Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Biomedicina Comparata ed Alimentazione (BCA)
Dipartimento di Agronomia, Animali, Alimenti, Risorse Naturali ed Ambiente (DAFNAE)

CORSO DI DOTTORATO DI RICERCA IN ANIMAL AND FOOD SCIENCE
CICLO XXXI

BREEDING IN EUROPEAN SEA BASS (*Dicentrarchus labrax* L.): GENETIC ASPECTS OF RESISTANCE TO VNN AND SEX DETERMINATION, AND DEVELOPMENT OF GENOMIC PREDICTION TOOLS

Tesi redatta con il contributo finanziario di Valle Cà Zuliani Società Agricola s.r.l.

Coordinatore: Ch.mo Prof. Stefano Schiavon

Supervisore: Ch.mo Prof. Paolo Carnier

Co-Supervisore: Ch.ma Prof.ssa Daniela Bertotto

Dottorando: Sara Faggion

To Aria

Table of contents

Abstract ... **1**
Riassunto ... **3**
List of abbreviations and acronyms ... **6**
1. General introduction ... **8**
1.1 Aquaculture: the fastest growing food-producing sector and its challenges ... 9
1.1.1 Disease outbreaks: one of the major issues for aquaculture ... 9
1.2 Biology and production of the European sea bass (*Dicentrarchus labrax* L.) ... 11
1.2.1 One of the major challenges in farmed European sea bass: viral nervous necrosis ... 11
1.2.2 Sex determination systems and the unusual case of the European sea bass ... 13
References ... 14
2. Selection for disease and stress resistance in farmed fish: a review ... **17**
Abstract ... 18
2.1 Introduction ... 19
2.2 Background: disease and stress response ... 20
2.2.1 The mechanism of response to infection and disease ... 20
2.2.2 The response to stress ... 21
2.2.3 Stress: why it could be interesting for disease susceptibility? ... 21
2.3 Phenotyping of disease and stress resistance ... 22
2.3.1 Experimental challenge tests to assess disease resistance ... 22
2.3.2 Assessment of stress response ... 24
2.4 Selection for disease and stress resistance: state of the art ... 25
2.4.1 Genetic parameters for disease resistance ... 25
2.4.2 Selection approaches for disease resistance: traditional and genomic approaches ... 33
2.4.3 Indicator traits for disease resistance and indirect selection approaches ... 35
2.4.4 Selection for stress resistance ... 36
2.5 Fish reproductive technology: the key for successful selective breeding programs ... 37
References ... 38
3. Genetic and genomic parameters for VNN resistance, body weight, cortisol concentrations and antibody titer in European sea bass (*Dicentrarchus labrax* L.) ... **47**
Abstract ... 48
3.1 Introduction ... 49
3.2 Materials and Methods ... 50
3.2.1 Production and rearing of experimental fish ... 50
3.2.2 NNV challenge test ... 50
3.2.3 Samples management ... 51
3.2.4 Indirect ELISA assay ... 51
3.2.5 Cortisol extraction and measurement ... 52
3.2.6 Genomic DNA extraction, 2b-RAD libraries construction and sequencing ... 52
3.2.7 SNP discovery, genotyping, parentage assignment and pedigree reconstruction ... 53
3.2.8 Statistical analyses ... 54
3.3 Results ... 58
3.3.1 NNV challenge test ... 58

3.3.2 Parentage assignment and pedigree reconstruction 59
3.3.3 Statistical analyses 60
3.4 Discussion 71
3.5 Conclusions 73
References 73
4. Genomic prediction of VNN resistance, body weight, cortisol concentrations and antibody titer in European sea bass (*Dicentrarchus labrax* L.) 77
Abstract 78
4.1 Introduction 79
4.2 Materials and Methods 80
4.2.1 Experimental fish: production, rearing and NNV challenge test 80
4.2.2 Antibody titer against VNN and cortisol concentrations 81
4.2.3 Genotyping, parentage assignment and pedigree reconstruction 81
4.2.4 Estimation of "traditional" breeding values 82
4.2.5 Genomic predictions 83
4.2.5.1 Genomic prediction models 83
4.2.5.2 Assessment of prediction accuracy and performance in classification 84
4.3 Results 85
4.3.1 Genomic prediction and comparison of models 85
4.4 Discussion 86
4.5 Conclusions 88
References 97
5. Population-specific variations of the genetic architecture of sex determination in wild European sea bass (*Dicentrarchus labrax* L.) 100
Abstract 101
5.1 Introduction 102
5.2 Materials and Methods 103
5.2.1 Broodstock origin, production and rearing of experimental fish 103
5.2.2 Genotyping, parentage assignment and descriptive statistics 104
5.2.3 Principal component analysis 104
5.2.4 Heritability, genetic and phenotypic correlations 104
5.2.5 Genome-wide association study (GWAS) 105
5.3 Results 106
5.3.1 Parentage assignment and descriptive statistics 106
5.3.2 Principal component analysis 107
5.3.3 Heritability, genetic and phenotypic correlations 109
5.3.4 Genome-wide association study (GWAS) 109
5.4 Discussion 112
References 114
Supplementary material: cross-validation of QTLs 117
6. General conclusions 122
Appendix 125
Acknowledgements - Ringraziamenti 136

Abstract

During the recent years, the amount of fish for human consumption provided by aquaculture has impressively increased, and this led to a growing attention towards the development and the improvement of production techniques, including selective breeding programs, that are seen as principal drivers in the improvement of production traits (growth, morphology) and efficiency traits (processing yield, feed efficiency, product quality and disease resistance) of the farmed species.
Disease outbreaks in aquaculture may have severe consequences, interfering with the progress and sustainability of the farm and causing huge production and economic losses. For these reasons, the interest towards the genetic improvement of disease resistance has significantly increased, thanks to the availability of new and more advanced genomic tools and techniques, the development of genome-wide sequencing and genotyping technologies and the generation of high-density marker data.
In aquaculture species, selective breeding with the aim to establish disease resistant lines is an effective approach, in particular for those viral and bacterial infections that are not treatable with chemotherapeutics and antibiotics or for which vaccines are not currently available on a commercial scale. Selective breeding allows a long-term control of the disease and a cumulative and permanent improvement of the trait over generation at the population level.
Betanodaviruses (nervous necrosis viruses, NNV) cause one of the major and severe infections (viral nervous necrosis, VNN) in more than 50 fish species, including the European sea bass (*Dicentrarchus labrax* L.). The effects of VNN outbreaks in sea bass hatcheries and farms are serious, with mortality rates up to 100% in the larval and juvenile stages.
The first aim of the project was to estimate the genetic and genomic parameters for VNN resistance traits. A sea bass experimental population derived from a commercial breeding stock (N = 650) was subjected to a VNN challenge test and a genome-wide SNP dataset for sea bass was generated through a high-throughput sequencing approach (2b-RAD).
Estimates of heritability for mortality as a binary trait (0/1) and survival time were moderate or low (genetic h^2_{mort} = 0.14-0.23, genomic h^2_{mort} = 0.06; genetic h^2_{surv} = 0.07-0.14, genomic h^2_{surv} = 0.03).
Due to the possible application of indirect approaches to improve VNN resistance, we investigated the variation in a physiological parameter (post-stress cortisol concentration) and an immunological parameter (antibody titer to NNV) to determine the additive genetic variation of these traits, and the genetic correlation with disease resistance. Both traits showed a significant genetic and genomic variability (genetic h^2_{HC} = 0.19-0.23, genomic h^2_{HC} = 0.19; genetic h^2_{AT} = 0.28-0.39, genomic h^2_{AT} = 0.26). Cortisol concentration showed a null genetic correlation with VNN mortality, while the genetic relationship between antibody titer and mortality was negative ($r_{a\ mort/AT}$ = -0.39).
The genetic and genomic heritability estimates for body weight (548 d post-hatching) were confirmed high (genetic h^2_{BW} = 0.45-0.60, genomic h^2_{BW} = 0.45); moreover, body weight was positively genetically correlated with cortisol concentrations and antibody titer ($r_{a\ BW/HC}$ = 0.12 and $r_{a\ BW/AT}$ = 0.49), while negatively correlated with mortality ($r_{a\ BW/mort}$ = -0.39).
The second aim of the project was to investigate the potential and the effectiveness of genomic tools in predicting the genetic merit for the traits of concern.
Five Bayesian models (BayesA, BayesB, BayesC, Bayesian LASSO, Bayesian Ridge Regression) were compared and their accuracy (in terms of correlation between observed value and prediction) when predicting the phenotype or the EBV (estimated breeding value) for VNN mortality, post-stress cortisol concentrations, antibody titer or body weight was estimated, without observing any

differences between the five models. The prediction accuracy of the EBV for VNN mortality reached values equal to 0.89; the prediction accuracies of the EBV and the phenotype for post-stress cortisol concentration were 0.88 and 0.22, respectively, for antibody titer 0.76 and 0.26, respectively, and for body weight 0.69 and 0.39, respectively.
Different metrics (the area under the ROC curve, the proportion of true results among the total number of samples, the Matthew's correlation coefficient) were employed to evaluate the classification performance of prediction of the EBV for VNN mortality when used to classify the phenotype for the same trait, resulting in better classification performance, as measured by the AUC metric, in comparison with the classification based on genomic predictions of the phenotype.
Overall, the results indicate that genomic prediction could have a great potential for traits like VNN mortality, body weight, cortisol concentration and antibody titer in European sea bass, leading to the reduction of the need of routine phenotyping and outperform traditional approaches in the enhancement of disease resistance, although further investigations on a larger experimental population are needed. Moreover, further studies should elucidate the effects of the method of infection used in challenge tests (immersion or injection) on the variation in the infection response observed and genetic parameters thereof.
The third part of the project addressed the genetic architecture of sex determination in European sea bass. Sea bass sex determination model is complex and consistent with a polygenic model, where phenotypic sex is determined by the sum of small individual gene effects, but it is significantly affected by environmental temperature, as well.
The comprehension of the genetic architecture of sex tendency in sea bass could be interesting even in the context of aquaculture. Actually, the strong bias towards males under aquaculture conditions has been recognised by farmers as problematic for different reasons (lower growth rates of males compared to females, reduced flesh quality and general decrease of the commercial values of the product).
Because polygenic sex determination is considered evolutionarily unstable, the genetic architecture could be different in populations belonging to various geographical areas and exposed to different environmental/thermal conditions. Experimental progeny crosses (N = 927) were produced using wild dams from the Western Mediterranean and wild sires from four geographical areas (North Atlantic, Western Mediterranean, North-Eastern Mediterranean, South-Eastern Mediterranean), consistent with the natural range of the species in North Atlantic and Mediterranean Sea and characterized by different environmental conditions.
The sex-ratio was strongly skewed towards males both in the global dataset and within each paternal group, with significant variation among populations, dams and sires. Sex, body weight and body length showed moderate heritability ($h^2_{\mathrm{sex}} = 0.52 \pm 0.17$, $h^2_{\mathrm{BW}} = 0.46 \pm 0.17$, $h^2_{\mathrm{BL}} = 0.34 \pm 0.15$) and sex tendency was genetically correlated with body weight and length ($r_{\mathrm{a\ sex/BW}} = 0.69 \pm 0.12$, $r_{\mathrm{a\ sex/BL}} = 0.66 \pm 0.13$).
A weighted genome-wide association study (wGWAS) performed both on the global dataset and within each paternal group revealed a different genetic architecture of sex determination between Atlantic and Mediterranean populations, with a more similar genetic architecture among geographically close populations compared to geographically distant populations, consistent with the hypothesis of a population-specific evolution of the polygenic sex determination system in different environments.

Riassunto

Negli ultimi anni la richiesta di prodotti derivanti dall'acquacoltura è aumentata considerevolmente, comportando una maggiore attenzione volta allo sviluppo ed al miglioramento delle tecniche di produzione, compresi programmi di selezione genetica volti al miglioramento di caratteri produttivi (peso, morfologia) e di efficienza delle specie allevate (resa produttiva, efficienza alimentare, qualità del prodotto, resistenza alle malattie).

In particolare, la comparsa di infezioni virali e batteriche in avannotteria ed in allevamento, può avere gravi conseguenze, sia in termini economici e produttivi, sia dal punto di vista del progresso e della sostenibilità dell'allevamento stesso. Per questo motivo, l'interesse verso il miglioramento genetico della resistenza alle malattie in specie ittiche è notevolmente aumentato, anche grazie al costante progresso di strumenti e tecniche di biologia molecolare che consentono la genotipizzazione e la generazione di panel di marcatori *genome-wide*. La selezione per produrre lotti di animali geneticamente migliorati risulta una strategia particolarmente efficace e di specifico interesse per patologie virali o batteriche per le quali le comuni misure di gestione come farmaci ed antibiotici non sono contemplate, presentano scarsa efficacia, o come i vaccini, risultano essere al momento soltanto in fase di sviluppo sperimentale. La selezione genetica per la resistenza alle malattie consente invece un controllo a lungo termine della patologia attraverso un cambiamento permanente delle caratteristiche genetiche degli animali che può essere trasmesso di generazione in generazione.

L'encefalopatia e retinopatia virale (o *viral nervous necrosis*, VNN) è una delle maggiori e più gravi patologie riscontrata in più di 50 specie ittiche, tra cui il branzino (*Dicentrarchus labrax* L.). Le conseguenze di focolai di VNN in branzino sono decisamente significative, con tassi di mortalità fino al 100% negli stadi larvali e giovanili.

Il primo obiettivo del progetto è stato quello di valutare i parametri genetici e genomici per i caratteri di resistenza a VNN. 650 animali prodotti in un allevamento commerciale sono stati quindi sottoposti ad infezione sperimentale con Betanodavirus (*nervous necrosis virus*, NNV) ed è stato prodotto un dataset di marcatori *genome-wide* (SNP) mediante approcci di sequenziamento *high-throughput* (tecnica 2b-RAD).

L'ereditabilità per la mortalità intesa come carattere binario (0/1) o come tempo di sopravvivenza ha mostrato valori moderati o lievi (genetica: h^2_{mort} = 0,14-0,23, h^2_{surv} = 0,07-0,14; genomica: h^2_{mort} = 0,06, h^2_{surv} = 0,03).

Visto l'interesse verso possibili indicatori da utilizzare come criteri di selezione indiretta, sono stati studiati anche parametri fisiologici (livello di cortisolo ematico post-stress) ed immunitari (titolo anticorpale a Betanodavirus) con lo scopo di determinarne la variabilità genetica additiva e la correlazione genetica con i caratteri di resistenza a Betanodavirus.

Entrambi i parametri hanno mostrato valori di ereditabilità genetica e genomica interessanti (genetica: h^2_{HC} = 0,19-0,23, h^2_{AT} = 0.28-0.39, genomica: h^2_{HC} = 0,19, h^2_{AT} = 0,26). La correlazione genetica tra concentrazione di cortisolo e probabilità di morte è risultata nulla, mentre quella tra titolo anticorpale e probabilità di morte è risultata negativa ($r_{\mathrm{a\ mort/AT}}$ = -0,39). L'ereditabilità genetica e genomica per il peso (548 giorni *post-hatching*) si è confermata elevata (genetica: h^2_{BW} = 0,45-0,60; genomica: h^2_{BW} = 0,45); il peso, inoltre, è risultato positivamente correlato dal punto vista genetico ai livelli di cortisolo ed al titolo anticorpale ($r_{\mathrm{a\ BW/HC}}$ = 0,12 e $r_{\mathrm{a\ BW/AT}}$ = 0,49), mentre la correlazione genetica tra questo carattere e la mortalità è risultata negativa ($r_{\mathrm{a\ BW/mort}}$ = -0,39).

Il secondo obiettivo del progetto è stato quello di investigare il potenziale predittivo di strumenti genomici per i caratteri considerati.

Sono stati confrontati cinque modelli Bayesiani (BayesA, BayesB, BayesC, Bayesian LASSO, Bayesian Ridge Regression) e ne è stata stimata l'accuratezza (in termini di correlazione tra valori osservati e predetti) quando i modelli sono stati utilizzati per predire il fenotipo o l'EBV (*estimated breeding value*) per mortalità a VNN, livelli di cortisolo, titolo anticorpale o peso, senza però osservare differenze significative tra i modelli. L'accuratezza nel predire gli EBV per la mortalità ha raggiunto valori di 0,89; l'accuratezza nel predire gli EBV ed il fenotipo per i livelli di cortisolo è risultata 0,88 e 0,22, rispettivamente, per il titolo anticorpale 0,76 e 0,26, rispettivamente, per il peso 0,69 e 0,39, rispettivamente.
Diverse metriche (area sottesa alla curva ROC, proporzione di risultati corretti sul numero totale di casi, coefficiente di correlazione di Matthew) sono state impiegate per valutare le performance di predizione degli EBV per la mortalità con lo scopo di classificare il fenotipo del carattere stesso, risultando in una migliore performance di classificazione, come indicato dalla metrica AUC, rispetto alla classificazione basata sulle predizioni genomiche del fenotipo.
Nel complesso, questi risultati indicano che la predizione genomica potrebbe avere un enorme potenziale per caratteri come la mortalità dovuta a VNN, peso, concentrazione di cortisolo e titolo anticorpale in branzino, utile anche per ridurre la necessità di raccolte periodiche di dati fenotipici e fornendo migliori risultati rispetto agli approcci tradizionali per il miglioramento della resistenza alle patologie. Studi ulteriori su un maggior numero di animali sono comunque indispensabili, così come un approfondimento sulle variazioni nella risposta all'infezione, e quindi nei parametri genetici, dovute ai diversi metodi di infezione usati nei *challenge test* (immersione, iniezione).
Il terzo obiettivo di questo progetto è stato invece rappresentato dallo studio dell'architettura genetica del meccanismo di determinazione del sesso in branzino. Il determinismo sessuale in branzino è un meccanismo poligenico complesso, in quanto risultato dell'interazione tra numerosi geni (con effetto limitato ma aventi azione uguale e cumulativa sul valore fenotipico) e temperatura ambientale. Capire i meccanismi genetici complessi che stanno alla base del determinismo sessuale è comunque interessante anche per una eventuale applicazione in acquacoltura, dal momento che l'alta percentuale di produzione di maschi rispetto alle femmine rappresenta un problema per diversi motivi (i maschi presentano precoce maturazione sessuale osservata in allevamento, con conseguente tasso di crescita inferiore, qualità inferiore del prodotto e, in generale, inferiore valore commerciale rispetto alle femmine).
Partendo dal presupposto che un meccanismo di determinazione sessuale poligenico è evolutivamente instabile, l'architettura genetica potrebbe essere differente in popolazioni appartenenti a distinte aree geografiche caratterizzate da condizioni ambientali e termiche diverse. A questo scopo sono stati studiati 927 individui prodotti incrociando riproduttori femmine provenienti dalla zona Ovest del Mar Mediterraneo e riproduttori maschi provenienti da quattro diverse aree geografiche caratterizzate da condizioni ambientali differenti (Nord Atlantico, Ovest Mediterraneo, Nord-Est Mediterraneo e Sud-Est Mediterraneo) e corrispondenti alla distribuzione ed alla struttura genetica delle popolazioni naturali di branzino in Atlantico e Mar Mediterraneo.
In generale, la percentuale di femmine nel dataset globale ed entro ciascun gruppo di diversa origine è sempre risultata inferiore rispetto alla percentuale di maschi, con differenze significative tra popolazioni, madri e padri. I valori di ereditabilità per *sex tendency*, peso e lunghezza (180 giorni *post-hatching*) sono risultati moderati (h^2_{sex} = 0,52 ± 0,17, h^2_{BW} = 0,46 ± 0,17, h^2_{BL} = 0,34 ± 0,15) e la *sex tendency* è risultata geneticamente correlata con peso e lunghezza, mostrando valori significativi ($r_{\mathrm{a\ sex/BW}}$ = 0,69 ± 0,12, $r_{\mathrm{a\ sex/BL}}$ = 0,66 ± 0,13).

Uno studio di associazione (*weighted* GWAS) è stato effettuato sia sul dataset globale sia su ciascun gruppo di diversa origine con lo scopo di individuare marcatori SNP potenzialmente associati al determinismo sessuale. Lo studio ha rivelato una diversa architettura genetica per le popolazioni Atlantiche e Mediterranee, ma con affinità per individui appartenenti ad aree geografiche adiacenti, consistenti con l'ipotesi di evoluzione popolazione-specifica del meccanismo di determinazione sessuale in aree caratterizzate da condizioni ambientali differenti.

List of abbreviations and acronyms

°C	Degree Celsius
µg	Micrograms
µl	Microliters
µM	Micromolar
ACC	Accuracy
AGD	Amoebic gill disease
AT	Antibody titer against NNV
ATP	Adenosine triphosphate
AUC	Area under the ROC curve
BCWD	Bacterial cold water disease
BFNNV	Barfin flounder nervous necrosis virus
BKD	Bacterial kidney disease
BL	Body length
BLUP	Best linear unbiased prediction
bp	Base-pairs
BSA	Bovine serum albumin
BW	Body weight
CORR	Correlation
CV%	Coefficient of variation
CV	Cross-validation
CyHV-3	Cyprinid herpesvirus-3
DGV	Direct genomic value
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
dph	Days post-hatching
EBV	Estimated breeding value
ELISA	Enzime-linked immunosorbent assay
ESD	Environmental sex determination
FP	False positive
FN	False negative
g	Grams
GBLUP	Genomic best linear unbiased prediction
gDNA	Genomic DNA
GEBV	Genomic estimated breeding value
GS	Genomic selection
GSD	Genotypic sex determination
GWAS	Genome-wide association study
h	Hour
HC	Hydrocortisone
HPD%	Highest posterior density %
HPI axis	Hypothalamic-pituitary-interrenal axis
HSPs	Heat shock proteins
IHNV	Infectious hematopoietic necrosis virus
IPNV	Infectious pancreatic necrosis virus
ISAV	Infectious salmon anaemia virus
Kb	Kilo base-pairs
Kg	Kilograms
KHV	Koi herpesvirus
LG	Linkage group
LH-RHa	Luteinizing Hormone–Releasing Hormone analog
MAF	Minor allele frequency
MAS	Marker-Assisted Selection

MCC	Matthew's correlation coefficient
MCMC	Monte-Carlo Markov Chain
mg	Milligrams
MHC	Major histocompatibility complex
ml	Milliliters
mm	Millimeters
mM	Millimolar
MORT	VNN mortality as a binary trait (0/1)
MS-222	Tricaine methanesulfonate
NAbs	Natural antibodies
NAT	Northern Atlantic
NEM	North-Eastern Mediterranean
ng	Nanograms
nm	Nanometers
NNV	Nervous necrosis virus
OD	Optical density
PBS	Phosphate buffer
PCA	Principal component analysis
PCR	Polymerase chain reaction
pg	Picograms
ppm	Parts per million
PIT-tag	Passive integrated transponder tag
PSD	Polygenic sex determination
QTL	Quantitative trait locus
2b-RAD	2b-restriction site associated DNA
RAD	Restriction site associated DNA
RGNNV	Red-spotted grouper nervous necrosis virus
RIA	Radioimmunoassay
RNA	Ribonucleic acid
ROC	Receiver operating characteristics
SEM	South-Eastern Mediterranean
SJNNV	Striped jack nervous necrosis virus
SNP	Single nucleotide polymorphism
SPDv	Salmon pancreas disease virus
SRHC	Square root HC concentration
SURVT	Survival time after NNV infection
TAE	Tris base, acetic acid, EDTA buffer
TRs	Technical replicates
$TCID_{50}$	50% tissue culture infectious dose
TPNNV	Tiger puffer nervous necrosis virus
TP	True positive
TN	True negative
U	Unit
USD	United States dollars
UV	UltraViolet
VER	Viral encephalopathy and retinopathy
VHSV	Viral hemorrhagic septicaemia virus
VLP	Virus-like particle
VNN	Viral nervous necrosis
WEM	Western Mediterranean
wGWAS	Weighted genome-wide association study

1. General introduction

1.1 Aquaculture: the fastest growing food-producing sector and its challenges

Fish is classically considered nutritionally valuable and important as a source of protein, due to its properties, consisting of essential amino acids and essential fats (such as long chain omega-3 fatty acids), vitamins (D, A and B) and minerals (including calcium, iodine, zinc, iron and selenium; Tilami and Sampels 2018).

FAO statistics (http://www.fao.org/fishery/aquaculture/en) reported that in 2013 fish accounted for about 17% of the global population's intake of animal protein and 6.7% of all protein consumed. In the past five decades, the global supply of fish for human consumption have increased at an average annual rate of 3.2% in the period 1961-2013, double that of population growth, resulting in increasing average per capita availability. Preliminary estimates for 2014 and 2015 indicate that the apparent average fish consumption should be more than 20 kg per capita.

Aquaculture is progressively increasing its importance in terms of supply of fish for human consumption, even due to a relatively static capture fishery production since late 1980s. In 1974, aquaculture provided only 7% of fish for human consumption; this rate was subjected to an impressive growth throughout the years, reaching 39% in 2004.

In 2016, the overall world aquaculture production amounted to 80 million tons with an estimated value of USD 231.5 billion, and with an increase throughout all continents. Of the overall production, 54 million tons consisted of finfish (USD 138.5 billion), 17.1 million tons of molluscs (USD 29.2 billion), 7.9 million tons of crustaceans (USD 57.1 billion), and 938,558 tons of other aquatic animals (USD 6.7 billion).

In Europe, the aquaculture production in 2016 amounted to 2.9 million tons with a value of USD 13.5 billion. This represents 3.68% of the worldwide aquaculture production by mass and 5.84% by value. Production by mass in Europe has increased over the last 10 years by 40% (by value 114%).

Atlantic salmon (*Salmo salar*) is the most important farmed species in Europe, with 1.5 million tons (USD 8.9 billion) in 2016, and in the recent years the value of Atlantic salmon production exceeded the combined value of all other aquaculture production in Europe, reaching the 58% of the value of all the species. Beyond Atlantic salmon, other farmed species, such as rainbow trout (*Oncorhynchus mykiss*), European seabass (*Dicentrarchus labrax*) and gilthead seabream (*Sparus aurata*) have shown relatively consistent increases in production (FAO 2018).

A fast-growing sector such as aquaculture may undoubtedly faces many challenges, such as the improvement of environmental management with the reduction of environmental impacts, the appropriate use of technologies and more efficient resource use and farm management, in order to assure appropriate food safety and quality of products and meet the growing demand of food.

Significant issues for aquaculture are represented by an increasingly competition for resources (in terms of land/water/feed) with other users, the decrease of the quality of water supplies due to the aquatic pollution and the rise of water temperature due to global warming, and the increase of disease outbreaks that imply a better fish health management (FAO 2000).

1.1.1 Disease outbreaks: one of the major issues for aquaculture

Disease outbreaks may have important impacts on aquaculture, threatening the survival rate of farmed animals, interfering with the progress and sustainability of intensive systems, and, consequently, causing significant production and economic losses (Ødegård et al. 2011).

Pathogens can enter aquaculture farms through water intake, feed, or infected broodstock; more important, diseases that normally are not problematic for wild hosts can become an issue in

aquaculture, with unforeseeable outcomes, due to the fact that farmed fish are often not coevolved with local infectious agents. Moreover, the high stocking density of aquaculture farms increases the stress of animals and raises the possibility of contact between host and pathogen, therefore making aquaculture highly susceptible to disease outbreaks (Lafferty et al. 2015).

Thus, it is necessary that farms invest in prevention, control and treatment of diseases. The good management of the farms, including biosecurity and hygienic standards, routine monitoring of water quality, appropriate stock density, control of the transfers/introduction of animals and selection of disease-free broodstock, is the first step to prevent the entering and the spread of pathogens. Antibiotics, chemotherapeutics, vaccines, probiotics or other immunostimulants can be useful, but their cost affects the profits. Moreover, these treatments are not effective for all the possible viral and bacterial infections affecting aquaculture species (Gjedrem 2015), they may lead to the development of resistant pathogen strains (Defoirdt et al. 2011) and drug residues might have a negative impact on consumer's health (Heuer et al. 2009) and on environment (Kumar et al. 2012).

In this context, selective breeding for disease resistance has been recognized as an interesting preventive measure, due to a long-term control of the disease (Das and Sahoo 2014) and a cumulative and permanent improvement of the trait over generation at the population level (Doan et al. 2016). In addition, the improvement in disease resistance inherited by new generations may provide protection at the larval stage, when typically the immune system is not completely developed (Zapata et al. 1996).

Genetic improvement programs through selective breeding have been initiated in several aquaculture species, initially with the aim to improve growth performances and morphology traits. Later, the focus has been directed to other traits, like processing yield, feed efficiency, product quality, reproduction and, indeed, disease resistance (Janssen et al. 2017).

The response to selection is variable and it is strictly related to the magnitude of additive genetic variation of the trait, as well as phenotypic variation and selection intensity (Dunham 2011).

Disease resistance show moderate to high values of heritability, that are highly species and pathogen-specific, but that support selective breeding for disease resistance as a feasible and effective approach (see Chapter 2).

Besides the direct approaches, the enhancing of disease resistance can be achieved through indirect approaches, as well. In order to obtain successful results, the chosen marker trait (physiological or immunological parameters) must show both a significant additive genetic variation and a positive genetic correlation with disease resistance. Apparently, low-stress responsiveness animals show better performances (including disease resistance; MacKenzie et al. 2009; Castanheira et al. 2015); this makes post-stress cortisol level an interesting marker trait, due to its reliability in terms of stress assessment (Pottinger 2008) and a moderate genetic variability (heritability estimates range between 0.08 and 0.33; Volckaert et al. 2012; Vandeputte et al. 2016).

The availability of more advanced genomic tools and techniques, and the development of genome-wide sequencing and genotyping technologies, make possible the generation of high-density marker data. This is certainly peculiar for the selective breeding field, which is progressively improving, from the simpler mass/individual/family selection strategies, to marker-assisted selection (MAS) approaches, and, finally, to genomic selection (GS), recognized a more enhanced and appropriate strategy for complex and polygenic traits like disease and stress resistance (Meuwissen et al. 2016).

Chapter 2 provides an exhaustive background about disease and stress resistance in fish, explaining the mechanism of response to diseases and to stress, the connection between the immune and the neuroendocrine systems, how to estimate these two traits and how these traits can be improved through selective breeding, focusing on the major farmed fish species in Europe.

1.2 Biology and production of the European sea bass (*Dicentrarchus labrax* L.)

The European sea bass (*Dicentrarchus labrax* L.) is a marine teleost fish species, well known for its commercial importance in Europe, both as aquaculture and fishery resource.
It is a eurythermal (5 to 32 °C) and euryhaline fish (3‰ to over 30‰; Volckaert et al. 2008); its distribution reaches the North Atlantic, the Mediterranean and the Black Sea.
It is a demersal fish, more common in shallow waters, and it inhabits the littoral zone (estuaries, lagoons).
During summer it reaches coastal waters and river mouths, while during winter migrates offshore reaching deep waters. Larvae are planktonic, whereas juveniles move inshore, forming schools in brackish or estuarine areas and sometimes rivers.
It has carnivorous feeding habits (zooplankton and invertebrates during juvenile stages, while adults prefer worms, shrimp, squids, mollusks and small fish).
The European sea bass is a gonochoristic species. In the wild, sexual maturity of males and females is reached between 2 to 4 years in the Mediterranean and between 4 to 7 years (for males) or 5 to 8 years (for females) in the Atlantic. Under farming conditions, maturity is reached earlier. Annual spawning occurs during winter/early spring; temperature and photoperiod are the two factors influencing reproduction. Fertilization is external and the eggs are pelagic (Pickett and Pawson 1994).
Population genetics studies of European sea bass recognized clear genetic differentiations between the Atlantic and the Mediterranean populations, and substructures even within the Mediterranean sea (Naciri et al. 1999; Bahri-Sfar et al. 2000; Lemaire et al. 2005; Quéré et al. 2012; Tine et al. 2014).
Progresses in extensive and intensive farming of the European sea bass have been done throughout years; at the beginning it was farmed in coastal lagoons and tidal reservoirs, then in the late 1960s France and Italy succeeded to develop mass production techniques for juvenile sea bass. By 1970s, new techniques and technologies were achieved in most European countries, leading to the production of hundreds of thousands of larvae (Pickett and Pawson 1994). The aquaculture production of sea bass in Europe has increased in the last 10 years, reaching 157,698 annual tons in 2016 (FEAP), and to date it is one of the most important marine species widely cultured in the Mediterranean areas, representing 49% of the total production.

1.2.1 One of the major challenges in farmed European sea bass: viral nervous necrosis

Viral nervous necrosis (VNN), known also as viral encephalopathy and retinopathy (VER), is becoming one of the most important issues for aquaculture, affecting more than 50 different marine and freshwater fish species throughout the world.
Nodavirus (or nervous necrosis virus, NNV) is a non-enveloped icosahedral virus of the Nodaviridae family, genus *Betanodavirus*, whose genome is composed of two single-stranded linear RNA segments, one responsible for the production of the RNA-dependent RNA polymerase and of the subgenomic RNA3 (RNA1) and the second responsible of the encoding of the capsid protein precursor (RNA2; Yong et al. 2017). In fish, four betanodavirus genotypes have been detected, according to the RNA2 T4 variable region: striped jack nervous necrosis virus (SJNNV, the first isolated strain), tiger puffer nervous necrosis virus (TPNNV), barfin flounder nervous necrosis virus

(BFNNV) and red-spotted grouper nervous necrosis virus (RGNNV). Among these genotypes, RGNNV is the most common in the Mediterranean area and in warm-water in general (Munday et al. 2002; Bovo et al. 1999).
Evidence shows that the transmission of NNV might be both vertical (from broodstock to offspring; Breuil et al. 2002) and horizontal (from infected individuals, water or feed; Munday et al. 2002). In addition, virus persist for a long time in sub-clinically infected fish (Johansen et al. 2004).
The pathogenesis of NNV consists in the vacuolation and necrosis of the nervous cells of the brain, spinal cord and retina (Mori et al. 1992), causing in fish clear clinical signs of neurological damages (abnormal swimming patterns, lethargy, skin darkening, loss of appetite; Munday and Nakai 1997).
The ability of the virus to infect and induce clinical disease/mortality seems to be not influenced by water salinity (Pascoli et al. 2016), while water temperature affects both the mortality and the virus load in the brain, with differences depending on the virus (RGNNV: 25-30 °C; BFNNV: 15-20 °C; TPNNV: 20 °C; SJNNV: 20-25 °C; Toffan et al. 2016).
In European sea bass NNV show a very high infectivity, with mortality rates up to 80-100% at the larval stages (Munday et al. 2002) and moderate mortalities in advanced juveniles (11-20%; Le Breton et al. 1997). If fish survive, anyway, the infection might become chronic, causing an impairment of the general conditions, poor growth rates and the possibility to become asymptomatic carriers (Vendramin et al. 2014). Overall, the economic consequences of VNN outbreaks in sea bass farms are rather serious.
Due to the ineffective chemotherapeutical approaches (Doan et al. 2016) and the absence of NNV vaccines that can be cost-effective and simply administrable even to larval stages (Buonocore et al. 2017), the management strategies to prevent the entry and diffusion of NNV in aquaculture farms are represented mainly by the selection of VNN-free broodstock, the disinfection of the eggs (ozonation, UV) and the UV-treatment of the water; the effectiveness of these measures, anyway, is not always assured (Doan et al. 2016).
Recent results have directed the attention to selective breeding to improve the resistance of sea bass to nodavirus, due to the promising heritability estimated (0.26-0.27; Doan et al. 2017; Palaiokostas et al. 2018).

In **Chapter 3**, the genetic and genomic parameters for mortality and survival time after a VNN challenge test in a sea bass experimental population were investigated. In terms of alternative physiological or immunological parameters, we estimated the genetic and genomic parameters of cortisol concentration after stress exposure and antibody titer against NNV. Genetic and phenotypic correlations among traits were also assessed, to understand the real feasibility of indirect approaches to improve disease resistance.

In **Chapter 4**, five Bayesian models were compared with the aim to determine the potential of these genomic prediction approaches for the studied traits.
The accuracy of the models (in terms of correlation between observed value and prediction) when predicting the phenotype or the EBV (estimated breeding value) for the traits was estimated, and the classification performance of the model predictions when used to classify the phenotype for VNN mortality was evaluated.

1.2.2 Sex determination systems and the unusual case of the European sea bass

Sex determination in fish may exhibit different modalities (Heule et al. 2014; Mank et al. 2006); taking into account only gonochoristic species, possible sex determination systems include the chromosomal sex determination (male heterogamety, XX/XY, or female heterogamety, ZZ/ZW), polyfactorial (or polygenic) sex determination, and environmental sex determination (ESD; Bull 1983). The first two systems are referred as genotypic sex determination (GSD) and both of them are observed in teleost fish (Bachtrog et al. 2014).

Polygenic sex determination (PSD) was clarified by Bulmer and Bull (1982). They postulated a hypothetical continuous variable with phenotypic value (Y) and they supposed that the individual shows a determined phenotype (male or female) depending on the fact that Y lies below or beyond a fixed threshold (V). The phenotype has a polygenic determinism and can be influenced equally by environmental or genotypic effects, bringing its value below or beyond the threshold and determining the sex of the animal.

In teleost fish, polygenic sex determination has been suggested in various species (including the Atlantic silverside, *Menidia menidia*, Conover and Heins 1987; some populations of the zebrafish, *Danio rerio*, Liew et al. 2012; Wilson et al. 2014; the European sea bass, Vandeputte et al. 2007; Palaiokostas et al. 2015) but it could be more frequent than normally thought (Moore and Roberts 2013).

The European sea bass is a gonochoristic species; the histological differentiation of the gonads is completed at about the end of the first year of life and juveniles exhibit no external sexual dimorphism (Blázquez et al. 1995). Anyway, differences in growth rate between sexes are well-documented: "size-age" curves show that, at the same age, females are larger than males (Volckaert et al. 2008).

European sea bass sex determination model is consistent with the polygenic model previously described: the sex of an offspring is determined by an underlying sex tendency, affected both by polygenes and environmental effects, whose phenotypic value is greater or less than a threshold value (Vandeputte et al. 2007; Palaiokostas et al. 2015).

There are clear evidence that in sea bass early life rearing temperature has a strong influence on sex determination during all different ontogenetic stages up to metamorphosis, with low rearing temperature (13-17 °C) favouring the production of female offspring (Koumoundouros et al., 2002; Vandeputte and Piferrer 2018).

The genetic architecture of sex determination in wild European sea bass was investigated in **Chapter 5**. Due to the fact that polygenic sex determination is considered evolutionarily unstable (Bulmer and Bull 1982), the hypothesis in our study was that the sex determination system of the European sea bass could show variations between populations exposed or adapted to different environments.

The comprehension of the genetic architecture of sex tendency could be interesting even for sea bass aquaculture production. Indeed, while in natural populations sex-ratio is balanced, under aquaculture conditions it is strongly biased towards males (Vandeputte et al. 2012). This represents a problem for farmers for different reasons: 1. males grow between 25 and 35% less than females; 2. males become reproductively mature during their second year of life (one year earlier than females) and at harvest time they are in spermiating condition, thus investing a significant amount of energy into developing gonads; 3. 30% males mature precociously (by the end of the first year) and for this reason grow slower during the second year and weigh up to 18% less than non-precocious males by the time of marketing (Felip et al. 2006); moreover, testicular maturation leads to reduced flesh quality and the

spontaneous release of the sperm affects external appearance, reducing the commercial value of the product.

References

Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman TL, et al. (2014). Sex determination: why so many ways of doing it? *PLoS Biol* **12**: e1001899.

Bahri-Sfar L, Lemaire C, Ben Hassine OK, Bonhomme F (2000). Fragmentation of sea bass populations in the western and eastern Mediterranean as revealed by microsatellite polymorphism. *Proc R Soc Lond B Bio* **267**: 929-935.

Blázquez M, Piferrer F, Zanuy S, Carrillo M, Donaldson EM (1995). Development of sex control techniques for European sea bass (*Dicentrarchus labrax* L.) aquaculture: effects of dietary 17 α-methyltestosterone prior to sex differentiation. *Aquaculture* **135**: 329-342.

Bovo G, Nishizawa T, Maltese C, Borghesan F, Mutinelli F, Montesi F, De Mas S (1999). Viral encephalopathy and retinopathy of farmed marine fish species in Italy. *Virus Res* **63**: 143-146.

Bull JJ (1983). Evolution of sex determining mechanisms. Benjamin/Cummings, Menlo Park, CA.

Bulmer MG, Bull JJ (1982). Models of polygenic sex determination and sex ratio control. *Evolution* **36**: 13-26.

Buonocore F, Nuñez-Ortiz N, Picchietti S, Randelli E, Stocchi V, Guerra L, Toffan A, Pascoli F, Fausto AM, Mazzini M, Scapigliati G (2017). Vaccination and immune responses of European sea bass (*Dicentrarchus labrax* L.) against betanodavirus. https://doi.org/10.1016/j.fsi.2017.11.039

Castanheira, M. F., L. Conceição, S. Millot, S. Rey, M.-L. Bégout, B. Damsgård, T. Kristiansen, E. Höglund, Ø. Øverli, and C. I. M. Martins (2015). Coping styles in farmed fish: consequences for aquaculture. *Rev Aquacult* **9**: 23-41

Conover DO, Heins SW (1987). Adaptive variation in environmental and genetic sex determination in a fish. *Nature* **326**: 496-498.

Das S, Sahoo P (2014). Markers for selection of disease resistance in fish: a review. *Aquacult Int* **22**: 1793-1812.

Defoirdt T, Sorgeloos P, Bossier P (2011). Alternatives to antibiotics for the control of bacterial disease in aquaculture. *Curr Opin Microbiol* **14**: 251-258.

Doan Q K, Vandeputte M, Chatain B, Morin T, Allal F (2016). Viral encephalopathy and retinopathy in aquaculture: a review. *J Fish Dis* **40**: 717-742.

Doan QK, Vandeputte M, Chatain B, Haffray P, Vergnet A, Breuil G, Allal F (2017). Genetic variation of resistance to Viral Nervous Necrosis and genetic correlations with production traits in wild populations of the European sea bass (*Dicentrarchus labrax*). *Aquaculture* **478**: 1-8.

Dunham RA (2011). Aquaculture and fisheries biotechnology: genetic approaches. CABI, Wallingford, Oxfordshire, UK; Cambridge MA, VIII, 495 pp.

FAO (2000). Future challenges in fisheries and aquaculture. Into the next millennium: fishery perspective. Working Paper Series prepared by Veravat Hongskul.

FAO yearbook (2018). Fishery and Aquaculture Statistics 2016, Rome, 104 pp.

FEAP. Annual Report (2017). Available from URL: http://www.feap.info.

Felip A, Zanuy S, Carrillo N (2006). Comparative analysis of growth performance and sperm motility between precocious and non-precocious males in the European sea bass (*Dicentrarchus labrax*, L.). *Aquaculture* **256**: 570-578.

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

Gjedrem T, Robinson N (2014). Advances by selective breeding for aquatic species: a review. *Agric* Sci **5**: 1152-1158.

Heuer OE, Kruse H, Grave K, Collignon P, Karunasagar I, Angulo FJ (2009). Human health consequences of use of antimicrobial agents in aquaculture. *Clin Infect Dis* **49**: 1248-1253.

Heule C, Salzburger W, Böhne A (2014). Genetics of sexual development: an evolutionary playground for fish. *Genetics* **196**: 579-591.

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

Johansen R, Grove S, Svendsen A, Modahl I, Dannevig B (2004). A sequential study of pathological findings in Atlantic halibut, *Hippoglossus hippoglossus* (L), throughout one year after an acute outbreak of viral encephalopathy and retinopathy. *J Fish Dis* **27**: 327-41.

Kosswig C (1964). Polygenic sex determination. *Experientia* **20**: 190-199.

Kumar R, Lee J, Cho J (2012). Fate, occurrence, and toxicity of veterinary antibiotics in environment. *J Korean Soc Appl Biol Chem* **55**: 701-709.

Lafferty KD, Drew Harvell C, Conrad JM, Friedman CS, Kent ML, Kuris AM, Powell EN, Rondeau D, Saksida SM (2015). Infectious diseases affect marine fisheries and aquaculture economics. *Ann Rev Mar Sci* **7**: 471-496.

Le Breton A, Grisez L, Sweetman J, Ollevier F (1997). Viral nervous necrosis (VNN) associated with mass mortalities in cage-reared sea-bass, *Dicentrarchus labrax* (L.). *J Fish Dis* **20:** 145-151.

Lemaire C, Versini JJ, Bonhomme F (2005). Maintenance of genetic differentiation across a transition zone in the sea: discordance between nuclear and cytoplasmic markers. *J Evolution Biol* **18**: 70-80.

Liew WC, Bartfai R, Lim Z, Sreenivasan R, Siegfried KR, Orieux N (2012). Polygenic sex determination system in zebrafish. *PLoS One* **7**: e34397.

MacKenzie S, Ribas L, Pilarczyk M, Capdevila DM, Kadri S, Huntingford FA (2009). Screening for coping style increases the power of gene expression studies. *PLoS One* **4**: e5314

Mank JE, Promislow DE, Avise JC (2006). Evolution of alternative sex-determining mechanisms in teleost fishes. *Biol J Linn Soc* **87**: 83-93.

Meuwissen THE, Hayes BJ, Goddard ME (2016). Genomic selection: a paradigm shift in animal breeding. *Animal Front* **6**: 6.

Moore EC, Roberts RB (2013). Polygenic sex determination. *Curr Biol* **23**: R510-R512.

Mori KI, Nakai T, Muroga K, Arimot M, Mushiake K, Furusawa I (1992). Properties of a new virus belonging to nodaviridae found in larval striped jack (*Pseudocaranx dentex*) with nervous necrosis. *Virology* **187**: 368-371.

Munday BL, Nakai T (1997). Nodaviruses as pathogens in larval and juvenile marine finfish. *World J Microbiol Biotechnol* **13**: 375.

Munday BL, Kwang J, Moody N (2002). Betanodavirus infections of teleost fish: a review. *J Fish Dis* **27**: 127-142.

Naciri M, Lemaire C, Borsa P, Bonhomme F (1999). Genetic study of the Atlantic/Mediterranean transition in sea bass (*Dicentrarchus labrax*). *J Hered* **90**: 591-596.

Ødegård J, Baranski M, Gjerde B, Gjedrem T (2011) Methodology for genetic evaluation of disease resistance in aquaculture species: challenges and future prospects. *Aquacult res* **42**: 103-104.

Palaiokostas C, Bekaert M, Taggart JB, Gharbi K, McAndrew BJ, Chatain B, Penman DJ, Vandeputte M (2015). A new SNP-based vision of the genetics of sex determination in European sea bass (*Dicentrachus labrax*). *Genet Sel Evol* **47**: 68.

Palaiokostas C, Cariou S, Bestin A, Bruantn JS, Haffray P, Morin T, Cabon J, Allal F, Vandeputte M, Houston RD (2018). Genome-wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (*Dicentrarchus labrax*) using RAD sequencing. *Genet Sel Evol* **50**: 30.

Pascoli F, Serra M, Toson M, Pretto T, Toffan A (2016). Betanodavirus ability to infect juvenile European sea bass, *Dicentrarchus labrax*, at different water salinity. *J Fish Dis* **39**: 1061-1068.

Pickett GD, Pawson MG (1994). Seabass - Biology, exploitation and conservation, Chapman & Hall, London.

Pottinger TG (2008).The stress response in fish-mechanisms, effects and measurement, pp. 32-44. In: Fish Welfare (Branson EJ, Ed.) Blackwell Publishing Ltd, Oxford.

Quéré N, Desmarais E, Tsigenopoulos CS, Belkhir K, Bonhomme F, Guinand B (2012). Gene flow at major transitional areas in sea bass (*Dicentrarchus labrax*) and the possible emergence of a hybrid swarm. *Ecol Evol* **2**: 3061-3078.

Tilami SK, Sampels S (2018). Nutritional value of fish: lipids, proteins, vitamins, and minerals. *Rev Fish Sci Aquac* **26**: 243-253.

Tine M, Kuhl H, Gagnaire P-A, Louro B, Desmarais E, Martins RST et al. (2014). The European sea bass genome and its variation provide insight into adaptation to euryhalinity and marine speciation. *Nat Commun* **5**: 5770.

Toffan A, Panzarin V, Toson M, Cecchettin K, Pascoli F (2016). Water temperature affects pathogenicity of different betanodavirus genotypes in experimentally challenged *Dicentrarchus labrax*. *Dis Aquat Org* **119**: 231-238.

Vandeputte M, Dupont-Nivet M, Chavanne H, Chatain B (2007). A polygenic hypothesis for sex determination in the European sea bass *Dicentrarchus labrax*. *Genetics* **176**: 1049-1057.

Vandeputte M, Quillet E, Chatain B (2012). Are sex ratios in wild European sea bass (*Dicentrarchus labrax*) populations biased? *Aquat Living Resour* **25**: 77-81.

Vandeputte M, Porte JD, Auperin B, Dupont-Nivet M, Vergnet A, Valotaire C, Claireaux G, Prunet P, Chatain B (2016). Quantitative genetic variation for post-stress cortisol and swimming performance in growth-selected and control populations of European sea bass (*Dicentrarchus labrax*). *Aquaculture* **455**: 1-7.

Vandeputte M, Piferrer F (2018). Genetic and environmental components of sex determination in the European sea bass (*Dicentrarchus labrax*). In: Wang HP, Piferrer F, Chen SL (eds) Sex Control in Aquaculture. Wiley-Blackwell, ISBN: 978-1-119-12726-0 (in print).

Vendramin N, Toffan A, Mancin M, Cappellozza E, Panzarin V, Bovo G, Cattoli G, Capua I, Terregino C (2014). Comparative pathogenicity study of ten different betanodavirus strains in experimentally infected European sea bass, *Dicentrarchus labrax* (L.). *J Fish Dis* **37**: 371-383.

Volckaert FAM, Batargias C, Canario A, Chatziplis D, Chistiakov D, Haley C, Libertini A, Tsigenopoulos C (2008). The European sea bass (*Dicentrarchus labrax* L.) and its genomic resources. In: Genome Mapping and Genomics in Fishes and Aquatic Animals, Vol. 2, cap. 5. Kocher T.D. and Kole C. editors, Springer-Verlag, Berlin, Heidelberg.

Volckaert FAM, Hellemans B, Batargias C, Louro B, Massault C, Van Houdt JKJ, Haley C, de Koning D-J, Canario AVM (2012). Heritability of cortisol response to confinement stress in European sea bass *Dicentrarchus labrax*. *Genet Sel Evol* **44**: 15.

Wilson CA, High SK, McCluskey BM, Amores A, Yan Y, Titus TA et al. (2014). Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. *Genetics* **198**: 1291-1308.

Winge O (1932). The nature of sex chromosomes. *Proceedings of the 6th International Congress of Genetics,* Ithaca, New York, **1**: 343-355.

Yong CY, Yeap SK, Omar AR, Tan WS (2017). Advances in the study of nodavirus. *PeerJ* **5**:e3841.

Zapata A, Torroba M, Varas A, Jimenez A (1996). Immunity in fish larvae. *Dev Biol Stand* **90**: 23-32.

2. Selection for disease and stress resistance in European farmed fish: a review

Sara Faggion[1], Daniela Bertotto[1] and Paolo Carnier[1]

[1] Department of Comparative Biomedicine and Food Science (BCA), University of Padova, Italy

Abstract

Disease outbreaks are serious threats that can interfere with the progress and sustainability of intensive aquaculture systems, and genetic selection for improved disease resistance has been recognised as an interesting solution, due to a long-term control of the disease and a cumulative and permanent improvement in resistance over generation at the population level.

The interest towards stress and welfare in farmed fish has dramatically increased over the last years, and low-stress responsive fish seem to have better productive and reproductive performances compared to high-stress responsive ones. Genetic selection with the aim to establish stress-resistant lines is a noteworthy opportunity and could be related to variability in disease resistance, due to a link between stress-immunodepression-disease susceptibility.

This study will review: (i) the process of disease and stress resistance in fish; (ii) the methodological approaches used to assess disease and stress resistance; (iii) the most important and noteworthy selective breeding approaches for both traits; (iv) the most relevant studies and results from the literature concerning European farmed species. Moreover, we tried to open to new perspectives in terms of different selection criteria, in order to suggest areas for future research.

Finally, we emphasized the importance of reproductive technologies in fish species as a key to develop successful selective breeding programs. Reproductive technologies are strictly linked with selection: in fish species, the availability of gametes is particularly limited by the reproductive biology, ecobiology and seasonality; in this context, it seems of great importance being able to manage and improve artificial fertilization.

Keywords disease resistance, stress resistance, selective breeding, aquaculture, reproductive technologies

2.1 Introduction

Aquaculture is progressively increasing its contribution to global food production and its importance as high-quality source of proteins, with a worldwide production of 80 million tons in 2016, a European production of about 2.9 million tons (FAO yearbook 2018), and the expectation of overcoming capture fisheries very soon.

High aquaculture productive levels imply more intensive productive systems that expose fish to even higher number of stressors. Stressors are generally defined as unfavourable pressures, intrinsic or estrinsic, real or perceived, acute or chronic, that perturb the homeostatis of an individual (Chrousos 2009) and that are normal components of the aquaculture environment: they refer to physical and environmental stressors (confinement, manipulation, transport, moving, rearing density, including also chemical stressors, related to water quality or to the presence of contaminants/pollutants), social stressors (dominance, hierarchy, aggressive interaction, competition) and perceived stressors (external stimuli perceived as threats; Tort 2011).

This situation of threatened homeostasis is frequently correlated to an impairment of the immune functions (Wendelaar Bonga 1997) and, therefore, to a higher probability of disease emergence.

Maintaining a high survival rate in the context of animal production is crucial: disease outbreaks are serious threats that can interfere with the progress and sustainability of intensive systems. Common interventions to prevent or control diseases (antibiotics, chemotherapeutics, vaccines, probiotics or other immunostimulants), are not always effective, particularly for viral infections, but also for some bacterial infections. In addition, medical treatments are usually provided as pelleted feed to all fish in a tank, with the disadvantage of treating also non-infected animals (Gjedrem 2015). Reducing the use of drugs and antibiotics is a priority in order to avoid the development of resistant pathogen strains (Defoirdt et al. 2011) and to reduce the negative impact of drug residues on consumer's health (Heuer et al. 2009) and on environment (Kumar et al. 2012). An additional consideration is merely economic and it is linked to the high cost of antibiotics, vaccines and pharmaceutical products. Hence, the search for alternatives to pharmacological treatments and vaccinations may favourably have an effect on profits of aquaculture farming.

In such view, selective breeding for improved disease resistance has been recognised as an interesting solution, as host resistance to different diseases and pathogens plays a key role in hindering the spread of pathogens and lowering the infection pressure.

Studies provided evidence that resistance to pathogens is, to some extent, under genetic control: the average genetic gain in resistance has been assessed to be 12.5% per generation (Gjedrem and Robinson 2014). This testifies the potential effectiveness of selective breeding as a preventive strategy leading to a long-term control of the disease (Das and Sahoo 2014) and cumulative and permanent improvements in resistance at the population level over generations (Doan et al. 2016). Establishment of resistant populations may affect the general welfare of the fish with benefits for yield, costs of fish production (Doan et al. 2016) and economic returns (Stear et al. 2012).

In a recent survey on programs managed by European aquaculture breeding companies, Chavanne et al. (2016) highlighted that disease resistance is currently one of the three most intensely selected traits.

Selecting for low-stress responsiveness may play a role in the view of enhancing resistance to diseases. Although the effectiveness and feasibility of selective breeding programs for stress tolerance has been slightly investigated and reliable estimates of the genetic relationship between stress tolerance and disease resistance are not available, the increase in levels of physiological stress

indicators following exposure to stressors has been reported to be heritable and affected by a non trivial amount of genetic variation (Fevolden et al. 1994 and 1999; Pottinger and Carrick 1999). Low stress-responsive fish proved to be able to successfully cope with increasing stressors induced by more intensive aquaculture systems, showing better performances in terms of reproductive and productive parameters (Cnaani 2006; Trenzado et al. 2006).
The selective breeding field is progressively improving, from simple mass/individual/family selection strategies, to marker-assisted selection (MAS) approaches, and, finally, to genomic selection (GS), recognized as a more efficient selective strategy for complex traits like disease and stress resistance (Meuwissen et al. 2016). The implementation of highly efficient approaches becomes possible due to the development of high-density markers arrays for the most important farmed species and the availability of advanced genomic tools and techniques.
This review intends to create a background about disease and stress resistance in fish, explaining how these traits can be enhanced through selective breeding and focusing on relevant studies and results from the literature.
New perspectives in terms of different physiological or immunological selection criteria for disease and stress resistance are discussed and the role of reproductive technologies is emphasized in the view of developing efficient selective breeding programs.
We focused on the major farmed fish species in Europe, as highlighted by the FEAP Annual Report (2017): Atlantic salmon (*Salmo salar*), rainbow trout (*Oncorhynchus mykiss*), gilthead sea bream (*Sparus aurata*), European sea bass (*Dicentrarchus labrax*) and common carp (*Cyprinus carpio*) which account for 98.07% of the total European production in 2016. Although the production of Atlantic cod (*Gadus morhua*) has sharply decreased to a null production in 2016, for many years cod has been included in the group of the major European farmed fish species. For this reason, literature results from genetic analyses of disease resistance in cod are reported in this review.

2.2 Background: disease and stress response

2.2.1 The mechanism of response to infection and disease

A disease is the result of complex interactions among three main leading actors: host, pathogen and environment that interact together leading or not to infection (Hedrick 1998). When an infection occurs, three main processes explain how the host copes with it: resistance, tolerance and resilience.
Disease resistance can be defined as the ability to resist and control the infection, since all the animals are susceptible to the initial infection, but differ in their ability to hinder the entry of pathogens, their replication, release and survival. It can be measured as the level of viremia or pathogen load within an animal (count of viral or bacterial cells) for a given level of exposure.
Disease tolerance and resilience reflect the connection between productive performance and pathogen load or level of exposure, and become important in a context of animal production.
Tolerance can be seen as the relationship between pathogen load and productive performance: when the pathogen load increases, the productive performance of animals with higher levels of tolerance deteriorates less than the one of less tolerant animals.
Resilience (or robustness) is the ability to maintain productivity when coping with different environmental challenges. It can be evaluated as the productive performance exhibited at a given level of exposure and the capacity to limit chronic forms of the disease (Dekkers 2016).
Dekkers (2016) underlined that developing lines of highly resilient animals might be an efficient strategy to control and prevent diseases. Enhanced resilience can be achieved by improving resistance

or tolerance, or both. Increasing resistance seems to be preferable as high tolerance limits the infection severity, but its effect on the pathogen is neutral (Roy and Kirchner 2000). Moreover, tolerance is difficult to measure because the pathogen load, at a given level of exposure, depends upon resistance. Conversely, resistance can be measured as viral or bacterial cells per host or per unit of host tissue (Guy et al. 2012).

This review is focused on disease resistance, primarily because resistance is pathogen-specific and, in terms of genetic selection, it should be possible to identify genetic markers associated with resistance through the host genome for a specific disease (Periasamy et al. 2014). Furthermore, resistance exhibits a significant additive genetic variation, which is a required feature to successfully improve a trait through genetic selection (Doeschl-Wilson et al. 2012).

2.2.2 The response to stress

Stress is recognised as a situation of threatened homeostasis and organisms normally carry out a complex and various set of physiological and behavioural adaptive responses in order to recover (Chrousos 1998).

Stress response consists of a primary, or neuroendocrine response, that involves the release of catecholamines (from chromaffin tissue) and corticosteroids (through the stimulation of hypothalamic-pituitary-interrenal, HPI, axis); a secondary response, or physiological response, that affects metabolism, respiration, acid-base status, hydromineral balance, immune and cellular functions; a tertiary, or whole-animal response, is related to a chronic stress exposure and implies changes in the overall performance in terms of growth, reproductive ability, behaviour and disease susceptibility (Ellis et al. 2012).

Stress quantification can be achieved by measurements of the stress response itself, in other words monitoring the primary, neuroendocrine response, or by measurements of the secondary/tertiary response (physiological changes or performance variations). Stress response quantification commonly involves the assessment of levels of relevant indicators (steroid hormones, physiological parameters). The analysis of these indicators helps to discriminate between animals with a lower sensitivity to stress and a greater ability to cope with stressful events, and higher stress-sensitive animals.

2.2.3 Stress: why it could be interesting for disease susceptibility?

The development of even more intensive aquaculture productive systems means that fish can be exposed to an increasing number of acute and chronic stressors. Stressful events may negatively affect the immune system, causing high susceptibility to pathogens, facilitating the onset of disease outbreaks and increasing the percentage of mortality (Magnadóttir 2010).

The immune system is commonly divided into innate (non-specific) and adaptive (specific). The innate immune system is the first line of defense and it is constituted by physical factors (scales, skin, gills and mucosal barriers in general), humoral factors (cell associated receptors and soluble molecules of plasma and other body fluids, such as the complement system, the natural antibodies, NAbs, cytokines/chemokines) and cells (phagocytic cells, non-specific cytotoxic cells, epithelial and dendritic cells). The activation of the innate immune system is fundamental to trigger the adaptive immune system, whose main components are the immunoglobulins (antibodies), that in teleost fish are referred only to IgM class, the major histocompatibility complex (MHC, class I and II) and the T-cell receptors (Magnadóttir 2010).

Actually, the connection between the immune and the neuroendocrine systems in fish is proved by the fact that several immune and endocrine messengers/receptors are comprised in the same family of molecules (Tort 2011) and by the organization of the head kidney, which is a "shared" organ between immune (due to the presence of cytokine-producing lymphoid cells) and endocrine systems (because of the presence of endocrine cells secreting cortisol, catecholamines, and thyroid hormones; Geven and Klaren 2017).

Moreover, it has been reported that cortisol, the main corticosteroid that is released after stressful events, induces changes in lymphatic structures, with apoptosis of lymphocytes or variations of their mitogenesis, amount, distribution and chemotaxis (Tort 2011). Cortisol has depressive effects on phagocytosis and it is associated with decrease or suppression of some pathways of the immune response (complement activity, Mauri et al. 2011; lysozyme activity, agglutination activity, antibody producing cells activity, and antibody titers, Sunyer and Tort 1995). Moreover, cortisol inhibits the respiratory burst activity (Wang and Belosevic 1995), the pro-oxidative activity of leukocytes from head kidney (Esteban et al. 2004), the lipopolysaccharide-induced expression of acute phase protein serum amyloid S and pro-inflammatory cytokines, which subsequently affect HPI axis (Fast et al. 2008).

Recently the interest towards the different responses to stressful events in fish species and the consequences on performances has increased (MacKenzie et al. 2009; Castanheira et al. 2015). This has lead to a growth of studies on what have been named "coping styles": animals defined proactive (or "active coping") tend to have specific physiological and behavioural features suggesting a lower sensitivity to environmental stressors and a greater ability to cope with stressful events compared to reactive fish (called "passive coping"). As a consequence, proactive and reactive individuals show different performances in terms of production and fish welfare (Castanheira et al. 2015) with proactive animals characterised by higher disease resistance (MacKenzie et al. 2009), better reproductive and productive performances (growth, feed conversion rate; Øverli et al. 2006), better resilience and the ability of maintaining productivity when environmentally challenged.

Producing a permanent improvement in terms of low stress-sensitivity may become possible through selective breeding. Establishing low stress-responsive lines of fish could be a chance to improve resistance to specific diseases. The first step should be the individual phenotyping of the stress response through precise approaches and relevant markers. Secondly, reliable procedures to predict the individual genetic merit for stress resistance need to be developed. In order to make efficient use of stress resistance to enhance the selective response in disease resistance, the genetic relationships between the traits needs to be assessed.

2.3 Phenotyping of disease and stress resistance

2.3.1 Experimental challenge tests to assess disease resistance

A mandatory step that is preliminary to the selection of animals based on their predicted genetic merit for resistance to a given pathogen (both when traditional and genomic selective breeding methods are used) is phenotyping. In general, the ability to resist to an infection is rarely manifested as an all-or-none phenomenon and the level of resistance can be determined as the survival of individuals after a disease outbreak, reflecting the cumulative effects of all host-pathogen interactions occurring during the disease.

Collecting data on animals that are exposed to the pathogen in standard conditions provides consistent results and can be implemented through experimental challenge tests for specific pathogens. This

allows to detect genetic differences in resistance across individuals, but it remains to assess whether such differences can be translated into phenotypic variation in resistance to the pathogen in the case of field outbreaks. Actually, experimental challenge tests are particular experimental conditions: fish are subjected to a controlled infection, in a standardized environment, normally using one pathogen at a time. These features minimise the variation due to uncontrolled sources, maximise the reproducibility of the procedure and enhance interpretation of the data in terms of individual resistance to the pathogen. Conversely, data from field outbreaks cannot be considered precise and reliable, due to different reasons: the complexity of the variables involved (biotic and abiotic) that can interfere with the host response and could bias the observed survival, the variation in the level of exposure to the pathogen (making difficult the distinction between unexposed and real resistant individuals) and the possible co-infection by other pathogens.

Exposure in challenge tests is accomplished by immersion of animals in water added with a virus solution or through intramuscular or intraperitoneal injection of a determined dose of pathogen: these different methods may lead to different results. In the case of immersion, the pathogen must cross the physical and mechanical barriers of the host (epithelial and mucosal), which are the first protective mechanisms during infection, and, later, must succeed in penetrating the host cell barriers in adequate number, attacking target cells and replicating. This type of experimental infection is more representative of natural infections, but it results in higher heterogeneity in the pathogen load and contribute additional genetic variation to the measured trait. Under injection, animals are directly infected with injections of equal doses of pathogens, so the level of exposure is the same for all individuals; this assures great effectiveness, especially when large sized individuals are tested (Robinson et al. 2017), but the phenomenon is only partially considered (physical and mechanical barriers are bypassed), leading to a reduction of genetic vatiability and affecting the estimation of genetic parameters.

Still, challenge tests remain the most used and widespread approach to investigate the additive genetic variation in disease resistance traits that provide exploitable estimates of genetic merit of breeding candidates. In addition, the results could be helpful in a context of commercial production system, since challenge trials can be merged to breeding programmes aiming at the production of animals with higher resistance to viral, bacterial and parasitic diseases through the application of genomic resources.

The statistical analysis of challenge data is normally implemented through different studies, cross-sectional (threshold) or longitudinal, providing interesting information about different elements of the host response to infection.

Cross-sectional studies consider the trait as binary (alive/dead or 0/1); they are, in general, simple and cost-effective (only one measurement is necessary, at the end of the challenge trial). In longitudinal studies, survival time is taken into account and the trait is defined as days to death, with the need of frequently measurements during the trial (Ødegård et al. 2011).

The statistical models used in the case of binary trait are linear or threshold (probit), while hazard frailty models (Cox, Weibull) are more appropriate in case of survival-time data. Fish that are still alive at the end of the trial have a censored observation (in other words, it is assumed that they would die after the fixed point corresponding to the end of the experiment; Ducrocq and Casella 1996).

2.3.2 Assessment of stress response

The peculiarities that make a marker useful for assessment of stress response are its reliability, specificity and diagnostic relevance, the ability to detect changes in a simple and precise way and the relative simplicity and cheapness of the procedures to estimate the parameter itself (Pottinger 2008).
Cortisol is the most important glucocorticoid hormone and it is considered a measurable component of the primary (neuroendocrine) stress response in fish (Wendelaar Bonga 1997). Cortisol increase is fast enough, albeit delayed when compared with cathecolamines, large and, therefore, clearly indicative of the severity and length of the reaction following stressful events. Other advantages are related to the simpleness, rapidity, cheapness and accuracy of the analyses with RIA (RadioImmunoAssay) or ELISA, and the fact that the sampling of some matrices currently used for the detection of cortisol levels do not implies the sacrifice of animals, such as blood plasma or serum, scales and mucus (Aerts et al. 2015; Bertotto et al. 2010; Guardiola et al. 2016).
It is worth noting that the degree of response may be biased by internal or external factors that affect cortisol secretion (salinity, pollutant exposure, developmental stage or nutritional status; Mommsen et al. 1999). Hence, in order to use cortisol levels as a stress response measure, physiological and environmental conditions must be standardized as much as possible. Cortisol levels should be assessed in a short time after acute stress exposure; in chronic stress assessment, cortisol levels may not be indicative (Hontela et al. 1992).
Catecholamines (adrenaline, noradrenaline) are key indicators of the primary response, able to provide precise measures of acute stress response of an organism and sensitive to a wide variety of stressors. The disadvantage is the rapidity of depletion that makes their detection window rather short (Pottinger et al. 2008).
The overall effect of glucocorticoid and cathecolamines induces changes in a wide number of parameters included in the secondary response, which in turns can be used as plausible parameters. Each of these parameters seems to have some disadvantages that make them not comparable with cortisol in terms of precision and reliability.
Plasma glucose concentration, for example, can provide an indication of the metabolic turnover of the glucose activated by the primary response (catecholamines first, cortisol later) but it is non-specific, since a number of other factors may affect glycaemia levels (species, developmental stage, metabolic state, diet, time since last feeding, season of the year).
Blood lactate has been studied as an acute or chronic stress indicator of anaerobic metabolism (Wells and Pankhurst 1999; Caruso et al. 2005), but this metabolite can be affected by general metabolic processes, with a difficulty in the interpretation of the results.
Heat shock proteins (HSPs) are responsive to different types of stressors (Palmisano et al. 2000), but their expression is affected by the duration and the magnitude of stress, as well as the acclimation of the organism, that could distort the assessment of stress response (Iwama et al. 2004).
Changes in overall plasma osmolality or ion balance are indicative of acute stress, whereas, in the context of chronic stress, the results become of difficult interpretation because strictly context-specific and affected by a wide number of internal and external factors (McDonald and Milligan 1997).
Other proposed indicators are more doubtful: experimental studies on plasma proteins level and haematological values (haematocrit and haemoglobin), highlighted the limitation and the inconsistency of these parameters in monitoring stress (Caruso et al. 2005; Wells and Pankhurst 1999), while intracellular enzymes (alanine transaminase, aspartate transaminase, lactate

dehydrogenase, or creatine kinase) can be considered indicative of stress only in the case of physical injury or tissue damage (Wagner and Congleton 2004).

Lysozymes are part of the non-specific immune system and have been initially proposed as indicators of acute stress in rainbow trout (Demers and Bayne 1997) and salmon (Fevolden et al. 1994). Later, in the study published by Fevolden et al. (in 1999), the correlation between post-stress lysozyme levels and cortisol increases was not confirmed, and no differences in lysozyme levels were detected, neither between non-stressed and stressed fish, nor between fish exposed to three subsequent stressful events.

In conclusion, the advantages and disadvantages of stress markers show that cortisol can be considered the most workable and informative parameter compared to other stress indicators, because it meets all the requirements needed. Anyway, in order to give a complete estimate of stress response it would be preferable, when possible, to use different markers and different matrices processed in a proper way and to compare or combine multiple measurements.

2.4 Selection for disease and stress resistance: state of the art

2.4.1 Genetic parameters for disease resistance

Literature heritability estimates for traits related to resistance to bacterial, viral and parasitic diseases in aquaculture species are reported in Table 1, 2 and 3, respectively.

Comparisons of such estimates are not straightforward, mainly because they have been obtained using a variety of statistical methodologies and models, the trait definition is often different and expressed on different scales. For some studies, comparisons are even unfeasible due to the use of unsuitable models (e.g., linear models for binary traits or for survival time in presence of censoring). Details are provided in the tables to facilitate the identification of possible analogies across estimates. Significant additive genetic variation has been reported for resistance against diseases in salmonids, cod, carp, sea bream and sea bass under challenge conditions.

Despite the productive importance of European sea bass and gilthead sea bream, particularly in the Mediterranean area, studies about selection for disease and stress resistance in these species are still few and mainly preliminary. Actually, most of the published studies reporting data about genetic variation or genetic selection approaches for disease resistance or stress responsiveness focus on salmonids (Atlantic salmon and rainbow trout), due to their importance in the context of European aquaculture and their longer selection process. For these species, many authors identified the presence of a genetic basis for resistance to various pathogens (bacteria and viruses), as well as parasites, with moderate to high heritability values. Recently, strains of salmonids selected for increased resistance has been established.

Data from commercial selection experiments seem to support the effectiveness of genetic improvement as a potential technique that can be used in the control of diseases through the increase of fish resistance: salmons selected for low resistance to infectious pancreatic necrosis virus (IPNV) had a mortality that was more than two times as high as the one of the group selected for high resistance to the disease (Storset et al. 2007), whereas rainbow trouts selected for high resistance to viral hemorrhagic septicaemia virus (VHSV) showed a considerable lower mortality (0-10%) compared to the non-selected population (70-90%; Dorson et al. 1995).

Heritability for resistance to nervous necrosis virus (NNV) is interestingly high in Atlantic cod, both when measured as time to death ($h^2 = 0.81$) and as a binary trait ($h^2 = 0.43$-0.49 on the observed scale, $h^2 = 0.68$-0.91 on the underlying liability scale; Ødegård et al. 2010a; Bangera et al. 2011, 2013 and 2014), while in European sea bass, the heritability estimates recently obtained by Doan et al. (2017;

$h^2 = 0.14$ on the observed scale, $h^2 = 0.26 \pm 0.11$ on the underlying liability scale) and Palaiokostas et al. (2018; $h^2 = 0.27$ on the underlying liability scale, binary trait) were not as high as in cod. In sea bass, Palaiokostas et al. (2018) obtained also an estimated heritability of resistance to VNN equal to 0.43, in this case exploiting a genomic relationship matrix, but they suggest caution in interpreting such estimates as the linkage disequilibrium derived from selective breeding could be a reason for such an overrating of additive genetic variance (Table 2).

Only in the case of *Aeromonas hydrophila* resistance in common carp heritability resulted very low ($h^2 = 0.03$ on the observed scale, $h^2 = 0.04 \pm 0.03$ on the underlying liability scale; Ødegård et al. 2010b), probably due to specific features of the disease which, in nature, is normally a secondary infection associated with environmental changes or other viral/parasitic diseases that makes the assessment of resistance through one-disease experimental challenge test difficult (Table 1).

Genetic improvement with respect to one trait (in this case resistance to one pathogen) might indirectly affects other traits (resistance against other diseases or economically important traits) because of the presence of genetic relationships between traits (Ødegård et al. 2011). Traits like growth, feed conversion rate, fat content, productive yield in terms of fillet quality and quantity, are particularly important in a context of animal production. Attention to a possible trade-off between selection for economically important traits and for disease resistance is essential. Literature estimates about the genetic correlations between resistance to a specific pathogen and production traits (e.g., body weight, body length, feed conversion efficiency, fat content, fillet yields) seems to be, in general, not very large or significant, ranging from moderately negative to moderately positive (Table 4). In case of infectious hematopoietic necrosis virus (IHNV) in Atlantic salmon and *Photobacterium damselae* subs. *piscicida* in gilthead sea bream, the genetic correlations between resistance and body weight or length are positive and significant (0.52 and 0.61 ± 0.16, respectively; Overturf et al. 2010, Antonello et al. 2009), albeit these results should be interpreted with caution, because in both cases weight and length were measured at the end of the challenge trial, making impossible to figure out if more resistant fish were larger because had more time to growth or if larger fish were actually more resistant to the diseases.

The genetic correlations between resistances to two diseases range from moderately negative to moderately positive, but no one of such estimates resulted statistically significant (Table 5).

Table 1. Heritability estimates of resistance to bacterial diseases in aquaculture species. We reported the trait definition, the model used to estimate the heritability and the scale to whom values are referred

Species	Disease agent	Reference	Trait definition	Model	h^2 ± SE	Scale of heritability	Notes
Atlantic salmon (*S. salar*)	*Piscirickettsia salmonis*	Yáñez et al. 2013	Time to death	Linear	0.18 ± 0.03	Observed	Censoring
			Time to death	Cox	0.41	Logarithmic-time	Censoring
			Time to death	Weibull	0.25	Logarithmic-time	Censoring
			Binary	Linear	0.15 ± 0.03	Observed	
			Binary	Probit	0.24 ± 0.04	Underlying liability	
			Survival time transformed to binary record	Logit	0.11 ± 0.02	Underlying liability	
		Correa et al. 2015	Time to death	Linear	0.19 ± NA	Observed	No censoring declaration
			Binary	Logit	0.20 ± NA	Underlying liability	
	Aeromonas salmonicida	Gjøen et al. 1997	Binary	MINQUE [1]	0.34 ± 0.13	Observed	
			Binary	MINQUE [1]	0.53 ± NA	Underlying liability	
		Ødegård et al. 2007	Binary	Probit	0.43 ± 0.02	Underlying liability	
		Kjøglum et al. 2008	Binary	Probit	0.62 ± NA	Underlying liability	
		Gjerde et al. 2009	Binary	Linear	0.31 ± 0.03	Observed	
			Binary	Linear	0.47 ± 0.05	Underlying liability	
		Drangsholt et al. 2011	Binary	Linear	0.33 ± 0.04	Observed	
			Binary	Probit	0.51 ± 0.05	Underlying liability	
	Vibrio anguillarum	Gjøen et al. 1997	Binary	MINQUE [1]	0.38 ± 1.07	Observed	
			Binary	MINQUE[1]	0.69 ± NA	Underlying liability	
Rainbow trout (*O. mykiss*)	*Yersinia ruckeri*	Henryon et al. 2005	Time to death	Weibull	0.21 ± 0.05	Logarithmic-time	Censoring
			Binary	Probit	0.42 ± NA	Underlying liability	
	Flavobacterium psychrophilum	Henryon et al. 2005	Time to death	Weibull	0.07 ± 0.02	Logarithmic-time	Censoring
			Binary	Probit	0.46 ± NA	Underlying liability	
		Silverstein et al. 2009	Time to death	Cox (animal)	0.35 ± 0.09	Logarithmic-time	Censoring
			Time to death	Cox (sire-dam)	0.43 ± 0.03	Logarithmic-time	Censoring

		Leeds et al. 2010	Time to death	Weibull	0.22 ± 0.03	Logarithmic-time	Censoring
			Binary	Linear	0.23 ± 0.03	Observed	
			Binary	Linear	0.37 ± NA	Underlying liability	
		Wiens et al. 2013	Binary	Linear	0.23 ± 0.09 - 0.25 ± 0.03	Observed	
		Evenhuis et al. 2015	Binary	Linear	0.18 ± 0.03	Observed	
		Vallejo et al. 2016	Time to death	Linear	0.31 ± NA	Underlying liability	Censoring
			Binary	Probit	0.48 ± NA	Underlying liability	
	Flavobacterium columnare	Evenhuis et al. 2015	Binary	Linear	0.17 ± 0.09	Observed	
Atlantic cod (*G. morhua*)	*Vibrio anguillarum*	Kettunen Præbel et al. 2007b	Time to death	Cox	0.08 ± NA - 0.10 ± NA	Logarithmic-time	No censoring declaration
			Time to death	Linear	0.13 ± 0.04 - 0.16 ± 0.04	Observed	No censoring declaration
			Time to death	Linear	0.14 ± 0.05 - 0.17 ± 0.06	Observed	Censoring
		Bangera et al. 2011	Binary	Probit	0.16 ± 0.04	Underlying liability	
		Bangera et al. 2013	Binary	Probit	0.19 ± 0.03	Underlying liability	
Common carp (*C. carpio*)	*Aeromonas hydrophila*	Ødegård et al. 2010b	Binary	Probit	0.03 ± NA	Observed scale	
			Binary	Probit	0.04 ± 0.03	Underlying liability	
Gilthead seabream (*S. aurata*)	*Photobacterium damselae* subs. *piscicida*	Antonello et al. 2009	Time to death	Linear	0.12 ± 0.04	Observed	No censoring declaration
			Binary	Logit	0.18 ± 0.08 - 0.45 ± 0.04	Observed	Six threshold days
		Palaiokostas et al. 2016	Time to death	Linear	0.22 ± NA	Observed	Censoring
			Time to death	Linear	0.28 ± NA	Observed	Censoring

[1] MINQUE (minimum norm quadratic unbiased estimation of h^2).

Table 2. Heritability estimates of resistance to viral diseases in aquaculture species. We reported the trait definition, the model used to estimate the heritability and the scale to whom values are referred

Species	Disease agent	Reference	Trait definition	Model	h^2 ± SE	Scale of heritability	Notes
Atlantic salmon (*S. salar*)	Salmon pancreas disease virus (SPDv)	Norris et al. 2008	Binary	Linear	0.21 ± 0.01	Underlying liability	
		Gonen et al. 2015	Binary	Linear	0.23 ± 0.05 - 0.34 ± 0.05	Observed	
			Binary	Linear	0.37 ± NA - 0.55 ± NA	Underlying liability	
			Binary	Probit	0. 33 ± 0.07 - 0.54 ± 0.07	Underlying liability	
			Binary	Logit	0. 46 ± 0.06	Underlying liability	
	Infectious pancreatic necrosis disease virus (IPNV)	Guy et al. 2006	Time to death		0.16[1]	Underlying liability	Censoring
			Binary		0.08 ± NA - 0.81 ± NA[1]	Observed	
		Wetten et al. 2007; Storset et al. 2007	Binary	Linear	0.31 ± NA	Observed	
		Kjoglum et al. 2008	Binary	Probit	0.55 ± NA	Underlying liability	
		Guy et al. 2009	Binary	Linear	0.38 ± 0.02	Observed	
			Binary	Linear	1.06 ± NA	Underlying liability	
		Drangsholt et al. 2011	Binary	Linear	0.20 ± 0.05	Observed	
			Binary	Probit	0.39 ± 0.05	Underlying liability	
	Infectious salmon anaemia virus (ISAV)	Gjøen et al. 1997	Binary	MINQUE[2]	0.13 ± 0.03	Observed	
			Binary	MINQUE[2]	0.19 ± NA	Underlying liability	
		Olesen et al. 2007	Binary	Probit	0.24 ± 0.03	Underlying liability	
		Ødegård et al. 2007	Binary	Probit	0.32 ± 0.02	Underlying liability	
		Kjøglum et al. 2008	Binary	Probit	0.37 ± NA	Underlying liability	
		Gjerde et al. 2009	Binary	Linear	0.26 ± 0.03	Observed	
					0.40 ± 0.04	Underlying liability	
		Drangsholt et al. 2011	Binary	Linear	0.22 ± 0.03	Observed	
			Binary	Probit	0.33 ± 0.05	Underlying liability	

Rainbow trout (*O. mykiss*)	Viral hemorrhagic septicaemia virus (VHSV)	Henryon et al. 2005	Time to death Binary	Weibull Probit	0.11 ± 0.10 0.57 ± NA	Logarithmic-time Underlying liability	Censoring
Atlantic cod (*G. morhua*)	Nervous necrosis virus (NNV)	Ødegård et al. 2010a	Binary Binary	Linear Probit	0.43 ± 0.07 0.75 ± 0.11	Observed Underlying liability	
		Bangera et al. 2011	Binary	Probit	0.68 ± 0.14	Underlying liability	
		Bangera et al. 2013	Binary Binary	CURE Probit	0.91 ± 0.06 0.76 ± 0.08	Underlying liability Underlying liability	
		Bangera et al. 2014	Time to death Binary	Linear Linear	0.81 ± NA 0.49 ± NA	Observed Observed	No censoring declaration
Common carp (*C. carpio*)	Koi herpesvirus (KHV)	Ødegård et al. 2010b	Binary	Probit	0.79 ± 0.15	Underlying liability	
European sea bass (*D. labrax*)	Nervous necrosis virus (NNV)	Doan et al. 2017	Binary Binary	Linear Linear	0.14 ± NA 0.26 ± 0.11	Observed Underlying liability	
		Palaiokostas et al. 2018	Bayesian binary Bayesian binary	Probit Probit	0.27 (HPD 95% 0.14-0.40) 0.43 (HPD 95% 0.29-0.57)	Underlying liability Underlying liability	Genetic Genomic

[1] h^2 estimated based on heterogeneity of chi-square using incidence within half-sib families and the variance in incidence among full-sib families.
[2] MINQUE (minimum norm quadratic unbiased estimation of h^2).

Table 3. Heritability estimates of resistance to parasites in Atlantic salmon. We reported the trait definition, the model used to estimate the heritability and the scale to whom values are referred

Species	Parasite	Reference	Trait definition	Model	$h^2 \pm$ SE	Scale of heritability	Notes
Atlantic salmon (*S. salar*)	*Lepeophtheirus salmonis* (sea lice)	Kolstad et al. 2005	Lice count	Linear	0.14 ± 0.02	Observed	Log transformation
		Glover et al. 2005	Lice count	Linear	0.07 ± 0.02	Observed	
		Gjerde et al. 2011	Lice count	Linear	0.33 ± 0.05	Observed	
			Lice density	Linear	0.26 ± 0.05		
		Tsai et al. 2016b	Lice count	Linear	0. 22 ± 0.08 - 0.33 ± 0.08	Observed	Log transformation
	Neoparamoeba perurans	Taylor et al. 2007	Gross gill score[1]	Linear	0.16 ± 0.07	Observed	
			Image gill score[2]	Linear	0.35 ± 0.10	Observed	
			Histopathology[3]	Linear	0.30 ± 0.09	Observed	Square root transformation
		Kube et al. 2012	Gross gill score[a]	Linear	0.09 ± 0.03 - 0.56 ± 0.07	Observed	measures of gill signs during 4 years and multiple reinfections
	Gyrodactylus salaris	Salte et al. 2010	Time to death	Linear	0.29 ± 0.07	Underlying liability	
			Binary	Probit	0.32 ± 0.10	Underlying liability	
			Binary	Probit	0.17 ± NA	Underlying liability	
	Caligus rogercresseyi	Correa et al. 2017	Lice count	Linear	0.12 ± NA	Observed	

[1] gross number/coverage of white patches over all gill hemibranchs.
[2] area of gill affected by amoebic gill disease expressed as the ratio of lesion pixels to total hemibranch filament area pixels.
[3] counting of the percentage of gill filaments (primary lamellae) displaying hyperplastic lesions.

Table 4. Genetic correlations between resistance to various diseases and productive traits in salmonids, gilthead sea bream and European sea bass

Species	Reference	Disease agent or parasite	Trait	Genetic correlation
Atlantic salmon (*S. salar*)	Gheyas et al. 2010	IPNV	Growth, fat content, fillet and gutted yields	No correlation
	Tsai et al. 2016b	*L. salmonis* (sea lice)	Length Weight	-0.04 to -0.1 -0.06 to -0.1
Rainbow trout (*O. mykiss*)	Overturf et al. 2010	*F. psychrophilum*	Early body weight End body weight Growth rate	0.35 -0.33 0.004
		IHNV	Early body weight End body weight Growth rate	0.28 0.52* - 0.11
	Evenhuis et al. 2015	*F. psychrophilum*	Body weight (9 months) Body weight (12 months)	-0.15 ± 0.08 -0.19 ± 0.24*
		F. columnare		-0.16 ± 0.24
			Body weight (9 months) Body weight (12 months)	-0.19 ± 0.24
	Henryon et al. 2002	VHSV	Body weight Body length Feed conversion efficiency	-0.33 to -0.14 -0.20 -0.22 to -0.01
Gilthead sea bream (*S. aurata*)	Antonello et al. 2009	*P. damselae* subs. *piscicida*	Body length	0.61 ± 0.16*
European sea bass (*Dicentrarchus labrax*)	Doan et al. 2017	NNV	Daily growth coefficient Body weight Fillet adiposity	-0.28 ± 0.20 -0.35 ± 0.14 -0.13 ± 0.19

* Significant values at $P < 0.05$

Table 5. Genetic correlations between resistances to different diseases in salmonids

Species	Reference	Disease agents	Genetic correlation
Atlantic salmon (*S. salar*)	Kjøglum et al. 2008 Ødegård et al. 2007	ISAV and *A. salmonicida*	0.07-0.15
	Kjøglum et al. 2008	ISAV and IPNV IPNV and *A. salmonicida*	-0.10 -0.11
Rainbow trout (*O. mykiss*)	Overturf et al. 2010	*F. psychrophilum* and IHNV	No correlation
	Evenhuis et al. 2015	*F. psychrophilum* and *F. columnare*	0.35 ± 0.25
	Biacchesi et al. 2007	VHSV and ISAV	No correlation
	Henryon et al. 2005	VHSV and *Y. ruckeri*	-0.11 (binary trait) -0.06 (time to death)
		VHSV and *F. psychrophilum*	0.15 (binary trait) 0.12 (time to death)
		F. psychrophilum and *Y. ruckeri*	-0.07 (binary trait) -0.23 (time to death)

2.4.2 Selection approaches for disease resistance: traditional and genomic approaches

Over the last years, selective breeding programs have been initiated in several aquaculture species, including some of the most important farmed fish in Europe, due to the genetic variation in resistance to diseases detected in several studies and low or null unfavorable genetic correlations between disease resistance and commercially important traits (Table 1, Table 2, Table 3 and Table 5).

Once resistance of fish has been assessed through challenge tests, phenotypic data (i.e., the observed mortality or survival) are processed with the aim to select breeding candidates. As the survivors of challenge trials cannot be used in breeding, because of the risk of vertical transmission of pathogens, conventional selective breeding methods for disease resistance rely on the prediction of the genetic merit of breeding candidates based on the phenotypic information of sib families. Hence, the selection of individuals is based on family-wise estimated breeding values (EBV) predicted from phenotypes of the challenged fish, which are full or half-sibs of the breeding candidates. Each family member has identical EBV, with no distinction across members (Saura et al. 2017) and with no possibility of ranking individuals within families. This family selection method is currently the most used approach, but it is not very efficient or precise and results in a quite low accuracy in breeding value estimation (Gjedrem and Baranski 2009), the exploitation of maximum 50% of the total genetic variation (Ødegård and Meuwissen 2014), a low selective response and the susceptibility to different components that further affect its efficiency (environmental variations overcoming genetic differences across families; family size, that affects the correlation between mean phenotypic value and mean genetic value; the necessity of an adequate number of families to keep inbreeding rates at low levels; Farias et al. 2017).

Genome sequencing has been raising fast due to the availability of new genomic tools, making progressively possible the construction of high density SNP arrays for aquaculture species, such as Atlantic salmon (132,033 SNPs, Houston et al. 2014; 159,099 SNPs, Yáñez et al. 2014; 96,396 SNPs, Tsai et al. 2016a), rainbow trout (40,900 SNPs, Palti et al. 2015), Atlantic cod (1,200 SNPs assigned to 23 linkage groups, Yu et al. 2014), turbot (6,647 SNPs assigned to 22 linkage groups, Wang et al. 2015), gilthead seabream (12,085 SNPs on 24 linkage groups, Palaiokostas et al. 2016), European sea bass (6,706 SNPs on 24 linkage groups, Palaiokostas et al. 2015) and common carp (3,470 SNPs and 773 microsatellites, Xu et al. 2014).

The availability of high-density marker maps allows the advancement of genetic approaches, providing alternatives to traditional methods in breeding for disease resistance: MAS (marker-assisted selection) and GS (genomic selection).

Theoretically speaking, MAS should be more efficient than conventional selection because realized on live breeding candidates using genetic markers linked with quantitative trait loci (QTL) affecting the trait phenotypic variation, but MAS power seems to be strongly limited for different reasons.

MAS requires prior knowledge of the linkage phase between the marker and QTL alleles for each candidate family, which can be investigated detecting within-family linkage disequilibrium. Due to the power limitations in the statistical tests used, only QTLs with large effects can be detected, leading to the exploitation of a restricted group of genetic markers (Meuwissen et al. 2016) that pick up only a small part of the genetic variation of a trait in a sort of "missing heritability paradox" (Manolio et al. 2009).

Disease resistance traits have showed to be complex, with a polygenic architecture. Identifying genetic markers associated to many QTLs of small effect represents a restraint for MAS and explains the poor performance exhibited by MAS when applied in selective breeding for disease resistance.

Only in two cases, infectious pancreatic necrosis (IPNV) in salmon (Houston et al. 2010; Moen et al. 2009) and whirling disease in rainbow trout (Baerwald et al. 2011), resistance traits have been associated with one QTL (hypothetically corresponding to one gene) explaining a large proportion of the phenotypic variance, ranging from 25 to 86%. In the majority of the studies on viral or bacterial diseases in salmonids, Atlantic cod and gilthead sea bream, many QTLs (major or suggestive) have been associated with disease resistance, but with the ability of explain a very small proportion of the trait phenotypic variance (Baranski et al. 2010; Barroso et al. 2008; Gonen et al. 2015; Massault et al. 2010; Palaiokostas et al. 2016; Vallejo et al. 2014a and 2014b). In addition, QTLs detection has considerable limitations related to low repeatability: QTLs are usually identified, but when different studies are compared, the QTLs position tends to change or is not confirmed due to the small effect of the QTLs themselves, the recombination between marker and QTL or to the possible false positive QTLs detection (Beavis 1997).

In contrast to MAS, GS is a method that estimates the effects of a large number of chromosomal regions, defined by a dense panel of genetic markers spread in the genome, without paying attention to the statistical significance of the effects, but aggregating such effects in the prediction of the genetic merit of a breeding candidate. It considers all markers as potentially linked to at least one QTL affecting the trait. Selective breeding based on genomic information seems to be particularly interesting for disease resistance because overcomes some problems faced by MAS (polygenic architecture of the traits) and by traditional methods (unfeasibility of phenotyping on breeding candidates). It relies on genotyping, using genome-wide dense marker panels (mainly SNP) covering the whole genome, and phenotyping, in challenge trials, a reference or "training" population of individuals. Genotypes and phenotypes of the reference population are used to estimate the marker effects on the traits, which allow, when aggregated, the prediction of genomic EBV (GEBV). After these predictions based on the marker, estimated effects are validated in a "test" population and the marker effects can be used in the routine prediction of the genetic merit of breeding candidates which need to be only genotyped. Hence, with GS, phenotyping needs to be performed only once for the individuals of the reference population, the assessment of the individual genetic merit for disease resistance becomes possible and the within-family component of the genetic variance can be exploited (Daetwyler et al. 2007). In this way, the selection of breeding candidates can be greatly anticipated in life.

As suggested by Palaiokostas et al. (2016) for pasteurellosis in sea bream, GEBV proved to be more accurate than "traditional" EBV because pedigree-based relationships are replaced with genomic relationships, thus delivering significant improvements in selection efficiency compared to traditional pedigree-based approaches and MAS. In addition, GS contributes to maintain the rate of inbreeding at acceptable levels, due to the within-family differentiation and a lower risk of co-selection of sibs.

Vallejo et al. (2017) tested the feasibility and accuracy of GS for improving BCWD resistance in rainbow trout, finding a predictive ability of GEBV twice as high as the predictive ability of EBV (0.63-0.70 vs. 0.34-0.36).

Finding a balance between genotyping costs and accuracy of selection seems to be one of the major challenges: currently, only a minority of the breeding companies incorporated GS in their breeding programs due to the high cost of genotyping, albeit the percentage has increased compared to ten years ago (Chavanne et al. 2016; Ødegård et al. 2011). Cost-efficient sequencing techniques and genotyping strategies are interesting possibilities to explore.

Other strategies, such as using DNA pooling (pooling DNA collected from good/bad animals in two pools, thus analysing only two samples per trait instead of a sample per tested animal; Sonesson et al. 2010), or within-family GEBV estimation (Lillehammer et al. 2013), are two cost-effective strategies, apparently with no risk of reduction of the accuracy of selection.
The size of the reference population, the degree of relationship among individuals belonging to the training population and between them and the breeding candidates, and the amount of genetic variation in resistance traits are important features that may influence the accuracy of the predicted GEBV (Pszczola et al. 2012).

2.4.3 Indicator traits for disease resistance and indirect selection approaches

Further approaches that can be used in selection for disease resistance are indirect and imply a deep analysis of physiological and immunological parameters that must be genetically correlated with disease resistance and affected by an appropriate genetic variation.
The correlation between plasma cortisol level, the most common stress responsiveness marker, and improved disease resistance, will be further discussed below.
Several immunological parameters have been proposed as useful markers to identify resistant animals; they are mainly components of the innate, or non-specific, immune system, such as serum lysozyme, myeloperoxidase, respiratory burst, complement activity, haemolytic and bactericidal activities, but the low additive genetic variation detected make them of difficult application as selection criteria and ineffective to predict breeding values (Magnadóttir 2006; Mohanty and Sahoo 2007).
A new and interesting possibility comes from other key components of non-specific immune system, the natural antibodies (NAbs), which seem to be promising as biomarkers for disease resistance in model species (Subramaniam et al. 2010) and terrestrial species (Wondmeneh et al. 2015; Mallard et al. 1998; Ploegaert et al. 2011). They have been detected in the serum of unimmunized individuals, they are polyreactive and they provide an immediate and broad protection against infection (Ochsenbein and Zinkernagel 2000). Higher NAbs titres, enhanced immune responsiveness and better disease resistance have shown to be three related concepts, as a number of studies on mice (Mi et al. 2000; Ochsenbein et al. 1999; Subramaniam et al. 2010), chickens (Parmentier et al. 2004; Wondmeneh et al. 2015), swine (Mallard et al. 1998) and cattle (Ploegaert et al. 2011) have reported. In addition, moderate heritability values for NAbs have been found in cattle and chickens (Berghof et al. 2015; Ploegaert et al. 2010), but the existence of a negative genetic correlation between production traits and immune response traits in sheeps, pigs and chickens represents a possible drawback (van der Most et al. 2010).
Studies in fish species are few, but promising, suggesting, similarly to terrestrial species, a negative association between high natural antibodies activity and risk of being infected by pathogens (Magnadóttir 2006 and 2009; Sinyakov et al. 2002). It would be therefore interesting understanding and determining the exact role of NAbs in the immune system of fish species, searching for confirmation of the hypothesis that higher NAbs are associated with increased disease resistance, analysing the heritability of immune response traits and genetic correlations between immune function and productive traits, and, in case, opening the possibility of using them as selective parameters in fish.

2.4.4 Selection for stress resistance

Selective breeding approaches favouring low stress-responders seem attractive for different reasons: stress-resistant animals can better adapt to intensive aquaculture productive systems, show better performances (Cnaani 2006) and, potentially, greater resistance to diseases due to lower impairment of the immune system.

In this context, post-stress cortisol level seems to be the most interesting trait related to stress response because of its reliability in terms of stress assessment (Chapter 2.3.2) and the existence of a genetic component of significant variability. Heritability estimates for cortisol response to confinement stress in sea bass are low to intermediate, ranging from 0.08 (Volckaert et al. 2012) to 0.33 (Vandeputte et al. 2016), and only slightly lower than those in salmonids (from 0.18 to 0.56; Fevolden et al. 1999 and 2002; Pottinger and Carrick 1999; Weber et al. 2008). In cod, the heritability estimates for cortisol level are inconsistent and with very large standard errors (from 0.11 ± 0.13 to 0.35 ± 0.18; Kettunen Præbel et al. 2007a). Studies on breeding for stress resistance are scarce, relatively not recent and mainly about salmonids.

Traditional breeding programs have been exploited in order to produce stress-resistant lines through the estimation of family and individual EBV using BLUP procedures (Fevolden et al. 2002). Several suggestive QTLs for stress response, explaining a very low percentage of the phenotypic variance, were identified in rainbow trout (Massault et al. 2010; Rexroad et al. 2012) and suggest that stress response can be attributed to the group of complex traits. This means that GS may be, likewise disease resistance, a more accurate and appropriate selection strategy than MAS (Meuwissen et al. 2016).

A negative genetic relationship between cortisol stress response and body weight has been detected in sea bass, but the estimates show large variability, ranging from high to low negative values, and large standard errors (from -0.60 ± 0.44 to -0.36 ± 0.16; Vandeputte et al. 2016; Volckaert et al. 2012). The expected response in cortisol responsiveness resulting from selection for growth is rather ambiguous. The increase and persistence of cortisol levels after stress exposure may be different in fish of variable size, with higher values in smaller fish (Fatira et al. 2014). Other authors (Vandeputte et al. 2016) suggested that the modification of cortisol responsiveness as a correlated response to the selection for growth is expected to be slow due to the weak genetic relationships between these traits. Genetic relationships between cortisol responsiveness and growth were not detected or were very low in rainbow trout, as well (Quillet et al. 2014).

For what concerns the improvement of disease resistance through the employment of cortisol response to stress as an indicator trait, results from experimental studies are doubtful and species- or disease-specific. A trout line selected for low cortisol levels was more susceptible to *V. anguillarum*, albeit less susceptible to *A. salmonicida* challenge (Fevolden et al. 1992). In salmons selected for high cortisol levels, mortality under farming conditions was increased in case of *V. anguillarum* and *A. salmonicida* challenge, whereas in case of *Renibacterium salmoninarum* infection (causing bacterial kidney disease, BKD), survival of high and low cortisol responders was similar (Fevolden et al. 1993). In carps exposed to chronic stress and subsequently challenged with *A. hydrophila*, the comparison between mortality rates of animals with high and low post-stress cortisol levels were not significantly different (Yin et al. 1995). In cod, genetic correlations between vibriosis resistance and stress resistance have not been detected (Kettunen Præbel et al. 2007a).

Due to the low reproducibility across stress experiments, searching for different selection criteria becomes necessary. Post-stress lysozyme activity has been proposed as a conceivable selective criterium, due to a significant additive genetic variation detected (h^2 from 0.11 to 0.46 in salmonids;

Fevolden et al. 1994 and 1999; Røed et al. 1993). Some uncertainties related to the consistency of this trait in measuring stress response have to be solved (Fevolden et al. 1999) and experimental selection tests using low post-stress lysozyme activity to improve disease resistance are controversial (Røed et al. 2002).
More advanced studies are required, in order to determine the consistency of selection criteria and to assess, for different species and diseases, if stress resistance is properly related to enhanced disease resistance.

2.5 Fish reproductive technology: the key for successful selective breeding programs

Expertise in reproductive technologies is recognized fundamental in a context of aquaculture production, but it becomes peculiar in case of planning of selective breeding programs.
In general, fish species are known to have high fecundity and external fertilization, so theoretically a large number of gametes can be collected. Nevertheless, the availability of gametes is particularly limited by the reproductive biology, ecobiology and seasonality of species.
Application of advanced reproductive technologies may solve problems arising from the reproductive biology and seasonality due to a control and, when necessary, an elongation of the reproductive seasons, a synchronization of gamete availability for both males and females and an enhancement of eggs production and spermiation. Other advantages are related to a simplified and more efficient management of broodstock and gametes, with the advantage of higher fertilization and hatching rates. Technologies such as artificial fertilization and gametes conservation through refrigeration or cryopreservation can be efficiently combined with genetic selection practices, thus simplifying the design of effective breeding programs and achieving high rates of genetic progress. In the case of selective designs based on genetic data from training populations, reproductive technologies can help in acquiring equal representativeness of all candidate families.
The practice of artificial fertilization seems to be regularly used in salmon, rainbow trout and turbot, whereas for sea bass and sea bream these technologies are not common and mass spawning is still the method of choice (Chavanne et al. 2016). Artificial reproduction of cultured fish is usually controlled by environmental manipulations (modifications of the photoperiod, water temperature or spawning substrate) or through the administration of exogenous hormones. The latter is considered more efficient in inducing gamete maturation, thus facilitating artificial fertilization for genetic selection programs (Mylonas et al. 2010).
Sperm and eggs conservation through refrigeration or cryopreservation is often combined with artificial fertilization techniques (Chavanne et al. 2016): the correct application of conservation protocols is peculiar in order to maintain gamete viability and ensure reproductive success if artificial fertilization is performed some time after gamete collection.
In the case of sperm, the best solution for short-term use or before cryopreservation is refrigeration at 0 and 4 °C. Refrigeration can be applied for different storage durations according to species, dilution media, dilution rate and storage temperature. The addition to sperm of antibiotic cocktails is preferable to prevent bacterial developments (Bobe and Labbé 2010).
For long-term conservation, cryopreservation is preferred. Sperm cryopreservation is commonly applied for a number of farm terrestrial species, but the application in aquaculture is more complex due to the requirement of protocols adapted to the peculiarities of fish sperm. In this case, the dilution medium should prevent any motility of sperm to avoid the consumption of energy or limit it to basic metabolism. Specific cryoprotectants are also added (Fauvel et al. 2012). The choice of cryoprotective

agents is undoubtedly important (Cabrita et al. 2010) and the application of cryopreservation methods to marine fish sperm has obtained different results in terms of motility, viability and hatching rates. The success is affected by the intrinsic quality of sperm, but other extrinsic factors, such as the presence of urine in the semen (Dreanno et al. 1998) or the time lapse from sperm collection/dilution to cryopreservation (Sansone et al. 2001) seem to play a role.
The cryopreservation of the sperm presents clear advantages for genetic selection programs. First, storing genomic resources of valuable strains allows the maintenance of desired fish characteristics. The sperm can be stored and used when necessary (for example, when eggs are available) and, in case of difficulty in obtaining semen or when low volumes of semen are stripped, the cryopreservation allows the use of the total volume of available sperm. Broodstock maintenance is simplified, allowing the induction of off-season spawning only in females, and transport of only gametes is easier than breeders transportation (Cabrita et al. 2010).
Lack of standardized protocols, due to non-homogeneous results obtained when the same procedures have been applied to different populations, individuals, or even to samples from the same male, represent important drawbacks. Furthermore, DNA damage after sperm cryopreservation has been documented in some species, such as rainbow trout, gilthead sea bream and sea bass (Cabrita et al. 2005; Labbé et al. 2001; Pérez-Cerezales et al. 2010; Zilli et al. 2003). Further research is undoubtedly needed and future studies should be focused on the standardization of freezing protocols to ensure an aware, safe and successful use of cryopreservation techniques in the aquaculture sector.
Likewise, it has been demonstrated that storing unfertilized eggs at low temperatures is possible: experiments focused mainly on cold-water species (salmonids), whereas results for warm-water species are not satisfactory due to different reproductive temperature ranges (Bobe and Labbè 2010). The intrinsic quality of the eggs has a significant impact on the overall success of the chilled storage procedure, but it is not the only factor. Breeding factors, husbandry practices, transport or environmental factors may play a role too (Bonnet et al. 2007).
Appropriate storage medium is necessary: studies confirmed that coelomic fluid is a possibility, but attempts have been made in order to produce an artificial medium that could standardise the procedure, albeit results are not conclusive yet. The holding procedure should limit bacterial contamination of the eggs. In some cases, coelomic fluid has antibacterial activity like in trout (Coffman et al. 2000), but, if not, artificial mediums should contain antibiotics (Goetz et al. 2000).
As regard cryopreservation, studies on eggs are scarce and not successful, even though eggs seem to survive for a short time after cooling at temperatures below 0 °C. Ovarian follicles showed to be weakly permeable to water and cryoptectants (Routray et al. 2002; Valdez et al. 2005; Zhang et al. 2005). Hence, an improvement of the fish oocyte membrane permeability is mandatory in order to achieve a successful cryopreservation (Zhang and Lubzens 2010).

References

Aerts J, Metz J, Ampe B, Decostere A, Flik G, De Saeger S (2015). Scales tell a story on the stress history of fish. *PLoS One* **10:** e0123411.

Antonello J, Massault C, Franch R, Haley C, Pellizzari C, Bovo G, Patarnello T, de Koning DJ, Bargelloni L (2009). Estimates of heritability and genetic correlation for body length and resistance to fish pasteurellosis in the gilthead sea bream (*Sparus aurata* L.). *Aquaculture* **298:** 29-35.

Baerwald M, Petersen J, Hedrick R, Schisler G, May B (2011). A major effect quantitative trait locus for whirling disease resistance identified in rainbow trout (*Oncorhynchus mykiss*). *Heredity* **106:** 920-926.

Bangera R, Ødegård J, Kettunen Præbel A, Nielsen HM (2011). Genetic correlations between growth rate and resistance to vibriosis and viral nervous necrosis in Atlantic cod (*Gadus morhua* L.). *Aquaculture* **317:** 67-73.

Bangera R, Ødegård J, Nielsen HM, Gjoen HM, Mortensen A (2013). Genetic analysis of vibriosis and viral nervous necrosis resistance in Atlantic cod (*Gadus morhua* L.) using a cure model. *J Anim Sci* **91:** 3574-3582.

Bangera R, Baranski M, Lien S (2014). A genome-wide association study for resistance to viral nervous necrosis in Atlantic cod using a 12K single nucleotide polymorphism array. *Proceedings, 10th World Congress on Genetics Applied to Livestock Production*, August 17th-22th, Vancouver, BC, Canada.

Baranski M, Kettunen Præbel A, Sommer AI, Hånes Kirste K, Wesmajervi M (2010). Major quantitative trait loci for viral nervous necrosis resistance in Atlantic cod. *Proceedings, 9th World Congress on Genetics Applied to Livestock Production,* August 1st-6th, Leipzig, Germany.

Barroso R, Wheeler P, LaPatra S, Drew R, Thorgaard G (2008). QTL for IHNV resistance and growth identified in a rainbow (*Oncorhynchus mykiss*) × Yellowstone cutthroat (*Oncorhynchus clarki bouvieri*) trout cross. *Aquaculture* **277:** 156-163.

Beavis WD (1997). QTL analysis: power, precision and accuracy, pp. 145-162. In: *Molecular dissection of complex traits*. (Paterson AH, Ed.) CRC press, Taylor and Francis Group, Boca Raton.

Berghof TVL, van der Klein SAS, Arts JAJ, Parmentier HK, van der Poel JJ, Bovenhuis H (2015). Genetic and non-genetic inheritance of natural antibodies binding keyhole limpet hemocyanin in a purebred layer chicken line. *PLoS One* **10**: e0131088.

Bertotto D, Poltronieri C, Negrato E, Majolini D, Radaelli G, Simontacchi C (2010). Alternative matrices for cortisol measurement in fish. *Aquac Res* **41:** 1261-1267.

Biacchesi S, Le Berre M, Le Guillou S, Benmansour A, Brémont M, Quillet E, Boudinot P (2007). Fish genotype significantly influences susceptibility of juvenile rainbow trout, *Oncorhynchus mykiss* (Walbaum), to waterborne infection with infectious salmon anaemia virus. *J Fish Dis* **30:** 631-636.

Bobe J, Labbé C (2010). Chilled storage of sperm and eggs, pp. 219-236. In: *Methods in reproductive aquaculture marine and freshwater species*. (Cabrita E, Robles V, Herráez P, Eds.) CRC press, Taylor and Francis Group, Boca Raton.

Bonnet E, Fostier A, Bobe J (2007). Characterization of rainbow trout egg quality: a case study using four different breeding protocols, with emphasis on the incidence of embryonic malformations. *Theriogenology* **67:** 786-794.

Cabrita E, Robles V, Rebordinos L, Sarasquete C, Herráez MP (2005). Evaluation of DNA damage in rainbow trout (*Oncorhynchus mykiss*) and gilthead sea bream (*Sparus aurata*) cryopreserved sperm. *Cryobiology* **50:** 144-153.

Cabrita E, Sarasquete C, Martínez-Páramo S, Robles V, Beirão J, Pérez-Cerezales S, Herráez MP (2010). Cryopreservation of fish sperm: applications and perspectives. *J Appl Ichthyol* **26:** 623-635.

Caruso G, Genovese L, Maricchiolo G, Modica A (2005). Haematological, biochemical and immunological parameters as stress indicators in *Dicentrarchus labrax* and *Sparus aurata* farmed in off-shore cages. *Aquacult Int* **13:** 67-73.

Castanheira MF, Conceição L, Millot S, Rey S, Bégout M-L, Damsgård B, Kristiansen T, Höglund E, Øverli Ø, Martins CIM (2015). Coping styles in farmed fish: consequences for aquaculture. *Rev Aquacult* **9:** 23-41.

Chavanne H, Janssen K, Hofherr J, Contini F, Haffray P, Komen H, Nielsen EE, Bargelloni L (2016). A comprehensive survey on selective breeding programs and seed market in the European aquaculture fish industry. *Aquacult Int* **24:** 1287-1307.

Chrousos G (1998). Stressors, stress, and neuroendocrine integration of the adaptive response. *The 1997 Hans Selye Memorial Lecture Annals of the New York Academy of Sciences* **851:** 311-335.

Chrousos G (2009). Stress and disorders of the stress system. *Nat Rev Endocrinol* **5**: 374-381.

Cnaani A (2006). Genetic perspective on stress response and disease resistance in aquaculture. *Isr J Aquac* **58:** 375-383.

Coffman MA, Pinter JH, Goetz FW (2000). Trout ovulatory proteins: site of synthesis, regulation, and possible biological function. *Biol Reprod* **62:** 928-938.

Correa K, Lhorente J, López M, Bassini L, Naswa S, Deeb N, Di Genova A, Maass A, Davidson WS, Yáñez JM (2015). Genome-wide association analysis reveals loci associated with resistance against *Piscirickettsia salmonis* in two Atlantic salmon (*Salmo salar* L.) chromosomes. *BMC Genomics* **16:** 854.

Correa K, Lhorente J, Bassini L, López M, Di Genova A, Maass A, Davidson WS, Yáñez JM (2017). Genome-wide association study for resistance to *Caligus rogercresseyi* in Atlantic salmon (*Salmo salar* L.) using a 50K SNP genotyping array. *Aquaculture* **472:** 61-65.

Daetwyler HD, Villanueva B, Bijma P, Woolliams JA (2007). Inbreeding in genome-wide selection. *J Anim Breed Genet* **124:** 369-376.

Das S, P. Sahoo P (2014). Markers for selection of disease resistance in fish: a review. *Aquacult Int* **22:** 1793-1812.

Defoirdt T, Sorgeloo P, Bossier P (2011). Alternatives to antibiotics for the control of bacterial disease in aquaculture. *Curr Opin Microbiol* **14:** 251-258.

Dekkers JCM (2016). Blueprint: Breeding for resistance to infectious disease. *National Hog Farmer.* Available from URL: http://www.nationalhogfarmer.com/animal-well-being/blueprint-breeding-resistance-infectious-disease.

Demers NE, Bayne CJ (1997). The immediate effects of stress on hormones and plasma lysozyme in rainbow trout. *Dev Comp Immunol* **21:** 363-373.

Doan QK, Vandeputte M, Chatain B, Morin T, Allal F (2016). Viral encephalopathy and retinopathy in aquaculture: a review. *J Fish Dis* **40:** 717-742.

Doan QK, Vandeputte M, Chatain B, Haffray P, Vergnet A, Breuil G, Allal F (2017). Genetic variation of resistance to Viral Nervous Necrosis and genetic correlations with production traits in wild populations of the European sea bass (*Dicentrarchus labrax*). *Aquaculture* **478:** 1-8.

Doeschl-Wilson A, Villanueva B, Kyriazakis I (2012). The first step toward genetic selection for host tolerance to infectious pathogens: obtaining the tolerance phenotype through group estimates. *Front Genet* **3**: 265.

Dorson M, Quillet E, Hollebecq MG, Torhy C, Chevassus B (1995). Selection of rainbow trout resistant to viral haemorrhagic septicaemia virus and transmission of resistance by gynogenesis. *Vet Res* **26:** 361-368.

Drangsholt T, Gjerde B, Ødegård J, Finne-Fridell F, Evensen Ø, Bentsen H (2011). Quantitative genetics of disease resistance in vaccinated and unvaccinated Atlantic salmon (*Salmo salar* L.). *Heredity* **107**: 471-477.

Dreanno C, Suquet M, Desbruyères E, Cosson J, Le Delliou H, Billard R (1998). Effect of urine on semen quality in turbot (*Psetta maxima*). *Aquaculture* **169**: 247-262.

Ducrocq V, Casella G (1996). A Bayesian analysis of mixed survival models. *Genet Sel Evol* **28:** 505-529.

Ellis T, Yildiz H, López-Olmeda J, Spedicato M, Tort L, Øverli Ø, Martins CI (2012). Cortisol and finfish welfare. *Fish Physiol Biochem* **38**: 163-188.

Esteban M, Rodríguez A, Ayala A, Meseguer J (2004). Effects of high doses of cortisol on innate cellular immune response of sea bream (*Sparus aurata* L.). *Gen Comp Endocrinol* **137**: 89-98.

Evenhuis J, Leeds T, Marancik D, LaPatra S, Wiens G (2015). Rainbow trout (*Oncorhynchus mykiss*) resistance to columnaris disease is heritable and favorably correlated with bacterial cold water disease resistance. *J Anim Sci* **93**: 1546-1554.

FAO yearbook (2018). *Fishery and Aquaculture Statistics 2016*, Rome, 104 pp.

Farias TF, César JRDO, Silva LPD (2017). Methods of selection using the quantitative genetics in aquaculture - a short review. *Insights Aquac Cult Biotechnol* **1:**1.

Fast M, Hosoya S, Johnson S, Afonso L (2008). Cortisol response and immune-related effects of Atlantic salmon (*Salmo salar* L.) subjected to short- and long-term stress. *Fish Shellfish Immunol* **24**: 194-204.

Fatira E, Papandroulakis N, Pavlidis M (2014). Diel changes in plasma cortisol and effects of size and stress duration on the cortisol response in European sea bass (*Dicentrarchus labrax*). *Fish Physiol Biochem* **40**: 911-919.

Fauvel C, Boryshpolets S, Cosson J, Wilson Leedy J, Labbé C, Haffray P, Suquet M. (2012). Improvement of chilled seabass sperm conservation using a cell culture medium. *J Appl Ichthyol* **28**: 961-966.

FEAP. Annual Report (2017). Available from URL: http://www.feap.info

Fevolden S, Refstie T, Røed K (1992). Disease resistance in rainbow trout (*Oncorhynchus mykiss*) selected for stress response. *Aquaculture* **104**: 19-29.

Fevolden S, Nordmo R, Refstie T, Røed K (1993). Disease resistance in Atlantic salmon (*Salmo salar*) selected for high or low responses to stress. *Aquaculture* **109**: 215-224.

Fevolden S, Røed K, Gjerde B (1994). Genetic components of post-stress cortisol and lysozyme activity in Atlantic salmon - correlations to disease resistance. *Fish Shellfish Immunol* **4**: 507-519.

Fevolden S, Roed K, Fjalestad K, Stien J (1999). Poststress levels of lysozyme and cortisol in adult rainbow trout: heritabilities and genetic correlations. *J Fish Biol* **54**: 900-910.

Fevolden S, Røed K, Fjalestad K (2002). Selection response of cortisol and lysozyme in rainbow trout and correlation to growth. *Aquaculture* **205**: 61-75.

Geven E, Klaren P (2017). The teleost head kidney: Integrating thyroid and immune signalling. *Dev Comp Immunol* **66**: 73-83.

Gheyas A, Haley C, Guy D, Hamilton A, Tinch A, Mota-Velasco J et al. (2010). Effect of a major QTL affecting IPN resistance on production traits in Atlantic salmon. *Anim Genet* **41**: 666-668.

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

Gjedrem T, Baranski M (2009). Selection methods, pp. 93-102. In: *Selective breeding in aquaculture: an introduction.* Gjedrem T, Baranski M (Eds.), Springer Dordrecht Heidelberg London New York.

Gjedrem T, Robinson N (2014). Advances by selective breeding for aquatic species: a review. *Agric Sci* **5**: 1152-1158.

Gjerde B, Evensen Ø, Bentsen H, Storset A (2009). Genetic (co)variation of vaccine injuries and innate resistance to furunculosis (*Aeromonas salmonicida*) and infectious salmon anaemia (ISA) in Atlantic salmon (*Salmo salar*). *Aquaculture* **287**: 52-58.

Gjerde B, Ødegård J, Thorland I (2011). Estimates of genetic variation in the susceptibility of Atlantic salmon (*Salmo salar*) to the salmon louse *Lepeophtheirus salmonis*. *Aquaculture* **314**: 66-72.

Gjøen H, Refstie T, Ulla O, Gjerde B (1997). Genetic correlations between survival of Atlantic salmon in challenge and field tests. *Aquaculture* **158:** 277-288.

Glover K, Aasmundstad T, Nilsen F, Storset A, Skaala Ø (2005). Variation of Atlantic salmon families (*Salmo salar* L.) in susceptibility to the sea lice *Lepeophtheirus salmonis* and *Caligus elongatus*. *Aquaculture* **245**: 19-30.

Goetz F, Coffman M (2000). Storage of unfertilized eggs of rainbow trout (*Oncorhynchus mykiss*) in artificial media. *Aquaculture* **184**: 267-276.

Gonen S, Baranski M, Thorland I, Norris A, Grove H, Arnesen P et al. (2015). Mapping and validation of a major QTL affecting resistance to pancreas disease (salmonid alphavirus) in Atlantic salmon (*Salmo salar*). *Heredity* **115**: 405-414.

Guardiola FA, Cuesta A, Esteban MÁ (2016). Using skin mucus to evaluate stress in gilthead seabream (*Sparus aurata* L.). *Fish Shellfish Immunol* **59:** 323-330.

Guy D, Bishop S, Brotherstone S, Hamilton A, Roberts R, McAndrew B et al. (2006). Analysis of the incidence of infectious pancreatic necrosis mortality in pedigreed Atlantic salmon, *Salmo salar* L., populations. *J Fish Dis* **29**: 637-647.

Guy D, Bishop S, Woolliams J, Brotherstone S (2009). Genetic parameters for resistance to Infectious Pancreatic Necrosis in pedigreed Atlantic salmon (*Salmo salar*) post-smolts using a Reduced Animal Model. *Aquaculture* **290**: 229-235.

Guy S, Thomson P, Hermesch S (2012). Selection of pigs for improved coping with health and environmental challenges: breeding for resistance or tolerance? *Front Genet* **3**: 281.

Hedrick RP (1998). Relationships of the host, pathogen, and environment: implications for diseases of cultured and wild fish populations. *J Aquat Anim Health* **10:** 107-111.

Henryon M, Jokumsen A, Berg P, Lund I, Pedersen P, Olesen N et al. (2002). Genetic variation for growth rate, feed conversion efficiency, and disease resistance exists within a farmed population of rainbow trout. *Aquaculture* **209**: 59-76.

Henryon M, Berg P, Olesen N, Kjær T, Slierendrecht W, Jokumsen A et al. (2005). Selective breeding provides an approach to increase resistance of rainbow trout (*Onchorhynchus mykiss*) to the diseases, enteric redmouth disease, rainbow trout fry syndrome, and viral haemorrhagic septicaemia. *Aquaculture* **250**: 621-636.

Heuer O, Kruse H, Grave K, Collignon P, Karunasagar I, Angulo F (2009). Human health consequences of use of antimicrobial agents in aquaculture. *Clin Infect Dis* **49**: 1248-1253.

Hontela A, Rasmussen J, Audet C, Chevalier G (1992). Impaired cortisol stress response in fish from environments polluted by PAHs, PCBs, and mercury. *Arch Environ Con Tox* **22:** 278-283.

Houston R, Haley C, Hamilton A, Guy D, Mota-Velasco J, Gheyas A et al. (2010). The susceptibility of Atlantic salmon fry to freshwater infectious pancreatic necrosis is largely explained by a major QTL. *Heredity* **105**: 318-327.

Houston R, Taggart J, Cézard T, Bekaert M, Lowe N, Downing A et al. (2014). Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). *BMC Genomics* **15**: 90.

Iwama G, Afonso L, Todgham A, Ackerman P, Nakano K (2004). Are HSPs suitable for indicating stressed states in fish? *J Exp Biol* **207**: 15-19.

Kettunen Præbel A, Westgård JI, Peruzzi S, Fevolden SE (2007a). Genetic parameters for post-stress cortisol activity and vibriosis resistance in Atlantic cod (*Gadus morhua* L.). *Aquaculture* **272**: S275-S276.

Kettunen Præbel A, Serenius T, Fjalestad K (2007b). Three statistical approaches for genetic analysis of disease resistance to vibriosis in Atlantic cod (L.). *J Anim Sci* **85:** 305.

Kjøglum S, Henryon M, Aasmundstad T, Korsgaard I (2008). Selective breeding can increase resistance of Atlantic salmon to furunculosis, infectious salmon anaemia and infectious pancreatic necrosis. *Aquacult Res* **39**: 498-505.

Kolstad K, Heuch P, Gjerde B, Gjedrem T, Salte R (2005). Genetic variation in resistance of Atlantic salmon (*Salmo salar*) to the salmon louse *Lepeophtheirus salmonis. Aquaculture* **247**: 145-151.

Kube P, Taylor R, Elliott N (2012). Genetic variation in parasite resistance of Atlantic salmon to amoebic gill disease over multiple infections. *Aquaculture* **364-365**: 165-172.

Kumar R, Lee J, Cho J (2012). Fate, occurrence, and toxicity of veterinary antibiotics in environment. *J Korean Soc Appl Biol Chem* **55**: 701-709.

Labbé C, Martoriati A, Devaux A, Maisse G (2001). Effect of sperm cryopreservation on sperm DNA stability and progeny development in rainbow trout. *Mol Reprod Dev* **60**: 397-404.

Leeds T, Silverstein J, Weber G, Vallejo R, Palti Y, Rexroad C et al. (2010). Response to selection for bacterial cold water disease resistance in rainbow trout. *J Anim Sci* **88**: 1936-1946.

Lillehammer M, Meuwissen T, Sonesson A (2013). A low-marker density implementation of genomic selection in aquaculture using within-family genomic breeding values. *Genet Select Evol* **45**: 39.

MacKenzie S, Ribas L, Pilarczyk M, Capdevila D, Kadri S, Huntingford F (2009). Screening for coping style increases the power of gene expression studies. *PLoS ONE* **4**: e5314.

Magnadóttir B (2006). Innate immunity of fish (overview). *Fish Shellfish Immunol* **20:** 137-151.

Magnadóttir B (2010). Immunological Control of Fish Diseases. *Mar Biotechnol* **12:** 361-379.

Magnadottir B, Gudmundsdottir S, Gudmundsdottir B, Helgason S (2009). Natural antibodies of cod (*Gadus morhua* L.): specificity, activity and affinity. *Comp Biochem Phys B* **154:** 309-316.

Mallard B, Wilkie BN, Kennedy BW, Gibson JP, Quinton M (1998). Immune responsiveness in swine: eight generations of selection for high and low immune response in Yorkshire pigs. *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, 11th-16th January, Armidale, New England, Australia.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ et al. (2009) Finding the missing heritability of complex diseases. *Nature* **461:** 747-753.

Massault C, Franch R, Haley C, de Koning D, Bovenhuis H, Pellizzari C et al. (2010). Quantitative trait loci for resistance to fish pasteurellosis in gilthead sea bream (*Sparus aurata*). *Anim Genet* **42**: 191-203.

Mauri I, Romero A, Acerete L, MacKenzie S, Roher N, Callol A et al. (2011). Changes in complement responses in gilthead seabream (*Sparus aurata*) and European seabass (*Dicentrarchus labrax*) under crowding stress, plus viral and bacterial challenges. *Fish Shellfish Immunol* **30**: 182-188.

McDonald G, Milligan L (1997). Ionic, osmotic and acid-base regulation in stress, pp. 119-144. In: *Fish stress and health in aquaculture*. Iwama GK, Pickering AD, Sumpter JP, Schreck C (Eds.) Cambridge University Press, Cambridge.

Meuwissen THE, Hayes BJ, Goddard ME (2016). Genomic selection: a paradigm shift in animal breeding. *Animal Frontiers* **6:** 6.

Mi Q, Zhou L, Schulze D, Fischer R, Lustig A, Rezanka L et al. (2000). Highly reduced protection against Streptococcus pneumoniae after deletion of a single heavy chain gene in mouse. *Proc Natl Acad Sci USA* **97:** 6031-6036.

Miller K, Winton J, Schulze A, Purcell M, Ming T (2004). Major Histocompatibility Complex loci are associated with susceptibility of Atlantic salmon to infectious hematopoietic necrosis virus. *Environ Biol Fishes* **69**: 307-316.

Moen T, Baranski M, Sonesson A, Kjøglum S (2009). Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*): population-level associations between markers and trait. *BMC Genomics* **10**: 368.

Mohanty B, Sahoo P (2007). Edwardsiellosis in fish: a brief review. *J Biosci* **32:** 1331-1344.

Mommsen TP, Vijayan MM, Moon TW (1999). Cortisol in teleosts: dynamics, mechanisms of action, and metabolic regulation. *Rev Fish Biol Fisher* **9:** 211-268.

Mylonas C, Fostier A, Zanuy S (2010). Broodstock management and hormonal manipulations of fish reproduction. *Gen Comp Endocrinol* **165**: 516-534.

Norris A, Foyle L, Ratcliff J (2008). Heritability of mortality in response to a natural pancreas disease (SPDV) challenge in Atlantic salmon, *Salmo salar* L., post-smolts on a West of Ireland sea site. *J Fish Dis* **31**: 913-920.

Ochsenbein AF, Fehr T, Lutz C, Suter M, Brombacher F, Hengartner H, Zinkernagel RM (1999). Control of early viral and bacterial distribution and disease by natural antibodies. *Science* **286:** 2156-2159.

Ochsenbein AF, Zinkernagel RM (2000). Natural antibodies and complement link innate and acquired immunity. *Immunol today* **21**: 624-630.

Ødegård J, Olesen I, Gjerde B, Klemetsdal G (2007). Positive genetic correlation between resistance to bacterial (furunculosis) and viral (infectious salmon anaemia) diseases in farmed Atlantic salmon (*Salmo salar*). *Aquaculture* **271**: 173-177.

Ødegård, J, Sommer AI, Kettunen Præbel A (2010a). Heritability of resistance to viral nervous necrosis in Atlantic cod (*Gadus morhua* L.). *Aquaculture* **300:** 59-64.

Ødegård, J, Olesen I, Dixon P, Jeney Z, Nielsen H, Way K, Joiner C, Jeney G, Ardó L, Rónyai A, Gjerde B (2010b). Genetic analysis of common carp (*Cyprinus carpio*) strains. II: Resistance to koi herpesvirus and *Aeromonas hydrophila* and their relationship with pond survival. *Aquaculture* **304:** 7-13.

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

Ødegård J, Meuwissen T (2014). Identity-by-descent genomic selection using selective and sparse genotyping. *Genet Select Evol* **46**: 3.

Olesen I, Hung D, Ødegård J (2007). Genetic analysis of survival in challenge tests of furunculosis and ISA in Atlantic salmon. Genetic parameter estimates and model comparisons. *Aquaculture* **272**: S297-S298.

Øverli Ø, Sørensen C, Kiessling A, Pottinger T, Gjøen H (2006). Selection for improved stress tolerance in rainbow trout (*Oncorhynchus mykiss*) leads to reduced feed waste. *Aquaculture* **261**: 776-781.

Overturf K, LaPatra S, Towner R, Campbell N, Narum S (2010). Relationships between growth and disease resistance in rainbow trout, *Oncorhynchus mykiss* (Walbaum). *J Fish Dis* **33**: 321-329.

Palaiokostas C, Bekaert M, Taggart J, Gharbi K, McAndrew B, Chatain B et al. (2015). A new SNP-based vision of the genetics of sex determination in European sea bass (*Dicentrarchus labrax*). *Genet Select Evol* **47**: 68.

Palaiokostas C, Ferraresso S, Franch R, Houston R, Bargelloni L (2016). Genomic prediction of resistance to pasteurellosis in Gilthead Sea bream (*Sparus aurata*) using 2b-RAD sequencing. *G3-Genes Genom Genet* **6**: 3693-3700.

Palaiokostas C, Cariou S, Bestin A, Bruant J, Haffray P, Morin T et al. (2018). Genome-wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (*Dicentrarchus labrax*) using RAD sequencing. *Genet Select Evol* **50**: 30.

Palmisano AN, Winton JR, Dickhoff WW (2000). Tissue-specific induction of HSP90 mRNA and plasma cortisol response in chinook salmon following heat shock, seawater challenge, and handling challenge. *Mar Biotechnol* **2:** 329-338.

Palti Y, Gao G, Liu S, Kent M, Lien S, Miller M et al. (2015). The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout. *Mol Ecol Resour* **15**: 662-672.

Parmentier H, Lammers A, Hoekman J, Reilingh G, Zaanen I, Savelkoul H (2004). Different levels of natural antibodies in chickens divergently selected for specific antibody responses. *Dev Comp Immunol* **28**: 39-49.

Pérez-Cerezales S, Martínez-Páramo S, Beirão J, Herráez M (2010). Evaluation of DNA damage as a quality marker for rainbow trout sperm cryopreservation and use of LDL as cryoprotectant. *Theriogenology* **74**: 282-289.

Periasamy K, Pichler R, Poli M, Cristel S, Cetrá B, Medus D et al. (2014). Candidate gene approach for parasite resistance in sheep – variation in immune pathway genes and association with fecal egg count. *PLoS ONE* **9**: e88337.

Ploegaert T, Wijga S, Tijhaar E, van der Poel J, Lam T, Savelkoul H et al. (2010). Genetic variation of natural antibodies in milk of Dutch Holstein-Friesian cows. *Journal of Dairy Science* **93**: 5467-5473.

Ploegaert T, Tijhaar E, Lam T, Taverne-Thiele A, van der Poel J, van Arendonk J et al. (2011). Natural antibodies in bovine milk and blood plasma: variability among cows, repeatability within cows, and relation between milk and plasma titers. *Vet Immunol Immunopathol* **144**: 88-94.

Pottinger TG (2000). Genetic selection to reduce stress in animals, pp. 291-308. In: *The biology of animal stress basic principles and implications for animal welfare*, Moberg GP, Mench JA (Eds.) CABI Publishing, New York.

Pottinger TG (2008). The stress response in fish-mechanisms, effects and measurement, pp. 32-44. In: *Fish Welfare*, Branson EJ (Ed.) Blackwell Publishing Ltd, Oxford.

Pszczola M, Strabel T, van Arendonk J, Calus M (2012). The impact of genotyping different groups of animals on accuracy when moving from traditional to genomic selection. *J Dairy Sci* **95**: 5412-5421.

Quillet E, Krieg F, Dechamp N, Hervet C, Bérard A, Le Roy P et al. (2014). Quantitative trait loci for magnitude of the plasma cortisol response to confinement in rainbow trout. *Anim Genet* **45**: 223-234.

Rexroad C, Vallejo R, Liu S, Palti Y, Weber G (2012). QTL affecting stress response to crowding in a rainbow trout broodstock population. *BMC Genetics* **13**: 97.

Robinson NA, Gjedrem T, Quillet E (2017). Testing for resistance: natural outbreaks versus controlled challenge testing, pp. 22-23. In: *Fish Disease, prevention and control strategies*, Jeney G (Ed.) Academic press, Elsevier, London.

Røed K, Fjalestad K, Strømsheim A (1993). Genetic variation in lysozyme activity and spontaneous haemolytic activity in Atlantic salmon (*Salmo salar*). *Aquaculture* **114**: 19-31.

Røed K, Fevolden S, Fjalestad K (2002). Disease resistance and immune characteristics in rainbow trout (*Oncorhynchus mykiss*) selected for lysozyme activity. *Aquaculture* **209**: 91-101.

Routray P, Suzuki T, Strüssmann C, Takai R (2002). Factors affecting the uptake of DMSO by the eggs and embryos of medaka, *Oryzias latipes*. *Theriogenology* **58**: 1483-1496.

Roy BA, Kirchner JW (2000). Evolutionary dynamics of pathogen resistance and tolerance. *Evolution* **54:** 51-63.

Salte R, Bentsen H, Moen T, Tripathy S, Bakke T, Ødegård J et al. (2010). Prospects for a genetic management strategy to control *Gyrodactylus salaris* infection in wild Atlantic salmon (*Salmo salar*) stocks. *Can J Fish Aquat Sci* **67**: 121-129.

Sansone G, Fabbrocini A, Zupa A, Lavadera S, Rispoli S, Matassino D (2001). Inactivator media of sea bass (*Dicentrarchus labrax* L.) spermatozoa motility. *Aquaculture* **202**: 257-268.

Saura M, Villanueva B, Fernández J, Toro M (2017). Effect of assortative mating on genetic gain and inbreeding in aquaculture selective breeding programs. *Aquaculture* **472**: 30-37.

Silverstein J, Vallejo R, Palti Y, Leeds T, Rexroad C, Welch T et al. (2009). Rainbow trout resistance to bacterial cold-water disease is moderately heritable and is not adversely correlated with growth1. *J Anim Sci* **87**: 860-867.

Sinyakov M, Dror M, Zhevelev H, Margel S, Avtalion R (2002). Natural antibodies and their significance in active immunization and protection against a defined pathogen in fish. *Vaccine* **20**: 3668-3674.

Sonesson, AK, Meuwissen THE, Goddard ME (2010). The use of communal rearing of families and DNA pooling in aquaculture genomic selection schemes. *Genet Sel Evol* **42:** 41.

Stear MJ, Nikbakht G, Matthews L, Jonsson NN (2012). Breeding for disease resistance in livestock and fish. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, **7:** 1-10.

Storset A, Strand C, Wetten M, Kjøglum S, Ramstad A (2007). Response to selection for resistance against infectious pancreatic necrosis in Atlantic salmon (*Salmo salar* L.). *Aquaculture* **272**: S62-S68.

Subramaniam K, Datta K, Quintero E, Manix C, Marks M, Pirofski L (2010). The absence of serum IgM enhances the susceptibility of mice to pulmonary challenge with *Cryptococcus neoformans*. *J Immunol* **184**: 5755-5767.

Sunyer JO, Tort L (1995). Natural hemolytic and bactericidal activities of sea bream *Sparus aurata* serum are affected by the alternative complement pathway. *Vet Immunol Immunopathol* **45:** 333-345.

Taylor J, Needham M, North B, Morgan A, Thompson K, Migaud H (2007). The influence of ploidy on saltwater adaptation, acute stress response and immune function following seawater transfer in non-smolting rainbow trout. *Gen Comp Endocrinol* **152**: 314-325.

Tort L (2011). Stress and immune modulation in fish. *Dev Comp Immunol* **35:** 1366-1375.

Trenzado C, Morales A, de la Higuera M (2006). Physiological effects of crowding in rainbow trout, *Oncorhynchus mykiss*, selected for low and high stress responsiveness. *Aquaculture* **258**: 583-593.

Tsai H, Robledo D, Lowe N, Bekaert M, Taggart J, Bron J et al. (2016a). Construction and annotation of a high density SNP linkage map of the Atlantic Salmon (*Salmo salar*) genome. *G3-Genes Genom Genet* **6:** 2173-2179.

Tsai H, Hamilton A, Tinch A, Guy D, Bron J, Taggart J et al. (2016b). Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. *Genet Sel Evol* **48:** 47.

Valdez D, Miyamoto A, Hara T, Edashige K, Kasai M (2005). Sensitivity to chilling of medaka (*Oryzias latipes*) embryos at various developmental stages. *Theriogenology* **64**: 112-122.

Vallejo R, Palti Y, Liu S, Evenhuis J, Gao G, Rexroad C et al. (2014a). Detection of QTL in rainbow trout affecting survival when challenged with *Flavobacterium psychrophilum*. *Mar Biotechnol* **16**: 349-360.

Vallejo R, Palti Y, Liu S, Marancik D, Wiens G (2014b). Validation of linked QTL for bacterial cold water disease resistance and spleen size on rainbow trout chromosome Omy19. *Aquaculture* **432**: 139-143.

Vallejo R, Leeds T, Fragomeni B, Gao G, Hernandez A, Misztal I et al. (2016). Evaluation of genome-enabled selection for bacterial cold water disease resistance using progeny performance data in rainbow trout: insights on genotyping methods and genomic prediction models. *Front Genet* **7**: 96.

Vallejo R, Leeds T, Gao G, Parsons J, Martin K, Evenhuis J et al. (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 Select Evol* **49**:17.

van der Most P, de Jong B, Parmentier H, Verhulst S (2010). Trade-off between growth and immune function: a meta-analysis of selection experiments. *Functional Ecology* **25**: 74-80.

Vandeputte M, Porte J, Auperin B, Dupont-Nivet M, Vergnet A, Valotaire C et al. (2016). Quantitative genetic variation for post-stress cortisol and swimming performance in growth-selected and control populations of European sea bass (*Dicentrarchus labrax*). *Aquaculture* **455**: 1-7.

Volckaert F, Hellemans B, Batargias C, Louro B, Massault C, Van Houdt J et al. (2012). Heritability of cortisol response to confinement stress in European sea bass *Dicentrarchus labrax*. *Genet Select Evol* **44**: 15.

Wagner T, Congleton JL (2004). Blood chemistry correlates of nutritional condition, tissue damage, and stress in migrating juvenile chinook salmon (*Oncorhynchus tshawytscha*). *Can J. Fish Aquat Sci* **61:** 1066-1074.

Wang R, Belosevic M (1995). The in vitro effects of estradiol and cortisol on the function of a long-term goldfish macrophage cell line. *Dev Comp Immunol* **19**: 327-336.

Wang W, Hu Y, Ma Y, Xu L, Guan J, Kong J (2015). High-density genetic linkage mapping in turbot (*Scophthalmus maximus* L.) based on SNP markers and major sex- and growth-related regions detection. *PLoS ONE* **10**: e0120410.

Weber G, Vallejo R, Lankford S, Silverstein J, Welch T (2008). Cortisol response to a crowding stress: heritability and association with disease resistance to *Yersinia ruckeri* in rainbow trout. *N Am J Aquac* **70**: 425-433.

Wendelaar Bonga SE (1997). The stress response in fish. *Physiol Rev* **77:** 591-625.

Wetten M, Aasmundstad T, Kjøglum S, Storset A (2007). Genetic analysis of resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar* L.). *Aquaculture* **272**: 111-117.

Wiens G, Vallejo R, Leeds T, Palti Y, Hadidi S, Liu S et al. (2013). Assessment of genetic correlation between bacterial cold water disease resistance and spleen index in a domesticated population of rainbow trout: identification of QTL on chromosome Omy19. *PLoS ONE* **8**: e75749.

Wondmeneh E, Van Arendonk J, Van der Waaij E, Ducro B, Parmentier H (2015). High natural antibody titers of indigenous chickens are related with increased hazard in confinement. *Poult Sci* **94**: 1493-1498.

Xu P, Zhang X, Wang X, Li J, Liu G, Kuang Y et al (2014). Genome sequence and genetic diversity of the common carp, *Cyprinus carpio. Nat Genet* **46:** 1212-1219.

Yáñez J, Bangera R, Lhorente J, Oyarzún M, Neira R (2013). Quantitative genetic variation of resistance against *Piscirickettsia salmonis* in Atlantic salmon (*Salmo salar*). *Aquaculture* **414-415:** 155-159.

Yáñez J, Naswa S, López ME, Bassini L, Cabrejos ME, Gilbey J, Bernatchez L, Norris A, Soto C, Eisenhart J, Simpson B, Neira R, Lhorente JP, Schnable P, Newman S, Mileham A, Deeb N (2014). Development of a 200K SNP array for Atlantic salmon: exploiting across continents genetic variation. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*, August 17th-22th, 2014, Vancouver, BC, Canada.

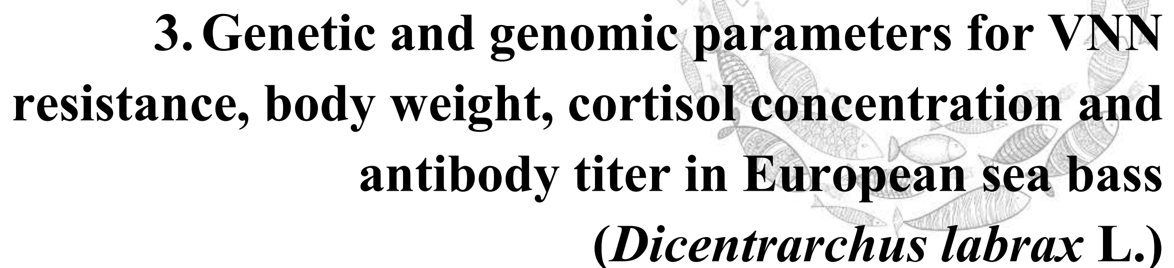
Yin Z, Lam T, Sin Y (1995). The effects of crowding stress on the non-specific immune response in fancy carp (*Cyprinus carpio* L.). *Fish Shellfish Immunol* **5**: 519-529.

Yu XZ, Meuwissen THE, Baranski M, Sonesson AK (2014). Selective breeding against infectious diseases in Atlantic cod with whole genome sequence data. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*, August 17th-22th, Vancouver, BC, Canada.

Zhang T, Isayeva A, Adams S, Rawson D (2005). Studies on membrane permeability of zebrafish (*Danio rerio*) oocytes in the presence of different cryoprotectants. *Cryobiology* **50**: 285-293.

Zhang T, Lubzens E (2010). Cryopreservation of fish oocytes. In: *Methods in reproductive aquaculture marine and freshwater species*. (Cabrita E, Robles V, Herráez P, Eds.) CRC press, Taylor and Francis Group, Boca Raton.

Zilli L, Schiavone R, Zonno V, Storelli C, Vilella S (2003). Evaluation of DNA damage in *Dicentrarchus labrax* sperm following cryopreservation. *Cryobiology* **47**: 227-235.



3. Genetic and genomic parameters for VNN resistance, body weight, cortisol concentration and antibody titer in European sea bass (*Dicentrarchus labrax* L.)

Sara Faggion[1], Daniela Bertotto[1], Massimiliano Babbucci[1], Giulia Dalla Rovere[1], Rafaella Franch[1], Francesco Pascoli[2], Anna Toffan[2], Luca Bargelloni[1] and Paolo Carnier[1]

[1] Department of Comparative Biomedicine and Food Science (BCA), University of Padova, Italy
[2] Istituto Zooprofilattico Sperimentale delle Venezie, National Reference Laboratory (NRL) for diseases of fish, mollusk and crustacean, Legnaro (PD), Italy

Abstract

The susceptibility of European sea bass (*Dicentrarchus labrax* L.) to viral nervous necrosis (VNN) is well-known and the interest towards selective breeding for improved genetic resistance in this species has sharply increased, due to the major threat represented by VNN for farmed sea bass and because of the lack of specific therapeutical measures.

We estimated the genetic and genomic parameters for resistance traits (mortality and survival time after a VNN challenge test) in a sea bass experimental population (N = 650) derived from a commercial breeding stock. A genome-wide SNP dataset for sea bass was generated through a high-throughput sequencing approach (2b-RAD). Estimates of heritabilities for mortality and survival time were moderate or low (genetic h^2_{mort} = 0.14-0.23, genomic h^2_{mort} = 0.06; genetic h^2_{surv} = 0.07-0.14, genomic h^2_{surv} = 0.03). Due to the possible application of indirect selection to improve VNN resistance, we investigated the variation in a physiological parameter (post-stress cortisol concentration) and an immunological parameter (antibody titer to nervous necrosis virus, NNV) to determine the additive genetic variation of these traits and the genetic correlation with disease resistance. The genetic relationship with body weight at a constant age was also investigated.

Cortisol concentration showed a significant genetic variability (genetic h^2 = 0.19-0.23, genomic h^2 = 0.19), but a null genetic correlation with disease resistance traits.

The heritability of antibody titer against NNV was moderate (genetic h^2 = 0.28-0.39, genomic h^2 = 0.26) and the genetic relationship (r_a) with mortality was negative (r_a = -0.39).

A negative genetic correlation was detected between mortality and body weight (r_a = -0.39), whereas body weight was positively correlated with antibody titer (r_a = 0.49).

In conclusion, mortality, body weight, cortisol concentration and antibody titer exhibited significant genetic variability, which makes selective breeding for VNN resistance a feasible approach in European sea bass, but specific strategies to implement such selection need to be investigated thoroughly.

Keywords European sea bass, viral nervous necrosis, cortisol, antibody titer, genetic parameters, genomic parameters, genetic correlation

3.1 Introduction

Disease outbreaks are, in general, severe threats that might interfere with the progress and sustainability of intensive aquaculture systems. Viral nervous necrosis (VNN), also known as viral encephalopathy and retinopathy (VER), affects more than 50 different marine and freshwater fish species. In recent years, Nodavirus (or nervous necrosis virus, NNV) infections have become one of the major diseases for aquaculture, due to serious consequences, both in terms of animal and economical losses (Doan et al. 2016).

NNV is a non-enveloped icosahedral virus of the Nodaviridae family, genus *Betanodavirus*, whose genome is composed of two single-stranded positive sense RNA segments (RNA1 and RNA2; Yong et al. 2017), characterized by both vertical and horizontal transmission (Munday et al. 2002). In fish, betanodaviruses are classified in four genotypes, according to the RNA2 T4 variable region: striped jack nervous necrosis virus (SJNNV, the first isolated strain), tiger puffer nervous necrosis virus (TPNNV), barfin flounder nervous necrosis virus (BFNNV) and red-spotted grouper nervous necrosis virus (RGNNV, very common in warm-water, such as the Mediterranean area; Munday et al. 2002; Bovo et al. 1999).

NNV has been detected in numerous organs, but major damages have been observed in the central nervous system (brain, spinal cord and retina; Mori et al. 1992), causing in fish recognizable clinical and behavioral signs (abnormal swimming patterns, lethargy, skin darkening, loss of appetite; Munday and Nakai 1997).

Among the species recognized as susceptible to NNV, European sea bass (*Dicentrarchus labrax* L.), one of the most common and valuable marine species widely cultured in the Mediterranean areas (FEAP Annual report 2016), is particularly threatened. Outbreaks of NNV at the larval stage lead to mortality rates up to 80-100% (Munday et al. 2002) and moderate mortalities have been reported even in advanced juveniles (11-20%; Le Breton et al. 1997). In survivor fish, NNV infection might become chronic, causing an impairment of the general conditions and poor growth rates (Vendramin et al. 2014).

Chemotherapeutics are not effective as a mean to control VNN (Doan et al. 2016). Different NNV vaccination approaches, such as those based on synthetic peptides of capsid proteins (Coeurdacier et al. 2003), virus-like particles (VLPs; Thiéry et al. 2006), inactivated viruses (Nuñez-Ortiz et al. 2016a) and DNA vaccines (Valero et al. 2016) have been investigated in sea bass. Despite some of these vaccination strategies showed relatively good results in terms of protection against NNV, only one commercially NNV vaccine based on inactivated RGNNV is currently available (Alpha Ject micro® 1Noda, PHARMAQ). The type of administration (intraperitoneal injection) and the relatively short time of protection (maximum 12 months) raise questions about the real applicability in a context of aquaculture production; moreover, the minimum weight of the fish required to inject the vaccine (12 g) precludes the protection of the larval stages, when the infectivity of the virus is the highest.

Recently, attention has been directed to selective breeding as a tool for enhancement of VNN resistance because, in general, host resistance plays a key role in hindering the pathogen spread or in lowering the infection pressure. Interest in selective breeding as a disease prevention action is supported by the magnitude of additive genetic variation estimated for disease resistance in different aquaculture species. Heritability of VNN resistance in sea bass has been reported to be comparable to that of some traits that, in terrestrial species, undergo artificial selection since many years (0.26-0.27; Doan et al. 2017; Palaiokostas et al. 2018), making selective breeding a possible option for both

a long-term control of the disease (Das and Sahoo, 2014) and a cumulative and permanent improvement in resistance over generations at the population level (Doan et al. 2016).
Indirect approaches exploiting non-trivial genetic relationships between disease resistance and other indicator traits of relevant additive genetic variation may facilitate selective breeding for resistance to pathogens.
Post-stress cortisol level seems to be an interesting marker trait, due to its reliability in terms of stress assessment (Pottinger 2008) and a significant genetic variability (in sea bass heritability estimates range between 0.08 and 0.33; Volckaert et al. 2012; Vandeputte et al. 2016). The use of cortisol levels in selective breeding has been scarcely investigated and the genetic relationship between this physiological trait and disease resistance is unclear. In sea bass, estimates of the genetic correlation between post-stress cortisol concentration and VNN resistance are lacking.
Likewise, the genetic component of variation in antibody response has been slightly studied in fish species. Two studies on Atlantic salmon (Strømsheim et al. 1994; Fjalestad et al. 1996) reported estimates of heritability for antibody titer against antigens of *Vibrio anguillarum* and *Vibrio salmonicida* ranging from 0.02 to 0.18 and from 0.03 to 0.12, respectively. In sea bass, the genetic variation in antibody titer after viral or bacterial infection or the genetic relationship between antibody titer and disease resistance has not been investigated yet.
The aim of this study was to estimate the heritability for mortality and survival time after a VNN challenge test in a sea bass experimental population derived from a commercial stock, as well as for cortisol concentration after stress exposure, antibody titer against VNN antigens and body weight at a fixed age. Genetic and phenotypic correlations among traits were also assessed. A high-throughput sequencing approach (2b-RAD) was applied to generate a genome-wide SNP dataset for sea bass, allowing the pedigree reconstruction of the experimental fish and the estimation of genomic heritability of the studied traits.

3.2 Materials and methods

3.2.1 Production and rearing of experimental fish

The experimental fish were produced in a commercial hatchery through artificial fertilization, using an incomplete factorial mating scheme, crossing every time 5 dams with 8 sires (20 dams and 32 sires in total). All the parents were NNV-free. The maturation stage of the dams was estimated through ovarian biopsy and the dams that reached the appropriate stage were hormonally injected (LH-RHa, 10 μg/kg). After 72 h, dams were stripped, eggs were gently mixed with the sperm (previously stripped and preserved following Fauvel et al. 2012) and then seawater was added (ratio: 0,5-2 ml sperm/100 ml eggs/50 ml seawater). All fertilized eggs were incubated in 2,000 l tanks at 13-14 °C, 35-38‰ salinity and oxygen between 8 and 10 ppm. Hatching occurred after 72 h. After hatching, temperature was gradually increased to 16 °C. At 40 d post-hatching, fish were transferred to 15,000 liters juvenile tanks (temperature of 18-19 °C, 35-38‰ salinity and oxygen between 8 and 10 ppm). At 164 d post-hatching, fish were transferred to a sea cage, where they were reared until the time of the experimental challenge test.

3.2.2 NNV challenge test

At the age of 539 d post-hatching, 652 randomly chosen experimental fish were transferred from the sea cage to the Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe, Legnaro, PD) and distributed into three close-system tanks of 2,500 litres (A, B, C) plus three close-system tanks of 380 litres (D, E, F) for an acclimation period of 9 d (pre-test period). Each tank was filled with artificial

salt water at 30‰ of salinity, temperature of 21 ± 1 °C, oxygen 6 ppm and artificial photoperiod of 10 h of light and 14 h of darkness. Fish were fed with a commercial diet (MRF Marine 3P, Skretting). After the pre-test period, fasting fish were subjected to an acute stress (confinement at high-density) for no less than 10 minutes. Fish were then anesthetized with 30 ppm of MS-222, weighed (body weight at 548 d post-hatching), individually tagged and a blood sample was collected from the caudal vein of each fish.
The virus strain used for the infection was RGNNV 283.2009, that was isolated from animals belonging to a commercial stock during a severe outbreak in the northern Adriatic sea and that showed a mortality rate of 35% in sea bass during an experimental infection by immersion (Vendramin et al. 2014). The strain RGNNV 283.2009 was propagated on a clone of SSN1 cell line, E11 cells (Iwamoto et al. 2000), and the titer was determined by endpoint dilutions assays.
The infection was then performed through intramuscular injection of 0.1 ml of viral suspension (RGNNV 283.2009, batch 7/16; $10^{8.30}$ $TCID_{50}$ per ml).
In general, the injection allows the direct infection of the animals with an equivalent dose of pathogen (i.e., the level of exposure is the same for all individuals); this particular method of infection is clearly less representative of natural disease outbreaks, but it assures great effectiveness and less variation and heterogeneity in pathogen load, particularly when large sized individuals are tested (Robinson et al. 2017).
At the beginning of the challenge test, temperature was increased to 25 ± 1 °C, whereas salinity and oxygen were maintained at 30‰ and 6 ppm, respectively. The photoperiod was the same as the pre-test period and fish were fed with the same commercial diet (MRF Marine 3P, Skretting). Fish were checked three times a day to identify typical clinical signs of the NNV infection and behaviours (abnormal swimming patterns, lethargy, skin darkening, loss of appetite) related to the disease and mortality and dead fish were removed each time. The experiment ended at 29 d post-challenge and all live fish, recorded as survivors, were euthanized with an overdose of anaesthetic (MS-222).
The experiment protocol was evaluated and approved by the Italian Ministry of Health (Law decree 26/2014 art. 31; permission number: 975/2016-PR of 13/10/2016).

3.2.3 Samples management

After collection, blood samples were kept at 4 °C for one night, centrifuged and the serum transferred to a new tube. All serum samples were then stored at -20 °C until use for indirect ELISA assay and cortisol analysis by radioimmunoassay (RIA).
Muscle and fin samples were collected both from dead fish throughout the experiment and from survivors at the end of the challenge test and preserved in absolute ethanol for subsequent genomic DNA extraction.

3.2.4 Indirect ELISA assay

The experimental fish were reared in a non-controlled environment (i.e., sea cage) before being used for the challenge test. For this reason, indirect ELISA for antibodies against NNV was performed at the Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe, Legnaro, PD) on the serum collected before the infection, following the protocol developed by Scapigliati et al. (2010) and Nuñez-Ortiz et al. (2016b). Each sample was measured in duplicate wells and optical density values (OD 450 nm) of control wells were automatically subtracted from OD 450 nm samples values. The status of the sample was evaluated through the ELISA OD sample to positive ratio (S/P ratio):

$$S/P\ ratio = \frac{sample\ mean\ (mean\ of\ OD) - negative\ control\ mean}{positive\ control\ mean - negative\ control\ mean}$$

3.2.5 Cortisol extraction and measurement

Cortisol from serum was extracted following the protocol described by Bertotto et al. (2010): 8 ml diethyl ether was added to 100 µl serum, shaken, centrifuged and kept at -20 °C. The supernatant was decanted in a new tube and dried under a stream of nitrogen at 40 °C. Dried samples were reconstituted in 1 ml of phosphate buffer (PBS, pH 7.2) with bovine serum albumin (BSA) 0.1%.
In order to perform solid-phase RIA in microplates as described by Simontacchi et al. (1995) with minor modifications, a 96-well microtiter polystyrene plate (Packard, optiplate 96 well, PerkinElmer Life and Analytical Sciences, Shelton, CT, USA) was firstly coated with anti-rabbit γ-globulin serum raised in goat diluted 1:1,000 in 0.15mM sodium acetate buffer (pH 9), incubated for 23 hours at 4 °C, emptied, washed with PBS, incubated again at 4 °C for 1 hour, emptied and preserved at -20 °C until required.
At the moment of analysis, the plate was incubated with the anti-cortisol serum solution 1:30,000 at 37 °C for 2 hours. The anti-cortisol serum raised in rabbit showed the following specificity features: cortisol 100%, prednisolone 44.3%, 11-deoxycortisol 13.9%, cortisone 4.9%, corticosterone 3.5%, prednisone 2.7%, 17-hydroxyprogesterone 1.0%, 11-deoxycorticosterone 0.3%, dexamethasone 0.1%, progesterone < 0.01%, 17-hydroxypregnenolone < 0.01% and pregnenolone < 0.01%.
Then the plate was washed once with PBS; standards (to create a calibration curve), quality controls, diluted extracts and 1,2,6,7-^{3}H tracer (PerkinElmer Life and Analytical Sciences) were added and the plate was incubated for one night at 4 °C. The plate was emptied and washed four times with PBS, scintillation cocktail (Microscint-20, PerkinElmer Life and Analytical Sciences) was added, the plate was closed with a heat seal (TOPSEAL-S, PerkinElmer Life and Analytical Sciences) and the radioactivity was counted on a β-counter (Top-Count NXT, Perkin Elmer Life and Analytical Sciences).
Assessed radioactivity values were analysed with the GraphPad Prism 5.0 software (La Jolla, CA, USA), that created a calibration curve and provided the cortisol concentration of each sample (pg/well). Cortisol concentration expressed as pg/well was then converted to ng per ml of serum.
To validate RIA cortisol, an extractive yield test, a parallelism test and inter/intra-assay precision tests were performed. The extractive yield test was performed adding to randomly chosen samples a determined quantity of 1,2,6,7-^{3}H tracer solution. The β-counter reading allowed the estimation of the percentage of 1,2,6,7-^{3}H tracer solution that has been regained after the extraction protocol. The tests of parallelism were achieved in two steps: firstly, by analysing the serially diluted extracts of two samples with high cortisol concentrations, and secondly, by analysing two extracted samples with high and low cortisol concentrations that were diluted together, each time using a different quantity of the “high” and the “low” sample. The inter/intra-assay precision test was performed analysing six repetitions of the same extracted samples between different plates and within the same plate.

3.2.6 Genomic DNA extraction, 2b-RAD libraries construction and sequencing

Genomic DNA was extracted from approximately 20 mg of tissue (muscle or fin; 652 offspring, 52 parents) using the commercial kits Invisorb® Spin Tissue Mini Kit and Invisorb® DNA Tissue HTS 96 Kit (Invitek, STRATEC Biomedical, Germany) following the manufacturer’s recommendations and its integrity was assessed by visualization on a 1% TAE agarose gel stained with SYBR® Safe DNA Gel Stain (Invitrogen - ThermoFisher Scientific). Genomic DNA concentration and purity were

quantified by using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) in terms of absorbance at 260/230 nm and at 260/280 nm. This procedure ensured comparable concentrations of high quality gDNA as required for 2b-RAD library preparation.
A 2b-RAD library was constructed for each individual following the protocol from Wang et al. (2012) with minor modifications.
Between 300 and 1000 ng of gDNA from each sample was cleaved with 2 U of the type 2b restriction endonucleases *Alf* I (Thermofisher Scientific) for 1 hour at 37 °C and 20 minutes at 65 °C for denaturation, producing a population of fragments of uniform length (37 bp) with protruding ends.
The digested products were ligated to partially double-stranded adaptors with compatible and fully degenerated overhangs in a 25 µl total volume reaction consisting of 0.4 µM of each adaptor, 0.2 mM ATP (New England Biolabs, NEB, Ipswich, Massachusetts, USA) and 1 U T4 DNA ligase (SibEnzyme Ltd., Academ town, Siberia).
Barcodes were designed by Barcode Generator (available at UC Davis Web site: http://comailab.genomecenter.ucdavis.edu/ index.php/Barcode_generator).
The 2b-RAD tags were amplified for 13 cycles (3 hours at 16 °C and 10 minutes at 65 °C) using two pairs of primers: the first pair (2bRAD_BC and 2bRAD_F; Eurofins Genomics SRL, Italy) to introduce sample-specific barcodes (7 bp) and the annealing sites for Illumina next-generation sequencing; the second pair (2bRAD_ampl_for and 2bRAD_ampl_rev; Eurofins Genomics SRL, Italy) to perform a bland amplification. The mixture consisted of 12.5 µl of ligated DNA, 0.5 µM of each 2bRAD_BC/2bRAD_F primer, 0.2 µM of each 2bRAD_ampl_for/2bRAD_ampl_rev primer, 0.3 mM dNTP (New England Biolabs, NEB, Ipswich, Massachusetts, USA), 1× Phusion high-fidelity buffer, 2 U Taq Phusion high-fidelity DNA polymerase (ThermoFisher Scientific) in a total volume of 50 µl. 2b-RAD tags were amplified splitting the 50 µl mixture in three wells in order to minimize PCR amplification bias (Mastretta-Yanes et al. 2015). Cycling conditions were as follows: 98 °C for 4 minutes; 98 °C for 5 seconds, 60 °C for 20 seconds, 72 °C for 5 seconds, for 14 cycles; 72 °C for 5 minutes. Amplification products were run on a 1.8% TAE agarose gel stained with SYBR® Safe DNA Gel Stain.
SPRIselect purification kit (Beckman Coulter, Pasadena, California, USA) consisting of magnetic beads was used to purify PCR products, in order to exclude any low-molecular weight DNA.
The concentration of each purified individual library was quantified using Qubit® ds DNA BR Assay kit (Invitrogen – ThermoFisher Scientific) and Mx3000P qPCR instrument (Agilent Technologies, Santa Clara, California, USA).
Individual libraries were pooled into equimolar amounts (52 for parents, 96-97 for offspring). Pooled libraries were analyzed with Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, California, USA) and then sequenced on an Illumina HiSeq4000 platform with a 50 bp single-read module at Fasteris SA (Plan-les-Ouates, Switzerland; http://www.fasteris.com) and UC Davis (http://comailab.genomecenter.ucdavis.edu).
To assess the robustness of the methods, two libraries were replicated from two samples (Technical Replicates, TRs).

3.2.7 SNP discovery, genotyping, parentage assignment and pedigree reconstruction

Demultiplexing and a first quality-check of raw data were performed by the same sequencing services. The software FastQC (http://bioinformatics.babraham.ac.uk) was used to visualize and confirm the quality of the raw demultiplexed reads.

A custom-made script was used to filter reads for the presence of *Alf* I recognition site and to trim adaptors, obtaining 34 bp fragments ready to be elaborated with STACKS software 2.0 (Catchen et al. 2013), which is a modular software pipeline designed to build loci from short-read sequences, like those generated by Illumina platform, thus allowing the identification of SNPs. The trimmed reads were firstly mapped against European sea bass genome (Tine et al. 2014; http://seabass.mpipz.mpg.de/DOWNLOADS/dicLab1_scaffold.fasta) using a length fraction of 1.0 and a similarity fraction equal to 0.9 (all remaining parameters as default). Mapping results were exported in SAM format and used as input for *refmap_map.pl* in STACKS. Firstly, the program runs *gstacks,* in order to build loci according to the alignment positions provided for each read before calling SNPs in each sample. To construct the stacks and the catalog of loci, a minimum coverage of 20X was used for parental samples, while for the offspring the minimum coverage was 5X. Then, the STACKS module "*populations"* was used to filter the outputs, excluding loci shared by less than 75% of the analysed individuals.

A total of 340 million demultiplexed and filtered quality reads were obtained for 714 samples (including 12 TRs). The number of reads per individual was on average 3,500,000. *Stacks* identified 351,203 loci (or tags) with 12,018,544 sites (253,999 variant sites). After filtering, the dataset consisted of 18,097 SNPs. Some SNP were discarded on the basis of a minor allele frequency (MAF) lower than 1%, genotype frequency deviating from Hardy-Weinberg equilibrium and missing genotype rate greater than 15%. Missing genotypes for the remaining SNP were then imputed using the FImpute software (Sargolzaei et al. 2014). The final number of available SNP genotypes per animal was 16,075. Two individuals of the challenged group were discarded from the used data because of bad genotyping results.

The parentage assignment was performed with a likelihood-based program, CERVUS 3.0 (Marshall et al. 1998; Kalinowski et al. 2007). For each tested individual, parentage was either assigned to the most-likely candidate parental pair with a level of confidence equal to 80%, or left unassigned otherwise. The R package sequoia (Huisman 2017) was used to identify sibling relationships between the 52 parents, assigning them dummy parents and reconstructing a high-likelihood pedigree of the 650 experimental fish.

3.2.8 Statistical analyses

Descriptive statistics

Descriptive statistics for the different traits were obtained using the MEANS and UNIVARIATE procedures in SAS 9.4 (SAS Institute Inc.). The frequency distribution of serum cortisol concentration, measured as ng per ml of serum, was skewed and then normalized through the square root transformation of the original data (Figure 1).

The Kaplan-Meier product-limit survival curve of the experimental fish challenged with NNV was estimated using the LIFETEST procedure in SAS 9.4.

Preliminary analyses of traits were performed with logistic regression and linear models (LOGISTIC and GLM procedures in SAS 9.4), to detect potential fixed effects. Tank effect was significant for mortality as a binary trait and survival time.

Genetic parameters

Variance components and genetic parameters were estimated using Bayesian procedures. Marginal posterior distributions for (co)variance components and related parameters were obtained through Monte-Carlo Markov Chain (MCMC) and Gibbs sampling methods implemented in the software TM (Legarra et al. 2008). A single Gibbs chain of 1,000,000 iterations was run for each trait analysis.

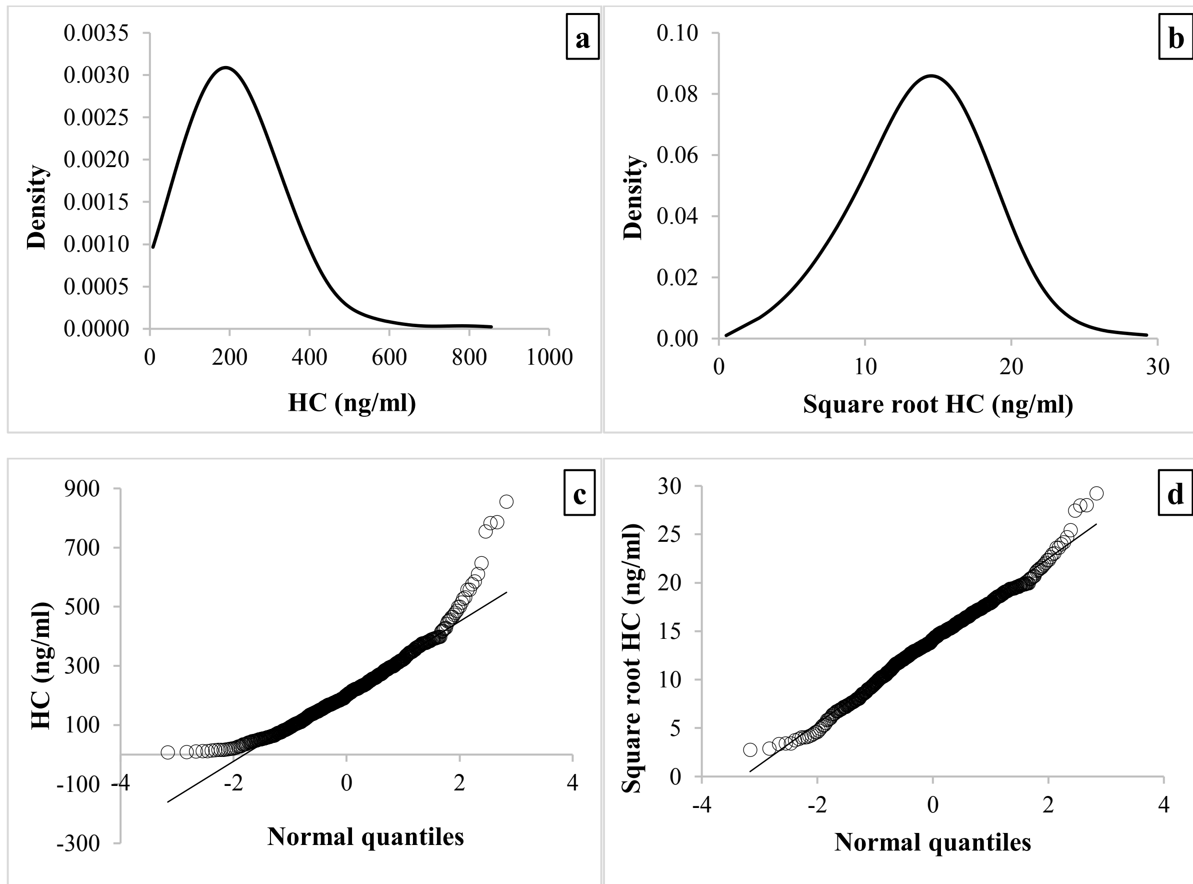


Figure 1. Density distribution and Q-Q plot of non-transformed values of cortisol concentration expressed as ng of cortisol per ml of serum (a and c) and of the square root of the ng of cortisol per ml of serum (b and d)

The burn-in period and thinning interval were chosen on the basis of Raftery and Lewis convergence diagnostic (1992). The Geweke convergence diagnostic based on Z-values criterion (Geweke 1992) was used to check convergence for each Markov chain. The samples included in the burn-in were not considered when the marginal posterior densities of parameters of concern were estimated and their features computed from the Gibbs samples.

Univariate genetic parameters. Variance components for mortality and survival time were estimated with the following univariate sire-dam model:

$$y_{ijk} = \mu + t_i + s_j + d_k + e_{ijk}$$

where y_{ijk} is the phenotypic record, μ is the model intercept, t_i is the fixed effect of the challenge test tank i, s_j is the random additive genetic effect of sire j, d_k is the random additive genetic effect of dam k, and e_i is the random residual effect.

Because mortality was a binary trait, we used a threshold (or probit) model to handle properly the dichotomous nature of the trait (Sorensen and Gianola 2002; Legarra et al. 2008). In threshold models, a latent unobservable variable, the liability, is modelled through the probit link function and the observed phenotype (0 or 1) depends on whether the underlying liability is lower or greater than a specific threshold. For identification purposes, the residual variance is set to 1 and the threshold is set to 0. In the Gibbs sampler for threshold models, the liability is taken into account as a nuisance parameter and integrated out in the Gibbs sampler. At each iteration, for each binary phenotype record

(0 or 1), a liability is generated below or beyond the threshold such that the observed phenotype is 0 or 1. The liability is sampled within the bounds -999 and +999.
For survival time, censoring was properly taken into account by the Bayesian analysis. In our study right-censoring occurred for all survived animals which had a survival time of 29 d. In the analysis, for these animals, the real unobserved phenotype was generated by sampling from a truncated normal distribution bounded at 29 d (the real phenotype of the survived animals cannot be less than 29 d) and defined by the effects fitted in the models and their variances. The censored nature of survival time is, then, properly treated by integrating out the conditional distribution of the censored data (Korsgaard et al. 2003).
Variance components for body weight, cortisol concentrations and antibody titer were estimated through the following univariate animal model:

$$y_i = \mu + a_i + e_i$$

where y_i is the phenotypic record, μ is the model intercept, a_i is the random additive genetic effect of the animal, and e_i is the random residual effect. Additive genetic relationships between animals used in this model were computed from the pedigree, which included all challenged animals, their parents and dummy grand-parents, assigned as sire or dam to the parents of the challenged fish in the pedigree reconstruction procedure.
In all Bayesian analyses, the *a priori* distribution for fixed effects was a uniform distribution, whereas the one for random effects was a normal distribution with zero mean and variance equal to σ_s^2, σ_d^2, $A\sigma_a^2$ and σ_e^2 for sire, dam, animal and residual effects, respectively, where A is the numerator relationship matrix between animals. Scaled inverted χ^2 distributions were assumed as prior distributions for the sire, dam, additive genetic and residual variance. The hyperparameters of scaled inverted χ^2 distributions (v and s^2) were equal to -2 and 0, making the prior distribution flat.
Heritability was then computed as $h^2 = [2(\sigma_s^2 + \sigma_d^2)]/\sigma_p^2$ (for sire-dam models, where σ_s^2 and σ_d^2 are the sire and dam components of the variance, respectively, σ_p^2 is the phenotypic variance) or as $h^2 = \sigma_a^2/\sigma_p^2$ (for animal models, where σ_a^2 is the additive genetic component of the variance and σ_p^2 is the phenotypic variance; Falconer and Mackay 1996).
The median of the marginal posterior density was used as a point estimate for variance components and heritability. Post-gibbs analyses on the posterior distribution of h^2 were performed through the R package BOA (Smith 2007) and the lower and upper bounds of the highest probability density interval (HPD 95%; Chen and Shao 1999) were obtained from the posterior distribution of h^2 as well as the probability of h^2 being greater than 0.1.
Bivariate genetic parameters. Bivariate Bayesian analyses were used to estimate (co)variance components and to investigate the relationships (genetic and phenotypic correlations) between the traits under consideration. The models were the same used in the univariate analyses. When mortality or survival time was one of the trait considered, a bivariate sire-dam model was used, otherwise a bivariate animal model was fitted to the data. Prior densities for the model effects were those described for univariate analyses. Inverse Wishart prior distributions, which are very vague priors, were used for all (co)variances.
Genetic (r_a) and phenotypic (r_p) correlations were computed as:

$$r_a = \frac{[2(cov_{s(1,2)} + cov_{d(1,2)})]}{\sqrt{[2(\sigma^2_{s(1)} + \sigma^2_{d(1)})][2(\sigma^2_{s(2)} + \sigma^2_{d(2)})]}}$$

$$r_p = \frac{(cov_{s(1,2)} + cov_{d(1,2)} + cov_{e(1,2)})}{\sqrt{\sigma^2_{p(1)}\sigma^2_{p(2)}}}$$

for sire-dam models, where $cov_{s(1,2)}$, $cov_{d(1,2)}$ and $cov_{e(1,2)}$ are the sire, the dam and the residual components of the covariance between trait 1 and 2, respectively, σ^2_s and σ^2_d are the sire and the dam components of the variance for each trait, respectively, σ^2_p is the phenotypic variance, or as

$$r_a = \frac{cov_{a(1,2)}}{\sqrt{\sigma^2_{a(1)}\sigma^2_{a(2)}}}$$

$$r_p = \frac{(cov_{a(1,2)} + cov_{e(1,2)})}{\sqrt{\sigma^2_{p(1)}\sigma^2_{p(2)}}}$$

for animal models, where $cov_{a(1,2)}$ and $cov_{e(1,2)}$ are the additive genetic and the residual components of the covariance between trait 1 and 2, respectively, σ^2_a is the additive genetic component of the variance for each trait, σ^2_p is the phenotypic variance.

As for univariate analyses, the median of the marginal posterior density was used as a point estimate for (co)variance components, h^2, r_a and r_p. R package BOA (Smith 2007) was used to perform post-gibbs analyses on the posterior distribution of h^2, r_a and r_p. The lower and upper bounds of the 95% highest probability density interval (HPD95%; Chen and Shao 1999) were obtained from the posterior distribution of h^2, r_a and r_p as well as the probability of a positive or negative correlation.

Genomic heritabilities

Genomic variance components and heritability for the different traits were estimated with Bayesian procedures using the R/BGLR software (Pérez and De Los Campos 2014) following the methodology described by de Los Campos et al. (2015). Five different regression models were implemented: BayesA (Meuwissen et al. 2001), BayesB, BayesC (Habier et al. 2011), Bayesian LASSO (Park and Casella 2008) and Bayesian Ridge Regression (Meuwissen et al. 2001). All models included an intercept and the effect of the genotype, coded as the number of copies of the MAF allele (0, 1 or 2), of the animal at 16,075 SNPs. Details on these models can be found in Pérez and De Los Campos (2014); the main differences between models are related to the prior distribution of the marker effects. Actually, the choice of the prior determines the type of shrinkage of estimates of the effects induced. Briefly, the scaled-t and double exponential densities are two types of prior with higher mass at zero and thicker tails than the normal distribution that induce size-of-effect dependent shrinkage (Gianola 2013). BayesA prior is the scaled-t density, which generates the shrinkage of the marker size effect and allows for variable marker effect sizes (Meuwissen et al. 2001), while the double-exponential (or Laplace) prior is the one used in the Bayesian LASSO (Park and Casella 2008).

In Bayes B and Bayes C, two finite mixture priors are implemented. In the former (BayesB), the prior is a mixture of a point of mass at zero and a scaled-t density, in the latter (BayesC) the prior is a mixture of a point of mass at zero and a Gaussian distribution (Habier et al. 2011). The priors used in

BayesB and BayesC can perform variable selection through the assignment of a nonnull prior probability for the marker effect to be equal to zero.
In the Bayesian Ridge Regression model, the Gaussian prior induces homogenous shrinkage across markers; in other words, all effects are shrunk to a similar extent (Meuwissen et al. 2001).
The median of the marginal posterior density was used as a point estimate for genomic variance components and genomic h^2; the lower and upper bounds of the highest probability density interval (HPD 95%; Chen and Shao 1999) were obtained from the posterior distribution of genomic h^2.

3.3 Results

3.3.1 NNV challenge test

Typical clinical signs related to NNV infection (Figure 2) were detected in each experimental tank starting from d 3 after injection. As indicated by the estimated Kaplan-Meier curve (Figure 3), the mortality reached its peak 4 d post-challenge, sharply decreased up to 9 d post-challenge and then decreased slowly. The survival rate at the end of the test was 52.2%.

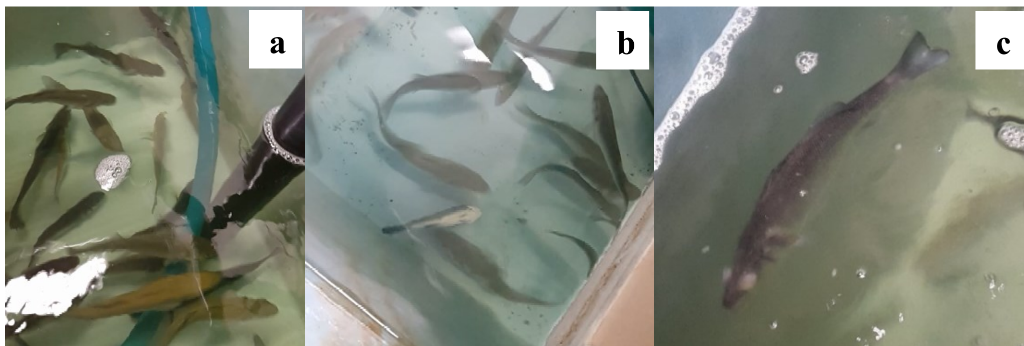


Figure 2. Clinical signs indicating the presence of neurological damages related to NNV infection in sea bass subjected to the experimental challenge test: a) changes in skin pigmentation; b) erratic swimming patterns; c) retinal necrosis

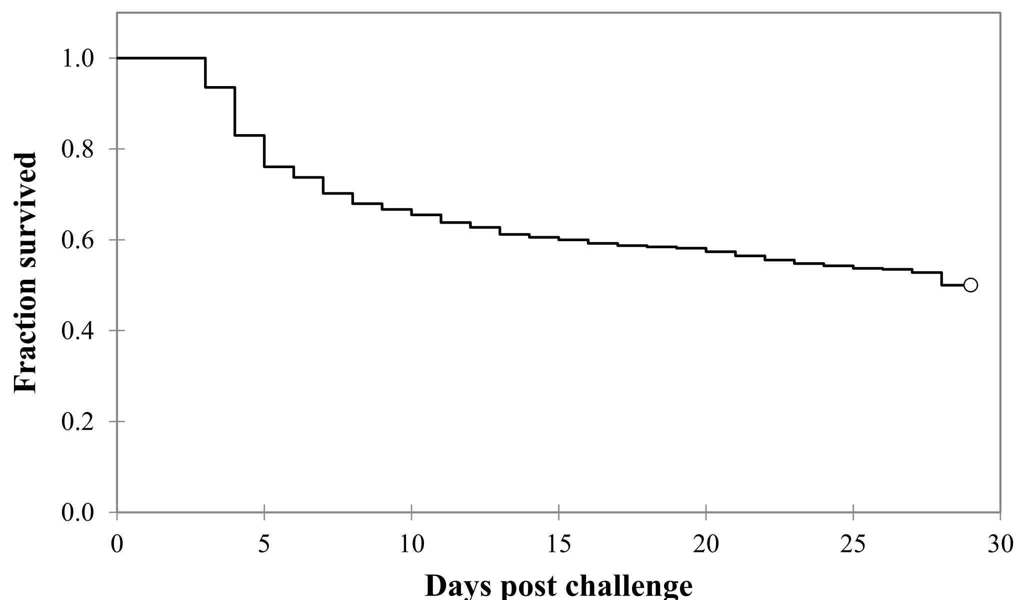


Figure 3. Estimated Kaplan-Meier curve describing survival of fish after the experimental infection with NNV

3.3.2 Parentage assignment and pedigree reconstruction

Parentage assignment to a unique parental pair was achieved for the 100% of the experimental fish. In addition, 41 parents were assigned to 22 dummy parental pairs (14 grandsires and 8 granddams), 3 to 3 single dummy parents (1 grandsires and 2 granddams) whereas 8 remained unassigned.

A total of 136 full-sib families were produced, with a number of offspring per family that varied from a minimum of 1 to a maximum of 17. Two males and one female of the parent group used to generate the experimental fish were not assigned to any individual in the challenged sample.

The number of offspring per sire ranged from 4 to 45, the number of offspring per dam varied from 10 to 68 (Figure 4).

		Dams																				
		D1	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**	**D11**	**D12**	**D13**	**D14**	**D15**	**D16**	**D17**	**D18**	**D19**	**D20**	**Total**
Sires	**M1**	4	5	3	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**18**
	M2	5	6	5	10	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**27**
	M3	3	3	5	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**15**
	M4	2	0	9	8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**20**
	M5	4	0	7	6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**18**
	M6	2	4	12	13	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**35**
	M7	8	5	17	14	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**45**
	M8	1	1	1	8	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**13**
	M9	0	0	0	0	0	6	1	6	1	0	0	0	0	0	0	0	0	0	0	0	**14**
	M10	0	0	0	0	0	8	4	3	1	0	0	0	0	0	0	0	0	0	0	0	**16**
	M11	0	0	0	0	0	9	9	13	9	0	0	0	0	0	0	0	0	0	0	0	**40**
	M12	0	0	0	0	0	8	1	6	8	0	0	0	0	0	0	0	0	0	0	0	**23**
	M13	0	0	0	0	0	14	6	8	13	0	0	0	0	0	0	0	0	0	0	0	**41**
	M14	0	0	0	0	0	12	3	11	8	0	0	0	0	0	0	0	0	0	0	0	**34**
	M15	0	0	0	0	0	6	2	12	8	0	0	0	0	0	0	0	0	0	0	0	**28**
	M16	0	0	0	0	0	3	3	4	3	0	0	0	0	0	0	0	0	0	0	0	**13**
	M17	0	0	0	0	0	0	0	0	0	1	0	12	4	4	0	0	0	0	0	0	**21**
	M18	0	0	0	0	0	0	0	0	0	1	0	5	2	2	0	0	0	0	0	0	**10**
	M19	0	0	0	0	0	0	0	0	0	3	0	7	3	2	0	0	0	0	0	0	**15**
	M20	0	0	0	0	0	0	0	0	0	2	3	5	3	2	0	0	0	0	0	0	**15**
	M21	0	0	0	0	0	0	0	0	0	2	2	1	0	0	0	0	0	0	0	0	**5**
	M22	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	**4**
	M23	0	0	0	0	0	0	0	0	0	2	4	13	9	15	0	0	0	0	0	0	**43**
	M24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	3	12	5	0	**24**
	M25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	2	6	3	0	**14**
	M26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	7	10	0	**23**
	M27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	8	2	2	0	**16**
	M28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	10	8	4	0	**25**
	M29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	8	3	7	5	0	**27**
	M30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	1	1	0	**8**
	M31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
	M32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
	Total	**29**	**24**	**59**	**68**	**11**	**66**	**29**	**63**	**51**	**12**	**10**	**44**	**22**	**25**	**11**	**20**	**33**	**43**	**30**	**0**	**650**

Figure 4. Heat-map showing the number of offspring per family, dam (in bold at the end of each column) and sire (in bold at the end of each line); darker colours correspond to greater numbers of individuals

Table 1. Descriptive statistics for body weight, cortisol concentration and antibody titer

Trait[1]	Mean	CV %	Min	Max
BW	146.42	29.18	35.00	318.00
HC	213.26	56.86	7.65	855.10
SRHC	13.97	30.34	2.76	29.24
AT	74.92	41.52	7.91	173.53

[1]BW: body weight at 548 d post-hatching (g); HC: serum cortisol concentration (ng/ml); SRHC: square root transformation of HC; AT: antibody titer: sample to positive ratio of the optical density values (OD 450 nm)

3.3.3 Statistical analyses

Descriptive statistics

Descriptive statistics for body weight (g) at 548 d post-hatching, cortisol concentration (ng of cortisol per ml of serum, non-transformed values), its square root transformation and antibody titer are presented in Table 1.

The body weight recorded in experimental fish at a constant age was variable (coefficient of variation, CV: 29.2%), ranging from a 35 to 318 g. The average level of cortisol assessed in the serum of fish after acute stress was high and highly variable across individuals (CV: 56.9%). Likewise, variation in antibody titer was large (CV: 41.5%). Besides normalisation of the frequency distribution, the square root transformation of cortisol concentrations decreased the variation in the data (CV: 30.3%).

Genetic parameters

Univariate genetic parameters. Point estimates and features of the marginal posterior densities of variance components and genetic heritability for the traits under investigation obtained in univariate analyses are presented in Table 2. The estimated marginal posterior densities of heritability for each trait are depicted in Figure 5.

Survival time after infection was the trait showing the lowest heritability. The heritability estimate for mortality was almost twice as high as the one for survival time and slightly lower than the estimate for cortisol concentration. The probability that such estimates are greater than 0.1, which is used as a critical threshold to identify a low heritability trait, is 0.7, 0.36 and 0.93 for mortality, survival time and cortisol concentration, respectively.

The estimated heritability for antibody titer is moderate (0.36), whereas the one for body weight is high (0.57). For both these estimates, the probability of being greater than the critical value of 0.1 is equal to one.

The 95% highest posterior density intervals of the estimated heritabilities, which reflect the uncertainty in the estimated parameter and are the narrowest intervals including parameter values with higher probability than values outside the interval, are heterogeneous. They are wider for body weight and antibody titer than for mortality, survival time and cortisol concentration.

Bivariate genetic parameters. Table 3 reports point estimates and features of the marginal posterior densities for variance components, heritability, genetic and phenotypic correlations estimated in bivariate analyses. The estimated marginal posterior densities of genetic and phenotypic correlations between traits are presented in Figure 6.

In general, the estimates of heritability of mortality and survival time are consistently greater than those from univariate models, whereas the estimates for body weight, cortisol concentration and antibody titer are similar to those provided by univariate analyses.

The phenotypic relationships between the investigated traits are weak and the estimated correlations are consistently close to 0 with the only exception of the point estimate of the phenotypic correlation

between antibody titer and body weight (0.24) which has a probability of being greater than 0.1 higher than 0.95.
The uncertainty in the estimation process of the genetic correlations is large and is reflected in the dispersion exhibited by the estimated marginal densities of these parameters (Figure 6) and in the size of the 95% highest posterior density intervals. However, the genetic antagonism or synergy between traits is detectable based on the estimated probability of a negative or a positive genetic correlation.
The genetic correlations between mortality and body weight or antibody titer are both moderately negative (-0.39), with a probability of being lower than 0 equal to 0.88, whereas the one between mortality and cortisol concentration is trivial and not statistically different from zero.
Survival time after the infection exhibits no genetic relationships with the other traits, with the only exception of body weight for which the estimated correlation has a probability of being positive which is slightly greater than 0.7.
Cortisol concentration and antibody titer are moderately genetically correlated (0.27), with a probability of the values being positive equal to 0.81. The genetic correlations between body weight and cortisol concentration and between body weight and antibody titer are both positive, with a low value in the first case (0.11), a fairly high value in the second case (0.49).

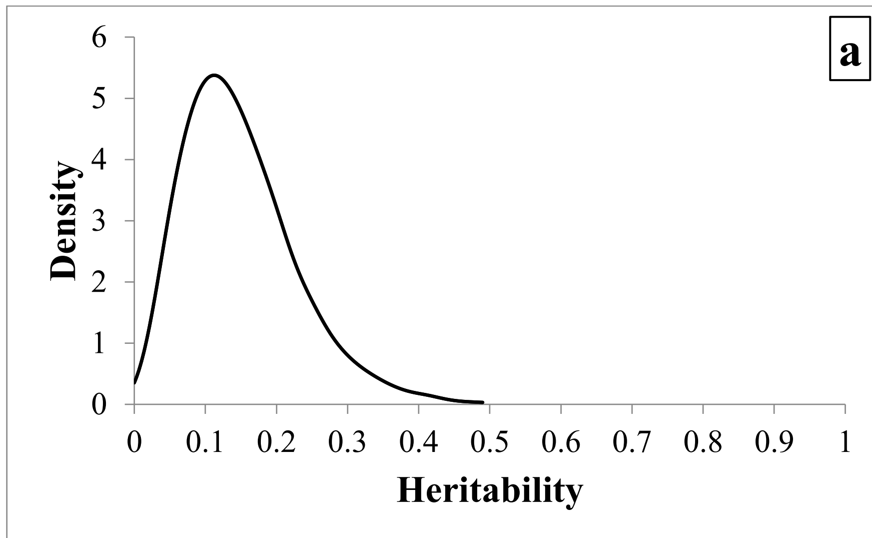
Genomic heritabilities

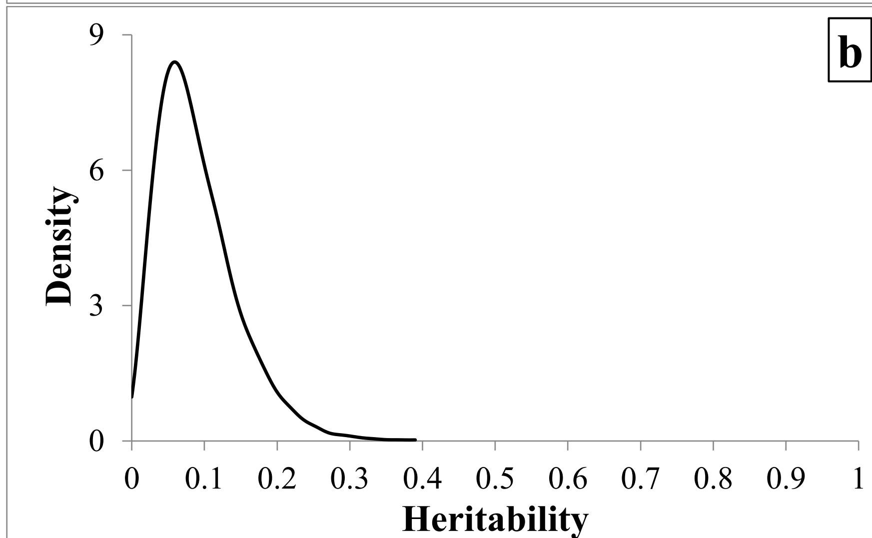
Point estimates and features of the marginal posterior densities for variance components and genomic heritability estimated with five Bayesian models for the traits under investigation are presented in Table 4. The estimated marginal posterior densities of the genomic heritability obtained with the different models, are depicted in Figure 7.
The estimates obtained are consistent throughout the five models used. For mortality, survival time, body weight and antibody titer, the estimated genomic heritability is lower than the estimated genetic heritability (0.06, 0.04, 0.45 and 0.26, respectively, compared to 0.14, 0.08, 0.57 and 0.36), while for cortisol concentration the genomic heritability estimate is consistent with the genetic heritability estimate (0.19 vs. 0.18).

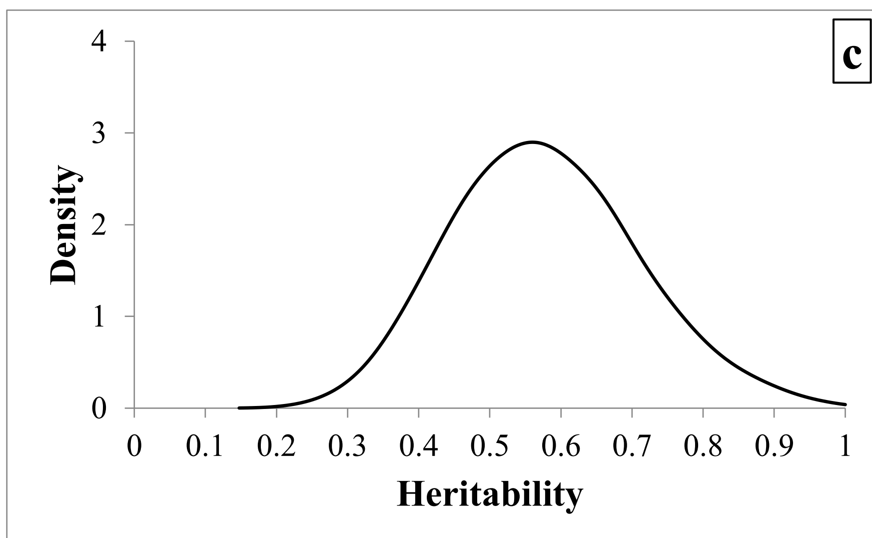
Table 2. Point estimates (median of the marginal posterior density) of variance components and genetic heritability (h^2), posterior probability of $h^2 > 0.1$ and 95% highest posterior density interval of h^2 (HPD95%) for post-challenge mortality (MORT) and survival time (SURVT), body weight at 548 d (BW), square root of serum cortisol concentration (SRHC) and antibody titer (AT)[1]

Trait[2]	σ_a^2	σ_p^2	h^2	$P(h^2 > 0.1)$	HPD95%
MORT	0.146	1.073	0.136	0.697	0.016, 0.307
SURVT	29.866	379.276	0.079	0.359	0.005, 0.198
BW	1198.250	2107.480	0.571	1.000	0.331, 0.838
SRHC	3.421	18.434	0.189	0.933	0.065, 0.342
AT	371.944	1049.240	0.357	0.999	0.163, 0.592

[1]σ_a^2: additive genetic variance; σ_p^2: phenotypic variance; variance components were estimated with a univariate sire-dam probit model for MORT, a univariate linear sire-dam model accounting for censoring for SURVT and univariate animal models for BW, SRHC and AT
[2]MORT: binary trait (0: alive, 1: dead); SURVT: right-censored trait.


a
Density
Heritability
0
1
2
3
4
5
6
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


b
Density
Heritability
0
3
6
9
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


c
Density
Heritability
0
1
2
3
4
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

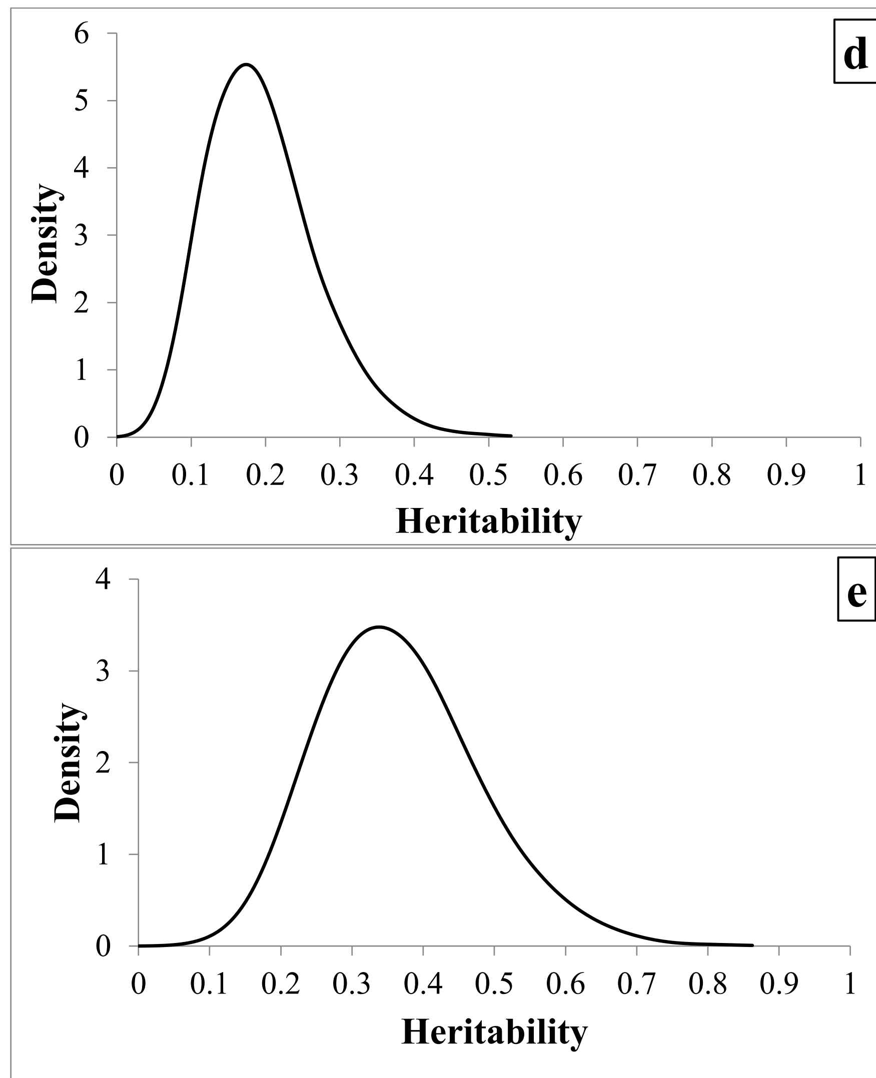


Figure 5. Estimated marginal posterior density of genetic heritability for a) VNN mortality, b) survival time, c) body weight, d) serum cortisol concentration and e) antibody titer against NNV[1,2]

[1]VNN mortality: binary trait (0: alive, 1: dead); survival time: right-censored trait; body weight (g) at 548 d post-hatching; serum cortisol concentration: square root of ng of cortisol per ml of serum; antibody titer: sample to positive ratio of the optical density values (OD 450 nm)

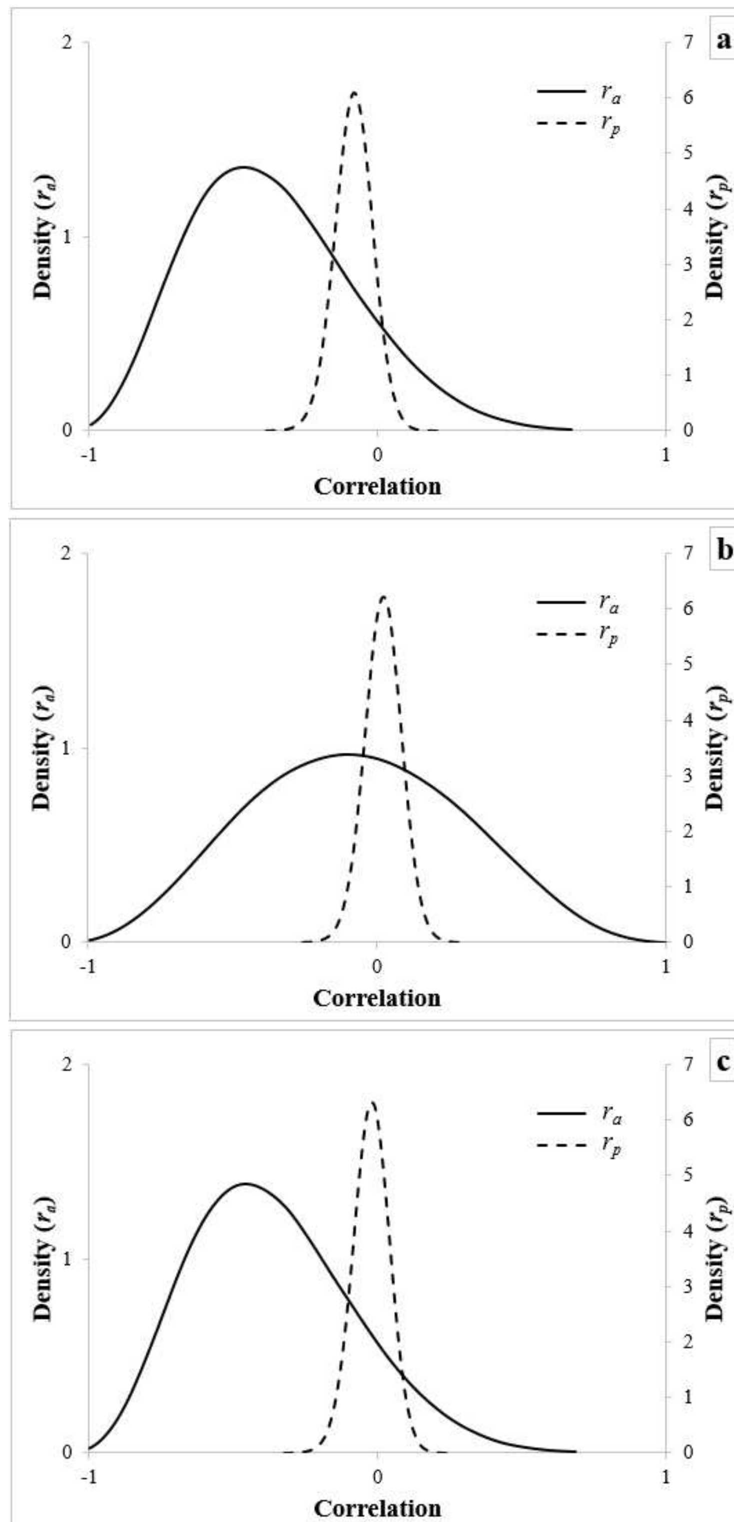
[2]Variance components were estimated with a univariate sire-dam probit model for mortality, a univariate linear sire-dam model accounting for censoring for survival time and univariate animal models for body weight, cortisol concentration and antibody titer

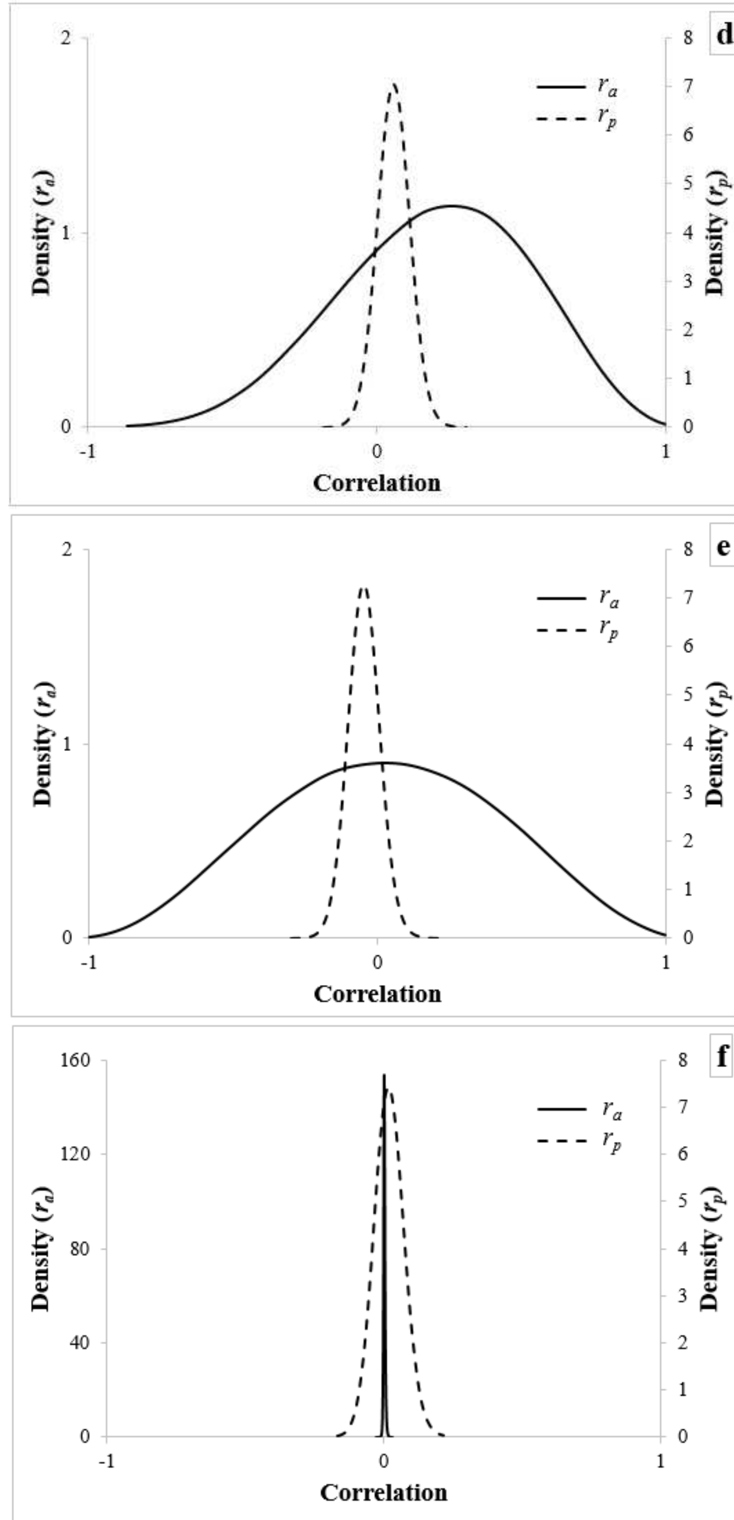
Table 3. Point estimates (median of the marginal posterior density) of additive genetic (σ_a^2) and phenotypic (σ_p^2) variance, genetic heritability (h^2), phenotypic (r_p) and genetic (r_a) correlations, 95% highest posterior density intervals (HPD95%) and probability (P) of a positive (for positive estimates) or negative (for negative estimates) genetic correlation obtained in bivariate analyses of post-challenge mortality and survival time, body weight at 548 d, square root of serum cortisol concentration and antibody titer against NVV[1]

Bivariate analysis[1,2]		Trait 1			Trait 2			Correlation		
Trait 1	**Trait 2**	σ_a^2	σ_p^2	h^2 (HPD95%)	σ_a^2	σ_p^2	h^2 (HPD95%)	r_p (HPD95%)	r_a (HPD95%)	P
Mortality	Body weight	0.2577	1.1288	0.2282 (0.0568, 0.4683)	878.5466	1953.3184	0.4512 (0.2448, 0.7020)	-0.0845 (-0.2153, 0.0440)	-0.3883 (-0.8777, 0.1984)	0.8840
Mortality	SRHC	0.2508	1.1254	0.2229 (0.0576, 0.4549)	4.5111	19.1939	0.2369 (0.0780, 0.4491)	0.0200 (-0.1065, 0.1475)	-0.0783 (-0.7461, 0.6247)	0.5778
Mortality	Antibody titer	0.2676	1.1338	0.2360 (0.0626, 0.4674)	335.5092	1030.4830	0.3276 (0.1583, 0.5391)	-0.0221 (-0.1487, 0.0994)	-0.3880 (-0.8657, 0.1909)	0.8864
Survival time	Body weight	56.3288	400.5311	0.1416 (0.0277, 0.3155)	868.3167	1942.3640	0.4495 (0.2423, 0.6956)	0.0587 (-0.0528, 0.1708)	0.2126 (-0.4302, 0.7992)	0.7266
Survival time	SRHC	55.9763	399.7729	0.1408 (0.0295, 0.3120)	4.5326	19.1417	0.2382 (0.0811, 0.4507)	-0.0474 (-0.1551, 0.0625)	0.0247 (-0.6991, 0.7475)	0.5230
Survival time	Antibody titer	59.1062	400.8923	0.1485 (0.0335, 0.3210)	282.9224	1027.7840	0.2766 (0.1206, 0.4783)	0.0190 (-0.0871, 0.1250)	0.0033 (-0.0027, 0.0106)	0.8957
SRHC	Antibody titer	3.6237	18.5562	0.1961 (0.0681, 0.3656)	424.1179	1078.5800	0.3941 (0.1890, 0.6348)	-0.0352 (-0.1362, 0.0702)	0.2676 (-0.2746, 0.7731)	0.8148
SRHC	Body weight	3.5652	18.5523	0.1931 (0.0662, 0.3579)	1267.9093	2147.7850	0.5917 (0.3543, 0.8513)	-0.0226 (-0.1343, 0.0899)	0.1151 (-0.4196, 0.6312)	0.6519
Antibody titer	Body weight	409.7458	1068.7879	0.3844 (0.1752, 0.6312)	1290.3643	2159.9085	0.5987 (0.3608, 0.8589)	0.2485 (0.1336, 0.3659)	0.4904 (0.0854, 0.8138)	1.0000

[1]Mortality: binary trait (0: alive, 1: dead); survival time: right-censored trait; body weight (g) at 548 d post-hatching; SRHC: square root of the ng of cortisol per ml of serum; antibody titer: sample to positive ratio of the optical density values (OD 450 nm);

[2](Co)variance components were estimated with a) bivariate sire-dam models when mortality or survival time was one of the analysed traits or b) bivariate animal models when mortality and survival time were not considered. For survival time, right censoring was properly accounted for in the model and a probit model was used for mortality.


a
2
1
0
7
6
5
4
3
2
1
0
-1
0
1
r_a
r_p
Density (r_a)
Density (r_p)
Correlation
b
2
1
0
7
6
5
4
3
2
1
0
-1
0
1
r_a
r_p
Density (r_a)
Density (r_p)
Correlation
c
2
1
0
7
6
5
4
3
2
1
0
-1
0
1
r_a
r_p
Density (r_a)
Density (r_p)
Correlation


d
r_a
r_p
Density (r_a)
Density (r_p)
Correlation
e
r_a
r_p
Density (r_a)
Density (r_p)
Correlation
f
r_a
r_p
Density (r_a)
Density (r_p)
Correlation

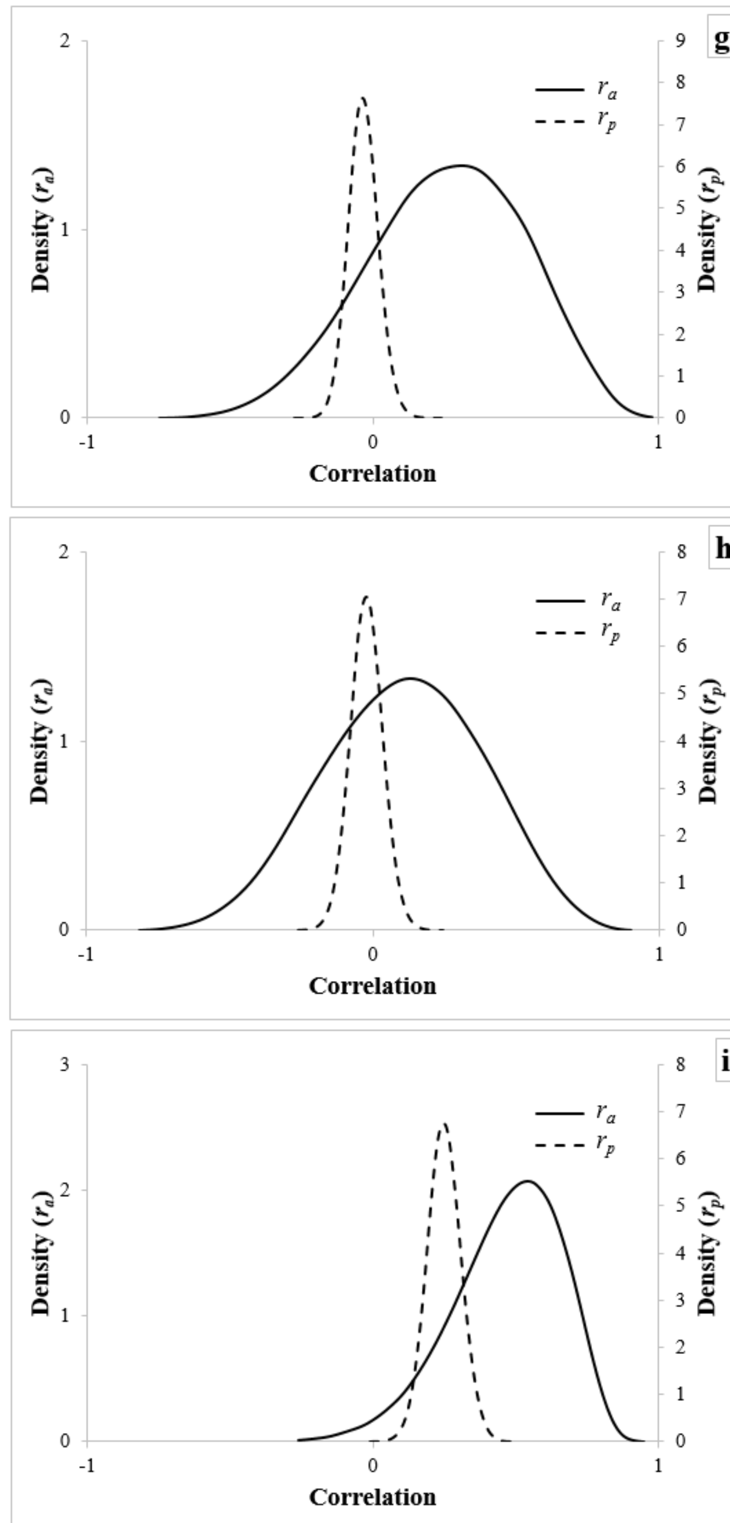


Figure 6. Estimated marginal posterior density of genetic (r_a) and phenotypic (r_p) correlations for a) mortality and body weight, b) mortality and cortisol concentration, c) mortality and antibody titer, d) survival time and body weight, e) survival time and cortisol concentration, f) survival time and antibody titer, g) cortisol concentration and antibody titer, h) cortisol concentration and body weight, i) antibody titer and body weight[1,2]

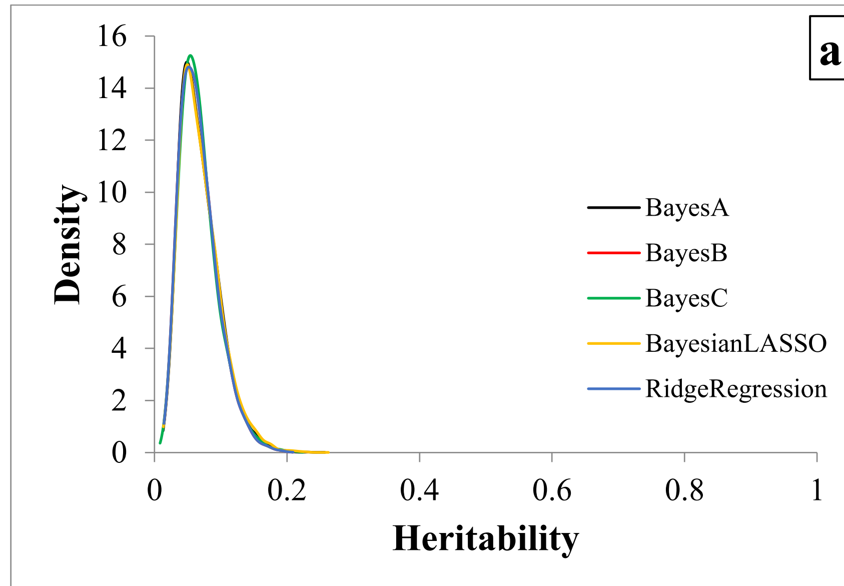
[1]Mortality: binary trait (0: alive, 1: dead); survival time: right-censored trait; body weight (g) at 548 d post-hatching; serum cortisol concentration: square root of ng of cortisol per ml of serum; antibody titer: sample to positive ratio of the optical density values (OD 450 nm)

[2](Co)variance components were estimated with a) bivariate sire-dam models when mortality or survival time was one of the analysed traits or b) bivariate animal models when mortality and survival time were not considered. For survival time, right censoring was properly accounted in the model and a probit model was used for mortality

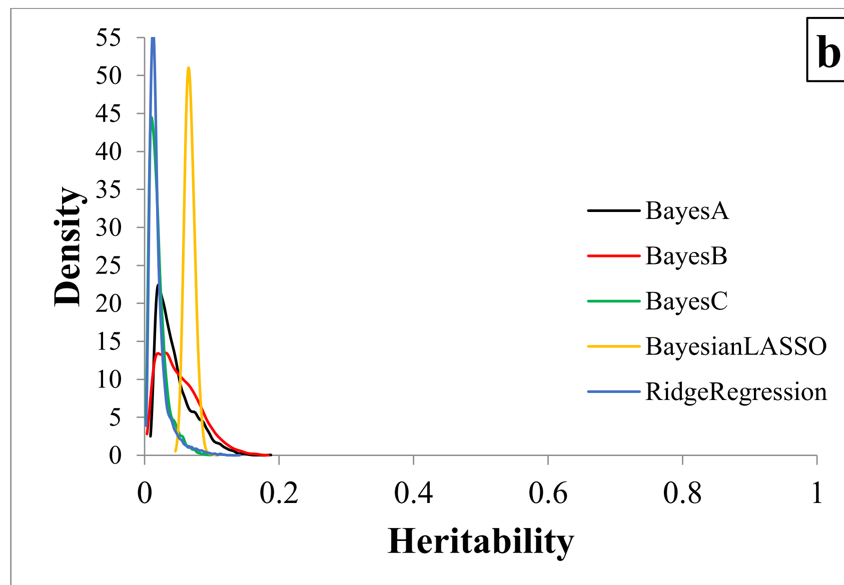
Table 4. Point estimates (median of the marginal posterior density) of genomic variance components and genomic heritability for post-challenge mortality (MORT) and survival time (SURVT), body weight at 548 d (BW), square root of serum cortisol concentration (SRHC) and antibody titer (AT) and 95% highest posterior density interval of genomic heritability (HPD95%)[1,2]

TRAIT	BayesA				BayesB				BayesC				Bayesian LASSO				Ridge Regression			
	σ_a^2	σ_p^2	h^2	HPD 95%	σ_a^2	σ_p^2	h^2	HPD 95%	σ_a^2	σ_p^2	h^2	HPD 95%	σ_a^2	σ_p^2	h^2	HPD 95%	σ_a^2	σ_p^2	h^2	HPD 95%
Mortality	0.07	1.07	0.06	0.02, 0.13	0.07	1.07	0.06	0.02, 0.13	0.07	1.07	0.06	0.02, 0.13	0.07	1.07	0.06	0.02, 0.13	0.07	1.07	0.06	0.02, 0.13
Survival time	14.78	396.98	0.04	0.01, 0.10	18.46	399.03	0.05	0.01, 0.11	6.14	393.53	0.02	0.00, 0.05	26.42	395.69	0.07	0.05, 0.08	5.97	393.79	0.02	0.00, 0.05
Body weight	802.96	1805.00	0.44	0.33, 0.55	805.20	1806.23	0.45	0.34, 0.56	803.20	1806.46	0.44	0.34, 0.55	842.10	1819.93	0.46	0.42, 0.50	814.96	1809.46	0.45	0.35, 0.55
SRHC	3.55	18.05	0.20	0.12, 0.29	3.43	17.99	0.19	0.09, 0.29	3.44	18.02	0.19	0.10, 0.28	3.43	18.01	0.19	0.12, 0.28	3.44	18.02	0.19	0.11, 0.28
Antibody titer	258.36	973.67	0.27	0.16, 0.36	236.70	968.03	0.24	0.15, 0.35	260.66	974.77	0.27	0.18, 0.37	265.43	979.08	0.27	0.17, 0.36	279.09	989.43	0.26	0.16, 0.37

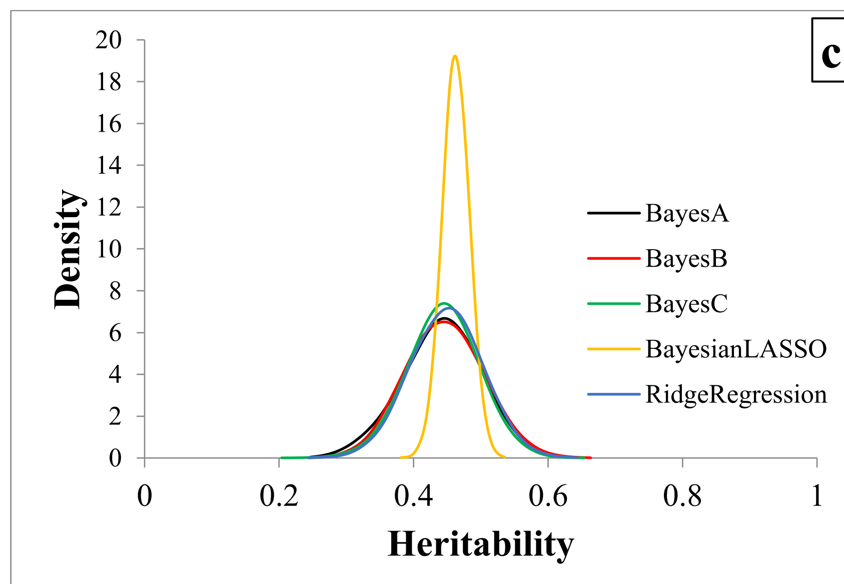
[1]σ_a^2 = additive genomic variance; σ_p^2 = phenotypic variance; h^2 = genomic heritability
[2]Mortality: binary trait (0: alive, 1: dead); survival time: right-censored trait; body weight (g) at 548 d post-hatching; SRHC: square root of the ng of cortisol per ml of serum; antibody titer: sample to positive ratio of the optical density values (OD 450 nm)



a
Density
16
14
12
10
8
6
4
2
0
0
0.2
0.4
0.6
0.8
1
Heritability
BayesA
BayesB
BayesC
BayesianLASSO
RidgeRegression



b
Density
55
50
45
40
35
30
25
20
15
10
5
0
0
0.2
0.4
0.6
0.8
1
Heritability
BayesA
BayesB
BayesC
BayesianLASSO
RidgeRegression



c
Density
20
18
16
14
12
10
8
6
4
2
0
0
0.2
0.4
0.6
0.8
1
Heritability
BayesA
BayesB
BayesC
BayesianLASSO
RidgeRegression

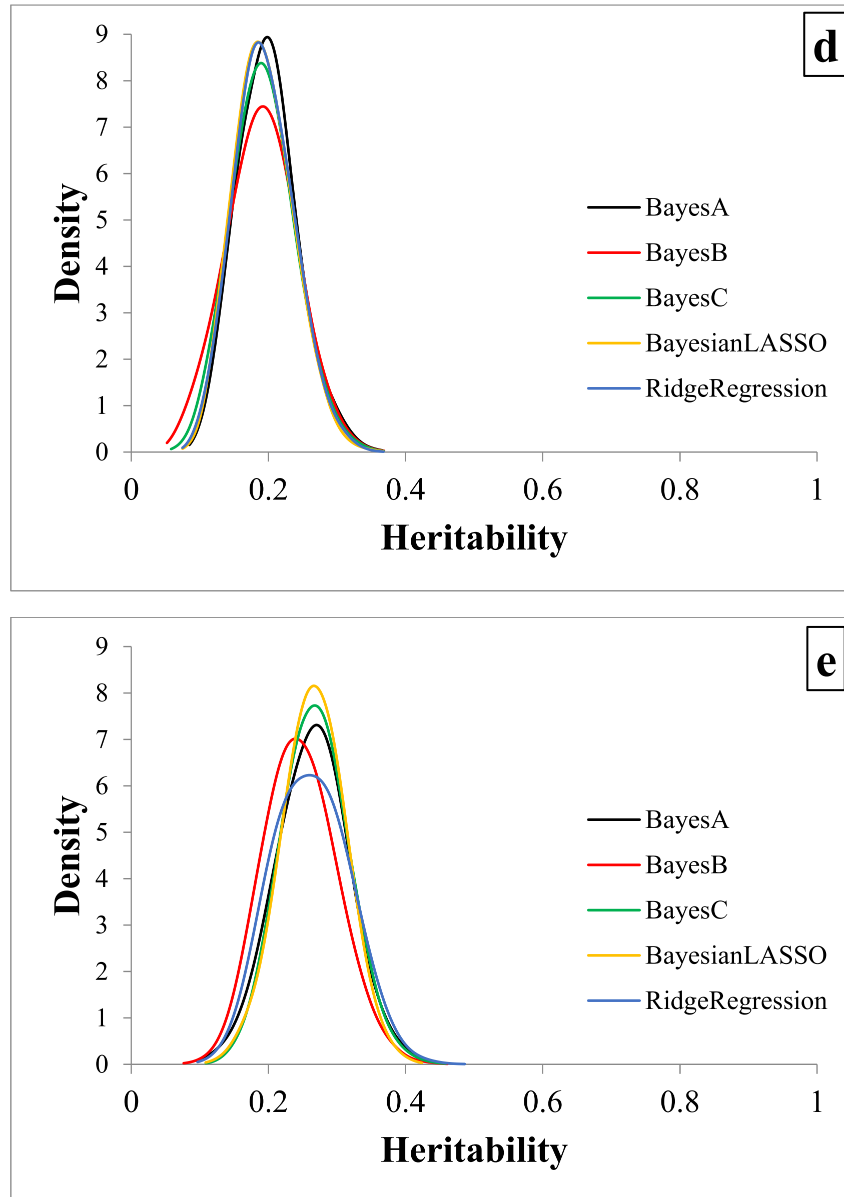


Figure 7. Estimated marginal posterior density of genomic heritability for a) VNN mortality, b) survival time, c) body weight, d) serum cortisol concentration, d) antibody titer against NNV, obtained with five Bayesian models[1,2]

[1]Models were BayesA, BayesB, BayesC, Bayesian LASSO, Bayesian Ridge Regression.
[2]Mortality: binary trait (0: alive, 1: dead); survival time: right-censored data; body weight (g) at 548 d post-hatching; serum cortisol concentration: square root of ng of cortisol per ml of serum; antibody titer: sample to positive ratio of the optical density values (OD 450 nm)

3.4 Discussion

The interest for selective breeding as a tool to enhance genetic resistance to VNN in European sea bass has markedly increased during recent years, because of the major threat represented by NNV for sea bass hatcheries and aquaculture farms combined with the lack of simply exploitable vaccines or effective therapeutical remedies specific for this disease (Doan et al. 2016; Buonocore et al. 2017).
The existence of a genetic basis for resistance to various diseases in fish species, evidenced by moderate to high heritability estimates, suggests the feasibility of such approach as a potential disease preventive measure (Ødegård et al. 2011; Doan et al. 2016). For some fish species (mainly salmonids), the genetic selection of strains showing improved resistance is well established and supported by the results from commercial selection experiments (Dorson et al. 1995; Storset et al. 2007).
Alternatively, selective breeding for improved disease resistance may be performed exploiting indirect approaches based on physiological (e.g., post-stress cortisol levels) or immunological (e.g., antibody titer) indicator traits, showing non-trivial genetic correlations with resistance and significant additive genetic variation.
This study investigated the genetic parameters for mortality after a VNN challenge test performed in a sea bass experimental population produced using breeding animals belonging to a commercial stock and the genetic relationships of resistance with body weight and cortisol concentration after stress exposure in the search of indicator traits.
Moreover, this study presented, for the first time, estimates of heritability for antibody titer against NNV in sea bass and of genetic correlations between this trait and resistance to VNN. Due to great difficulties in performing individual phenotyping of these traits and "traditional" selection procedures, we implemented a high-throughput sequencing approach (2b-RAD) to generate a genome-wide SNP panel in the view of exploring a genomic selection approach.
The brand-new result came from the heritability of antibody titer against NNV, with moderate to high genetic and genomic values (0.26-0.39). These estimates were higher compared to other fish species (0.2-0.18 in salmonids; Strømsheim et al. 1994; Fjalestad et al. 1996) and indicated the presence of a genetic basis for this trait in sea bass.
Antibody titer was not genetically nor phenotipically correlated with the survival time recorded in the challenge test, but it exhibited a negative genetic relationship with mortality, indicating that selective breeding based on the genetic merit for antibody titer should result in a favorable correlated response in the resistance to VNN.
The heritability of mortality estimated through a univariate sire-dam model was low, but when it was estimated through bivariate models, the estimates were higher and comparable to recent estimates reported by other authors on larger datasets of individuals (Doan et al. 2017; Palaiokostas et al. 2018).
The genomic heritability estimate obtained for mortality was lower than the one estimated using a polygenic sire-dam model, indicating that SNP effects were not able to pick up all the variation due to additive gene effects and due to the presence of the "missing heritability" phenomenon as described by De Los Campos et al. (2015). Palaiokostas et al. (2018) used a genomic relationship matrix to estimate heritability and achieved a value of 0.43, but the linkage disequilibrium derived from selective breeding could be a reason for such an overrating of additive genetic variance.
Our estimate of heritability for survival time was greatly lower than the estimates obtained for Atlantic cod (*Gadus morhua*) challenged with NNV (0.81; Bangera et al. 2014), but in agreement with the estimated heritability for survival time in other species challenged with different viral pathogens, such as rainbow trout (*Oncorhynchus mykiss*) challenged with viral hemorrhagic

septicaemia virus (0.11; Henryon et al. 2005) and Atlantic salmon (*Salmo salar*) challenged with infectious pancreatic necrosis disease virus (0.16; Guy et al. 2009).

Mortality and body weight exhibited a significant negative genetic relationship. These results imply that current selection for increased body weight at a fixed age or daily gain, which are major traits in the selection goal of sea bass breeding strains, should have a favorable impact on VNN resistance. Our results are not consistent with those by Doan et al. (2017) who reported an unfavorable genetic correlation between body weight at 180 d post-hatching and VNN resistance. Such inconsistencies may be due to the different sizes of infected fish, the quantity of viral copies injected, the survival rate after infection and the use of different statistical methodologies when estimating (co)variance components. In comparison to our study, Doan et al. (2017) infected fish of smaller size, had a survival rate greater than 70% at the end of the challenge test and estimated the genetic correlation using a linear model for survival in place of a probit model, which is the method of choice for binary traits. Doan et al. argued that the estimated genetic correlation had high standard errors indicating large uncertainty in the estimation process. In our study, the probability of the estimated genetic relationship between mortality and body weight at 548 d post-hatching being negative was close to 0.9.

Post-stress cortisol concentration showed a significant genetic variability, with estimated values in the range of the previous estimates for sea bass (Volckaert et al. 2012; Vandeputte et al. 2016). The genetic and phenotypic correlations between post-stress cortisol concentration and mortality or survival time were not significant. This is an interesting and new result for sea bass, which is consistent with what has observed in cod for vibriosis resistance (Kettunen Præbel et al. 2007). Other experimental studies, focused on different species (rainbow trout, salmon and carp; Fevolden et al. 1992; Fevolden and Røed, 1993; Fevolden et al. 1993; Yin et al. 1995), provided inconsistent results about performance of fish selected for low cortisol concentration and challenged with pathogen experimental infections. Based on the results of this study, the effectiveness of selection based on the genetic merit for post-stress cortisol concentration as an indirect strategy to enhance VNN disease resistance in sea bass seems to be questionable.

As expected, genetic and genomic heritability estimates for body weight were high and consistent with recent literature estimates (0.64, Massault et al. 2010; 0.29-0.62, Dupont-Nivet et al. 2010; 0.54, Volckaert et al. 2012; 0.62, Vandeputte et al. 2016). Age (548 d post-hatching) and size ($\bar{x} \pm s$: 146 ± 43 g) of the experimental fish at the infection time were, in our study, rather unusual compared to previous investigations on sea bass (Doan et al. 2017; Palaiokostas et al. 2018) or other species (Atlantic cod, Bangera et al. 2013; rainbow trout, Henryon et al. 2005; Vallejo et al. 2014; Atlantic salmon, Guy et al. 2009). Commonly, challenge tests are performed using young juveniles (age: 200 d post-hatching, weight: 10 to 20 g). Viral nervous necrosis affects sea bass larvae and juveniles with very high infectivity and mortality rates that may reach 100% (Munday et al. 2002), but our study confirms that VNN might have significant consequences even in advanced juveniles, which were previously reported only in one paper about a VNN natural outbreaks (Le Breton et al. 1997).

The use of smaller fish in challenge test is preferred because it allows the infection through the immersion of the animals in water containing the pathogen rather than the injection of a determined quantity of pathogen suspension. In experimental trials using immersion, the conditions are much more similar to natural infections, where pathogen must cross the first protective systems of the host (physical or mechanical barriers, epithelium and mucus), penetrate the host cell barriers in adequate number, attack target cells and replicate. Conversely, when infection is realized through injection, the trial is less representative in terms of natural disease outbreaks, but advantages are represented by a

systematic infection due to an equal level of exposure to the virus for all individuals and greater effectiveness when large sized individuals are tested (Robinson et al. 2017).
The choice of the method, however, is largely arbitrary. Infection through intraperitoneal injection is fairly common across studies, even when the size of the tested animals is small (e.g. Doan et al. 2017 and Bangera et al. 2013). Only in one study on VNN resistance in sea bass (Palaiokostas et al. 2018) animals were infected through immersion.
We can hypothesize that the use of different methods to infect the animals may lead to different results, particularly when the focus is on the investigation of the genetic variation in resistance. Comparison of results from studies where resistance traits are measured after infection carried out with different modalities (immersion *vs*. injection) is not straightforward because, in case of injection, the mechanism of resistance is considered only partially, taking into account only the components of the non-specific immune system related to cells, proteins or responses (granulocytes and macrophages, complement system, inflammation) and the specific immunity (lymphocytes, antibodies), in fact disregarding external barriers to infection (skin, mucous membranes). Considering only a part of the whole phenomenon may lead to a reduction of the variability that, in turn, has an impact on the heritability estimates. Actually, in the case of immersion, physical and mechanical barriers may play a significant role in preventing the penetration of the virus and contribute additional variation to the measured trait.

3.5 Conclusions

Our results suggest the feasibility of selective breeding approaches to improve the investigated traits in European sea bass, evidenced by significant heritability estimates for mortality, body weight, cortisol concentration and antibody titer, and indicating the presence of a genetic basis for these traits. In particular, selective breeding may be an effective approach to be integrated in the pool of VNN prevention tools for farmed sea bass.
Antibody titer against NNV offers the opportunity of indirect selection approaches to improve VNN resistance through the exploitation of a non-null genetic relationship with mortality. However, in a traditional selective breeding program, both direct and indirect selection would imply the need of individual phenotyping for mortality and antibody titer, which would be difficult or unfeasible to implement as a routine. Such difficulties can be overcome when developing genomic predictions of genetic merit for the traits of concern. In this view, the effectiveness of genomic tools, like the one used in this study to estimate heritabilities, in predicting the genetic merit for VNN mortality and antibody titer deserves specific investigations. Future studies are also required to elucidate the effects of different methods currently used in challenge tests to perform infection (immersion, injection) on the observed variation, and its genetic component for VNN resistance.

References

Bangera R, Ødegård J, Nielsen HM, Gjøen HM, Mortensen A (2013). Genetic analysis of vibriosis and viral nervous necrosis resistance in Atlantic cod (*Gadus morhua* L.) using a cure model. *J Anim Sci* **91**: 3574-3582.

Bangera R, Baranski M, Lien S (2014). A genome-wide association study for resistance to viral nervous necrosis in Atlantic cod using a 12K single nucleotide polymorphism array. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*, August 17th-22th, Vancouver, BC, Canada.

Bertotto D, Poltronieri C, Negrato E, Majolini D, Radaelli G, Simontacchi C (2010). Alternative matrices for cortisol measurement in fish. *Aquacult Res* **41**: 1261-1267.

Bovo G, Nishizawa T, Maltese C, Borghesan F, Mutinelli F, Montesi F, De Mas S (1999). Viral encephalopathy and retinopathy of farmed marine fish species in Italy. *Virus Res* **63**: 143-146.

Buonocore F, Nuñez-Ortiz N, Picchietti S, Randelli E, Stocchi V, Guerra L et al. (2017). Vaccination and immune responses of European sea bass (*Dicentrarchus labrax* L.) against betanodavirus. *Fish Shellfish Immunol* https://doi.org/10.1016/j.fsi.2017.11.039.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013). Stacks: an analysis tool set for population genomics. *Mol Ecol* **22**: 3124-3140.

Chen M-H, Shao Q-M (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *J Comput Graph Stat* **8**: 69-92.

Coeurdacier JL, Laporte F, Pepin JF (2003). Preliminary approach to find synthetic peptides from nodavirus capsid potentially protective against sea bass viral encephalopathy and retinopathy. *Fish Shellfish Immunol* **14**: 435-447.

Das S, Sahoo PK (2014). Markers for selection of disease resistance in fish: a review. *Aquacult Int* **22**: 1793-1812.

De Los Campos S, Sorensen D, Gianola D (2015). Genomic heritability: what is it? *PLoS Genet* **11**: e1005048. doi:10.1371/journal.pgen.1005048.

Doan QK, Vandeputte M, Chatain B, Morin T, Allal F (2016). Viral encephalopathy and retinopathy in aquaculture: a review. *J Fish Dis* **40**: 717-742.

Doan QK, Vandeputte M, Chatain B, Haffray P, Vergnet A, Breuil G, Allal F (2017). Genetic variation of resistance to Viral Nervous Necrosis and genetic correlations with production traits in wild populations of the European sea bass (*Dicentrarchus labrax*). *Aquaculture* **478**: 1-8.

Dorson M, Quillet E, Hollebecq MG, Torhy C, Chevassus B (1995). Selection of rainbow trout resistant to viral haemorrhagic septicaemia virus and transmission of resistance by gynogenesis. *Vet. Res.* **26:** 361-368.

Dupont-Nivet M, Karahan-Nomm B, Vergnet A, Merdy O, Haffray P, Chavanne H, Chatain B, Vandeputte M (2010). Genotype by environment interactions for growth in European seabass (*Dicentrarchus labrax*) are large when growth rate rather than weight is considered. *Aquaculture* **306**: 365-368.

Falconer DS, Mackay TFC (1996). Introduction to quantitative genetics. Longman, Harlow, England.

Fauvel C, Boryshpolets S, Cosson J, Wilson Leedy JG, Labbé C, Haffray P,Suquet M (2012). Improvement of chilled seabass sperm conservation using a cell culture medium. *J Appl Ichthyol* **28**: 961-966.

Fevolden SE, Refstie T, Røed KH (1992). Disease resistance in rainbow trout (*Oncorhynchus mykiss*) selected for stress response. *Aquaculture* **104**: 19-29.

Fevolden SE, Røed KH (1993). Cortisol and immune characteristics in rainbow trout (*Oncorhynchus mykiss*) selected for high or low tolerance to stress. *J Fish Biol* **43**: 919-930.

Fevolden SE, Nordmo R, Refstie T, Røed KH (1993). Disease resistance in Atlantic salmon (*Salmo salar*) selected for high or low responses to stress. *Aquaculture* **109**: 215-224.

Fjalestad KT, Larsen HJS, Røed KH (1996). Antibody response in Atlantic salmon (*Salmo salar*) against *Vibrio anguillarum* and *Vibrio salmonicida* O-antigens: heritabilities, genetic correlations and correlations with survival. *Aquaculture* **145**: 77-89.

Geweke J (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments (with discussion), pp. 169-193. In: *Bayesian statistics*, Bernardo JM, Berger JO, Dawid AP, Smith AF (Eds.) Oxford University Press (Oxford, UK).

Gianola D (2013). Priors in whole-genome regression: the Bayesian alphabet returns. *Genetics* **90**: 525-540.

Guy D, Bishop SC, Woolliams JA, Brotherstone S (2009). Genetic parameters for resistance to infectious pancreatic necrosis in pedigreed Atlantic salmon (*Salmo salar*) post-smolts using a Reduced Animal Model. *Aquaculture* **290:** 229-235.

Habier D, Fernando R, Kizilkaya K, Garrick D (2011). Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics* **12**: 186.

Henryon, Berg P, Olesen NJ, Kjær TE, Slierendrecht WJ, Jokumsen A, Lund I (2005). Selective breeding provides an approach to increase resistance of rainbow trout (*Onchorhynchus mykiss*) to the diseases, enteric redmouth disease, rainbow trout fry syndrome, and viral haemorrhagic septicaemia. *Aquaculture* **250**: 621- 636.

Huisman J (2017). Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. *Mol Ecol Res* **17**: 1009-1024.

Iwamoto T, Nakai T, Mori K, Arimoto M, Furusawa I (2000). Cloning of the fish cell line SSN- 1 for piscine nodaviruses. *Dis Aquat Organ* **43**: 81-89.

Kalinowski ST, Taper ML, Marshall TC (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* **16**: 1099-1106.

Kettunen Præbel A, Westgård JI, Peruzzi S, Fevolden SE (2007). Genetic parameters for post-stress cortisol activity and vibriosis resistance in Atlantic cod (*Gadus morhua* L.). *Aquaculture* **272**: S275-S276.

Korsgaard IR, Lund MS, Sorensen D, Gianola D, Madsen P, Jensen J (2003). Multivariate Bayesian analysis of Gaussian, right censored Gaussian, ordered categorical and binary traits using Gibbs sampling. *Genet Sel Evol* **35**: 159-183.

Le Breton A, Grisez L, Sweetman J, Ollevier F (1997). Viral nervous necrosis (VNN) associated with mass mortalities in cage-reared sea-bass, *Dicentrarchus labrax* (L.). *J Fish Dis* **20**: 145-151.

Legarra A, Varona L, López de Maturana E (2008). TM Threshold Model. http://snp.toulouse.inra.fr/~alegarra/

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* **7**: 639-655.

Massault C, Franch R, Haley C, de Koning DJ, Bovenhuis H, Pellizzari C, Patarnello T, Bargelloni L (2010). Quantitative trait loci for resistance to fish pasteurellosis in gilthead sea bream (*Sparus aurata*). *Animal Genet* **42**: 191-203.

Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Mol Ecol Resour* **15**: 28-41.

Meuwissen THE, Hayes BJ, Goddard ME (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**: 1819-1829.

Mori KI, Nakai T, Muroga K, Arimot M, Mushiake K, Furusawa I (1992). Properties of a new virus belonging to nodaviridae found in larval striped jack (*Pseudocaranx dentex*) with nervous necrosis. *Virology* **187**: 368-371.

Munday BL, Nakai T (1997). Nodaviruses as pathogens in larval and juvenile marine finfish. *World J Microbiol Biotechnol* **13**: 375.

Munday BL, Kwang J, Moody N (2002). Betanodavirus infections of teleost fish: a review. *J Fish Dis* **27**: 127-142.

Nuñez-Ortiz N, Pascoli F, Picchietti S, Buonocore F, Bernini C, Toson M, Scapigliati G, Toffan A (2016a). A formalin-inactivated immunogen against viral encephalopathy and retinopathy (VER) disease in European sea bass (*Dicentrarchus labrax*): immunological and protection effects. *Vet Res* **47**: 89.

Nuñez-Ortiz N, Stocchi V, Toffan A, Pascoli F, Sood N, Buonocore F, Picchietti S, Papeschi C, Taddei AR, Thompson KD, Scapigliati G (2016b). Quantitative immunoenzymatic detection of viral encephalopathy and retinopathy virus (betanodavirus) in sea bass *Dicentrarchus labrax*. *J Fish Dis* **39**: 821-831.

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

Palaiokostas C, Cariou S, Bestin A, Bruantn JS, Haffray P, Morin T, Cabon J, Allal F, Vandeputte M, Houston RD (2018). Genome-wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (*Dicentrarchus labrax*) using RAD sequencing. *Genet Sel Evol* **50**: 30.

Park T, Casella G (2008). The Bayesian Lasso. *J Am Stat Assoc* **103**: 681-686.

Pérez P, de Los Campos G (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **198**: 483-495.

Pottinger, TG (2008). The stress response in fish - mechanisms, effects and measurement, pp. 32-44. In: *Fish Welfare*, Branson EJ (Ed.), Blackwell Publishing Ltd, Oxford.

Raftery AE, Lewis SM (1992). How many iterations in the Gibbs Sampler? pp. 763-774. In: *Bayesian statistics*, Bernardo JM, Berger JO, Dawid AP, Smith AF (Eds.). Oxford University Press (Oxford, UK).

Robinson NA, Gjedrem T, Quillet E (2017). Testing for resistance: natural outbreaks versus controlled challenge testing, pp. 22-23. In: *Fish disease, prevention and control strategies*, Jeney G (Ed.) Academic press, Elsevier, London.

Sargolzaei M, Chesnais JP, Shenkel FS (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* **15**: 478-489.

Scapigliati G, Buonocore F, Randelli E, Casani D, Meloni S, Zarletti G, Tiberi M, Pietretti D, Boschi I, Manchado M, Martin-Antonio B, Jimenez-Cantizano R, Bovo G, Borghesan F, Lorenzen N, Einer-Jensen K, Adams A, Thompson K, Alonso C, Bejar J, Cano I, Borrego JJ, Alvarez MC (2010). Cellular and molecular immune responses of the sea bass (*Dicentrarchus labrax*) experimentally infected with betanodavirus. *Fish Shellfish Immunol* **28**: 303-311.

Simontacchi C, Bongioni G, Ferasin L, Bono G (1995). Messa a punto di un metodo RIA su micropiastra per il dosaggio diretto del progestrerone ematico. Atti XLIX Convegno Nazionale S.I.S.Vet., pp. 343-344.

Smith BJ (2007). boa: an R Package for MCMC Output Convergence Assessment and Posterior Inference. *J Stat Softw* **21**: 1-37.

Sorensen D, Gianola D (2002). Likelihood, Bayesian, and MCMC methods in quantitative genetics. New York, Springer.

Storset A, Strand C, Wetten M, Kjøglum S, Ramstad A (2007). Response to selection for resistance against infectious pancreatic necrosis in Atlantic salmon (*Salmo salar* L.). *Aquaculture* **272:** S62-S68.

Strømsheim A, Eide DM, Hofgaard PO, Larsen HJS, Refstie T, Røed KH (1994). Genetic variation in the humoral immune response against *Vibrio salmonicida* and in antibody titre against *Vibrio anguillarum* and total IgM in Atlantic salmon (*Salmo salar*). *Vet Immunol Immunopathol* **44**: 85-95.

Thiéry R, Cozien J, Cabon J, Lamour F, Baud M, Schneemann A (2006). Induction of a protective immune response against viral nervous necrosis in the European sea bass *Dicentrarchus labrax* by using Betanodavirus virus-like particles. *J Virol* **80**: 10201-10207.

Tine M, Kuhl H, Gagnaire P-A, Louro B, Desmarais E, Martins RST et al. (2014). The European sea bass genome and its variation provide insight into adaptation to euryhalinity and marine speciation. *Nat Commun* **5**: 5770.

Valero Y, Awad E, Buonocore F, Arizcun M, Esteban MA, Meseguer J, Chaves-Pozo E, Cuesta A (2016). An oral chitosan DNA vaccine against nodavirus improves transcription of cell-mediated cytotoxicity and interferon genes in European sea bass juveniles gut and survival upon infection. *Dev Comp Immunol* **65**: 64-72.

Vallejo R, Palti Y, Liu S, Evenhuis J, Gao G, Rexroad C et al. (2014). Detection of QTL in rainbow trout affecting survival when challenged with *Flavobacterium psychrophilum*. *Mar Biotechnol* **16**: 349-360.

Vandeputte M, Porte JD, Auperin B, Dupont-Nivet M, Vergnet A, Valotaire C, Claireaux G, Prunet P, Chatain B (2016). Quantitative genetic variation for post-stress cortisol and swimming performance in growth-selected and control populations of European sea bass (*Dicentrarchus labrax*). *Aquaculture* **455**: 1-7.

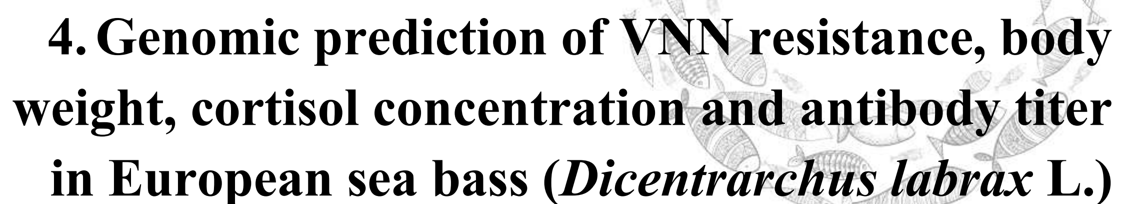
Vendramin N, Toffan A, Mancin M, Cappellozza E, Panzarin V, Bovo G, Cattoli G, Capua I, Terregino C (2014). Comparative pathogenicity study of ten different betanodavirus strains in experimentally infected European sea bass, *Dicentrarchus labrax* (L.). *J Fish Dis* **37**: 371-383.

Volckaert FAM, Hellemans B, Batargias C, Louro B, Massault C, Van Houdt JKJ, Haley C, de Koning D-J, Canario AVM (2012). Heritability of cortisol response to confinement stress in European sea bass *Dicentrarchus labrax. Genet Sel Evol* **44:** 15.

Wang S, Meyer E, McKay JK, Matz MV (2012). 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nat Methods* **9**: 808-810.

Yin Z, Lam T, Sin Y (1995). The effects of crowding stress on the non-specific immunoresponse in fancy carp (*Cyprinus carpio* L.). *Fish Shellfish Immunol* **5**: 519-529.

Yong CY, Yeap SK, Omar AR, Tan WS (2017). Advances in the study of nodavirus. *Peer J* **5**: e3841.



4. Genomic prediction of VNN resistance, body weight, cortisol concentration and antibody titer in European sea bass (*Dicentrarchus labrax* L.)

Sara Faggion[1], Daniela Bertotto[1], Massimiliano Babbucci[1], Giulia Dalla Rovere[1], Rafaella Franch[1], Luca Bargelloni[1] and Paolo Carnier[1]

[1] Department of Comparative Biomedicine and Food Science (BCA), University of Padova, Italy

Abstract

Selective breeding programs based on genomic data are still not a common practice in aquaculture, although genomic selection has widely demonstrated its superiority over traditional pedigree-based approaches. In European sea bass, the significant heritability estimates obtained for traits like viral nervous necrosis (VNN) mortality, post-stress cortisol concentrations, antibody titer against nervous necrosis virus (NNV) and body weight make genomic selection for these traits an interesting possibility with a potential practical application in commercial farms that allows overcoming issues related to a problematic phenotyping.

In this study, we exploited an experimental population (N = 650) produced using a commercial sea bass broodstock. All the individuals were phenotyped for the traits of interest and genotyped with a genome-wide SNP dataset generated applying a high-throughput sequencing approach (2b-RAD).

We compared five Bayesian models (BayesA, BayesB, BayesC, Bayesian LASSO and Bayesian Ridge Regression) through a cross-validation process and we estimated the accuracy of the models when predicting the phenotype or the EBV (correlation between observed value and prediction) for VNN mortality, post-stress cortisol concentrations, antibody titer and body weight. We did not observe differences in the prediction accuracy between the models; the prediction accuracy of the EBV for VNN mortality reached values of 0.89; the prediction accuracies of the EBV and the phenotype for post-stress cortisol concentration were 0.88 and 0.22, respectively, for antibody titer 0.76 and 0.26, respectively, and for body weight 0.69 and 0.39, respectively.

Different metrics (area under the ROC curve, the proportion of true results among the total number of samples, Matthew's correlation coefficient), allowed us to evaluate the classification performance of the model predictions when used to classify the phenotype for VNN mortality. Using the prediction of the EBV for VNN mortality to classify the phenotypes of the same trait resulted in better classification performance in comparison with the classification based on genomic predictions of the phenotype.

Overall, these results indicate that genomic prediction could have a great potential for traits like VNN mortality, body weight, cortisol concentration and antibody titer in European sea bass, although further investigations on a larger experimental population are needed.

The analysis repeated using 9,987 SNP or 16,075 SNP provided similar results. Finding the correct balance between number of markers and prediction accuracy could be interesting with the aim to develop cost-effective genotyping, which might be important for driving the practical implementation of genomic selection in the sea bass industry.

Keywords European sea bass, genomic prediction, viral nervous necrosis, cortisol, antibody titer

4.1 Introduction

Selective breeding aiming at enhancing economically important complex traits in aquaculture species is a relatively young scientific field compared to that in terrestrial species, and it has followed mainly the industrialization and the importance increase of the aquaculture sector (Chavanne et al. 2016; Janssen et al. 2016).

Traditional selective breeding approaches are based on the estimated breeding value (EBV), commonly predicted from the individual phenotypes of the breeding candidates or those of their full- or half-sib families and computed with best linear unbiased prediction (BLUP) methods accounting for additive genetic relationships between animals. Currently, such selection approaches are the most widespread throughout the European breeding companies (Chavanne et al. 2016; Janssen et al. 2016) and their application ensures the achievement of relevant genetic gains for many economically important traits. Individual phenotyping of breeding candidates or their relatives plays a key role in the application of traditional selective approaches. Family-based selective procedures have significant limitations in accuracy (e.g., identical EBV is assigned to different family members), exploit only 50% of the total genetic variation (Ødegård and Meuwissen 2014), have unfavourable impact on inbreeding and restrain the selection response and efficiency (Gjedrem and Baranski 2009).

The availability of new genomic tools and techniques enabling the development of high-density markers arrays makes possible the incorporation of dense markers genotypes into fish breeding programs. Gene-assisted and marker-assisted selection (MAS), based on sparse maps of genetic markers associated with quantitative trait loci (QTL), have been early attempts to exploit genomic information in selective breeding. Limitations of these approaches in artificial selection for complex traits, where many QTL of small effect exist, are the small amount of genetic variance picked up by the genomic information used (Meuwissen et al. 2016). For MAS, the limitation is represented by the need of assessing the linkage phase between each marker and QTL for each family and the instability of the detected marker-QTL associations due to a weak linkage resulting from large genome distances between a marker and QTL. The underlying genetic architecture of the trait of interest plays a significant role in the design of selective breeding programs and the majority of economically important traits shows a polygenic architecture whose genetic variation is determined by a large number of QTL of small effect.

Genomic selection (Meuwissen et al. 2001) overcomes the deficiencies of MAS because it is based on genome-wide dense marker maps, typically of single nucleotide polymorphisms (SNP), covering the whole genome and making the exploitation of population-wide linkage disequilibrium feasible. Phenotypic records and genotypes of a reference population, representative of the genetic characteristics of the animals candidated to breeding, are employed to estimate SNP effects on the traits of interest and then used to predict the genetic merit (genomic estimated breeding values, GEBV) of breeding candidates which are subjected to genotyping only (Meuwissen et al. 2001). In aquaculture species, better performance of genomic selection for complex polygenic traits compared with traditional approaches have been recently reported by various authors (Bangera et al. 2017; Palaiokostas et al. 2016, 2018a and 2018b; Tsai et al. 2016; Vallejo et al. 2017) who detected an improvement in selection accuracy and magnitude of expected genetic gain.

Disease outbreaks may have important impacts on aquaculture, threatening the survival rate of farmed animals, interfering with the progress and sustainability of intensive systems, and, consequently, causing significant production and economic losses. Selective breeding for disease resistance is

viewed as a feasible and a sustainable approach to prevent and control disease outbreaks in aquaculture farms (Ødegård et al. 2011).
For the European sea bass (*Dicentrarchus labrax* L.) industry one of the major threat is represented by viral nervous necrosis (VNN) and, recently, the genetic enhancement of resistance against VNN has raised great interest (Munday et al. 2002). Significant heritability (h^2) estimates for VNN resistance, reported in Chapter 3 of this thesis (h^2 = 0.14-0.23) and in other studies (h^2 = 0.26-0.27; Doan et al. 2017; Palaiokostas et al. 2018a) indicate the presence of a genetic basis for this trait and the opportunity for the development of selective breeding approaches. Employment of traditional procedures in selection for VNN resistance is hampered by difficulties in phenotyping, which should be carried out, as a routine, in experimental challenge test on relatives (full- and half-sibs) of the breeding candidates.
In Chapter 3, we investigated the variation of antibody titer against nervous necrosis virus (NNV) antigens in sea bass, obtaining a moderate heritability (h^2 = 0.28-0.39) and a negative genetic correlation (r_a) with VNN mortality (r_a = -0.39). These results suggest the exploitation of VNN antibody titer in indirect selection procedures to enhance VNN resistance. However, the feasibility of routine phenotyping for antibody titer raises a number of critical concerns.
Despite a relationship between stress response and disease resistance has been hypothesized (MacKenzie et al. 2009), the role of post-stress cortisol level (the most relevant physiological stress-response measure; Pottinger 2008) as an indicator trait for VNN resistance seems to be questionable due to a null genetic relationship with mortality (this thesis, Chapter 3). However, the improvement of fish stress resistance is still interesting in a context of production, where attention to different responses to stressful events and possible consequences on productive performance has increased in the last years. Low stress-sensitive fish show increased efficiency in aquaculture environment, higher resilience and enhanced productivity in face of environmental challenges typical of intensive systems (Castanheira et al. 2015). In sea bass, significant additive genetic variation in post-stress cortisol level has been reported (h^2 = 0.19-0.23, Chapter 3; 0.08 to 0.33, Volckaert et al. 2012 and Vandeputte et al. 2016), making selective breeding for this trait feasible.
When traits like disease resistance, stress resistance and antibody titer against a specific pathogen are among the breeding objectives, direct recording of individual phenotypes of the animals, which is often difficult or sometimes unfeasible, costly and time-consuming to implement routinely, is a critical issue. The development of genomic tools, to predict the genetic merit for such traits, is helpful in overcoming these criticalities.
The aims of this study were to assess the accuracy in the prediction of the phenotype or "traditional" EBV for VNN mortality, antibody titer against NNV, post-stress serum cortisol concentration and body weight at a constant age provided by five Bayesian regression models exploiting the genomic information of a genome-wide dense panel of SNP, and to compare the performance of these genomic predictions when used in the classification of the phenotype for VNN mortality.

4.2 Material and methods

4.2.1 Experimental fish: production, rearing and NNV challenge test

The procedures of production, rearing and NNV challenge test of the experimental fish were detailed in Chapter 3. Briefly, artificial fertilization was performed, using a commercial NNV-free breeding

stock of 20 dams and 32 sires, by mating every time, according to a factorial scheme, 5 dams to 8 sires.
Before transfer to the experimental facility of the Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe, Legnaro, PD) for NNV infection, fish were reared in a sea cage.
The challenge test with NNV was carried out using 652 randomly-chosen fish, which had an age of 548 d post-hatching and an average body weight of 146.4 g (CV% 29.18). Before being infected through intramuscular injection of 0.1 ml of viral suspension (RGNNV 283.2009, batch 7/16; $10^{8.30}$ $TCID_{50}$ per ml; the virus strain was isolated from animals belonging to a commercial stock during an outbreak in the northern Adriatic sea and propagated on E11 cells; Iwamoto et al. 2000; Vendramin et al. 2014), fish were subjected to a stress-test (acute stress, confinement at high-density for a minimum of 10 minutes), anesthetized (MS-222, 30 ppm), individually tagged with passive integrated transponders (PIT-tag) and individually blood-sampled to perform indirect ELISA assay and cortisol analysis by radioimmunoassay (RIA).
Fish were distributed into six close-system tanks with different capacity (380 litres or 2500 litres). After infection, they were checked three times a day to record clinical signs of VNN and mortality. The experiment ended at 29 d post-challenge, when a cumulative survival rate of 52.2% was recorded. All fish that were alive at d 29 were euthanized with an overdose of anaesthetic (MS-222) and classified as survivors. Tissue samples (muscle and fin) were collected from each fish (dead or survived) and preserved in absolute ethanol for genomic DNA extraction.
The experiment protocol was evaluated and approved by the Italian Ministry of Health (Law decree 26/2014 art. 31; permission number: 975/2016-PR of 13/10/2016).

4.2.2 Antibody titer against NNV and cortisol concentrations

Indirect ELISA assay for antibodies against NNV and cortisol extraction and measurement were described in Chapter 3. Briefly, indirect ELISA assay was performed at the Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe, Legnaro, PD) on the blood serum collected from each fish before the infection, following the protocol by Scapigliati et al. (2010) and Nuñez-Ortiz et al. (2016). The ELISA optical density values (OD 450 nm) sample to positive ratio (S/P ratio) was applied to evaluate the status of the sample.
Cortisol was extracted from blood serum, following the protocol by Bertotto et al. (2010), and the concentration of cortisol of each sample was analysed through solid-phase RIA (Simontacchi et al. 1995) with minor modifications. Radioactivity values counted on a β-counter (Top-Count NXT, Perkin Elmer Life and Analytical Sciences) were analysed with GraphPad Prism 5.0 software (La Jolla, CA, USA), which provided the cortisol concentration of each sample expressed as pg/well. The concentration in pg/well was later converted to ng per ml of serum and normalized through the square root transformation.

4.2.3 Genotyping, parentage assignment and pedigree reconstruction

Genomic DNA was extracted from muscle or fin samples of the experimental fish (N = 652) and their parents (N = 52) using the commercial kits Invisorb® Spin Tissue Mini Kit and Invisorb® DNA Tissue HTS 96 Kit (Invitek, STRATEC Biomedical, Germany), quality checked and quantified. A 2b-RAD library was constructed for each individual following the protocol by Wang et al. (2012) with minor modifications. Individual libraries were pooled into equimolar amounts (52 for parents, 96-97 for offspring). Pooled libraries were analyzed with Agilent 2100 Bioanalyzer (Agilent

Technologies, Santa Clara, California, USA) and then sequenced on an Illumina HiSeq4000 platform with a 50 bp single-read module at Fasteris SA (Plan-les-Ouates, Switzerland; http://www.fasteris.com) and UC Davis (http://comailab.genomecenter.ucdavis.edu), which performed also demultiplexing and a first quality-check of raw data. After filtering and trimming, the fragments were elaborated with STACKS software 2.0 (Catchen et al. 2013), firstly mapping the trimmed reads against European sea bass genome (Tine et al. 2014; http://seabass.mpipz.mpg.de/DOWNLOADS/dicLab1_scaffold.fasta), then allowing the identification of SNPs (to construct the stacks and the catalog of loci, a minimum coverage of 20X and 5X were used for parental samples and for the offspring, respectively). Results were filtered excluding loci shared by less than 75% of the individuals, thus resulting in a dataset of 18,097 SNPs. Some SNPs were discarded on the basis of a minor allele frequency (MAF) lower than 1%, genotype frequency deviating from Hardy-Weinberg equilibrium and missing genotype rate greater than 15%. Missing genotypes for the remaining SNP were imputed using the FImpute software (Sargolzaei et al. 2014). The final number of available SNP genotypes per animal was 16,075. Moreover, two individuals of the experimental population were discarded because of bad genotyping results.
The parentage assignment and the pedigree reconstruction were performed combining the software CERVUS 3.0 (Marshall et al. 1998; Kalinowski et al. 2007) and the R package sequoia (Huisman 2017). 100% of the experimental fish were assigned to a unique parental pair, 41 parents were assigned to 22 dummy parental pairs (14 grandsires and 8 granddams), 3 to 3 single dummy parents (1 grandsires and 2 granddams) whereas 8 remained unassigned. The overall number of full-sib families was 136 as progeny of 19 dams and 30 sires (two sires and one dam were not assigned to any individual in the challenged sample).

4.2.4 Estimation of “traditional” breeding values

“Traditional” EBV for VNN mortality coded as a binary trait (0 for the fish that were alive at the end of the NNV challenge test, 1 for the dead ones), body weight at 548 d post-hatching, cortisol concentration and antibody titer were estimated using mixed animal models. A Bayesian approach employing Monte-Carlo Markov Chain (MCMC) and Gibbs sampling methods was implemented using the software TM (Legarra et al. 2008) and the following models:

$$g(\pi_{ij}) = \lambda_{ij} = \mu + t_i + a_{ij} + e_{ij} \text{ (for VNN mortality)}$$

$$y_j = \mu + a_j + e_j \text{ (body weight, cortisol concentration and antibody titer)}$$

where $\pi_{ij}$is the probability of death of animal ij in tank i, $g(\cdot)$ is the probit link function, λ_{ij} is an unobservable latent variable (liability) assumed to be normally distributed, y_j is an observed phenotype, μ is the model intercept, t_i is the fixed effect of the challenge test tank i, a_j and a_{ij} are the random additive genetic effects of animals assumed to be $N(0, \mathrm{A}\sigma_\mathrm{a}^2)$ where $N()$ indicates a normal distribution, A is the numerator relationship matrix and σ_a^2 is the additive genetic variance, e_{ij} and e_i are random residual effects assumed to be $N(0, \sigma_e^2)$ where σ_e^2 is the residual variance with $\sigma_e^2 = 1$ for e_{ij}. Prior densities used in the Bayesian analysis were uniform densities for the fixed effects, normal densities for additive genetic and residual effects and scaled inverted χ^2 distributions for the additive genetic and residual variance. The hyperparameters of scaled inverted χ^2 distributions (v and s^2) were equal to -2 and 0, making the prior distribution flat.

The Gibbs sampler was run using 1,000,000 iterations, with a burn-in of 100,000 iterations and thinning interval of 20 samples. More details about the Bayesian analysis can be found in Chapter 3.

4.2.5 Genomic predictions

4.2.5.1 Genomic prediction models

Five Bayesian regression models were fitted to the data using the BGLR package in R software (Pérez and de Los Campos 2014): BayesA (Meuwissen et al. 2001), BayesB, BayesC (Habier et al. 2011), Bayesian LASSO (Park and Casella 2008) and Bayesian Ridge Regression (Meuwissen et al. 2001). These models differ in the prior distribution used for the marker genotype effects accounted for by the model; the prior distribution determines the type of shrinkage of estimates of the effects induced (for details see Pérez and de Los Campos 2014).
The scaled-t density (used as prior in BayesA models; Meuwissen et al. 2001) and the double exponential density (typical of Bayesian LASSO models; Park and Casella 2008) are two types of prior with higher mass at zero and thicker tails than the normal distribution that induce size-of-effect dependent shrinkage (Gianola 2013).
The Bayesian Ridge Regression model includes a Gaussian prior, which assume that all marker effects are shrunk to a similar extent (Meuwissen et al. 2001).
Two finite mixture priors are implemented in BayesB and BayesC: in the first case, the prior is a mixture of a point of mass at zero and a scaled-t density, in the second case the prior is a mixture of a point of mass at zero and a Gaussian distribution (Habier et al. 2011); in both cases, variable selection can be performed through the assignment of a nonnull prior probability for the marker effect to be equal to zero.
Two genetic architectures for the trait under investigation were considered for each Bayesian model: 1) additive genetic effects only or 2) additive genetic and dominance effects. Consequently, the general form of each Bayesian model was as follows:

1) additive genetic effects

$$y_i = \mu + \sum_{s=1}^{S} b_s X_{si} + e_i$$

2) additive genetic + dominance effects

$$y_i = \mu + \sum_{s=1}^{S} b_s X_{si} + \sum_{s=1}^{S} c_s H_{si} + e_i$$

where y_i was, for the trait under investigation, the observed phenotype or the "traditional" EBV of animal i, μ is the model intercept, X_{si} is the genotype, coded as the number of copies of the MAF allele (0, 1 or 2), of animal i at SNP s, H_{si} is a dominance component (0 for homozygous genotypes and 1 for heterozygous genotypes) at SNP s, b_s and c_s are unknown additive genetic and dominance random effects (to be estimated), S is the total number of SNP, and e_i is a random residual effect.
In the analysis of the phenotypes for VNN mortality, the probit function was used as a link function to properly account for the binary nature of the trait.
Each Bayesian analysis was carried out using a minimum MAF of 1% (S = 16,075) or of 5% (S = 9,987) to filter out the SNP whose effects were fitted in the regression model and by generating a

single Gibbs chain of 200,000 samples, with a burn-in of 20,000 iterations and a thinning interval of 5 samples.

4.2.5.2 Assessment of prediction accuracy and performance in classification

The accuracy in the prediction of phenotypes or “traditional” EBV for the traits under investigation and performance metrics of such predictions when used in classification of the phenotypes for VNN mortality were assessed, for all models, by means of 16 independently-generated 5-fold cross-validations (CV). The performance in the classification of the phenotype for VNN mortality was evaluated using, as classifiers, the genomic predictions of the phenotype for VNN mortality and those of the “traditional” EBV of all the four investigated traits.

The total number of Bayesian analyses was equal to 12,800 (4 traits × 2 SNP minimum MAF × 2 predicted values × 2 genetic architectures × 5 Bayesian models × 16 CV × 5 CV iterates).

In each CV, the data (N = 650) were randomly split into five equally-sized data segments. In each of the five iterates performed in a single CV, four of these data segments (80% of the data) were used as a training set to obtain model solutions (parameters b_s and c_s) and the remaining segment served as a test set. In the test set, the phenotypes or “traditional” EBV for the trait under study were masked and set to missing. The phenotypes or EBV of the test set were then predicted using model solutions, provided by the analysis of the training data, and the genomic information of the test set animals. The predictions for each test set were saved, aggregated with the others at the end of the CV and used to compute a pool of metrics to assess the accuracy of the predictions and their performance when exploited in the classification of the phenotype for VNN mortality.

For continuous traits (body weight, post-stress cortisol serum concentration and antibody titer), the correlation (CORR) between the observed and predicted values was computed as a measure of model accuracy in the prediction of the phenotype or EBV. This metric was also used when evaluating the accuracy in the prediction of EBV for VNN mortality.

Three metrics were used to evaluate the performance of the genomic Bayesian models when their predictions of the “traditional” EBV (for all traits) or of the observed phenotype (for VNN mortality only) were used to classify the phenotype for VNN mortality: the area (AUC) under the ROC (receiving operating characteristics) curve, the prediction accuracy (ACC) for binary traits and the Matthew’s correlation coefficient (MCC).

The AUC is the most appropriate metric to evaluate model performance in classification (Fawcett 2006). This metric measures the ability of a classifier (model predictions) in discriminating the two classes for VNN mortality (0: alive, 1: dead): values for AUC range between 1 (perfect class separation capacity) and 0 (classification perfectly reversed), with 0.5 as a value indicating a performance comparable to that of a classification strategy based on random guessing a class. Further details on ROC curves and their use can be found in Fawcett (2006).

The ACC is a measure of how well a binary classifier correctly identifies the two classes and it is computed as:

$$ACC = \frac{TP + TN}{n}$$

where TP is the number of true positive, TN is the number of true negative and n is the total number of classified samples.

The Matthew's correlation coefficient (Matthew 1975) is the correlation coefficient between the observed and the predicted binary classifications; it is a measure of the quality of the binary classification and it is calculated as:

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

where *TP* is the number of true positive, *TN* is the number of true negative, *FP* is the number of false positive and *FN* is the number of false negative. The coefficient ranges between -1 (total disagreement between predicted and observed values) and +1 (perfect prediction); a value of 0 suggests that the model prediction is not better than random prediction.
All the metrics were computed in R using the package ROCR (Sing et al. 2005).

4.3 Results

4.3.1 Genomic prediction and comparison of different models

The accuracy (CORR) in the prediction of the phenotype or EBV of the investigated trait ensured by the five Bayesian genomic models and the metrics to evaluate the performance of such predictions in the classification of the VNN mortality phenotype are reported in Tables 1, 2, 3, and 4 for VNN mortality, body weight at 548 d post-hatching, serum cortisol concentration and antibody titer against NNV antigens, respectively.
Overall, the five Bayesian genomic models showed similar accuracies (CORR) in the prediction of the phenotype for a given trait. Such accuracies were higher for body weight (average CORR = 0.39; Table 2) than for serum cortisol concentration (average CORR = 0.22; Table 3) and antibody titer (average CORR = 0.26; Table 4). The SNP marker density, based on different minimum MAF (1 or 5%) to filter out the SNP whose effects were included in models, and the assumed architectures for the genetic determinism of the traits (additive genetic effects only or additive genetic plus dominance effects), did not affect the accuracies of the models when predicting the phenotype of the traits.
When the predicted value was the "traditional" EBV for the traits, the prediction accuracies of the genomic models increased with trivial differences across models. The trait whose EBV were predicted with the largest accuracy was VNN mortality (average CORR = 0.89; Table 1) which showed a slightly higher prediction accuracy than that for cortisol concentration (average CORR = 0.88; Table 3). The prediction accuracies for the EBV were somewhat lower for body weight (average CORR = 0.69; Table 2) and antibody titer (average CORR = 0.76; Table 4). Also in the EBV prediction, the SNP marker density and the type of genetic effects fitted had no effect on the prediction accuracies of the models.
Using the genomic predictions of the phenotype for VNN mortality, provided by the five models, as classifiers of the observed VNN mortality resulted in unsatisfactory classification performance: the average CV-AUC (Table 1) was very close to the expected value for this metric (0.5) when a random guessing of the classes is used as a classification strategy. The ACC (average ACC = 0.52; Table 1) and MCC (average MCC = 0.07; Table 1) metrics also indicated poor classification performance. Conversely, when the classification of VNN mortality phenotypes was carried out using the genomic predictions of the EBV for the trait, the metrics were more favourable (average AUC = 0.60; average ACC = 0.58; average MCC = 0.17; Table 1), indicating that the genomic classifiers contributed an

added value to the prediction of the mortality class. These results were consistent across models, marker densities and types of genetic effects accounted for in the analysis.
The genomic predictions of the EBV for body weight, cortisol concentration and antibody titer showed performance in the classification of the phenotype of mortality, which were inferior to those of the genomic-predicted EBV for mortality. The classification based on the genomic prediction of the EBV for body weight or antibody titer resulted in a CV-AUC which was slightly lower than 0.5 (body weight: average AUC = 0.48; Table 2; antibody titer: average AUC = 0.48; Table 4) whereas the one based on the genomic-predicted EBV for cortisol concentration had an AUC of 0.5 (Table 3). Again, the metrics were consistent across models, marker densities and types of genetic effects used.

4.4 Discussion

In European sea bass, the phenotypic variation in VNN mortality, body weight, post-stress cortisol concentration and antibody titer against NNV has a non-null additive genetic component which has been consistently detected in different studies (Chapter 3 of this thesis; Doan et al. 2017; Dupont-Nivet et al. 2008; Palaiokostas et al. 2018a; Vandeputte et al. 2016; Volckaert et al. 2012). Traditional breeding programs for complex traits require individual phenotyping of the breeding candidates or of their relatives. For some of the traits investigated in this thesis, like VNN mortality and antibody titer, routine phenotyping is unfeasible or, like for cortisol concentration, difficult and expensive when the number of breeding candidates is large.
Hence, the practical implementation of genomic selection procedures is greatly valuable for traits whose enhancement cannot be efficiently addressed with traditional methods (Ødegård and Meuwissen 2014; Meuwissen et al. 2016). The exploitation of genomic information, provided by genome-wide dense marker panels, offers the opportunity of developing innovative tools for the prediction of the genetic merit of breeding candidates, which reduce the need of individual phenotyping. Improvement in selection accuracy due to genomic predictions and the benefits of genomic selection over pedigree-based methods have been documented both with simulation studies (Sonnesson and Meuwissen 2009) and in application studies in various fish species for traits like growth and disease resistance (Tsai et al. 2016; Vallejo et al. 2017; Palaiokostas et al. 2016, 2018a, 2018b). Genome-wide dense marker data are typically required and have become increasingly available due to the development of high-throughput genotyping techniques (such as RAD and 2b-RAD sequencing; Robledo et al. 2017).
In this study, we exploited a genome-wide SNP panel for European sea bass, generated by applying a 2b-RAD sequencing approach, with the aim of developing genomic predictions of the phenotype or "traditional" EBV of the investigated traits. "Traditional" EBV were considered because, as estimates of the additive genetic component of the phenotype, their genomic prediction is easier when compared to the one of raw phenotypes. Moreover, for VNN mortality, which is a major trait in our study, the "traditional" EBV, estimated from the data of the challenge test exhibited good performance (Figure 1) in the prediction of the phenotype of mortality. Hence, it was worth to exploit EBV availability in a genomic prediction framework.
The accuracy of the tested genomic models in the prediction of the trait phenotype were satisfactory, ranging from 0.22 to 0.39, depending upon the trait. It is worth noting that the genomic information exploited by the models can actually predict only the genetic component of the phenotype, which, for the investigated traits, explains a proportion of the phenotypic observed variation ranging from 14 to 57% (see heritability estimates presented in Chapter 3). If the true values of the additive genetic

component of the individual phenotypes were known without error and used in the prediction of the phenotype, the expected accuracy of such prediction would be equal to the square root of the heritability of the trait. By assuming that the estimated heritabilities obtained in Chapter 3 are reliable, the accuracy of the genomic models in the prediction of the additive genetic component of the phenotype can be computed as $CORR/\sqrt{h^2}$ where CORR is the accuracy of the genomic model predictions of the phenotype and h^2 is the heritability. Hence, such accuracies were moderately high in our study, being 0.52, 0.51 and 0.44 for body weight, cortisol concentration and antibody titer, respectively. Being a recent field of investigation for aquaculture species, literature estimates of accuracy of prediction genomic models are scarce. Palaiokostas et al. (2018b) reported, in a study performed on common carp, an accuracy in the prediction of the genetic component of the phenotype for growth ranging from 0.60 to 0.70.

As expected, the accuracies of the genomic models in the prediction of the “traditional” EBV were always greater than those observed in the prediction of phenotypes and were very high for the EBV of VNN mortality and cortisol concentration and moderately high for body weight and antibody titer. Using the prediction of the EBV for VNN mortality to classify the phenotypes of the same trait resulted in better classification performance, as measured by the AUC metric, in comparison with the classification based on genomic predictions of the phenotype. However, the value for the AUC metric was lower than the one (70%) reported by Palaiokostas et al. (2018a) who used the genomic prediction of the phenotype for VNN resistance in sea bass to classify the observed survival in a challenge test where fish of small size were infected by immersion. These inconsistencies across studies may be ascribed to the different nature of resistance assessed in the two studies because of the methods of infection used in the challenge test. In our study, due to the size of the experimental fish, NNV infection was carried out through injection. The probability of surviving was then influenced by the response of the specific components of the individual immune system (antibodies, lymphocytes) or other non-specific components such as cells (granulocytes and macrophages), proteins (related to the complement system) and responses (inflammation). When immersion is used as infection method, the outcome of the test is affected by a larger variety of defense mechanisms, which involves also an unspecific response provided, for instance, by mechanical barriers, mucus and other components. This is partly confirmed by the magnitude of the estimated heritabilities for VNN resistance reported in the two studies which, in Palaiokostas et al. (2018a), was approximately two times as high as the one detected in our study. Hence, the observed phenotypes in the two studies may actually be observations of different resistance traits. Comparisons with literature studies on different species and diseases (Atlantic salmon, Tsai et al. 2016 and Bangera et al. 2017; gilthead sea bream, Palaiokostas et al. 2016) are not straightforward due to the use of a different metric (accuracy in the prediction of the genetic component of phenotype) which ranged from 0.4 to 0.7.

No one of the genomic predictions of the EBV for the continuous traits contributed an added value to the prediction of VNN mortality. This unsatisfactory performance might be explained by the sample size of our experimental population or by weaker “true” genetic correlations. The sample size of the reference population used to estimate SNP effects is a major factor for the prediction accuracy, but it is a key feature also for the precision of the estimated genetic correlations. Increasing the number of genotyped animals has been proved to enhance the prediction accuracy (Goddard and Hayes 2009; Vallejo et al. 2017). The large uncertainty in the estimation of the genetic correlations

between the investigated traits observed in Chapter 3 might also be greatly reduced by a greater sample size.

The marker density and the type of genetic effects are currently another point of discussion in relation to their effects on the prediction accuracy of genomic models. High-density marker panels are thought to be more efficient and to give better prediction accuracy (Meuwissen et al. 2001). In our study, neither the marker density nor the type of genetic effects accounted for by the models influenced the prediction accuracy and the classification performance.

4.5 Conclusions

Our results indicate that genomic predictions of the phenotype or of “traditional” EBV can be exploited in selective breeding for VNN resistance, response to stressors and growth performance of European sea bass, with a potential of important practical applications, which lead to the reduction of the need for routine phenotyping and outperform traditional approaches in the enhancement of disease resistance. The results obtained in this study can benefit from further investigations on a larger experimental population, in order to be confirmed and to become applicative in a context of selective breeding programs. Further studies should elucidate the effects of the method of infection used in challenge tests on the variation in the infection response observed and genetic parameters thereof.

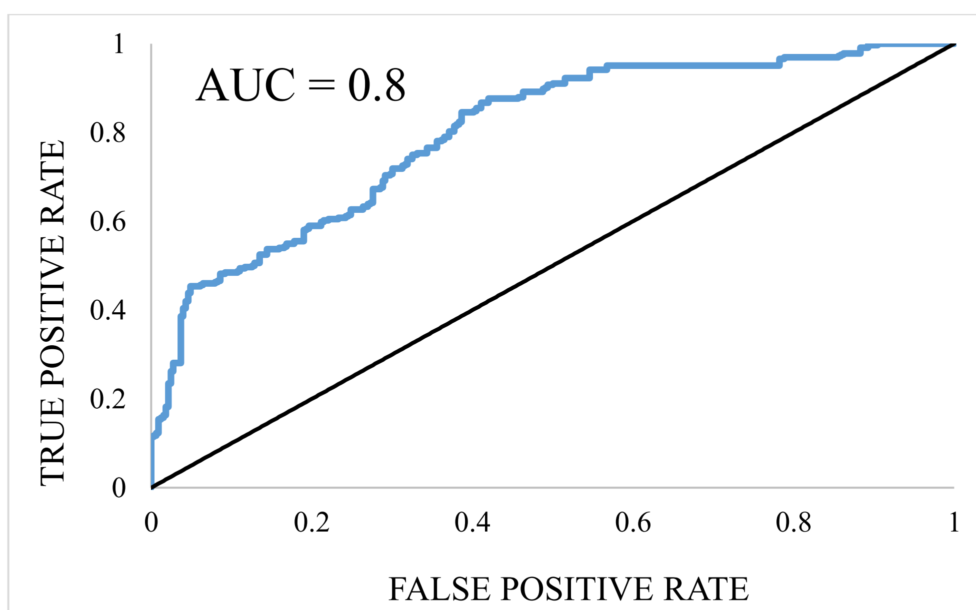


Figure 1. ROC curve for the prediction of VNN mortality provided by "traditional" estimated breeding value

Table 1. Accuracy (CORR) when predicting the estimated breeding value (EBV) for VNN mortality with five Bayesian models exploiting genotypes of a dense SNP panel and their performance (AUC, ACC, MCC) when model predictions of phenotype or EBV were used to classify the phenotype for VNN mortality in 16 independent 5-fold cross-validations

SNP MAF[1]	Predicted value[2]	Genetic effect fitted	Model[3]	Average (SD) of 16 cross-validations[4]			
				AUC	ACC	MCC	CORR
≥ 0.01	EBV	Additive	BA	0.5965 (0.0028)	0.5796 (0.0040)	0.1664 (0.0087)	0.8922 (0.0021)
			BB	0.5959 (0.0028)	0.5795 (0.0034)	0.1662 (0.0084)	0.8917 (0.0021)
			BC	0.5965 (0.0028)	0.5797 (0.0039)	0.1674 (0.0086)	0.8922 (0.0021)
			BL	0.5966 (0.0028)	0.5799 (0.0039)	0.1668 (0.0091)	0.8925 (0.0021)
			BRR	0.5969 (0.0028)	0.5804 (0.0039)	0.1678 (0.0089)	0.8925 (0.0021)
		Additive + Dominance	BA	0.5958 (0.0027)	0.5800 (0.0033)	0.1658 (0.0074)	0.8936 (0.0020)
			BB	0.5957 (0.0027)	0.5804 (0.0033)	0.1660 (0.0078)	0.8935 (0.0022)
			BC	0.5962 (0.0027)	0.5801 (0.0023)	0.1660 (0.0062)	0.8935 (0.0021)
			BL	0.5963 (0.0029)	0.5799 (0.0034)	0.1668 (0.0086)	0.8927 (0.0021)
			BRR	0.5963 (0.0029)	0.5801 (0.0033)	0.1664 (0.0079)	0.8938 (0.0022)
	Phenotype	Additive	BA	0.4982 (0.0175)	0.5257 (0.0120)	0.0714 (0.0274)	
			BB	0.4969 (0.0226)	0.5246 (0.0144)	0.0721 (0.0295)	
			BC	0.4876 (0.0249)	0.5195 (0.0128)	0.0680 (0.0231)	
			BL	0.4912 (0.0231)	0.5226 (0.0130)	0.0698 (0.0246)	
			BRR	0.4918 (0.0248)	0.5218 (0.0138)	0.0708 (0.0318)	
		Additive + Dominance	BA	0.4977 (0.0178)	0.5256 (0.0119)	0.0696 (0.0235)	
			BB	0.4926 (0.0203)	0.5216 (0.0123)	0.0684 (0.0276)	
			BC	0.4900 (0.0239)	0.5183 (0.0127)	0.0621 (0.0242)	
			BL	0.4925 (0.0216)	0.5226 (0.0136)	0.0718 (0.0273)	
			BRR	0.4899 (0.0233)	0.5202 (0.0121)	0.0670 (0.0277)	
≥ 0.05	EBV	Additive	BA	0.5979 (0.0027)	0.5816 (0.0030)	0.1694 (0.0081)	0.8907 (0.0028)
			BB	0.5973 (0.0027)	0.5813 (0.0037)	0.1704 (0.0088)	0.8902 (0.0027)
			BC	0.5981 (0.0027)	0.5814 (0.0032)	0.1698 (0.0081)	0.8909 (0.0028)
			BL	0.5981 (0.0027)	0.5810 (0.0031)	0.1689 (0.0082)	0.8911 (0.0028)
			BRR	0.5984 (0.0027)	0.5818 (0.0032)	0.1710 (0.0087)	0.8911 (0.0027)
		Additive + Dominance	BA	0.5974 (0.0028)	0.5809 (0.0028)	0.1685 (0.0069)	0.8921 (0.0026)
			BB	0.5969 (0.0027)	0.5813 (0.0031)	0.1687 (0.0070)	0.8916 (0.0026)
			BC	0.5976 (0.0028)	0.5805 (0.0030)	0.1676 (0.0073)	0.8922 (0.0026)
			BL	0.5979 (0.0027)	0.5812 (0.0031)	0.1692 (0.0069)	0.8912 (0.0027)
			BRR	0.5976 (0.0028)	0.5811 (0.0028)	0.1677 (0.0065)	0.8924 (0.0026)
	Phenotype	Additive	BA	0.5004 (0.0207)	0.5270 (0.0122)	0.0707 (0.0271)	
			BB	0.4940 (0.0216)	0.5225 (0.0108)	0.0680 (0.0314)	
			BC	0.4950 (0.0208)	0.5263 (0.0118)	0.0745 (0.0262)	
			BL	0.4849 (0.0222)	0.5213 (0.0094)	0.0664 (0.0251)	
			BRR	0.4970 (0.0225)	0.5279 (0.0117)	0.0780 (0.0243)	
		Additive + Dominance	BA	0.4962 (0.0196)	0.5244 (0.0096)	0.0711 (0.0251)	
			BB	0.4946 (0.0207)	0.5253 (0.0107)	0.0690 (0.0181)	
			BC	0.4916 (0.0201)	0.5247 (0.0098)	0.0721 (0.0217)	
			BL	0.4819 (0.0217)	0.5188 (0.0101)	0.0615 (0.0219)	
			BRR	0.4963 (0.0208)	0.5256 (0.0116)	0.0719 (0.0216)	

[1]MAF: minimum minor allele frequency of the SNP whose effects were accounted for by the models (≥ 1%, 16,670 SNP; ≥ 5%, 9,987 SNP)
[2]The 5 models analysed the phenotype (0: survived, 1: dead) or the estimated breeding value (EBV) for VNN mortality
[3]BA: BayesA; BB; BayesB; BC: BayesC; BL; Bayesian LASSO; BRR: Bayesian Ridge Regression
[4]AUC: area under the ROC curve; ACC: accuracy computed as (true positives + true negatives)/number of samples; MCC: Matthew's correlation; CORR: correlation between the observed value and the model prediction

Table 2. Accuracy (CORR) when predicting the phenotype or the estimated breeding value (EBV) for body weight at 548 d post-hatching with five Bayesian models exploiting genotypes of a dense SNP panel and classification performance (AUC, ACC, MCC) of the model predictions of the body weight EBV when used to classify the phenotype for VNN mortality in 16 independent 5-fold cross-validations

SNP MAF[1]	Predicted value[2]	Genetic effect fitted	Model[3]	Average (SD) of 16 cross-validations[4]			
				AUC	ACC	MCC	CORR
≥ 0.01	EBV	Additive	BA	0.4764 (0.0059)	0.5093 (0.0029)	0.0660 (0.0158)	0.6886 (0.0082)
			BB	0.4770 (0.0062)	0.5095 (0.0030)	0.0626 (0.0172)	0.6884 (0.0083)
			BC	0.4765 (0.0059)	0.5093 (0.0025)	0.0625 (0.0172)	0.6888 (0.0081)
			BL	0.4754 (0.0061)	0.5087 (0.0025)	0.0617 (0.0186)	0.6864 (0.0090)
			BRR	0.4763 (0.0060)	0.5093 (0.0025)	0.0642 (0.0176)	0.6887 (0.0082)
		Additive + Dominance	BA	0.4804 (0.0067)	0.5120 (0.0052)	0.0659 (0.0181)	0.6906 (0.0079)
			BB	0.4802 (0.0067)	0.5112 (0.0040)	0.0635 (0.0174)	0.6905 (0.0081)
			BC	0.4799 (0.0061)	0.5119 (0.0046)	0.0646 (0.0178)	0.6909 (0.0079)
			BL	0.4793 (0.0065)	0.5114 (0.0041)	0.0585 (0.0201)	0.6888 (0.0085)
			BRR	0.4803 (0.0065)	0.5110 (0.0050)	0.0607 (0.0179)	0.6904 (0.0081)
	Phenotype	Additive	BA				0.3985 (0.0147)
			BB				0.3986 (0.0146)
			BC				0.3975 (0.0147)
			BL				0.3959 (0.0135)
			BRR				0.3975 (0.0144)
		Additive + Dominance	BA				0.3932 (0.0177)
			BB				0.3931 (0.0164)
			BC				0.3937 (0.0176)
			BL				0.3926 (0.0196)
			BRR				0.3942 (0.0174)
≥ 0.05	EBV	Additive	BA	0.4731 (0.0036)	0.5079 (0.0026)	0.0612 (0.0219)	0.6888 (0.0097)
			BB	0.4735 (0.0037)	0.5082 (0.0030)	0.0629 (0.0185)	0.6888 (0.0100)
			BC	0.4733 (0.0037)	0.5082 (0.0027)	0.0631 (0.0195)	0.6890 (0.0094)
			BL	0.4727 (0.0052)	0.5079 (0.0029)	0.0592 (0.0222)	0.6855 (0.0093)
			BRR	0.4731 (0.0035)	0.5081 (0.0025)	0.0626 (0.0187)	0.6892 (0.0093)
		Additive + Dominance	BA	0.4762 (0.0038)	0.5081 (0.0020)	0.0608 (0.0166)	0.6918 (0.0087)
			BB	0.4762 (0.0041)	0.5078 (0.0023)	0.0623 (0.0143)	0.6921 (0.0089)
			BC	0.4762 (0.0041)	0.5083 (0.0027)	0.0621 (0.0186)	0.6919 (0.0084)
			BL	0.4756 (0.0040)	0.5092 (0.0032)	0.0595 (0.0207)	0.6886 (0.0096)
			BRR	0.4767 (0.0040)	0.5087 (0.0030)	0.0603 (0.0173)	0.6921 (0.0084)
	Phenotype	Additive	BA				0.3889 (0.0133)
			BB				0.3890 (0.0134)
			BC				0.3891 (0.0135)
			BL				0.3826 (0.0144)
			BRR				0.3897 (0.0139)
		Additive + Dominance	BA				0.3813 (0.0173)
			BB				0.3827 (0.0157)
			BC				0.3828 (0.0163)
			BL				0.3729 (0.0172)
			BRR				0.3832 (0.0162)

[1]MAF: minimum minor allele frequency of the SNP whose effects were accounted for by the models (≥ 1%, 16,670 SNP; ≥ 5%, 9,987 SNP)
[2]The 5 models analysed the phenotype or the estimated breeding value (EBV) for body weight at 548 d post-hatching
[3]BA: BayesA; BB; BayesB; BC: BayesC; BL; Bayesian LASSO; BRR: Bayesian Ridge Regression
[4]AUC: area under the ROC curve; ACC: accuracy computed as (true positives + true negatives)/number of samples; MCC: Matthew's correlation; CORR: correlation between the observed value and the model prediction

Table 3. Accuracy (CORR) when predicting the phenotype or the estimated breeding value (EBV) for the square root of serum cortisol concentration with five Bayesian models exploiting genotypes of a dense SNP panel and classification performance (AUC, ACC, MCC) of the model predictions of cortisol concentration EBV when used to classify the phenotype for VNN mortality in 16 independent 5-fold cross-validations

SNP MAF[1]	Predicted value[2]	Genetic effect fitted	Model[3]	Average (SD) of 16 cross-validations[4]			
				AUC	ACC	MCC	CORR
≥ 0.01	EBV	Additive	BA	0.5008 (0.0013)	0.5176 (0.0029)	0.0634 (0.0117)	0.8818 (0.0035)
			BB	0.5009 (0.0014)	0.5175 (0.0030)	0.0649 (0.0098)	0.8818 (0.0037)
			BC	0.5007 (0.0013)	0.5178 (0.0031)	0.0642 (0.0098)	0.8818 (0.0035)
			BL	0.5007 (0.0013)	0.5172 (0.0029)	0.0641 (0.0099)	0.8820 (0.0035)
			BRR	0.5007 (0.0013)	0.5176 (0.0029)	0.0642 (0.0092)	0.8818 (0.0035)
		Additive + Dominance	BA	0.5004 (0.0013)	0.5170 (0.0030)	0.0652 (0.0086)	0.8833 (0.0034)
			BB	0.5006 (0.0012)	0.5176 (0.0028)	0.0645 (0.0115)	0.8829 (0.0035)
			BC	0.5004 (0.0012)	0.5171 (0.0027)	0.0642 (0.0095)	0.8831 (0.0033)
			BL	0.5008 (0.0014)	0.5177 (0.0033)	0.0654 (0.0108)	0.8826 (0.0035)
			BRR	0.5004 (0.0012)	0.5182 (0.0035)	0.0659 (0.0104)	0.8832 (0.0033)
	Phenotype	Additive	BA				0.2216 (0.0180)
			BB				0.2184 (0.0166)
			BC				0.2180 (0.0160)
			BL				0.2196 (0.0188)
			BRR				0.2199 (0.0173)
		Additive + Dominance	BA				0.2175 (0.0161)
			BB				0.2151 (0.0170)
			BC				0.2131 (0.0169)
			BL				0.2151 (0.0154)
			BRR				0.2148 (0.0184)
≥ 0.05	EBV	Additive	BA	0.4999 (0.0035)	0.5175 (0.0044)	0.0629 (0.0088)	0.8809 (0.0023)
			BB	0.5002 (0.0034)	0.5172 (0.0045)	0.0616 (0.0089)	0.8808 (0.0024)
			BC	0.4997 (0.0036)	0.5166 (0.0042)	0.0619 (0.0085)	0.8810 (0.0023)
			BL	0.4999 (0.0035)	0.5164 (0.0039)	0.0628 (0.0082)	0.8812 (0.0023)
			BRR	0.4998 (0.0034)	0.5164 (0.0042)	0.0604 (0.0086)	0.8811 (0.0023)
		Additive + Dominance	BA	0.4995 (0.0037)	0.5172 (0.0045)	0.0638 (0.0073)	0.8821 (0.0023)
			BB	0.4996 (0.0035)	0.5167 (0.0042)	0.0612 (0.0082)	0.8819 (0.0022)
			BC	0.4995 (0.0036)	0.5174 (0.0040)	0.0628 (0.0081)	0.8821 (0.0022)
			BL	0.4996 (0.0035)	0.5173 (0.0041)	0.0641 (0.0075)	0.8815 (0.0023)
			BRR	0.4995 (0.0036)	0.5174 (0.0040)	0.0633 (0.0068)	0.8824 (0.0022)
	Phenotype	Additive	BA				0.2183 (0.0218)
			BB				0.2207 (0.0210)
			BC				0.2196 (0.0205)
			BL				0.2151 (0.0194)
			BRR				0.2200 (0.0204)
		Additive + Dominance	BA				0.2096 (0.0205)
			BB				0.2117 (0.0205)
			BC				0.2125 (0.0212)
			BL				0.2066 (0.0210)
			BRR				0.2124 (0.0214)

[1]MAF: minimum minor allele frequency of the SNP whose effects were accounted for by the models (≥ 1%, 16,670 SNP; ≥ 5%, 9,987 SNP)
[2]The 5 models analysed the phenotype or the estimated breeding value (EBV) for the square root of serum cortisol concentration (square root of the ng of cortisol per ml of serum)
[3]BA: BayesA; BB; BayesB; BC: BayesC; BL; Bayesian LASSO; BRR: Bayesian Ridge Regression
[4]AUC: area under the ROC curve; ACC: accuracy computed as (true positives + true negatives)/number of samples; MCC: Matthew's correlation; CORR: correlation between the observed value and the model prediction

Table 4. Accuracy (CORR) when predicting the phenotype or the estimated breeding value (EBV) for antibody titer with five Bayesian models exploiting genotypes of a dense SNP panel and classification performance (AUC, ACC, MCC) of the model predictions of antibody titer EBV when used to classify the phenotype for VNN mortality in 16 independent 5-fold cross-validations

SNP MAF[1]	Predicted value[2]	Genetic effect fitted	Model[3]	Average (SD) of 16 cross-validations[4]			
				AUC	ACC	MCC	CORR
≥ 0.01	EBV	Additive	BA	0.4769 (0.0040)	0.5042 (0.0016)	0.0274 (0.0123)	0.7605 (0.0056)
			BB	0.4766 (0.0039)	0.5040 (0.0020)	0.0258 (0.0117)	0.7601 (0.0056)
			BC	0.4771 (0.0041)	0.5038 (0.0016)	0.0280 (0.0121)	0.7609 (0.0053)
			BL	0.4770 (0.0043)	0.5041 (0.0021)	0.0344 (0.0166)	0.7606 (0.0067)
			BRR	0.4769 (0.0041)	0.5046 (0.0013)	0.0321 (0.0121)	0.7611 (0.0055)
		Additive + Dominance	BA	0.4767 (0.0033)	0.5042 (0.0018)	0.0301 (0.0128)	0.7629 (0.0046)
			BB	0.4767 (0.0035)	0.5042 (0.0023)	0.0279 (0.0138)	0.7631 (0.0048)
			BC	0.4769 (0.0033)	0.5046 (0.0023)	0.0276 (0.0158)	0.7630 (0.0047)
			BL	0.4759 (0.0037)	0.5031 (0.0012)	0.0294 (0.0141)	0.7621 (0.0051)
			BRR	0.4771 (0.0035)	0.5045 (0.0022)	0.0300 (0.0135)	0.7632 (0.0046)
	Phenotype	Additive	BA				0.2664 (0.0173)
			BB				0.2660 (0.0172)
			BC				0.2673 (0.0168)
			BL				0.2612 (0.0183)
			BRR				0.2684 (0.0165)
		Additive + Dominance	BA				0.2608 (0.0165)
			BB				0.2591 (0.0155)
			BC				0.2607 (0.0159)
			BL				0.2566 (0.0160)
			BRR				0.2614 (0.0168)
≥ 0.05	EBV	Additive	BA	0.4776 (0.0034)	0.5051 (0.0026)	0.0395 (0.0175)	0.7555 (0.0077)
			BB	0.4773 (0.0034)	0.5045 (0.0024)	0.0415 (0.0169)	0.7548 (0.0077)
			BC	0.4778 (0.0036)	0.5051 (0.0023)	0.0421 (0.0151)	0.7561 (0.0077)
			BL	0.4783 (0.0034)	0.5054 (0.0028)	0.0392 (0.0128)	0.7539 (0.0096)
			BRR	0.4779 (0.0034)	0.5055 (0.0027)	0.0399 (0.0160)	0.7561 (0.0076)
		Additive + Dominance	BA	0.4774 (0.0033)	0.5055 (0.0033)	0.0433 (0.0164)	0.7570 (0.0073)
			BB	0.4772 (0.0033)	0.5052 (0.0032)	0.0427 (0.0173)	0.7568 (0.0076)
			BC	0.4776 (0.0032)	0.5049 (0.0024)	0.0441 (0.0153)	0.7577 (0.0075)
			BL	0.4779 (0.0040)	0.5060 (0.0040)	0.0396 (0.0135)	0.7557 (0.0078)
			BRR	0.4776 (0.0031)	0.5057 (0.0032)	0.0431 (0.0149)	0.7579 (0.0075)
	Phenotype	Additive	BA				0.2635 (0.0164)
			BB				0.2646 (0.0166)
			BC				0.2656 (0.0165)
			BL				0.2597 (0.0175)
			BRR				0.2660 (0.0164)
		Additive + Dominance	BA				0.2571 (0.0174)
			BB				0.2562 (0.0190)
			BC				0.2570 (0.0155)
			BL				0.2544 (0.0144)
			BRR				0.2580 (0.0161)

[1]MAF: minimum minor allele frequency of the SNP whose effects were accounted for by the models (≥ 1%, 16,670 SNP; ≥ 5%, 9,987 SNP)
[2]The 5 models analysed the phenotype or the estimated breeding value (EBV) for antibody titer (sample to positive ratio of the optical density values, OD 450 nm)
[3]BA: BayesA; BB; BayesB; BC: BayesC; BL; Bayesian LASSO; BRR: Bayesian Ridge Regression
[4]AUC: area under the ROC curve; ACC: accuracy computed as (true positives + true negatives)/number of samples; MCC: Matthew's correlation; CORR: correlation between the observed value and the model prediction

References

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

Bertotto D, Poltronieri C, Negrato E, Majolini D, Radaelli G, Simontacchi C (2010). Alternative matrices for cortisol measurement in fish. *Aquacult Res* **41**: 1261-1267.

Castanheira MF, Conceição L, Millot S, Rey S, Bégout M-L, Damsgård B, Kristiansen T, Höglund E, Øverli Ø, Martins CIM (2015). Coping styles in farmed fish: consequences for aquaculture. *Rev Aquacult* **9**: 23-41.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013). Stacks: an analysis tool set for population genomics. *Mol Ecol* **22**: 3124-3140.

Chavanne H, Janssen K, Hofherr J, Contini F, Haffray P, Komen H, Nielsen EE, Bargelloni L (2016). A comprehensive survey on selective breeding programs and seed market in the European aquaculture fish industry. *Aquacult Int* **24**: 1287-1307.

Doan QK, Vandeputte M, Chatain B, Haffray P, Vergnet A, Breuil G, Allal F (2017). Genetic variation of resistance to Viral Nervous Necrosis and genetic correlations with production traits in wild populations of the European sea bass (*Dicentrarchus labrax*). *Aquaculture* **478**: 1-8.

Dupont-Nivet M, Vandeputte M, Vergnet A, Merdy O, Haffray P, Chavanne H, Chatain B (2008). Heritabilities and GxE interactions for growth in the European sea bass (*Dicentrarchus labrax* L.) using a marker-based pedigree. *Aquaculture* **275**: 81-87.

Fawcett T (2006). An introduction to ROC analysis. *Pattern Recognit Lett* **27**: 861-874.

Gianola D (2013). Priors in whole-genome regression: the Bayesian alphabet returns. *Genetics* **90**: 525-540.

Goddard ME, Hayes BJ (2009). Mapping genes for complex traits in domestic animals and their use in breeding programs. *Nat Rev Genet* **10**: 381-391.

Gjedrem T, Baranski M (2009). Selection methods, pp. 93-102. In: *Selective breeding in aquaculture: an introduction,* Gjedrem T. and M. Baranski (Eds.) Springer Dordrecht Heidelberg London New York.

Habier D, Fernando R, Kizilkaya K, Garrick D (2011). Extension of the Bayesian alphabet for genomic Selection. *BMC Bioinformatics* **12**: 186.

Huisman J (2017). Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. *Mol Ecol Res* **17**: 1009-1024.

Iwamoto T, Nakai T, Mori K, Arimoto M, Furusawa I (2000). Cloning of the fish cell line SSN- 1 for piscine nodaviruses. *Dis Aquat Organ* **43**: 81-89.

Janssen K, Chavanne H, Berentsen P, Komen H (2016). Impact of selective breeding on European aquaculture. *Aquaculture* **472**: 8-16.

Kalinowski ST, Taper ML, Marshall TC (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* **16**: 1099-1106.

Legarra A, Varona L, López de Maturana E (2008). TM Threshold Model. http://snp.toulouse.inra.fr/~alegarra/

MacKenzie S, Ribas L, Pilarczyk M, Capdevila DM, Kadri S, Huntingford FA (2009). Screening for coping style increases the power of gene expression studies. *PLoS One* **4:** e5314.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* **7**: 639-655.

Matthews B (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim biophys acta* **405**: 442-451.

Meuwissen THE, Hayes BJ, Goddard ME (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**: 1819-1829.

Meuwissen THE, Hayes BJ, Goddard ME (2016). Genomic selection: a paradigm shift in animal breeding. *Animal Frontiers* **6:** 6.

Munday BL, Kwang J, Moody N (2002). Betanodavirus infections of teleost fish: a review. *J Fish Dis* **27**: 127-142.

Nuñez-Ortiz N, Stocchi V, Toffan A, Pascoli F, Sood N, Buonocore F, Picchietti S, Papeschi C, Taddei AR, Thompson KD, Scapigliati G (2016). Quantitative immunoenzymatic detection of viral encephalopathy and retinopathy virus (betanodavirus) in sea bass Dicentrarchus labrax. *J Fish Dis* **39**: 821-831.

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

Ødegård J, Meuwissen THE (2014). Identity-by-descent genomic selection using selective and sparse genotyping. *Genet Sel Evol* **46:** 3.

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

Palaiokostas C, Cariou S, Bestin A, Bruantn JS, Haffray P, Morin T, Cabon J, Allal F, Vandeputte M, Houston RD (2018a). Genome-wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (*Dicentrarchus labrax*) using RAD sequencing. *Genet Sel Evol* **50**: 30.

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

Park T, Casella G (2008). The Bayesian Lasso. *J Am Stat Assoc* **103**: 681-686.

Pérez P, de Los Campos G (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **198**: 483-495.

Pottinger TG (2008). The stress response in fish-mechanisms, effects and measurement, pp. 32-44. In: *Fish Welfare,* Branson EJ (Ed.) Blackwell Publishing Ltd, Oxford.

Robledo D, Palaiokostas C, Bargelloni L, Martínez P, Houston R (2017). Applications of genotyping by sequencing in aquaculture breeding and genetics. *Reviews in Aquaculture* **10**: 670-682.

Sargolzaei M, Chesnais JP, Shenkel FS (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* **15**: 478-489.

Scapigliati G, Buonocore F, Randelli E, Casani D, Meloni S, Zarletti G, Tiberi M, Pietretti D, Boschi I, Manchado M, Martin-Antonio B, Jimenez-Cantizano R, Bovo G, Borghesan F, Lorenzen N, Einer-Jensen K, Adams A, Thompson K, Alonso C, Bejar J, Cano I, Borrego JJ, Alvarez MC (2010). Cellular and molecular immune responses of the sea bass (*Dicentrarchus labrax*) experimentally infected with betanodavirus. *Fish Shellfish Immunol* **28**: 303-311.

Simontacchi C, Bongioni G, Ferasin L, Bono G (1995). Messa a punto di un metodo RIA su micropiastra per il dosaggio diretto del progestrerone ematico. Atti XLIX Convegno Nazionale S.I.S.Vet., pp. 343-344.

Sing T, Sander O, Beerenwinkel N, Lengauer T (2005). ROCR: visualizing classifier performance in R. *Bioinformatics* **21**: 3940-3941.

Sonesson AK, Meuwissen THE (2009). Testing strategies for genomic selection in aquaculture breeding programs. *Genet Select Evol* **41**: 37.

Tine M, Kuhl H, Gagnaire P-A, Louro B, Desmarais E, Martins RST, Hecht J, Knaust F, Belkhir K, Klages S, Dieterich R, Stueber K, Piferrer F, et al. (2014). The European sea bass genome and its variation provide insight into adaptation to euryhalinity and marine speciation. *Nat Commun* **5**: 5770.

Tsai H, Hamilton A, Tinch AE, Guy DR, Bron JE, Taggart JB, Gharbi K, Stear M, Matika O, Pong-Wong R, Bishop SC, Houston RD (2016). Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. *Genet Sel Evol* **48**: 47.

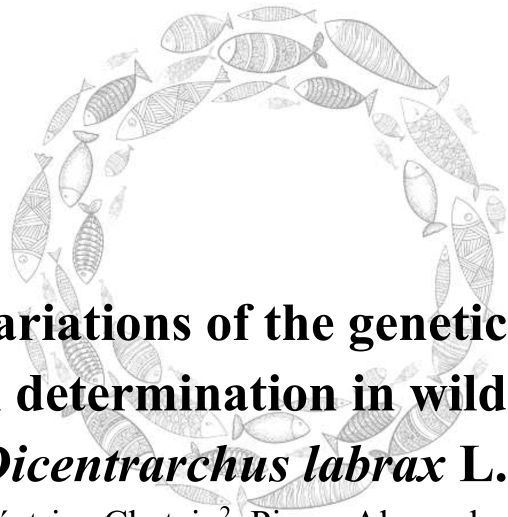
Vallejo R, Leeds T, Gao G, Parsons J, Martin K, Evenhuis J, Fragomeni BO, Wiens GD, Palti Y (2017). Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture. *Genet Sel Evol* **49**: 17.

Vandeputte M, Porte JD, Auperin B, Dupont-Nivet M, Vergnet A, Valotaire C, Claireaux G, Prunet P, Chatain B (2016). Quantitative genetic variation for post-stress cortisol and swimming performance in growth-selected and control populations of European sea bass (*Dicentrarchus labrax*). *Aquaculture* **455:** 1-7.

Vendramin N, Toffan A, Mancin M, Cappellozza E, Panzarin V, Bovo G, Cattoli G, Capua I, Terregino C (2014). Comparative pathogenicity study of ten different betanodavirus strains in experimentally infected European sea bass, *Dicentrarchus labrax* (L.). *J Fish Dis* **37**: 371-383.

Volckaert FAM, Hellemans B, Batargias C, Louro B, Massault C, Van Houdt JKJ, Haley C, de Koning D-J, Canario AVM (2012). Heritability of cortisol response to confinement stress in European sea bass *Dicentrarchus labrax. Genet Sel Evol* **44:** 15.

Wang S, Meyer E, McKay JK, Matz MV (2012). 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nat Methods* **9**: 808-810.



5. Population-specific variations of the genetic architecture of sex determination in wild European sea bass *Dicentrarchus labrax* L.

Sara Faggion[1,2], Marc Vandeputte[2,3], Béatrice Chatain[2], Pierre-Alexandre Gagnaire[4] and François Allal[2]

[1] Department of Comparative Biomedicine and Food Science (BCA), University of Padova, Italy
[2] MARBEC, Univ. Montpellier, Ifremer, CNRS, IRD, Palavas-les-Flots, France
[3] GABI, INRA, AgroParisTech, Université Paris-Saclay, 78350, Jouy-en-Josas, France
[4] ISEM, University of Montpellier, CNRS, IRD, EPHE, UM, 34095, Montpellier, France

Published in the journal Heredity (Springer Nature)
https://doi.org/10.1038/s41437-018-0157-z

Oral presentation, AQUA 2018 conference (European Aquaculture Society and World Aquaculture Society conference), Montpellier (FR), August, 25th - 29th, 2018

Abstract

Polygenic sex determination (PSD) may show variations in terms of genetic and environmental components between populations of fish species exposed/adapted to different environments. The European sea bass (*Dicentrarchus labrax*) is an interesting model, combining both a PSD system and a genetic subdivision into an Atlantic and a Mediterranean lineage, with genetic substructures within the Mediterranean Sea. Here, we produced experimental progeny crosses (N = 927) from broodstock sampled in four wild populations (North Atlantic, NAT; Western Mediterranean, WEM; North-Eastern Mediterranean, NEM; South-Eastern Mediterranean, SEM). We found less females than males in the progeny, both in the global dataset (32.5%) and within each paternal group (from 25.1% for NEM to 39.0% for WEM), with significant variation among populations, dams and sires. Sex, body weight (BW) and body length (BL) showed moderate heritability (0.52 ± 0.17, 0.46 ± 0.17, 0.34 ± 0.15, respectively) and sex was genetically correlated with BW and BL ($r_{a\ sex/BW}$ = 0.69 ± 0.12, $r_{a\ sex/BL}$ = 0.66 ± 0.13).

A weighted GWAS performed both on the global dataset and within each paternal group revealed a different genetic architecture of sex determination between Atlantic and Mediterranean populations (9 QTLs found in NAT, 7 in WEM, 5 in NEM and 4 in SEM, with a cumulated variance explained of 27.04%, 21.87%, 15.89% and 12.10%, respectively) with a more similar genetic architecture among geographically close populations compared to geographically distant populations, consistent with the hypothesis of a population-specific evolution of polygenic sex determination systems in different environments.

Keywords European sea bass, polygenic sex determination, wGWAS, heritability, genetic correlations

5.1 Introduction

Sex determination is a mechanism of major evolutionary importance that exhibits a high variety of modalities, and this is especially true in fish (Mank et al. 2006; Heule et al. 2014). These modalities are generally classified as genotypic sex determination systems (or GSD, which includes male/female heterogamety and polygenic sex determination, or PSD) and environmental sex determination systems (ESD; Bull 1983).

PSD has been initially formalised by Bulmer and Bull (1982): they proposed an underlying "sex tendency" phenotype, with a polygenic determinism influencing the observed phenotype (male or female) depending on whether it lies below or beyond a fixed threshold. Under this model, any environmental or genotypic effect can equally affect the phenotype (here sex tendency) and bring its value below or beyond the threshold, therefore determining sex.

In the classical view, polygenic sex determination is thought to be unstable, and should evolve either towards GSD in a fluctuating environment generating biased sex-ratios, or towards ESD if some environments increase the fitness of a specific sex (Bulmer and Bull 1982). Modelling approaches have shown that the orientation of polygenic sex determination in one or the other direction depends on complex combinations of environmental variation between and within environmental patches, and on migration rates between patches (Van Dooren and Leimar 2003; Bateman and Anholt 2017). Therefore, starting with the same ancestral polygenic sex determination system, we may hypothesize that there could be a population-specific evolution of the sex determination mechanism, in other words the components of the model may balance differently between sub-populations of the same species that are exposed to different environmental conditions (Vandeputte et al. 2012; Guinand et al. 2017).

In teleost fish, polygenic sex determination has been well-documented in the Atlantic silverside, *Menidia menidia* (Conover and Heins 1987), some populations of the zebrafish, *Danio rerio* (Liew et al. 2012; Wilson et al. 2014) and in the European sea bass, *Dicentrarchus labrax* (Vandeputte et al. 2007; Palaiokostas et al. 2015). However, some authors consider that it may be more frequent than classically thought (Moore and Roberts 2013) since it is difficult to characterize.

The European sea bass (*Dicentrarchus labrax* L.) offers an interesting model to investigate the evolution of polygenic sex determination. This species combines both a polygenic sex determination system (Vandeputte et al. 2007; Palaiokostas et al. 2015) and a clear genetic subdivision into an Atlantic and a Mediterranean lineage, in addition to population genetic structure within the Mediterranean Sea (Naciri et al. 1999; Bahri-Sfar et al. 2000; Lemaire et al. 2005; Quéré et al. 2012; Tine et al. 2014).

Experimental evidence show that early life rearing temperature has a strong influence on the sex-ratio of sea bass, with low rearing temperature (between 13 and 17 °C) during the first 60 days of life, favouring the production of female offspring (reviewed by Vandeputte and Piferrer 2018), which are preferred in aquaculture due to their higher growth rate (Saillant et al. 2001; Felip et al. 2006). In the natural environment, observed variation in yearly cohort sex-ratio may be indicative of natural variations in temperature influencing sex determination also in the wild populations of sea bass (Vandeputte et al. 2012).

Variation of sex-ratios between families in sea bass is consistent with the hypothesis that phenotypic sex is determined by an underlying continuous sex tendency combining the effects of polygenes and temperature (Vandeputte et al. 2007). However, at least three genome-wide significant sex determining quantitative trait loci (QTLs) were recently found by Palaiokostas et al. (2015), showing

that some specific loci could have a stronger effect on sex tendency, at least in some populations (in this case the western Mediterranean).
Here, we hypothesize that the sex determination system of European sea bass could have a different genetic architecture in different populations exposed to different environmental conditions (especially temperature) in the wild. Experimental progeny crosses were produced from broodstock sampled in four wild populations (corresponding to the whole distribution of natural populations, from Northern Atlantic to Eastern Mediterranean Sea). We estimated the additive genetic variation for sex tendency and the correlation between sex tendency and growth-related traits (weight and length) and performed a weighted genome-wide association study (wGWAS), both on the global dataset and within each group of paternal origin, to assess possible population-specific variations in the architecture of sex determination in sea bass.

5.2 Materials and methods

5.2.1 Broodstock origin, production and rearing of experimental fish

The male broodstock used in this study belonged to four different origins, matching with most of the natural range of the species: North Atlantic (NAT), Western Mediterranean (WEM), North-Eastern Mediterranean (NEM) and South-Eastern Mediterranean (SEM). The female broodstock belonged to the WEM population. The origin and collection of the broodstock has been detailed by Vandeputte et al. (2014).
The NAT and WEM sires were reared at the IFREMER facilities of Palavas-les-Flots (France) and the sperm stripped and cryopreserved following Fauvel et al. (1998), while the sperm of SEM and NEM sires were both cryopreserved in 2005 at IOLR, Eilat, Israel (SEM sperm) and at Beymelek Lagoon, Turkey (NEM sperm) following the protocol of Sansone et al. (2002). The dams were reared at the IFREMER station of Palavas-les-Flots under natural photoperiod which was 11L:13D at the time of spawning (3rd of March, 2014) and natural temperature, which decreased from 23.5 °C in August 2013 to 13.5 °C at the time of spawning; dams with a suitable stage of development of eggs (determined after ovarian biopsy) were hormonally injected and stripped 72 h after the injection. Artificial fertilization was performed at the IFREMER station, using a full factorial mating scheme: 15 sires per origin (60 sires in total) were crossed with 9 WEM dams. The fertilization protocol and the rearing of experimental fish were described previously by Doan et al. (2017a). Briefly, after hatching, larvae were reared in a common garden at a temperature of 16.5 °C and 25‰ salinity until 58 d post-hatching; the following seven days the temperature was gradually increased to 20 °C. Fish were then reared at a mean temperature of 21.5 °C (18.1-22.4 °C) and 30‰ salinity until 102 d post-hatching. Afterwards, fish were divided into five juvenile tanks (A, B, C, D, E), with a mean temperature of 22.1 °C (15.5-27.9 °C). Fish were fed using a classical hatchery feeding sequence (*Artemia* nauplii, Le Gouessant Marine Start and Neo Start pellets). At 180 d post-hatching, fish were individually tagged and measured for body weight (BW) and fork length (BL). At the same time, fin samples were collected for genomic DNA extraction. At 226 d post-hatching, 927 randomly chosen experimental fish were euthanized with an overdose of benzocaine, dissected, and the sex was recorded by visual observation of the gonads or using the squash technique (Menu et al. 2005) when macroscopic observation was ambiguous. The reliable identification of the phenotypic sex was possible for all 927 fish. Sex was coded as a binary trait, 1 for males and 2 for females.

5.2.2 Genotyping, parentage assignment and descriptive statistics

Fin clips from the 927 experimental fish, from the 60 sires and from the 9 dams were sent to LABOGENA (Jouy-en-Josas, France) for genomic DNA extraction and genotyping. Genotyping was performed with an iSelect Custom Infinium Illumina® European sea bass 3K SNP array. The design of this SNP array was done by selecting 2,722 SNPs from an initial genome-wide variation map containing 2,628,725 SNPs phased into chromosome-wide haplotypes. These SNPs were discovered from 14 wild individuals from both the Atlantic and Mediterranean areas, using whole-genome sequencing as described by Duranton et al. (2018). A first filtering of the SNPs was made to remove variants closer than 80 bp from another known variant. Then, as recommended by Illumina®, we filtered the remaining SNPs to avoid A/T and C/G variants, that need a particular design (Infinium I Probe Design) involving setting up two probes instead of one. Among the remaining candidates, SNPs were chosen to cover all the chromosomes with a variable SNP density depending on the local nucleotide diversity (π), as reported by Tine et al. (2014). To do so, five π-classes were defined depending on the π estimated in non-overlapping 50kb windows: class 1 for $\pi < 10^{-3}$; class 2 for $10^{-3} < \pi < 2\cdot10^{-3}$; class 3 for $2\cdot10^{-3} < \pi < 3\cdot10^{-3}$; class 4 for $3\cdot10^{-3} < \pi < 4\cdot10^{-3}$; class 5 for $\pi > 4\cdot10^{-3}$. Based on "best quality criterion", one SNP was selected in class 1 windows (in interval 22.5-27.5 kb of the window), 2 SNPs in class 2 windows (in intervals 12.5-17.5 kb and 35-40 kb), 3 SNPs in class 3 windows (in intervals 5.8-10.8 kb, 22.5-27.5 kb and 39-44 kb), 4 SNPs in class 4 windows (in intervals 3.7-8.7 kb, 16.2-21.2 kb, 28.7-33.7 kb and 41.2-46.2 kb) and 5 SNPs in class 5 windows (in intervals 2.5-7.5 kb, 12.5-17.5 kb, 22.5-27.5 kb, 32.5-37.5 kb and 42.5-47.5 kb). Since nucleotide diversity is negatively correlated to the local recombination rate in sea bass (which was estimated by Tine et al. 2014), this local adjustment in the density of SNPs aimed at homogenizing the density of markers along the recombination map instead of the physical map.

Parentage assignment was performed with an exclusion-based software, VITASSIGN (Vandeputte et al. 2006), using 2,722 markers and allowing 29 allelic mismatches to recover pedigree.

Proportion of individuals, males and females in the global dataset, per origin, per tank, per dam and per sire, pairwise comparisons and χ^2 tests were performed in R version 3.4.3 using the packages *stats* (R Core Team 2017) and *gmodels* (Warnes et al. 2015). All *P*-values were adjusted for multiple testing with the Bonferroni correction method.

5.2.3 Principal component analysis

To describe the overall genetic structure among the 927 individuals genotyped on the basis of genome-wide SNP data, we performed a principal component analysis (PCA) using *–pca* function in PLINK (Purcell et al. 2007). A two-dimension scatter plot of individuals coordinates on the two first principal components was generated, indicating the percentages of variance explained.

5.2.4 Heritability, genetic and phenotypic correlations

Heritability was estimated on the entire dataset through a linear mixed sire model using the software VCE 6.0 (Groeneveld et al. 2010). The model was the following:

$$y_{ijkl} = o_i + t_j + s_{k(i)} + d_l + e_{ijkl}$$

where y_{ijkl} is the phenotype for the studied trait (coded as a binary trait, 1 for male and 2 for female in the case of sex); o_i is the fixed effect of the population of origin of the sires i; t_j is the fixed effect of the rearing tank j; $s_{k(i)}$ is the random additive genetic effect of sire k within origin i; d_l is the random effect of dam l; e_{ijkl} is the random residual.

As explained by Falconer and Mackay (1996), the sire component accounts for ¼ of the additive genetic variance; for this reason, the heritability was estimated as $h^2 = 4\sigma_s^2/\sigma_p^2$, with σ_s^2 being the sire component of variance and σ_p^2 the phenotypic variance. When the trait was sex, heritability on the observed (binary) scale was transformed to the value on the underlying liability scale (Dempster and Lerner 1950; Lynch and Walsh 1998) following the formula:

$$h_u^2 = h_o^2\, p\, (1\text{-}p) / z^2$$

where h_u^2 is the heritability on the liability scale, h_o^2 is the heritability on the observed scale, p is the incidence (proportion of females) in the population and z is the value of the normal distribution density at the point where the cumulative distribution function of the normal distribution reaches incidence.

Genetic and phenotypic correlations between sex and growth-related traits (body weight and length at 180 d post-hatching) were assessed using VCE 6.0 software (Groeneveld et al. 2010) applying a three traits sire model with sex, weight and length as variables.

5.2.5 Genome-wide association study (GWAS)

GWAS was performed through the BLUPf90 family of programs for mixed-model computations (Misztal et al. 2015) in order to identify possible SNPs associated with phenotypic sex.

Owing to the genetic subdivision between Atlantic and Mediterranean sea bass lineages and the finer scale differentiation between Western and Eastern Mediterranean populations (Naciri et al. 1999; Bahri-Sfar et al. 2000; Lemaire et al. 2005; Quéré et al. 2012; Tine et al. 2014; Duranton et al. 2018), the dataset was split into four groups depending on paternal origin (NAT, WEM, NEM and SEM) and the analyses were performed separately within each group. The origin of the sire was taken into account as a fixed effect when we performed GWAS on the global dataset.

The raw SNP dataset was quality-filtered before running the GWAS. We firstly removed 8 individuals with a call rate lower than 0.9. We then applied variant filters to exclude SNPs showing either significant Mendelian distortions, a minor allele frequency (MAF) lower than 0.05, or a proportion of missing genotypes greater than 0.1. This resulted in a final dataset containing 1,205 retained markers out of the total 2,722 SNPs which were genotyped in 919 offspring and 69 parents (988 individuals in total). The high number of discarded SNPs is mainly due to technical design problems, which lead to target the wrong (i.e. non variable) base position for 50% of the markers. The mapping of the SNPs used for further analyses was thus reported by Doan (2017b), reconstructed from the genotypes of the same individuals.

For the weighted GWAS (wGWAS) the following model was applied:

$$\mathrm{y} = \mathbf{X}\mathrm{b} + \mathbf{W}\mathrm{u} + \mathrm{e}$$

where y is the vector of phenotypes, b vector of the fixed effects (intercept, tank and origin of the sires), $\boldsymbol{X}$ the incidence matrix relating phenotypes with the fixed effects, $\boldsymbol{W}$ the incidence matrix relating phenotypes with the random animal effects, u the vector of random animal effects ~ N(0, $G\sigma_g^2$) with $\boldsymbol{G}$ being the genomic relationship matrix (VanRaden 2008), σ_g^2 the additive genetic variance, e the vector of residuals ~N(0, I) and the residual variance. The genomic relationship matrix $\boldsymbol{G}$ was established as follow:

$$\mathbf{G} = \mathbf{ZDZ'}/q$$

where ***D*** is a diagonal matrix with weights for SNP effects, ***Z*** is a matrix of gene content adjusted for allele frequencies and q is a weighting factor equal to $2\Sigma p_i (1 - p_i)$, where p_i is the MAF of SNP i.
The wGWAS was implemented through an iterative process (Zhang et al. 2010; Wang et al. 2012; Zhang et al. 2016) and using Gibbs sampling (THRGIBBS1F90) to estimate the *GEBV*s (genomic estimated breeding values) since this is specifically adapted to the analysis of binary traits. The following steps were performed:
1) in the first iteration, ***D*** was first set equal to ***I***, the identity matrix (VanRaden, 2008);
2) ***G*** was calculated as $\mathbf{G} = \mathbf{ZDZ'}/q$;
3) *DGV*s (direct genomic values) were obtained from *GEBV*s as $DGV_i = -(\Sigma_{j, j \neq i} g^{ij} GEBV^j / g^{ii})$ with g^{ij} elements of the $\mathbf{G}^{-1}$ matrix (Lourenco et al. 2015), and converted to SNP effects as $a_i = \mathbf{DZ'G^{-1}}DGV_i$;
4) new SNP weights over a 5 SNP window as $d_i = \Sigma_i a_i^2/5$ were calculated and normalized so that the total genetic variance remained constant;
5) the process was then iteratively repeated from step 2 with the new ***D*** matrix.
After 3 rounds there was no further modification of the variance explained by the SNPs (i.e. the correlation coefficient R^2 between round 1 and 2 was equal to 0.93, 0.97 between round 2 and 3 and 0.99 between round 3 and 4).
We calculated the regional variance explained by summing neighbouring SNP variance in overlapping windows of 5 adjacent SNPs, suggested by Habier et al. (2011) as the more appropriate method to infer the effect of QTLs. We considered as QTLs the genomic segments that explained a proportion of genetic variance higher than 2%.

5.3 Results

5.3.1 Parentage assignment and descriptive statistics

Assignment to a unique parental pair was achieved for all 927 offspring. The number of fish per paternal origin varied from a minimum of 150 (NAT) to a maximum of 328 individuals (SEM). The number of offspring per sire varied from 1 to 36 and the number of offspring per dam varied from 2 to 250.
The proportion of females in the global dataset was 32.5% and was variable among groups of paternal origin ($\chi^2 = 12.27$, df = 3, *P*-value = $7 \cdot 10^{-3}$), ranging from 25.1% for NEM to 39.0% for WEM (Table 1). Pairwise comparisons showed significant sex-ratio differences between NAT and NEM and between WEM and NEM (Table 1).
There were neither significant differences in the proportion of males and females between the five rearing tanks ($\chi^2 = 9.48$, df = 4, *P*-value = $5.02 \cdot 10^{-2}$) nor in the proportion of animals belonging to the four paternal origins between tanks ($\chi^2 = 16.37$, df = 12, *P*-value = $17.47 \cdot 10^{-2}$; Table 2). On the contrary, the proportion of females in the offspring strongly differed per sire and per dam ($\chi^2 = 124.56$, df = 59, *P*-value = 10^{-6} for sires, $\chi^2 = 34.72$ df = 8, *P*-value = $3 \cdot 10^{-6}$ for dams; Figure 1). The proportion of females ranged from 0 to 100% in paternal half-sib families and from 0 to 55% in maternal half-sib families.
Female offspring were, on average, heavier and longer than males at 180 d post-hatching; this was true both for the global dataset and within groups of paternal origin (Table 1).

5.3.2 Principal component analysis

The principal component analysis performed on the global dataset revealed that genotypic variance was mainly explained by paternal origin (first principal component explaining 5.9% of the variance, Figure 2a). Furthermore, variations among dams were also detected (second axis explaining 4.1% of the variance, Figure 2b).

More precisely, the first PC axis distinguished three groups corresponding to the population of origin of the sires; North Atlantic, Western and Eastern Mediterranean groups were clearly separated, while the difference between North and South Eastern Mediterranean group was more subtle. This stratification explained by the population of origin of the sires was properly taken into account as a fixed effect in the models used to estimate heritability and to perform the wGWAS on the global dataset.

Table 1. Number of individuals (N), sex-ratio, mean body weight (BW, g) and length (BL, mm) at 180 d post-hatching (globally and separately for males and females) with the coefficient of variation expressed in percentage (CV%); the results are showed for the global dataset and separately per group of paternal origin; different letters indicate a significant difference at $P \leq 0.05$

		Paternal origin				
		NAT N = 150	**WEM** N = 182	**NEM** N = 267	**SEM** N = 328	**Global** N = 927
Sex-ratio	Proportion of males	0.620[a]	0.610[a]	0.749[a]	0.677[ab]	0.675
	Proportion of females	0.380[a]	0.390[a]	0.251[b]	0.323[ab]	0.325
Mean $BW_{180\ dph}$ (CV%)		17.57 (36.4%)	14.70 (30.3%)	16.22 (34.5%)	16.09 (34.7%)	16.09 (34.7%)
	Males	16.00 (38.4%)	13.02 (28.1%)	14.90 (31.5%)	14.19 (30.5%)	14.48 (32.7%)
	Females	20.11 (29.9%)	17.32 (25.0%)	20.17 (30.8%)	20.08 (28.9%)	19.46 (29.4%)
Mean $BL_{180\ dph}$ (CV%)		113.01 (11.4%)	108.89 (9.4%)	111.48 (10.5%)	110.46 (10.7%)	110.86 (10.6%)
	Males	109.46 (11.4%)	105.23 (8.9%)	108.90 (10.0%)	106.54 (9.4%)	107.50 (9.9%)
	Females	118.79 (9.6%)	114.62 (7.8%)	119.18 (9.0%)	118.68 (9.4%)	117.85 (9.1%)

Table 2. Proportion of individuals, sex-ratio and proportion of individuals belonging to the different groups of paternal origin (NAT, WEM, NEM and SEM) per each rearing tank (A, B, C, D, E); different letters indicate a significant difference at $P \leq 0.05$

		Tank				
		A	B	C	D	E
Proportion of individuals		0.223[a]	0.180[a]	0.219[a]	0.187[a]	0.191[a]
Sex-ratio	Proportion of males	0.667[ab]	0.587[a]	0.729[b]	0.676[ab]	0.706[ab]
	Proportion of females	0.333[ab]	0.413[a]	0.271[b]	0.324[ab]	0.294[ab]
Proportion of individuals per origin	NAT	0.097[a]	0.186[a]	0.172[a]	0.168[a]	0.198[a]
	WEM	0.256[a]	0.210[a]	0.172[a]	0.168[a]	0.169[a]
	NEM	0.266[a]	0.251[a]	0.315[a]	0.301[a]	0.305[a]
	SEM	0.382[a]	0.353[a]	0.340[a]	0.364[a]	0.328[a]

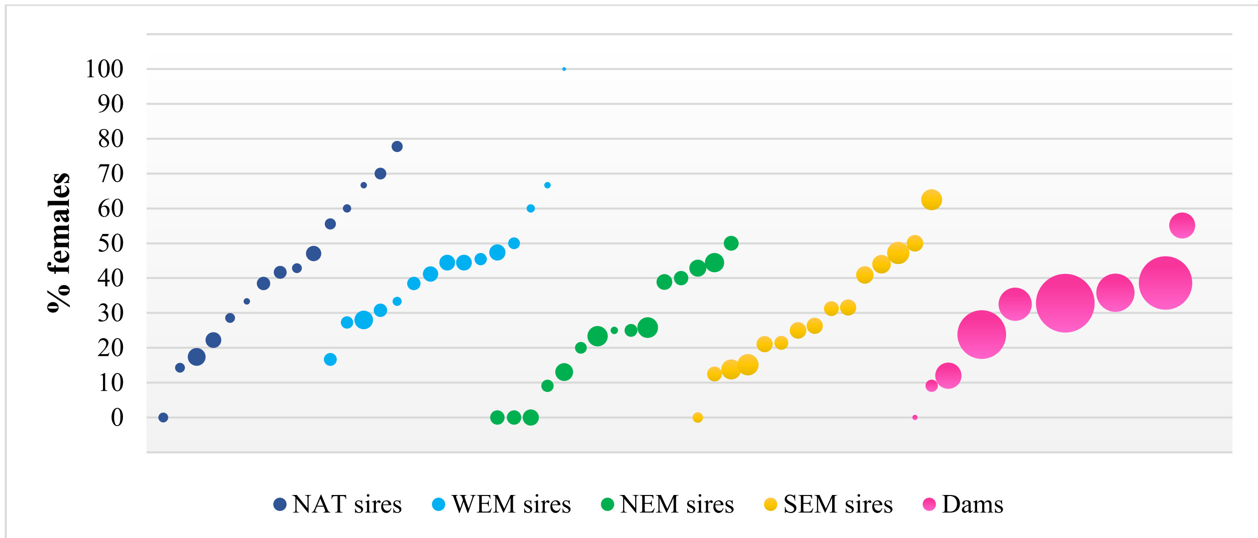


Figure 1. Percentage of females per sire and per dam; blue, light blue, green and yellow bubbles identify the different origins of the sires, the pink ones identify the dams; the size of the bubble represents the total number of offspring per sire/dam

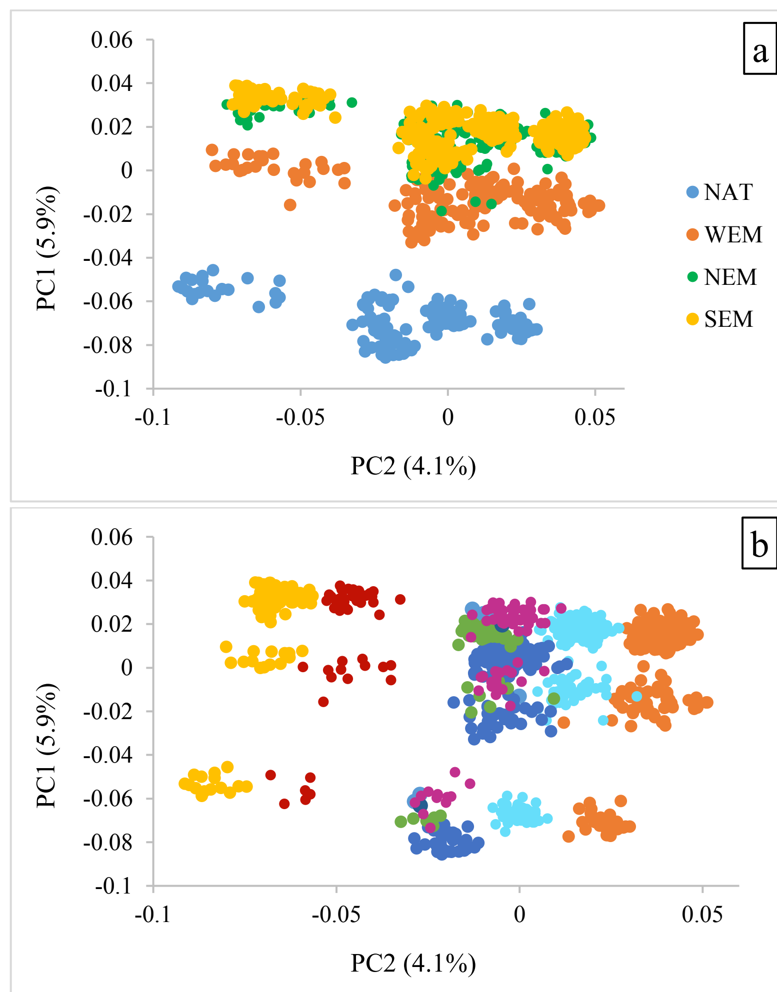


Figure 2. Two-dimension scatterplots showing the population stratification in the global dataset (N = 927) by paternal origin (a) and by dams (b). The first principal component was plotted against the second; the percentages of variance explained by each axis is indicated

5.3.3 Heritability, genetic and phenotypic correlations

Heritability was moderately high for all variables (sex, body weight and body length at 180 d post-hatching, Table 3). The genetic and phenotypic correlations between sex and growth-related traits were moderately high and the genetic correlations were higher compared to the phenotypic correlations (> 0.65 *vs.* 0.42). The genetic and phenotypic correlations between body weight and length were close to unity (Table 3).

Table 3. Heritability (± *s.e*, in bold on the diagonal) for sex (on the liability scale) and growth-related traits (body weight, BW, and body length, BL, at 180 d post-hatching), genetic (± *s.e.*; below the diagonal) and phenotypic correlations (above the diagonal) among traits, estimated with VCE 6.0 (Groeneveld et al. 2010)

Trait	Sex	$BW_{180\ dph}$	$BL_{180\ dph}$
Sex	**0.52 ± 0.17**	0.42	0.42
$BW_{180\ dph}$	0.69 ± 0.12	**0.46 ± 0.17**	0.97
$BL_{180\ dph}$	0.66 ± 0.13	0.99 ± 0.005	**0.34 ± 0.15**

Table 4. Identification of European sea bass chromosomes with a QTL explaining more than 2% of the variance in the global dataset and in each of the four offspring groups with the same paternal origin

	1A	1B	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18-21	19	20	22-25	24	*x*
Global			●				●	●					●			●								●
NAT		●		●		●			●	●	●		●				●			●				
WEM					●	●		●	●			●								●	●			
NEM						●				●	●	●				●								
SEM		●						●													●			●

Table 5. Total number of QTL for each paternal origin (in bold on the diagonal), total number of shared QTL between origins (below the diagonal) and total number of QTL that differ between origins (above the diagonal)

	NAT	WEM	NEM	SEM
NAT	**9**	10	8	11
WEM	3	**7**	8	7
NEM	3	2	**5**	9
SEM	1	2	0	**4**

5.3.4 Genome-wide association study (GWAS)

Results from the wGWAS performed on the global dataset identified one major group of SNPs on LG6 explaining up to 3.41% of the variance for sex. Other important groups of SNPs were detected on LG7, LG12, LG15 and LG*x*, explaining up to 2.73% of variance, while a minor group explaining slightly more than 2% of variance was located on LG2 (Figure 3a; Table 4).

The comparisons of the Manhattan plots of the wGWAS performed separately within each group of paternal origin (Figure 3b, 3c, 3d, 3e; Table 4), revealed a clear pattern of similarity between samples belonging to adjacent paternal origins. Nevertheless, taking into account genomic regions explaining at least 2% of the variance showed a very variable architecture of sex determination, with some peaks

being shared among populations, while others being clearly population-specific (Table 5 and Figure 4).

A group of SNPs on LG5 explained 4.37% of variance for sex in ♀WEM×♂NAT, 3.12% in ♀WEM×♂WEM and 4.6% in ♀WEM×♂NEM, while in ♀WEM×♂SEM this peak was not observed. Peaks shared only between ♀WEM×♂NAT and ♀WEM×♂WEM were identified on LG8, explaining 2.18% and 4.06% of the variance, respectively, and on LG19, explaining 4.25% and 2.80% of the variance, respectively. The crossings ♀WEM×♂NAT and ♀WEM×♂NEM shared two peaks on LG9 (2.17% and 2.38% of variance explained, respectively) and on LG10 (4.97% and 3.36% of variance explained, respectively).

One group of SNPs, which was shared between ♀WEM×♂NAT and ♀WEM×♂SEM was identified on LG1B, explaining 2.18% and 3.53%, respectively. One group of SNPs that was in common between ♀WEM×♂WEM and ♀WEM×♂SEM was identified on LG11 (3.48% and 2.54% of variance explained, respectively). Furthermore, ♀WEM×♂WEM cross share two groups of SNPs with ♀WEM×♂SEM cross, on LG7 (variance explained of 2.41% and 2.22%, respectively) and on LG20 (3.47% and 3.78% of the variance, respectively).

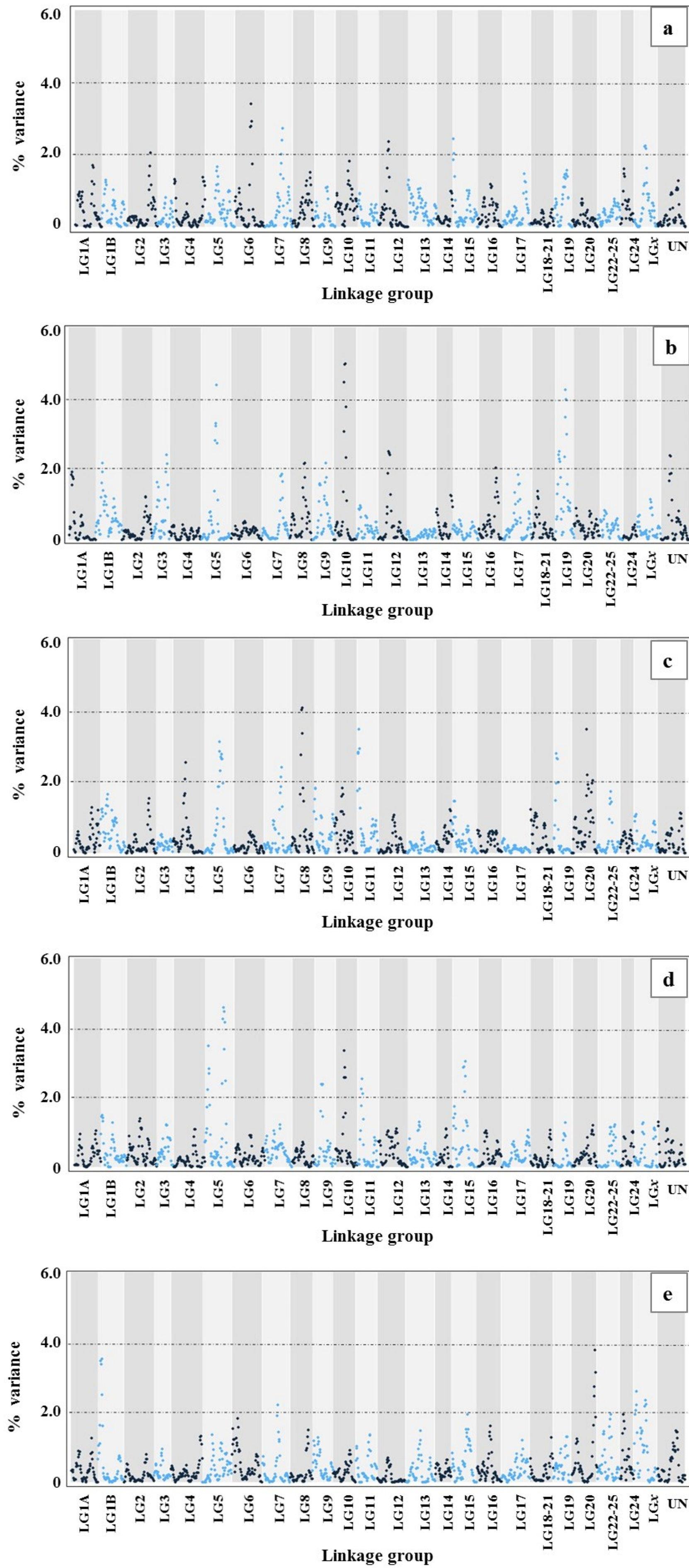
The ♀WEM×♂NAT cross showed specific peaks, that were not shared with any other paternal origin, on LG3 (2.41% of variance explained), LG12 (2.51%), LG16 (2.03%). One specific peak was found in the ♀WEM×♂WEM cross, as well (LG4, 2.55% of variance explained).

Two groups of SNPs were identified as specific to the Eastern Mediterranean populations, one in the ♀WEM×♂NEM cross (LG15, 3.04% of variance explained) and one in the ♀WEM×♂SEM cross (LG*x*, 2.59% of variance explained).

Interestingly, we did not identify any sex QTL explaining more than 2% of the variance that was common to all populations.

In the following page:

Figure 3. Manhattan plots showing the percentage of variance explained for sex. The results were obtained with a weighted GWAS in a sliding window of 5 adjacent SNPs in BLUPf90 (Misztal et al. 2015). a. global dataset; b. WEM×NAT; c. WEM×WEM; d. WEM×NEM; e. WEM×SEM



a
b
c
d
e
6.0
4.0
2.0
0
% variance
Linkage group
LG1A
LG1B
LG2
LG3
LG4
LG5
LG6
LG7
LG8
LG9
LG10
LG11
LG12
LG13
LG14
LG15
LG16
LG17
LG18-21
LG19
LG20
LG22-25
LG24
LGx
UN

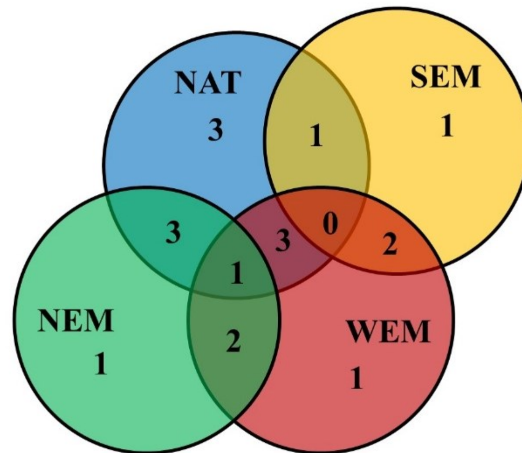


Figure 4. Venn diagram showing the number of QTLs for sex explaining more than 2% of the variance that were specific for each paternal group or that were shared between groups of paternal origin.

5.4 Discussion

In this study, we explored the genetic basis of the sex determination system in the European sea bass by implementing a genome-wide association study approach in a factorial crossing experiment. For the first time, sea bass belonging to different origins across the whole distribution range of natural populations were compared to assess variation in the genetic architecture of sex, including a comparison between the Atlantic and Mediterranean sea bass lineages. We found different QTLs underlying sex determination between Atlantic and Mediterranean populations, with a gradient of similarities from Western to Eastern Mediterranean populations, reflecting the previously documented introgression of Atlantic genes within the Mediterranean genetic background (Guinand et al. 2017, Duranton et al. 2018). This finding is consistent with the hypothesis of a population-specific evolution of polygenic sex determination systems in different evolutionary lineages occupying different environments.

An important result was the increased sharing of QTLs for sex determination in adjacent populations, which could result from an ongoing admixture between two evolutionary lineages (i.e. Atlantic and Mediterranean) characterized by different genetic architectures of sex determination systems. The detected geographical gradient in the architecture, from NAT to SEM, would then reflect the level of introgression and indeed corresponds to the admixture gradient recently found in sea bass population genomic studies (Duranton et al. 2018).

The ancestral architecture of the sex determination in sea bass might have evolved differently during the 300,000 years of divergence between Atlantic and Mediterranean lineages, explaining the origin of the variation that now has population-specific influences on sexual determination. Indeed, we did not find any linkage group common to all populations with groups of SNPs explaining more than 2% of the variance, which support the hypothesis put forward by Guinand et al. (2017) that the most important genes affecting sex may differ between sea bass populations.

♀WEM×♂NAT cross showed some similarities compared to ♀WEM×♂WEM and ♀WEM×♂NEM, that have gradually reduced in ♀WEM×♂SEM. This finding can be more likely explained by the recent history of inter-basins connectivity, since Atlantic alleles have been progressively diffused from the Western to the Eastern Mediterranean since the end of the last glacial maximum (Tine et al. 2014; Duranton et al. 2018). The resulting longitudinal gradient of admixture across the

Mediterranean populations makes the WEM population (31% of Atlantic ancestry) more similar to the Atlantic than the NEM and SEM population (13% of Atlantic ancestry) in most of the genome (Duranton et al. 2018).
Therefore, a gradient in similarity of genomic architecture is expected if sex determination QTLs introgress similarly to neutral genes. We do not reject, however, the possibility that differential adaptations between Atlantic and Mediterranean environments have also contributed to the patterns we observed, although this hypothesis is difficult to distinguish from historical admixture. Finally, the presence of a biogeographical barrier to gene flow located in the Siculo-Tunisian Strait (Quignard 1978; Bahri-Sfar et al. 2000), which limits the connections between Western and Eastern Mediterranean, may explain the further reductions of similarities between Western and Eastern Mediterranean populations.
In our case, the four paternal groups are all related by the WEM dams (i.e. all the individuals have 50% of the genome coming from WEM), with the result that even the Eastern Mediterranean group contains a higher level of Atlantic ancestry than what is expected in "pure" wild ♀NEM×♂NEM or ♀SEM×♂SEM crosses. This leads to the conclusion that the real differences existing in nature could be even stronger than what we observed here, due to our experimental design.
We interpreted our results as differences between male origins with the implicit assumption that they mostly reflect additive QTLs effects from the sire population of origin. Still, some QTLs could be due to the dam population (WEM). This is not the preferred hypothesis as no QTLs are shared between all paternal group, although they all share the same dams. Another possibility is that some of the QTLs observed are not linked to additive genetic variation but to dominance. The higher number of QTLs in the ♀WEM×♂NAT cross could be indicative of dominant alleles involved in sex determination (especially since some heterosis in sex-ratio has been shown by Guinand et al. 2017 when mating Atlantic and Mediterranean individuals).
Within each paternal origin, we performed a cross validation analysis of QTLs, by removing four times 25% of the offspring along the second axis of the PCA (representative of variation between dams, see Figure 2b). Most of the QTLs were identified in all subgroups, suggesting they were linked to the sire origin studied, but several of them were absent in some of the subgroups, which may be indicative of sire origin by dam interaction, i.e. dominance variation (see Supplementary material).
Parental effects on phenotypic sex were clearly significant: the dams (χ^2 = 34.72, df = 8, *P*-value = $3 \cdot 10^{-6}$) and sires (χ^2 = 124.56, df = 59, *P*-value = 10^{-6}) variation for the proportion of females in the offspring was strongly different, similar to Vandeputte et al. (2007), where both sires and dams had a similar-size effect on the sex-ratio of the progeny.
The heritability of sex tendency we estimated in the present study through a linear mixed sire model was relatively high (h^2 = 0.52 ± 0.17), similar to the estimate obtained for sire heritability by Vandeputte et al. (2007) on a larger dataset consisting of individuals of Northern Atlantic origin (0.52 ± 0.13), suggesting that the influence of the genetic and the environmental components on sex-ratio variance should be roughly equivalent.
The genetic correlation between sex and growth-related traits was significant (h^2 = 0.69 ± 0.12 between sex and weight, h^2 = 0.66 ± 0.13 between sex and length) and higher compared to previous studies (r_A between sex and size in the range of 0.23 and 0.59; Vandeputte et al. 2007; Palaiokostas et al. 2015). Overall, these results confirm the hypothesis of a strong link between genes affecting sex and growth (reviewed by Vandeputte and Piferrer 2018), with a clear sexual growth dimorphism (at the age of 180 d post-hatching, females were 34.4% heavier and 9.63% longer than males).

The sex-ratio in the global dataset was strongly skewed towards males, with a percentage of females less than half the percentage of males (32.5% versus 67.5%, respectively). This is consistent with the general observation that cultured sea bass, because of the hatchery environment (especially temperature), show an unbalanced sex-ratio in favour of males (Saillant et al. 2003; Piferrer et al. 2005), different if compared to wild-born sea bass, where in younger fish the sex-ratio seems to be balanced (Vandeputte et al. 2012).
The percentage of females was slightly higher in Atlantic/West-Med populations compared to Eastern Mediterranean populations, suggesting a possible different tendency in sex-ratio related to the origin of the individuals under aquaculture conditions. This is consistent with the study by Guinand et al. (2017), where the ♀WEM×♂Atlantic and ♀WEM×♂WEM crosses showed a higher mean proportion of females compared to the ♀WEM×♂NEM and ♀WEM×♂SEM crosses.
A limitation in our study is represented by the fact that the between-populations variation of sex-ratio could be confounded by non-additive genetic effects, as NAT, NEM and SEM broodstock were used only as sires. As previously reported by Guinand et al. (2017), sex-ratio can show non-additive components of genetic variance, and we have no possibility to disentangle additive and non-additive genetic effects in our case.
Finally, a better understanding of the genetic architecture of sex tendency in sea bass could have applications in aquaculture production. European sea bass is one of the most important marine species widely cultured in the Mediterranean areas and represents 49% of the marine Mediterranean aquaculture production (FEAP Annual report 2016). The strong bias towards males under aquaculture condition has been recognised by farmers as a problem for different reasons (lower growth rates of males compared to females, reduced flesh quality and general decrease of the commercial values of the product; Felip et al. 2006). Uncovering the population-specific sex determination system may help to produce stocks with higher proportions of females, through selective breeding and genomic selection. Moreover, the choice of broodstock coming from a specific origin could be interesting to start new breeding programs, due to the between-population differences in sex-ratio we found.

Data archiving

The dataset underlying our findings is available in the institutional public data repository (SEANOE: http://www.seanoe.org/), http://doi.org/10.17882/55576

References

Bahri-Sfar L, Lemaire C, Ben Hassine OK, Bonhomme F (2000). Fragmentation of sea bass populations in the western and eastern Mediterranean as revealed by microsatellite polymorphism. *Proc R Soc Lond B Bio* **267**: 929-935.

Bateman AW, Anholt BR (2017). Maintenance of polygenic sex determination in a fluctuating environment: an individual-based model. *Journal of Evolutionary Biology* **30**: 915-925.

Bull JJ (1983). Evolution of sex determining mechanisms. Benjamin/Cummings, Menlo Park, CA.

Bulmer MG, Bull JJ (1982). Models of polygenic sex determination and sex ratio control. *Evolution* **36**: 13-26.

Conover DO, Heins SW (1987). Adaptive variation in environmental and genetic sex determination in a fish. *Nature* **326**: 496-498.

Dempster ER, Lerner IM (1950). Heritability of threshold characters. *Genetics* **35**: 212-236.

Doan QK, Vandeputte M, Chatain B, Haffray P, Vergnet A, Breuil G et al. (2017a). Genetic variation of resistance to viral nervous necrosis and genetic correlations with production traits in wild populations of the European sea bass (*Dicentrarchus labrax*). *Aquaculture* **478**: 1-8.

Doan QK (2017b). Genetic and genomic variation of resistance to viral nervous necrosis in wild populations of European seabass (*Dicentrachus labrax*). PhD thesis, Université de Montpellier.

Duranton M, Allal F, Fraïsse C, Bierne N, Bonhomme F, Gagnaire P-A (2018). The origin and remolding of genomic islands of differentiation in the European sea bass. *Nat Commun* **9**: 2518.

Falconer DS, Mackay TFC (1996). Introduction to quantitative genetics. Longman, Harlow, England.

Fauvel C, Suquet M, Dreanno C, Menu B (1998). Cryopreservation of sea bass (*Dicentrarchus labrax*) spermatozoa in experimental and production simulating conditions. *Aquat Living Resour* **11**: 387-394.

Felip A, Zanuy S, Carrillo N (2006). Comparative analysis of growth performance and sperm motility between precocious and non-precocious males in the European sea bass (*Dicentrarchus labrax*, L.). *Aquaculture* **256**: 570-578.

Groeneveld E, Kovac M, Mielenz N (2010). VCE User's Guide and Reference Manual Version 6.0. Available at: https://vce.tzv.fal.de/

Guinand B, Vandeputte M, Dupont-Nivet M, Vergnet A, Haffray P, Chavanne H et al. (2017). Metapopulation patterns of additive and nonadditive genetic variance in the sea bass (*Dicentrarchus labrax*). *Ecol Evol* **7**: 2777-2790.

Habier D, Fernando RL, Kizilkaya K, Garrick DJ (2011). Extension of the bayesian alphabet for genomic selection. BMC Bioinformatics **12**: 186.

Heule C, Salzburger W, Böhne A (2014). Genetics of sexual development: an evolutionary playground for fish. *Genetics* **196**: 579-591.

Lemaire C, Versini JJ, Bonhomme F (2005). Maintenance of genetic differentiation across a transition zone in the sea: discordance between nuclear and cytoplasmic markers. *J Evolution Biol* **18**: 70-80.

Liew WC, Bartfai R, Lim Z, Sreenivasan R, Siegfried KR, Orieux N (2012). Polygenic sex determination system in zebrafish. *PLoS One* **7**: e34397.

Lourenco DAL, Tsuruta S, Fragomeni BO, Masuda Y, Aguilar I, Legarra A, Bertrand JK et al. (2015). Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. *J Anim Sci* **93**: 2653-2662.

Lynch M, Walsh B (1998). Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA.

Mank JE, Promislow DE, Avise JC (2006). Evolution of alternative sex-determining mechanisms in teleost fishes. *Biol J Linn Soc* **87**: 83-93.

Menu B, Peruzzi S, Vergnet A, Vidal MO, Chatain B (2005). A shortcut method for sexing juvenile European sea bass, *Dicentrarchus labrax* L. *Aquac Res* **36**: 41-44.

Misztal I, Tsuruta S, Lourenco D, Aguilar I, Legarra A, Vitezica Z (2015). Manual for BLUPF90 family of programs. University of Georgia, Athens, USA.

Moore EC, Roberts RB (2013). Polygenic sex determination. *Curr Biol* **23**: R510-R512.

Naciri M, Lemaire C, Borsa P, Bonhomme F (1999). Genetic study of the Atlantic/Mediterranean transition in sea bass (*Dicentrarchus labrax*). *J Hered* **90**: 591-596.

Palaiokostas C, Bekaert M, Taggart JB, Gharbi K, McAndrew BJ, Chatain B et al. (2015). A new SNP-based vision of the genetics of sex determination in European sea bass (*Dicentrachus labrax*). *Genet Sel Evol* **47**: 68.

Piferrer F, Blazquez M, Navarro L, Gonzalez A (2005). Genetic, endocrine, and environmental components of sex determination and differentiation in the European sea bass (*Dicentrarchus labrax* L.). *Gen Comp Endocrinol* **142**: 102-110.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D et al. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**: 559-575.

Quéré N, Desmarais E, Tsigenopoulos CS, Belkhir K, Bonhomme F, Guinand B (2012). Gene flow at major transitional areas in sea bass (*Dicentrarchus labrax*) and the possible emergence of a hybrid swarm. *Ecol Evol* **2**: 3061-3078.

Quignard JP (1978). La Méditerranée: creuset ichtyologique. *Boll Zool* **45**: 23-26.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/.

Saillant, E, Fostier A, Menu B, Haffray P, Chatain B (2001). Sexual growth dimorphism in sea bass *Dicentrarchus labrax*. *Aquaculture* **202**: 371-387.

Saillant E, Fostier A, Haffray P, Menu B, Laureau S, Thimonier J et al. (2003). Effects of rearing density, size grading and parental factors on sex ratios of the sea bass (*Dicentrarchus labrax* L.) in intensive aquaculture. *Aquaculture* **221**: 183-206.

Sansone G, Fabbrocini A, Ieropoli S, Langellotti AL, Occidente M, Matassino D (2002). Effects of extender composition, cooling rate, and freezing on the motility of sea bass (*Dicentrarchus labrax* L.) spermatozoa after thawing. *Cryobiology* **44**: 229-239.

Tine M, Kuhl H, Gagnaire P-A, Louro B, Desmarais E, Martins RST et al. (2014). The European sea bass genome and its variation provide insight into adaptation to euryhalinity and marine speciation. *Nat Commun* **5**: 5770.

Van Dooren TJM, Leimar O (2003). The evolution of environmental and genetic sex determination in fluctuating environments. *Evolution* **57**: 2667-2677.

VanRaden PM (2008) Efficient methods to compute genomic predictions. *J Dairy Sci* **91**: 4414-4423.

Vandeputte M, Mauger S, Dupont-Nivet M (2006). An evaluation of allowing for mismatches as a way to manage genotyping errors in parentage assignment by exclusion. *Mol Ecol Notes* **6**: 265-267.

Vandeputte M, Dupont-Nivet M, Chavanne H, Chatain B (2007). A polygenic hypothesis for sex determination in the European sea bass *Dicentrarchus labrax*. *Genetics* **176**: 1049-1057.

Vandeputte M, Quillet E, Chatain B (2012). Are sex ratios in wild European sea bass (*Dicentrarchus labrax*) populations biased? *Aquat Living Resour* **25**: 77-81.

Vandeputte M, Garouste R, Dupont-Nivet M, Haffray P, Vergnet A, Chavanne H et al. (2014). Multi-site evaluation of the rearing performances of 5 wild populations of European sea bass (*Dicentrarchus labrax*). *Aquaculture* **424-425**: 239-248.

Vandeputte M, Piferrer F (2018). Genetic and environmental components of sex determination in the European sea bass (*Dicentrarchus labrax*). In: Wang HP, Piferrer F, Chen SL (eds) *Sex Control in Aquaculture*. Wiley-Blackwell, ISBN: 978-1-119-12726-0 (in print).

Wang H, Misztal I, Aguilar I, Legarra A, Muir W (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet Res* **94**: 73-83.

Warnes GR, Bolker B, Lumley T, Johnson RC (2015). *gmodels*: various R programming tools for model fitting. Available at: https://CRAN.R-project.org/package=gmodels.

Wilson CA, High SK, McCluskey BM, Amores A, Yan Y, Titus TA et al. (2014). Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. *Genetics* **198**: 1291-1308.

Zhang X, Lourenco D, Aguilar I, Legarra A, Misztal I (2016). Weighting strategies for single-step genomic BLUP: an iterative approach for accurate calculation of GEBV and GWAS. *Front Genet* **7**: 151.

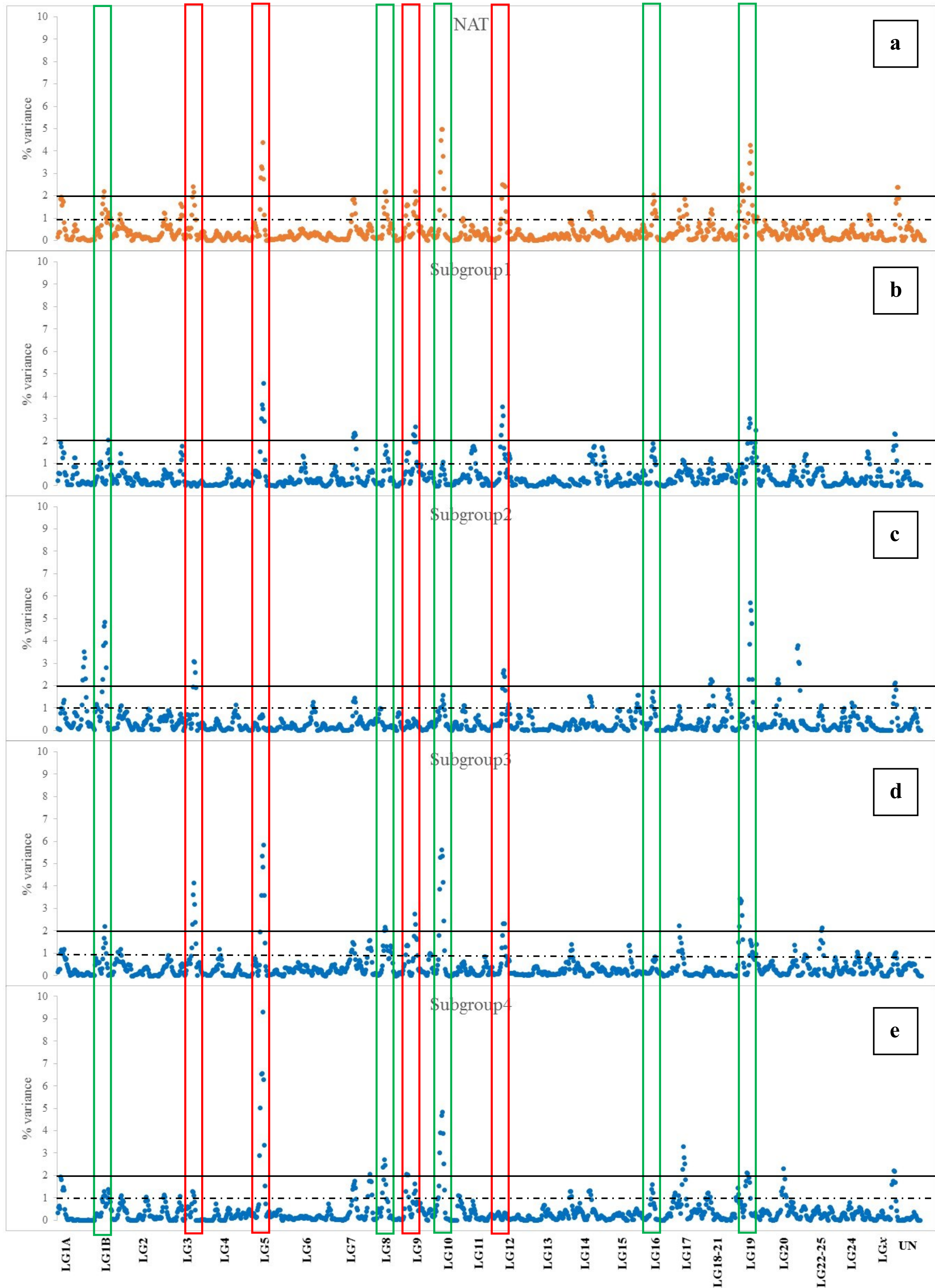
Zhang Z, Liu J, Ding X, Bijma P, de Koning D, Zhang Q (2010). Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. *PLoS ONE* **5**: e12648.

Supplementary material: cross-validation of QTLs

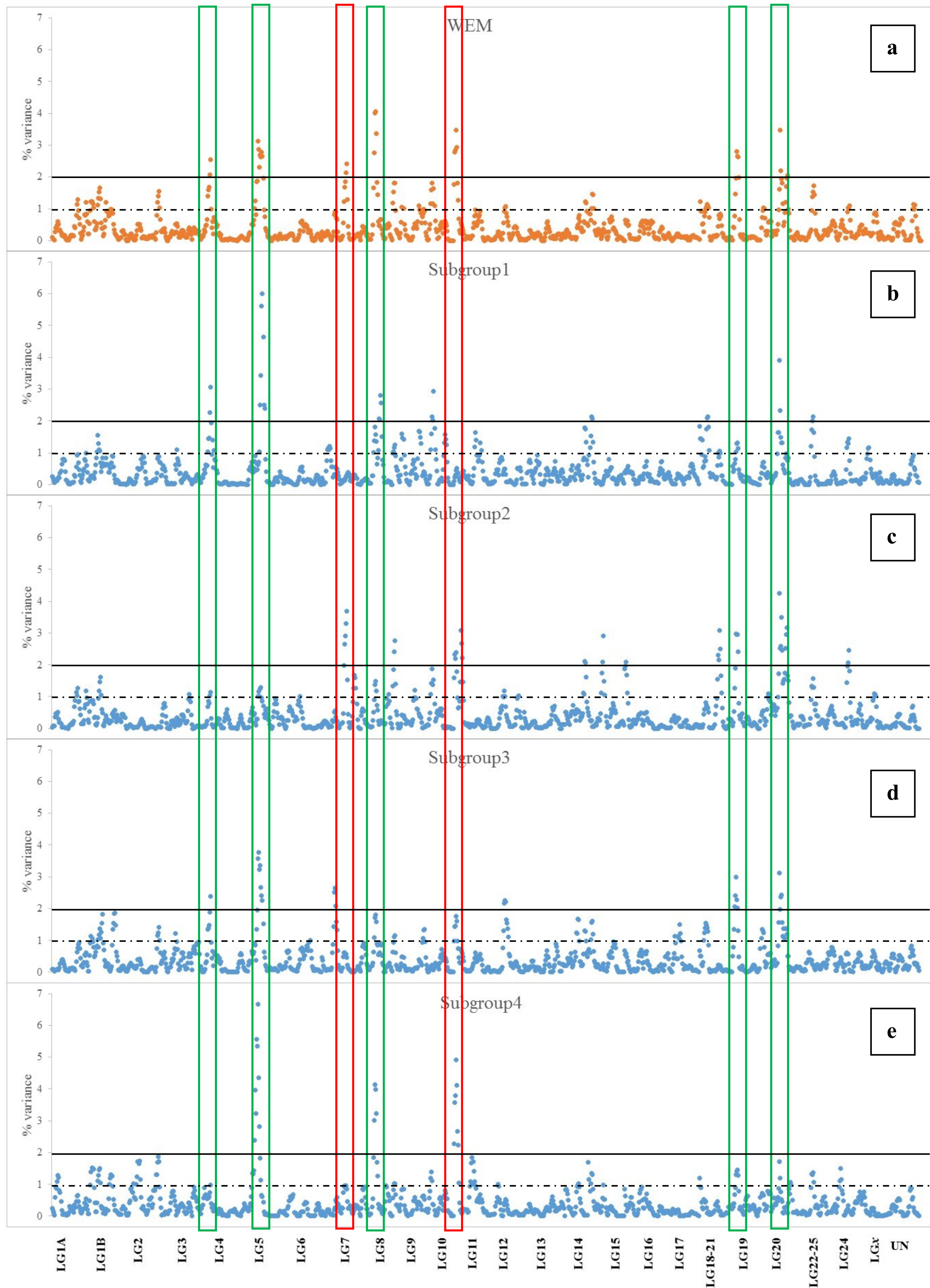
Method: in order to assess the stability of the QTLs examined, for each paternal origin, we generated four subgroups of offspring by removing for each 25% of the fish along the second axis of the PCA presented in Figure 2b. Fish were then removed along an axis representing variation between dams, but on average all shared the same proportion of the different sires. For each of the subsets, which represented 75% of the offspring of a given paternal origin, we performed the wGWAS through the same iterative process and the same model described in the material and methods.

Results: in the following figures (supplementary figures 1 to 4), we compare the QTLs identified on the whole sample of the given sire origin (> 2% variance explained) to the peaks observed in the subgroups. QTLs present in the whole sample of the given sire origin (>2% variance explained) and present in all subgroups tested within this population (>1% variance explained) are highlighted in green. In red QTLs absent in at least one of the subgroups.

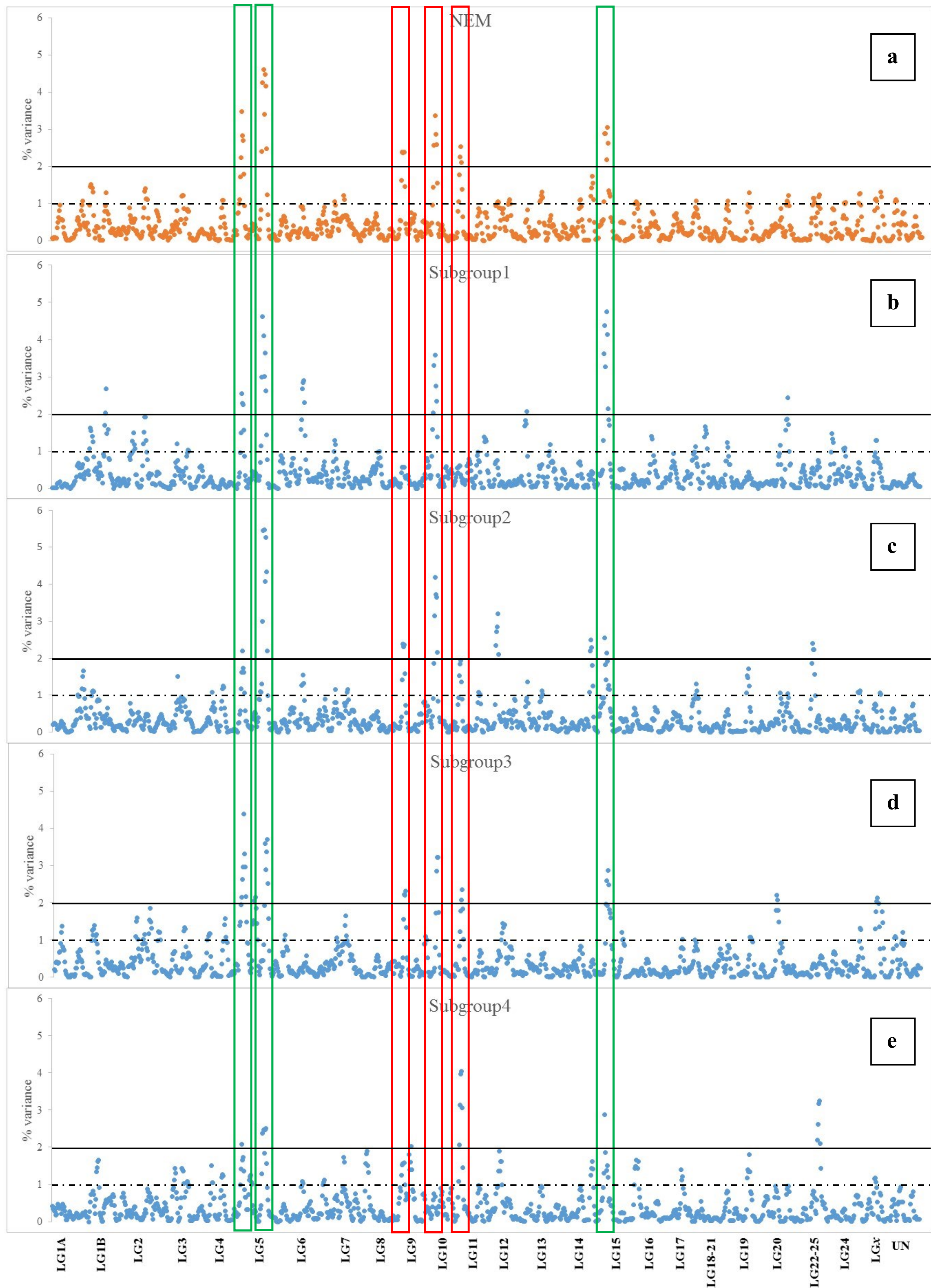
Discussion: in this cross validation, the subgroups were generated to maximize the genetic divergence (mainly due to the variation between dams) between them. Most of the QTLs found from the offspring of a given sire origin were also observed in all subgroups, suggesting they were linked to the sire origin studied. However, several of them were absent in some of the subgroups and some additional QTLs (> 2% variance explained) were detected. This may indicate sire origin by dam interaction, i.e. dominance variation.



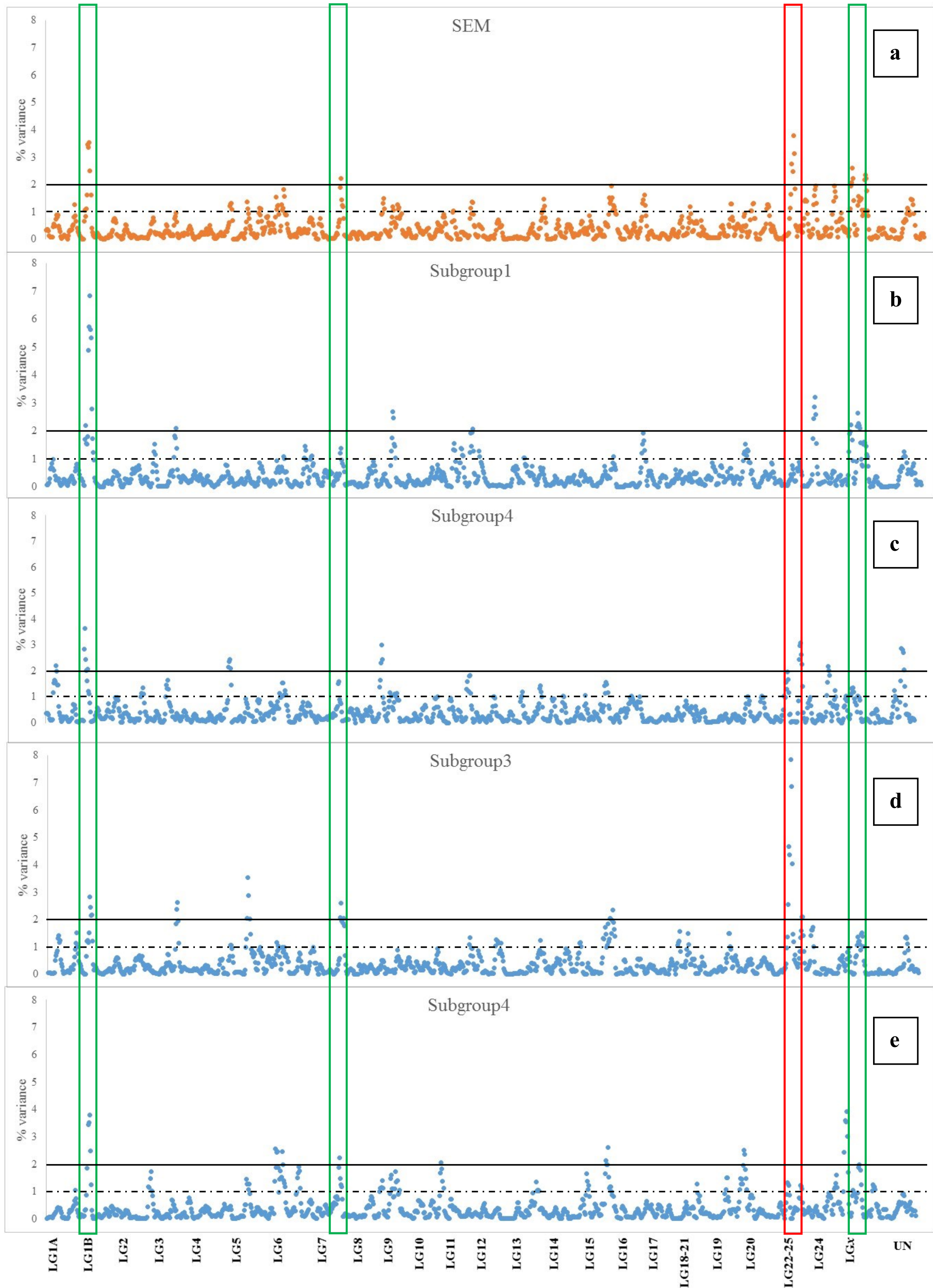
Supplementary Figure 1. Manhattan plots showing the percentage of variance explained for sex in North Atlantic (NAT) progenies. The results were obtained with a weighted GWAS in a sliding window of 5 adjacent SNPs in BLUPf90 (Misztal et al. 2015). a. all NAT sires offspring; b. subgroup 1; c. subgroup 2; d. subgroup 3; e. subgroup 4



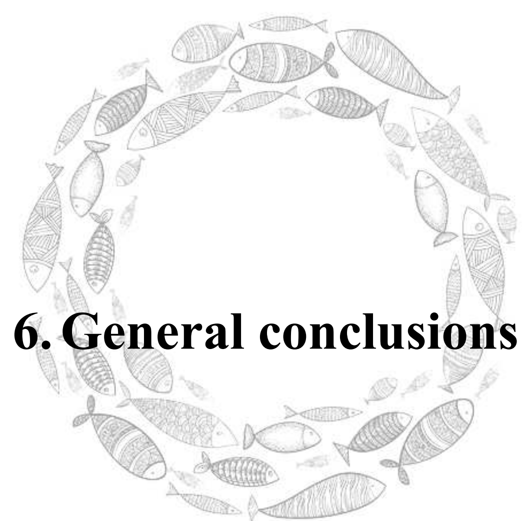
Supplementary Figure 2. Manhattan plots showing the percentage of variance explained for sex in Western Mediterranean (WEM) progenies. The results were obtained with a weighted GWAS in a sliding window of 5 adjacent SNPs in BLUPf90 (Misztal et al. 2015). a. all WEM sires offspring; b. subgroup 1; c. subgroup 2; d. subgroup 3; e. subgroup 4



Supplementary Figure 3. Manhattan plots showing the percentage of variance explained for sex in Northern-Eastern Mediterranean (NEM) progenies. The results were obtained with a weighted GWAS in a sliding window of 5 adjacent SNPs in BLUPf90 (Misztal et al. 2015). a. all NEM sires offspring; b. subgroup 1; c. subgroup 2; d. subgroup 3; e. subgroup 4



Supplementary Figure 4. Manhattan plots showing the percentage of variance explained for sex in Southern-Eastern Mediterranean (SEM) progenies. The results were obtained with a weighted GWAS in a sliding window of 5 adjacent SNPs in BLUPf90 (Misztal et al. 2015). a. all SEM sires offspring; b. subgroup 1; c. subgroup 2; d. subgroup 3; e. subgroup 4



6. General conclusions

This project addressed two topical issues in the European sea bass, the first concerning viral nervous necrosis resistance and the second related to the unusual sex determination mechanism of this commercially important species.

Due to the increasing interest in selective breeding for disease resistance, we aimed firstly at the estimation of the genetic and genomic parameters for VNN resistance in sea bass, with satisfactory results, that reinforce other outcomes in the same species and which outline the potential of selective breeding strategies as VNN prevention tools.
The possible exploitation of other indicator traits as indirect selective parameters to improve VNN resistance drove our attention to cortisol levels (the major physiological indicator of stress) and antibody titer to nervous necrosis virus, providing very interesting and new results in terms of genetic and genomic parameters and elucidating the genetic relationships between these traits and VNN resistance.
Antibody titer against VNN represents the novelty in our study and it is a not commonly measurable trait. Heritability estimates are surprisingly high and the significant negative genetic correlation with the mortality trait offers the opportunity of effective indirect selection strategies to improve VNN resistance in sea bass.
In contrast, based on our results, post-stress cortisol levels cannot be taken into account as an indirect parameter to improve VNN resistance in sea bass, due to the trivial genetic correlation with mortality, but the significant heritability we found suggests that the improvement of sea bass stress resistance can be possible. This could be interesting in a context of commercial production, due to the better efficiency in aquaculture environment that characterized low stress-sensitive fish, with higher resilience and productivity in face of the environmental challenges typical of intensive systems.
Body weight at a fixed age was included in the list of analysed traits, because of the importance of growth traits in the context of production. As expected, heritability estimates are high and the genetic correlation between body weight and VNN mortality is significantly negative, implying that an increment in body weight should improve the resistance to VNN.
The clear issues related to the routinely recording of the individual phenotypes in case of disease resistance, post-stress cortisol levels and antibody titer, motivated us to investigate the effectiveness of genomic tools in predicting the genetic merit of these traits.
The comparison between commonly used Bayesian models provide significant information:

i) in our case, none of the model outperform the others;
ii) the prediction accuracy of the phenotype or the EBV (correlation between observed value and prediction) for VNN mortality, post-stress cortisol concentrations, antibody titer and body weight is highly satisfactory, and drive the practical implementation of genomic selection in sea bass aquaculture for these traits;
iii) using the prediction of the EBV for VNN mortality to classify the phenotypes of the same trait resulted in better classification performance in comparison with the classification based on genomic predictions of the phenotype;
iv) the utilization of a lower number of markers seems not compromise the prediction accuracy.

Overall, these results indicate that genomic prediction could have a great potential for traits like VNN mortality, body weight, cortisol concentration and antibody titer in European sea bass, although further investigations on a larger experimental population are needed. The comparison of different

SNP marker densities could be interesting, as well, with the aim to develop cost-effective genotyping for a future practical application in a context of commercial aquaculture.
Moreover, during the experimental challenge test, questions concerning the modalities to infect the animals were raised, which could be addressed in future studies. More in detail, our hypothesis is that the use of different methods (immersion in water containing the pathogen rather than direct injection of viral suspension) might lead to different results when the focus is the investigation of the genetic variation in resistance, mainly because in the case of immersion the first protective systems of the fish may prevent the penetration of the virus, contributing to additional variation.

The unusual sex determination mechanism of European sea bass, consistent with a polygenic model affected by environmental effects, is another interesting topic, which has been investigated in terms of different genetic architecture exhibited by populations belonging to distinct geographical areas and therefore exposed to variable environmental/thermal conditions, such as the Atlantic Ocean, the Western and the Eastern Mediterranean.
Significant heritability estimates indicate the presence of a genetic basis for sex, that is however roughly equivalent to the environmental component. Moreover, sex tendency resulted genetically correlated with growth-related traits (body weight and length), confirming the hypothesis of a strong link between the genes affecting these traits. Parental effects on phenotypic sex were significant, as well, with sires and dams having a clear effect on the sex-ratio of the offspring.
The wGWAS revealed different QTLs underlying sex determination between Atlantic and Mediterranean lineages, with analogies between individuals belonging to adjacent origins that might reflect the introgression of Atlantic genes within the Mediterranean genetic background, and an interesting gradient of similarities from Western to Eastern Mediterranean populations. These findings are therefore consistent with the hypothesis of a population-specific evolution of the sex determination system in different environments. In addition, a different tendency in sex-ratio according to the origin of the individuals was suggested by the significant differences in terms of proportion of females between Atlantic/Western Mediterranean and Eastern Mediterranean populations. This study was another step in the comprehension of the genetic architecture of sex determination in sea bass, with a clear evolutionary interest, but even a practical interest, due to the strong bias towards males under aquaculture conditions and the concern directed in finding a strategy to control sex-ratio and to produce a higher number of females.

Appendix

The following article is the one published in the journal Heredity (Springer Nature, https://www.nature.com/hdy/).

Heredity
https://doi.org/10.1038/s41437-018-0157-z

the geneticssociety

ARTICLE

Population-specific variations of the genetic architecture of sex determination in wild European sea bass *Dicentrarchus labrax* L.

Sara Faggion[1,2] · Marc Vandeputte[2,3] · Béatrice Chatain[2] · Pierre-Alexandre Gagnaire[4] · François Allal[2]

Received: 25 May 2018 / Revised: 10 September 2018 / Accepted: 28 September 2018
© The Genetics Society 2018

Abstract

Polygenic sex determination (PSD) may show variations in terms of genetic and environmental components between populations of fish species exposed/adapted to different environments. The European sea bass (*Dicentrarchus labrax*) is an interesting model, combining both a PSD system and a genetic subdivision into an Atlantic and a Mediterranean lineage, with genetic substructures within the Mediterranean Sea. Here, we produced experimental progeny crosses ($N = 927$) from broodstock sampled in four wild populations (North Atlantic, NAT; Western Mediterranean, WEM; North-Eastern Mediterranean, NEM; South-Eastern Mediterranean, SEM). We found less females than males in the progeny, both in the global dataset (32.5%) and within each paternal group (from 25.1% for NEM to 39.0% for WEM), with significant variation among populations, dams, and sires. Sex, body weight (BW), and body length (BL) showed moderate heritability (0.52 ± 0.17, 0.46 ± 0.17, 0.34 ± 0.15, respectively) and sex was genetically correlated with BW and BL, with $r_{\mathrm{Asex/BW}} = 0.69 \pm 0.12$ and $r_{\mathrm{A\ sex/BL}} = 0.66 \pm 0.13$. A weighted GWAS performed both on the global dataset and within each paternal group revealed a different genetic architecture of sex determination between Atlantic and Mediterranean populations (9 QTLs found in NAT, 7 in WEM, 5 in NEM, and 4 in SEM, with a cumulated variance explained of 27.04%, 21.87%, 15.89%, and 12.10%, respectively) and a more similar genetic architecture among geographically close populations compared to geographically distant populations, consistent with the hypothesis of a population-specific evolution of polygenic sex determination systems in different environments.

Introduction

Sex determination is a mechanism of major evolutionary importance that exhibits a high variety of modalities, and this is especially true in fish (Mank et al. 2006; Heule et al. 2014). These modalities are generally classified as genotypic sex determination systems (or GSD, which includes male/female heterogamety and polygenic sex determination, or PSD) and environmental sex determination systems (ESD; Bull 1983).

PSD has been initially formalized by Bulmer and Bull (1982): they proposed an underlying "sex tendency" phenotype, with a polygenic determinism influencing the observed phenotype (male or female) depending on whether it lies below or beyond a fixed threshold. Under this model, any environmental or genotypic effect can equally affect the phenotype (here sex tendency) and bring its value below or beyond the threshold, therefore determining sex.

In the classical view, PSDis thought to be unstable, and should evolve either towards GSD in a fluctuating environment generating biased sex-ratios, or towards ESD if some environments increase the fitness of a specific sex (Bulmer and Bull 1982). Modeling approaches have shown that the orientation of PSD in one or the other direction depends on complex combinations of environmental

Electronic supplementary material The online version of this article (https://doi.org/10.1038/s41437-018-0157-z) contains supplementary material, which is available to authorized users.

✉ François Allal
francois.allal@ifremer.fr

1 Department of Comparative Biomedicine and Food Science, University of Padua, Agripolis, Viale dell'Università, 16, 35020 Legnaro, PD, Italy

2 MARBEC, Univ. Montpellier, Ifremer, CNRS, IRD, Palavas-les-Flots, France

3 GABI, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, 78350 Paris, France

4 ISEM, Univ. Montpellier, CNRS, IRD, EPHE, Montpellier, France

Published online: 24 October 2018

variation between and within environmental patches, and on migration rates between patches (Van Dooren and Leimar 2003; Bateman and Anholt 2017). Therefore, starting with the same ancestral PSD system, we may hypothesize that there could be a population-specific evolution of the sex determination mechanism, in other words the components of the model may balance differently between sub-populations of the same species that are exposed to different environmental conditions (Vandeputte et al. 2012; Guinand et al. 2017).

In teleost fish, PSD has been well-documented in the Atlantic silverside, *Menidia menidia* (Conover and Heins 1987), some populations of the zebrafish, *Danio rerio* (Liew et al. 2012; Wilson et al. 2014), and in the European sea bass, *Dicentrarchus labrax* (Vandeputte et al. 2007; Palaiokostas et al. 2015). However, some authors consider that it may be more frequent than classically thought (Moore and Roberts 2013) since it is difficult to characterize.

The European sea bass (*D. labrax* L.) offers an interesting model to investigate the evolution of PSD. This species combines both a PSD system (Vandeputte et al. 2007; Palaiokostas et al. 2015) and a clear genetic subdivision into an Atlantic and a Mediterranean lineage, in addition to population genetic structure within the Mediterranean Sea (Naciri et al. 1999; Bahri-Sfar et al. 2000; Lemaire et al. 2005; Quéré et al. 2012; Tine et al. 2014).

Experimental evidence shows that early life rearing temperature has a strong influence on the sex-ratio of sea bass, with low rearing temperature (between 13 and 17 °C) during the first 60 days of life, favouring the production of female offspring (reviewed by Vandeputte and Piferrer 2018), which are preferred in aquaculture due to their higher growth rate (Saillant et al. 2001; Felip et al. 2006). In the natural environment, observed variation in yearly cohort sex-ratio may be indicative of natural variations in temperature influencing sex determination also in the wild populations of sea bass (Vandeputte et al. 2012).

Variation of sex-ratios between families in sea bass is consistent with the hypothesis that phenotypic sex is determined by an underlying continuous sex tendency combining the effects of polygenes and temperature (Vandeputte et al. 2007). However, at least three genome-wide significant sex determining quantitative trait loci (QTL) were recently found by Palaiokostas et al. (2015), showing that some specific loci could have a stronger effect on sex tendency, at least in some populations (in this case the western Mediterranean).

Here, we hypothesize that the sex determination system of European sea bass could have a different genetic architecture in different populations exposed to different environmental conditions (especially temperature) in the wild. Experimental progeny crosses were produced from broodstock sampled in four wild populations (corresponding to the whole distribution of natural populations, from Northern Atlantic to Eastern Mediterranean Sea). We estimated the additive genetic variation for sex tendency and the correlation between sex tendency and growth-related traits (weight and length) and performed a weighted genome-wide association study (wGWAS), both on the global dataset and within each group of paternal origin, to assess possible population-specific variations in the architecture of sex determination in sea bass.

Materials and methods

Broodstock origin, production, and rearing of experimental fish

The male broodstock used in this study belonged to four different origins, matching with most of the natural range of the species: North Atlantic (NAT), Western Mediterranean (WEM), North-Eastern Mediterranean (NEM) and South-Eastern Mediterranean (SEM). The female broodstock belonged to the WEM population. The origin and collection of the broodstock has been detailed by Vandeputte et al. (2014).

The NAT and WEM sires were reared at the IFREMER facilities of Palavas-les-Flots (France) and the sperm stripped and cryopreserved following Fauvel et al. (1998), while the sperm of SEM and NEM sires were both cryopreserved in 2005 at IOLR, Eilat, Israel (SEM sperm) and at Beymelek Lagoon, Turkey (NEM sperm) following the protocol of Sansone et al. (2002). The dams were reared at the IFREMER station of Palavas-les-Flots under natural photoperiod which was 11L:13D at the time of spawning (3 March 2014) and natural temperature, which decreased from 23.5 °C in August 2013 to 13.5 °C at the time of spawning; dams with a suitable stage of development of eggs (determined after ovarian biopsy) were hormonally injected and stripped 72 h after the injection. Artificial fertilization was performed at the IFREMER station, using a full factorial mating scheme: 15 sires per origin (60 sires in total) were crossed with 9 WEM dams. The fertilization protocol and the rearing of experimental fish were described previously by Doan et al. (2017a). Briefly, after hatching, larvae were reared in a common garden at a temperature of 16.5 °C and 25‰ salinity until 58 dph (days post-hatching); the following 7 days the temperature was gradually increased to 20 °C. Fish were then reared at a mean temperature of 21.5 °C (18.1–22.4 °C) and 30‰ salinity until 102 dph. Afterwards, fish were divided into five juvenile tanks (A–E), with a mean temperature of 22.1 °C (15.5–27.9 °C). Fish were fed using a classical hatchery feeding sequence (*Artemia* nauplii, Le Gouessant Marine

Start and Neo Start pellets). At 180 dph, fish were individually tagged and measured for body weight (BW) and fork length (BL). At the same time, fin samples were collected for genomic DNA extraction. At 226 dph, 927 randomly chosen experimental fish were euthanized with an overdose of benzocaine, dissected, and the sex was recorded by visual observation of the gonads or using the squash technique (Menu et al. 2005) when macroscopic observation was ambiguous. The reliable identification of phenotypic sex was possible for all 927 fish. Sex was coded as a binary trait, 1 for males and 2 for females.

Genotyping, parentage assignment, and descriptive statistics

Fin clips from the 927 experimental fish, from the 60 sires and from the 9 dams were sent to LABOGENA (Jouy-en-Josas, France) for genomic DNA extraction and genotyping. Genotyping was performed with an iSelect Custom Infinium Illumina® European sea bass 3K SNP array. The design of this SNP array was done by selecting 2722 SNPs from an initial genome-wide variation map containing 2,628,725 SNPs phased into chromosome-wide haplotypes. These SNPs were discovered from 14 wild individuals from both the Atlantic and Mediterranean areas, using whole-genome sequencing as described by Duranton et al. (2018). A first filtering of the SNPs was made to remove variants closer than 80 bp from another known variant. Then, as recommended by Illumina®, we filtered the remaining SNPs to avoid A/T and C/G variants, that need a particular design (Infinium I Probe Design) involving setting up two probes instead of one. Among the remaining candidates, SNPs were chosen to cover all the chromosomes with a variable SNP density depending on the local nucleotide diversity (π), as reported by Tine et al. (2014). To do so, five π-classes were defined depending on the π estimated in non-overlapping 50 kb windows: class 1 for $\pi < 10^{-3}$; class 2 for $10^{-3} < \pi < 2\times10^{-3}$; class 3 for $2\times10^{-3} < \pi < 3\times10^{-3}$; class 4 for $3\times10^{-3} < \pi < 4\times10^{-3}$; class 5 for $\pi > 4\times10^{-3}$. Based on "best quality criterion", one SNP was selected in class 1 windows (in interval 22.5–27.5 kb of the window), 2 SNPs in class 2 windows (in intervals 12.5–17.5 and 35-40 kb), three SNPs in class 3 windows (in intervals 5.8–10.8, 22.5–27.5, and 39–44 kb), four SNPs in class 4 windows (in intervals 3.7–8.7, 16.2–21.2, 28.7–33.7, and 41.2–46.2 kb), and five SNPs in class 5 windows (in intervals 2.5–7.5, 12.5–17.5, 22.5–27.5, 32.5–37.5, and 42.5-47.5 kb). Since nucleotide diversity is negatively correlated to the local recombination rate in sea bass (which was estimated by Tine et al. 2014), this local adjustment in the density of SNPs aimed at homogenizing the density of markers along the recombination map instead of the physical map.

Parentage assignment was performed with an exclusion-based software, VITASSIGN (Vandeputte et al. 2006), using 2722 markers and allowing 29 allelic mismatches to recover pedigree.

Proportion of individuals, males and females in the global dataset, per origin, per tank, per dam and per sire, pairwise comparisons and χ^2 tests were performed in R version 3.4.3 using the packages *stats* (R Core Team 2017) and *gmodels* (Warnes et al. 2015). All *P*-values were adjusted for multiple testing with the Bonferroni correction method.

Principal component analysis (PCA)

To describe the overall genetic structure among the 927 individuals genotyped on the basis of genome-wide SNP data, we performed a PCA using *–pca* function in PLINK (Purcell et al. 2007). A two-dimension scatter plot of individuals coordinates on the two first principal components was generated, indicating the percentages of variance explained.

Heritability, genetic, and phenotypic correlations

Heritability was estimated on the entire dataset through a linear mixed sire model using the software VCE 6.0 (Groeneveld et al. 2010). The model was the following:

$$y_{ijkl} = o_i + t_j + s_{k(i)} + d_l + e_{ijkl}$$

where y_{ijkl} is the phenotype for the studied trait (coded as a binary trait, 1 for male and 2 for female in the case of sex); o_i is the fixed effect of the population of origin of the sires i; t_j is the fixed effect of the rearing tank j; $s_{k(i)}$ is the random additive genetic effect of sire k within origin i; d_l is the random effect of dam l; e_{ijkl} is the random residual.

As explained by Falconer and Mackay (1996), the sire component accounts for 1/4 of the additive genetic variance; for this reason, the heritability was estimated as $h^2 = 4\sigma_s^2/\sigma_p^2$, with σ_s^2 being the sire component of variance and σ_p^2 the phenotypic variance. When the trait was sex, heritability on the observed (binary) scale was transformed to the value on the underlying liability scale (Dempster and Lerner 1950; Lynch and Walsh 1998) following the formula:

$$h_u^2 = h_o^2 p(1-p)/z^2$$

where h_u^2 is the heritability on the liability scale, h_o^2 is the heritability on the observed scale, p is the incidence (proportion of females) in the population and z is the value of the normal distribution density at the point where the cumulative distribution function of the normal distribution reaches incidence.

Genetic and phenotypic correlations between sex and growth-related traits (body weight and length at 180 dph)

were assessed using VCE 6.0 software (Groeneveld et al. 2010) applying a three traits sire model with sex, weight, and length as variables.

Genome-wide association study (GWAS)

GWAS was performed through the BLUPf90 family of programs for mixed-model computations (Misztal et al. 2015) in order to identify possible SNPs associated with phenotypic sex.

Owing to the genetic subdivision between Atlantic and Mediterranean sea bass lineages and the finer scale differentiation between Western and Eastern Mediterranean populations (Naciri et al. 1999; Bahri-Sfar et al. 2000; Lemaire et al. 2005; Quéré et al. 2012; Tine et al. 2014; Duranton et al. 2018), the dataset was split into four groups depending on paternal origin (NAT, WEM, NEM, and SEM) and the analyses were performed separately within each group. The origin of the sire was taken into account as a fixed effect when we performed GWAS on the global dataset.

The raw SNP dataset was quality-filtered before running the GWAS. We firstly removed eight individuals with a call rate lower than 0.9. We then applied variant filters to exclude SNPs showing either significant Mendelian distortions, a minor allele frequency (MAF) lower than 0.05, or a proportion of missing genotypes greater than 0.1. This resulted in a final dataset containing 1205 retained markers out of the total 2722 SNPs which were genotyped in 919 offsprings and 69 parents (988 individuals in total). The high number of discarded SNPs is mainly due to technical design problems, which lead to target the wrong (i.e. non-variable) base position for 50% of the markers. The mapping of the SNPs used for further analyses was reported by Doan (2017b).

For the weighted GWAS (wGWAS) the following model was applied:

$$y = \mathbf{X}b + \mathbf{W}u + e$$

where y is the vector of phenotypes, b vector of the fixed effects (intercept, tank and origin of the sires), $\mathbf{X}$ the incidence matrix relating phenotypes with the fixed effects, $\mathbf{W}$ the incidence matrix relating phenotypes with the random animal effects, u the vector of random animal effects ~$N(0, \mathbf{G}\sigma_g^2)$ with $\mathbf{G}$ being the genomic relationship matrix (VanRaden 2008), σ_g^2 the additive genetic variance, e the vector of residuals ~$N(0, I\ \sigma_e^2)$ and σ_e^2 the residual variance. The genomic relationship matrix $\boldsymbol{G}$ was established as follows:

$$\mathbf{G} = \mathbf{ZDZ}'/q$$

where $\mathbf{D}$ is a diagonal matrix with weights for SNP effects, $\mathbf{Z}$ is a matrix of gene content adjusted for allele frequencies and q is a weighting factor equal to $2\Sigma p_i\,(1-p_i)$, where p_i is the MAF of SNP i.

The wGWAS was implemented through an iterative process (Zhang et al. 2010, 2016; Wang et al. 2012) and using Gibbs sampling (THRGIBBS1F90) to estimate the genomic estimated breeding values (GEBVs), since this is specifically adapted to the analysis of binary traits. The following steps were performed:

(1) in the first iteration, $\mathbf{D}$ was first set equal to $\boldsymbol{I}$, the identity matrix (VanRaden 2008);

(2) $\mathbf{G}$ was calculated as $\mathbf{G} = \mathbf{ZDZ}'/q$;

(3) Direct genomic values (DGVs) were obtained from GEBVs as $\mathrm{DGV_i} = -(\Sigma_{j,\, j\neq i}\, g^{ij}\mathrm{GEBV}^j/g^{ii})$ with g^{ij} elements of the $\mathbf{G}^{-1}$ matrix (Lourenco et al. 2015), and converted to SNP effects as $a_\mathrm{i} = \mathbf{DZ}'\mathbf{G}^{-1}\mathrm{DGV}_i$;

(4) new SNP weights over a 5 SNPs window as $d_i = \Sigma_i a_i^2/5$ were calculated and normalized so that the total genetic variance remained constant;

(5) the process was then iteratively repeated from step 2 with the new $\mathbf{D}$ matrix.

After three rounds there was no further modification of the variance explained by the SNPs (i.e. the correlation coefficient R^2 between rounds 1 and 2 was equal to 0.93, 0.97 between rounds 2 and 3, and 0.99 between rounds 3 and 4).

We calculated the regional variance explained by summing neighboring SNP variance in overlapping windows of five adjacent SNPs, suggested by Habier et al. (2011) as the more appropriate method to infer the effect of QTLs. We considered as QTLs the genomic segments that explained a proportion of genetic variance higher than 2%.

Results

Parentage assignment and descriptive statistics

Assignment to a unique parental pair was achieved for all 927 offspring. The number of fish per paternal origin varied from a minimum of 150 (NAT) to a maximum of 328 individuals (SEM). The number of offspring per sire varied from 1 to 36 and the number of offspring per dam varied from 2 to 250.

The proportion of females in the global dataset was 32.5% and was variable among groups of paternal origin ($\chi^2 = 12.27$, df = 3, P-value = 7×10^{-3}), ranging from 25.1% for NEM to 39.0% for WEM (Table 1). Pairwise comparisons showed significant sex-ratio differences between NAT and NEM and between WEM and NEM (Table 1).

Table 1 Number of individuals (N), sex-ratio, mean body weight (g) and length (mm) at 180 dph (total and separately for males and females) with the coefficient of variation expressed in percentage (CV%); the results are showed for the global dataset and separately per group of paternal origin; different letters indicate a significant difference at $P \leq 0.05$

	Paternal origin				Global $N = 927$
	NAT $N = 150$	WEM $N = 182$	NEM $N = 267$	SEM $N = 328$	
Sex-ratio					
Proportion of males	0.620[a]	0.610[a]	0.749[a]	0.677[ab]	0.675
Proportion of females	0.380[a]	0.390[a]	0.251[b]	0.323[ab]	0.325
Mean $BW_{180\ dph}$ (CV%)	17.57 (36.4%)	14.70 (30.3%)	16.22 (34.5%)	16.09 (34.7%)	16.09 (34.7%)
Males	16.00 (38.4%)	13.02 (28.1%)	14.90 (31.5%)	14.19 (30.5%)	14.48 (32.7%)
Females	20.11 (29.9%)	17.32 (25.0%)	20.17 (30.8%)	20.08 (28.9%)	19.46 (29.4%)
Mean $BL_{180\ dph}$ (CV%)	113.01 (11.4%)	108.89 (9.4%)	111.48 (10.5%)	110.46 (10.7%)	110.86 (10.6%)
Males	109.46 (11.4%)	105.23 (8.9%)	108.90 (10.0%)	106.54 (9.4%)	107.50 (9.9%)
Females	118.79 (9.6%)	114.62 (7.8%)	119.18 (9.0%)	118.68 (9.4%)	117.85 (9.1%)

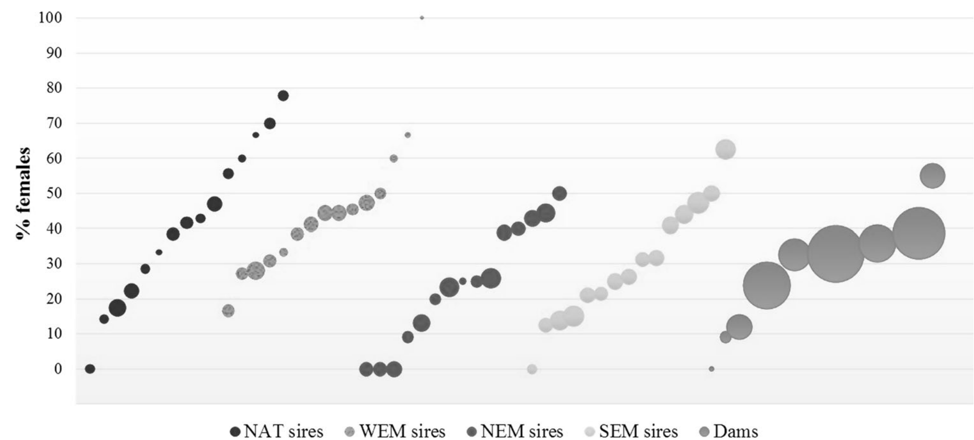


Fig. 1 Percentage of females per sire and per dam; blue, light blue, green and yellow bubbles identify the different origins of the sires, the pink ones identify the dams; the size of the bubble represents the total number of offspring per sire/dam

There were neither significant differences in the proportion of males and females between the five rearing tanks ($\chi^2 = 9.48$, df = 4, P-value = 5.02×10^{-2}) nor in the proportion of animals belonging to the four paternal origins between tanks ($\chi^2 = 16.37$, df = 12, P-value = 17.47×10^{-2}; Table S1). On the contrary, the proportion of females in the offspring strongly differed per sire and per dam ($\chi^2 = 124.56$, df = 59, P-value = 10^{-6} for sires, $\chi^2 = 34.72$, df = 8, P-value = 3×10^{-6} for dams; Fig. 1). The proportion of females ranged from 0% to 100% in paternal half-sib families and from 0% to 55% in maternal half-sib families.

Female offspring were, on average, heavier and longer than males at 180 dph; this was true both for the global dataset and within groups of paternal origin (Table 1).

Principal component analysis

The PCA performed on the global dataset revealed that genotypic variance was mainly explained by paternal origin (first principal component explaining 5.9% of the variance, Fig. 2a). Furthermore, variations among dams were also detected (second axis explaining 4.1% of the variance, Fig. 2b).

More precisely, the first PC axis distinguished three groups c orresponding to the population of origin of the sires; NAT, Western, and Eastern Mediterranean groups were clearly separated, while the difference between North and South Eastern Mediterranean group was more subtle. This stratification explained by the population of origin of the sires was properly taken into account as a fixed effect in the models used to estimate heritability and to perform the wGWAS on the global dataset.

Heritability, genetic, and phenotypic correlations

Heritability was moderately high for all the variables (sex, BW, and body length at 180 dph, Table 2). The genetic and phenotypic correlations between sex and growth-related traits were moderately high and the genetic correlations were higher compared to the phenotypic correlations (> 0.65 vs. 0.42). The genetic and phenotypic correlations between body weight and length were close to unity (Table 2).

Genome-wide association study (GWAS)

Results from the wGWAS performed on the global dataset identified one major group of SNPs on LG6 explaining up to 3.41% of the variance for sex. Other important groups of SNPs were detected on LG7, LG12, LG15, and LG*x*,

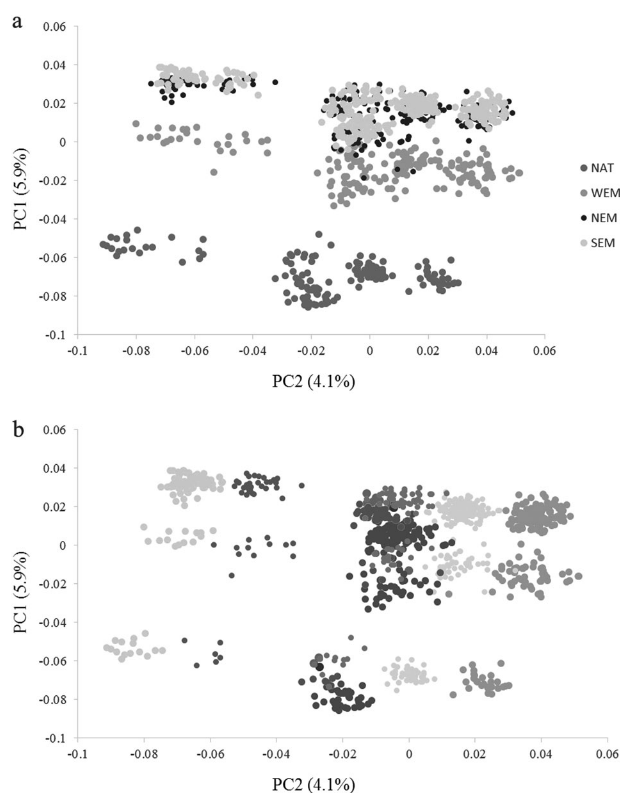


Fig. 2 Two-dimension scatterplots showing the population stratification in the global dataset ($N = 927$) by paternal origin (**a**) and by dams (**b**). The first principal component was plotted against the second; the percentages of variance explained by each axis is indicated

Table 2 Heritability (± s.e., in bold on the diagonal) for sex (on the liability scale) and growth-related traits (body weight, BW, and body length, BL, at 180 dph), genetic (± s.e.; below the diagonal), and phenotypic correlations (above the diagonal) among traits, estimated with VCE 6.0 (Groeneveld et al. 2010)

Trait	Sex	$BW_{180\ dph}$	$BL_{180\ dph}$
Sex	**0.52 ± 0.17**	0.42	0.42
$BW_{180\ dph}$	0.69 ± 0.12	**0.46 ± 0.17**	0.97
$BL_{180\ dph}$	0.66 ± 0.13	0.99 ± 0.005	**0.34 ± 0.15**

explaining up to 2.73% of variance, while a minor group explaining slightly more than 2% of variance was located on LG2 (Fig. 3a and Table 3).

The comparisons of the Manhattan plots of the wGWAS performed separately within each group of paternal origin (Fig. 3b–e and Table 3), revealed a clear pattern of similarity between samples belonging to adjacent paternal origins. Nevertheless, taking into account genomic regions explaining at least 2% of the variance showed a very variable architecture of sex determination, with some peaks shared among populations, while others were clearly population-specific (Table 4 and Fig. 4).

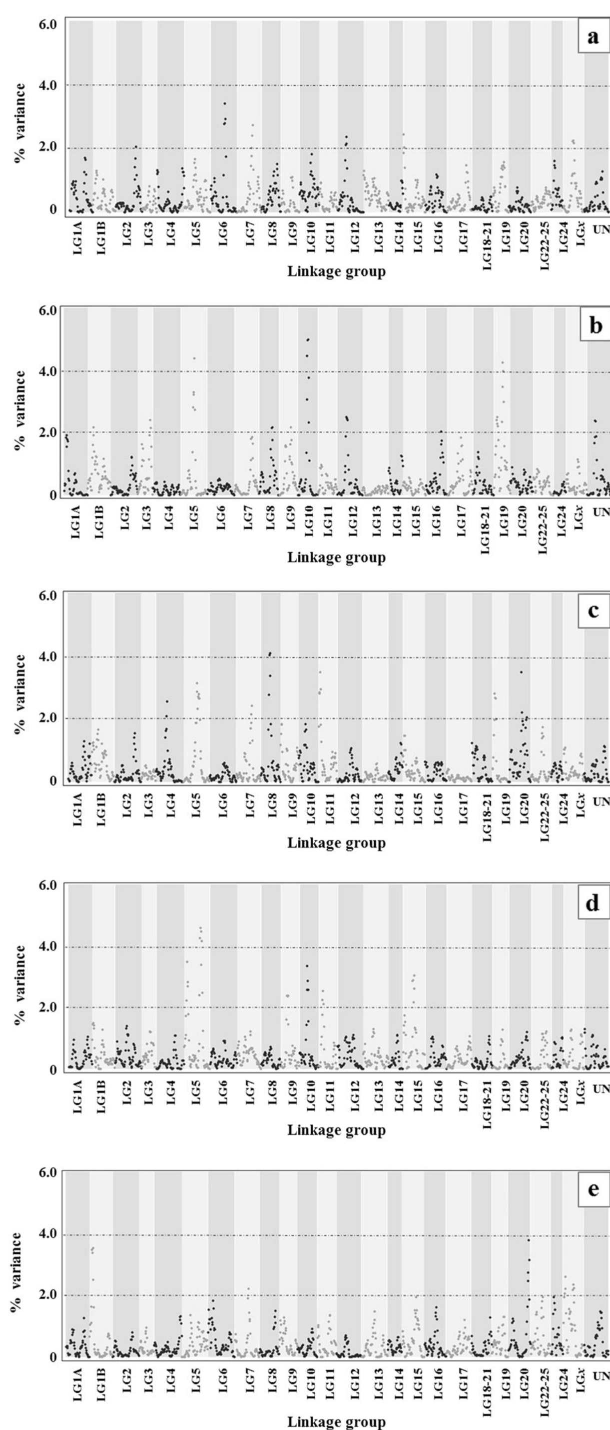


Fig. 3 Manhattan plots showing the percentage of variance explained for sex. The results were obtained with a weighted GWAS in a sliding window of five adjacent SNPs in BLUPf90 (Misztal et al. 2015). **a** global dataset; **b** WEM × NAT; **c** WEM × WEM; **d** WEM × NEM; and e WEM × SEM

A group of SNPs on LG5 explained 4.37% of variance for sex in ♀WEM × ♂NAT, 3.12% in ♀WEM × ♂WEM, and 4.6% in ♀WEM × ♂NEM, while in ♀WEM × ♂SEM this peak was not observed. Peaks shared only between

Table 3 Identification of European sea bass chromosomes with a QTL explaining more than 2% of the variance in the global dataset and in each of the four offspring groups with the same paternal origin

	1A	1B	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18–21	19	20	22–25	24	*x*
Global			●				●	●					●			●								●
NAT		●		●		●			●	●	●		●				●			●				
WEM					●	●		●	●			●								●	●			
NEM						●				●	●	●				●								
SEM		●						●													●			●

Table 4 Total number of QTL for each paternal origin (in bold on the diagonal), total number of shared QTL between origins (below the diagonal) and total number of QTL that differ between origins (above the diagonal)

	NAT	WEM	NEM	SEM
NAT	**9**	10	8	11
WEM	3	**7**	8	7
NEM	3	2	**5**	9
SEM	1	2	0	**4**

♀WEM × ♂NAT and ♀WEM × ♂WEM were identified on LG8, explaining 2.18% and 4.06% of the variance, respectively, and on LG19, explaining 4.25% and 2.80% of the variance, respectively. The crossings ♀WEM × ♂NAT and ♀WEM × ♂NEM shared two peaks on LG9 (2.17% and 2.38% of variance explained, respectively) and on LG10 (4.97% and 3.36% of variance explained, respectively).

One group of SNPs, which was shared between ♀WEM × ♂NAT and ♀WEM × ♂SEM was identified on LG1B, explaining 2.18% and 3.53%, respectively. One group of SNPs that was in common between ♀WEM × ♂WEM and ♀WEM × ♂SEM was identified on LG11 (3.48% and 2.54% of variance explained, respectively). Furthermore, ♀WEM × ♂WEM cross share two groups of SNPs with ♀WEM × ♂SEM cross, on LG7 (variance explained of 2.41% and 2.22%, respectively) and on LG20 (3.47% and 3.78% of the variance, respectively).

The ♀WEM × ♂NAT cross showed specific peaks, that were not shared with any other paternal origin, on LG3 (2.41% of variance explained), LG12 (2.51%), and LG16 (2.03%). One specific peak was found in the ♀WEM × ♂WEM cross, as well (LG4, 2.55% of variance explained).

Two groups of SNPs were identified as specific to the Eastern Mediterranean populations, one in the ♀WEM × ♂NEM cross (LG15, 3.04% of variance explained) and one in the ♀WEM × ♂SEM cross (LG*x*, 2.59% of variance explained).

Interestingly, we did not identify any sex QTL explaining more than 2% of the variance that was common to all populations.

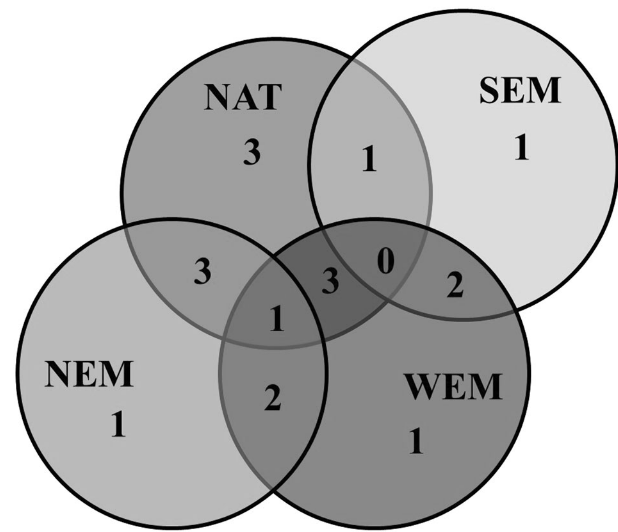


Fig. 4 Venn diagram showing the number of QTLs for sex explaining more than 2% of the variance that were specific for each paternal group or that were shared between groups of paternal origin

Discussion

In this study, we explored the genetic basis of the sex determination system in the European sea bass by implementing GWAS genome-wide association study approach in a factorial crossing experiment. For the first time, sea bass belonging to different origins across the whole distribution range of natural populations were compared to assess variation in the genetic architecture of sex, including a comparison between the Atlantic and Mediterranean sea bass lineages. We found different QTLs underlying sex determination between Atlantic and Mediterranean populations, with a gradient of similarities from Western to Eastern Mediterranean populations, reflecting the previously documented introgression of Atlantic genes within the Mediterranean genetic background (Guinand et al. 2017; Duranton et al. 2018). This finding is consistent with the hypothesis of a population-specific evolution of PSD polygenic sex determination systems in different evolutionary lineages occupying different environments.

An important result was the increased sharing of QTLs for sex determination in adjacent populations, which could result from an ongoing admixture between two evolutionary lineages (i.e. Atlantic and Mediterranean) characterized by different genetic architectures of sex determination systems. The detected geographical gradient in the architecture, from NAT to SEM, would then reflect the level of introgression and indeed corresponds to the admixture gradient recently found in sea bass population genomic studies (Duranton et al. 2018).

The ancestral architecture of the sex determination in sea bass might have evolved differently during the 300, 000 years of divergence between Atlantic and Mediterranean lineages, explaining the origin of the variation that now has population-specific influences on sexual determination. Indeed, we did not find any linkage group common to all populations with groups of SNPs explaining more than 2% of the variance, which support the hypothesis put forward by Guinand et al. (2017) that the most important genes affecting sex may differ between sea bass populations.

♀WEM × ♂NAT cross showed some similarities compared to ♀WEM × ♂WEM and ♀WEM × ♂NEM, that have gradually reduced in ♀WEM × ♂SEM. This finding can be more likely explained by the recent history of inter-basins connectivity, since Atlantic alleles have been progressively diffused from the Western to the Eastern Mediterranean since the end of the last glacial maximum (Tine et al. 2014; Duranton et al. 2018). The resulting longitudinal gradient of admixture across the Mediterranean populations makes the WEM population (31% of Atlantic ancestry) more similar to the Atlantic than the NEM and SEM population (13% of Atlantic ancestry) in most of the genome (Duranton et al. 2018).

Therefore, a gradient in similarity of genomic architecture is expected if sex determination QTLs introgress similarly to neutral genes. We do not reject, however, the possibility that differential adaptations between Atlantic and Mediterranean environments have also contributed to the patterns we observed, although this hypothesis is difficult to distinguish from historical admixture. Finally, the presence of a biogeographical barrier to gene flow located in the Siculo-Tunisian Strait (Quignard 1978; Bahri-Sfar et al. 2000), which limits the connections between Western and Eastern Mediterranean, may explain the further reductions of similarities between Western and Eastern Mediterranean populations.

In our case, the four paternal groups are all related by the WEM dams (i.e. all the individuals have 50% of the genome coming from WEM), with the result that even the Eastern Mediterranean group contains a higher level of Atlantic ancestry than what is expected in "pure" wild ♀NEM × ♂NEM or ♀SEM × ♂SEM crosses. This leads to the conclusion that the real differences existing in nature could be even stronger than what we observed here, due to our experimental design.

We interpreted our results as differences between male origins with the implicit assumption that they mostly reflect additive QTLs effects from the sire population of origin. Still, some QTLs could be due to the dam population (WEM). This is not the preferred hypothesis as no QTLs are shared between all paternal group, although they all share the same dams. Another possibility is that some of the QTLs observed are not linked to additive genetic variation but to dominance. The higher number of QTLs in the ♀WEM × ♂NAT cross could be indicative of dominant alleles involved in sex determination (especially since some heterosis in sex-ratio has been shown by Guinand et al. (2017) when mating Atlantic and Mediterranean individuals).

Within each paternal origin, we performed a cross validation analysis of QTLs, by removing four times 25% of the offspring along the second axis of the PCA (representative of variation between dams, see Fig. 2b). Most of the QTLs were identified in all subgroups, suggesting they were linked to the sire origin studied, but several of them were absent in some of the subgroups, which may be indicative of sire origin by dam interaction, i.e. dominance variation (see Supplementary material S2).

Parental effects on phenotypic sex were clearly significant: the dams ($\chi^2 = 34.72$, df = 8, P-value $= 3 \times \cdot 10^{-6}$) and sires ($\chi^2 = 124.56$, df = 59, P-value $= 10^{-6}$) variation for the proportion of females in the offspring was strongly different, similar to Vandeputte et al. (2007), where both sires and dams had a similar-size effect on the sex-ratio of the progeny.

The heritability of sex tendency we estimated in the present study through a linear mixed sire model was relatively high ($h^2 = 0.52 \pm 0.17$), similar to the estimate obtained for sire heritability by Vandeputte et al. (2007) on a larger dataset consisting of individuals of Northern Atlantic origin (0.52 ± 0.13), suggesting that the influence of the genetic and the environmental components on sex-ratio variance should be roughly equivalent.

The genetic correlation between sex and growth-related traits was significant ($h^2 = 0.69 \pm 0.12$ between sex and weight, $h^2 = 0.66 \pm 0.13$ between sex and length) and higher compared to previous studies (r_A between sex and size in the range of 0.23 and 0.59; Vandeputte et al. 2007; Palaiokostas et al. 2015). Overall, these results confirm the hypothesis of a strong link between genes affecting sex and growth (reviewed by Vandeputte and Piferrer 2018), with a clear sexual growth dimorphism (at the age of 180 dph, females were 34.4% heavier and 9.63% longer than males).

The sex-ratio in the global dataset was strongly skewed towards males, with a percentage of females less than half the percentage of males (32.5% versus 67.5%, respectively).

This is consistent with the general observation that cultured sea bass, because of the hatchery environment (especially temperature), show an unbalanced sex-ratio in favor of males (Saillant et al. 2003; Piferrer et al. 2005), different if compared to wild-born sea bass, where in younger fish the sex-ratio seems to be balanced (Vandeputte et al. 2012).

The percentage of females was slightly higher in Atlantic/West-Med populations compared to Eastern Mediterranean populations, suggesting a possible different tendency in sex-ratio related to the origin of the individuals under aquaculture conditions. This is consistent with the study by Guinand et al. (2017), where the ♀WEM × ♂Atlantic and ♀WEM × ♂WEM crosses showed a higher mean proportion of females compared to the ♀WEM × ♂NEM and ♀WEM × ♂SEM crosses.

A limitation in our study is represented by the fact that the between-populations variation of sex-ratio could be confounded by non-additive genetic effects, as NAT, NEM, and SEM broodstock were used only as sires. As previously reported by Guinand et al. (2017), sex-ratio can show non-additive components of genetic variance, and we have no possibility to disentangle additive and non-additive genetic effects in our case.

Finally, a better understanding of the genetic architecture of sex tendency in sea bass could have applications in aquaculture production. European sea bass is one of the most important marine species widely cultured in the Mediterranean areas and represents 49% of the marine Mediterranean aquaculture production (FEAP Annual report 2016). The strong bias towards males under aquaculture condition has been recognized by farmers as a problem for different reasons (lower growth rates of males compared to females, reduced flesh quality and general decrease of the commercial values of the product; Felip et al. 2006). Uncovering the population-specific sex determination system may help to produce stocks with higher proportions of females, through selective breeding and genomic selection. Moreover, the choice of broodstock coming from a specific origin could be interesting to start new breeding programs, due to the between-population differences in sex-ratio we found.

Data archiving

The dataset underlying our findings is available in the institutional public data repository (SEANOE: http://www.seanoe.org/), https://doi.org/10.17882/55576

Acknowledgements S.F. was supported by a grant from Valle Cà Zuliani Società Agricola s.r.l. (Pila di Porto Tolle, Italy). The experimentaldata and genotyping were derived from RE-SIST project “Improvement of disease resistance of farmed fish by selective breeding” selected at the 15th “Fonds interministériel”.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Bahri-Sfar L, Lemaire C, Ben Hassine OK, Bonhomme F (2000) Fragmentation of sea bass populations in the western and eastern Mediterranean as revealed by microsatellite polymorphism Proc R Soc Lond B Biol Sci 267:929–935

Bateman AW, Anholt BR (2017) Maintenance of polygenic sex determination in a fluctuating environment: an individual-based model J Evolut Biol 30:915–925

Bull JJ (1983) Evolution of sex determining mechanisms. Benjamin/Cummings, Menlo Park, CA

Bulmer MG, Bull JJ (1982) Models of polygenic sex determination and sex ratio control. Evolution 36:13–26

Conover DO, Heins SW (1987) Adaptive variation in environmental and genetic sex determination in a fish. Nature 326:496–498

Dempster ER, Lerner IM (1950) Heritability of threshold characters. Genetics 35:212–236

Doan QK, Vandeputte M, Chatain B, Haffray P, Vergnet A, Breuil G et al. (2017a) Genetic variation of resistance to viral nervous necrosis and genetic correlations with production traits in wild populations of the European sea bass (*Dicentrarchus labrax*). Aquaculture 478:1–8

Doan QK (2017b). Genetic and genomic variation of resistance to viral nervous necrosis in wild populations of European seabass (*Dicentrachus labrax*). PhD thesis, Université de Montpellier.

Duranton M, Allal F, Fraïsse C, Bierne N, Bonhomme F, Gagnaire P-A (2018) The origin and remolding of genomic islands of differentiation in the European sea bass. Nat Commun 9:2518

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longman, Harlow, England

Fauvel C, Suquet M, Dreanno C, Menu B (1998) Cryopreservation of sea bass (*Dicentrarchus labrax*) spermatozoa in experimental and production simulating conditions. Aquat Living Resour 11:387–394

Felip A, Zanuy S, Carrillo N (2006) Comparative analysis of growth performance and sperm motility between precocious and non-precocious males in the European sea bass (*Dicentrarchus labrax*, L.). Aquaculture 256:570–578

Groeneveld E, Kovac M, Mielenz N (2010). VCE user’s guide and reference manual version 6.0. https://vce.tzv.fal.de/

Guinand B, Vandeputte M, Dupont-Nivet M, Vergnet A, Haffray P, Chavanne H et al. (2017) Metapopulation patterns of additive and nonadditive genetic variance in the sea bass (*Dicentrarchus labrax*). Ecol Evol 7:2777–2790

Habier D, Fernando RL, Kizilkaya K, Garrick DJ (2011) Extension of the bayesian alphabet for genomic selection BMC Bioinforma 12:186

Heule C, Salzburger W, Böhne A (2014) Genetics of sexual development: an evolutionary playground for fish. Genetics 196:579–591

Lemaire C, Versini JJ, Bonhomme F (2005) Maintenance of genetic differentiation across a transition zone in the sea: discordance between nuclear and cytoplasmic markers. J Evol Biol 18:70–80

Liew WC, Bartfai R, Lim Z, Sreenivasan R, Siegfried KR, Orieux N (2012) Polygenic sex determination system in zebrafish. PLoS ONE 7:e34397

Lourenco DAL, Tsuruta S, Fragomeni BO, Masuda Y, Aguilar I, Legarra A, Bertrand JK et al. (2015) Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. J Anim Sci 93:2653–2662

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA

Mank JE, Promislow DE, Avise JC (2006) Evolution of alternative sex-determining mechanisms in teleost fishes. Biol J Linn Soc 87:83–93

Menu B, Peruzzi S, Vergnet A, Vidal MO, Chatain B (2005) A shortcut method for sexing juvenile European sea bass, *Dicentrarchus labrax* L Aquacult Res 36:41–44

Misztal I, Tsuruta S, Lourenco D, Aguilar I, Legarra A, Vitezica Z (2015) Manual for BLUPF90 family of programs. University of Georgia, Athens, USA

Moore EC, Roberts RB (2013) Polygenic sex determination. Curr Biol 23:R510–R512

Naciri M, Lemaire C, Borsa P, Bonhomme F (1999) Genetic study of the Atlantic/Mediterranean transition in sea bass (*Dicentrarchus labrax*). J Hered 90:591–596

Palaiokostas C, Bekaert M, Taggart JB, Gharbi K, McAndrew BJ, Chatain B et al. (2015) A new SNP-based vision of the genetics of sex determination in European sea bass (*Dicentrachus labrax*). Genet Sel Evol 47:68

Piferrer F, Blazquez M, Navarro L, Gonzalez A (2005) Genetic, endocrine, and environmental components of sex determination and differentiation in the European sea bass (*Dicentrarchus labrax* L.). Gen Comp Endocrinol 142:102–110

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D et al. (2007) PLINK: aA tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575

Quéré N, Desmarais E, Tsigenopoulos CS, Belkhir K, Bonhomme F, Guinand B (2012) Gene flow at major transitional areas in sea bass (*Dicentrarchus labrax*) and the possible emergence of a hybrid swarm. Ecol Evol 2:3061–3078

Quignard JP (1978) La Méditerranée: creuset ichtyologique. Boll Zool 45:23–26

R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/

Saillant E, Fostier A, Menu B, Haffray P, Chatain B (2001) Sexual growth dimorphism in sea bass *Dicentrarchus labrax*. Aquaculture 202:371–387

Saillant E, Fostier A, Haffray P, Menu B, Laureau S, Thimonier J et al. (2003) Effects of rearing density, size grading and parental factors on sex ratios of the sea bass (*Dicentrarchus labrax* L.) in intensive aquaculture. Aquaculture 221:183–206

Sansone G, Fabbrocini A, Ieropoli S, Langellotti AL, Occidente M, Matassino D (2002) Effects of extender composition, cooling rate, and freezing on the motility of sea bass (*Dicentrarchus labrax* L.) spermatozoa after thawing. Cryobiology 44:229–239

Tine M, Kuhl H, Gagnaire P-A, Louro B, Desmarais E, Martins RST et al. (2014) The European sea bass genome and its variation provide insight into adaptation to euryhalinity and marine speciation. Nat Commun 5:5770

Van Dooren TJM, Leimar O (2003) The evolution of environmental and genetic sex determination in fluctuating environments. Evolution 57:2667–2677

VanRaden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci 91:4414–4423

Vandeputte M, Mauger S, Dupont-Nivet M (2006) An evaluation of allowing for mismatches as a way to manage genotyping errors in parentage assignment by exclusion. Mol Ecol Notes 6:265–267

Vandeputte M, Dupont-Nivet M, Chavanne H, Chatain B (2007) A polygenic hypothesis for sex determination in the European sea bass *Dicentrarchus labrax*. Genetics 176:1049–1057

Vandeputte M, Quillet E, Chatain B (2012) Are sex ratios in wild European sea bass (*Dicentrarchus labrax*) populations biased? Aquat Living Resour 25:77–81

Vandeputte M, Garouste R, Dupont-Nivet M, Haffray P, Vergnet A, Chavanne H et al. (2014) Multi-site evaluation of the rearing performances of 5 wild populations of European sea bass (*Dicentrarchus labrax*). Aquaculture 424-425:239–248

Vandeputte M, Piferrer F (2018). Genetic and environmental components of sex determination in the European sea bass (*Dicentrarchus labrax*). In: Wang HP, Piferrer F, Chen SL (eds) Sex control in aquaculture: theory and practice. Wiley-Blackwell Hoboken, NJ, USA (in press).

Wang H, Misztal I, Aguilar I, Legarra A, Muir W (2012) Genome-wide association mapping including phenotypes from relatives without genotypes. Genet Res 94:73–83

Warnes GR, Bolker B, Lumley T, Johnson RC (2015). *gmodels*: various R programming tools for model fitting. Available at: https://CRAN.R-project.org/package=gmodels.

Wilson CA, High SK, McCluskey BM, Amores A, Yan Y, Titus TA et al. (2014) Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. Genetics 198:1291–1308

Zhang X, Lourenco D, Aguilar I, Legarra A, Misztal I (2016) Weighting strategies for single-step genomic BLUP: an iterative approach for accurate calculation of GEBV and GWAS. Front Genet 7:151

Zhang Z, Liu J, Ding X, Bijma P, de Koning D, Zhang Q (2010) Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix PLoS ONE 5: e12648

Acknowledgements – Ringraziamenti

Prima di tutto un ringraziamento speciale ai miei supervisori, Paolo Carnier, che ha saputo guidarmi con determinazione e che in poco tempo è riuscito ad insegnarmi una quantità di cose che non avrei mai immaginato; solo grazie al tuo aiuto prezioso questa tesi ha acquisito forma e contenuto, ed avrei tanto voluto avere più tempo per imparare di più da te; Daniela Bertotto, il mio angelo buono, non so nemmeno come ringraziarti tanto è stato il tuo incoraggiamento durante questi 3 anni; mi hai insegnato tantissimo, ma soprattutto mi hai trasmesso tutta la passione possibile per questo lavoro e mi hai motivata a non mollare mai, le giornate/nottate insieme in laboratorio ad aliquotare sieri o in impianto a campionare branzini rimarranno per sempre tra i miei ricordi.
Un dovuto ringraziamento a chi poi ha contribuito con il proprio lavoro alla realizzazione di questo progetto: Stan Laureau, Franco Furlan, Mauro Bovolenta e Matteo Freguglia di Valle Cà Zuliani; Francesco Pascoli, Anna Toffan, Marco Penzo ed il laboratorio di virologia dell'Istituto Zooprofilattico Sperimentale delle Venezie; Rafaella Franch, Massimiliano Babbucci, Giulia Dalla Rovere ed il laboratorio di Biologia Molecolare BCA, che sono stati iper efficienti, consentendoci di avere tutti i dati in tempo utile.

I ringraziamenti ora vanno a tutte le persone che in questi anni hanno speso parte del loro tempo ad aiutarmi, ascoltarmi, consolarmi, incoraggiarmi, consigliarmi, supportarmi e sopportarmi.
Ringrazio i miei genitori, Alessandro e Graziella, mio fratello Davide e Valentina. Spesso (purtroppo) distanti fisicamente, ma sempre presenti con il cuore e sempre pronti ad ascoltarmi, a consolarmi, ad asciugare le mie (tante) lacrime ed a supportarmi nelle mie scelte. Vi voglio bene.
Nicola. Non è semplice accettare le mie scelte, e nonostante tutto ci sei stato sempre, nei momenti di debolezza e sconforto più totale, così come nei momenti di felicità che questi anni hanno portato. Il tuo conforto ed il tuo incoraggiamento hanno spesso cambiato il mio atteggiamento verso i problemi, diecimila volte grazie per la tua pazienza, la tua bontà ed il tuo amore.
Laura, Elena, Tommaso, Giovanni e Carlo, non siete stati solo dei colleghi, siete diventati dei veri amici, vi ringrazio per quanto mi avete insegnato e per l'incoraggiamento che avete saputo darmi in ogni momento. Una menzione speciale va a Matteo, il mago dei server ed il risolutore dei problemi informatici, ma che nel corso del tempo si è rivelato anche un animo affine.
Gianfranco Gabai. Forse il primo a credere in me, tanti anni fa quando ancora ero una studentessa di triennale. Mi hai consigliato ed ascoltato talmente tanto durante questi anni, facendomi spesso vedere i problemi da una prospettiva diversa, grazie.
Camilla ed Andrea, my soul mates. Nonostante la distanza, la nostra amicizia si è rafforzata tantissimo ed ormai vi considero come una sorella ed un fratello. Mi avete sempre capita ed incoraggiata, avete condiviso con me gioie, dolori e follie di questo dottorato, grazie di esserci stati, ha significato moltissimo per me.
Francesca, in questi lunghi 11 anni ti ho sempre sentita vicina, anche se lavoro, vita ed impegni spesso ci tengono distanti. Grazie per il conforto che sei sempre in grado di darmi e sarai per sempre la testimone del primo paper "Accepted for publication".
I have to switch to English now, because this part is dedicated to the IFREMER family. There are so many people I would like to thank, Helena, Killian, Mathieu, Fred, Alain, Stéphane, Bastien, Benjamin, David, Marie Claire… I miss you guys, I think I fell in love with you the first day I came to Palavas and my experience in IFREMER, although not so long, meant a lot to me and I felt a little bit part of this family. A big big thanks to Béatrice Chatain, who allowed me to work in this amazing experimental station with the crazy Geneteam, Béatrice you are a pillar, I admire you so much and you are an inspiration to me; to Marc Vandeputte, what can I say? When I started my PhD I read so many papers written by you and I remember I thought "He's absolutely the God of the sea bass genetics", working with you has been a dream came true; to François Allal, you are not just a supervisor or a colleague, you are a friend and a great person; we've worked together a lot, and you've taught me a lot, but, more important, you've supported me so much and made me believe I could do something. I'm too sensitive, I know, but I would like to let you know that I appreciated everything and you will always be the "Superstar" to me.
Ho tenuto per ultima la stupenda persona a cui è dedicata questa tesi. Aria, tu sei il futuro, ed a te vorrei fare non una dedica ma un augurio. Ti auguro di trovare qualcosa che possa appassionarti e che ti faccia sentire felice nonostante tutto come ho avuto la fortuna di trovare io, e quando lo troverai lotta per questo. La vita spesso non andrà come ti saresti immaginata o come l'avevi programmata, non aver paura a cambiare i tuoi piani, non gettare mai la spugna e ricordati che per la zia sarai sempre una piccola guerriera.