Borrell, YJ.; Carleos, CE.; Sanchez, JA.; Vázquez, E.; Gallego Albiach, V.; Asturiano Nemesio, JF.; Blanco, G. (2011). Heterozygosity-fitness correlations in the gilthead sea bream Sparus aurata using microsatellite loci from unknown and gene-rich genomic locations. Journal of Fish Biology. 79:1111-1129. doi:10.1111/j.1095-8649.2011.03099.x.

Title

Heterozygosity-fitness correlations in the Gilthead Sea bream *Sparus aurata* using microsatellite loci from unknown and gene-rich genomic locations.

Authors

Y. J. Borrell[a], C.E. Carleos[b], J.A. Sánchez[a], E. Vázquez[a], V. Gallego[c], J. F. Asturiano[c]. and G. Blanco[a,*]

Postal addresses

[a]*Laboratorio de Genética Acuícola, Departamento de Biología Funcional, Universidad de Oviedo, IUBA, 33071 Oviedo, Spain.*

[b]*Grupo de Estadística Genómica, Departamento de Estadística, Universidad de Oviedo, 33007, Oviedo, Spain.*

[c]*Grupo de Acuicultura y Biodiversidad. Instituto de Ciencia y Tecnología Animal. Universidad Politécnica de Valencia. Camino de Vera s/n. 46022, Valencia, Spain.*

Corresponding Author

Gloria Blanco Lizana: * Corresponding author: Tel: 34 985 103889; Fax: 34 985 103534; E.mail: gloriablanco@uniovi.es.

Running Title

HFCs in *S. aurata*.

Abstract

Heterozygosity-fitness correlations (*HFCs*) were assessed for a sample of a Gilthead Sea bream *Sparus aurata* population. Two hundred and seventy one fish were genotyped at 22 known and novel microsatellite loci, from which correlations between the multilocus heterozygosity index (*MLH*) and various fitness traits (length, weight and specific growth rates) were calculated. Significant global *HFC*s were found in this sample ($0.02 \leq r^2 \leq 0.08$). In addition, all the significant correlations found in this work were negative, indicating that heterozygotes had lower fitness than their homozygote counterparts. Marker location could not explain the observed *HFC*s. Evidence of inbreeding, outbreeding, or population/family structuring was not found in this work. However, the presence of undetected general effects that may lead to the appearance of *HFC*s cannot be ruled out. In any case, these results seem to be best explained by the occurrence of local effects (due to linkage) or even by possible direct locus advantages.

Keywords

HFC, fish, aquaculture, growth, inbreeding, overdominance

Introduction

Heterozygosity-fitness correlations (*HFCs*) have been defined as the empirical observation of a correlation between heterozygosity, as measured at a marker locus or a set of marker loci, and a fitness-related trait (David, 1998). They have historically provided an opportunity to use marker-based genetics to assess two major issues: the genetic basis of inbreeding depression and heterosis and the neutralist-selectionist controversy (Ohta, 1971; Zouros, 1993; David, 1998). Thus, their existence makes *HFCs* a focus of attention, not only for biologists and evolutionists, but also for farm and hatchery managers. The debate is still vigorous, and in the last 5 years, *HFCs* have been reported for several taxa, including more than 40 plant and animal species (Grueber *et al.*, 2008). However, the findings regarding *HFC*s are not universally consistent (see Chapman *et al.* (2009) for a review).

Heterozygosity-fitness correlations may primarily be the result of direct or associative overdominance (Ohta, 1971; Zouros, 1993; David, 1997). Correlations may arise from a heterozygous advantage at certain loci (direct overdominance). Alternatively, an association may result from differences in inbreeding among individuals within a population (associative overdominance via a general effect) or as a consequence of loci being in gametic phase disequilibrium with loci that affect the traits being studied (associative overdominance via a linkage effect) (David, 1997; 1998; Hansson & Westerberg, 2002).

Early studies could detect *HFC*s using alloenzyme data, but not by using noncoding DNA markers (e.g. Zouros & Pogson, 1994). The main assumption at the time was that enzyme loci have a direct influence on fitness components through the control of metabolic reactions (Haldane, 1954; Koehn *et al*., 1988; Pogson, 1991; Mitton, 1993; 1997). Additionally, these findings were the basis for Thelen & Allendorf's reasoning (2001) that allozyme loci could be located in gene-rich regions (i.e., would have a greater probability of linkage with relevant genes), whereas noncoding DNA markers would be mainly located in gene-poor regions of the genome. Despite this thinking, significant positive *HFCs* have also been documented using noncoding markers (Pogson & Fevolden, 1998; Coulson *et al.,* 1998; 1999; Coltman *et al.*, 1999; Rowe *et al.*, 1999; etc.; reviewed by Chapman *et al.* (2009)). The associations between noncoding DNA markers and fitness traits have been interpreted as showing that some correlations are due to factors other than the direct effects of marker genes on the phenotype. Therefore, they support the associative overdominance hypothesis that *HFCs* result from inbreeding or local effects. It has generally been assumed that mean heterozygosity reflects the global level of heterozygosity, which should in turn correlate with individual inbreeding levels, as revealed by noncoding markers (Coltman & Slate, 2003; Pemberton, 2004; Slate *et al.*, 2004; Aparicio *et al*., 2007). However, with analytical and empirical approaches, several authors (e.g., Balloux *et al.* (2004), Slate *et al.* (2004), De Woody & De Woody (2005), Hansson & Westerberg (2008) and Väli *et al.* (2008)) have suggested that *HFC*s are better explained by relationships at the level of the individual markers, or genes linked to them, because DNA marker heterozygosity (mainly at microsatellites) does not reflect genome-wide inbreeding.

Szulkin *et al*. (2010) have published a comprehensive review about this controversial subject recently. They argue that all data supporting *HFC*s is consistent with inbreeding-based theory. In particular, they attempted to demonstrate that linkage disequilibrium is not an alternative to inbreeding but is rather a consequence of some forms of inbreeding and is not restricted to closely linked loci (Szulkin *et al*., 2010). Moreover, they argue that local chromosomal effects on *HFC*s are expected to be small and have rarely, if ever, proven to be statistically significant using adequate tests (Szulkin *et al*., 2010). Curiously, from Szulkin *et al.´*s perspective (2010) the possibility of direct effects has been obviated when using microsatellites for *HFC* studies. Microsatellites are ubiquitous and can sometimes seriously affect gene function. It appears that repeats inside genes are usually negative *cis* modulators of transcription (Streelman & Kocher, 2002; Li *et al.,* 2004; Almuly *et al.,* 2005; Xu *et al.,* 2006; De-Santis & Jerry, 2007; Sharma *et al.,* 2007).

The increased prevalence of inbreeding in small populations (such as those found in aquaculture) facilitates the detection of *HFCs* (Grueber *et al.,* 2008). Under these circumstances, a small number of loci may provide useful information about genomic heterozygosity and inbreeding for *HFC* studies (DeWoody & DeWoody, 2005). Despite this possibility, not much work has been performed regarding *HFCs* in cultured fish populations, and recent studies have focused on only a few species (e.g., *Salmo salar* L. (Borrell *et al.,* 2004), *Anguilla anguilla* (Pujolar *et al.,* 2005; 2006; 2009), and *Salmo trutta* (Tiira *et al.,* 2006)). This is perhaps due to the difficulties

involved in conducting experiments using these species, including the duration of fish life cycles and the space and number of individuals required, among other factors.

Sparus aurata (Linnaeus 1758), a member of the Sparidae family, is one of the most important farmed fish in Europe (especially in the Mediterranean area). *S. aurata* is a protandrous hermaphroditic mass spawning species. Individuals are males for the first two years of life and then become females. Members of the species breed once a year during a six- to eight-week period (Zohar *et al.*, 1978; 1995). However, not all Gilthead Sea bream males seem to follow this pattern; some animals either delay or never attain sex reversal, possibly in relation to social, environmental, and/or genetic factors (Zohar *et al.,* 1995). Spawning in the wild takes place in large groups or schools. Therefore, a minimum of five to seven fish appears to be necessary for reduced stress and natural spawning in artificial environments (Brown, 2003 and references therein). As far as one knows, no previous research has been conducted to assess *HFC*s in *S. aurata*.

This research have conducted to pursue the following aims: 1) to determine if *HFCs* exist in the Gilthead Sea bream *S. aurata* using 22 microsatellites and the phenotypic parameters length, mass and specific growth rate (which are generally accepted as fitness-related traits (Chapman *et al.*, 2009)) and 2) to assess which mechanism(s) that could explain *HFCs* (if present) in the Gilthead Sea bream. The microsatellite sets used here contain both classical microsatellites from genomic

libraries (*CM*s; genomic locations unknown) and microsatellites markers located in gene-rich regions (gene-rich microsatellites, *GRM*).

Materials and Methods

Samples and fitness measurements

A population of two-year-old Gilthead Sea bream individuals of unknown origin that had reached commercial size was sampled from the marine farm Safor, S.L. (Gandía, Valencia). A total of 271 fish were selected based on phenotype (no evidence of diseases and/or physical deformities) as potential breeders for the research project (*CRIOGEN 2003*) at the beginning of 2004 (January). These individuals where tagged using microchips. Small caudal fin pieces were collected, stored in absolute ethanol, and sent to the University of Oviedo.

Masses were recorded in June 2004 (Mass 1, M1), June 2005 (Mass 2, M2), and June 2006 (Mass 3, M3), at which time the length of each fish was also measured (Length, L). Mortality was extremely low during this period (~4%), and individuals that did not survive were excluded from all the analyses. Growth rates were measured in terms of Specific Growth Rates (*SGR)* (% Mt/day) (Ricker, 1975; Boyer *et al.*, 1994) using the formula $SGR = 100 \times ((\ln Mt_1 - \ln Mt_0)/t)$ (for *SGR1*: Mt_1 =M2, Mt_0 =M1; for

SGR2: Mt_1 =M3, Mt_0 =M2; and for *SGRtot*: Mt_1 =M3, Mt_0 =M1). More than 200 fish were weighed, and their chips were correctly identified during both the 2004 (218) and 2005 (238) weighing processes in the hatchery. During the study, all fish were maintained in an 8-m-diameter sea cage, where they were fed at a rate of 0.7%. As the result of a hatchery management decision, over 100 fish were lost in 2006 before any fitness measures could be obtained; thus, only 107 fish were measured and weighed in June 2006. A principal component analysis (PCA) for the seven fitness traits measured was performed using the R software package (Hornik, 2006).

Microsatellite analysis

DNA was extracted from tissue stored in ethanol using the rapid Chelex® protocol (Walsh *et al.,* 1991). This protocol involved heating approximately 1 mg of fish caudal fins for 1 hour at 55ºC with 500 μL of 10% Chelex® (which was previously heated to 60ºC) and 7.5 μL of Proteinase K (P-K, 20 mg mL^{-1}). The P-K was inactivated by heating the samples at 100ºC for 15 min.

The samples were analyzed using two groups of microsatellites:

1) Eleven "classic" microsatellites whose genomic locations were unknown (*CM*), were obtained from several published genomic libraries: *SaGT1*, *SaGT26*, *SaGT41b* (Batargias *et al.,* 1999), *SauI41INRA* (Launey *et al.,* 2003), *Pb-OVI-A3*, *Pb-OVI-B2*, *Pb-OVID102*, *Pb-OVI-D22*, *Pb-OVI-D106* (Piñera *et al.,* 2006) *Dxd44*,

208 *Dpt3* (last two from De la Herrán *et al.,* 2005). Specific PCR procedures were
209 performed as previously described by Borrell *et al.* (2007) (Table I).
210
211 2) Eleven microsatellites located in gene-rich regions (*GRM*) were obtained,
212 including 2 markers from *EST* libraries and 9 markers identified from the publically-
213 available sequences for genes related to somatotopic axis (growth) (De Santis &
214 Jerry, 2007).
215
216 a) Five microsatellite primer pairs were obtained from previous reports
217 (summarized in Franch *et al.* (2006) and Table I):
218 *μ184*: *S. aurata* pituitary cDNA *EST*-library (Power *et al.*, 2003);
219 *μ190: S. aurata* pituitary cDNA *EST*-library (Power *et al.*, 2003);
220 *saGHpCA: S. aurata* Growth Hormone gene (5´-UTR) (Almuly *et al.*,
221 2005);
222 *2G*: *S. aurata* myostatin gene (Intron II) (Maccatrozzo *et al.*, 2001);
223 *G4*: *S. aurata* myogenic factor 1 *MYOD1* gene (3´-UTR) (Tan & Du,
224 2002).
225
226 b) Six other microsatellite primer pairs were developed (denoted with a *)
227 in this study (see Table I for details):
228 *saPROpCA**: *S. aurata* prolactin gene (5´-UTR) (Astola *et al.*, 2003);
229 *saMT2pCA*: S. aurata* myostatin gene (3´-UTR) (Maccatrozzo *et al.*,
230 2001);
231 *saMGpCA**: *S. aurata* myogenin gene (5´-UTR) (Codina *et al.*, 2008);

saMD2pAC: S. aurata* myogenic factor 2 *MYOD2* gene (3´-UTR) (Tan & Du, 2002);

*saGHR1pCTC**: *S. aurata* growth hormone receptor type I gene (*GHR-I*) (Intron-Exon II) (Saera-Vila *et al.*, 2005; 2007);

*saGHR2pGCT**: *S. aurata* growth hormone receptor type II gene (*GHR-II*) (5´-UTR-Exon I) (Saera-Vila *et al.*, 2005; 2007).

To develop the new microsatellite primers, genes sequences were downloaded from GenBank and screened for repeats using the *Tandem Repeats Finder* software (Benson, 1999). Forward and reverse primers were designed for effective amplification of microsatellites using the *FastPCR Professional* package (Kalendar *et al.*, 2009) (Table I). The specific PCR procedures for each microsatellite are described in Table I.

All amplification products were analyzed on polyacrylamide gels using the OpenGene™ automated DNA sequencing system. One primer from each primer set was end-labeled with the fluorescent dye *CY5.5*, and the internal size marker was end-labeled with the fluorescent dye *CY5*. PCR reactions were diluted (0.6 μL of the PCR reaction was mixed with 0.6 μL of the marker), heated at 95ºC for two minutes, placed immediately on ice and loaded onto a Surefill® 6% denaturing polyacrylamide gel (Visible Genetics, Ontario). Allele sizes were calculated using *GeneObjects™* version 3.1 (Visible Genetics, Ontario).

Genetic analyses of sample characteristics.

The number of alleles at each microsatellite locus (N_a), the proportion of individual samples that were heterozygous (direct count heterozygosity, H_o) and the unbiased estimate of expected heterozygosity (H_e) were assessed using a modified version of the *Biosys-1* program capable of accepting more than 20 alleles per locus (Swofford & Selander, 1981). The *Fstat* statistical package (Goudet, 1995; 2001) was used to estimate the total variance in gene frequencies (F_{IT}), which were partitioned into components of variance occurring within (F_{IS}) and among (F_{ST}) samples for each locus, as described by Weir & Cockerham (1984). Significance levels for F_{IS} were assessed through randomizing alleles within samples 1000 times followed by Bonferroni correction (Rice, 1989). Tests for global population differentiation were applied using the log-likelihood G as the test statistic, with no assumption of random mating within samples (Goudet *et al.,* 1996). All loci were tested for linkage disequilibrium using *Fstat*. To check for genotyping errors, the data were analyzed with the *Microchecker* software (Van Oosterhout *et al.,* 2004). A test for recent bottlenecks (the Wilcoxon sign-rank test, Luikart & Cornuet, 1997) was conducted using the *Bottleneck* software. A two-phase model of mutation (TPM) assuming 90% SMM and 10% IAM, with 20,000 iterations was used (Cournet & Luikart, 1996).

Two different approaches were used to assess the extent of population or family structuring within this sample. Population structure was assessed using *Structure* (Pritchard *et al.,* 2000; 2007). The number of possible clusters (*K*) within the data set

using a Bayesian approach based on genotype data was assessed. The parameters for each run were set according to Pritchard *et al.* (2000; 2007) and Evanno *et al.* (2005). An admixture model (to infer the degree of admixture from the data, α) and correlated allele frequencies among populations were used. The burn-in was set to 50,000, and the MCMC was set to 1,000,000 chains. For each data set, four runs were conducted to estimate ΔK, as suggested by Evanno *et al.* (2005), as this was a better indication of the true value of *K* than *LnP(D*). The range of *K*s tested was 1 to 20.

Possible family structuring in the sample was tested using *Colony* (version 2) (Wang & Santure, 2009). Family structure will go undetected when the data is similar to data from a single population with no internal structure because independent panmictic subsets are not detectable using programs such as *Structure*. *Colony* implements a maximum likelihood method to jointly assign sibship and parentage using individual multilocus genotypes at a number of codominant marker loci. It will infer full-sibs, paternal half-sibs, and maternal half-sibs among "offspring" (i.e., these 271 individuals). *Colony* also estimates the current effective population size ($N_{\hat{e}}$) from sibship assignments (Wang, 2009). The logic behind the method is simple. A small population (small $N_{\hat{e}}$) will contain a high proportion of sibs because the smaller the $N_{\hat{e}}$, the greater the probability that two individuals drawn at random from the same cohort within a population are sibs that share one parent or both parents. Three replicates of short runs (Wang & Santure, 2009) using polygamy (both parents) and the full likelihood method (Wang, 2004), with no genotyping errors at the markers allowed, were performed.

Estimating MLH and testing HFCs.

The multilocus heterozygosity index (*MLH*) was calculated according to the commonly used formula recommended by Chapman *et al.* (2009) (i.e., *MLH* = (the number of heterozygous loci/the total number of loci)).

To determine if heterozygosity is correlated across loci within individuals the loci were randomly subdivided into two groups and then it was assessed whether the multilocus heterozygosity (*MLH*) of the first group of loci was correlated with the *MLH* of the second group of loci (Balloux *et al.*, 2004). This strategy has been previously proposed as a way to test for global inbreeding (Balloux *et al.*, 2004, Lieutenant-Gosselin & Bernatchez, 2006). The loci were randomly subdivided into two groups of 11, the individual *MHL* values was recalculated for both groups and then the correlation between those measurements was assessed using a simple linear regression. This procedure was repeated 10,000 times using R (*http://www.r-project.org*) (Hornik, 2006) to obtain the mean and standard deviation of the correlation coefficients. The g2 value, the overall correlation in heterozygosity among all loci, was also measured. This eliminates the need to divide the set of loci into two arbitrary halves. David *et al.* (2007) describe the method and its implementation in the *RMES* software package, which also computes the significance of the observed g2 value, as well as selfing rates when inbreeding is believed to stem from partial selfing. The g2 estimates are still valid even if the inbreeding is not due to selfing (David *et al.*, 2007).

Regression analyses were performed using the *SPSS* 15.0 statistical software. Two analyses were performed: a simple linear regression model using the *MLH* index and the logarithms of the fitness traits, and a multiple linear regression model using the 22 microsatellite loci data instead of *MLH*. For the multiple regression model, a stepwise regression model was used to identify those loci that were most predictive of fitness using an automatic successive steps procedure and F tests with a significance threshold of $P<0.05$ (Hocking, 1976; *SPSS* 15.0). Quadratic regression models were conducted following the reasoning of Blanchet *et al.* (2009), to uncover signals of stabilizing selection (see also Neff (2004)). Both models (linear and quadratic) were compared using ANOVA and AIC-BIC tests (as implemented in R; Hornik, 2006). Similar regression analyses were conducted, this time using the principal components *PC1* and *PC2*, which contained around an 80% of the information of the fitness variables assessed here. The relationship between those principal components and single locus heterozygosity was assessed using regression analyses to confirm relevant single locus-*HFC*s.

It was also assessed whether dividing the microsatellites into two classes, classic microsatellites with unknown locations (*CM*) and gene-rich regions microsatellites (*GR*M), resulted in a significant improvement in terms of the amount of variance explained as compared to the simple regression model ($F = [(SS_{SIMPLE} - SS_{CM+GRM})] / [(SS_{CM+GRM} / (N-3)]$). Significance was assessed using tables of the F-distribution with 1 and N-3 degrees of freedom (a classical approach) and AIC and BIC values as well.

Results

Genetic description of the sample.

Of the 271 individuals analyzed, only eight produced samples that failed to amplify during PCR. These individuals were excluded from subsequent analyses. Genetic variation (N_a, H_e, H_o) among individuals was high (Table I). There was no significant difference between the two groups of microsatellites in terms of genetic variation levels (*CM* group: N_a=17.36, H_e=0.800, H_o=0.777; *GRM* group: N_a=14.09, H_e=0.785, H_o=0.735). The 22 microsatellite loci used here were not in linkage disequilibrium, although Bonferroni correction may have resulted in a number of false negatives, as there were 231 possible comparisons and 4620 permutations in total, resulting an adjusted P-value threshold for significance at the 5% nominal level of 0.0002. However, linkage between microsatellites was not found even for microsatellites located within the same gene (e.g., for microsatellites *2G* and *saMT2pCA** within the *S. aurata* myostatin gene, P=0.0781*).* A significant departure from Hardy-Weinberg expectations was found in this sample ($F_{IS\ (22\ microsatellites)}$ = +0.047, P=0.0023). This result was driven by seven loci (*SaGT26, SauI41INRA, Pb-OVI-A3, saGHpCA, µ184, saMT2pCA** and *saMD2pAC**) (Table I). Evidence for possible null alleles was found at four of these loci using *Microchecker*: *SauI41INRA*, *µ184, saMT2pCA** and *saMD2pAC** (Table I). A global test for population

differentiation using the log-likelihood G as a statistic (Goudet *et al.*, 1996) and an assumption of no random mating within samples did not detect any significant global genetic structuring in this sample, ($F_{ST\ (22\ microsatellites)}$ = 0.001 n.s.). The sample had a normal L-shape allele distribution, indicating that bottlenecks are very unlikely to have occurred recently in the sampled population (P=0.5253).

Analysis using *Structure* failed to identify a biologically sensible or likely number of the populations (*K*) within the sample. The procedure of Evano *et al.* (2005) for estimating *K* demonstrated that ln (Pr (X/*K*)) was bimodal (*K* = 1, *K* = 8), which indicates that the MCMC scheme was finding different solutions. Longer runs did not fix this problem. Moreover, the proportion of the sample assigned to each population was roughly symmetric and most individuals were admixed when testing the different *K* clusters (1-20). Three short runs of the *Colony* software converged toward the same result, that the 271 individuals are likely comprised of 111 males and 105 females from 229 full sib families (190 families with one individual, 36 with two individuals and 3 with 3 individuals) and 467 and 45 half and full sib pairs, respectively. A minimum of 14 independent sib clusters were inferred. Two individuals that are not directly related can still be included in the same sib cluster if they share a half-sib (*Colony* user guide notes). The effective population size estimated by *Colony* was $N_{\hat{e}}$ = 257, with a 95% CI of 210-314.

Fitness parameters.

The mean mass, length, and growth rate in this sample are shown in Table II. All fish weighed approximately 825 g in 2004. A mass increase of approximately 43% (a mean of 0.12% per day) was observed during the first year of this study (the third year of life for the fish). This trend did not continue into the second year of cultivation where a minor, but still significant, mass increase was observed (13%, a mean of 0.04% per day). The mean length of the fish at the end of the experiment (June 2006) was approximately 37 cm (Table II). Body masses were correlated with each other, and the specific body mass (M3) and length (L) were also correlated (F values $P<0.05$). The *SGR1* and *SGR2* indices were correlated with *SGRtot* ($P<0.05$), but not with each other ($P>0.05$). An ANOVA analysis comparing the means of all of the fitness traits under study (M1, M2, M3, L, *SGR1*, *SGR2*, *SGR*tot) between the 14 independent sibs clusters identified by *Colony* found no significant differences between them ($0.025<F<1.448$, $0.232<P<0.874$). The PCA indicated that one component (*PC1*) contains almost all of the variance in fitness traits (58.8% of the variance in the fitness data), except for *logSGR2,* which was mainly separated by the second PCA component (*PC2;* 20.6% of the variance in fitness).

Multilocus heterozygosity (MLH) and correlations within individuals.

The mean multilocus heterozygosity, $MLH_{(22\ \text{microsatellites})}$, was 0.9551 in this sample. When the loci were randomly subdivided into two groups of 11, and individual *MLH* values were recalculated for both groups and measured the correlation between

those measurements using linear regression 10,000 times, it was found that the *MLH* values within an individual were not correlated at all (r^2 = 0.002). The overall correlation in heterozygosity among all loci (based on g2 values) was zero (P=0.436 after 10,000 iterations).

Heterozygosity-Fitness correlations (HFCs).

Simple regression of *MLH* and the logarithms of the fitness measurements revealed several significant associations (Figure 1, Table III). Significant negative correlations between *MLH* (using all 22 microsatellites) and *L* (r^2 = 0.066, F = 7.12, P=0.009), *SGR1* (r^2 = 0.026, F = 5.13, P=0.025) and *SGRtot* (r^2 = 0.071, F = 7.14, P=0.009) were found. The *PC1* component from the PCA was negatively correlated with *MLH* (r^2 =0.083 F=7.57 P=0.0072) while *PC2* was not (P>0.05) (Figure 1, Table III). In all the cases, the quadratic models of regression did not explain the data variance any better than did any of the linear models assayed (P>0.05). These results indicate that heterozygotes had lower fitness than their homozygous counterparts in this work.

The results of multiple regression analyses (using all 22 loci as independent variables) are shown in Table III. Regression coefficients were not significant, and in all cases, loci with both positive and negative regression coefficients were observed (Table III). These models did not better explain the dependent variables (fitness traits) than did the global model previously assessed (no significant improvement was indicated by the

F values, AIC or BIC values). The automatic procedure for identifying loci that were most predictive of fitness found 6 loci, *SaGT26*(+), *Pb-OVI-B2*(-), *Pb-OVI-D22*(-), *2G*(-), *saMT2pCA**(+) and *saGHR1pCTC**(-) where heterozygosity could be used to predict some of the dependent fitness variables measured (Table III). All of these loci (with the exception of *saMT2pCA**) were also detected as loci related to *HFC*s after single locus analyses using simple regressions with the *PC1* and *PC2* variables (Table III).

Finally, multiple regression analyses using the *MLH* values of the two previously described groups of microsatellites ("potentially differentiated" in terms of their genomic locations: *CMs*-unknown *vs. GRM*s-vicinity of genes) as two independent variables, MLH_{CM} and MLH_{GRM}, were performed. The results obtained indicate that little gain in the fraction of variance in the assessed fitness traits explained is obtained, as compared to the simple regression model (Table III). F tests, comparing the simple and the two variables models, always yielded P-values greater than 0.05. A similar result was obtained using AIC and/or BIC values (data not shown).

Discussion

Two comprehensive reviews on heterozygosity-fitness correlations have recently been published (Chapman *et al*., 2009; Szulkin *et al*., 2010). The subject seems to have engendered serious and contentious debate. Chapman *et al.* (2009) conducted a quantitative review of *HFC*s in animal populations and concluded that *HFC*s studies do

not generally agree with the patterns predicted by population genetic theory and explain only small effects (less than 1%). This is the reason they have been largely discarded as a useful tool in selection and breeding strategies in aquaculture (Fjalestad, 2005). Szulkin *et al.* (2010) affirmed that quantitative and qualitative *HFC* studies are consistent with inbreeding-based theory. They rejected linkage disequilibrium and local effects as an alternative to inbreeding and described how *HFC*s can be used to quantify inbreeding load and unravel the structure of natural populations (Szulkin *et al.*, 2010). In this study, there were two main results: 1) late life-stage *HFC*s were observed in a sample from a Gilthead Sea bream population, and 2) the significant correlations between heterozygosity and growth found were negative.

Do HFCs in a late life-stage of a sample from a Gilthead Sea bream population represent general or local effects?

Szulkin *et al*. (2010) strongly support the idea that heterozygosity at neutral markers is correlated with heterozygosity at selected loci, both linked and unlinked, through genetic associations that arise in the context of a form of *sensu lato* inbreeding, such as when there is a small population size, nonrandom mating, population admixture or bottlenecks (see also Slate & Pemberton (2006)). The use of the approaches proposed by Balloux *et al.* (2004) and David *et al.* (2007) (the g values) demonstrate that marker heterozygosities are not correlated in this sample. Thus, marker heterozygosity did not reflect the level of inbreeding in this work. The family analysis revealed more than 200 full sib families (229) among the 271 fish being studied and the effective population

size as estimated from analysis of sibs was $N_{\hat{e}}$=257 (95% confidence intervals of 210-314), accounting for the sample size used. In light of this, an underlying familial structure within the sample does not appear to be responsible for the significant correlations found.

Some of loci genotyped in this work varied from Hardy-Weinberg expectations (in the form of heterozygote deficits). However, under the Wahlund effect, genetic markers are expected to show Hardy-Weinberg disequilibrium at most loci. This is not the case for the majority of the loci in this study. Hardy-Weinberg disequilibrium was observed at only 7/22 loci (31%), with evidence of the possible existence of null alleles in four of these loci (the sample would be in global HW equilibrium without those four loci). It would also expected to find evidence of linkage disequilibrium among the loci if there is population stratification (Wang *et al.,* 1998) and/or if a high percentage of individuals were asymmetrically assigned to putative subpopulations after Bayesian clustering with a software package such as *Structure* (Pritchard *et al.,* 2000). Neither of these possibilities occurred (this was also assessed using *Partition* 2 (Dawson & Belkhir, 2001), but the software used failed to determine the *K* number of subpopulations; data not shown)). The global population stratification test performed using the *Fstat* program was not significant. Together, the genetic analyses indicate that this sample lacks clear genetic structure.

Overall, this data does not point to a “general effect” (inbreeding, or family or population structure) being responsible for the correlations found. However, it should be

noted that finding a lack of significance when attempting to detect these general effects does not means they are absent (Waples, 1998; Hedrick, 1999; Szulkin *et al.*, 2010). In fact, Szulkin *et al*. (2010) said that one of the possible utilities of *HFC* studies is that they can serve as a "warning signal" for genetic erosion in unpedigreed captive or wild populations. *HFC* studies have suggested inbreeding in large and open populations of marine bivalves, which had been missed by previous studies that had used other methods of genetic analysis (Szulkin *et al*. (2010) and references therein).

David P. (pers. comm.) and Szulkin *et al*. (2010) have affirmed that local effects due to linkage or direct action are unlikely to be detected due to a dilution effect, and they have almost never been tested for using an adequate statistical approach (see David *et al*., 1997). This work has been tested for correlations that could be a result of linkage. Two types of microsatellites have been used; one with unknown genomic locations and another located in gene-rich regions. Thelen & Allendorf (2001) reasoned that allozyme loci might tend to be located in gene-rich regions, whereas microsatellite loci (which are commonly obtained from genomic libraries and have unknown genomic locations) would probably be located in gene-poor regions of the genome. Thus, it was suggested that the *HFC*s found when using allozymes but not microsatellites, such as those found in the deep-sea scallop *Placopecten magellanicus* (Zouros & Pogson, 1994), the rainbow trout *Oncorhynchus mykiss* (Thelen & Allendorf, 2001), the Atlantic salmon *Salmo salar* L. (Borrell *et al.*, 2004) and the eel *Anguilla anguilla* L. (Pujolar *et al.*, 2005), could be more influenced by a linkage effect than a direct overdominance effect (a hypothesis widely accepted at the time). However, it seems that the genomic locations of the markers used in this work had little to do with the *HFC*s found. It was

not observed a significant increase in the amount of variance in fitness traits that could be explained after dividing the loci into microsatellites from gene-rich regions (*GRM*) and classic microsatellites whose positions were unknown (*CM*). In fact, 3 loci from each "class" were identified as predictive variables for fitness traits. Similar to the work of Thelen & Allendorf (2001), this result is limited by the fact that the true location of the *CM*s in the genome cannot be determined (they could be also in gene-rich regions). Recently, Pujolar *et al.* (2009) found a complete lack of *HFC*s in eels using 22 expressed sequence tag-derived microsatellite loci that were all located in gene-rich regions. This result seems to confirm that a marker's location in a gene-rich region of the genome is not sufficient on its own to explain the existence of *HFC*s.

A more appropriate procedure to test for local effects seems to be to test whether a multiple regression incorporating specific effects for each locus explains more variance than a simple regression model (David, 1997; David *et al.*, 1997; David, 1998; Szulkin *et al.*, 2010). It does not seem correct to perform a separate regression for each locus because they are not independent from each other due to identity disequilibrium, as was affirmed by Szulkin *et al.* (2010), or to use Fisher tests, as fitness traits are not independent (see Fisher (1948), Sokal & Rohlf (1995) and Szulkin *et al.* (2010)). It was performed the same tests as David *et al.* (1997), and the multiple regression models did not help to explain significantly more variance in fitness traits. David *et al.* (1997) and Szulkin *et al.* (2010) remarked that although this is the more rigorous way to test for local effects, they are extremely difficult to detect if the *MLH* regression results are already weak. Sample size also limits the ability of David's tests (David, 1997) to statistically demonstrate direct or associative overdominance. All three tests (A, B and

C tests) proposed by David (1997) require large samples sizes (n≈1000 individuals; see also David *et al.,* 1997). In any case, some of the loci assayed here seem to be associated with some of the fitness traits under study. There is not information about the genomic location of the *SaGT26, Pb-OVI-B2,* or *Pb-OVI-D22* markers. However, it is not clear why the two microsatellites located in the vicinity of the myostatin gene, *saMT2pCA**(+) and *2G*(-), a gene that negatively regulates muscle development and growth (Maccatrozzo *et al.,* 2001), and the *saGHR1pCTC** (-) locus located in the growth hormone receptor gene showed significant single locus *HFC*s. They may be linked to relevant genes that influence fitness or they may themselves directly influence a role on fitness and growth. More research is required to answer these questions.

Negative HFC in a late life-stage Gilthead Sea bream sample.

Positive *HFC*s based DNA markers have been previously reported (Pogson & Fevolden, 1998) and continue to be reported today (Chapman *et al.,* 2009). However, negative correlations and heterosis can occur simultaneously in a single population (Den & Fu, 1998). Empirically, there are some examples of negative relationships between fitness and microsatellite heterozygosity in the literature (Zouros & Pogson, 1994; Borrell *et al.*, 2004; Lieutenant-Gosselin & Bernatchez, 2006; Van Dongen *et al.,* 2007), although many cases may go unreported (Chapman *et al.,* 2009).

606 Two ideas have been invoked to explain negative *HFCs.* The first involves
607 exogamic depression (Keller *et al.,* 2000; Fenster & Galloway, 2000; Burke & Arnold,
608 2001), whereby a breakdown of coadapted gene complexes occurs with increasing
609 heterozygosity, leading to reduced growth (Marshall & Spalton, 2000; Van Dongen *et*
610 *al.,* 2007). The second idea is that ribosomal (or other) genes that participate in protein
611 turnover are involved, (Mitton & Koehn, 1985; Hawkins & Day, 1999; Hedgecock *et*
612 *al.,* 2007) resulting in the dominance of low expression and/or underdominance, and
613 heterozygotes having poorer performance in some genic functions by because they are
614 more efficient individuals that consume less energy (Hedgecock *et al.,* 2007).
615
616
617 There is not evidence of outbreeding in this sample. There is neither excess of
618 heterozygotes, a finding that is commonly related to bottleneck events (Luikart &
619 Cournet, 1997), and also to exogamy (Von Ahsen *et al.,* 2010). Nevertheless, it could be
620 present, as the origin of the base population is not known. Thus the negative *HFCs*
621 found in this study could be a sign, perhaps the tip of the iceberg, indicating a previous
622 exogamic process. On the other hand, it has been previously observed that growth is not
623 consistent fitness in fish at all life-stages. In Atlantic salmon, the more heterozygous
624 salmon grow faster during early life-stages, but afterwards some fish initiate maturation
625 and divert more energy to gonad development and gamete production than to somatic
626 tissue development (Blanco *et al.,* 1998; Borrell, 2002; Pineda *et al.,* 2003; Fernandez,
627 2005). Thus, if some of the microsatellites assayed here are related (by linkage or in a
628 direct manner) to more efficient negative regulation of the growth pathways in sea
629 breams, then it is possible that individuals with greater heterozygosity could then be
630 more likely to have both sexual revert to females, as well as have a more active sexual

maturation process (in both sexes), which would result in lower growth than that demonstrated by their less heterozygous counterparts.

In summary, negative *HFC*s have been detected in a sample from the late life-stage of the Gilthead Sea bream *S. aurata* using microsatellite loci from unknown and gene-rich genomic locations. There was not any evidence of inbreeding, outbreeding or population/family structuring in this data. However, these results cannot rule that undetected general effects may exist that cause the appearance of *HFC*s. In fact, *HFC*s themselves may provide a way to reveal *sensu lato* forms of inbreeding or earlier exogamic processes in populations, even when genetic data indicates otherwise (the "tip of the iceberg" referred to by Szulkin *et al*. (2010)). In any case, these results seem to be best explained by the occurrence of local effects due to linkage with proximal or distal functional loci or even by possible direct locus advantages.

Acknowledgements

Thanks to F. Funkenstein who gave us sequences for the primers pairs for the *saGHpCA* locus. J. Wang helped with the use of the *Colony* 2 software. P. David helped with the *RMES* software. This work was carried out in collaboration with the fish farm Safor, S.L. (Gandía, Valencia) and the hatchery Piscimar de Burriana, Valencia. It has been financed by JACUMAR and the Department of Science and Technology in Spain (National Program of Resources and Food and Agriculture Technologies, *AGL2003-*

05362 and *AGL2007-64040* including European Regional Development Funds).

References

Almuly, R., Poleg-Danin, Y., Gorshkov, S., Gorshkova, G., Rapoport, B., Soller, M., Kashi, Y., Funkenstein, B. (2005). Characterization of the 5´flanking region of the growth hormone gene of the marine teleost, gilthead sea bream (*Sparus aurata*): analysis of the polymorphic microsatellite in the proximal promoter. *Fisheries Science* **71**, 479-490.

Aparicio, J.M., Ortego, J., Cordero, P.J. (2007). Can a Simple Algebraic Analysis Predict Markers–Genome Heterozygosity Correlations? *Journal of Heredity* **98**(1), 93–96.

Astola, A., Ortiz, M., Calduch-Giner, J.A., Pérez-Sánchez, J., Valdivia, M.M. (2003). Isolation of *Sparus aurata* prolactin gene and activity of the cis-acting regulatory elements. *General and Comparative Endocrinology* **134**, 57–61.

Balloux, F., Amos, W., Coulson, T. (2004). Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology* **13**, 3021–3031.

Batargias, C., Dermitzakis, E., Magoulas, A., Zouros, E. (1999). Characterization of six polymorphic microsatellite markers in gilthead sea bream, *Sparus aurata* (Linnaeus 1758). *Molecular Ecology Notes* **8**, 897-899.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27**(2), 573-580.

680 Blanchet, S., Bernatchez, L., Dodson, J.J. (2009). Does interspecific competition
681 influence relationships between heterozygsoity and fitness-related behaviors in
682 juvenile Atlantic salmon (*Salmo salar*)? *Behavioral Ecology and Sociobiology*
683 **63**, 605-615.
684 Blanco, G., Presa, P., Vázquez, E., Sánchez, J.A. (1998). Allozyme heterozygosity and
685 development in Atlantic salmon. *Fish Physiology and Biochemistry* **19**, 163–
686 169.
687 Borrell, Y.J. (2002). Loci microsatélites como marcadores genéticos para la mejora del
688 rendimiento en acuicultura de especies marinas. PhD thesis. Asturias.
689 Universidad de Oviedo. Dpto Biología Funcional. 2002. online.
690 http://www.tesisenred.net/TDR-0119109-103338. Last accessed 1 July 2011.
691 Borrell, Y.J., Pineda, H., McCarthy, I., Vazquez, E., Sánchez, J.A., Lizana, G.B. (2004).
692 Correlations between fitness and heterozygosity at allozyme and microsatellite
693 loci in the Atlantic salmon, *Salmo salar* L. *Heredity* **92**(6), 585-593.
694 Borrell, Y.J., Carleos, C.E., Asturiano, J.F., Bernardo, D., Vazquez, E., Corral, N.,
695 Sanchez, J.A., Blanco, G. (2007). Use of microsatellites and a combinatorial
696 optimization approach in the acquisition of gilthead sea bream (*Sparus aurata L*)
697 broodstocks for hatcheries. *Aquaculture* **269**, 200-210.
698 Boyer, N.J., Van Toever, W., Jansen, M.E. (1994). Effect of photoperiodon growth of
699 Arctic charr under commercial production conditions. *Progressive Fish-Culturist*
700 **56**, 44-46.
701 Brown, R.C. (2003). Genetic Management and Selective Breeding in Farmed
702 Populations of Gilthead Sea bream (*Sparus aurata*). PhD thesis. Stirling,
703 University of Stirling. UK.

Burke, J.M., Arnold, M.L. (2001). Genetics and the fitness of hybrids. *Annual Review of Genetics* **35**, 31–52.

Chapman, J.R., Nakagawa, S., Coltman, D.W., Slate, J., Sheldon, B.C. (2009). A quantitative review of heterozygosity–fitness correlations in animal populations. *Molecular Ecology* **18**(13), 2746 – 2765.

Codina, M., Bian, Y., Gutiérrez, J., Du, S. (2008). Cloning and characterization of myogenin from sea bream (*Sparus aurata*) and analysis of promoter muscle specificity. *Comparative Biochemistry and Physiology. Part D, Genomics & Proteomics* **3**(1), 128-139.

Coltman, D.W., Pilkington, J.G., Smith, J.A., Pemberton, J.M. (1999). Parasite-mediated selection against inbred soay sheep in free-living, island population. *Evolution* **53**, 1259–1267.

Coltman, D.W., Slate, J. (2003). Microsatellites measures of inbreeding: a meta-analysis. *Evolution* **57**(5), 971-983.

Cornuet, J.M., Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001–2014.

Coulson, T.N., Pemberton, J., Albon, S., Beaumont, M., Marshall, T.C., Slate, J. (1998). Microsatellite reveal heterosis in red deer. *Proceedings of the Royal Society of London B*: *Biological Sciences* **265**, 489–495.

Coulson, T.N., Albon, S., Slate, J., Pemberton, J. (1999). Microsatellite loci reveal sex-dependent responses to inbreeding and outbreeding in red deer calves. *Evolution* **53**, 1951–1960.

David, P. (1997). Modeling the genetic basis of heterosis: Tests of alternative hypotheses. *Evolution* **51**, 1049–1057.

David, P., Delay, B., Jarne, P. (1997). Heterozygosity and growth in the marine bivalve *Spisula ovalis*: testing alternative hypothesis. *Genetical Research Cambridge* **70**, 215-223.

David, P. (1998). Heterozygosity-fitness correlations: new perspectives on old problems. *Heredity* **80**(5), 531–537.

David, P., Pujol, B., Viard, F., Castella, V., Goudet, J. (2007). Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology* **16**, 2474-2487.

Dawson, K.J., Belkhir, K. (2001). A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genetical Research* **78**(1), 59-77.

De la Herrán, R., Magoulas, A., Garrido-Ramos, M.A., Ruiz-Rejón, C., Ruiz-Rejón, M., Zorros, E. (2005). Desarrollo de microsatélites en tres especies de espáridos con interés en acuicultura. In: La acuicultura como actividad económica de las zonas costeras. IX Congreso nacional de Acuicultura pp 113-136. Eds Consejeria de Agricultura y Pesca. Junta de Andalucia. Sevilla. Spain.

Deng, H.W., Fu, Y.X. (1998). Conditions for positive and negative correlations between fitness and heterzygosity in equilibrium populations. *Genetics* **148**, 1333-1340.

De-Santis, C., Jerry, D.R. (2007). Candidate growth genes in finfish-Where should we be looking? *Aquaculture* **272,** 22-38.

DeWoody, Y.D., DeWoody, J.A. (2005). On the estimation of genome-wide heterozygosity using molecular markers. *Journal of Heredity* **96,** 85–88.

Evanno, G., Regnaut, S., Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

754 Fenster, C.B., Galloway, L.F. (2000). Inbreeding and outbreeding depression in natural
755 populations of *Chamaecrista fasciculata* (Fabaceae). *Conservation Biology*
756 **14**(5), 1406-1412.
757 Fernández, A. (2005). Estrategias de crecimiento en salmón atlántico (*Salmo salar* L.)
758 durante el primer año de vida: Influencia ambiental y genética. University of
759 Oviedo. Spain. PhD Thesis.
760 Fisher, R.A. (1948). Combining independent tests of significance. *The American*
761 *Statistician* **2**(5), 30.
762 Fjalestad, K.T. (2005). Breeding strategies. In: Gjedrem, T. (Ed.), Selection and
763 Breeding Programs in Aquaculture. Springer, Dordrecht, The Netherlands.
764 Franch, R., Louro, B., Tsalavouta, M., Chatziplis, D., Tsigenopoulos, C.S.,
765 Sarropoulou, E., Antonello, J., Magoulas, A., Mylonas, C.C., Babbucci, M.,
766 Patarnello, T., Power, D.M., Kotoulas, G., Bargelloni, L. (2006). A genetic
767 linkage map of the hermaphrodite teleost fish *Sparus aurata L*. *Genetics* **174**(2),
768 851-861
769 Goudet, J. (1995). FSTAT (vers. 1.2): a computer program to calculate *F*-statistics.
770 *Journal of Heredity* **86**, 485-486.
771 Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation
772 indices (version 2.9.3). Available from
773 http://www2.unil.ch/popgen/softwares/fstat.htm. Last accessed 1 July 2011
774 Goudet, J., Raymond, M., Demeeus, T. Rousset, F. (1996). Testing differentiation in
775 diploid populations. *Genetics* **144**, 1933-1940.
776 Grueber, C.E., Wallis, G.P., Jamieson, I.G. (2008). Heterozygosity–fitness correlations
777 and their relevance to studies on inbreeding depression in threatened species.
778 *Molecular Ecology* **17**, 3978–3984.

779 Haldane, J.B.S. (1954). An exact test for randomness of mating. *Journal of Genetics* **52**,
780 631-635.
781 Hansson, B., Westerberg, L. (2002). On the correlation between heterozygosity and
782 fitness in natural populations. *Molecular Ecology* **11**, 2467–2474.
783 Hansson, B., Westerberg, L. (2008). Heterozygosity–fitness correlations within
784 inbreeding classes: local or genome-wide effects? *Conservation Genetics* **9**, 73–
785 83.
786 Hawkins, A.J.S., Day, A.J. (1999). Metabolic interrelations underlying the
787 physiological and evolutionary advantages of genetic diversity. *American*
788 *Zoologist* **39**, 401-411.
789 Hedgecock, D., Lin, J., DeCola, S., Haudenschild, C.D., Meyer, E., Manahan, D.T.,
790 Bowen, B. (2007). Transcriptional analysis of growth heterosis in larval pacific
791 oysters (*Crassostrea gigas*). *Proceedings of the National Academy of Sciences*
792 **104**(7), 2313-2318.
793 Hedrick, P.W. (1999). Highly variable genetic loci and their interpretation in evolution
794 and conservation. *Evolution* **53,** 313-318.
795 Hocking, R.R. (1976). The Analysis and Selection of Variables in Linear Regression.
796 *Biometrics* **32**(1), 1-49.
797 Hornik, K. (2006). The R.F.A.Q. www.r-project.org . ISBN 3-900051-08-9.
798 Kalendar, R. (2009). FastPCR software for PCR design and repeat search.
799 [http://primerdigital.com/fastpcr.html]. Last accessed 01 July 2011.
800 Keller, M., Kollmann, J., Edwards, P.J. (2000). Genetic introgression from distant
801 provenances reduces fitness in local weed populations. *Journal of Applied*
802 *Ecology* **37**, 647-659.

803 Koehn, R.K., Diehl, W.J., Scott, T.M. (1988). The differential contribution by
804 individual enzymes of glycolysis and protein catabolism to the relationship
805 between heterozygosity and growth rate in the coot calm *Mulinia lateralis*.
806 *Genetics* **118**, 121–130.
807 Launey, S., Krieg, F., Haffray, P., Bruant, J.S., Vanniers, A., Guyomard, R. (2003).
808 Twelve new microsatellites markers for gilted sea bream *(Sparus aurata L.):*
809 characterization, polymorphism and linkage. *Molecular Ecology Notes* **3,** 457-
810 459.
811 Li, Y.C., Korol, A.B., Fahima, T., Nevo, E. (2004). Microsatellites Within Genes:
812 Structure, Function, and Evolution. *Molecular Biology and Evolution* **21**(6),
813 991–1007.
814 Lieutenant-Gosselin, M., Bernatchez, L. (2006). Local heterozygosty–Fitness
815 correlations with global positive effects on fitness in threespine stickleback.
816 *Evolution* **60**(8), 1658–1668.
817 Luikart, G., Cornuet, J.M. (1997). Empirical evaluation of a test for identifying recently
818 bottlenecked populations from allele frequency data. *Conservation Biology*
819 **12**(1), 228–237.
820 Maccatrozzo, L., Bargelloni, L., Radaelli, G., Mascarello, F., Patarnello, T. (2001).
821 Chaacterization of the myostatin gene in the gilthead sea bream (*Sparus aurata*):
822 Sequence, genomic structure and expresión patterns. *Marine Biotechnology* **3,**
823 224-230.
824 Marshall, T.C., Spalton, J.A. (2000). Simultaneous inbreeding and outbreeding
825 depression in reintroduced Arabian oryx. *Animal Conservation* **3,** 241–248.
826 Mitton, J.B. (1993). Enzyme heterozygosity, metabolism, and developmental stability.
827 *Genetica* **89**, 47–65.

828 Mitton, J.B. (1997). Selection in natural populations. Oxford University Press, Oxford.
829 Mitton, J.B., Koehn, R.K. (1985). Shell shape variation in the blue mussel Mytulis
830 edulis L., and its association with enzyme heterozygosity. *Journal of*
831 *Experimental Marine Biology and Ecology* **90,** 73–80.
832 Neff, B.D. (2004). Stabilizing selection on genomic divergence in a wild population.
833 *PNAS* **101**(8), 2381-2385.
834 Ohta, T. (1971). Associative overdominance caused by linked detrimental mutations.
835 *Genetical Research* **18**, 277–286.
836 Pemberton, J. (2004). Measuring inbreeding depression in the wild: the old ways are the
837 best. *Trends in Ecology and Evolution* **19**(12), 613-616.
838 Pineda, H., Borrell, Y.J., McCarthy, I., Vazquez, E., Sanchez, J.A., Blanco, G. (2003).
839 Timing of first feeding and life-history strategies in salmon: genetic data.
840 *Hereditas* **139**, 41-48.
841 Piñera, J.A., Bernardo, D., Blanco, G., Vázquez, E., Sánchez, J.A. (2006). Isolation and
842 characterization of polymorphic microsatellite markers in *Pagellus bogaraveo*,
843 and cross-species amplification in *Sparus aurata* and *Dicentrarchus labrax*.
844 *Molecular Ecology Notes* **6**(1), 33-35.
845 Pogson, G.H. (1991). Expression of overdominance for specific activity at the
846 phosphoglucomutase-2 locus in the pacific oyster, Crassostrea gigas. *Genetics*
847 **128**, 133–141.
848 Pogson, G.H., Fevolden, S.E. (1998). DNA heterozygosity and growth rate in the
849 Atlantic Cod *Gadus morhua* (L). *Evolution* **52,** 915–920.
850 Power, D., Almeida, S., Louro, B.E.P. (2003). Sea bream Pituitary C-DNA *EST*-
851 Library. Genbank references: DN048405, DN048399.

852 Pritchard, J.K., Stephens, M., Donnelly, P. (2000). Inference of population structure
853 using multilocus genotype data. *Genetics* **155,** 945–959.
854 Pritchard, J.K., Wen, X., Falush, D. (2007). Documentation for structure software:
855 version 2.2 April 3, 2007. http://pritch.bsd.uchicago.edu/software. Last accessed
856 01 July 2011.
857 Pujolar, J.M., Maes, G.E., Vancoillie, C., Volckaert, F.A.M. (2005). Growth rate
858 correlates to individual heterozygosity in the european eel, *Anguilla anguilla* L.
859 *Evolution* **59**(1), 189-199.
860 Pujolar, J.M., Maes, G.E., Vancoillie, C., Volckaert, F.A.M. (2006). Environmental
861 stress and life-stage dependence on the detection of heterozygosity-fitness
862 correlations in the European eel, *Anguilla anguilla*. *Genome* **49**, 1428-1437.
863 Pujolar, J.M., Bevacqua, D., Capoccioni, F., Ciccotti, E., De Leo, G.A., Zane, L. (2009).
864 Genetic variability is unrelated to growth and parasite infestation in natural
865 populations of the European eel (*Anguilla anguilla*). *Molecular Ecology* **18**(22),
866 4604-4616.
867 Rice, W.R. (1989). Analysing tables of statistical tests. *Evolution* **43**, 223-225.
868 Ricker, W.E. (1975). Computation and interpretation of biological statistics of fish
869 populations. *Fisheries Research Board of Canada Bulletin* **191**: 1-382.
870 Rowe, G.., Beebee, T.J.C., Burke, T. (1999). Microsatellite heterozygosity, fitness and
871 demography in natterjack toads *Bufo calamita*. *Animal Conservation* **2**, 85–92.
872 Saera-Vila, A., Calduch-Giner, J.A., Pérez-Sánchez, J. (2005). Duplication of growth
873 hormone receptor (GHR) in fish genome: gene organization and transcriptional
874 regulation of GHR type I and II in gilthead sea bream (*Sparus aurata*). *General
875 and Comparative Endocrinology* **142**, 193-203.

876 Saera-Vila, A., Calduch-Giner, J.A., Pérez-Sánchez, J. (2007). Co-Expression of IGFs
877 and GH receptors (GHRs) in gilthead sea bream (*Sparus aurata* L.*):* sequence
878 analysis of the GHR-flanking region. *Journal of Endocrinology* **194**, 361-372.
879 Sharma, V.K., Kumar, N., Brahmachari, S.K., Ramachandran, S. (2007). Abundance of
880 dinucleotide repeats and gene expression are inversely correlated: a role for gene
881 function in addition to intron length. *Physiological Genomics* **31**, 96-103.
882 Slate, J., Pemberton, J. (2006). Does reduced heterozygosity depress sperm quality in
883 wild rabbits (*Oryctolagus cuniculus*)? *Current Biology* **16**, 790–791.
884 Slate, J., David, J., Dodds, K.G., Veenvliet, B.A., Glass, B.C., Broad, T.E., McEwan,
885 J.C. (2004). Understanding the relationship between the inbreeding coefficient
886 and multilocus heterozygosity: theoretical expectations and empirical data.
887 *Heredity* **93**, 255–265.
888 Sokal, R.R., Rohlf, F.J. (1995). Miscellaneous methods. Pp. 794–797 *in* Biometry: the
889 principles and practice of statistics in biological research. 3rd ed. Freeman and
890 Company, New York.
891 Streelman, J.T., Kocher, T.D. (2002). Microsatellite variation associated with prolactin
892 expression and growth of salt-challenged tilapia. *Physiological Genomics* **9**, 1-4,
893 Swofford., D.L., Selander, R.B. (1981). BIOSYS: a Fortran program for the
894 comprehensive analysis of electrophoretic data in population genetics and
895 systematics. *Journal of Heredity* **72,** 281-287.
896 Szulkin, M., Bierne, N., David, P. (2010). Heterozygosity-fitness correlations: a time for
897 reappraisal. *Evolution* **64**(5), 1202-1217.
898 Tan, X., Du, S.J. (2002). Differential expresión of two MyoD genes in fast and slow
899 muscles of gilthead sea bream (*Sparus aurata*). *Development, Genes and*
900 *Evolution* **212**, 207-217.

Thelen, G.C., Allendorf, F.W. (2001). Heterozygosity- fitness correlation in rainbow trout: Effects of allozyme loci or associative overdominance? *Evolution* **55**(6), 1180-1187.

Tiira, K., Laurila, A., Enberg, K., Piironen, J., Aikio, S., Ranta, E., Primmer, C.R. (2006). Do dominants have higher heterozygosity? Social status and genetic variation in brown trout, *Salmo trutta*. *Behavioral Ecology and Sociobiology* **59**, 657–665.

Väli, Ü., Einarsson, A., Waits, L., Ellegren, H. (2008). To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Molecular Ecology* **17**, 3808-3817.

Van Dongen, S., Backeljau, T., Matthysen, E., Dhondt, A.A. (2007). Fitness-heterozygosity associations differ between male and female winter moths *Operophtera brumata L*. *Belgian Journal of Zoology* **137**(1), 41-45.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P. (2004). Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**(3), 535-538.

Von Ahsen, N., Tzvetkov, M., Karunajeewa, H.A., Gomorra, S., Ura, A., Brockmoller, J., Davis, T.M.E., Mueller, I., Llett, K.F., Oellerich, M. (2010). CYP2D6 and CYP2C19 in Papua New Guinea: High frequency of previously uncharacterized CYP2D6 alleles and heterozygote excess. *International Journal of Molecular Epidemiology and Genetics* **1**(4), 310-319.

Walsh, P.S., Metzger, D.A., Higuchi, R. (1991). Chelex^{R} 100 as a médium for simple extraction of DNA for PCR-based typing form forensic material. *Biotechnics* **10,** 506-510.

925 Wang, J.L. (2004). Sibship reconstruction from genetic data with typing errors. *Genetics*
926 **166,** 1963-1979.
927 Wang, J.L. (2009). A new method for estimating effective population sizes from a
928 single sample of multilocus genotypes. *Molecular Ecology* **18**, 2148-2164.
929 Wang, J.L., Santure, A.W. (2009). Parentage and sibship inference from multi-locus
930 genotype data under polygamy. *Genetics* **181,** 1579-1594.
931 Wang, J.L., Caballero, A., Hill, W.G. (1998). The effect of linkage disequilibrium and
932 deviation from Hardy–Weinberg proportions on the changes in genetic variance
933 with bottlenecking. *Heredity* **81,** 174–186.
934 Waples., R.S. (1998). Separating the wheat from the chaff: patterns of genetic
935 differentiation in high gene flow species. *Journal of Heredity* **89,** 438-450.
936 Weir, B.S., Cockerham, C.C. (1984). Estimating *F*-statistics for the analysis of
937 population structure. *Evolution* **38**, 1358–1370.
938 Xu, Y., Zhu, Z., Lo, L., Wang, C., Lin, G., Feng, F., Yue, G. (2006). Characterization of
939 two parvalbumin genes and their association with growth traits in asian seabass
940 (*Lates calcarifer*). *Animal Genetics* **37**, 266-268.
941 Zohar. Y,, Abraham, M., Gordin, H. (1978). The gonadal cycle of the captivity-reared
942 hermaphroditic teleost *Sparus aurata* (L.) during the first two years of life.
943 *Annals of Biology and Animal Biochemistry and Biophysics* **18**, 877–882.
944 Zohar, Y., Harel, M., Hassain, S., Tandler, A. (1995). Gilthead sea bream (*Sparus*
945 *aurata*) In, Bromage, NR and Roberts, RJ (Eds) *Broodstock Management and*
946 *Egg and Larval Quality*. Blackwell Science, Oxford.
947 Zouros, E. (1993). Associative overdominance: Evaluating the effects of inbreeding and
948 linkage disequilibrium. *Genetica* **89,** 35–46.

949 Zouros, E., Pogson, G.H. (1994). The present status of the relationship between
950 heterozygosity and heterosis. In: Beaumont AR (eds.) Genetics and the
951 Evolution of Aquatic Animals. Chapman & Hall: London, pp 135–146.