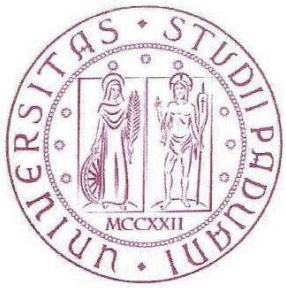


UNIVERSITY OF PADOVA

DOCTORAL SCHOOL IN
VETERINARY SCIENCE

PHD THESIS

GENOMIC PREDICTION OF RESISTANCE TO *PHOTOBACTERIUM DAMSELAE* SUBSP. *PISCICIDA* AND *SPARICOTYLE CHRYSOPHRII* IN GILTHEAD SEABREAM (*SPARUS AURATA*) USING 2B-RAD SEQUENCING

Candidate: Roberta Carraro

Roberta Carraro

Tutor: Tomaso Patarnello

XXX cycle

Index

Summary

Context: Gilthead seabream (*Sparus aurata*) is a highly important farmed fish species specifically in the Mediterranean aquaculture industry. Infectious diseases present a significant threat to the sustainability of aquaculture with high economic losses due to mortalities, reduced productivity, and the necessity of additional treatments/vaccinations. Specific and sensitive methods for the detection of fish pathogens represent useful tools to investigate infection dynamics and enable early detection of the disease for better prevention. Selection and breeding for resistance against infectious diseases is also a highly valuable tool to help prevent or diminish disease outbreaks, and applying genomic information to the currently available advanced selection methods could accelerate the response to selection. The gram-negative bacteria *Photobacterium damselae* subsp. *piscicida* (*Phdp*) and the ectoparasite *Sparicotyle chrisophrii* (*Sc*) are two of the most important pathogens affecting seabream cultivation. **Purpose of the study:** The aims of this work are: (i) to investigate the genomic prediction of resistance to two highly problematic diseases in seabream through the application of 2b-RAD with the objective of achieving selective breeding goals and (ii) to design an effective assay for the detection and quantification of *Phdp*. **Materials and methods:** (i) 1233 and 1001 seabream individuals were challenged trough intramuscular injection with a virulent strain of *Phdp* and by co-habitation with naturally *Sc*-infected seabreams, respectively. Animals were monitored daily and data of *dead/survived fish* and *number of parasites in the gills/body length* were recorded. Genomic DNA was extracted from the finfish of all individuals and used to construct 2b-RAD libraries. Data were analyzed in order to find SNP based genotypes (GATK, SAMtools), perform Genome-Wide Association Studies, estimate breeding values (ASReml 4.0) and construct linkage-maps (Lep-Map v2). (ii) A primer set was designed from a partial sequence of the *bamB* gene (Primer3 web) considering two SNPs that discriminate between *Phdp* and its strictly correlated subspecies *Phdd*. The assay was tested for specificity/sensitivity on laboratory-generated samples as well as on previous experimentally infected seabream tissue samples. **Results and discussion:** (i) The reference catalogue contained 175,725 and 269,660 tags for the *Phdp* and the *Sc* challenge, respectively. The SNP detection process yielded genotypic data for 19,313 and 21,773 quality SNPs for *Phdp* and *Sc*, respectively, both grouped into 24 linkage groups (LG), which are consistent with the karyotype of this species. Genomic heritability for resistance to photobacteriosis was 0.31-0.33 and genomic heritability for tolerance to *Sc* was 0.11-0.22, suggesting potential to enhance both resistances through family-based selection. Estimated breeding values (EBV) using genomic (GBLUP) information presented 5-43% higher accuracy in comparison to those measured using the only pedigree information (PBLUP). GWAS revealed a quantitative trait locus (QTL) including 7 SNPs at LG17 which presented significant association with resistance to *Phdp*, while one SNP (LG17) was found affecting tolerance to *Sc.* (ii) The molecular method proposed for *P. damselae* diagnosis, with high specificity and sensitivity, proved to be suitable for detection, quantification and subspecies identification in one-step, overcoming the limitations of previous assays. **Conclusions:** The SNPs discovered through 2b-RAD genotyping could be used to implement new marker-assisted selection programs for the generation of more resistant fish, preventing important disease outbreaks in fish farms. In addition, the original molecular method proposed holds the potential to improve the current knowledge of *Phdp* infection dynamics and the development of better strategies to control this important fish disease.

1. Introduction

1.1 How to perform selective breeding in aquaculture?

Animal breeding aim is '*the development of animals that will result in animal products economically more advantageous under the current environmental, social and economic conditions*' (Groen et al., 1995).
The selection of desirable characteristics is an essential tool to ensure continuous improvement of animals and plants and to achieve the maximum productivity and a high profit in farming. The main steps involved in a selective breeding program are:

A) Defined interested trait(s) genetic parameters
B) Establish base (parental) population
C) Define breeding strategy

A) The estimation of genetic and phenotypic parameters are prerequisites for starting effective breeding programs, specific to a given population, and aimed at improving the economic value and quality of the fish (Powell et al., 2008). Important genetic parameters are the *heritability* and the *genetic correlation*.
The heritability concept can be explained through the phenotypic and the genetic variance. Because quantitative phenotypes exhibit continuous distributions in a population, the only way to work with and to improve these traits is to analyze their variance and to divide it into the heritable and the non-heritable components. Phenotypic variance (VP), the variability that a phenotype exhibits in a population, is the sum of three components: genetic variance (VG), environmental variance (VE), and genetic-environmental interaction variance (VG-E). This can be represented by the following formula (Russel, 2002):

$$\mathbf{VP = VG + VE + VG\text{-}E}$$

Genetic variance is the component that breeders try to manipulate during a breeding program and it is the sum of additive genetic variance (VA), dominance genetic variance (VD), and epistatic genetic variance (VI). As before, this can be represented by a formula:

$$\mathbf{VG = VA + VD + VI}$$

Heritability is a concept which summarizes how heritable a phenotype of interest is, in particular with reference to the resemblance of offspring and parents (Visscher et al., 2008). Because additive genetic

variance is transmitted from a parent to its offspring in a predictable and reliable manner, if the percentage of phenotypic variance that is due to additive genetic variance is known, a farmer can predict the amount of improvement that can be made as a result of selection, and he can even customize selection to achieve a pre-determined amount of improvement per generation.

The proportionate amount of additive genetic variance is called *heritability*, and it can be represented by the following formula:

$$\mathbf{h^2 = VA/VP}$$

where: h^2 is the symbol for heritability, VA is additive genetic variance, and VP is phenotypic variance. Heritability is expressed as a percentage (0-100% or 0.0-1.0).

The *genetic correlation* is the proportion of variance that two traits share due to genetic causes. It can be positive or negative by assigning a number from +1 to -1 with 0 indicating no genetic correlation; the greater the displacement of the value from 0, the greater the correlation between traits. The genetic correlation of traits is independent of their h^2, thus two traits can have a very high genetic correlation even when the h^2 of each is low and viceversa (Russel, 2002).

B) A necessary condition of a successful fish genetic improvement program is starting with a population with ample genetic variation. Failure of some attempts to achieve genetic improvement with aquatic animals was likely due more to weaknesses in the base population than to the selection method utilized (Lind et al 2012). The base population for a selective breeding program can be chosen from wild populations or from the existing wild stocks, which may differ for commercially important traits such as growth rate, but also, for instance, in adaptation to specific environmental conditions, sex ratio, or resistance to infections (Vandeputte et al. 2014)

C) Define the best strategy. The current strategies involve:

- Individual selection (within-family selection or between-family selection)
- Combined selection

In *Individual selection*, family relationships are ignored, and comparisons are made among individuals, based only on the phenotype. While within-family selection uses the best individual from each family for breeding, between-family selection uses the whole family for selection. This kind of selection is most

effective when heritability is medium-high, for traits that can be measured on selection candidates (e.g. body size, shape), but it is not feasible for traits that can only be measured post-mortem (e.g. processing traits) or are risky to estimate in selection candidates (e.g. disease resistance) (www.fao.org). *Combined selection*, in a broad sense, means selection that is based on individual information as well as on information coming from relatives (*e.g.* full and half sibs, progeny). In this case, all of the additive genetic variance is available for selection and the use of pedigree information increases the accuracy of the estimation of breeding values (EBV), a statistical numerical prediction of the relative value of a particular individual available for breeding (www.fao.org). Accuracy of prediction can be estimate through a Best Linear Unbiased Prediction approach (BLUP) with the implementation of a pedigree-based matrix (pBLUP). In the case of aquaculture, all systematic (fixed) effects (e.g. batch, sex, production environment, age variation etc) associated with traits of interest can be accounted in the model to fit the data. Furthermore, in particular, relatives records can be used to estimate breeding values for traits that require slaughter of the animals (e.g. carcass and flesh quality traits) or that entail a risky challenge (e.g. disease resistance, tolerance to some environmental component).

1.2 Marker-Assisted Selection

During the 1990s, there was an attempt to incorporate DNA information, and genomic markers in particular, in breeding values predictions. This is what is called Marker-Assisted Selection (MAS) and the theoretical background was firstly introduced by Fernando and Grossman in 1989. More specifically, MAS is a process whereby a marker (morphological, biochemical or based on DNA/RNA variation) is used for indirect selection of a genetic trait related/responsible for the phenotype of interest (such as, for example, productivity, disease resistance, abiotic stress tolerance). Although the majority of the markers has no direct effects on cell function (i.e they do not change the protein sequence, or alter transcription), many are closely associated (in Linkage Disequilibrium; a non-random association of alleles at two or more loci) with alleles or genes responsible for quantitative traits. A quantitative trait is one that has measurable phenotypic variation owing to genetic and/or environmental influences. This variation can consist of discrete values or can be continuous; sometimes a threshold must be crossed for the quantitative trait to be expressed; this is common among complex diseases. A QTL is a genetic locus, the alleles of which affect this variation. Generally, quantitative traits are multi-factorial and are influenced by several polymorphic genes and environmental conditions, so one or many QTL can influence a trait or a phenotype (The Complex Trait Consortium, 2007).

The markers, acting as molecular "tags", allow producing genetic maps, therefore the identification and subsequent mapping of QTL through Genome-Wide Association Studies (GWAS). GWAS consist of a genetic association case-control study, which compare the frequency of the SNPs in individuals from a given population, with and without the trait, in order to determine whether a statistical association exists between the trait and the genetic marker (Clarke et al., 2011). However, in the 1990's, specific limitations of MAS such as i) small number of markers available ii) low fraction of genetic variation explained by QTLs and iii) the fact that results were only descriptive for the specific sire family where QTL analysis had been performed, resulted in difficulties in implementing MAS in practice. The limitations of MAS have been overcome by the recent developments in molecular technology. Nowadays, it is provided the possibility of sequencing the entire/reduced genome and discover a large number of genetic markers in the form of Single Nucleotide Polymorphisms (SNPs) at low/medium cost. In MAS, genetic markers and BLUP model are combined for the prediction of the EBV of the animals, resulting in a higher accuracy of prediction (gBLUP).

In practice, MAS is applied in a population that is different from the reference population in which the marker effects were estimated. It uses two types of datasets: a training set and a validation set. The training set is the reference population in which the marker effects were estimated; it contains: (1) phenotypic information (2) molecular marker scores and (3) pedigree information. Hence, marker effects are estimated based on the training set using certain statistical methods to incorporate this information; the genomic breeding value or genetic values of new genotypes are predicted based only on the marker effect. The validation set contains the selection candidates (derived from the reference population) that have been genotyped (but not phenotyped) and selected based on marker effects estimated in the training set (Meuwissen et al., 2001b).

The investigation of markers of traits of interests holds the potential to establish and develop new marker-assisted selection programs in aquaculture, as well as the identification of a large number of genes, elucidating their functions and interactions.

1.3 Disease resistance in aquaculture species

From literature, there is a high potential for increasing the efficiency and profitability of European aquaculture through genetic improvement but currently only less than the 10% of the aquaculture production considers this approach (Gjedrem et al., 2012; Gjedrem et al., 2016). In addition, the traits under selection are limited and incomplete, considering the high growth rate as the (only) preferential trait. Satisfying results have been achieved for the trait *body weight* (*growth*), which is easily selected using the individual's own performance, with high % of genetic gain (the amount of increase in performance) per generation in many species such as

Atlantic salmon, Tilapia, Common carp and others (Gjedrem et al., 2016). However, the increasing of the growth rate and thus the intensification of fish production, led to take into account other important characteristics: diseases resistance, mortality, fillet quality and adaptability to alternative feeds (plant-based). These traits are only selected in some species, due to complications in recording and selection methods that require established pedigrees. Table 1 reports the results of an online survey about the most frequently selected traits by the major aquaculture breeding companies operating in Europe (Chavanne et al., 2016), which are growth and morphology (35 and 23 breeding programs, respectively). Table 2 describes the current status of the aquaculture breeding industries for the six main species in Europe.

Table 1. Types and number of selected traits in the surveyed European breeding programs per each of the six main reared finfish species in the Mediterranean (Chavanne et al., 2016).

	Growth	Morphology	Disease resistance	Processing yield	Product quality	Maturity, fecundity	Feed efficiency[a]
Sea bream	6	4	2	0	2	0	1
Sea bass	6	4	1	2	1	1	1
Turbot	2	0	1	0	0	0	0
Rainbow trout	10	8	5	5	3	5	2
Salmon	7	3	7	6	6	2	2
Carp	4	4	2	1	0	1	0
All species	35	23	18	14	12	9	6

[a] Indirect genetic selection response assumed

Table 2. Aquaculture breeding industries status for the six main finfish species in Europe (www.fishboost.eu).

	A. salmon	R. trout	E. seabass	G. seabream	Turbot	C. carp
History of modern breeding	40 years	30 years	20 years	15 years	20 years	0
Regions of production in Europe	North and West	North and West	South and West	South	South and West	Est
Level of sophistication	High	High	Moderate	Moderate-High	High	Low
Main breeding traits	Growth deformation, fillet color, disease resistance	Growth, maturity age, carcass yield, deformation, fillet color	Growth, shape, sex ratio	Growth, shape, lipid deposition	Growth, shape	None
Type of pedigree tracing	Separate families	Separate families or genotyping	Genotyping	Genotyping	Separate families for genotyping	Genotyping possible
To enhance production	Diseases, FE, adaptation to novel diets	Processing yields, FE, diseases, adaptation to novel diets	Diseases, FE, fillet %	FE, fillet%, diseases	Fillet yield, diseases, adaptation to novel diets	Growth, winter survival, fillet yield, diseases
Need for genomics for next level of breeding	High	High	For diseases	For diseases	Moderate-high	Low

The most advanced specie is Atlantic salmon, for which selective breeding programs already exist for disease resistance, product quality and processing yields using pedigrees and sometimes molecular genetics (MAS). At the other end of the scale are common carp current breeding programs that do not go beyond the traditional approach of crossbreeding of lines with no selective improvement for industry related traits (www.aquabreeding.eu). Anyway, as evidenced in Table 3, there is a positive trend concerning the use of molecular tools in a larger number of programs from 2008 to 2016 (Chavanne et al., 2016).

Table 3. Number of programs using the molecular tools as reported in the 2016 (first number) and 2008 (second number) surveys, grouped by species (Chavanne et al., 2016).

Species	Genomic selection	Marker-assisted selection	Molecular fingerprinting	Genetic traceability
Gilthead sea bream	0–0	4–0	2–2	2–2
European sea bass	0–0	3–0	1–1	2–1
Turbot	0–0	0–0	2–0	0–0
Rainbow trout	0–0	1–0	6–1	3–0
Atlantic salmon	5–0	6–2	6–4	5–3
Common carp	0–0	0–0	0–0	0–1
Total	5–0	14–2	17–8	12–7

In particular, maintaining a high survival rate in aquaculture is crucial to economy, animal welfare and sustainability of the industry (Odegard et al., 2011). Currently, the most significant losses in aquaculture are due to infectious diseases caused by specific viral, bacterial or parasitic agents, causing mortality, productivity reduction, and the necessity of additional treatments and vaccinations. Selection and breeding for resistance against infectious diseases represent a highly valuable tool to help prevent or diminish disease outbreaks, and currently available advanced selection methods with the application of genomic information could pace up response to selection. The selection of genetically resistant fish has two potential advantages compared with the normal practices of prevention: the fish have low or no chance of developing the disease, and the presence of a uniform and limited proportion of resistant fish reduces the opportunity of transmission of the disease, decreasing the risk of infection for individuals susceptible to the bacterium (Antonello et al., 2009). In recent years, selective breeding for disease resistance has received more attention including species such as Atlantic salmon, rainbow trout, Atlantic cod, rohu carp and Pacific white shrimp (Odegard et al., 2011).

Heritability represents the potential for short-term response of a quantitative trait to selection and traditionally, it has been estimated through pedigree analyses or using individuals of known relationships established by experimental crosses (Stanton-Geddes et al., 2013). In recent years, multiple methods have been developed in the animal breeding literature that use large-scale genomic data to predict phenotypes (Meuwissen et al., 2001a; Van Raden 2008; Goddard et al. 2009; Campos et al. 2012) and estimate heritability based on the proportion of phenotypic variance explained by genomic information (genomic markers). Microsatellites have been the most widely used genomic markers to calculate the (genomic-based) heritability of traits since recently

(Antonello et al., 2009). However, microsatellites are low-resolution markers and have shown to be sufficient to detect the presence of genetic variance for highly heritable traits, but they are not sufficiently reliable to estimate genetic parameters (Coltman, 2005). On the contrary, genome-wide SNPs have proved to largely improving the heritability estimation (Speed et al., 2012). More than the past, encouraging levels of heritability have been found for resistance to important bacterial, viral and parasitic pathogens in many species (Table 4).

Table 4. Recent heritability estimates of resistance to bacterial, viral and parasitic pathogens in marine/aquaculture species.

Species	Pathogen/Disease	h^2	Authors
European Sea Bass	Viral Nervous Necrosis (VNN)	0.26	Doan et al., 2017
Blunt snout bream	Aeromonas hydrophila	0.33	Xiong et al., 2017
Rainbow trout	Bacterial cold water disease (BCDW)	0.23-0.35	Vallejo et al., 2017
Rainbow trout	Infectious pancreatic necrosis (IPN)	0.39	Flores-Mara et al., 2017
Haliotis rufescens	Withering syndrome disease	0.21-0.36	Brokordt et al., 2017
Nile Tilapia	*Streptococcus iniae*	0.52	Shoremaker et al., 2017
Nile Tilapia	*Streptococcus agalactiae*	0.38	Shoremaker et al., 2017
Meretrix petechialis	*Vibrio parahaemolyticus*	0.31-0.32	Liang et al., 2017
Nile Tilapia	*Flavobacterium columnare*	0.30	Wonmongkol et al., 2017

Selection based on results of challenge tests (resistance to a specific disease measured as survival under challenge testing with the pathogen) has shown to be very efficient (Gjøen et al. 1997; Storset et al. 2007) and many QTL for disease resistance have been discovered in different species (Table 5).

Table 5. Some QTL mapped for disease resistance traits in aquaculture species, with authors.

Species	Pathogen	References
Atlantic salmon (*Salmo salar*)	IPNV	Houston et al. (2008); Houston et al. (2009); Moen et al. (2009); Houston et al. (2012)
Atlantic salmon (*Salmo salar*)	ISAV	Moen et al. (2007)
Atlantic salmon (*Salmo salar*)	*Gyrodactylus salaris* parasite	Gilbey et al. (2006)
Turbot (*Scophthalmus maximus*)	*Aeromonas salmonicida*	Rodríguez-Ramilo et al. (2011)
Seabream (*Sparus aurata*)	*Photobacterium damselae* subsp. *piscicida*	Massault et al. (2011)
Asian seabass (*Lates calcarifer*)	VNN	Liu et al. (2015)
Rainbow trout (*Oncorhynchus mykiss*)	IPNV	Ozaki et al. (2011)

Thus, selection for disease resistance is very promising as it shows that it is possible to develop highly tolerant or possibly truly resistant strains to specific pathogen. The use of MAS with searches for QTL, is auspicious for reduction in mortality in the future.

On this basis, the present study could contribute to improve the (genetic) selective breeding. In particular, this study has the object to better understand the genetic mechanisms that govern the resistance (defined as the host's ability to reduce pathogen invasion or replication) to fish photobacteriosis and the tolerance (defined as an infected host's ability to maintain performance and fitness, Lipschutz-Powell et al., 2012) to *Sparicotyle chrysophrii* in the teleost fish seabream (*Sparus aurata*).

1.4 Seabream and fish photobacteriosis

Seabream (*Sparus aurata*) is a marine teleost belonging to the family of Sparidae. It is the most commonly farmed fish species in the Mediterranean, covering a major role in the economy of the fishing and aquaculture, with high commercial value conferred by the precious food quality. The majority of the commercialized seabreams come from aquaculture. The UE is the world's larger producer with Greece as the main producer followed by Turkey, Spain and Italy (Figure 1).

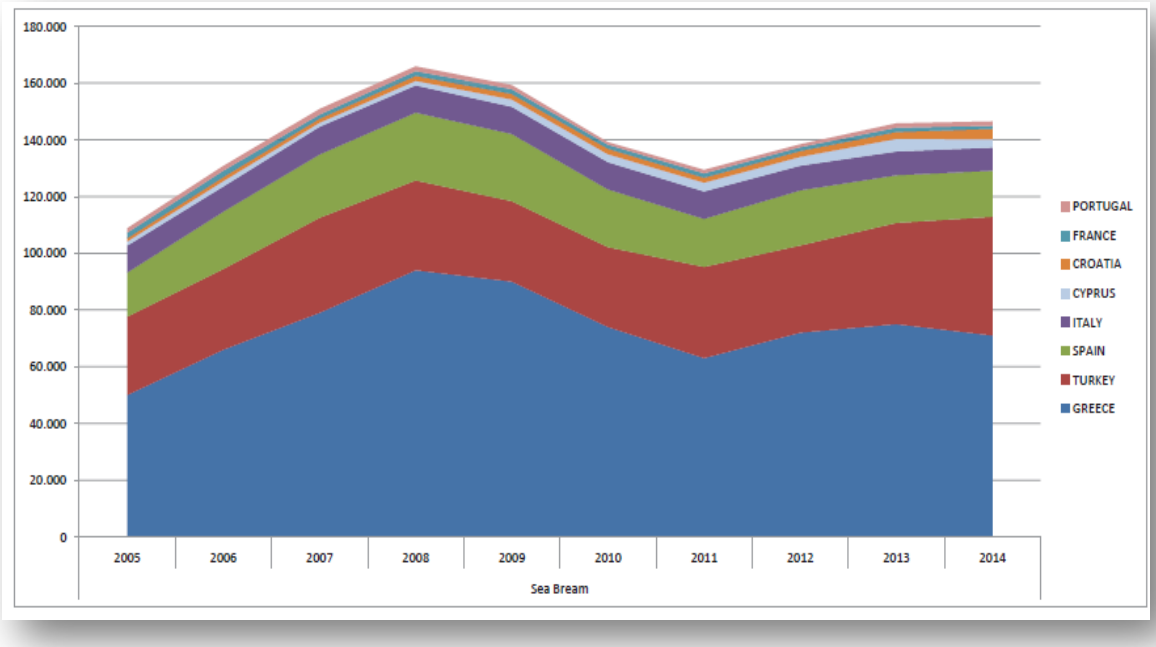


Figure 1. European production of seabream by aquaculture from 2005 to 2014 (FEAP, 2016).

The major health problem regarding the intensive farming of seabream is the infection with *Photobacterium damselae* subsp. *piscicida* (*Phdp*).

PHOTOBACTERIUM DAMSELAE SUBSP. PISCICIDA

Phylum Proteobacteria

Class Gammaproteobacteria

Order Vibrionales

Family Vibrionaceae

Genus Photobacterium

Specie P. damselae

Subsp. piscicida

This gram-negative bacteria is the responsible of fish photobacteriosis (once upon a time, Pasteurellosis), a septicemic disease which can cause a mortality rate of 90-100% of the infected fish, leading to serious economic losses for the breeders (Hawke et al., 2003). This septicemic disease may have either an acute or a chronic form (Noga J.E., 2011) and primarily affects young individuals, more sensitive to the bacteria (Andreoni and Magnani, 2014). As symptoms of the infection, fish show ataxic swimming, lethargy, absence of reaction to the human presence and signs of anorexia. At the level of the viscera (liver, spleen and kidney) necrotizing lesions, hemorrhages and free bacterial aggregates or within phagocytes are visible, in the capillaries and the interstitial spaces (Magarinos et al., 1994). Conditions favoring and aggravating the infection are high water temperature, low salinity, high population density, low dissolved oxygen, high level of pollution and the concomitant presence of opportunistic pathogens (Andreoni and Magnani, 2014).

1.4.1 Diagnosis of fish photobacteriosis

Because of this important disease, the diagnosis of the *Phdp* bacteria and its discrimination between its strictly correlated subsp. *damselae* (*Phdd*) and other co-infectious species is a big challenge. Although *Phdd* does not cause photobacteriosis, the bacteria is a primary pathogen causing ulcers and hemorrhagic septicemia in a variety of marine species as sharks, dolphins and shrimps, as well as wild and cultivated fish (Rivas et al., 2011). In addition, it has been associated to cause diseases in human (Abdel-Aziz et al., 2013); most of the reported infections in humans have their origin in wounds inflicted during the handling of fish, exposure to seawater and marine animals, and ingestion of raw seafood (Rivas et al., 2011). Despite causing different pathologies, *Phdd* and *Phdp* share their major epitopes and have identical 16S rRNA gene sequences (Osorio et al., 1999). Many molecular techniques have been proposed for the diagnosis of this fish-pathogen bacteria (Osorio et al., 2000; Romalde, 2002; Rajan et al., 2003; Zappulli et al., 2005; Amagliani et al., 2009; Martins et al., 2015) but currently none of these consists in a single step, since they involve two or more steps or, alternatively, a post-PCR manipulation. A single-step method able to simultaneously identify and quantify the presence of *Pdhp* without any post amplification manipulation would allow for a quicker and more accurate diagnosis as well as precise

monitoring of *Phdp* in the infected host. To date, no quantitative evidence is available on the dynamics of infection, neither during the acute form nor for the chronic one.

1.4.2 Control strategies

Antibiotics have been the first line of defense in fish aquaculture to control photobacteriosis outbreaks, but after only a few years, the pathogen acquired resistance to various antibiotics (kanamycin, sulphonamide, tetracycline, ampicillin, chloramphenicol, florfenicol and erythromycin, Andreoni & Magliani, 2014). Chemotherapy is indeed not feasible, due to the intracellular parasitism of the bacteria within macrophages. Vaccination based on inactivated bacterial pathogens plays an important role in large-scale commercial fish farming and has been successful for cultivation of many species like salmon, trout, European seabass and other species (Skjold et al., 2016). With particular reference to fish photobacteriosis in seabream, convention *Phdp* vaccines are based on inactivated products containing cellular (heat-o formalin killed bacteria) and soluble antigens (LPS and ribosomal formulations) for immersion and injection administration (Andreoni & Amagliani, 2014). However, these vaccines seem to have poor effect and to be ineffective against photobacteriosis (Andreoni & Amagliani, 2014). In addition, the vaccination itself might be laborious and costly, and vaccine induced side effects might occur (Midtlyng et al., 1996).

1.4.3 Genetic architecture of resistance to fish photobacteriosis in seabream

Selective breeding could help in reduce the disease pressure in the farmed and wild fish environment through increase of genetic gains. Despite the fact that the improvement of host resistance is gradual, such that several generations might be needed to obtain adequate protection towards the disease, the application of genomic to breeding is particularly valuable for disease resistance, which is typically expensive or impossible to measure on the selection candidates themselves. While MAS for major disease resistance loci has been well documented in Atlantic salmon (*Salmo salar*) breeding programs (Houston et al. 2008, 2010; Moen et al. 2009), few successful examples exist for other farmed finfish species (Palaiokostas et al., 2016). Resistance to fish photobacteriosis has been shown to be moderately heritable (h^2=0.22-0.28, Palaiokostas et al., 2016), highlighting the potential for selective breeding programs to improve this trait. In 2010, Massault and others identified two significant QTL for resistance to the disease. The results derived from a challenge experiment on 500 individuals by testing 151 microsatellite loci used as markers. The first QTL was located on linkage group (LG) 3 affecting late survival (survival at day 15). The second one, for overall survival, was located on LG21, which allowed to highlight a

potential marker (*Id13*) linked to disease resistance. A significant QTL was also found for body length at death on LG6 explaining 5–8% of the phenotypic variation (Massault et al., 2010). This study has also allowed researchers to improve the seabream genomic map previously produced by Franch and colleagues in 2006 (204 microsatellites markers, 1241 cM) that describes the genetic loci and interesting markers for disease resistance, as well as the distance between loci on the chromosomes. Tsigenopoulos et al (2013) produced an updated second-generation linkage map for seabream, by the use of microsatellites, Expressed Sequence Tags and a few SNPs (321 genetic markers, 1769.7 cM). However, these studies were restricted by the low resolution of microsatellite markers, as compared to the high-density SNPs genotypes offered by SNP arrays (e.g., Houston et al. 2014) or genotyping-by-sequencing approaches (Davey et al. 2011). Recently, Palaiokostas and collaborators (2016) presented a new linkage-map based on 12.085 SNPs discovered through a restriction-based genomic technique named 2b-RAD (Wang et al., 2012) with some modifications (Pecoraro et al., 2016). The challenge test was conducted on 825 seabream individuals by immersion with a virulent strain of *Phdp*. Although no significant SNPs associated with disease resistance were found, some putative suggestive QTL were identified on LGs 1-3, 10, 17, 20 and 21.

The results highlighted the potential for selective breeding programs to improve this trait and advances in selection accuracy and genetic gain could be achieved through the use of genetic markers.

Furthermore, data from literature show that there is space for better investigate the genetic architecture of *Phdp* resistance and locate major QTL as candidates for MAS. The identification of genes underlying these major QTL may also lead to an improved understanding of the biology of bacterial resistance in seabream.

1.5 Seabream and *Sparicotyle chrysophrii*

As a consequence of the intensification of aquaculture in almost all the regions of the world, parasites have recently been highlighted at serious pathogenic problems in cultured fish (Rabet et al., 2016). In particular, *Sparicotyle chrysophrii (Sc)*, originally called *Microtyle chrysophrii*, is considered as one of the main disease threats to Mediterranean bream aquaculture (Rigos et al., 2015).

SPARICOTYLE CHRYSOPHRII
Phylum Platyhelminthes
Class Monogenea
Order Mazocraeidea

Family Microcotylidae

Genus Sparicotyle

Specie Chrysophrii

From a morphological point of view, this ectoparasite shows an elongated and flatted body with the dimensions of approximately 3.2-3.7x0.3-0.4 mm and a robust specialized organ with numerous clamps for the attack in the gills' host (50 clamps/side). Its pathological effects have been well documented, and symptoms include lethargy due to hypoxia and severe anaemia leading to a reduction in the health of stocks, and active feeding on mucus and epithelial cells of infected host fish by large populations of monogeneans which can induce histopathological damage and haemorrhages, necrosis and deletion of filaments due to the insertion points of hooks (Antonelli et al*.,* 2010). Among the factors that determine the infection by *Sc*, the temperature is the most important with the highest prevalence during the winter (Gonzalez-Lanza et Al., 1991; Robertsen et al., 2008). The parasite has been reported to be responsible for reduced catches in some wild populations, by altering their behavior and making them more susceptible to predation (Shirakashi et al., 2008). Mortality levels induced by the parasite in farmed conditions are typically low, but the probability of fish massive mortality can increase in case of mixed-infections with other parasites and secondary bacterial infections. A serious case of mortality by *Sc*, aggravated by the concomitant infection with *Polysporoplasma sparis*, occurred in June 2005 in a seabream intensive breeding in floating cages along the Greek coasts (Caffara et al., 2005). Some years later, Susini et al (2016) isolated both *Sc* and *Phdp* from farmed dead seabream in July 2010, Italy. Other co-infections events were reported also in the 1990's (Padrós & Crespo 1995; Cruz et al., 1997). Cases of mortality induced by the only parasites have also been reported (Sanz 1992; Alvarez-Pellitero 2004; Vagianou et al., 2006). However, the more common losses in aquaculture associated to sparicotylosis have to be attributed to a notable growth reduction of the stock due to the emaciated and anemic state of the survivors (Sitjà-Bobadilla & Alvarez-Pellitero, 2009).

Parasite invasion in sea cages is more difficult to control than in other type of facilities. The limited use and efficacy of chemicals and drugs for treatment makes the situation even more difficult to handle (Sitjà-Bobadilla et al. 2006). Effective vaccines do not exist and the ubiquity of the parasite in the seawater suggests that management and biosecurity strategies for control can be of limited benefit. Therefore, economic losses result from both the reduction of growth and general immunosuppression of the stocks, in addition to the costs of treatments. Since no estimates of genetic variation or known QTL are available at this point in time, this gill parasite is a new candidate for evaluation of the genetic basis of fish tolerance, and inclusion in selective breeding goals.

1.6 2b-RAD

In the (post) Next-Generation Sequencing era, discovering SNPs as molecular markers of disease resistance/tolerance necessarily requires the implementation of a rapid, cost-effective genomic tool able, at the same time, to identify a high number of SNPs (10^3). In recent years, Restriction-site-associated DNA (RAD) technology has become a powerful and useful approach for the identification, at reasonable costs, of thousands of genetic markers randomly distributed across the target genome, using Illumina technology (Davey and Blaxter, 2010). RAD-seq covers a range of Genotyping-By-Sequencing (GBS) techniques which combine the use of genome complexity reduction with restriction enzymes and the high sequencing output of NGS technologies (Robledo et al., 2017). The original RAD-seq protocol, described by Miller et al (2007) and Baird et al (2008), successfully found large applications in genotyping and SNP discovery, as well as in quantitative genetics and phylo-geographic studies (Davey and Blaxter, 2010). Several variants of this methodology have subsequently been described: 2b-RAD (Wang et al., 2012), dd-RAD (Peterson et al., 2012), ezRAD (Toonen et al., 2013) and SLAF-seq (Sun et al., 2013). While 2b-RAD and dd-RAD have been extensively used in aquaculture genetics research, ezRAD and SLAF-seq introduced minor modifications, which do not confer a major advantage for aquaculture application (Robledo et al., 2017). The technical protocols and the main features of RAD, 2b-RAD and dd-RAD are described in Table 5.

Table 5. **Description of the technical protocols, advantages and disadvantages of the most used RAD-technologies (Puritz et al., 2014; Robledo et al., 2017).**

Technique	Main steps	Advantages	Disadvantages
RAD	Genomic DNA is first digested with a single restriction enzyme. Barcode containing adapters are then ligated onto digested 5′ ends. Ligated genomic DNA is then sonicated, and a 3′ adapter is ligated to the randomly sheared end. After ligation, the library is size-selected. Finally, RAD fragments with both adapters properly ligated are enriched with PCR.	Paired-end contigs PCR duplicate removal	Complex library preparation
2b-RAD	Genomic DNA is digested with a IIb-restriction enzyme generating fragments of identical size. Adapters are	No size-selection step High reproducibility Easy library preparation Strand bias detection	Short fragments Removal of PCR duplicates not possible

	ligated into protruding ends and 2b-RAD tags are enriched with PCR, during with a barcode is inserted.		
dd-RAD	Genomic DNA is simultaneously digested with two restriction enzymes. Barcoded P1 adapters and P2 adapters are ligated onto digested fragments in a single sticky-end ligation. Samples are then pooled and size-selected. Lastly, PCR is used to enrich the library and also to introduce a second barcode in the form of an Illumina index, increasing multiplexing potential.	Can multiplex many samples Easy library preparation Flexibility over SNP density	Repeatability dependent on size-selection step

The 2b-RAD variant, in particular, has been widely used for genetic studies in aquaculture species for its flexibility and simplicity. For instance, it avoids the time-consuming and potentially error-prone size-selection step, which is substituted by the use of IIb-restriction enzymes facilitating the sampling and sequencing of identical sites across individuals (Robledo et al., 2017). Also, the procedure can be customized to represent less loci in the genome by the use of base-selective adapters (Puritz et al., 2014). This method permits parallel and multiplexed samples sequencing that can be customized as well, by multiplexing different numbers of samples, based on the purpose. The production of individually barcoded libraries allows having a major control on equal representation of individual samples during the pooling (Robledo et al., 2017). The 2b-RAD technique has been the protagonist in many different genetic studies involving important species for aquaculture. The potential of the 2b-RAD technique in investigating population genetic structure in the non-model specie yellowfin tuna (*Tunnus albacores*, Pecoraro et al., 2016) assessed that 2b-RAD is a powerful tool for assigning genetic divergence in this marine fish. 2b-RAD was applied to test genomic prediction in a limited number of Yesso Scallop (*Patinopecten yessoensis*) families (Duo et al., 2016). With particular reference to QTL detection and GWAS, the technique was used to identify putative markers/QTL associated to mandibular prognatism in the European seabass (*Dicentrarchus labrax*, Babbucci et al., 2016), leading to the identification of three significant QTL and two candidate SNPs associated to this skeletal anomaly. In 2016, Palaiokostas and others applied the technique on 825 seabream (*Sparus aurata*) individuals for genomic prediction of resistance of this specie to pasteurellosis. Beyond aquaculture, 2b-RAD was described as suitable for other kind of applications such as microbial ecology (Pauletto et al., 2016), population genetics (Maroso et al., 2016) and identification of SNPs

associated to traits of interest in other species such as the resistance to ivermectin in the nematode *Haemonchus contortus* (Luo et al., 2017).

1.7 The FISHBOOST project

This work is part of the E.U project FISHBOOST which aim is to improve the efficiency and profitability of European aquaculture by advancing selective breeding to the next level for each of the six main finfish species through collaborative research with industry. By the involvement of 26 partners from the whole Europe (universities, research facilities and industries), this 5-years project is investigating the genetic basis to resistance to the main problematic diseases in the six main finfish Mediterranean species. In particular, breeding programs will be developed for resistance to Pancreas Disease (PD) in Atlantic salmon, Koi Herpes Virus (KHV) in Common carp (*Cyprinus caprio*), Viral Nervous Necrosis (VNN) virus in European Seabass (*Dicentrarchus labrax*), *Sparicotyle chrysophrii* and fish photobacteriosis in Seabream, Flavobacteriosis in Rainbow trout (*Oncorhynchus mykiss*) and Scuticociliatosis in Turbot (*Scophthalmus maximus*). Other important goals for the project are to search and validate non-lethal ways of improving feed efficiency, fillet% and lipid deposition on the same species and to quantify the potential for improving rainbow trout performance on multiple future diets composed of alternative plant- protein sources. Economic assessment of breeding programs and training, communication and dissertation of delivering impact are therefore aspects that are considered in FISHBOOST.

2. Objectives

- To develop a fast and accurate method for detection and quantification of *Phdp* in fish tissues

- To investigate the genomic prediction of resistance/tolerance of two highly problematic diseases in seabream in a view of achieving selective breeding goals through the application of new genomic tools (2b-RAD)

3. Materials and Methods

3.1 Challenge experiments and data collection

Two challenge experiments were performed. For better clearness, the number (1) refers to *Phdp* experiment, while number (2) to the *Sc* challenge.

(1) Nearly 1300 *S. aurata* individuals of similar body weight (3-5 gr) derived from 109 parents (84 males x 25 females) were transferred to the experimental aquarium of the IZSVe (Legnaro, Italy) and kept for 30 days in rectangular fiberglass 4000 L tanks supplied with re-circulating aerated seawater (30 ppt salinity). The animals were fed with a commercial pellet diet (Neo Supra Al4g, LeGoussant) supplemented with vitamins, antioxidants and oligo-elements. Some (n= 15) sea bream individuals were randomly sampled within the experimental population and euthanized by an overdose of MS-222 (100 mg/L, Sigma) to check for the presence of *P.damselae* sp. and other bacteria, parasites, or viruses. In addition, total DNA was extracted from the spleen following standard procedures. All DNA samples were subjected to a PCR-RFLP test to detect the presence of *Phdp* as described in Zappulli et al. (2005).

The specific dose to inject was pre-determined based on both immersion and injection of serial dilutions of *Phdp* strain 249/ITT/99, a field pathogenic isolate, on a total of 200 animals, in order to determine the Lethal Dose 50, that is the specific dose that induces the mortality of the 50% of the fish. No mortality was recorded during the immersion challenge test; on the contrary, the injection method has shown to be suitable to carry out the whole experiment. Thus, experimental infection was carried out on 1233 sea breams by intramuscular injection with 100 µl per fish of 10^3 CFU of *Phdp* 249/ITT/99. A group of 67 uninfected fish represented the control group and received intramuscular injection with PBS 0.01 M. After pathogen exposure, fish were monitored twice a day over an 18 days period in order to register mortalities. Survivor animals were euthanized with an overdose of MS-222 (100 mg/L) at 18 days post-infection. The finfish was collected from all individuals (Dead, DD and survived, SV) and stored in ethanol 85% at +4°C. In addition, during the period with the peak of mortality, target organs including spleen, liver, and gills were collected from 15 dead fish (DD, 5 fishes/day) and stored in RNAlater at -20°C to test the real-time PCR assay. For the same reason, eight days after the end of the mortality peak (18 days post challenge), the spleen was collected from 130 survived fish (SV) and stored in RNAlater at -20°C.

(2) Nearly 1200 offspring derived from 119 parents were transferred to the facilities of the Hellenic Centre of Marine Research (HCMR, Athens, Greece) for the experiment. Fish were splitted into two group: one group (n=

1001) was used as a challenge population and the other one (n=199) was used as unchallenged controls in the disease experiment. Individuals were placed in specifically-designed nets for the experiment and reared under commercial conditions with provided diets. Fish were tagged in the farm and the weight of challenge fish and controls were measured before their transfer into the experimental facilities (BW1) and after the challenge (BW2). The naïve bream were challenged following the infection model proposed by Rigos et al. (2015), that is using a cohabitation model with 50-60g bream (n 250) naturally infected with the monogenean. A small number of challenge fish were slaughtered throughout experiment to monitor parasite load and disease progression. The experiment lasted for 68 days and the number of adults parasites in the two external gills arches (ParCount) was measured per each SV fish at the end of the experiment over ten days (~80-100 fish per day). ParCount values were log transformed to make them normally distributed, and these values were then used in all the analyses. The Specific Growth Rate per day (SGR) was also calculated by using the following formula:

$$\text{Specific growth rate} = (\ln BW2 - \ln BW1)/68$$

where 68 is the duration of challenge test. Fin tissue samples were collected from all fish (including parents) for DNA extraction and genotyping.

3.2 DNA extractions and 2b-RAD libraries preparation

Prior to the library construction, a "in silico" digestion with the *AlfI* enzyme (based on the recognition site) on a draft of the seabream reference genome was performed to evaluate the theoretical number of expected tags/SNPs and to set the best pooling strategy for sequencing.

Total DNA was extracted from all tissue samples (~20 mg) using a commercial kit (DNA Tissue HTS 96 Kit, Invisorb, Germany) following the manufacturer's instructions. DNA concentration was determined using a Qubit fluorimeter with a dsDNA BR Assay (Invitrogen, California) and DNA quality was assessed by agarose gel electrophoresis. The 2b-RAD libraries (n ~ 3000) were constructed for each individuals following the protocol from Wang et al. (2012) with minor modifications described below. Ninety-six individuals were processed at a time and for two individuals per plate, libraries were repeated three times. Template DNA for each individual (300 ng) was digested in 6 µl reaction volume using 1.4 U *AlfI* (Thermo Fisher Scientific, USA) at 37° for 1 h, followed by enzyme heat inactivation at 65° for 20 min. The ligation reaction was performed by combining 6 µl of digested DNA with 20 µl of a ligation master mix containing 0.4 µM each of two library-specific adaptors with fully degenerate cohesive ends (5' -NN- 3'), 10 mM ATP (New England Biolabs, USA), and 1 U T4 DNA ligase (SibEnzyme Ltd., Siberia). Ligation was carried out at 16° for 3 h, with subsequent heat inactivation for 10 min at

65°. Sample-specific barcodes were designed through a Barcode Generator program (http://comailab.genomecenter.ucdavis.edu/index.php/Barcode_generator). PCR reactions were firstly prepared in a volume of 50 µl containing 12 µl of ligated DNA product, 0.5 µM of each primer (P4 and P6-BC, Eurofins Genomics S.r.l, Italy), 0.2 µM each primer (P5 and P7, Eurofins Genomics), 25 mM dNTPs (New England Biolabs, NEB, Ipswich, Massachusetts, USA), 1X Phusion HF buffer, and 1 U TaqPhusion high-fidelity DNA polymerase (New England Biolabs)and then split into three separate reactions of 16.6 µl each.

2b-RAD tags were amplified using the following cycling conditions: 98 °C for 4 min, 13 cycles of. 98 °C for 5 s, 60 °C for 20 s, 72 °C for 5 followed by 5 min at 72 °C. Adaptor and primer sequences were those reported in Wang et al., 2012. The quality of all amplicon libraries was checked at 1.8% agarose gel and then purified using the SPRIselect purification kit (Beckman Coulter, Pasadena, CA). The concentration of the purified libraries was quantified using a Qubit dsDNA BR Assay Kit (Invitrogen, USA) and Mx3000P qPCR Instrument. Additionally, the quality of 10% of randomly selected libraries was also assessed by running them on an Agilent 2100 Bioanalyzer.

3.3 Multiplexing strategy

In order to decide the best pooling strategy, 7 libraries randomly sampled from both the experiments were mixed following two different strategies (1X and 2.4X). More specifically, two pools were prepared. The first one (pool_1) consisted of 7 libraries pooled into equimolar amounts, while in the second one (pool_2) the 7 same libraries were mixed with other 10 random libraries. Finally, the two pools were merged into equimolar amounts (pool_test) and sequenced on a single lane of an Illumina HiSeq2500 platform (Illumina, San Diego, CA) using a 50 base single-end approach. The final “representation” of the libraries was 1X and 2.4X when inserted in pool_2 and in pool_1, respectively. Figure 2 summarizes the experimental scheme.

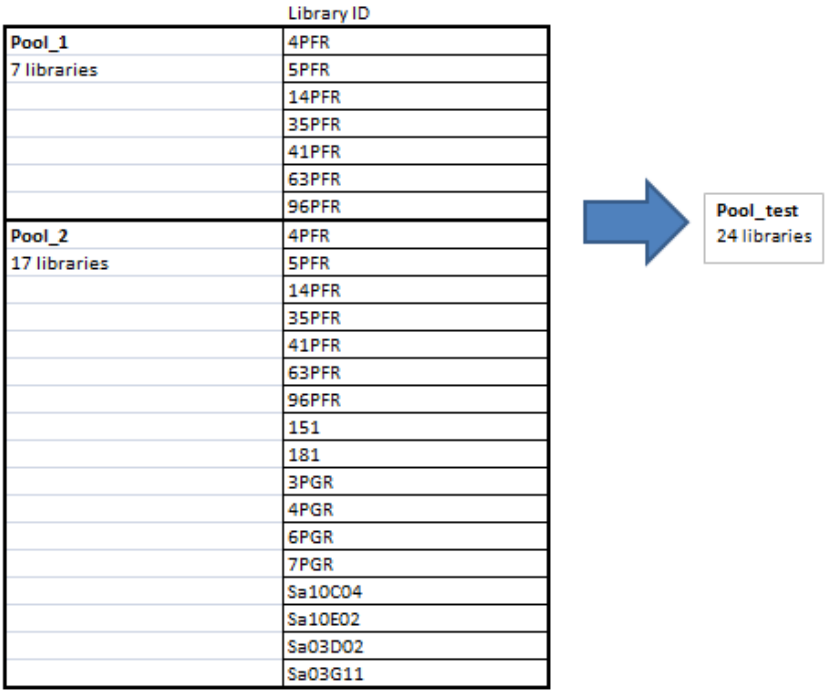


Figure 2. Multiplexing strategy test: scheme of the two pools and IDs of the libraries composing each pool.

Demultiplexed reads returned by the sequencing facility were quality-checked with the software FastQC (http://bioinformatics.babraham.ac.uk) and trimmed with a customized script from 2b-RAD pipeline v2.0 (Wang et al., 2012; Babbucci et al., 2016; Palaiokostas et al., 2016; Pauletto et al., 2016; Pecoraro et al., 2016) with a final read length of 36-bp. An exploratory analysis through the software Stacks v.1.32 (Catchen et al., 2011, 2013) was carried out to understand the loss of information in terms of number of SNPs between the two multiplexing strategies. Filtered reads were analyzed by the script denovo_map.pl which allows genotype inference through the identification of SNP loci without the reference genome using the following parameters: -m 10, -m 3, -n 2.

3.4 Sequencing

Based on the test's results, all the individual libraries were pooled into equimolar amounts by adopting two different multiplexing strategies for parents (64 libraries per pool) and offspring (128 libraries per pool). The quality of each pool was verified on Agilent 2100 Bioanalyzer. Finally, pooled libraries were sequenced on an Illumina NextSeq500 platform (Illumina, San Diego, CA) using 50 base single-end sequencing (v6 chemistry) at the BMR Genomics facility (Padova, Italy).

In a view of achieving a reasonable coverage, a threshold for acceptability of the sequencing results was decided considering the *AlfI* expected tags (n 160.000) and all the samples that did not exceed the threshold were re-sequenced.

Threshold for parents 3.8 Million of reads ~ 24X mean theoretical coverage

Threshold for offspring 1.8 Million of reads ~12X mean theoretical coverage

3.5 Design of the Real-time assay and sensitivity/specificity assessment

All proprietary and publicly available (GenBank accession number AY191100-AY191121) nucleotide sequences of partial *bamB* gene, (responsible for the outer membrane protein assembly factor bamB) from *P. damselae* were employed in order to find single nucleotide polymorphisms (SNPs) differentiating *Phdp* and *Phdd* subspecies. A total of 31 sequences belonging to *Phdd* (n 13) and to *Phdp* (n 18) were retrieved including, for both subspecies, strains isolated in Europe, USA and Japan. Multiple alignment was performed using Mega software (Molecular Evolutionary Genetics Analysis, Tamura et al. 2013) with the ClustalW algorithm (Supplementary file 1). Two SNPs at position 121 and 142 of the alignment were identified as able to discriminate all strains of *Phdp* from the closely related subsp. *Phdd*. Two sets of primers, Ph_PiscA and Ph_PiscB (Figure 3 and Table 6) were designed using Primer3web (4.0.0) targeting the region containing the two SNPs.

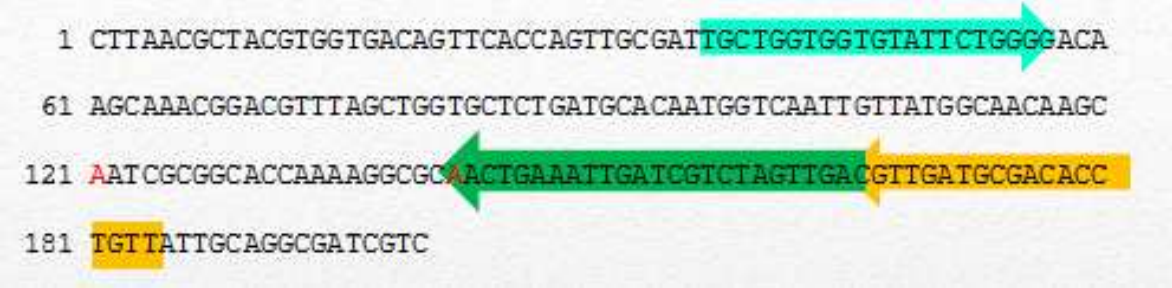


Figure 3. Scheme of the primer design. SNPs with red color are those discriminating between *Phdd-Phdp* subspecies in the context of the portion of the *bamB* gene considered. Forward primer is highlighted with light blue color and is common for Ph_Pisc.A and Ph_Pisc.B; Reverse primer for Ph_Pisc.A is highlighted with green color and drops onto one of the two discriminating SNPs. Reverse primer for Ph_Pisc.B is evidenced with yellow color.

Table 6. Primer sets for *Phdp* amplification employed in the present study.

		Start	Length	Ta	Amplicon length	Primer sequence (5' - 3')
Ph_Pisc.A	For	37	20	59.67	130	TGCTGGTGGTGTATTCTGGG
	Rev	166	25	57.65		GTCAACTAGACGATCAATTTCAGTT
Ph_Pisc.B	For	37	20	59.67	148	TGCTGGTGGTGTATTCTGGG
	Rev	184	20	60.53		AACAGGTGTCGCATCAACGT

Start: primer start position (5') in relation to the alignment in Supplementary file 1
Length: primer length (nt)
Ta: primer annealing temperature

A primer set for the amplification of *Sparus aurata* TRL-9 gene (Sa_TLR9_Forward 5'-GAAATTGTCTGCCAGGTCGC-3' and Sa_TRL9_Reverse 5'-AGGTATACGGGTGAGGCTGT-3') was used as internal control in PCR reactions. Real-time PCR amplification was performed using the Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen, California) in 10-µl reaction volumes containing 1X Platinum SYBR Green qPCR SuperMix-UDG, 10 uM each primer and 2.5 µl of template DNA. The amplification reaction was conducted using the LightCycler 480 System instrument (Roche, Basilea). The cycling conditions were as follows: initial incubation at 50°C for 2 min, followed by 2 min at 95°C, and 45 cycles at 95°C for 10 s and 60°C for 1 min. A melting curve between 40 °C and 95 °C was determined by adding a dissociation step after the last amplification cycle at a temperature transition rate of 4.4 °C/s. For each sample, the Cp (Crossing point) was used to determine the amount of target gene; each measurement was made in duplicate.

Two different standard curves were constructed to evaluate the efficiency and sensitivity of the Ph_PiscA and Ph_PiscB primer sets. The first one was obtained by using 5-fold serial dilutions of purified genomic DNA from the *Phdp* 249/ITT/99 strain. Genomic DNA was diluted in molecular-grade water in order to obtain concentrations ranging from 5 ng/µl (25.5 * 10^5 copies) to 0.00000256 ng/µl (1 copy). In order to test both primer sets in a scenario as close as possible to field conditions, a second standard curve was constructed in spiked conditions. Briefly, 2-fold serial dilutions of *Phdp* 249/ITT/99 DNA ranging from 0.038 ng/µL (19.4 * 10^3 copies) to 0.000018 ng/µl (9 copies), mixed with a fixed concentration of purified genomic DNA from *S. aurata* (100 ng/µl). For each dilution, the bacterial copy number was estimated by dividing DNA quantity (ng) with the molecular weight of the *Phdp* genome (5.2 fg), as described by Pathak et al. 2012. Standard curves were generated by plotting the Cp vs the number of bacterial copies of serially diluted samples. Ph_PiscB primer set was also tested for specificity on a collection of target (19 *Phdp* strains and 11 *Phdd* ones) and non-target bacterial species (17 strains belonging to the genus *Vibrio*, *Photobacterium*, *Yersinia*, *Aeromonas* and

Tenacibaculum, see Table 7). In order to test the repeatability of the assay, three technical replicates for each target and non-target bacteria were performed. Specificity of real-time PCR products was determined by the analysis of amplification profiles and melting curves. Mann-Whitney test was applied to assess the statistical significance of melting temperatures differences between *Phdp* and *Phdd* strains.

Table 7. Species/strains used in the assay, name and origin.

	Photobacterium damselae subsp. piscicida	**Origin**
1	R 198 DSMZ 22834	Not reported
2	4/19/ITT	Italy
3	TFP88012	Japan
4	ATCC17911	USA
5	557/ITT	Italy
6	331/ITT	Italy
7	332/ITT	Italy
8	B3	Greece
9	I736	Italy
10	4/18/ITT	Italy
11	B7	Greece
12	SU7	Greece
13	M28561	Japan
14	NCIMB2058	Japan
15	Sa071194	Israel
16	XXX0696	Israel
17	SP98037	Japan
18	K12	Italy
19	249/ITT/99	Italy
	Photobacterium damselae subsp. damselae	
1	NCIMB2181	USA
2	NCIMB2182	USA
3	NCIMB2183	USA
4	ATCC33539	USA
5	NTCT11648	UK
6	JCM8969	Japan
7	ATCC35083	USA
8	ATCC97299	USA
9	NCIMB13351	Japan
10	JCM8967	Japan

11	NTCT11646	UK
	Other bacterial species	
1	*Aeromonas hydrophila* R 71 CECT 398	USA
2	*Vibrio parahaemolyticus* ATCC 17802	Japan
3	*Vibrio harveyi* ATCC 14126	USA
4	*Vibrio harveyi* LM 64044	Not reported
5	*Vibrio furnissii* NCTC 11218	Not reported
6	*Vibrio furnissii* R 193 CECT 4203	Japan
7	*Vibrio alginolyticus* ATCC 17749	Japan
8	*Yersinia ruckeri* A	Not reported
9	*Yersinia ruckeri* B	Not reported
10	*Photobacterium leiognati* CCUG16229	Malaysia
11	*Tenacibaculum maritimum* R 206 CECT 4276	Japan
12	*Vibrio anguillarum* 5894/83	UK
13	*Vibrio anguillarum* R 12 LMG 10861	Denmark
14	*Vibrio natriegens* ATCC 14048	USA
15	*Vibrio carcharie* CCUG 19116	USA
16	*Vibrio campbelli* CECT 523	Not reported
17	*Photobacterium fischeri* ATCC 7744	USA

3.6 Real time PCR assay on experimentally infected samples

Samples from DD and SV fish (see paragraph 3.1) were used to validate the real-time PCR assay. For each sample, bacterial copy number was determined by interpolation of the resulting Cp values to the standard curve constructed with *S. aurata* DNA spiked with *Phdp* DNA.

Data, expressed in number of copies, were statistically analyzed using Spearman's rank-order correlation test and Wilcoxon signed-rank test to assess the correlation of estimated bacterial loads between different tissues. Differences were considered statistically significant when $p < 0.01$.

3.7 SNPs discovery and genotyping

Demultiplexed reads returned by the sequencing facility were checked for quality and trimmed as described in paragraph 3.3. Reads were filtered out if average quality within a sliding window of 4-bp was less than 15. As reference sequence for this species was not available, therefore reference was developed de novo by clustering of reads. Perl based custom developed scripts in combination with CD-HIT program were used to develop the

reference sequence (Wang et al., 2012; Fu et al., 2012)] which was further used to call genotypes by aligning individual specific reads to the reference sequence. Alignments of short reads to the reference sequence were performed using bwa samse (V = 0.7.13-r1126) (Li et al., 2009). Data analysis for SNPs discovery was performed using the GATK package (McKenna et al., 2011) with the following criteria: a minimum genotype quality of 20, a minimum read depth of 5, and a population wise observed minor allele counts for a particular site that must be at-least 50. A minimum of 40 individuals in a population needed to have a genotype call that met these criteria at a specific position. A SNP that passed the above mentioned criteria was considered as a putative SNP for further analyses. Two catalogues (a catalog for each experiment) of unique tags were developed from the sequence data of all the parents and used as a reference sequence; the individual specific reads were aligned to this reference sequence to detect and call SNP based genotypes. The mpileup function of SAMtools version was used to call variants and the call option of bcftools (Li et al., 2009) was used to call the genotype at each variant site for each animal.

3.8 Genetic linkage analysis

Genotypic data was further used to develop two linkage maps for seabream (one per each challenged population). Quality control (QC) was performed at individual and marker levels. Individuals were filtered out based on two criterions: with less than 30% missing rate and family size of ≥ 5, offspring were kept to construct the linkage map. QC at marker level was performed for each full sibling family by excluding SNPs with minor allele frequency (MAF) ≤ 0.05 and those deviating from expected Mendelian segregation (P ≤ 0.001). Linkage groups (LG) were built using minimum LOD threshold value of 46 in "SeparateChromosomes" module of Lep-Map v2 (Rastas et al., 2015) by allowing maximum distance of 20 cM between consecutive SNPs. "JoinSingles" module of Lep-Map was used to join singular markers to the already defined linkage groups applying LOD score limit of 5 in combination with LOD score difference of 2 between the best LG and the second best LG of each joined marker. The module "OrderMarkers" was then used to estimate the order and distance between the markers in centiMorgans (cM). "OrderMarkers" implements hidden Markov model to compute likelihood for the order of markers (Rastas et al., 2013).

3.9 Genetic parameters

Genetic parameters of the traits were estimated using ASReml 4.0 (Gilmour et al., 2015) implementing genomic or pedigree-based relationship matrices (G-matrix and A-matrix, respectively). The genomic relationship matrix was constructed using VanRaden method (VanRaden et al., 2008) while pedigree relationship matrix was computed using efficient algorithm of Meuwissen and Luo method (1992). The pedigree used for the construction of the A-matrix was developed by assigning parents to offspring using the SNPs based molecular markers. SNPs were filtered by applying criteria of minor allele frequency (MAF) ≥ 0.35 and genotyping rate of ≥ 90 %. The remaining selected set of highly informative SNPs were then used to construct pedigree applying likelihood ratio method implemented in CERVUS version 3.0 (Kalinowski et al., 2007).

For the binary trait dead/survived and day to death (*phdp* experiment), the following mixed model was applied:

$$y = \mu + Zu + e$$

where y is a vector of 'n' phenotypic records, μ is an overall mean, u is a vector of additive genetic effects distributed as $u \sim N(0, G\sigma^2 u)$, or $u \sim N(0, A\sigma^2 u)$, were $\sigma^2 u$ is the additive genetic variance, G and A are genomic and pedigree relationships matrices, respectively, Z is the corresponding incidence matrix to additive effects aand e is the vector of random residual effects with $e \sim N(0, I\sigma^2 e)$. Regarding the traits ParCount, SGR and BW2 (*Sc* experiment), the same model was used but a fixed effect for initial body weight was also included, therefore the model changed as follow:

$$y_1, y_2 = \mu + Xb + Zu + e$$

where *X* is the corresponding incidence matrix for fixed effect (BW1) and *b* is the vector of fixed effect.

3.10 GWAS

To test the association between individual SNPs and resistance to photobacteriosis or to *Sc*, a GWAS was performed by applying the same model of genetic parameters (see paragraph 3.9) with SNP effects also included as additional variable. Thus, the two models were:

$$y = \mu + Ma + Zu + e$$

$$y_1, y_2 = \mu + Xb + Ma + Zu + e$$

where *M* and *a* are the corresponding matrix and the vector for the marker effect, respectively.

The GTCA program (Yang et al., 2011) with –mlma-loco function was used to detect marker associations with traits. SNPs were considered genome wide significant when they exceed the Bonferroni threshold for multiple testing (alpha 0.05) of 0.05/*tg* where *tg* is the total number of genome-wide SNPs.

3.11 EBV and accuracy of prediction

The same models described under the GWAS were applied for the estimation of breeding values and the predictions were performed using PBLUP and GBLUP using the R/BGLR (Pérez et al., 2014) program. A BayesB estimation (Habier et al., 2011) was also included for comparison in the context of the *Phdp* population. Accuracy of prediction was computed using cross validation scheme by masking the phenotypes of ~ 30% (1) and ~ 10% (2) of the offspring. The population was split in training and validation groups, while the random sampling was repeated 10 times. The accuracy was computed as the correlation of the estimated breeding value (pedigree/genomic, PEBV/GEBV) with the pre-corrected phenotype scaled by the square root of heritability, or/and as the prediction error variance method.

4. Results

4.1 Challenge experiments

(1) Before the challenge, no evidence of *Phdp* was observed in all fish tested. The challenged population consisted of 1233 individuals, originating from 109 broodfish. Mortalities in challenged fish started at day 3 and occurred during 11 days after the injection with peaks on days 4, 5 and 6 (Table 8). The cumulative mortality at the end of the challenge was 36,4 % (449 specimens; survival 61,5%, 784 specimens). In many fish, external lesions at the injection site were observed. No mortality was recorded in control fish over the entire experimental period.

Table 8. Mortality of fish measured during the experiment.

Day	0	1	2	3	4	**5**	**6**	**7**	8	9	10	11
n° fish	0,0	0,0	0,0	5,0	118,0	**182,0**	**88,0**	**26,0**	11,0	9,00	8,00	2,00
Mortality per day (%)	0,0	0,0	0,0	0,41	9,61	**16,4**	**9,48**	**3,1**	1,35	1,12	1,01	0,25
Cumulative mortality (%)	0,0	0,0	0,0	0,41	9,98	**24,74**	**31,87**	**33,9**	34,87	35,6	36,3	36,42
Tot. of dead fish	0,0	0,0	0,0	5,0	123,0	**305,0**	**393,0**	**419,0**	430,0	439,0	447,0	449,0

(2) The challenged population consisted of 1001 individuals, originating from 119 broodfish. Sampled fish were found to be heavily infected with all parasitic stages (100% prevalence, 4-15 adults in the left external gill arch), as early as two weeks following the transfer of the donor fish. Low mortalities (4-10 fish / day) along with a reduction of fish appetite were initiated on week 4; few fish appeared lethargic and often anaemic gills were observed in dead individuals. At the end of the challenge the cumulative mortality was 15% (150 specimens; survival 85%, 851 specimens), while no mortality was observed in the control group. The average weight of the recipient fish at the final sampling was 43 ± 11 g while that of the control fish reached 58 ± 12 g. 7 % of the

examined population was pathogen free and the mean intensity was calculated to be 6.4 parasites / fish at the final sampling (Figure 4).

Percentage (%)	Number of parasites	Mean intensity**
7	0	6.4
75	1-10	
15	10-20	
3	>20	

** Total count of adults parasites in the two external gill arches*

*** Mean intensity was expressed as the mean number of parasites found in the infected hosts excluding zeros of uninfected hosts*

Figure 4. Percentage of infection in relation to *Sc* intensity* in survived (SV) infected seabreams (n = 851).

4.2 Real-time PCR assay/ sensitivity

The efficiency of the two candidate primer pairs (Ph_PiscA and Ph_PiscB, see methods) was assessed using either pure bacterial DNA or bacterial DNA spiked with fish genomic DNA. In the first case, efficiencies calculated from standard curve were 95% and 94% for Ph_PiscA and Ph_PiscB, respectively, with a limit of detection (LOD) of 1 copy of bacterial DNA for both assays. In the second case, 2-fold serial dilutions of 249/ITT/99 were mixed with purified seabream genomic DNA and the calculated efficiencies of the primer pairs were 97% for Ph_PiscA and 99% for Ph_PiscB. Ph_PiscB showed higher sensitivity with a theoretical limit of quantification (LOQ), under "spiked" conditions, of 9 copies in comparison to Ph_PiscA, whose LOQ was 18.9 copies. Based on the obtained results, the Ph_PiscB primer pair was used in all subsequent analyses.

4.3 Real-time PCR assay/ specificity

Ph_PiscB primer set was tested on equal amounts of all bacterial species/strains listed in Table 7. The qPCR assay amplified both *Phdp* and *Phdd* strains with high efficiency, although the latter was detected 2-3 cycles later. All ”non target” species/strains (see Table 7) were either not amplified or amplified with very low efficiency, being barely detectable after 35 cycles of amplification (~15 cycles later than *Phdp* strains). Melting curve analysis was then applied to discriminate between *Phdp* and *Phdd* strains. All *Phdp* strains were characterized by a melting temperature (Tm) of 83.3-84 °C (83.8 ± 0.15) while all *Phdd* strains showed a Tm of 84.3-84.9 °C (84.5 ± 0.15) thus making the two subspecies easily distinguishable. In addition, all the technical replicates showed a good concordance (mean coefficient of variation CV of 0.08 %), proving the reliability of the assay. A Mann-Whitney test revealed a very high statistical significance between *Phdp-Phdd* melting temperatures, with a p-value < 2*10-16. The melting curve analysis also allowed further discrimination of the majority of non-target bacteria.

4.4 Diagnostic effectiveness of qPCR assay

Spleen, liver, and gills collected from a total of 15 fish dead during the mortality peak (days 4 to 6 post-challenge, 5 fish/day) were tested with Ph_PiscB primer set. Positivity to *Phdp* in the real-time PCR assay was registered in all samples, where the number of *Phdp* DNA copies detected ranged from 41.5 to 348,500 (Figure 5A).

Quantitative PCR revealed a great variation in bacterial load across individuals and tissues, although spleen samples generally showed a higher load with a mean value of 42,935 copies, compared to 1,184 and 1,282 copies for liver and gills, respectively. Boxplots of Log_2-transformed copy numbers clearly show the different distribution of *Phdp* in the three tissues tested (Figure 5C). Among tissues from the same individual, a Wilcoxon signed-rank test showed that the spleen always had a significantly greater *Phdp* load ($p < 0.001$) compared to gills and liver. No statistical significance was found for the difference in bacterial distribution between liver and gills. In order to measure the correlation of bacterial load across different tissues from the same dead animal, a Spearman’s rank correlation test was carried out on estimated copy numbers of tissue pairs (spleen-liver; spleen-gills; liver-gills). Spearman’s test showed extremely strong positive correlation, with rho coefficients greater than 0.8 (p-value<0.001) for all pairs tested (Figure 5B), reflecting a linear relationship on bacterial load between tissues of the same infected fish.

The same qPCR assay was employed to evaluate the residual presence of *Phdp* in fish that survived the experimental infection. In this case, a larger number of individuals (N= 130) was tested, and only spleen samples were analyzed. The vast majority of surviving fish were negative (128 specimens, 98.4 %) for *Phdp*, while only two specimens (1.6 %) still showed presence of the pathogen DNA.

A.

Day of death	Sample ID	Spleen	Liver	Gills
4	75086	3,580	343.5	759.5
	75088	1,270	78.35	246
	75089	58,050	2200	2110
	75092	255	82.35	41.5
	75105	119,000	1300	4100
5	75150	44,600	1270	592.5
	75177	23,000	1845	1115
	75211	348,500	9200	7930
	75231	23,950	364.5	399.5
	75265	8,550	234.5	657
6	75331	3,710	427.5	618.5
	75363	2,455	212.5	348.5
	75364	595	54.3	135.5
	75383	1,985	113.5	100.3
	75388	4,530	39.65	86.65

B.

	Spleen	Liver	Gills
Spleen	1		
Liver	0.907**	1	
Gills	0.828**	0.900**	1

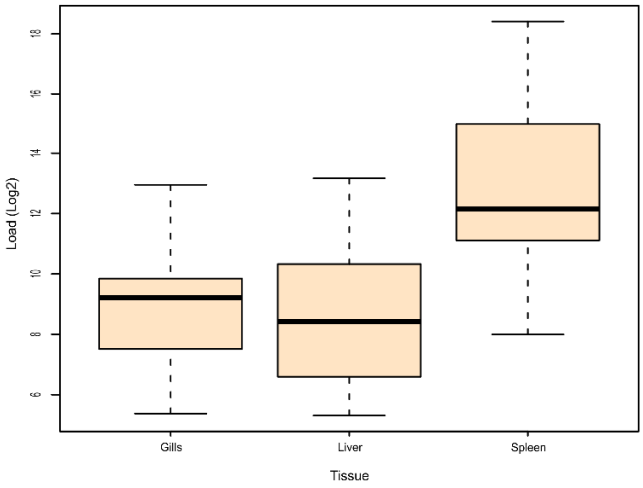


Figure 5. A *Phdp* bacterial copy numbers of spleen, liver and gills and concerning Sample IDs. B Spearman's rank-order correlation coefficients between tissue sample pairs across individuals **p-value <0.001 C Box-plots of Log2-transformed copy numbers in the three tissues

4.5 2b-RAD libraries

A total of 3072 (32 96-well plates) 2b-RAD libraries were generated. Quality and quantity of the libraries were generally good with a mean of 5 samples/plate (5/96, 5.2 %) that failed and were repeated. A total of 8 samples (8/2462, 0.3%) failed the second or the third attempt. A descriptive statistic about 2b-RAD libraries is summarized in Table 9.

Table 9. Means of the concentration yield values (ng/µl) per each plate containing 96 2b-RAD libraries with the median, the standard deviation and the number of failed libraries*.

Plate ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Mean (ng/µl)	42,2	56,8	36,8	36,6	18,7	34,3	60,9	19,1	36,8	85,8	63,3	82,7	49,6	46,9	42,0	39,1
Median (ng/µl)	45,1	53,2	38,2	32,1	16,9	33,2	58,6	16,1	38,5	88,5	58,4	83,3	49,8	47,0	42,3	39,7
Standard dev.	15,9	27,0	10,8	15,7	8,5	14,1	21,7	13,8	13,9	25,8	30,6	33,1	11,2	9,5	11,7	14,8
N° failed samples	0	0	2	0	7	2	0	7	6	1	2	1	0	0	0	5
Plate ID	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**
Mean (ng/µl)	32,3	37,6	33,8	25,8	11,0	33,9	23,4	38,3	17,2	17,2	34,5	30,1	14,1	30,7	26,5	72,2
Median (ng/µl)	30,7	36,1	33,3	25,6	10,1	35,3	21,9	38,3	19,2	17,9	33,7	29,5	13,2	32,6	29,3	74,2
Standard dev.	11,9	17,0	14,7	15,6	6,5	9,1	10,3	14,8	7,9	4,4	9,8	10,1	9,1	12,0	13,8	14,0
N° failed libraries	2	3	3	12	20	0	0	2	12	2	2	2	9	7	2	0

*Libraries were considered as failed with a DNA concentration < 8 ng/µl

4.6 Multiplexing strategy

In order to investigate the best pooling strategy to be applied to the ~3000 libraries, 24 samples by both experiments were tested in a HiSeq2500 machine including 7 replicates. Samples of pool_2 and samples of pool_1 were pooled together by using the proportion 1 : 2.4. The number of trimmed reads ranged between 3.000.000 and 9.000.000 for individuals in pool_2, while it ranged between 14.000.000 and 30.000.000 for individuals in pool_1. Data were in accordance with the pooling strategy, with an increase in terms of raw reads which reached the 71% for the individuals of pool_1 in comparison to pool_2 (Figure 6).

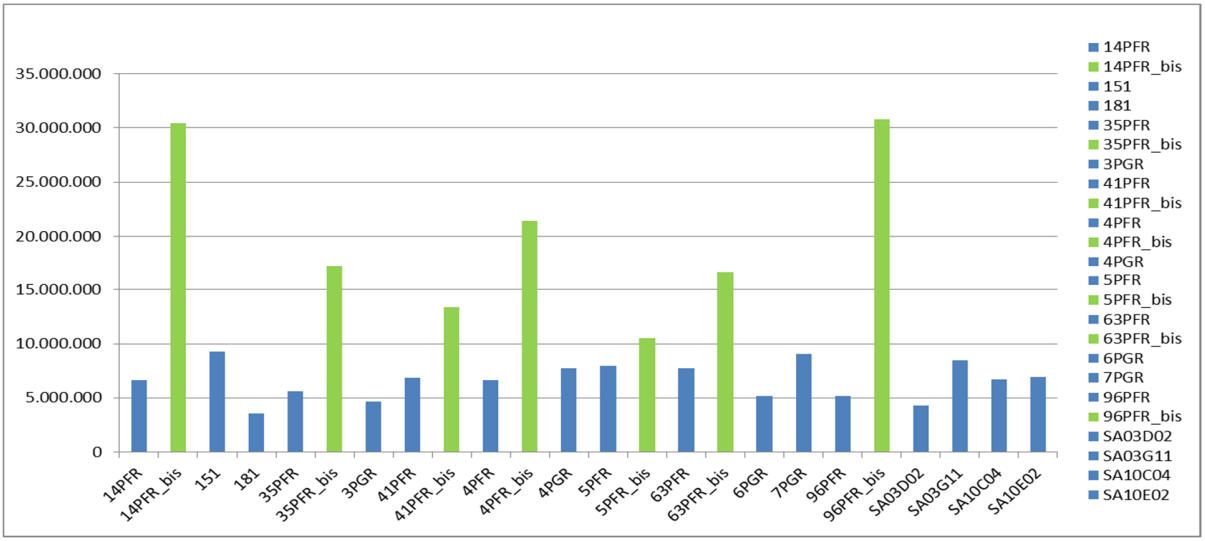


Figure 6. Number of demultiplexed trimmed reads per each sample. Individuals in pool_1 are represented by the green bar while individuals in pool_2 are represented by the blue bar.

However, no significant differences in terms of number of quality SNPs were discovered between the two pools, which varies between +1% and +8% in pool_1 in comparison to pool_2 (Figure 7).

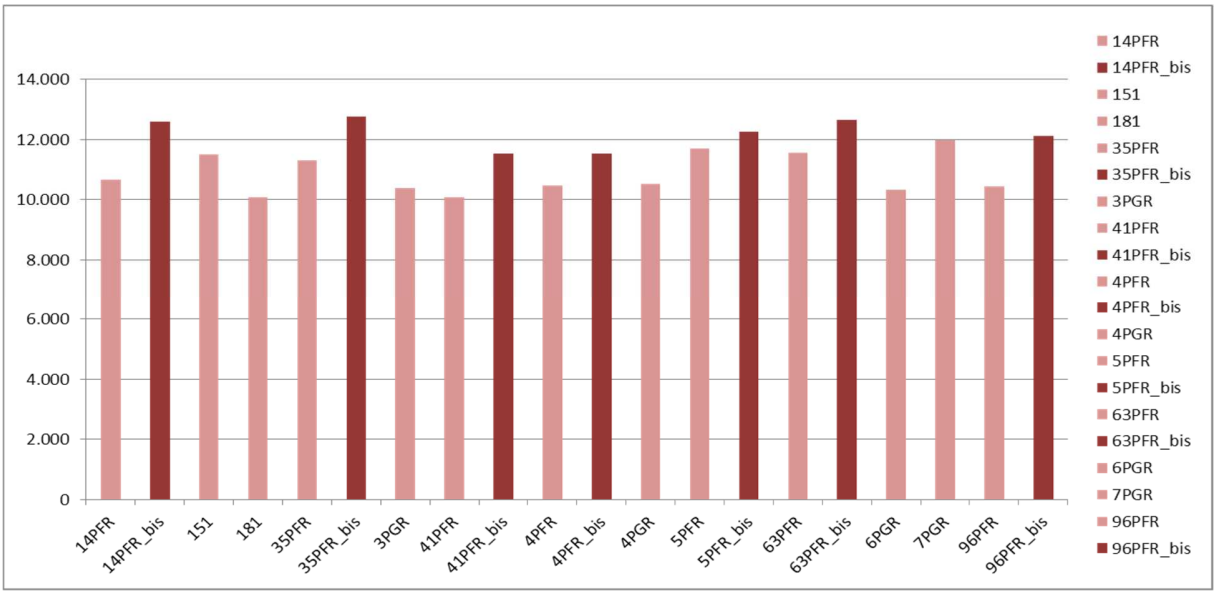


Figure 7. Number of SNPs discovered after filtering among individuals of pool_1 (dark bar) and pool_2 (light bar).

4.7 Sequencing results

24 demultiplexed lanes referring to both experiments returned from BMR sequencing. A summary of the sequencing results is the reported in Table 10.

Table 10. Sequencing output in terms of million of raw reads for the 24 pools sequenced on the NexSeq500 platform.

Pool/Lane ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
N. raw reads (million)	290*	393	452	454	501	446	408	482
Pool/Lane ID	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
N. raw reads (million)	455	498	493	411	430	412	415	321
Pool/Lane ID	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
N. raw reads (million)	250	414	456	490	313	381	400	430

*repeated twice

Lanes 1,2, 16, 17, 21 and 22 yielded less then as expected (400M) with an output of 290M, 393M, 321M, 250M, 313M and 381M respectively. The first run was repeated leading to a total output of 500M. For approximately 300 individuals, a second sequencing-step was performed in order to reach the established threshold for coverage. The overall number of raw reads was ~ 1.5 billion for parents and ~3 billion for offspring for both trials. Demultiplexed reads were checked by FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/) which provided quality summary graphs and tables. Similar qualities were obtained among different pools/different samples. The sequence_lenght_distribution picture showed that the majority of the reads had a length of 36 pb demonstrating that the Illumina adaptor was recognized by the NeqSeq machine and trimmed. A minority of the reads did not have the adaptor and kept a length of 51pb. The per_base_sequence_quality was generally good showing a Phred>30 in all positions except for the first nucleotide following *AlfI* recognition site and at the end of the reads. The percent of the reads lost during filtering steps with a customized script (Pecoraro et al., 2016; Babbucci et al., 2016; Pauletto et al., 2016) was about 0.17% (75 million).

4.8 SNPs detection/genotyping

The reference catalogues developed using a total of 3 billion reads from parents contained 230,500 and 269,660 2b-RAD tags for the first and the second challenge, respectively. The SNPs detection process yielded genotypic data for 19,313 and 21,773 quality SNPs (with Minor Allele Frequency, MAF ≥ 0.05 and locus specific genotyping rate ≥30). Individuals were also removed if genotyping rate of an individual was < 30%, and the level of heterozygosity > 65%. Out of total 1233 (1) and 1001 (2) offspring individuals, 46 individuals and 81 individuals respectively did not meet the criterion, and hence were filtered out. Indeed, all the parents individuals met the criterion and were used for the analysis. After the above described filtration steps, SNPs based genotyping data consisted of 19,313 and 21,773 loci typed on 1187 offspring x 109 parents for the *Phdp* challenge and 920 offspring x 119 parents for the *Sc* challenge.

4.9 Pedigree construction and linkage maps

The filtering SNPs process for pedigree construction produced 750 (1) and 916 (2) selected highly informative SNP markers, which produced 177 full-sib families with 1-49 sibs per family (1) and 154 full-sib families with 1-22 individuals per family (2). Out of total 177 and 154 families, 74 and 17 families respectively, had a minimum of 5 sibs and were used for the construction of linkage map. SNPs were grouped into 24 linkage groups (Sa LGs),

which are consistent with the karyotype of this species. Since each experiment analysis was carried on separately, LGs numbering is different between the two linkage maps. Specific details of the two seabream linkage maps are reported in Supplementary file 2. The two total linkage maps lengths were 1947 cM (1) and 1406 cM (2). There were 5785 and 6589 SNP singleton markers for the *Phdp* and *Sc* maps, respectively, which did not get assignments to any group.

4.10 Genetic parameters estimation

(1) Genomic heritability for resistance to photobacteriosis measured as *dead/alive* phenotype was 0.33 (0.052), explaining the 33% of the total variance, while it was 0.31 (0.052) if measured as *day of survival* (days to death), explaining the 31% of the total variance, with genetic correlation ~1, indicating that the traits are very similar. Both traits were calculated including the genomic information (see methods). When only pedigree information were used, h^2 was 0.30 (0.074) for *dead/alive* and 0.32 (0.076) for *days to death* traits. Figure 8 summarizes all values for the traits.

Trait	Pedigree	Genomic
P_{DS} (h^2)	0.308 (0.074)	0.332 (0.052)
P_{D2D} (h^2)	0.320 (0.076)	0.317 (0.052)

Figure 8. Estimates of h^2 on dead/survive and days to death phenotype using pedigree vs. genomic information. P_{DS} = dead/survive phenotype, binary trait; P_{D2D}= day to death phenotype; h^2=heritability.

(2) Genomic heritability for tolerance to *Sc* was measured on SV fish considering 3 different traits, the *parasite count* (ParCount), the *weight of the fish* at the end of the challenge (BW2) and the *specific growth rate* (SGR). SGR is calculated based on the fish before the start and after the end of the challenge (see methods). A descriptive statistic for recorded phenotypes (mean, minimum value, maximum value and standard deviation calculated on a total of 807 passed-filtering survivors) is reported in Table 11.

Table 11. **Descriptive statistics for recorded phenotypes .**

Traits	Mean	Min	Max	SD
ParCount	5,75	0	45	5,88
BW1	17,19	2	60	8,96
BW2	43,93	11,6	87,8	11,65
SGR	0,016	-0,0012	0,033	0,0042

BW1= weight of fish before the start of challenge
BW2= weight of fish at the end of challenge
SGR = specific growth rate
SGR = (ln BW2 – ln BW1)/68

Heritability was 0.129 (0.068), 0.135 (0.066) and 0.141 (0.066) for ParCount, SGR and BW2 respectively, using the pedigree information, explaining nearly the 10 % of the total variance each. When genomic information were used, h^2 was 0.119 (0.062), 0.223 (0.081) and 0.192 (0.075) for the same traits (Figure 9).

Components	Pedigree	Genomic
Parasite Count (h^2)	0.129 (0.068)	0.119 (0.062)
SGR (h^2)	0.135 (0.066)	0.223 (0.081)
BW2 (h^2)	0.141 (0.066)	0.192 (0.075)

Figure 9. Heritability of the traits related to resistance to *Sc* with the implementation of the A- and the G-matrix.

A strong negative genetic correlation was found between ParCount and SGR/BW2, which values are shown in Table 12.

Table 12. Genetic correlation values between *Sc* traits using pedigree or genomic information.

Trait	Pedigree	Genomic
ParCount vs SGR	-0.807±0.24	-0.769±0.22
ParCount vs BW2	-0.549±0.27	-0.701±0.21

4.11 GWAS

(1) GWAS revealed a QTL including 7 SNPs at LG17 presenting significant association with the resistance to *Phdp*, with p-value crossing genome-wide Bonferroni corrected threshold (P-value 2.218e-06, Figure 10).

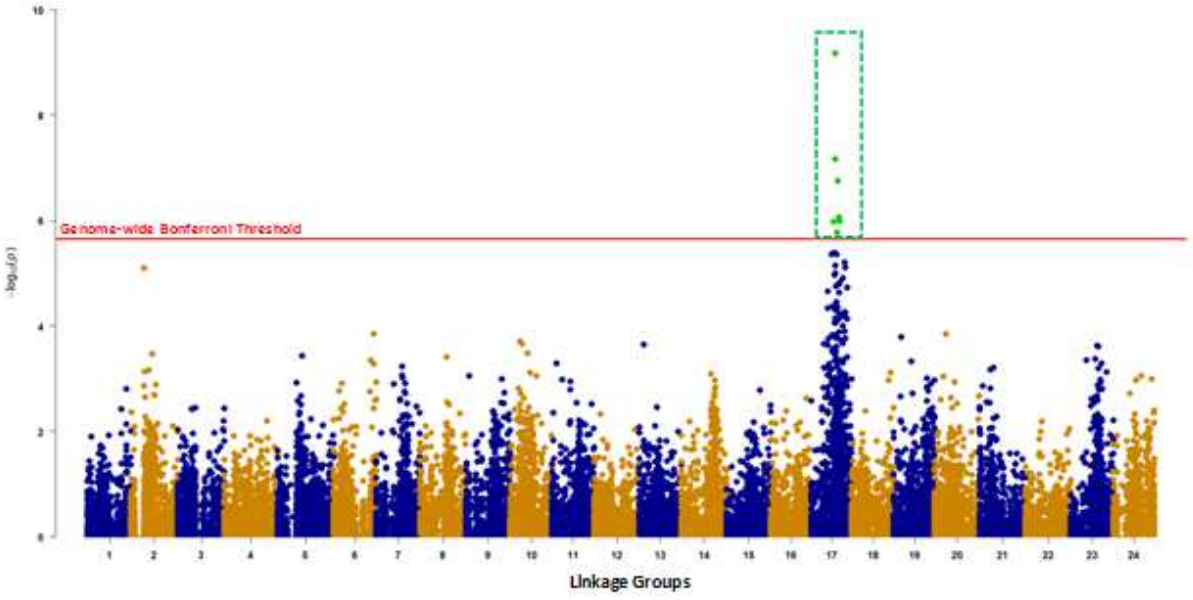


Figure 10. **Manhattan plot showing the distribution of p-values across linkage groups 1-24. Highlighted green dots represent genome-wide significant SNPs and horizontal solid line represents the Bonferroni significance threshold.**

Detected QTL region (including 7 genome-wide significant SNPs) of SA17 spans from 40.5 to 53.1 cM with peak SNP positioned at 44.728 cM.

(2) One SNP located in LG17 surpassed the Bonferroni-corrected genome-wide significance threshold (p value 4.02e-05) for the tolerance to *Sc*. In addition, one QTL in LG19 (one SNP) and two SNPs located in Unknown-LG were found as putative (suggestive) for the tolerance to the parasite (Figure 11).

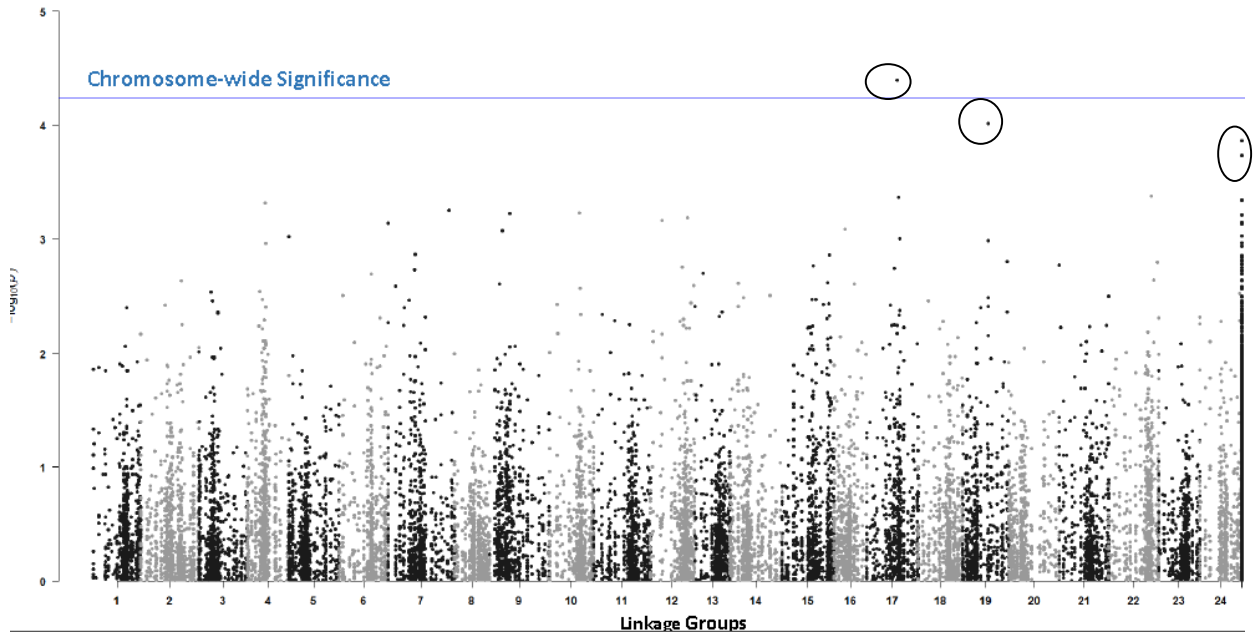


Figure 11. Manhattan plot showing the distribution of p-values across linkage groups 1-24. Circled dots represent genome-wide significant/putative SNPs and horizontal solid line represents the Bonferroni significance threshold.

4.12 EBV & genomic prediction

Genomic prediction was tested as a means of obtaining breeding values, and compared to prediction using a pedigree-based approach and a BayesB method.

(1) The prediction was conducted using genotype information from 19,313 markers (passed QC filters set for genomic prediction) for the 1187 disease challenged animals. Figure 12 shows the selection candidates within each family based on EBV and GEBV for the phenotype dead/survive.

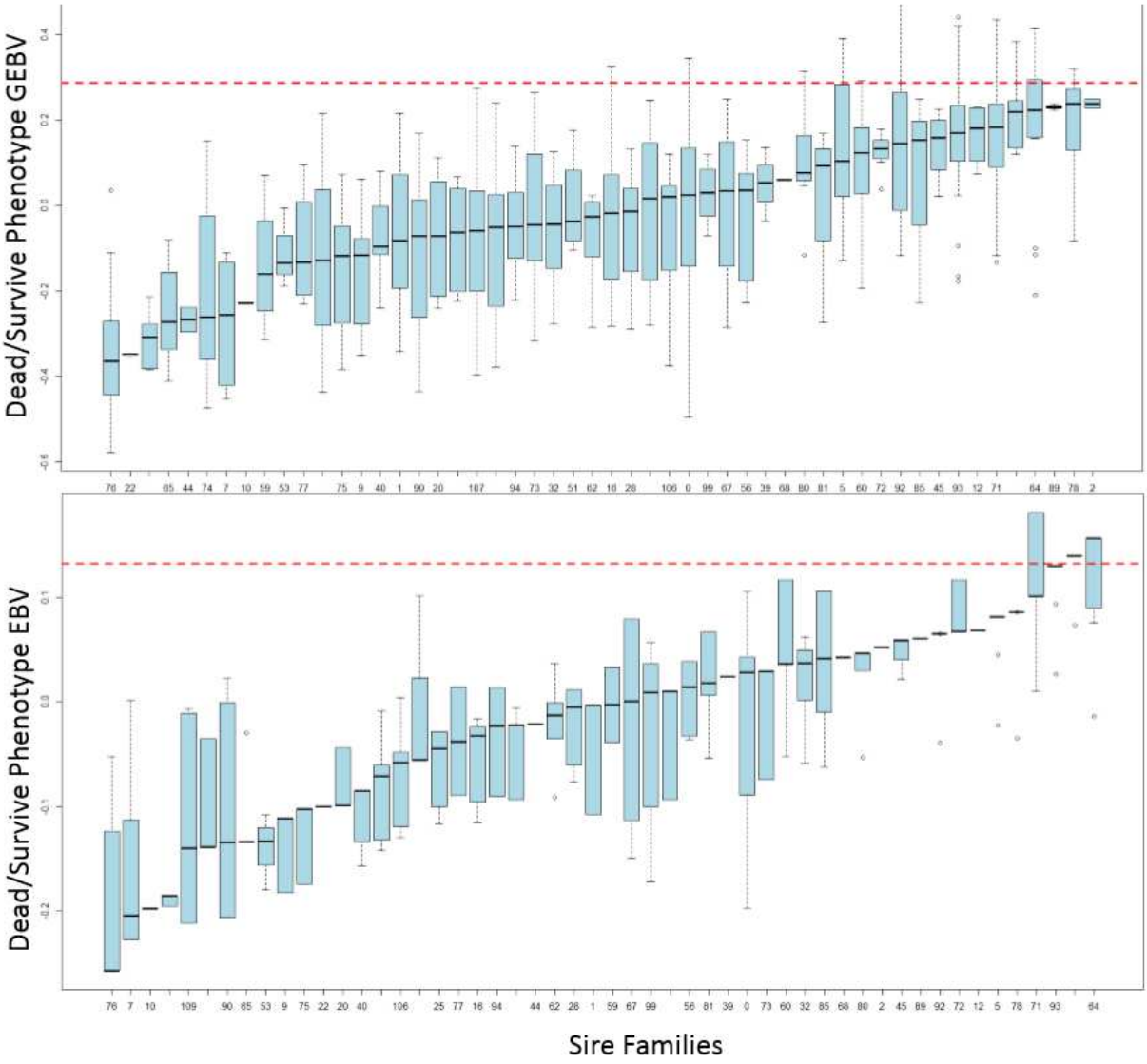


Figure 12. Selection candidates within each family based on EBV/GEBV. Bars represent each selection candidate, dark line within each bar is the median.

Prediction accuracy was measured as correlation between breeding value (PEBV/GEBV) and pre-corrected phenotype ($rcorr$) by using the following formula (I):

$$r_{corr} = \frac{\rho(G[P]EBV, y_{adj})}{sqrt(h^2)}$$

Animals were randomly split in training (n = 709) and validation (n = 320) datasets. The random sampling was repeated 10 times. As summarized in Figure 13, prediction accuracy with PBLUP was 0.39 ± 0.087 and 0.56 ± 0.075 for the genomic prediction models for the trait *days to death*. Similar results were found for the binary trait *dead/survive*. The BayesB method resulted in accuracy of 0.55±0.067.

Accuracy	PBLUP	GBLUP	BayesB
r_{corr}	0.39± 0.087	**0.56± 0.075**	0.55 ± 0.067

Figure 13. **Prediction accuracy calculated with PBLUP, GBLUP and BayesB methods.**

(2) The prediction was conducted using genotype information from 21,773 markers (passed QC filters set for genomic prediction) for the 841 disease challenged animals. Figure 14 shows the selection candidates within each family based on EBV and GEBV for the phenotype parasite count.

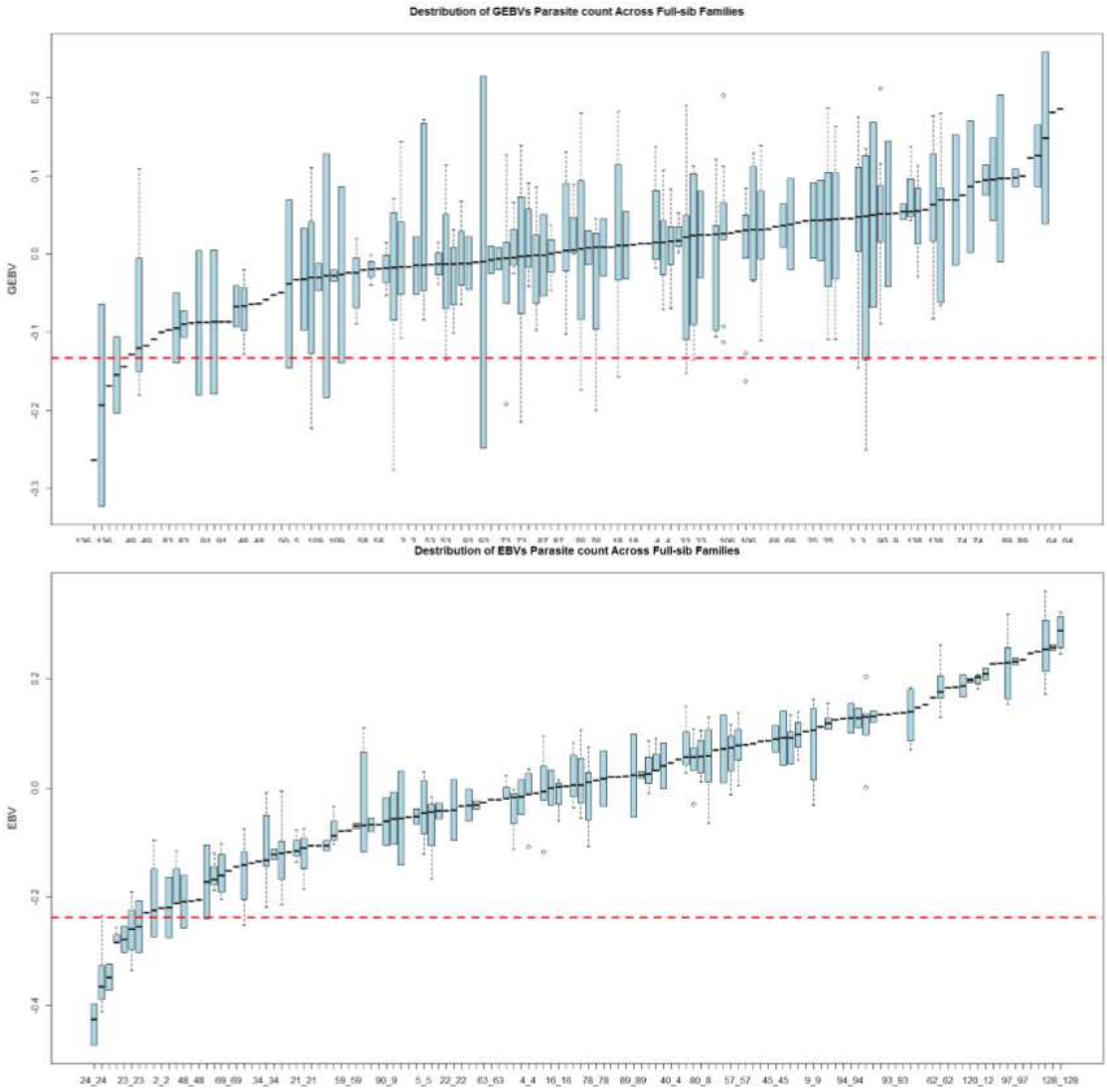


Figure 14. Selection candidates within each family based on EBV/GEBV. Bars represent each selection candidate, dark line within each bar is the median.

The prediction accuracy was measured in two ways. The first one was the same as (I) $rcorr$, that is the correlation between estimated breeding value (PEBV/GEBV) and pre-corrected phenotype, while the second one was the prediction error variance method $rpev$ (II).

$$r_{pev} = \sqrt{1 - \frac{SEP^2}{\sigma_g^2}}$$

Animals were randomly split in training (n = 412) and validation (n = 138) datasets. The random sampling was repeated 10 times. As reported in Figure 15, the accuracies of prediction were 0.52 ± 0.051 (pedigree information) and 0.64 ± 0.039 (genomic information) calculated as $rpev$, while $rcorr$ of 0.51 ± 0.147 and 0.55 ± 0.165 were found as accuracies for the pedigree and the genomic matrix, respectively.

Accuracy	Pedigree	Genomic
$r_{pev-LPC}$	0.52 ± 0.051	**0.64 ± 0.039**
$r_{corr-LPC}$	0.51± 0.147	**0.55± 0.165**

Figure 15. Prediction accuracy with PBLUP & GBLUP methods measured in two different ways.

5 Discussion

5.1 Challenge experiments

In the present work, the genetic bases of the resistance/tolerance against two of the most important pathogens for seabream cultivation were investigated. The study aimed at exploring genetic variation contributing resistance against fish photobacteriosis and *Sc*, detecting QTL for resistance as well as determine the consistency in accuracies for genomic vs. pedigree based prediction methods.

In order to represent the traits under study, two challenge experiments were performed. (1) Differently to previous works, where immersion was used as the way of infection (Antonello et al. 2009; Palaiokostas et al. 2016), this challenge was performed through intramuscular injection. Previous studies reported that the absorption through gills is the most representative modality of the natural infection (Chang et al., 2014). Anyway, the infection by immersion could not be complete (Zimmer et al., 2014). On the contrary, the intramuscular infection is generally fast and complete (Chang et al., 2014). The experimental intramuscular infection performed in this work induced a significant mortality (38.5 %) with peak mortalities at day 5 post infection which reached to asymptote around day 8, whereas studies from Antonello et al. (2009) and Palaiokostas et al. (2016) showed much wider distribution of mortalities with high mortalities falling between day 7 to 14. The challenge test conducted adopted model of this study ultimately produced relatively more variation and informative mortality curve leaving higher proportion of survivors compared previously adopted challenge test models where survival was very low, rather a suitable based population to investigate the genetic architecture to *Phdp* resistance (mortality range 30-70%). External lesions were observed only at the injection site, otherwise no other external lesions were observed in accordance with previous challenge on gilthead seabream (Noya et al., 1995; Cerezuela et al., 2012). (2) Challenge was successful and mimic (in a faster profile) the progression of the disease observed in the field. The parasitic intensity and the observed disease symptoms (anorexia, lethargy, anaemia) reflected the characteristics of the pathology found in the production units and in previous challenge experiments (De Vico et al., 2008). The initiation of the infection was faster compared to what has been noticed in the field where naïve fish are regularly infected with *Sc* within a two-three month period after their introduction to the sea cages. Perhaps the progression of the disease is accelerated in a controlled environment, which is in agreement with the findings of previous challenge studies in tanks (5-6 weeks, Sitjà-Bobadilla et al., 2009) and previous challenge where the same ratio of donor/recipient was used (0.25, Rigos et al., 2015; Henry et al., 2015). Cumulative mortality in the present study (15%) was relatively higher than that documented from field infections with the parasite (Sanz 1992; Alvarez-Pellitero 2004;

Vagianou et al., 2006) but reflected the experimental infection described by Rigos et al. in 2015. Fish losses are due to the reduced capacity for oxygen transport and to mechanical lesions of gills induced by the parasitic attachment (Henry et al., 2015). The lower final biomass of *Sc* -challenged gilthead sea bream (43 vs 58 g) closely reflects what is usually observed in the field, which is undoubtedly the most crucial constrain for the production of caged gilthead sea bream. The infestation level observed, expressed as the mean intensity of *Sc*, approaches published values (Rigos et al., 2015), although occasionally, peaked values from field outbreaks can be by far higher (< 70 in the first gill arch).

5.2 Real-time PCR assay for detection and identification of *Phdp*

The development of fast and sensitive tools for the detection of viruses and bacteria is crucial for disease control and management in fish farms. In the last decade, many efforts have been devoted to developing diagnostic assays for the identification of *Photobacterium* sp. (Osorio et al., 2000, Romalde, 2002, Rajan et al., 2003, Zappulli et al., 2005, Amagliani et al., 2009, Martins et al. 2015). While they demonstrate good sensitivity in *Photobacterium* sp detection, the existing tests are either not quantitative (Osorio et al., 2000, Romalde et al, 2002, Rajan et al., 2003, Zappulli et al. 2005, Amagliani et al., 2009) or require further PCR product manipulation to properly discriminate *P. damselae damselae* and *P. damselae piscicida* subspecies (Osorio et al., 2000, Zappulli et al. 2005, Amagliani et al., 2009, Martins et al. 2015).

In the present study, a qPCR assay has been developed for the simultaneous detection and quantification of *P. damselae*. As a consequence of the two SNPs located in the PCR target region (see Methods), the qPCR assay generates two separated melting profiles for each subspecies that allow a rapid and accurate differentiation between *Phdp* and *Phdd* by simply analyzing the qPCR dissociation curve. To our knowledge, this is the first report of a molecular method for the diagnosis of *P. damselae* that couples detection, quantification, and subspecies identification in a single step. The ability to discriminate between both *Photobacterium* subspecies was validated on a total of 19 *Phdp* and 11 *Phdd* strains isolated from Europe, USA, and Japan, proving the assay is able to overcome the documented limitation due to the genetic variability of isolates from diverse hosts and geographic origins (Costas et al. 2013; Magarinos et al. 2000). The observed difference of at least 0.3°C between the two subspecies has been already proved sufficient/suitable for differentiating virus and bacterial species of medical and veterinary importance (Gelaye et al. 2017, Miller et al. 2015) as well as for the monitoring of bacterial communities in food (Sardaro et al. 2016). In fact, both classical melting curve analysis and High Resolution Melting (HRM) analysis, have been described as powerful techniques for variant scanning and genotyping, enabling to analyze even single genetic variants in nucleic acid sequences (Wittwer, 2008).

Since the common occurrence of co-infections in aquatic animals (Kotob et al., 2016), the assay specificity was tested on a collection (N= 17) of important fish-pathogens bacteria such as *Vibrio anguillarum*, *Vibrio furnissi*, *Aeromonas hydrophila*, *Yersinia ruckeri, Tenacibaculum maritimum*. In addition, some aquatic environmental bacteria were also included (*Vibrio fisheri*, *Vibrio alginolyticus*, *Vibrio campbelli*). The obtained results, consistently validated in all technical replicates, demonstrated the specificity of the assay that thus overcomes either co-infection and natural contamination issues.

Photobacterium damselae subspecies damselae, isolated for the first time in the '80s, has been associated to diseases in many marine fish species, especially the gilthead sea bream and European sea bass (Abdel-Aziz et al., 2013, Labella et al. 2011, Terceti et al. 2016). In addition, it has been reported as a human pathogen capable of causing fatal infections (Shin et al. 1996, Yamane et al. 2004, Rivas et al. 2013); for these reasons, the identification of this subspecies may hold the advantage to monitor its occurrence and evaluate proper management procedures.

The extremely high mortality induced by *Phdp* has increased the need to develop effective strategies to prevent disease outbreaks. In this context, the development of diagnostic tools that are able to rapidly detect *Phdp* at very low concentrations is of crucial importance for early intervention. With a theoretical LOQ of 9 copies of target DNA in spiked conditions (1 copy when pure bacterial DNA is employed), the assay developed in the present study shows a sensitivity of at least one or two orders of magnitude higher than the assays proposed by Amagliani et al. (i.e. 100 molecules of pure bacterial DNA) and by Martins et al. (i.e. $1x10^3$ molecules of pure bacterial DNA), respectively. The only previous assay with a LOD lower than 10 copies was the one proposed by Zappulli et al. (2005), however its sensitivity has only been tested on pure bacterial DNA and an additional RFLP protocol on amplified DNA is needed to discriminate between *Phdp* and *Phdd* subspecies. The use of closed-tube, SYBR-Green chemistry represents a practical and robust tool for rapid screening and diagnosis of bacterial fish diseases (e.g. Duodu et al., 2012), which is cost-effective and avoids contamination-prone post-PCR manipulations.

In addition to providing a highly sensitive, cost-effective, and robust diagnostic test, the proposed real-time PCR protocol guarantees accurate quantification of the target bacteria. The ability to quantify *Phdp* is important to improve our knowledge on infection dynamics. In this context, the developed qPCR assay was applied to 175 samples of farmed sea bream experimentally infected with *Phdp*. All fish dead during the challenge showed high bacterial load in the three tissues analyzed (i.e spleen, liver and gills) with spleen consistently showing higher levels of bacterial colonization than gills and liver. Recently, the early process of *Phdp* infection was investigated in *Seriola quinqueradiata* by immersion challenge (Nagano et al. 2009). This study clearly indicated the gills as the site of election to initiate infection, while internal organs (spleen, kidney) appeared to be the final tissues

where bacterial proliferation occurs. In the present study, the method employed for experimental infection (intramuscular injection) does not reflect the natural way of bacterial entry in the host organism, yet the results obtained confirm the role of the spleen as an important target organ for bacterial colonization. A similar result was reported also by Mosca and colleagues (2014), who demonstrated the presence of *Phdp* colonies by immunohistochemistry only 24 h post-infection, particularly in the spleen of infected *D. labrax*. Interestingly, quantitative analysis of dead fish showed a very broad range of bacterial load (*e.g.* in the spleen, <50 to >3x10^5 *Phdp* DNA copies). These results may suggest variable sensitivity to infection, with some individuals being more resistant to *Phdp*, which is in agreement with the reported genetic heritability of survival after experimental infection (Antonello et al. 2009, Palaiokostas et al. 2016).

The ability of *Phdp* to survive within phagocytic and non-phagocytic host cells is also well documented, and such a persistence likely explains the chronic form of the infection (Avci et al. 2013, Abu-Elala et al. 2015). This has led to the hypothesis of either an asymptomatic carrier state for the pathogen in adult fish, or of a subclinical infection (Toranzo et al. 2005), however the frequency of such condition(s) in the field remains to be investigated. In general, the ability to fully eliminate the pathogen after the resolution of clinical symptoms is relevant to assess the risk posed by surviving fish after an outbreak. In a first attempt to evaluate such a risk, the assay developed here was used to investigate the positivity to *Phdp* in the spleen of 130 *S. aurata* individuals that survived an experimental infection. No residual bacterial DNA was found in the great majority (98.4%) of fish, suggesting that most survivors might not develop a chronic form, having completely eradicated the invading pathogen. However, much work remains to be done to fully understand several aspects of photobacteriosis, especially in field conditions, and the quantitative test reported herein might provide an important tool to investigate infection dynamics and at the same time enable early detection of the pathogen for better prevention.

5.3 Genetic parameters estimation

With the inclusion of the SNPs-based genomic information, measured heritability values for resistance to *Phdp* in this study (0.31-0.33) were similar/slightly higher to data previously reported by Palaiokostas et al. in 2016 (0.22 and 0.28 for different models) and, anyway, falling within the range reported by Antonello et al. in 2009 (h^2 of 0.12 ± 0.04 for days of survival post challenge, defined as a continuous trait, while it ranged from 0.45 ± 0.04 to 0.18 ± 0.08 for the binary trait dead/alive at a specific day, by using nine microsatellites loci as predictors). No significant difference in heritability estimates of the two traits (dead/survived and day to death) were found, and the genetic correlation was ~1, giving the impression that these were the same traits, possibly

because mortalities happened in a very small window of time and hence making both traits behave the same. In addition, the inclusion of genomic information in the model did not lead to significant differences between the values. Even if the estimated heritability of resistance was moderate compared to those previously reported for disease resistance traits in various aquaculture species (Ødegård et al. 2011), the reported values highlighted the existence of genetic variation for resistance against *Phdp* in seabream and the potential to enhance this trait through selection programs. Regarding the heritability estimation of *Sc*, genomic heritability for ParCount was slightly lower than pedigree information based heritability (0.119 vs 0.129). However, genomic heritability estimates for SGR and BW2 were higher than pedigree estimates (0.223 vs 0.135 and 0.192 vs 0.141). Based on these values, the traits seemed to be low/moderate-heritable (0.11-0.22) showing that there is space for improving this trait through selection breeding programs. In addition, Parasite count showed a strong negative genetic correlation of -0.807±0.24 and -0.769±0.22 with SGR using pedigree vs. genomic information respectively. Similar direction and magnitude of genetic correlation was obtained for PC vs. BW2 with values ranging from -0.549±0.27 and -0.701±0.21 using pedigree vs. genomic information, respectively. These strong negative genetic correlations suggest that indirect selection for growth may help to improve resistance against *Sc* parasite. Similar negative genetic correlations were found for resistance to pathologies and body weight/growth in other aquaculture species, such as, for example, Viral Nervous Necrosis resistance vs body weight in European seabass (− 0.35 ± 0.14, Doan et al., 2017) or resistance against *Piscirickettsia salmoni* vs weight in coho salmon (*Oncorhynchus kisutch*, −0.50 ± 0.13, Yáñez et al., 2016). However, the correlations found in the present work showed large standard errors, which could be due to small available data size.

5.4 2b-RAD: evaluation of this technique's power for genomic selection

As reported in previous studies in many species (scallop, Dou et al., 2016; seabream, Palaiokostas et al., 2016; seabass, Babbucci et al., 2016), the present work confirmed the utility and the flexibility of 2b-RAD for cost-effective genotyping in aquaculture species, generating results which were very similar to those "in silico" obtained and discovering tags at sufficient density to capture SNPs linked to QTL in the two populations. As far as the technical aspect, we observed a low fraction of failed libraries and a low quality/quantity variation between the simultaneously processed libraries (n 96), proving that this protocol can be easily performed and standardized for a high number of samples. Reaching a high coverage is an important to reduce genotyping error rates (Fountain et al., 2016) and the 2b-RAD method allows a fine customization to produce high-quality reads with reasonable coverage per each individual. In general, the *Sc* experiment resulted in a higher number of RAD-tags as well as a higher number of SNPs in comparison to the *Phdp* challenge, a substantial difference probably due to the genetic diversity (different geographical origin) of the two seabream populations. The genomic

analysis yielded a significant peak for seven SNPs crossing the genome-wide Bonferroni threshold and located in LG 17, thus indicating a QTL for resistance to fish photobacteriosis, farther one SNP crossing the Bonferroni threshold for the resistance to Sc. The QTL found are encouraging for practical implementation of selective breeding programs for genetic *Phdp*/*Sc* resistance/tolerance in seabream and demonstrate that there is space for additional investigations. The map length obtained for *Phdp*-challenged population (1970 cM) was slightly longer than the map reported by Tsigenopoulos et al. in 2014 (1769.7 cM) but was half of the length of the one reported by Palaiokostas et al. in 2016 (3899 cM). On the other end of the scale there was the map obtained from *Sc*-challenged population with a total length of 1406 cM; these differences could be due to differences in adopted parameters/methods.

In this work, genomic prediction approaches were tested to investigate the potential of the 2b-RAD-obtained SNPs for estimating resistance and tolerance and assessing accuracy of genomic prediction versus pedigree-based prediction methods. To this aim, two of the most widely used GS methods were applied: GBLUP (genomic best linear unbiased prediction), using realized genomic relationship matrix calculated from the genome-wide SNP markers, and the Bayesian method BayesB. The application of all the genomic prediction models resulted in higher accuracies compared to those achieved using traditional BLUP (PBLUP) and their output varied based on the considered phenotype trait. For *Phdp* resistance, the accuracy of the EBV was 43% higher when using GBLUP in comparison to the sole use of pedigree information. As far as *Sc* tolerance, the obtained results were similar: GBLUP analysis showed an accuracy rate 8-23% higher than those observed using the pedigree-based approach. These results are in accordance with the very recent literature suggesting that the combined use of genome-wide markers (SNPs) and phenotype information of full-sibs families to predict EBV will increase accuracy and response to selection over traditional pedigree-based Best Linear Unbiased Prediction (PBLUP). For example, a work by Correa et al. (2017) reported a comparison between breeding value predictions accuracy for sea louse (*Caligus rogercresseyi*) resistance in Atlantic Salmon (*Salmo salar*) by using different genomic prediction approaches (G-BLUP, Bayesian Lasso and Bayes C) against P-BLUP methodology. The accuracy of genomic predictions increased with increasing SNP density and was higher than pedigree-based BLUP predictions by up to 22%. Another study performed on the same specie (*Salmo salar*) by Bangera et al. (2017) reported that the heritability for resistance against *Piscirickettsia salmoni* estimated from GS methods was significantly higher than PBLUB, thus proving that genomic predictions can accelerate the genetic improvement for resistance to problematic diseases affecting Atlantic salmon.

In this work, similar accuracies were estimated implementing G-BLUP/BayesB approaches. Both BayesianB and G-BLUP methods can predict breeding values with higher accuracies than pedigree-based BLUP but G-BLUP may be the preferred method because of reduced computational time and simple implementation (Correa et al.,

2017, Bangera et al., 2017). The advantage of genomic information over pedigree is double: first, pedigree-based relationships ignore Mendelian sampling; secondly, genomic selection in aquaculture populations allows predictions with higher accuracy especially for traits that cannot be phenotypically measured on the selection candidates (e.g. disease resistance and fillet quality). Nonetheless, relying on genomic predictions, the potential within family can be fully exploited with the use of genomic information.

As far as the *Phdp* experiment, the panel of SNPs discovered in this study seemed to have higher prediction values (0.56) in comparison to that implemented by Palaiokostas et al. in 2016 (0.44). Makowsky and collaborators (2011), focused on several factors that may affect the prediction accuracy of complex traits, reporting as influential factors the number of SNPs and the size of the training population, while no drastic differences between any of the statistical methods considered were found. Similarly, Berenos et al (2014) estimated the heritability of body-size trait in a population of Soay sheep by using three different approaches (pedigree-based and SNPs-based) and reported that the proportion of genetic variance captured by genomic markers increased with marker density. Therefore, the higher number of individuals implicated in the training (n=709) in comparison to those used by Palaiokostas et al. (2016) (n=578) and in the validation populations (320 and 200, respectively), as well as the higher number of SNPs, seemed to have increased the genomic prediction accuracy. Instead, prediction accuracies similar to those obtained by Palaiokostas et al. (2017) were found when the sole pedigree information were implemented.

The results here obtained by applying the genomic prediction approach were encouraging for practical implementation of selective breeding for genetic resistance in seabream, with the genomic prediction models outperforming the traditional BLUP approach and showing the potential to enhance resistance through family-based selection. The higher percentage of accuracy achieved with GBLUP, ultimately promises higher genetic gain, as the selection accuracy is directly proportional to the rate of genetic gain. In conclusion, our data suggest that candidate genes putatively associated with resistance to *Phdp*/*Sc* in sea bream might be identified by following a genome-wide SNP approach using 2b-RAD sequencing.

6 Conclusions

The results obtained in this work could help to improve the performance of seabream cultivation, through a faster/accurate diagnosis of diseases and through the inclusion of selected markers of resistance in new breeding programs. The proposed assay, with high specificity and sensitivity, is able to quickly diagnose the disease, so that positive results on production efficiency can be achieved by the implementation of prompt resolution measures. On the other hand, the 2b-RAD technique has proved to be efficient and cost-effective for discovering SNPs linked to QTL in a large number of samples (ca. 2000 seabream individuals), identifying a quantitative trait locus affecting resistance to fish photobacteriosis and a quantitative trait locus linked to the tolerance to *Sparicotyle Chrysophrii*. Estimated breeding values using genomic (GBLUP) vs. pedigree (PBLUP) information presented 8-43% higher accuracy, proving that genomic selection gives the opportunity to perform within and between family selections with relatively higher accuracy, which ultimately promise higher genetic gain. The insertion of significant favorable SNPs in MAS selection may prove efficient cost effective strategy. In conclusion, these results hold the potential to advance in sophistication and efficiency the breeding industry for gilthead seabream, allowing it to address complex breeding goals, as well as the aquaculture industry in general, with the availability of a new assay for rapid diagnosis of *Phdp*. The impacts will concern a more accurate selection for productivity and disease resistance and greater economic performance for seabream cultivation. Reducing the impact of *Phdp*/*Sc* incidence will prompt an expansion of the industry as these diseases are currently hampering development. Further work should be carried out in the future to confirm these promising results as well as to locate markers in the seabream genome, thus investigating interesting genes for resistance.

Acknowledgments

Challenge tests were performed at the Istituto Profilattico Sperimentale delle Venezie (IZS, Legnaro, Italy) and at the Hellenic Centre for Marine Research (HCMR, Athens, Greece). All the experiments and analysis have been discussed and conducted in close cooperation with all the partners FISHBOOST, in particular with the Norwegian University of Life Science (NOFIMA, Norway). I had the opportunity and the pleasure to spend a few time of my PhD at the HCMR of Crete under the supervision of prof. Costas Tsigenopoulos where I learnt the dd-RAD technique and lot of bioinformatics tricks. I am grateful to prof. Tomaso Patarnello, prof. Luca Bargelloni and all the people belonging to the research group to have allowed this work.

References

Abdel-Aziz M., Eissa A.E., Hanna M. & Okada M.A. (2013) Identifying some pathogenic Vibrio/Photobacterium species during mass mortalities of cultured Gilthead seabream (Sparus aurata) and European seabass (Dicentrarchus labrax) from some Egyptian coastal provinces. International Journal of Veterinary Science and Medicine 1, 87-95.

Abu-Elala N.M., Abd-Elsalam, R.M. & Marzouk, M.S. (2015) Molecular and Immunohistochemical Diagnosis of Photobacterium damselae Subspecies piscicida During Naturally Occurring Disease in Egypt. Journal of the World Aquaculture Society 46, 583-595.

Alvarez-Pellitero, P. (2004) Report about fish parasitic diseases. Etudes et Recherches, Options Mediterranennes. CIHEAM/FAO, Zaragoza (2004): 103-130.

Amagliani G., Omiccioli E., Andreoni F., Boiani R., Bianconi I., Zaccone R., Mancuso M. & Magnani M. (2009) Development of a multiplex PCR assay for Photobacterium damselae subsp. piscicida identification in fish samples. Journal of Fish Diseases 32, 645-653.

Andreoni F. & Magnani M. (2014) Photobacteriosis: prevention and diagnosis. Journal of immunology research 2014.

Antonelli L., Quilichini Y. & Marchand B. (2010) Biological study of Furnestinia echeneis Euzet and Audouin 1959 (Monogenea: Monopisthocotylea: Diplectanidae), parasite of cultured gilthead sea bream Sparus aurata (Linnaeus 1758) (Pisces: Teleostei) from Corsica. Aquaculture 307.3, 179-186.

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

Avci H., Birincioglu S., Epikmen E. & Dereli M. (2013) Comparative histopathological and immunohistochemical evaluations in juvenile sea bass (Dicentrarhus labrax) and gilthead sea bream (Sparus aurata) naturally infected with Photobacterium damselae subsp. piscicida. Revue de Medicine Veterinaire 164, 72-79.

Babbucci M., Ferraresso S., Pauletto M, Franch R., Papetti C., Patarnello T., Carnier P. & Bargelloni L. (2016) An integrated genomic approach for the study of mandibular prognathism in the European seabass (Dicentrarchus labrax). Scientific Reports 6.

Baird N.A., Etter P.D., Atwood T.S., Currey M.C., Shiver A.L., Lewis Z.A., Selker E.U., Cresko W.A. & Johnson E.A. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PloS one 3.10, e3376.

Bangera R., Correa K., Lhorente J.P., Figuero, R. & Yáñez J.M. (2017) Genomic predictions can accelerate selection for resistance against Piscirickettsia salmonis in Atlantic salmon (Salmo salar). BMC genomics,

18.1, 121.

Bérénos C., Ellis P.A., Pilkington J.G. & Pemberton J.M. (2014) Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. Molecular ecology, 23.14, 3434-3451.

Caffara M., Quaglio F., Fioravanti M.L., Gustinelli A., Marcer F., Moscato M. & Caggıano M. (2005) Coinfezione da Polysporoplasma sparis (Myxozoa) e Sparicotyle chrysophrii (Monogenea) in orata (Sparus aurata). Atti Del Xii Convegno Nazionale S.I.P.I., Cesenatico (FC), 47.

Campos de los G., Hickey J.M., Pong-Wong R., Daetwyler H.D. & Calus M.P.L. (2012) Whole genome regression and prediction methods applied to plant and animal breeding. Genetics 193, 327-345.

Catchen J.M., Amores A., Hohenlohe P., Cresko W. & Postlethwait J.H. (2011) Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes, genomes, genetics 1(3), 171-182.

Catchen J., Hohenlohe P.A., Bassham S., Amores A. & Cresko W.A. (2013) Stacks: an analysis tool set for population genomics. Molecular ecology 22(11), 3124-3140.

Cerezuela R., Guardiola F.A., Meseguer J. & Esteban M.Á. (2012) Increases in immune parameters by inulin and Bacillus subtilis dietary administration to gilthead seabream (Sparus aurata L.) did not correlate with disease resistance to Photobacterium damselae. Fish & shellfish immunology 32.6, 1032-1040.

Chang C.J., Robertsen C., Sun B. & Robertsen B. (2014) Protection of Atlantic salmon against virus infection by intramuscular injection of IFNc expression plasmid. Vaccine 32, 4695-4702.

Chavanne H., Janssen K., Hofherr J., Contini F., Haffray P., Aquatrace Consortium, Komen H., Nielsen E.E. & Bargelloni L. (2016) A comprehensive survey on slective breeding programs and seed market in the European aquaculture fish industry. Aquaculture International 24, 1287-1307.

Clarke G.M., Anderson C.A., Petterson F.H., Cardon L.R., Morris A.P. & Zondervan K.T. (2011) Basic statistical analysis in genetic case-control studies. Nature protocols 6, 121-133.

Coltman, David W. (2005) Testing marker-based estimates of heritability in the wild. Molecular Ecology 14.8, 2593-2599.

Complex Trait Consortium (2007) The nature and identification of quantitative trait loci. Nature Reviews. Genetics 4.11, 911-916.

Correa K., Bangera R., Figueroa R., Lhorente, J. P. & Yáñez, J. M. (2017) The use of genomic information increases the accuracy of breeding value predictions for sea louse (Caligus rogercresseyi) resistance in Atlantic salmon (Salmo salar). Genetics Selection Evolution 49.1, 15.

Costas B., Rêgo, P.C.N.P., Simões I., Marques J.F., Castro-Cunha M. & Alfonso A. (2013) Cellular and humoral immune responses of Senegalese sole, Solea Senegalensis (Kaup), following challenge with two Photobacterium damselae subsp. piscicida strains from different geographical origins. Journal of Fish Diseases 36, 543-553.

Cruz E., Silva M.P., Freitas M.S. & Orge M.L. (1997) Co-infection by monogenetic trematodes of the genus Microcotyle, Beneden & Hesse, 1863, Lamellodiscus ignoratus Palombi, 1943, the protozoan Trichodina sp. Ehrenberg, 1838 and the presence of epitheliocystis, Vibrio alginolyticus and V. vulnificus in cultured seabream (Sparus aurata L.) in Portugal. Bulletin of the European Association of Fish Pathologists 17, 40-42.

Davey J.W. & Blaxter M.L. (2010) RADSeq: next-generation population genetics. Briefings in functional genomics 9.5-6, 416-423.

Davey J.W., Hohenlohe P.A., Etter P.D., Boone J.Q., Catchen J.M. & Blaxter M.L. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature reviews. Genetics 12.7, 499.

De Vico G., Cataldi M., Carella F., Marin, F. & Passantino A. (2008) Histological, histochemical and morphometric changes of splenic melanomacrophage centers (Smmcs) in sparicotyle-infected cultured sea breams (Sparus aurata). Immunopharmacology and immunotoxicology, 30.1, 27-35.

Doan K.Q., 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.

Dou J., Li X., Fu Q., Jiao W., Li Y., Li T., Wang J., Hu X., Wang S. & Bao, Z. (2016) Evaluation of the 2b-RAD method for genomic selection in scallop breeding. Scientific reports, 6, 19244.

Duodu S., Larsson P., Sjödin A., Soto E., Forsman M. & Colquhoun D.J. (2012) Real-time PCR assays targeting unique DNA sequences of fish-pathogenic Francisella noatunensis subspecies noatunensis and orientalis. Diseases of aquatic organisms 101, 225-234.

FEAP, 2016. European Aquaculture Production Report.

Fernando R.L. & Grossman M., 1989. Marker assisted selection using best linear unbiased prediction. Genetics Selection Evolution 21, 467-477.

Fountain E.D., Pauli J.N., Reid B.N., Palsbøll P.J. & Peery M.Z. (2016) Finding the right coverage: the impact of coverage and sequence quality on single nucleotide polymorphism genotyping error rates. Molecular ecology resources, 16.4, 966-978.

Franch R., Louro B., Tsalavouta M., Chatziplis D., Tsigenopoulos C.S., Sarropoulou E., Antonello J., Magoulas A., Mylonas C.C., Babbucci M., Patarnello T., Power D.M., Kotoulas G. & Bargelloni L. (2006) A genetic linkage map of the hermaphrodite teleost fish Sparus aurata L. Genetics 174.2, 851-861.

Fu L., Niu B., Zhu Z., Wu S. & Li W. (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28.23, 3150-3152.

Gelaye E., Mach L., Kolodziejek J., Grabherr R., Loitsch A., Achenbach J.E., Nowotny N., Diallo A. & Lamien C.E. (2017) A novel HRM assay for the simultaneous detection and differentiation of eight poxviruses of medical and veterinary importance. Scientific Reports 7, 42892.

Gilmour A.R., Gogel B.J., Cullis B.R., Welham S.J. & Thompson R. (2015) ASReml User Guide Release 4.1 Structural Specification. VSN International Ltd 2015.

Gjedrem T. (2012) Genetic improvement for the development of efficient global aquaculture: a personal opinion review. Aquaculture 344, 12-22.

Gjedrem T. & Rye M. (2016) Selection response in fish and shellfish: a review. Reviews in Aquaculture, 2016.

Gjøen H.M. & Bentsen H.B. Past, present and future of genetic improvement in salmon aquaculture (1997) Ices journal of marine science 54, 1009-1014.

Goddard, M.E. & Hayes B.J. (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nature Reviews Genetics 10(6), 381-391.

González-Lanza C., Alvarez-Pellitero P. & Stijá-Bobadilla (1991) Diplectanidae (Monogenea) infestations of sea bass, Dicentrarchus labrax (L.), from the Spanish Mediterranean area. Parasitology Research 77, 307-314.

Groen A. & Van Arendok J., 1995. Programmes Lecture notes for E250-210. Department of Animal Breeding, Wageningen university.

Habier D., Fernando R.L., Kizilkaya K., Garrick D.J. (2011) Extension of the bayesian alphabet for genomic selection. BMC Bioinformatics 12, 186-186.

Hawke J.P., Thune R.L., Cooper R.K., Judice E. & Kelly-Smith M. (2003). Molecular and phenotypic characterization of strains of Photobacterium damselae subsp. piscicida isolated from hybrid striped bass cultures in Louisiana, USA. Journal of Aquatic Animal Health 15,189-201.

Henry, M.A., Nikoloudaki C., Tsigenopoulos C. & Rigos G. (2015) Strong effect of long-term Sparicotyle chrysophrii infection on the cellular and innate immune responses of gilthead sea bream, Sparus aurata. Developmental & Comparative Immunology 51(1), 185-193.

Houston R.D., Taggart J.B., Cézard T., Bekaert M., Lowe N.R., Downing A., Talbot R., Bishop S.C., Archibald A.L., Bron J.E., Penman D.J., Davassi A., Brew F., Tinch A., Gharbi K. & Amilton A. (2014) Development and validation of a high density SNP genotyping array for Atlantic salmon (Salmo salar). BMC genomics 15.1, 90.

Houston R.D., Haley C.S., Hamilton A., Guy D.R., Tinch A.E., Taggart J.B., McAndrew B.J. & Bishop S.C. (2008) Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic salmon (Salmo salar). Genetics 178.2, 1109-1115.

Houston R.D., Haley C.S., Hamilton A., Guy D.R., Mota-Velasco J.C., Gheyas A.A., Tinch A.E., Taggart J.B., Bron J.E., Starkey W.G., McAndrew B.J., Verner-Jeffreys D.W., Paley R.K., Rimmer G.S.E., Tew i.J &Bishop S.C. (2010) The susceptibility of Atlantic salmon fry to freshwater infectious pancreatic necrosis is largely explained by a major QTL. Heredity 105.3, 318-327.

Kalinowski S.T., Taper M.L., Marshall T.C. (2007) Revising how the computer program cervus

accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16.5, 1099-1106.

Kotob M.H, Menanteau-Ledouble S., Kumar G., Abdelzaher M. & El-Matbouli M. (2016) The impact of co-infections on fish: a review. Veterinary Research 47, 98-110.

Labella A., Berbel C., Manchado M., Castro D. & Borrego J.J. (2011) Photobacterium damselae subsp. damselae, an emerging pathogen affecting new cultured marine fish species in southern Spain. Recent Advances in Fish Farms, eds Aral Faruk, Doğu Zafer editors. (New York: InTech), 135-152.

Li H, Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G. & Durbin R. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25.16, 2078-2079.

Lind C.E., Ponzoni R.W., Nguyen N.H. & Khaw H.L. (2012). Selective breeding in fish and conservation of genetic resources for aquaculture. Reproduction in Domestic Animals 47, 255-263.

Lipschutz-Powell D., Woolliams J.A., Bijma P., Pong-Wong R., Bermingham M.L. & Doeschl-Wilson A.B. (2012) Bias, accuracy, and impact of indirect genetic effects in infectious diseases. Frontiers in Genetics 3, 215.

Luo X., Shi X., Yuan C., Ai M., Ge C., Hu M., Feng X. & Yang X. (2017) Genome-wide SNP analysis using 2b-RAD sequencing identifies the candidate genes putatively associated with resistance to ivermectin in Haemonchus contortus. Parasites & vectors 10.1, 31.

Magarinos B., Romalde J.L., Lemos M.L., Barja J.L. & Toranzo A.E. (1994) Iron uptake by Pasteurella piscicida and its role in pathogenicity for fish. Applied and Environmental Microbiology 60, 2990-2998.

Magarinos B., Toranzo A.E., Barja J.L. & Romalde J.L. (2000) Existence of two geographically-linked clonal lineages in the bacterial fish pathogen Photobacterium damselae subsp. piscicida evidenced by random amplified polymorphic DNA analysis. Epidemiology and Infection 125, 213-221

Makowsky R., Pajewski N.M., Klimentidis Y.C., Vazquez A.I., Duarte C.W., Allison D.B. & de Los Campos G. (2011) Beyond missing heritability: prediction of complex traits. PLoS genetics 7(4), e1002051.

Maroso F., Franch R., Dalla Rovere G., Arculeo M. & Bargelloni L. (2016) RAD SNP markers as a tool for conservation of dolphinfish Coryphaena hippurus in the Mediterranean Sea: Identification of subtle genetic structure and assessment of populations sex-ratios. Marine Genomics 28, 57-62.

Martins P., Navarro R.V., Coelho F.J. & Gomes N.C. (2015) Development of a molecular methodology for fast detection of Photobacterium damselae subspecies in water samples. Aquaculture 435, 137-142.

Massault C., Franch R., Haley C., de Koning D.J., Bovenhuis H., Pellizzari C., Patarnello T. & Bargelloni L (2011) QTLs for resistance to fish pasteurellosis in gilthead seabream Sparus aurata. Animal Genetics 42, 191-203.

Mastretta-Yanes A., Arrigo N., Alvarez N., Jorgensen T.H., Piñero D., & Emerson B.C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. Molecular Ecology Resources 15(1), 28-41.

McKenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K., Altshuler D., Gabriel S., Daly M., DePristo & M.A. (2011) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 20, 1297-1303.

Meuwissen T.H. & Luo Z. (1992) Computing inbreeding coefficients in large populations. Genetics, Selection, Evolution 24.4, 305-313.

Meuwissen T.H. & Goddard M.E. (2001) Prediction of identity by descent probabilities from marker-haplotypes. Genetics Selection Evolution 33.6, 605. (a)

Meuwissen T.H., Hayes B.J. & Goddard M.E. (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157.4, 1819-1829. (b)

Midtlyng P.J., Reitan L.J. & Speilberg L. (1996) Experimental studies on the efficacy and side-effects of intraperitoneal vaccination of Atlantic salmon (Salmo salar L.) against furunculosis. Fish & Shellfish immunology 6, 335-350.

Miller M., Zorn J. & Brielmeier M. (2015) High-Resolution Melting Curve Analysis for Identification of Pasteurellaceae Species in Experimental Animal Facilities. PLoS One 10, e0142560.

Miller M.R., Dunham J.P., Amores A., Cresko W.A. & Johnson E.A. (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome research 17.2, 240-248.

Moen T., Baranski M., Sonesson A.K. & 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(1), 368.

Mosca F., Ciulli S., Volpatti D., Romano N., Volpe E., Bulfon C., Massimini M., Caccia E., Galeotti M. & Tiscar P.G. (2014) Defensive response of European sea bass (Dicentrarchus labrax) against Listonella anguillarum or Photobacterium damselae subsp. piscicida experimental infection. Veterinary immunology and immunopathology 162, 83-95.

Nagano I., Inoue S., Kawai K. & Oshima S. (2009) Repeatable immersion infection with Photobacterium damselae subsp. piscicida reproducing clinical signs and moderate mortality. Fisheries Science 75, 707-714.

Noga J.E. (2011) Fish disease: diagnosis and treatment. Wiley-Blackwell ed.

Noya M., Magarinos B., Toranzo A. E. & Lamas, J. (1995) Sequential pathology of experimental pasteurellosis in gilthead seabream Sparus aurata. A light-and electron-microscopic study. Diseases of Aquatic Organisms 21.3, 177-186.

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

Osorio C., Toranzo A., Romalde J. & Barja J. (2000) Multiplex PCR assay for ureC and 16S rRNA genes

clearly discriminates between both subspecies of Photobacterium damselae. Diseases of aquatic organisms 40, 177-183.

Osorio C.R., Collins M.D., Toranzo A.E., Barja J.L. & Romalde J.L. (1999) 16S rRNA gene sequence analysis of Photobacterium damselae and nested PCR method for rapid detection of the causative agent of fish pasteurellosis. Applied and Environmental Microbiology 65, 2942-2946

Padrós F., & Crespo S. (1995) Proliferative epitheliocystis associated with monogenean infection in juvenile seabream Sparus aurata in the North East of Spain. Bulletin of the European Association of Fish Pathologists (United Kingdom).

Palaiokostas C., Ferraresso S., Franch R., Houston R.D. & Bargelloni L. (2016) Genomic Prediction of Resistance to Pasteurellosis in Gilthead Sea Bream (Sparus aurata) Using 2b-RAD Sequencing. G3: Genes|Genomes|Genetics 6.11, 3693-3700.

Pathak S., Awuh J.A., Leversen N.A, Flo T.H. & Åsjø B. (2012) Counting mycobacteria in infected human cells and mouse tissue: a comparison between qPCR and CFU. PLoS One 7, e34931.

Pauletto M., Carraro L., Babbucci M., Lucchini R., Bargelloni L & Cardazzo B. (2016) Extending RAD tag analysis to microbial ecology: a comparison between MultiLocus Sequence Typing and 2b-RAD to investigate Listeria monocytogenes genetic structure. Molecular ecology resources 16(3), 823-835.

Pecoraro C., Babbucci M., Villamor A., Franch R., Papetti C., Leroy B., Ortega-Garcia S., Muir J., Rooker J., Arocha F., Zudaire I., Chassot E., Bodin N., Tinti F., Bargelloni L. & Cariani A. (2016) Methodological assessment of 2b-RAD genotyping technique for population structure inferences in yellowfin tuna (Thunnus albacares). Marine genomics 25, 43-48.

Pérez P. & de los Campos G. (2014) Genome-Wide Regression and Prediction with the BGLR Statistical Package. Genetic*s* 198.2, 483-495.

Peterson B.K., Weber J.N., Kay E.H., Fisher H.S. & Hoekstra H.E. (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS one, *7*.5, e37135.

Powell J., White I., Guy D. & Brotherstone S. (2008) Genetic parameters of production traits in Atlantic salmon (Salmo salar). Aquaculture 274, 225-231.

Puritz J.B., Matz M.V., Toonen R.J., Weber J.N., Bolnick D.I. & Bird C.E. (2014) Demystifying the RAD fad. Molecular Ecology 23, 5937-5942.

Rabet C.M., Ensibi C., Dhaouadi R. & Yahia O.K. (2016) A preliminary study on gill parasites of gilthead seabream Sparus aurata (Linnaeus 1758) (Pisces: Teleostei) from the eastern Tunisian sea-cage aquaculture. GERF Bull Biosc 7, 1-5.

Rajan P., Lin J., Ho M. & Yang H. (2003) Simple and rapid detection of Photobacterium damselae ssp. piscicida by a PCR technique and plating method. Journal of applied microbiology 95, 1375-1380.

Rastas P., Calboli F.C.F., Guo B., Shikano T. & Merilä J. (2015) Construction of ultra-dense linkage maps

with Lep-MAP2: stickleback F2 recombinant crosses as an example. Genome Biology and Evolution 2015.

Rigos G., Henry M. & Tsigenopoulos C. (2015) Sparicotyle chrysophrii and gilthead sea bream-potential experimental infection model. Bulletin European Association of Fish Pathologists 35, 49-53.

Rivas A.J., Balado M., Lemos M.L. & Osorio C.R. (2011) The Photobacterium damselae subsp. damselae hemolysins and HlyA are encoded within a new virulence plasmid. Infection and Immunity 79.11, 4617-4627.

Rivas A.J., Lemos M.L. & Osorio C.R. (2013) Photobacterium damselae subsp. damselae, a bacterium pathogenic for marine animals and humans. Frontiers in microbiology 4, 283.

Robertsen G., Oldstad K., Plaisance L., Bachmann L. & Bakke T.A. (2008) Gyrodactylus salaris (Monogenea, Gyrodactylidae) infections on resident Arctic charr (Salvelinus alpinus) in southern Norway. Environmental Biology of Fishes 83, 99.

Robledo D., Palaiokostas C., Bargelloni L., Martínez P. & Houston R. (2017) Applications of genotyping by sequencing in aquaculture breeding and genetics. Reviews in Aquaculture 2017.

Romalde J.L. (2002) Photobacterium damselae subsp. piscicida: an integrated view of a bacterial fish pathogen. International Microbiology 5, 3-9.

Russel P.J (2002) iGENETICS. Pearson Education, Inc. by Benjamin Cumming Publishing Company,Inc.

Sanz, F. (1992) Mortality of cultured seabream (Sparus aurata) caused by an infection with a trematode of the genus Microcotyle. Bulletin-european association of fish pathologists 12, 186-186.

Sardaro M.L.S., Levante A., Bernini V., Gatti M., Neviani E. & Lazzi C. (2016) The spxB gene as a target to identify Lactobacillus casei group species in cheese. Food Microbiology 59, 57-65.

Shin J.H., Shin M.G., Suh S.P., Ryang D.W., Rew J.S. & Nolte F.S. (1996) Primary Vibrio damsela septicemia. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 22, 856-857.

Shirakashi S., Kazuhisa T. & Kazuo O. (2008) Altered behaviour and reduced survival of juvenile olive flounder, Paralichthys olivaceus, infected by an invasive monogenean, Neoheterobothrium hirame. International journal for parasitology 38.13, 1513-1522.

Sitjà-Bobadilla A. & Alvarez-Pellitero P. (2009) Experimental transmission of Sparicotyle chrysophrii (Monogenea: Polyopisthocotylea) to gilthead seabream (Sparus aurata) and histopathology of the infection. Folia parasitologica 56.2, 143.

Sitjà-Bobadilla A., de Felipe M.C. & Alvarez-Pellitero P. (2006) In vivo and in vitro treatments against Sparicotyle chrysophrii (Monogenea: Microcotylidae) parasitizing the gills of gilthead seabream (Sparus aurata L.). Aquaculture 261.3, 856-864.

Skjold P., Sommerset I., Frost P. & Villoing S. (2016) Vaccination against pancreas disease in Atlantic

salmon, Salmo salar L., reduces shedding of salmonid alphavirus. Veterinary Research 47.1, 78.

Speed D., Hemani G., Johnson M.R. & Balding D.J. (2012) Improved heritability estimation from genome-wide SNPs. The American Journal of Human Genetics 91.6, 1011-1021.

Stanton-Geddes J., Yoder J.B., Briskine R., Young N.D. & Tiffin P. (2013) Estimating heritability using genomic data. Methods in Ecology and Evolution, 4.12, 1151-1158.

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

Sun X., Liu D., Zhang X., Li W., Liu H., Hong W., ... & Xu, C. (2013). SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. PloS one 8.3, e58700.

Susini F., Fichi G., Macori G., Cocumelli C., Cardeti G., Alimonti C., Forletta R., Gustinelli A. & Fioravanti M.L (2011) Episodio di mortalità da Sparicotyle Chrysophrii in orate (Sparus aurata) in gabbia. Convegno Nazionale della Società Italiana di Patologia Ittica 2011, 57-57.

Tamura K., Stecher G., Peterson D., Filipsk, A. & Kumar S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular biology and evolution 30.12, 2725-2729.

Terceti M.S., Ogut H. & Osorio C.R. (2016) Photobacterium damselae subsp. damselae, an emerging fish pathogen in the Black sea: evidence of a multiclonal origin. Applied and Environmental Microbiology 83, 3736-3745.

Toonen R.J., Puritz J.B., Forsman Z.H., Whitney J.L., Fernandez-Silva I., Andrews K.R. & Bird C.E. (2013) ezRAD: a simplified method for genomic genotyping in non-model organisms. PeerJ 1, e203.

Tsai H.T., Matika O., Edwards S.M.K., Antolin-Sánchez R., Hamilton A., Guy D.R., Tinch A.E., Gharbi K., Stear M.J., Taggart J.B., Bron J.E., Hickey J.M. & Houston R.D. (2017) Genotype imputation to improve the cost-efficiency of genomic selection in farmed Atlantic salmon. G3: Genes, Genomes, Genetics 7.4, 1377-1383.

Tsigenopoulos C. S., Louro B., Chatziplis D., Lagnel J., Vogiatzi E., Loukovitis D., ... & Mylonas C.C. (2014) Second generation genetic linkage map for the gilthead sea bream Sparus aurata L. Marine genomics 18, 77-82.

Vagianou S., Athanassopoulou F., Ragias V., Di Cave D., Leontides L., & Golomazou, E. (2006) Prevalence and pathology of ectoparasites of Mediterranean Seabream and sea bass reared under different environmental and aquaculture conditions, 2006.

Vandeputte M., Garouste R., Dupont-Nivet M., Haffray P., Vergnet A., Chavanne H., Laureau S., T. Benny Ron., Pagelson G. & Mazorra C. (2014) Multi-site evaluation of the rearing performances of 5 wild populations of European sea bass (Dicentrarchus labrax). Aquaculture 424, 239-248.

VanRaden, P.M. (2008) Efficient methods to compute genomic predictions. Journal of dairy science 91.11, 4414-4423.

Visscher P.M., Hill W.G., Wrai N.R. (2008) Heritability in the genomics era – concepts and misconceptions. Nature Reviews Genetics 9, 255-266.

Wang S., Meyer E., McHay J.K. & Matz M.E. (2012) 2b-RAD: a simple and flexible method for genome-wide genotyping. Nature methods 9.9, 808-810.

Wittwer C.T. (2008) High-resolution melting analysis: advancements and limitations. Human mutation 30, 857-859.

Yamane K., Asato J., Kawade N., Takahashi H., Kimura B. & Arakawa Y. (2004) Two cases of fatal necrotizing fasciitis caused by Photobacterium damsela in Japan. Journal of clinical microbiology 42, 1370-1372.

Yáñez J.M., Bangera R., Lhorente J.P., Barría A., Oyarzún M., Neira R. & Newman S. (2016) Negative genetic correlation between resistance against Piscirickettsia salmonis and harvest weight in coho salmon (Oncorhynchus kisutch). Aquaculture 459, 8-13.

Yang J., Lee S.H., Goddard M. E. & Visscher P. M. (2011) GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics 88.1, 76-82.

Zappulli V., Patarnello T., Patarnello P., Frassineti F., Franch R., Manfrin A., Castagnaro M. & Bargelloni L. (2005) Direct identification of Photobacterium damselae subspecies piscicida by PCR-RFLP analysis. Diseases of aquatic organisms 65, 53-61.

Zimmer A.M., Wright P.A. & Wood C.M. (2014) What is the primary function of the early teleost gill? Evidence for Na+/NH+4 exchange in developing rainbow trout (Oncorhynchus mykiss). Proceedings of the Royal Society of London B: Biological Science 2014, 20141422.

Web sources

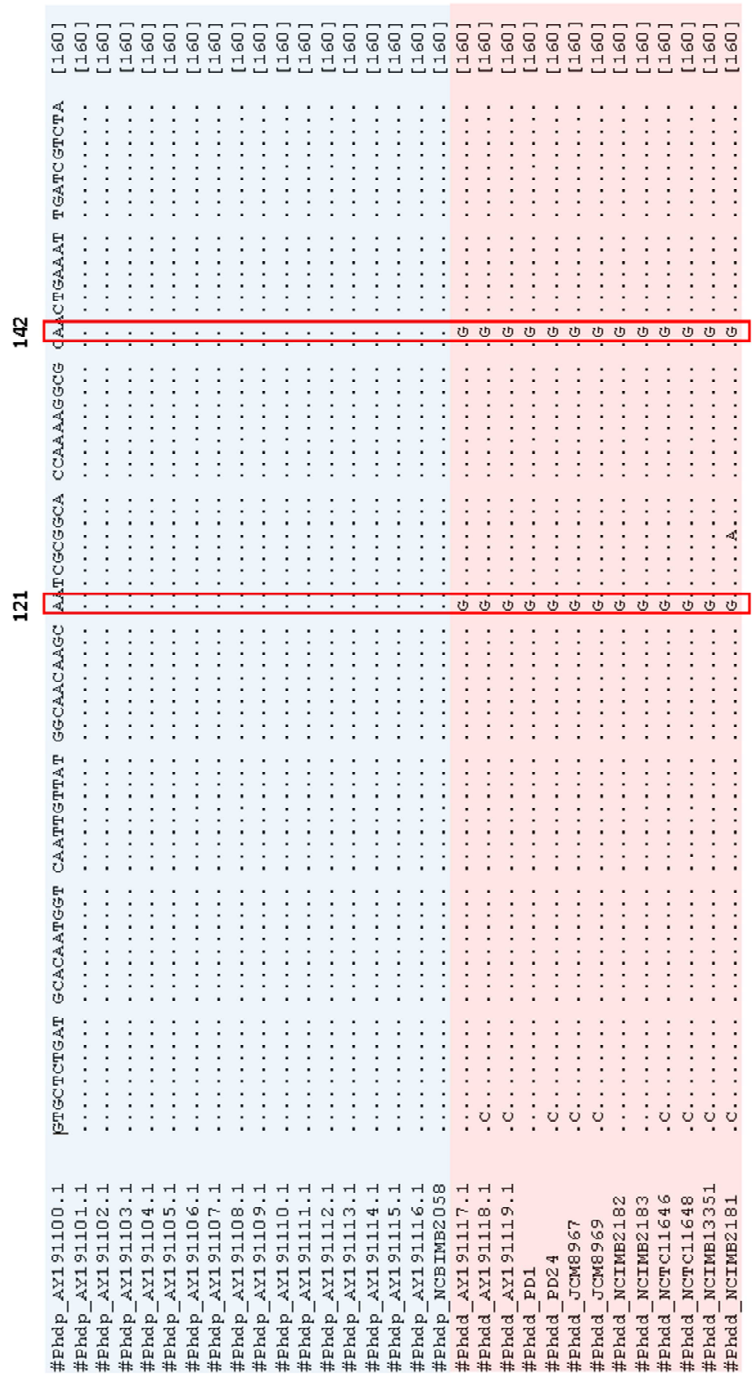
www.FISHBOOST.eu

www.fao.org

www.aquabreeding.eu

Supplementary file 1

Multiple alignment of *P. damselae piscicida* (*Phdp*)/*P. damselae damselae* (*Phdd*) BamB sequences



Supplementary file 2

A. Linkage map based on *Phdp* challenge experiment: LGs IDs with number of markers per each LG, sex-based maps and average length.

Linkage Groups	Number of Markers	Male Map Length (cM)	Female Map Length (cM)	Average Length (cM)
SA01	590	83,76	89,09	82,20
SA02	575	114,24	122,06	112,65
SA03	562	102,31	128,40	112,31
SA04	562	77,55	95,53	82,94
SA05	526	85,57	99,24	84,59
SA06	506	99,13	157,39	112,97
SA07	468	61,85	103,69	76,56
SA08	529	91,42	102,35	92,28
SA09	482	71,15	100,33	81,85
SA10	487	79,84	123,99	92,23
SA11	461	87,35	73,09	77,58
SA12	447	60,61	93,52	76,26
SA13	478	83,13	94,93	81,98
SA14	458	88,42	88,89	83,59
SA15	474	65,61	89,62	76,33
SA16	482	88,90	102,03	88,80
SA17	465	62,77	77,51	67,03
SA18	465	73,35	69,52	67,46
SA19	430	88,03	77,97	79,24
SA20	445	71,01	78,38	71,99
SA21	430	53,40	89,49	71,24
SA22	455	69,23	63,19	63,27
SA23	397	47,07	66,60	56,45
SA24	379	53,23	63,28	55,62
Uknkown	5805	-	-	-
Total	**17358**	**1858,93**	**2250,09**	**1947,44**

B. Linkage map based on *Sc* challenge experiment: LGs IDs with number of markers per each LG, sex-based maps and average length.

Linkage Groups	Number of Markers	Male Map Length (cM)	Female Map Length (cM)	Average Length (cM)
SA01	710	55,44	62,71	57,75
SA02	754	65,70	74,91	70,83
SA03	639	59,83	60,82	59,69
SA04	681	33,89	68,39	51,03
SA05	744	48,74	73,79	61,12
SA06	642	53,30	67,07	60,39
SA07	715	81,19	81,97	80,97
SA08	661	32,68	55,03	43,37
SA09	649	73,21	75,09	72,78
SA10	630	44,21	67,02	54,22
SA11	649	58,26	80,85	69,20
SA12	594	52,79	61,15	54,54
SA13	619	26,17	60,60	43,13
SA14	603	62,05	64,88	63,86
SA15	701	61,72	67,14	63,71
SA16	614	23,80	57,72	40,02
SA17	622	53,92	71,62	62,98
SA18	613	54,90	55,97	54,32
SA19	575	49,93	63,26	56,61
SA20	581	57,17	67,12	61,83
SA21	580	58,93	61,93	60,56
SA22	600	53,69	69,42	61,09
SA23	512	22,84	77,52	50,14
SA24	496	52,00	52,21	51,87
Uknkown	6589	-	-	-
Total	**21773**	**1236,35**	**1598,19**	**1406,02**