org/10.1016/j.aquaculture.2021.737592

Yin, X., Arias-Pérez, A., Kitapci, T.H., Hedgecock, D., 2020a. High-density linkage maps based on genotyping-by-sequencing (GBS) confirm a chromosome-level genome assembly and reveal variation in recombination rate for the Pacific oyster *Crassostrea gigas*. G3: Genes Genomes Genet. 10, 4691–4705. https://doi.org/10.1534/g3.120.401728

Yoshida, G.M., Carvalheiro, R., Lhorente, J.P., Correa, K., Figueroa, R., Houston, R.D., Yáñez, J.M., 2018. 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. https://doi.org/10.1016/j.aquaculture.2018.03.004

Zenger, K.R., Khatkar, M.S., Jones, D.B., Khalilisamani, N., Jerry, D.R., Raadsma, H.W., 2019. Genomic selection in aquaculture: Application, limitations and opportunities with special reference to marine shrimp and pearl oysters. Front. Genet. 9. https://doi.org/10.3389/fgene.2018.00693

Zhang, Guofan, Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H., Xiong, Z., Que, H., Xie, Y., Holland, P.W.H., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang, Jiafeng, Peng, C., Meng, J., Yang, L., Liu, J., Wen, B., Zhang, N., Huang, Z., Zhu, Q., Feng, Y., Mount, A., Hedgecock, D., Xu, Z., Liu, Y., Domazet-Lošo, T., Du, Y., Sun, X., Zhang, Shoudu, Liu, B., Cheng, P., Jiang, X., Li, J., Fan, D., Wang, W., Fu, W., Wang, T., Wang, B., Zhang, J., Peng, Z., Li, Yingxiang, Li, Na, Wang, Jinpeng, Chen, M., He, Y., Tan, F., Song, X., Zheng, Q., Huang, R., Yang, Hailong, Du, X., Chen, L., Yang, M., Gaffney, P.M., Wang, S., Luo, L., She, Z., Ming, Y., Huang, W., Zhang, Shu, Huang, B., Zhang, Y., Qu, T., Ni, P., Miao, G., Wang, Junyi, Wang, Q., Steinberg, C.E.W., Wang, H., Li, Ning, Qian, L., Zhang, Guojie, Li, Yingrui, Yang, Huanming, Liu, X., Wang, Jian, Yin, Y., Wang, Jun, 2012. The

oyster genome reveals stress adaptation and complexity of shell formation. Nature 490, 49–54. https://doi.org/10.1038/nature11413

Zheng, H., Zhang, G., Liu, X., Zhang, F., Guo, X., 2004. Different responses to selection in two stocks of the bay scallop, Argopecten irradians irradians Lamarck (1819). J. Exp. Mar. Biol. Ecol.313, 213–223. https://doi.org/10.1016/j.jembe.2004.04.015

Zhong, X., Li, Q., Kong, L., Yu, H., 2017. Estimates of linkage disequilibrium and effective population size in wild and selected populations of the Pacific oyster using single-nucleotide polymorphism markers. J. World Aquac. Soc. 48, 791–801. https://doi.org/10.1111/jwas.12393